

统计物理和复杂系统专题编者按

$\{\text{article.authorCnNames}\}$

$\{\text{articleTitleEn}\}$

引用信息 Citation: *Acta Physica Sinica*, 69, 080101 (2020) DOI: 10.7498/aps.69.080101

在线阅读 View online: <https://doi.org/10.7498/aps.69.080101>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

复杂系统重构

Complex system reconstruction

物理学报. 2020, 69(8): 088906 <https://doi.org/10.7498/aps.69.20200001>

小型化铯光钟物理系统的研制

Miniaturization of physics system in Sr optical clock

物理学报. 2018, 67(5): 050601 <https://doi.org/10.7498/aps.67.20172584>

全固态磁制冷系统物理模型的研究进展

Research progress of physical model of full-solid-state magnetic refrigeration system

物理学报. 2019, 68(21): 217501 <https://doi.org/10.7498/aps.68.20191139>

具有多物理特性的X射线脉冲星导航地面验证系统

Ground verification system of X-ray pulsar navigation with multi-physical properties

物理学报. 2019, 68(8): 089701 <https://doi.org/10.7498/aps.68.20182232>

基于光纤微结构加工和敏感材料物理融合的光纤传感技术

Optical fiber sensing technologies based on femtosecond laser micromachining and sensitive films

物理学报. 2017, 66(7): 070703 <https://doi.org/10.7498/aps.66.070703>

转动双星同步和轨道圆化的物理过程研究

Physical process of tidal synchronization and orbital circularization in rotating binaries

物理学报. 2018, 67(19): 199701 <https://doi.org/10.7498/aps.67.20181056>

专题：统计物理和复杂系统

统计物理和复杂系统专题编者按

DOI: 10.7498/aps.69.080101

从 20 世纪中叶至今, 复杂系统研究迅速发展, 成为了引人注目并具有广泛应用的新领域. 复杂系统要么具有结构的复杂性, 要么具有演化的复杂性, 在多数情况下二者兼具. 不同于传统物理学通常处理的规则介质, 许多复杂系统具有复杂结构, 近年来受到极大关注的复杂网络结构就是其中最典型的代表. 同时复杂系统也可表现为演化行为的多样性和复杂性. 即便系统结构并不复杂, 系统中的非线性相互作用可能产生复杂的演化行为, 包括: 形形色色的不稳定性; 丰富的斑图动力学; 各种各样的自组织、涌现及进化行为等等. 物理学从一开始就深深进入了复杂系统研究领域, 其中统计物理无疑是研究和理解复杂系统最主要的工具.

复杂系统研究紧密联系着当前科学发展的两大趋势. 一是不同学科的交叉和融合. 近年来物理学和数学越来越深入地进入其他学科领域, 特别是生物学和社会科学, 使这些传统大多以定性描述为主的学科开始了以数据为依托的定量研究, 而这些交叉领域研究几乎都处于复杂系统的研究范畴. 二是大数据科学的迅猛发展和应用. 基于互联网和物联网数据采集和存储技术的突飞猛进, 现在可利用的数据量正在爆炸性的增长. 这些数据中包含了极大量对自然和社会的有用信息, 能合理利用会带来巨大并不断增长的财富. 但产生这些数据的系统和可能被这些数据所影响的系统, 往往都是复杂系统, 其行为具有高度的不可预测性, 使这笔财富并不容易获取. 深入研究复杂系统, 发展有效的数据分析手段是成功使用这笔潜在财富的关键和核心.

要研究和处理所有以上困难和问题, 统计物理是强有力的手段. 长期以来统计物理在处理各种不可确切预见的轨道和状态中发展了丰富的思想、方法和技术手段, 这些必然将会和已经为复杂系统的研究提供了强有力的工具. 同时复杂系统由于结构和行为的大量新特点又为统计物理的创新提供强大推动.

本专题邀请了在领域前沿活跃工作的专家学者撰写了 18 篇研究和综述论文, 介绍了作者们在该领域的最新进展和成果. 内容包括对物理领域以及生物、经济、工业和其他交叉领域的复杂系统的研究; 既有宏观经典系统的讨论, 也有量子系统复杂行为的探索; 有论文讨论了复杂系统行为的基础统计理论, 也有论文分析了复杂系统演化的同步化、斑图动力学及其调控. 专题中多篇论文涉及复杂网络问题: 有关于网络结构形成和稳定性分析, 也有利用网络产生的数据分析网络结构, 网络上信息传播, 网络结构下人文活动, 经济演化, 社会运行规律等等. 统计物理和复杂系统是一个内涵宏大的领域, 专题论文都是作者兴趣所在的课题研究成果和心得, 只涉及领域中的点点滴滴. 但我们期望专题中介绍的成果能加强国内学者在这一领域的交流, 吸引对该领域有兴趣的青年学者和学生进来钻研, 推动我国在这一领域的研究水平更上一层楼.

(客座编辑: 北京师范大学 胡岗; 电子科技大学 周涛; 中国科学院物理研究所 叶方富)

SPECIAL TOPIC—Statistical physics and complex systems

Preface to the special topic: Statistical physics and complex systems

DOI: 10.7498/aps.69.080101

Since the middle of the twentieth century the study of complex systems has been developing rapidly, and now has become a new scientific field of broad applications. Complex systems are defined as systems that have complex structures, complicated dynamics, or, as in most cases, both. Unlike

typical physical systems where regular media are considered, complex systems often have complex structures (or media), of which complex networks, attracting great attention in recent decades in both natural and social sciences, are representative examples. On the other hand, regardless of whether the structure is complex or regular, systems can present various complicated behaviors due to their dynamical nonlinearities, such as: various instabilities; rich pattern formation and dynamics; diverse emergent and evolutionary behaviors, and so on. Physics has been involved deeply in the development of this novel field from the beginning. In particular, statistical physics is the main tool for studying and understanding the structural and evolutionary rules of complex systems.

The study of complex systems is closely related to two important frontiers. The first is the development of interdisciplinary research. In recent years, physics and mathematics-based tools and thinking have been used more and more extensively in many other fields, such as the biological and social sciences, bringing quantitative computational analyses into these areas which traditionally relied on qualitative descriptions. In this aspect, the analyses of complex systems, in particular complex networks, often serve as an important and in some cases even an irreplaceable foundation. The second frontier is the rapid growth of big data science in recent decades. Due to the fast development of measurement, recording, and storage techniques, data are collected and accumulated at an explosive pace. These data contain an enormous amount of information from both natural and social systems. Extracting and making use of this information are of great interest. However, this is a challenging task, because discernible patterns are typically deeply buried or masked in data produced by complex systems. To realize the full potential of big data, the development of theoretical frameworks and techniques for data analyses of complex systems is critical.

Statistical physics is a powerful approach to tackle the challenges described above. In the past few centuries the field of statistical physics has developed profound and far-reaching ideas, methods, and techniques, for analyzing complicated problems with nondeterministic characteristics. This can offer and has offered powerful tools for studying problems of complex systems. Conversely, novel phenomena, features, and behaviors in complex systems, due to their structural and dynamical complexities, can stimulate conceptual development of statistical physics itself, and help it to explore its capacity further.

For the special issue “statistical physics and complex systems”, a number of leading scientists and experts working actively in this field were invited to contribute research and review papers, and to present their recent research achievements. This body of work investigates complex systems in a wide range of disciplines, including the fields of physics, biology, economical activities, and other natural and social systems. The collection includes discussions on classical systems as well as quantum complex behaviors; descriptions of basic theories of statistical physics of complex systems as well as rich behaviors of pattern formation, pattern dynamics and their controls. Many contributions in this issue address complex network problems, including problems of formation, stability, and reconstruction of network structures, and also dynamical problems of information transport and social activities in networks. Most of the above investigations are based on analyses of available data, both structural data and dynamical data. As the field of statistical physics and complex systems is vast, only a small number of selected areas and topics can be covered in one issue. Nevertheless, we hope that this special issue will enhance academic exchanges among scientists in the field, attract young scientists and students interested in this field to join the research community, and effectively promote research of this exciting emerging field in our country.

Guest editor: Hu Gang (*Beijing Normal University, China*); Zhou Tao (*University of Electronic Science and Technology of China*); Ye Fang-Fu (*Institute of Physics, Chinese Academy of Sciences*)

排名聚合算法在少量长列表聚合中的性能比较分析

陈玟宇 朱章黔 王晓蒙 贾韬

Comparison of performance of rank aggregation algorithms in aggregating a small number of long rank lists

Chen Wen-Yu Zhu Zhang-Qian Wang Xiao-Meng Jia Tao

引用信息 Citation: *Acta Physica Sinica*, 69, 080201 (2020) DOI: 10.7498/aps.69.20191584

在线阅读 View online: <https://doi.org/10.7498/aps.69.20191584>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

电场和温度对聚合物空间电荷陷阱性能的影响

Simulations of the effects of electric field and temperature on space charge traps in polymer

物理学报. 2017, 66(8): 087201 <https://doi.org/10.7498/aps.66.087201>

时域混合算法在一维海面与舰船目标复合电磁散射中的应用

Composite electromagnetic scattering from a ship located on one-dimensional sea surface with time-domain hybrid method

物理学报. 2017, 66(18): 180301 <https://doi.org/10.7498/aps.66.180301>

塑料类高分子聚合物材料水中目标声学参数反演

Inversion for acoustic parameters of plastic polymer target in water

物理学报. 2019, 68(21): 214301 <https://doi.org/10.7498/aps.68.20190991>

单分子光学探针揭示易混聚合物受限纳米区域的动力学

Single-molecule probes revealed dynamics of confined nano-regions in miscible polymer blends

物理学报. 2019, 68(14): 148201 <https://doi.org/10.7498/aps.68.20190423>

基于窄带隙聚合物的高性能可见-近红外光伏探测器

High performance visible-near infrared photovoltaic detector based on narrow bandgap polymer

物理学报. 2017, 66(22): 228501 <https://doi.org/10.7498/aps.66.228501>

专题：统计物理和复杂系统

排名聚合算法在少量长列表聚合中的性能比较分析*

陈玟宇¹⁾ 朱章黔²⁾ 王晓蒙¹⁾ 贾韬^{1)†}

1) (西南大学计算机与信息科学学院 软件学院, 重庆 400715)

2) (中国人民解放军陆军勤务学院国防经济系, 重庆 500106)

(2019年10月17日收到; 2019年11月24日收到修改稿)

排名聚合将多个排名列表聚合成一个综合排名列表, 可应用于推荐系统、链路预测、元搜索、提案评选等. 当前已有工作从不同角度对不同排名聚合算法进行了综述、比较, 但存在算法种类较少、数据统计特性不清晰、评价指标不够合理等局限性. 不同排名聚合算法在提出时均声称优于已有算法, 但是用于比较的方法不同, 测试的数据不同, 应用的场景不同, 因此何种算法最能适应某一任务在很多情况下仍不甚清楚. 本文基于 Mallows 模型, 提出一套生成统计特性可控的不同类型的排名列表的算法, 使用一个可应用于不同类型排名列表的通用评价指标, 介绍 9 种排名聚合算法以及它们在聚合少量长列表时的表现. 结果发现启发式方法虽然简单, 但是在排名列表相似度较高、列表相对简单的情况下, 能够接近甚至超过一些优化类方法的结果; 列表中平局数量的增长会降低聚合排名的一致性并增加波动; 列表数量的增加对聚合效果的影响呈现非单调性. 整体而言, 基于距离优化的分支定界方法 (FAST) 优于其他各类算法, 在不同类型的排名列表中表现非常稳定, 能够很好地完成少量长列表的排名聚合.

关键词: 排名聚合, 不等长列表, Mallows 模型, 有偏等级重叠**PACS:** 02.50.-r, 89.75.-k, 87.15.Qt, 87.15.nr**DOI:** 10.7498/aps.69.20191584

1 引言

排序是在复杂系统的研究中经常使用的方法, 例如通过节点中心性对重要节点排序^[1,2], 链路预测中对可能存在的边排序^[3,4]. 在很多场景中, 我们会面对基于不同参数, 采用不同机制所得到的不同排名, 如不同人群对候选人给出的投票排名, 不同推荐系统给出的推荐排名, 不同算法给出的链路预测排名. 将多个排名列表聚合成一个综合排名列表, 就是排名聚合 (rank aggregation, RA) 要解决的问题. 排名聚合也称为 Kemeny 排名聚合^[5]、偏好聚合^[6]、共识排名问题^[7], 可以应用于推荐系统、

元搜索、期刊排名、提案评选、复杂网络等领域^[2,3,8-20]. 作为一个研究主题, 排名聚合已有超过两百多年的研究历史, 最早可追溯到 1781 年法国数学家 Borda^[21] 解决的法国科学院选举问题. 基于“总体大于各部分之和”的基本思想^[22], 不同研究领域提出了大量排名聚合方法, 大致可分为启发式方法和优化类方法. 启发式方法基于某种直观经验或规则将多个排名综合为一个排名, 具有简单、快速的特点. 优化类方法通过优化某一目标函数, 得到一个与已有排名之间存在某种优化量的排名. 虽然优化类方法在理论上更加完备, 但是因为其对应的优化问题是 NP (non-deterministic polynomial) 难的, 难以保证得到最优解. 对不同排

* 国家自然科学基金 (批准号: 61603309) 资助的课题.

† 通信作者. E-mail: tjia@swu.edu.cn

名聚合方法进行对比分析,能帮助在不同应用场景和数据条件下选择最合适的聚合方法,提高效率和可信度,具有重要的意义.

少量长列表的排名聚合是很多实际场景中需要解决的问题,这些排名列表相互不等长,且可能包含平局,例如不同机构给出的大学排名、基于不同基因数据库的靶点预测排名、不同搜索引擎对相关主题给出的推荐排名.选择何种算法能最好地处理少量长列表的排名聚合尚没有明确的答案.虽然当前已有综述类工作^[23,24]对不同聚合算法进行了列举和介绍,但是一般缺少不同方法之间的性能比较.一些工作^[25–30]虽然涉及了一部分排名聚合算法的性能比较,但是用于比较的聚合算法较少,使用的排序列表类型不够丰富,评价指标也有不够合理之处.同时,当前的绝大多数算法对比工作均是基于少量实证数据,由于数据的统计特征缺失,难以区分算法优劣是源自特定数据集,还是基于一系列数据集的共同统计特征,算法性能的泛化能力难以估测.最后,不同的聚合方法在提出时,均会声称优于已有的算法,但是用于比较基础算法不尽相同,用于测试的数据也往往大相径庭,使得我们往往无法真实知晓在具体应用场景下何种算法性能最佳.

针对这些局限性,本文基于 Mallows 模型提出一套用于生成统计特性可控的不同类型的排名列表的算法,使用一个可应用于不同类型排名列表的通用评价指标,介绍 9 种排名聚合算法,比较它们在少量长列表排名聚合下的表现.结果发现启发式方法虽然简单,但是在排名列表相似度较高、列表相对简单的情况下,能够接近甚至超过一些优化类方法的结果;列表中平局数量的增长会降低聚合排名的一致性并增加波动;列表数量的增加对各聚合结果的影响呈现非单调性.整体而言,基于距离优化的分支定界方法 (FAST) 优于其他各类算法,在不同类型的排名列表中表现非常稳定,能够很好地完成少量长列表的排名聚合.

2 基本概念

给定包含 $|S|$ 个对象的集合 S , 一个块序列 (bucket order) 是一种可传递二元关系 \prec , 存在块集合 $B_1, \dots, B_t (1 \leq t \leq |S|)$ 分割 S , 即块数量为 t . 如果 $x \in B_i$, 称 B_i 是 x 的块. 如果 $i < j$, 则称块 B_i

位于 B_j 之前. 对象 $x \prec y$ 当且仅当对于 $i < j$ 有 $x \in B_i$ 且 $y \in B_j$. 如果一个给定块中包含多个对象, 则形成一个平局, 即块内的对象具有相同的位置. 直观上看, 一个块序列就是一个可能包含平局的严格线性序列, 块中对象 x 的位置定义为 $p(x|x \in B_i) = \sum_{j < i} |B_j| + 1$, 即采用“1224”列表类型来表达对象的排名偏好信息, 而非采用块中对象的平均位置^[31]. “1224”列表类型表示为第 2 和第 3 个对象具有相同的位置, 形成一平局. 本工作主要针对少量长列表的聚合, 根据相关应用场景, 主要针对如下 3 种类别的列表^[32].

完全列表 (full list, FL): 块序列中所有块大小均为 1 且 $t = |S|$, 即包含 S 中所有对象且不含平局;

平局列表 (full list with ties, TL): 块序列中至少有一个块的大小大于 1 且 $1 \leq t \leq |S|$, 即包含 S 中所有对象且存在平局;

不完全列表 (incomplete list, IL): 由 S 中部分对象形成的块序列, 可能包含平局.

例如, 假设 S 包含 A, B, C 和 D 四个对象, 现有如下 5 个块序列:

块序列 1: [A], [B], [C], [D];

块序列 2: [A], [B, C], [D];

块序列 3: [A], [C];

块序列 4: [A], [B], [D];

块序列 5: [B], [C, A],

其中, 一个 “[]” 代表一个块, 即一个位置. 块序列 1 包含 S 中所有对象, 并且块数量等于 S 中对象数, 因此是 FL 类型. 块序列 2 包含 S 中所有对象, 但 B 与 C 都在块 2 中, 即 B 与 C 形成一平局, 是 TL 类型. 块序列 3—5 都只包含 S 中部分对象, 均为 IL 类型.

3 相关研究

Ali 和 Meila^[25] 重点探讨了不同排名聚合方法在搜索时间与结果表现之间的均衡性, 发现决定均衡的重要因素是数据集的一致程度 (degree of consensus). Schalekamp 和 Zuylen^[26] 也针对相似问题开展了工作, 并且将平局情况引入比较过程, 给出多个聚合方法表现的下界, 但是在计算过程中一些方法没有考虑打破平局的代价. 上述两个工作均使用真实数据集测试不同算法, 但是所使用数据量较小. Brancotte 等^[27] 针对平局比较了一系列排

名聚合方法,标注了不同方法是否适用于 TL 数据,研究了平局等数据集特征对所讨论方法的影响.但是以上几个工作都没有考虑更为复杂的 IL 数据.

Cohen-boulakia 等^[29]基于广义肯德尔距离,提出了一个新的启发式排名聚合算法,此方法可适用于所有列表类型,但是算法验证和对比仅仅基于一个小型生物医学数据集和人工数据(4—8 个对象).Lin^[24]从排名聚合方法在生物信息学中的应用角度,综述和对比了多个聚合算法,但是该工作使用的数据较小并且对 IL 数据的情况考虑不充分.Sculley^[33]比较了不同算法在大量短列表聚合下表现,但是其比较的方法较少,且均为启发式算法;Xiao 等^[30]基于其提出的数据生成模型对四种排名聚合算法进行了比较,并通过一个预设的真实排名(ground truth ranking)比较不同算法结果的质量.由于该数据生成模型对于平局数量、IL 数据生成过程不可控,同时真实排名在现实场景中难以获得,其算法质量评估在现实场景中的应用具有局限性.

在评价指标的选择上,大多数工作使用斯皮尔曼等级相关系数或肯德尔 τ 距离.这两个经典量只适用于排名列表包含所有对象的情况,不能应用在 IL 数据中,同时它们也没有考虑不同排名位置的不同权重.在真实场景中,靠前的排名应比靠后的排名具有更高的权重,例如第 1 名和第 2 名、第 50 名与 51 名之间均只相差 1 个排位,但前者的排名差距权重比后者更大.

综上所述,当前很多排名聚合算法比较工作存在数据量小、评估指标不够合理、数据统计特征缺失或列表类型单一等问题.针对以上问题,本文提出一个数据生成模型,生成列表类型、列表长度(即对象数量)、列表数量、一致程度等特征可控数据用于算法比较.本文同时采用更为合理的评价指标,更好地处理平局、排序权重等之前工作中未能充分考虑的问题,从而更合理地比较不同算法在不同情况下的表现.

4 排名聚合方法及其分类

排名聚合方法可根据相关特征从不同角度进行分类^[23,24,27,31,34].文献^[23]是最早对排名聚合方法进行分类的排名聚合公理派综述文章,将排名聚合方法分为启发式方法和优化类方法两类.文献^[24]将所有排名聚合方法分为基于分布的方法、启发式

方法和随机搜索方法三类.文献^[27]从能否处理平局的角度分为基于广义肯德尔距离的方法、基于传统肯德尔距离的方法和位置类方法三类.文献^[31]按是否需要训练数据分为监督类方法^[35,36]和非监督类方法两类.本文采用第一种分类标准.

4.1 启发式方法

4.1.1 KwikSort

KwikSort^[37]是一种分治算法,其核心思路是给定一组对象,随机选择一个对象作为中心点,并将其他所有对象按照一定规则置于该中心点前后的两个块中,从而使每一个对象与该中心点的违例数最少(为与下文排名聚合方法 MVR 中违例数相区分,此处违例数定义为对于两个对象,如果其在两个列表中的相对位置不一致,则形成一个违例).van Zuylen 和 Williamson^[38]提出了 KwikSort 去随机化版本.事实上,对象有一定概率与中心点形成平局.Brancotte 等^[27]通过这种策略对 KwikSort 进行了改进使其可以处理平局.

4.1.2 FaginSmall

Fagin 等^[31]针对平局提出几种指标并利用动态规划方法提出一种新的近似算法.Cohen-boulakia 等^[29]对上述方法做了修改得到 FaginLarge 和 FaginSmall,前者得到的结果包含平局,而后者不包含平局,因此,本文选用 FaginSmall.

4.1.3 BioConsert

给定 M 个对象和 N 个排名列表 $R_{\mu_1}, \dots, R_{\mu_N}$ (μ_1, \dots, μ_N 是每个列表对应的排名机制),通用目标函数为

$$R_C = \arg \min_{R_\mu} \left\{ \sum_{k=1}^N w_k d(R_\mu, R_{\mu_k}), R_\mu \in U \right\}, \quad (1)$$

其中 w 是权重向量,用于指定有关排名列表相对重要性或可靠性的先验信息.本文取为元素全为 1 的向量,以表征所有排名列表一样重要,此时该问题也称为 Kemeny 排名聚合问题.本文不考虑最终得到的聚合排名中包含平局的情况,所以 U 表示由这 M 个对象组成的所有 FL 集合,则 $|U| = M!$. d 是某种距离函数,如肯德尔 τ 距离^[39]、斯皮尔曼简捷距离^[40]、KS 距离^[23]以及 Hausdorff 距离^[41]等. R_C 在本文中叫聚合排名,而在社会选择和离散

数学文献中叫共识排名 (consensus ranking), 在统计学文献中称为中值排名 (median ranking).

BioConsert^[29] 的主要思路是从一个初始排名列表 R 开始, 通过不断迭代选择执行两类操作以减少上述通用目标函数值. 当任何操作均不能再减少目标函数值时, 则将此时的排名列表作为最终的聚合排名. 这两类操作是: 1) 将一个对象从原来的块中取出放入另一个块中; 2) 将一个对象从原来的块中取出并在另一个新的位置新建单对象块. 如果执行上述某一操作后目标函数变小, 则执行该操作; 反之, 则不执行. 不同初始排名列表 R 会产生不同的聚合排名结果, 极大影响算法的表现^[27]. 本文选用波达计数法的聚合排名 R_C 作为初始排名列表.

4.1.4 波达计数法 (BordaCount)

波达计数法^[21] 是最简单直观的排名聚合方法: 在每个排名列表中, 根据排名顺序对每个对象赋值一个分数, 将每个列表中该对象分数相加并对每个对象总分数进行排序就得到了最终的聚合排名. 同一对象在不同排名列表中的总分数除使用简单相加以外, 也可使用其他聚合函数, 例如中值函数、几何平均函数、 p 范数等^[24]. 本文采用最简单的波达计数法, 即某一列表中对象分数为该对象所击败的对象数量 (例如在长度为 M 的排列中, 第 1 名的分数为 $M - 1$, 而第 M 名的分数为 0), 并采用求和的方法计算各对象总分数.

4.1.5 MedRank

MedRank^[42] 的核心思路是给定阈值 $q \in [0, 1]$, 逐一并行读取所有 N 个排名列表, 一旦有对象出现次数首次超过 $N \times q$, 则根据对象出现的先后顺序依次添加至最终排名列表, 直至达到指定列表长度. 本文将阈值 q 取为 0.5.

4.1.6 马尔科夫链方法 (MC3)

Dwork 等^[11] 使用成对比较信息, 基于马尔科夫链提出了一系列的排名聚合方法, 其核心思路是将所有对象当作状态, 构建一各态历经的马尔科夫链转移矩阵, 从而其稳态分布会给予排在前面的状态更高的概率, 对象转移概率的不同赋值方法取决于我们的目标. Lin^[24] 对上述方法进行改进以使其可以适用于不同类型列表. 本文采用文献^[24] 中的 MC3 方法, 给定 M 个对象和 N 个排名列表

$R_{\mu_1}, \dots, R_{\mu_N}$ (μ_1, \dots, μ_N 是每个列表对应的排名机制), 对于 u, v 属于 M , 且 u 不等于 v , 则 $u \rightarrow v$ 的概率为

$$P(u \rightarrow v) = \frac{\sum_l^N I[R_{\mu_l}(u) > R_{\mu_l}(v)]}{N|M|}, \quad (2)$$

其中当 $I(\cdot)$ 所包含的条件满足时, $I(\cdot) = 1$, 否则 $I(\cdot) = 0$. 此时, 定义 $P(u \rightarrow u) = 1 - \sum_{v \neq u} P(u \rightarrow v)$. 同时, 对象转移概率与将待转向对象排在当前对象前面的列表数量成正比. 根据 (2) 式建立相应的状态之间的概率转移矩阵, 同时建立以状态为节点, 状态间转移概率作为边权重的加权有向图 $G(V, E)$, V 为状态集合, E 为边集合, 此时, 将问题转变为寻找最大的有向非循环连通子图.

4.1.7 PageRank

依据经典 PageRank 算法^[43] 的排名聚合算法 PageRank^[44] 的主要思想是给定排名列表构造图 $G(V, E)$, V 为节点集合, E 为边集合, 将每个对象看作图 G 中一个节点, 对于每一个排名, 如果对象 u 排名高于对象 v , 则建立一条加权有向边 $e(v, u)$, 其权重为所有给定排名列表的差值. 此外, 对所有权重进行归一化处理, 以便每个节点的出度边权重总和为 1. 对每个节点 u 构建 $\text{Pg}(u)$ 值作为排名得分, $\text{Pg}(u)$ 定义为

$$\text{Pg}(u) = (1 - \alpha)p_u + \alpha \sum_{e(v,u) \in E} \frac{\text{Pg}(v) \times w(u,v)}{\text{outdeg}(v)}, \quad (3)$$

其中 outdeg 是节点的出度, indeg 为节点的入度, $p_u = \frac{\text{indeg}(u)}{\sum_{v \in V} \text{indeg}(v)}$. 本文 α 取值为 0.85, 将算法给出的节点排序列表作为最终的聚合排名.

4.2 优化类方法

上述启发式方法尽管在运算速度上有优势, 但是并不能在理论上保证最终排名的性能最优性. 针对这一不足, 一些学者提出了优化类方法, 通过优化基于某一性能指标的目标函数, 获得聚合排名. 在衡量两个排名之间一致性情况下, 采用不同的性能指标 (如距离函数、等级相关系数和违例数等) 会得到不同的优化方法^[7,12,13,45-47]. 不同性能指标之间有时可相互转化, 比如在 FL 情况下, KS 距离和肯德尔 τ 距离完全等价^[48]; 此外, KS 距

离函数和 τ_x 等级相关系数是等价的, 具有线性变换关系 [45].

4.2.1 分支定界方法 (FAST)

肯德尔提出了 τ_a 和 τ_b 等级相关系数 [39], 前者适用于 FL, 而后者还适用于 TL. 因为 τ_b 在处理平局方面存在问题, Emond 和 Mason [45] 基于 τ_b 提出 τ_x 等级相关系数. 在一个包含 M 个对象的列表中, 定义分数矩阵 \mathbf{A} 为一个方阵, 对于任意两个对象 u 和 v , 如果 u 排在 v 前面或者与 v 形成平局, 则 $a_{uv} = 1$; 如果 u 排在 v 后面, 则 $a_{uv} = -1$; 如果 u 和 v 相等, 则 $a_{uv} = 0$, 即 \mathbf{A} 对角线上元素始终为 0. 因此, τ_b 在处理平局时将其分数矩阵对应元素置为 0, 而 τ_x 置为 1, 故而分数矩阵中除对角线以外的 0 元素表征无比较信息. 给定两个列表 R_{μ_1} 和 R_{μ_2} , τ_x 等级相关系数定义为

$$\tau_x(R_{\mu_1}, R_{\mu_2}) = \frac{\sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ij}}{M(M-1)}. \quad (4)$$

Emond 和 Mason [45] 基于 τ_x 提出一个分支定界算法来寻找聚合排名, 即寻找一个聚合排名 R_C 使平均加权 τ_x 等级相关系数最大 (或平均加权 KS 距离最小), 其目标函数为

$$\max \left[\sum_{k=1}^N w_k \tau_x(R_C, R^{(k)}) / \sum_{k=1}^N w_k \right].$$

注意: 此处的目标函数与通用目标函数 (1) 式是一致的, 因为 τ_x 与 KS 距离具有线性变换关系. 本文所有基础排名列表的权重都取 1, 以表征所有基础排名列表一样重要. 上述目标函数化简为

$$\max \sum_{i=1}^m \sum_{j=1}^m s_{ij} c_{ij}, \quad (5)$$

其中 $c_{ij} = \sum_{k=1}^n r_{ij}^{(k)}$. Amodio 等 [7] 基于 Emond 和 Mason 的分支定界算法提出 FAST 算法, 效果更好, 计算时间更短, 计算效率得到了很大提高, 是本文采用的方法.

4.2.2 最少违例数方法 (MVR)

在给定对象两两比较信息情况下, Pendings 等 [49] 利用 0—1 线性整数规划来寻找一个聚合排名以使对象不一致数量最少. 如果一个方阵每一行从左到右元素依次增大, 每一列从上至下依次减小, 那么该方阵就称为坡型矩阵. 对于一个方阵,

违反坡型结构的元素对数就是违例数 (注意: 此处违例数与 KwikSort 违例数 [27] 定义不一样). MVR 旨在尽可能寻找这样一个坡型结构, 从而使违例数最少. MVR 目标函数为

$$\min \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij}, \quad (6)$$

其中方阵 \mathbf{X} 是决策矩阵, 其值 $x_{ij} = 1$ 表示将 m_i 置于 m_j 前面, 否则置于后面, 并且需满足三个条件: 1) $x_{ij} \in \{0, 1\}$; 2) $x_{ij} + x_{ji} = 1$; 3) $x_{ij} + x_{jk} + x_{ki} \leq 2$. 方阵 \mathbf{C} 用于计算与坡型矩阵的违例数. 对任意 i 和 j , $c_{ij} = \#\{k | d_{ik} < d_{jk}\} + \#\{k | d_{ki} > d_{kj}\}$, 其中 d_{ij} 表示将 m_i 排在 m_j 前面的列表数量与将 m_j 排在 m_i 前面的列表数量的差值, 以衡量输入排名列表之间的一致程度. 值得注意的是, MVR 采用对象成对比较方式来表达偏好信息, 其最原始的应用场景是各类循环赛, 但在排名聚合背景下, 可将基础排名列表转换为两两比较信息. 如果出现平局或比较的两个对象至少有一个并未参与某一排名, 则不计算这两个对象在该排名中的得分, 因为未参与该排名, 故而无法得知孰强孰弱.

4.3 算法比较

本文主要基于 Mallows 模型生成的完全列表、包含平局的列表和不完全列表, 利用相似性指标来考察不同类型的列表数据对算法的影响. 从算法自身设计角度来说, 除少数算法以外, 多数都可以直接处理不同类型的列表. 例如, 除了 BioCount, MedRank 和 MVR 方法需要对列表作稍微调整外, 其余方法均可以处理包含平局的列表; 除 BioConsert, BordaCount 和 MedRank 需要对列表作稍微调整外, 其余方法均能处理 IL 数据.

以上 KwikSort, FaginSmall, BioConsert, BordaCount 和 MedRank 聚合算法使用文献 [27] 所提供的在线平台进行相关实验, MC3, PageRank, FAST 和 MVR 使用相关程序进行本地运算.

5 评价指标

为对比不同的排名聚合方法在不同列表类型上的表现, 需要一个通用评价指标来表征各排名之间的相似性. 一个合理的相似性度量指标需要能够处理对象未同时出现在排名中的情况, 即列表不等

长; 赋予高排名对象比低排名对象更多的权重; 同时相似度取值随着排名列表长度的增长而最终收敛. 在排名聚合领域, 有大量指标可用于计算列表之间相似性^[50–54], 但是很多都不能同时满足以上三个要求. 比如, 经常用于排名相似度计算的斯皮尔曼等级相关系数只能处理完全列表, 当列表不等长或列表元素存在不同时不再适用; 肯德尔 τ 距离以及广义肯德尔距离都可计算将一个排名转换为另一个排名所需的相邻对象交换次数, 但是却不能给排在前面的对象更高权重, 同时取值随着列表长度的增加而发散. Webber 等^[55] 基于简单概率用户模型实现了一个满足上述三个条件的指标——有偏等级重叠 (rank-biased overlap, RBO). RBO 通过在给定评估深度下计算一个基本分数 (下界) 和一个最大分数 (上界) 来提供单调性. 需要点估计时, 也可以计算出一个介于上下界之间的分数. $RBO \in [0, 1]$, 0 表示两个列表中对象完全不同, 1 表示两个列表包含相同对象且相对顺序一致. RBO 中包含一个参数 p , 其决定对象被加权的程度, 决定权重下降陡峭程度, p 越小, 指标对前面的对象加权越大. 本文根据文献^[55], 选取参数 $p = 0.9$.

为考察聚合算法的效能, 本文根据不同的列表类型, 使用 2 种指标. 对于 FL 数据, 由于列表生成模型基于一个中心排名均匀的产生随机样本, 因此可以认为中心排名即为最佳的聚合排名. 聚合排名与中心排名的差异性则体现出算法效能:

$$RBO = RBO(R_C, R_0), \quad (7)$$

其中 R_C 为聚合排名, R_0 为中心排名.

对于 TL 和 IL 数据, 由于在添加平局和构造不等长列表的过程中, 不可避免地使得列表不再基于中心排名列表随机分布, (7) 式不再适用. 因此我们使用聚合排名与原排名的平均相似度来度量聚合效果:

$$\overline{RBO} = \frac{1}{N} \sum_i^N RBO(R_i, R_C), \quad (8)$$

其中 $RBO(R_i, R_C)$ 代表原排名列表 R_i 与聚合排名列表 R_C 的相似度. 这一度量也同样适用于 FL 数据类型.

对于算法搜索时间比较方面, 由于各方法使用不同的平台、语言和优化工具, 本文不做比较.

6 数据生成模型

为评估和比较各种排名聚合方法, 首先需要生成具有不同统计特征的数据集, 包含多个相似但不相同的排名列表. TL 和 IL 数据都可作为 FL 数据的变体, 因此首先介绍 FL 数据的生成模型. 当前已有多种 FL 类型数据生成模型^[56–62], 本文选用理论和应用研究中应用广泛的 Mallows 模型^[56,57], 以更好地分析数据一致程度对排名聚合方法表现的影响. Mallows 模型是一个基于排名列表之间距离的指数模型, 包含两个参数: 中心排名 R_0 和离散分布参数 θ . Mallows 模型会给每一个排列赋予一个概率值

$$P(R|\theta, R_0) \propto e^{-\theta d(R, R_0)}, \quad (9)$$

其中 d 代表某种距离函数, 如肯德尔 τ 距离、海明距离、Cayley 距离和 Ulam 距离等^[58]. 本文使用肯德尔 τ 距离. θ 控制生成的排名与中心排名的距离. 当 $\theta = 0$ 时, 生成的排名 R 与 R_0 无关; θ 越大, 概率衰减越快, 生成的排名 R 越集中于 R_0 附近.

(9) 式虽然理论上非常简洁, 但是在实际操作中却存在难度, 一方面归一化常数的解析解难以获得, 另一方面也很难直接通过距离随机地生成一个排名. 因此在生成随机排名样本的过程中, 我们具体使用公式

$$P(R|\theta, n, d(R, R_0) = d) \propto S(n, d)e^{-\theta d}, \quad (10)$$

其中 $S(n, d)$ 为中心列表长度为 n 时, 所有与其距离为 d 的列表的数量. 在样本生成过程中, 首先基于中心排名, 穷举出所有相关的排名列表组合, 获得 $S(n, d)$ (这一步骤通常为 $O(n^3)$ 的复杂度). 根据 (10) 式随机生成距离 d , 再从所有与中心排名距离为 d 的排名列表中随机选择一个排名. 由于 (10) 式中的 $S(n, d)$ 随 d 增长, $e^{-\theta d}$ 随 d 下降, 因此最终获得的随机样本, 与中心排名的距离应该满足一钟型分布, 即概率存在一个峰值, 并在峰值距离两端迅速下降.

由于在 Mallows 模型中使用肯德尔 τ 距离生成排名序列, 而在聚合效果的度量中使用 RBO 相似度, 为验证生成的排名序列在 RBO 度量下也具有同样的统计特性, 基于不同的参数 θ 生成 3 组序列, 计算其与中心排名的相似度 RBO (图 1), 获得了预期中的钟形分布, 随着参数 θ 的增长, 生成的随机排名与中心排名相似度逐渐增加.

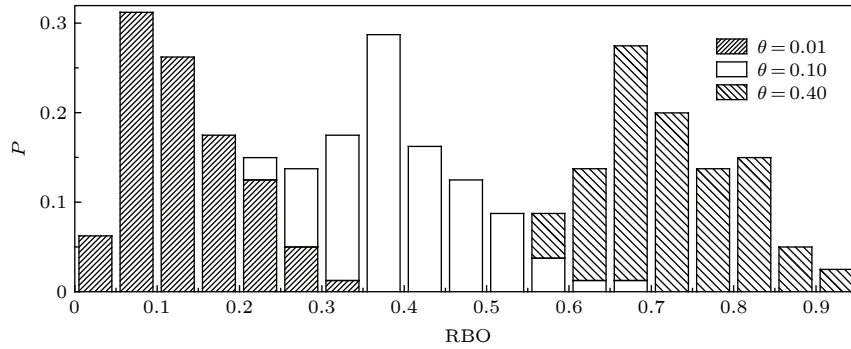


图 1 Mallows 模型在 RBO 度量下的表现
Fig. 1. The Mallows model under the RBO metric.

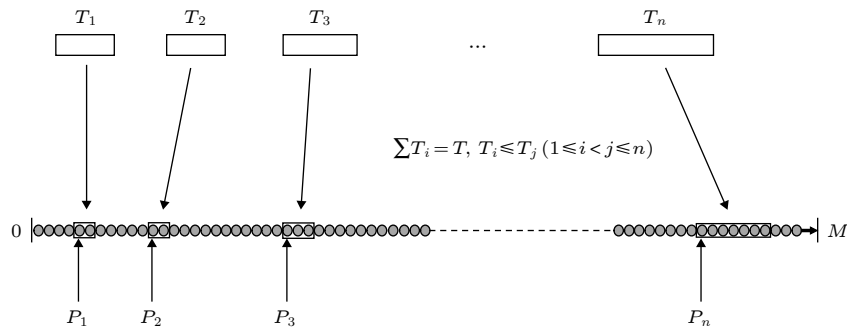


图 2 FL2 TL 示意图
Fig. 2. An illustration of FL2TL.

基于 Mallows 模型生成的 FL 数据, 本文提出两种方法将其转换为 TL 和 IL 类型数据.

1) FL2TL(r_t)

由 (10) 式获得长度为 M 的 FL 后, 生成一随机变量 $T \sim U[0, r_t * M]$, 作为列表中的平局总量, 其中 r_t 定义为平局比例, 取值范围为 $[0, 1]$. 令 $T' = T$, 在 2 至 T' 之间生成一随机数 t , 将一平局子块的长度设为 $\min(t, T' - t)$. 令 $T' = \max(t, T' - t)$, 重复以上步骤, 直到 $T' \leq 3$. 假设总共获得了 n 个和为 T 的随机数, 代表着对总长度为 T 的平局块的随机划分. 将 n 个随机数由小到大排序, 获得平局子块长度序列 $\{T_1, T_2, T_3, \dots, T_n\}$. 在排序列表中随机选择 n 个对象, 按照排名先后获得对象序列 $\{P_1, P_2, P_3, \dots, P_n\}$. 结合序列 T_i 与 P_i , 将序列中排在 P_i 到 $P_i + T_i - 1$ 范围内的对象设置为平局, 若两个平局之间存在相同的对象, 则合并这两个平局为一大平局. 整个过程如图 2 所示.

此方法在 FL 数据中加入了位置随机、长度随机的平局块, 并且兼顾了经验规律: 排名位置靠后的对象更容易形成包含对象数量较多的平局. 同时此方法只有一个参数, 便于进行相关分析. 为验证模型的可行性, 我们基于不同的参数 θ 和 r_t 生成

100 个排序, 计算了平均平局对 (平均的总个数) 和平局块 (平局集中出现在多少分块中) 的数量 (图 3), 发现通过控制 r_t 可以很好地控制平局对与平均块的数量. 作为验证, 我们也使用了一些多参数的复杂模型, 控制平局块长度和位置, 发现不同的算法并不改变本文所获得的聚合算法性能的整体结论.

2) TL2IL($r_k, \Delta k$)

考虑到真实数据中, 不同排名列表对靠前的排名个体差异并不大, 因此使用前端截取的方法生成不完全列表. 由 1) 中获得长度为 M 的 TL 数据后, 可以从列表前端截取一个长度为 L 的列表获得 IL 数据, 其中随机变量 $L \sim U[r_k * M - \Delta k, r_k * M + \Delta k]$, $r_k \in [0, 1]$ 定义为列表长度比例. 参数 r_k 控制 IL 数据的平均长度, 参数 Δk 控制列表长度的波动区间.

7 实验与讨论

本文主要讨论少量不等长列表的聚合情况, 如无特殊说明, 则使用列表长度 $M = 100$, 列表数量 $N = 20$, 初始数据一致程度 $\theta = 0.7$, $r_t = 0.2$, $r_k = 0.8$ 和 $\Delta k = 0.2$ 为参数. 本文对数据生成过程、聚合排名算法实现以及评价指标重复进行

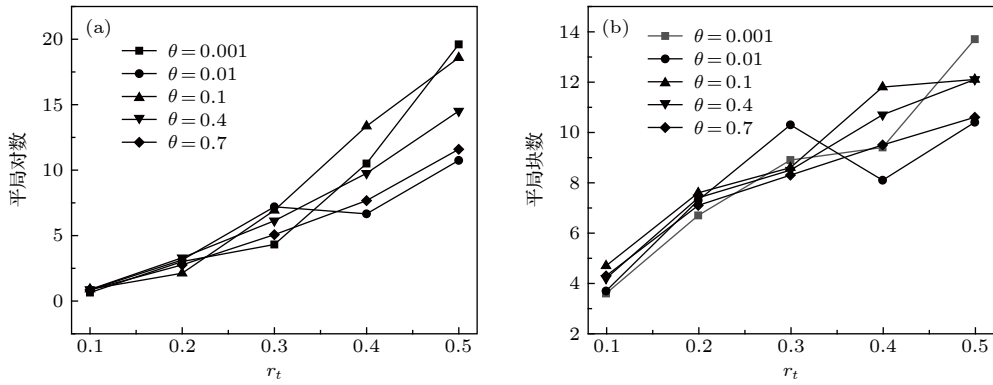


图 3 (a) r_t 与平均对数数量的关系; (b) r_t 与平均块数量的关系

Fig. 3. (a) The relationship between r_t and the number of ties; (b) the relationship between r_t and the block number of ties.

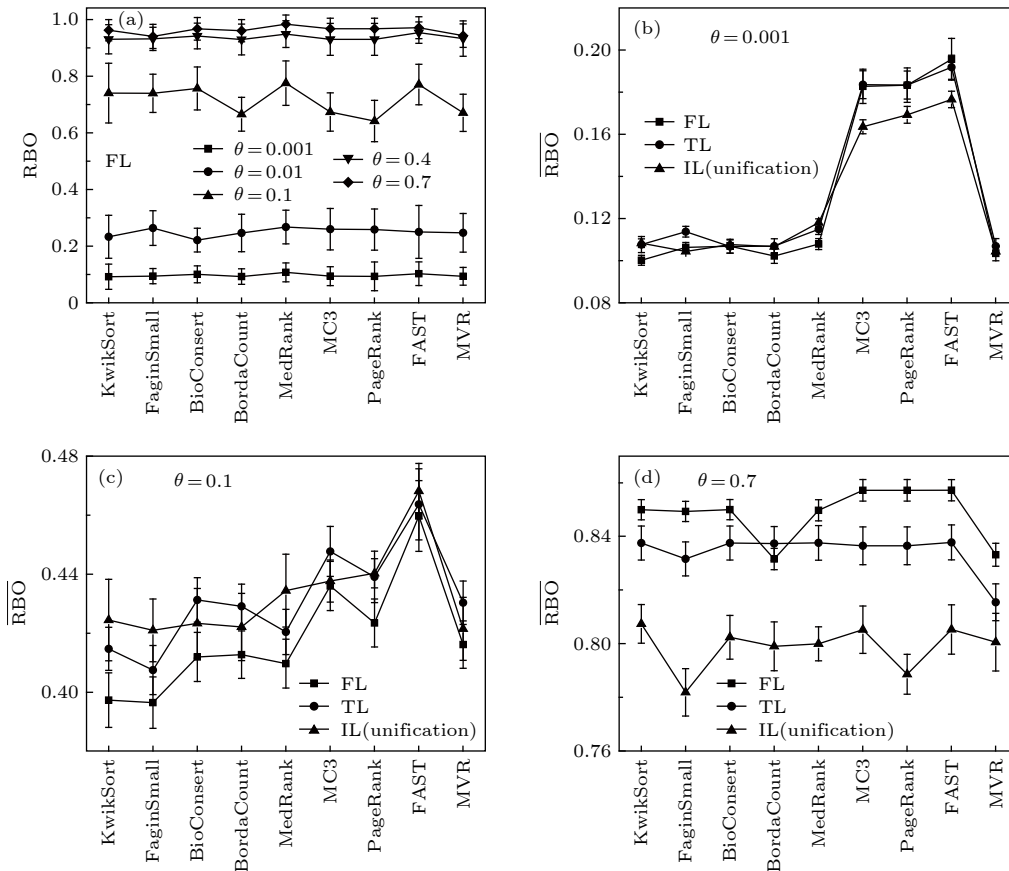


图 4 θ 值对算法表现的影响

Fig. 4. Impact of the degree of consistency θ on the algorithm performance.

10 次实验, 以确保结果更稳定. 由于 Xiao 等^[30] 探究了在不同 r_k 值下, 列表长度不等对排名聚合方法表现的影响, 故本文不再考察 r_k 和 Δk 组合对聚合方法的影响. 对于个别不能直接处理 IL 数据的算法, 我们将 IL 数据进行了规范化 (unification) 处理, 将其变换为算法可以处理的数据类型, 再将结果与其他算法结果对比. 在 7.4 节中将具体讨论规范化处理对结果的影响.

7.1 θ 值对各方法表现的影响

整体而言, 算法的表现主要由数据的发散程度决定. 初始数据一致程度越高 (θ 越大), 列表的发散程度越低, 最终的聚合效果也越好 (图 4). 对于 FL 数据类型, 如以聚合排名与中心排名的距离作为衡量标准 (图 4(a)), 各算法之间的表现差异较大, 即使在排名列表一致性较高的情况下, 启发式

算法也有可能不能给出与中心排名相似的结果. 优化类算法 MVR 更适用于例如循环比赛成绩综合等列表长度短、数量多的场景, 在少数长列表的情况下, 表现较差, 往往还不及启发式算法. 当 θ 过小时, 列表非常发散, 任何算法均无法获得与中心排名相似的聚合排名. 如以聚合排名与其他排名的平均距离 \overline{RBO} 作为衡量标准 (图 4(b)—图 4(d)), 在排名列表发散程度较低的情况下, 平局的引入 (由 FL 到 TL) 一定程度上减少了 \overline{RBO} , 但是在发散程度较高的情况下, 平局对于 \overline{RBO} 的影响较小. 同时, 对于 TL 和 IL 数据, 考察聚合排名与中心排名的 RBO 值时, 发现在不同的一致性程度下, 得到了与 \overline{RBO} 类似的结论, 这里只显示了 \overline{RBO} 的结果. 整体而言, 优化类方法 FAST 在不同参数 θ 下, 其表现均优于其他算法.

7.2 N 值对各方法表现的影响

为探究数据一致程度相同的情况下, 列表数量 N 对聚合效果的影响, 我们固定列表特征参数 ($M=100, \theta=0.7, r_t=0.2, r_k=0.8, \Delta k=0.2$), 改变

列表数量 N , 结果如图 5 所示.

对于 FL 数据, 聚合排名与中心排名的相似度基本随列表数量的增加而增加. 这与我们的直观感知一致, 数据量越大, 则生成的排名列表越均匀, 得到的综合排名也越接近于中心排名. 但也存在少数情况, 如波达计数法 (BordaCount) 与 MVR 表现随列表数量增加呈现不规则变化. 对于 TL 和 IL 数据, 聚合排名与中心排名的相似度与列表数量的关系与 FL 数据的结果类似, 不能较好地体现聚合算法的差异, 因此在考察 N 值对算法表现的影响时, 主要考虑聚合排名与其他排名的平均相似度 \overline{RBO} 的变化情况.

对于所有数据类型, 聚合排名与其他排名的平均相似度 \overline{RBO} 随列表数量的变化呈现不规则变化. 我们推测这是由于两方面的原因共同影响. 一方面列表数量的增加使得 N 个排名之间的差异性呈现不规则变化, 即图 5(b)—(d) 中 N 个排名与聚合排名的方差与列表数量呈非单调的关系; 而另一方面列表数量的增加提供更多的数据点, 使得聚合排名更加稳定. 这两个相互竞争的因素导致了平均相似

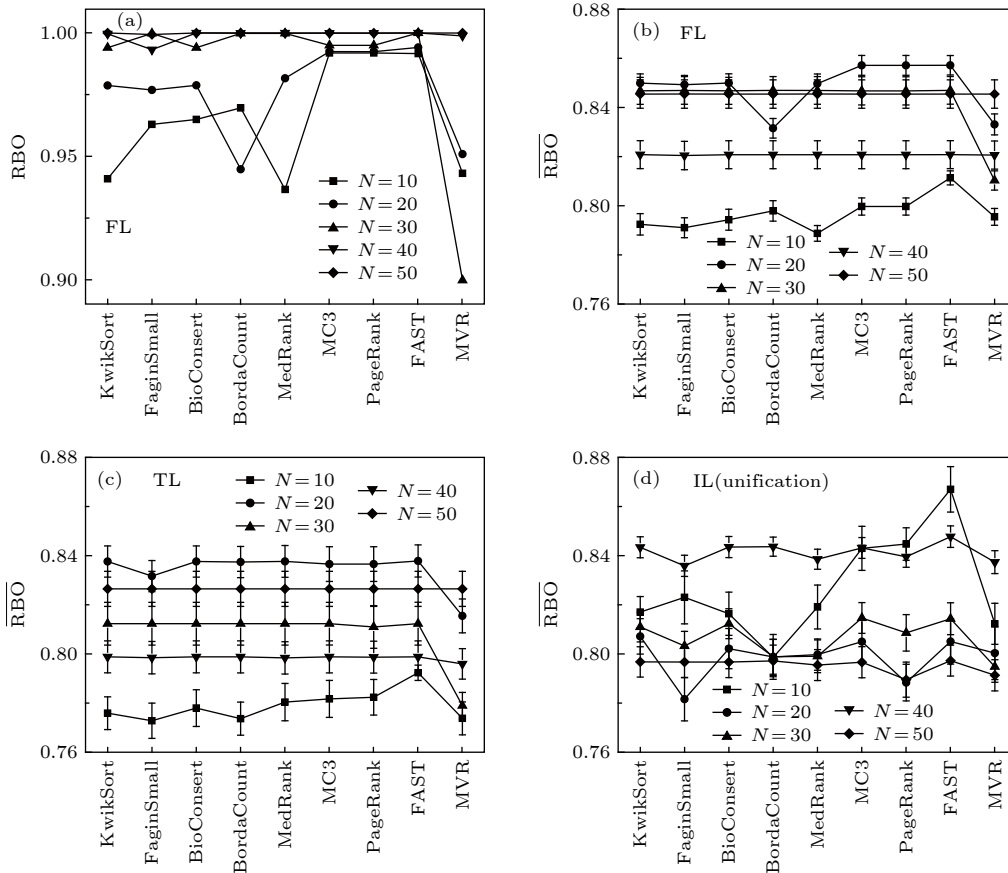


图 5 N 值对算法表现的影响

Fig. 5. Impact of the number of rank lists on the algorithm performance.

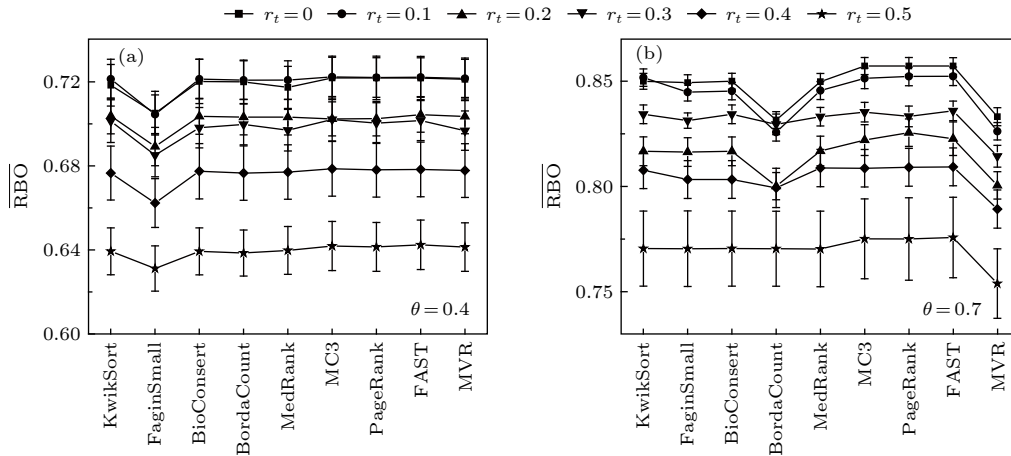


图 6 r_t 值对算法表现的影响
Fig. 6. Impact of the parameter r_t on algorithm performance.

度 \overline{RBO} 与列表数量的非单调关系。

同时, 对于所有列表类型, 随着列表数量的增加, 各聚合算法之间的相对差异也逐渐减少. 整体上优化类方法 FAST 表现也都较为稳定, \overline{RBO} 与 \overline{RBO} 表现普遍更好.

7.3 平局数量对各方法表现的影响

FL 数据和 TL 数据的区别在于 TL 包含平局. 在实际问题中, TL 比 FL 更常见. 我们调节参数 r_t 以分析平局数量对聚合效果的影响, 结果如图 6 所示.

图 6(a) 和图 6(b) 分别表示在 $\theta = 0.4$ 和 0.7 下各聚合方法的最终结果. 当 $r_t = 0$ 时, 对应数据类型是完全列表, 随着 r_t 逐渐增大, 平局块大小及其数量都在增加, 聚合排名与其他排名的平均相似度 \overline{RBO} 逐渐下降, 且波动性增加 (方差增大). 不同算法对平局情况的处理结果存在一定差异, 例如 FaginSmall 方法受整体序列一致性程度影响较大, 而波达计数法 (BordaCount) 在序列整体一致性较高但存在平局时的表现较弱, 考虑到波达计数法被广泛运用于简单的投票统计, 这也一定程度上说明为何一般投票中均没有平局.

7.4 Unification 规范化对算法表现的影响

对于一些不能直接处理 IL 类型数据的排名聚合算法, 如 MedRank, BordaCount, BioConsert 等, 可以通过对数据的规范化处理, 将 IL 数据变为 TL 数据, 使得这些算法可以发挥作用. 一般而言有两种规范化处理方法: Projection 和 unification. Projection 只考虑同时出现在所有列

表中的对象, 将未同时出现在所有给定列表中的对象从每个列表中删除. Unification 将未出现在本列表中的对象全部放到一个块中, 置于该列表底部, 形成一个多对象平局. 例如 2 节示例中的块序列 3—5 经过 unification 处理后变为:

- 块序列 3: [A], [C],
- 块序列 3 (unification): [A], [C], [B, D];
- 块序列 4: [A], [B], [D],
- 块序列 4 (unification): [A], [B], [D], [C];
- 块序列 5: [B], [C, A],
- 块序列 5 (unification): [B], [C, A], [D].

在 7.1 节与 7.2 节的性能比较中, 为了使所有算法均可以参与比较, 我们对排名列表进行了 unification 处理, 将不同算法应用到 unification 后的数据进行比较. 对于不能处理 IL 数据的算法, 利用 unification 处理数据是合理且必须的, 但是对于可以处理 IL 数据的算法, unification 可能会带来算法性能的高估或低估. 从另一个角度来看, 初始数据决定着可用信息量, 而经过规范化或简单推理以后, 实际上增加了一些人为规则信息, 必然会影响最终的聚合排名结果. 为探究 unification 规范化过程对算法表现的影响, 我们选用可以处理 IL 数据的优化类方法 FAST 和启发式方法 PageRank, 分析在处理前和处理后的算法性能变化.

为消除列表不等长对结果的影响, 令 $\Delta k = 0$, 改变参数 r_k , 比较 PageRank 和 FAST 不执行规范化和执行规范化的结果 (图 7). 结果发现, 聚合排名与其他排名的 \overline{RBO} 均值随列表长度呈不规则变化, 同时, 不执行规范化的结果总是优于执行规范化的结果, 这也说明 7.1 节和 7.2 节中的结果,

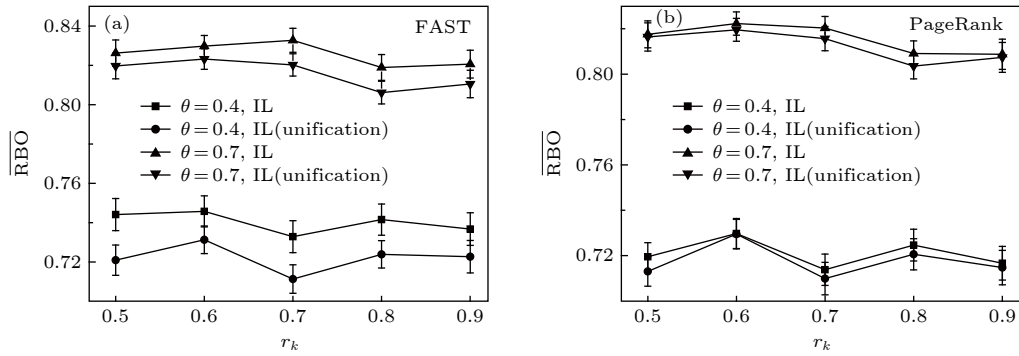


图 7 Unification 对算法表现的影响

Fig. 7. Impact of unification on algorithm performance.

一定程度上低估了如 FAST 这类可以直接处理 IL 数据的算法, 进一步说明了 FAST 算法在整体表现上的优越性. 与此同时, 尽管是否进行 unification 所带来的 \overline{RBO} 值差异不大, 但是对于需要更精细结果的应用而言, 谨慎处理初始数据和规范化也非常重要.

8 结 论

使用 Mallows 模型生成 FL 排名数据, 并提出一套 TL 和 IL 数据生成机制以生成人工可控的排名数据, 同时, 结合有偏等级重叠指标来探究数据特征对不同排名聚合方法的影响. 通过本文提出的数据生成机制, 可以生成具有不同统计特征的可控数据集. 实验结果表明, 启发式方法虽然简单, 但是在排序列表相似度较高、列表相对简单的情况下, 能够接近甚至超过一些优化类方法的结果, 例如在完全列表情况下, 波达计数法简单可行, 最终结果也基本可以接受. 整体而言, 基于距离优化的分支定界方法 (FAST) 优于其他各类算法, 在不同类型的排序列表中表现非常稳定, 能够很好地完成少量长列表的排名聚合.

参考文献

[1] Liao H, Mariani M S, Medo M, Zhang Y C, Zhou M Y 2017 *Phys. Rep.* **689** 154
 [2] Liu J G, Ren Z M, Guo Q, Wang B H 2013 *Acta Phys. Sin.* **62** 178901 (in Chinese) [刘建国, 任卓明, 郭强, 汪秉宏 2013 物理学报 **62** 178901]
 [3] Pujari M, Kanawati R 2012 *Proceedings of the 21st International Conference on World Wide Web* Lyon, France, April 16–20, 2012 p11
 [4] Tabourier L, Libert A S, Lambiotte R 2016 *EPJ Data Sci.* **5** 1
 [5] Snell J L, Kemeny J G 1962 *Mathematical Models in the Social Sciences* (Boston: Introduction to Higher Mathematics)

pp3–23
 [6] Davenport A J, Kalagnanam J 2004 *Conference on 19th National Conference on Artificial Intelligence* San Jose, USA, July 25–29, 2004 p697
 [7] Amodio S, D’ambrosio A, Siciliano R 2016 *Eur. J. Oper. Res.* **249** 667
 [8] Meila M, Phadnis K, Patterson A, Bilmes J 2012 arXiv: 1206.5265 [cs.LG]
 [9] Baskin J P, Krishnamurthi S 2009 *Proceedings of the Third ACM Conference on Recommender Systems* New York, USA, October 23–25, 2009 p337
 [10] Lü L Y, Medo M, Yeung C H, Zhang Y C, Zhang Z K, Zhou T 2012 *Phys. Rep.* **519** 1
 [11] Dwork C, Kumar R, Naor M 2001 *Proceedings of the 10th International Conference on World Wide Web* Hong Kong, May 1–5, 2001 p613
 [12] Cook W D, Raviv T A L, Richardson A J 2010 *Accounting Perspectives* **9** 217
 [13] Cook W D, Golany B, Penn M 2007 *Comput. Oper. Res.* **34** 954
 [14] Guo C H, Li M Q 2018 *Data Analysis and Knowledge Discovery* **2** 10 (in Chinese) [郭崇慧, 李敏谦 2018 数据分析与知识发现 **2** 10]
 [15] Jia T, Wang D, Szymanski B K 2017 *Nat. Hum. Behav.* **1** 0078
 [16] Zhang H X, LÜ Z, Zhang C T 2018 *Journal of University of Electronic Science and Technology of China* **47** 112 (in Chinese) [张海霞, 吕振, 张传亭 2018 电子科技大学学报 **47** 112]
 [17] Jia T, Xia F 2019 *Big Data Res.* **04** 38 (in Chinese) [贾韬, 夏锋 2019 大数据 **04** 38]
 [18] Wang X, Ran Y, Jia T 2020 *Chaos: An Interdisciplinary Journal of Nonlinear Science* **30** 013101
 [19] Liu W, Wang Y B 2011 *Acta Phys. Sin.* **60** 070301 (in Chinese) [刘文, 王永滨 2011 物理学报 **60** 070301]
 [20] Han Z M, Wu Y, Tan X S, Duan D G, Yang W J 2015 *Acta Phys. Sin.* **64** 58902 (in Chinese) [韩忠明, 吴杨, 谭旭升, 段大高, 杨伟杰 2015 物理学报 **64** 58902]
 [21] Borda J C de 1781 *Histoire de l’Academie Royale des Sciences* 657
 [22] Langville A N, Meyer C D 2012 *Who’s # 1?: The Science of Rating and Ranking* (Princeton: Princeton University Press) pp159–231
 [23] Cook W D 2006 *Eur. J. Oper. Res.* **172** 369
 [24] Lin S 2010 *Wiley Interdiscip. Rev. Comput. Stat.* **2** 555
 [25] Ali A, Meila M 2012 *Math. Soc. Sci.* **64** 28
 [26] Schalekamp F, Zuylen A 2009 *Proceedings of the Eleventh*

- Workshop on Algorithm Engineering and Experiments* New York, USA, January 3, 2009 p38
- [27] Brancotte B, Yang B, Blin G, Cohen B S, Denise A, Hamel S 2015 *Proceedings of the VLDB Endowment* **8** 1202
- [28] Fagin R, Kumar R, Sivakumar D 2003 *SIAM J. Discrete Math.* **17** 134
- [29] Cohen-boulakia S, Denise A, Hamel S 2011 *International Conference on Scientific and Statistical Database Management* Portland, USA, July 20–22, 2011 p73
- [30] Xiao Y, Deng Y, Wu J 2017 *Nav. Res. Logist.* **64** 556
- [31] Fagin R, Kumar R, Mahdian M 2004 *Proceedings of the Twenty-third ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* Paris, France, June 14–16, 2004 p47
- [32] Li X, Wang X, Xiao G 2017 *Brief. Bionform.* **20** 178
- [33] Sculley D 2007 *Proceedings of the 2007 SIAM International Conference on Data Mining* Minneapolis, Minnesota, USA, April 26–28, 2007 p587
- [34] Deng K, Han S, Li K J 2014 *J. Am. Stat. Assoc.* **109** 1023
- [35] Liu Y T, Liu T Y, Qin T 2007 *Proceedings of the 16th international conference on World Wide Web* Banff, Alberta, Canada, May 8–12, 2007 p481
- [36] Freund Y, Iyer R, Schapire R E 2003 *J. Mach. Learn. Res.* **4** 933
- [37] Ailon N, Charikar M, Newman A 2008 *J. ACM* **55** 23
- [38] Van Zuylen A, Williamson D P 2009 *Math. Oper. Res.* **34** 594
- [39] Kendall M G 1948 *Rank correlation methods* (London: Griffin)
- [40] Diaconis P, Graham R L 1977 *J. R. Stat. Soc. B* 262
- [41] Fagin R, Kumar R, Mahdian M, Sivakumar D, Vee E 2006 *SIAM J. Discrete. Math.* **20** 628
- [42] Fagin R, Kumar R, Sivakumar D 2003 *Proceedings of the 2003 ACM SIGMOD International Conference on Management of data* San Diego, California January 9–12, 2003 p301
- [43] Brin S, Page L 1998 *Comput. Networks ISDN Syst.* **30** 107
- [44] Adali S, Hill B, Magdon-Ismael M 2007 *J. Digital Information Management(JDIM)* **5** 292
- [45] Emond E J, Mason D W 2002 *J. Multi-Crit. Decis. Anal.* **11** 17
- [46] Ailon N 2010 *Algorithmica* **57** 284
- [47] Lin S, Ding J 2009 *Biometrics* **65** 9
- [48] Heiser W J, D’ambrosio A (edited by Lausen B, Dirk V P) 2013 *Algorithms from and for Nature and Life* (New York: Springer) pp19–31
- [49] Pedings K E, Langville A N, Yamamoto Y 2012 *Optim. Eng.* **13** 349
- [50] Bar-Ilan J, Mat-Hassan M, Levene M 2006 *Comput. Networks* **50** 1448
- [51] Lin Z W, Yi L, Guo X L 2017 arXiv: 1704.08464 [cs.AI]
- [52] Ekstrom C T, Gerds T A, Jensen A K 2018 *Biostatistics* **20** 582
- [53] Kumar R, Vassilvitskii S 2010 *Proceedings of the 19th International Conference on World Wide Web* Raleigh, North Carolina, USA, April 26–30, 2010 p571
- [54] Sakai T, Nicola F 2014 *Metrics, Statistics, Tests in: Bridging between Information Retrieval and Databases* (Heidelberg: Springer) pp116–163
- [55] Webber W, Moffat A, Zobel J 2010 *ACM T. Inform. Syst.* **28** 1
- [56] Mallows C L 1957 *Biometrika* **44** 141
- [57] Critchlow D E, Fligner M A, Verducci J S 1991 *J. Math. Psychol.* **35** 294
- [58] Irurozki E, Calvo B, Lozano J A 2016 *J. Stat Softw.* **71** 1575
- [59] Fligner M A, Verducci J S 1986 *J. R. Stat. Soc. B* 359
- [60] Smith B B 1950 *J. R. Stat. Soc. B* 12 41
- [61] Luce R D 2012 *Individual Choice Behavior: A Theoretical Analysis* (North Chelmsford: Courier Corporation) pp68–70
- [62] Thurstone L L 1927 *Psychol. Rev.* **34** 273

SPECIAL TOPIC—Statistical physics and complex systems

Comparison of performance of rank aggregation algorithms in aggregating a small number of long rank lists^{*}

Chen Wen-Yu¹⁾ Zhu Zhang-Qian²⁾ Wang Xiao-Meng¹⁾ Jia Tao^{1)†}¹⁾ (*College of Computer & Information Science, Southwest University, Chongqing 400715, China*)²⁾ (*Department of National Defense Economy, Army Logistics University of Chinese People's**Liberation Army, Chongqing 500106, China*)

(Received 17 October 2019; revised manuscript received 24 November 2019)

Abstract

Rank aggregation aims to combine multiple rank lists into a single one, which has wide applications in recommender systems, link prediction, metasearch, proposal selection, and so on. Some existing studies have summarized and compared different rank aggregation algorithms. However, most of them cover only a few algorithms, the data used to test algorithms do not have a clear statistical property, and the metric used to quantify the aggregated results has certain limitations. Moreover, different algorithms all claim to be superior to existing ones when proposed, the baseline algorithms, the testing samples, and the application scenario are all different from case to case. Therefore, it is still unclear which algorithm is better for a particular task. Here we review nine rank aggregation algorithms and compare their performances in aggregating a small number of long rank lists. We assume an algorithm to generate different types of rank lists with known statistical properties and cause a more reliable metric to quantify the aggregation results. We find that despite the simplicity of heuristic algorithms, they work pretty well when the rank lists are full and have high similarities. In some cases, they can reach or even surpass the optimization-based algorithms in performance. The number of ties in the list will reduce the quality of the consensus rank and increase fluctuations. The quality of aggregated rank changes non-monotonically with the number of rank lists that need to be combined. Overall, the algorithm FAST outperforms all others in three different rank types, which can sufficiently complete the task of aggregating a small number of long rank lists.

Keywords: rank aggregation, incomplete rank list, Mallows model, rank-biased overlap**PACS:** 02.50.-r, 89.75.-k, 87.15.Qt, 87.15.nr**DOI:** 10.7498/aps.69.20191584

* Project supported by the National Natural Science Foundation of China (Grant No. 61603309).

† Corresponding author. E-mail: tjia@swu.edu.cn

人脑默认模式网络的动力学行为

姚楠 苏春旺 李尤君 王珏 周昌松 黄子罡

Dynamics of the default mode network in human brain

Yao Nan Su Chun-Wang Li You-Jun Wang Jue Zhou Chang-Song Huang Zi-Gang

引用信息 Citation: *Acta Physica Sinica*, 69, 080203 (2020) DOI: 10.7498/aps.69.20200170

在线阅读 View online: <https://doi.org/10.7498/aps.69.20200170>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

静息态功能磁共振成像评估健康老年人认知行为的多尺度熵模型研究

Study of Multiscale Entropy Model to Evaluate the Cognitive Behavior of Healthy Elderly People Based on Resting State Functional Magnetic Resonance Imaging

物理学报. 优先发表 <https://doi.org/10.7498/aps.69.20200051>

基于复杂网络动力学模型的无向加权网络节点重要性评估

Evaluation methods of node importance in undirected weighted networks based on complex network dynamics models

物理学报. 2018, 67(9): 098901 <https://doi.org/10.7498/aps.67.20172295>

利用神经网络识别高分子链在表面的吸附相变

Recognition of adsorption phase transition of polymer on surface by neural network

物理学报. 2019, 68(20): 200701 <https://doi.org/10.7498/aps.68.20190643>

在具有排斥耦合的神经网络中有序斑图的熵测量

Entropy measurement of ordered patterns in neuronal network with repulsive coupling

物理学报. 2019, 68(11): 110503 <https://doi.org/10.7498/aps.68.20190231>

非一致通信时滞动力学网络上的接连滞后同步

Successive lag synchronization on dynamical networks with non-uniform communication delays

物理学报. 2018, 67(1): 018901 <https://doi.org/10.7498/aps.67.20171630>

专题：统计物理和复杂系统

人脑默认模式网络的动力学行为*

姚楠^{1) #} 苏春旺^{2) 3) #} 李尤君^{2) †} 王珏²⁾ 周昌松⁴⁾ 黄子罡^{2) ‡}

1) (西安理工大学理学院, 西安 710048)

2) (西安交通大学生命学院, 生物医学信息工程教育部重点实验室, 西安 710049)

3) (兰州大学物理科学与技术学院, 兰州 730000)

4) (香港浸会大学物理系, 非线性研究中心, 计算及理论研究所, 香港)

(2020年2月2日收到; 2020年3月22日收到修改稿)

大脑具有自适应、自组织、多稳态等重要特征, 是典型的复杂系统. 人脑在静息态下的关键功能子网络——默认模式网络 (DMN) 的激活处于多状态间持续跳转的非平衡过程, 揭示该过程背后的动力学机制具有重要的科学意义和临床应用前景. 本文基于功能磁共振获得的血氧水平依赖 (BOLD) 信号, 建立了 DMN 吸引子跳转非平衡过程的能量图景、吸引子非联通图、跳转关系网络等; 以高级视觉皮层和听觉等皮层活动为例, 通过对应激活 DMN 状态空间的分布, 以及 XGBoost、深度神经网络等算法验证了 DMN 状态变化与外部脑区状态的密切依赖关系; 通过偏相关、收敛交叉映射等方法分析了 DMN 内各个脑区之间的相互作用. 本文结果有助于理解静息态下大脑内在非平衡过程的动力学机制, 以及从动力学的角度探索具有临床意义的脑功能障碍生物标志物.

关键词: 脑功能网络, 静息态, 默认模式网络, 能量图景, 状态跳转**PACS:** 02.50.-r, 64.60.aq, 87.18.-h**DOI:** 10.7498/aps.69.20200170

1 引言

近年来, 非侵入式脑功能成像技术的发展促使脑科学研究进入了新的阶段^[1]. 其中, fMRI (functional magnetic resonance imaging) 可通过采集大脑的低频 (< 0.1 Hz) 血氧水平依赖 (blood oxygen level dependent, BOLD) 信号, 来反映各个脑区在特定任务态或静息态下的动态激活状态. 因无创和高空间分辨率等优势, fMRI 被广泛应用于脑的工作机制和脑功能障碍的研究^[2-5]. 基于

fMRI 对大脑在任务态和静息态等状态下激活模式的研究发现了大脑的多种功能子网络, 包括默认模式网络 (default mode network, DMN)、突显网络 (salience network, SN)、执行控制网络 (executive control network, ECN)、注意网络 (attention network, AN)、视觉网络 (visual network, VN)、听觉网络 (auditory network, Aud) 等^[6].

DMN 的正常运转对大脑功能的维持具有重要的意义. 2001 年, Raichle 等^[7] 提出了大脑的“默认模式” (default mode) 的概念, 其对应的脑区在任务态或受到明显外部刺激时被抑制, 而在静息态时

* 国家自然科学基金 (批准号: 11975178, 11647052)、认知神经科学与学习国家重点实验室 (批准号: CNLYB1802)、陕西省自然科学基金基础研究计划 (批准号: 2018JQ1010, 2020JM-058, 2020JQ-096)、陕西省教育厅科研计划 (批准号: 17JK0553) 和陕西省高校科协青年人才托举计划 (批准号: 20170606) 资助的课题.

共同第一作者.

† 通信作者. E-mail: liyoujun1@xjtu.edu.cn

‡ 通信作者. E-mail: huangzg@xjtu.edu.cn

恢复并维持激活,“默认模式”被认为是大脑神经系统的基线状态. 2003年, DMN 被明确提出^[8], 它是维持该基线状态最核心的功能子网络, DMN 网络内不同脑区的静息态神经活动具有特定的时间相干性. 此外, 静息态下脑区之间的功能连接也被发现与对应的白质纤维束物理投射关系具有一定程度的一致性^[9]. DMN 的主要特征是在静息态下呈现出连续的高水平激活, 而在任务态时受到抑制^[10], 已经证实它与维持觉醒和监测外部刺激等功能有关^[11]. 在 DMN 的概念提出后, 更多与大脑基线水平, 即默认模式相关的脑区及脑区的具体功能被逐渐发现 (见图 1^[12]). 目前认为 DMN 主要包括: 腹内侧前额叶皮质 (Ventral Medial Prefrontal Cortex, VMPC)、背内侧前额叶皮质 (Dorsal Medial Prefrontal Cortex, DMPC)、后扣带回皮质 (Posterior Cingulate Cortex, PCC)、楔前叶 (Precuneus) 和横向顶叶皮质 (Lateral Parietal Cortex, LPC).

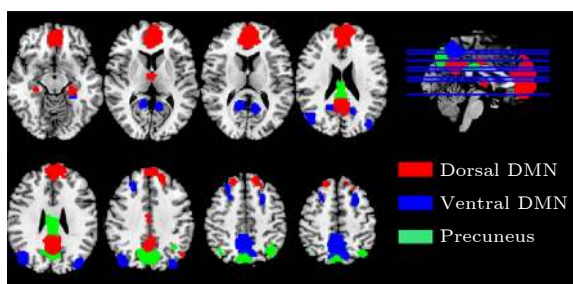


图 1 大脑默认模式网络 (DMN)^[12]

Fig. 1. DMN of the brain^[12].

DMN 是大脑基线状态的核心功能网络, 一些常见的脑功能障碍往往同时对应 DMN 激活模式及功能连接的异常, 这也使得 DMN 活动的异常指标成为进行脑疾病的诊断和评估的潜在生物标志物^[13-15], 如与正常衰老相比, 早期阿尔兹海默症 (AD) 患者的 DMN 中就存在显著的异常激活^[16], 轻度认知障碍 (ADHD) 患者在任务态下表现出 DMN 去激活率明显降低^[17] 和 DMN 内网络相关性降低^[18], 精神分裂症的阳性症状严重程度直接与 DMN 激活模式及功能连接的异常相关^[19,20], 脑外伤导致的轻度脑损伤患者的 DMN 则无法在任务态实现去激活, 显著影响任务表现^[21].

近年来, 对静息态脑活动的研究发现, DMN 单个脑区的激活看似随机, 但从网络的层次却有规律可循. 静息态下 DMN 的状态存在更细致的子状态, 对应于不同的激活模式, 系统在多种激活模式

间持续跳转, 是典型的非平衡过程, 其表现出的规律及背后的机理适合运用统计物理和非线性动力学的方法进行研究. 然而, 在这一方向上更进一步的深入研究到目前为止还尚未见到进展, 例如, 静息态 DMN 的子状态是否和其他脑区的功能活动有对应关系等, 需要做进一步的考察. 此外, 由于脑功能障碍对应的 DMN 激活模式及跳转行为与健康状态存在显著差异^[13,19,20,22], 对静息态下 DMN 所处非平衡过程的研究具有重要的科学意义和临床应用前景, 能够帮助我们理解人脑静息态的动力学特征和工作机制, 从功能连接的角度揭示脑功能障碍的发病机理, 并为相关疾病的预防、诊断和治疗提供启发^[22,23,24].

鉴于上述原因, 本文将基于静息态 DMN 的 BOLD 数据, 围绕 DMN 的动力学行为及其背后的机理展开研究, 所采用的方法主要包括能量图景构造^[25,26]、共变行为分析、基于压缩感知的偏相关分析^[27] 与收敛交叉映射方法^[28] 等. 研究的问题包括: DMN 激活行为的非平衡过程的主要特征, DMN 与外部其他脑区活动的交互, DMN 内部各脑区之间的相互作用, 以及对其非平衡过程动力学机理的探讨.

2 方法及结果

2.1 方法

2.1.1 数据及预处理

本文的研究采用静息态功能磁共振成像 (rs-fMRI) 数据来源于人类连接组计划 (human connectome project, HCP) 网站的公开数据集, fMRI 空间分辨率为 $2\text{ mm} \times 2\text{ mm} \times 2\text{ mm}$ (单个体素尺寸), 时间分辨率为 0.72 s (采样时间间隔), 包括 55 个被试, 每个被试的采样点数为 $T = 1200$, 采样持续时间为 $0.72\text{ s} \times 1200 = 14\text{ min}$. 原始影像数据的预处理流程为: 功能与结构图像配准, 配准到标准空间, 去除白质和脑脊液信号, 头动校正, 带通滤波 ($0.01-0.10\text{ Hz}$), 空间高斯平滑, 去漂移等. 经上述预处理后, 再根据脑区划分模板将默认模式网络相关的区域划分为 23 个感兴趣区 (regions of interest, ROIs), 并通过 ROI 内所有体素时间序列求平均, 得到每个 ROI 的 BOLD (血氧水平依赖) 信号时间序列. 23 个 ROI 的相关基本信息见表 1.

表 1 默认模式网络 ROI 的名称、标签、Brodmann 分区 (BA) 编号及位置坐标信息

Table 1. Name, label, Brodmann area (BA) number, and location information of ROIs belonging to the default mode network.

ROI	Label	L/R	BA	X	Y	Z
Posterior Cingulate	vDMN_1	L	31	-12	-62	10
Middle Frontal Gyrus	vDMN_2	L	10	-27	-6	59
Culmen	vDMN_3	L	37	-30	-39	-20
Superior Occipital Gyrus	vDMN_4	L	19	-36	-88	28
Posterior Cingulate Gyrus	vDMN_5	R	31	15	-56	13
Precuneus	vDMN_6		7	-6	-61	56
Middle Frontal Gyrus	vDMN_7	R	10	24	26	47
Culmen	vDMN_8	R	37	27	-33	-23
Angular Gyrus	vDMN_9	R	39	43	-79	28
Cerebellum	vDMN_10	R		12	-47	-63
Ventral Posterior Cingulate Gyrus	pDMN_1		23	0	-35	28
Precuneus	pDMN_2		7	0	-76	38
Inferior Parietal Lobule	pDMN_3	L	40	-39	-64	46
Inferior Parietal Lobule	pDMN_4	R	40	39	-64	46
Middle Frontal Gyrus	dDMN_1		9	0	49	12
Angular Gyrus	dDMN_2	L	39	-48	-73	32
Superior Frontal Gyrus	dDMN_3	R	6	18	38	51
Dorsal Posterior Cingulate Gyrus	dDMN_4		31	0	-57	30
Ventral Anterior Cingulate Gyrus	dDMN_5		24	0	-17	35
Angular Gyrus	dDMN_6	R	39	48	-66	29
Thalamus	dDMN_7			-6	-6	3
Parahippocampal Gyrus	dDMN_8	L	36	-24	-37	-9
Parahippocampal Gyrus	dDMN_9	R	36	24	-21	-23

2.1.2 DMN 的能量图景构造方法

以背侧默认模式网络 (dDMN, 为 DMN 的核心区域) 作为代表, 将其 9 个 ROI 的 BOLD 信号根据某阈值 (一般取均值线或以均值线为中心的窄带) 二值化为 0 或 1 序列, 分别对应于不活跃状态和活跃状态, 如图 2 所示, 编号为 i 的 ROI 的状态为 $\sigma_i = 0, 1 (i = 1, 2, \dots, 9)$, 那么 t 时刻 dDMN 状态可表示为 $\sigma = [\sigma_1, \sigma_2, \dots, \sigma_9]$. 系统状态 σ 所对应的“能量”定义为

$$E(\sigma) = -\frac{1}{2} \sum_{i \neq j} J_{ij} \sigma_i \sigma_j - \sum_i h_i \sigma_i,$$

其中 $\mathbf{J} \in \mathbb{R}^{9 \times 9}$ 和 $\mathbf{h} \in \mathbb{R}^9$ 是模型的两组参数, $J_{ii} = 0 (i = 1, 2, \dots, 9)$. 以上能量函数给出系统状态的玻尔兹曼分布:

$$P(\sigma) = \frac{1}{Z} e^{-E(\sigma)},$$

配分函数为

$$Z = \sum_{\sigma} e^{-E(\sigma)}.$$

在真实 rs-fMRI 数据 dDMN 脑区激活占比的约束条件下, 通过梯度上升算法, 不断调节 (即训练) 模型的两个参数 \mathbf{J} 和 \mathbf{h} , 使得玻尔兹曼分布的熵

$$S = -\sum_{\sigma} P(\sigma) \log P(\sigma)$$

最大化 (训练的约束条件为 $\langle \sigma_i \rangle_{\text{data}} = \sum_{\sigma} (\sigma_i) P(\sigma)$, $\langle \sigma_i \sigma_j \rangle_{\text{data}} = \sum_{\sigma} \sigma_i \sigma_j P(\sigma)$, $i, j = 1, 2, \dots, 9, i \neq j$), 即得到优化的模型. 根据优化的模型可得系统的能量图景.

局部极小态及能量壁垒的计算: 将 dDMN 状态矢量中只有一位存在差异的两个状态视为相邻状态, 因此每个 dDMN 状态都有 9 个相邻状态. 若某状态的能量小于所有 9 个相邻状态, 则称为局部极小态.

局部极小态的非联通图构建: 1) 首先建立状态矢量构成的超晶格, 其中每个状态都与其相邻状态相连; 2) 设定能量阈值, E_{th} , 它等于当前所有状

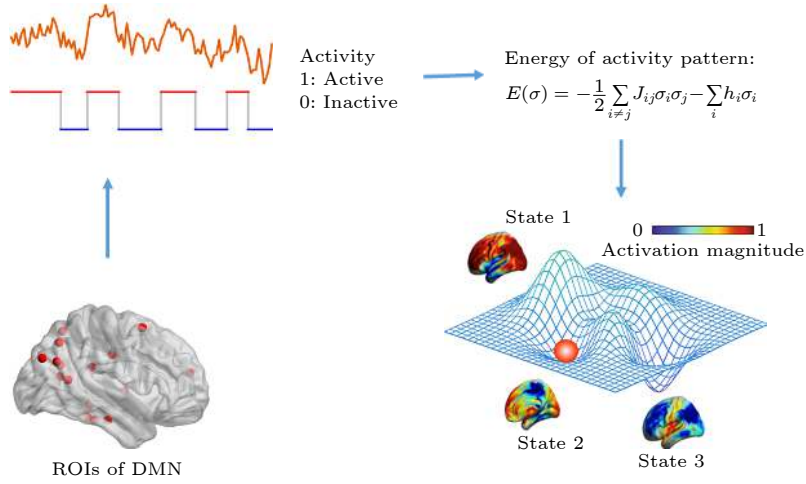


图 2 DMN 能量图景构造方法示意图

Fig. 2. Schematic diagram of the construction of DMN energy landscape.

态中最大的能量; 3) 移除所有能量 $\geq E_{th}$ 的状态及其连边; 4) 检测超晶格中是否每个局部极小态都有至少一条路径相互联通; 5) 重复步骤 3) 和 4), 将 E_{th} 设为剩余格点中的能量最大值, 直到所有的局部极小态都互不联通. 记录下每两个局部极小态第一次不连通时所设定的 E_{th} , 作为两者之间的势垒高度, 即可构建局部极小态的非联通图.

2.1.3 基于压缩感知的偏相关方法

基于脑区 BOLD 信号的相关性分析是构建脑功能网络的常用方法. 但相关性分析所得两变量 (如 DMN 内部某两个 ROI) 间关系无法排除来自其他变量 (即除这两个 ROI 之外的 DMN 内部其他 ROI) 的干扰, 而偏相关分析则可以在一定程度上去除其他变量的影响. 压缩感知 (包括 LASSO) 利用系统稀疏性的先验知识辅助求解线性回归问题, 在小样本数据集的情况下优势更加突出, 甚至可解线性欠定方程组. 脑功能网络具有稀疏性这一基本特征, 而且核磁共振数据的时间序列长度一般都比较短, 利用压缩感知辅助进行偏相关分析构建脑功能网络, 不仅可以凸显网络中的重要边、自动剔除弱的伪连边, 而且还可以对基于短时间序列的相关性 (或线性回归) 计算的准确性起到一定的补偿. 以下将简要介绍基于压缩感知的偏相关方法^[27].

假设 $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_p$ 是所测 p 个对象的数据矢量 (时间序列), 长度为 n , 且已经被中心化 (centered, 即 $\mathbf{f}_i^T \mathbf{1}_n = 0$, 其中 $\mathbf{1}_n \in \mathbb{R}^{n \times 1}$ 为所有元素均为 1 的 n 维列矢量)、标准化 (normalized, 即 $\mathbf{f}_i^T \mathbf{f}_i = 1$). 那么协方差矩阵为

$$\Sigma = [\sigma_{ij}] = \left[\frac{\mathbf{f}_i^T \mathbf{f}_j}{n-1} \right] \in \mathbb{R}^{p \times p},$$

协方差矩阵的逆矩阵 (称为精度矩阵, precision matrix) 为

$$\Pi = \Sigma^{-1} = [\pi_{ij}] \in \mathbb{R}^{p \times p},$$

则 \mathbf{f}_i 与 \mathbf{f}_j 之间的相关系数和偏相关系数分别为

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}} = \mathbf{f}_i^T \mathbf{f}_j, \quad \theta_{ij} = -\frac{\pi_{ij}}{\sqrt{\pi_{ii}\pi_{jj}}}.$$

如果对 \mathbf{f}_i 作如下线性回归:

$$\mathbf{f}_i = \alpha_{ij} \mathbf{f}_j + \epsilon_i,$$

那么回归系数 α_{ij} 就等于 \mathbf{f}_i 与 \mathbf{f}_j 之间的相关系数 ρ_{ij} . 而如果对 \mathbf{f}_i 做如下线性回归:

$$\mathbf{f}_i = \sum_{j \neq i} \beta_{ij} \mathbf{f}_j + \epsilon_i,$$

那么回归系数 β_{ij} 和 \mathbf{f}_i 与 \mathbf{f}_j 之间偏相关系数 θ_{ij} 的关系为

$$\theta_{ij} = \beta_{ij} \sqrt{\frac{\pi_{ii}}{\pi_{jj}}} = \beta_{ij} \sqrt{\frac{\text{var}(\epsilon_j)}{\text{var}(\epsilon_i)}},$$

其中 $\text{var}(\epsilon)$ 表示回归残差 ϵ 的方差.

在回归系数矢量 β 具有比较强的先验稀疏性 (在脑功能网络中这是一个合理的假设) 的情况下, 可被如下 LASSO 回归的方法很好地估计:

$$\hat{\beta}_{ij} = \underset{\beta_{ij}}{\text{argmin}} \sum_{i=1}^p \left\| \mathbf{f}_i - \sum_{j \neq i} \beta_{ij} \mathbf{f}_j \right\|_2^2 + \lambda \sum_{i,j} |\beta_{ij}|.$$

在通过上述方法得到 β 之后, 即可计算偏相关系数,

$$\theta_{ij} = \beta_{ij} \sqrt{\frac{\text{var}(\epsilon_j)}{\text{var}(\epsilon_i)}}$$

其中

$$\epsilon_i = f_i - \sum_{j \neq i} \beta_{ij} f_j.$$

2.1.4 收敛交叉映射方法 (CCM)

鉴于大脑活动的典型非线性特征及 fMRI 所得 BOLD 数据采集时长有限两方面因素, 我们还将采用收敛交叉映射法 (convergent cross mapping, CCM)^[28] 进行动力学因果性分析. CCM 基于泰肯嵌入定理 (Taken's embedding theorem), 通过相空间重构的方式对变量间的动力学因果性进行推断. 相较于目前广泛应用的格兰杰因果性分析 (Granger causality, GC) 和传递熵 (transfer entropy, TE) 等方法, CCM 的优势在于更加适合非线性系统的处理, 并且对数据量的要求相对较低.

通过 CCM 方法对两个脑区 BOLD 时间序列 X 和 Y 进行因果性分析, 可以得到脑区间有向的动力学影响强度 (directed dynamical influence), 进而构建 DMN 与其他脑区相互作用的脑网络. 基本步骤包括: 由 X 和 Y 时间序列分别进行相空间重构得到流形 M_X 与 M_Y ; 根据 M_X 的结构得到预测序列 \hat{Y} , 求得 \hat{Y} 与实际序列 Y 的皮尔森相关系数 $\rho_{Y|M_X}$, 此系数即是变量 Y 对变量 X 影响程度的度量; 同样地, $\rho_{X|M_Y}$ 是变量 X 对变量 Y 影响程度的度量. X 与 Y 之间的因果方向根据 $R = \rho_{X|M_Y} - \rho_{Y|M_X}$ 的符号进行判断 ($R > 0$ 则表明 X 对 Y 的动力学影响更强, $R < 0$ 则反之).

2.2 结果

2.2.1 DMN 静息态非平衡过程的动力学行为分析

2.2.1.1 DMN 的吸引子间跳转行为特征

基于前面所述方法, 可由每个被试的数据计算得到 dDMN 能量图景. 这里, 高维超晶格上的能量图景的主要特征可通过非连通图 (disconnectivity graph) 显示, 包括吸引子能量大小及相邻最低势垒连通关系等信息, 见图 3(a). 非连通图的构造是在训练得到系统能量函数表达式后, 通过搜索能量极小值, 分析各极小值对应的系统状态吸引子的位

置关系, 最终保留所有吸引子及它们之间的最低翻越势垒构成的树状结构图. 图 3(a)—图 3(c) 上下两行分别对应于被试 1 和被试 2 的结果. 以能量图景吸引域为标签对 DMN 活动时间序列进行着色可对其在状态空间中的分布进行观察, 见图 3(b), 其中左右图为不同时刻的状态激活情况 (为观察方便计, 这里使用了对同一状态轨道不同视角的展示, 即左右图观察角度不同). 三维投影空间根据 23 维数据主成分分析 (PCA) 得到的前三个主成分方向进行构造. 图 3(c) 给出了吸引子间状态跳转关系网络, 其中颜色与图 3(a) 中的吸引子一致, 节点大小等于 DMN 在该吸引域内出现的次数, 节点间有向边的权重表示对应方向的跳转频次, 总时长 $T = 1200$. 比较图 3(a)—图 3(c) 上下两行对应的两个不同被试的结果可以看到, 不同被试在静息态下其默认模式网络激活行为的动力学特征存在明显差异, 例如被试 1 具有较多的吸引子状态, 且存在更频繁的吸引子间切换行为, 而被试 2 的吸引子个数较少, 且跳转行为主要表现为局部切换 (例如 state 1 和 state 2 之间, state 3 和 state 4 之间), 这表明静息态下被试 1 的思维活动更加丰富、活跃, 而被试 2 的思维活动则相对单一且不易变化.

图 3(d)—图 3(i) 展示了 55 个被试能量图景的统计分布特征. 可以看到: 吸引子数目存在明显的峰值 (图 3(d), 峰值位于 4 附近), 表明静息态下吸引子典型数目的普遍存在. 吸引域占据率 (即被访问次数) 呈双峰分布 (图 3(e)), 表明静息态下 DMN 的状态多处在其中某些吸引域而很少在另外的吸引域, 存在两极分化现象. 类似地, 势阱深度 (即势阱最低点到最近鞍点的能量差) 也呈双峰分布 (图 3(f)), 势阱深度和吸引域占据率的正相关关系如图 3(g) 所示. 图 3(h) 给出了吸引域被访问次数 O 和吸引域面积 A (即吸引域内包含的状态总数) 之间具有很好的线性关系 $O = kA$ (其中 k 为常数), 而具有较大面积的吸引域也往往具有较大的势阱深度, 其具体关系曲线见图 3(i), 这与图 3(g) 中吸引域被访问次数和势阱深度之间的正相关对应. 结合 DMN 能量图景结构及跳转过程可以得到以下启发.

a) 不同被试的能量图景吸引子数目以及吸引域间跳转行为存在差异, 分别反映了其在静息态下脑活动的丰富程度和活跃程度的差异. 这可能与不

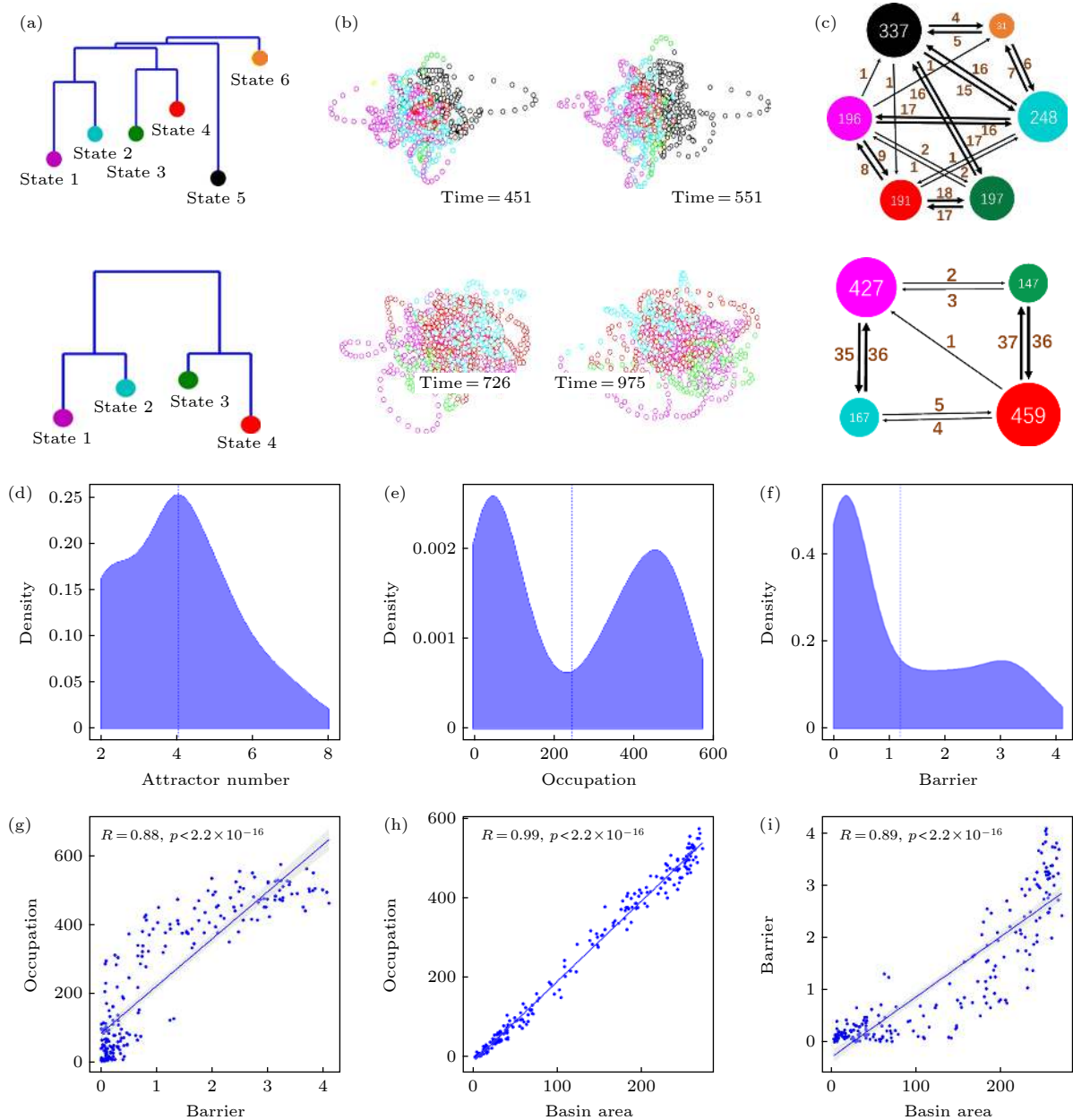


图 3 默认模式网络 (DMN) 激活行为的动力学特征 (a) dDMN 状态能量图景中吸引子状态的非联通图 (disconnectivity graph); (b) 根据吸引子吸引域标签进行着色的 DMN 状态轨迹在 3 维空间的投影, 左右两图显示了同一被试轨迹不同角度的观察; (c) 吸引子间跳转网络, 节点大小表示对应吸引域的访问次数, 点间有向边的权重表示对应的单向跳转频次, 总时长 $T = 1200$. (a)–(c) 中上下两行分别对应于被试 1 和被试 2 的结果; (d)–(i) 为 55 个被试能量图景的统计分布特征: (d) 吸引子数目的统计分布; (e) 吸引域占据率 (访问次数) 的统计分布; (f) 能量壁垒 (势阱深度) 的统计分布; (g) 吸引域占据率与能量壁垒的关系; (h) 吸引域占据率与吸引域面积的关系; (i) 能量壁垒和吸引域面积的关系

Fig. 3. Dynamic characteristics of the resting state default mode network (DMN): (a) Disconnectivity graph of the energy landscape for the dDMN; (b) projections of the DMN state trajectories, colored according to the labels of attraction basins, in a 3-dimensional space. The left and right panels show the trajectory of the same subject from different viewing angles, respectively; (c) attractors transition networks. The nodal size indicates the visiting frequency in the corresponding attracting domain, and the weight of directed edge indicates the corresponding directed switching frequency, with the recording time $T = 1200$. The upper and lower rows in the panels (a)–(c) correspond to the results of test subjects 1 and 2, respectively. Panels (d)–(i) are the statistical results of energy landscapes for 55 subjects: (d) The distribution of attractor numbers; (e) the distribution of basin occupations; (f) the distribution of energy barriers; (g) dependence of basin occupation and energy barrier; (h) dependence of basin occupation and area; (i) dependence of energy barrier and area.

同被试的认知及思维活动特征差异有关,也可能与被试在实验测量前的短期经历有关. 已有研究发现静息态数据采集之前被试所处环境对其静息水平存在遗留影响. 可以推测, 被试在进入静息态之前若经历了较为丰富的刺激, 所得到的 DMN 静息态能量图景应具有更多的吸引子状态和更频繁的跳转. 该推测还有待进一步的实验验证.

此外, 能量图景的吸引子势能、位置关系和吸引域大小等信息可提示潜在的激活模式间跳转方式. 如被试 1 的数据中存在跨越高势垒(见图 3(a))的状态 2—状态 5 间频繁交替的模式(状态 5 具有较大的吸引域), 跨越低势垒的状态 3—状态 4 间频繁交替的相对局部的模式(其状态较为接近), 以及由上述 2—5 相到 3—4 相过渡的状态 3—状态 5 交替的中间模式. 据此可以了解某被试在静息态下其 DMN 状态的特有跳转模式, 由于 DMN 的各个脑区激活行为与外部脑区的神经活动状态存在关联, 可以根据该跳转模式理解被试在静息态下的习惯思维活动.

b) 目前基于 BOLD 数据对 fingerprint(指纹)指标, 即稳定的个体间差异指标的探讨是一个重要的前沿科学问题^[29]. 从上述分析可以看到, 不同被试的吸引子个数、状态空间轨迹构型、跳转方式等均具有特异性, 其中对特定个体相对稳定同时在不同个体之间差异较大的动力学属性可以作为潜在的动力学指纹, 并运用于对被试的个体属性及脑功能障碍的识别.

2.2.1.2 DMN 与外部其他脑区活动的交互

以上介绍了静息态下 DMN 状态在不同吸引子间跳转的行为, 该非平衡过程的机理解释需要从 DMN 的内在动力学以及外部其他脑区的影响两方面考虑. DMN 作为众多脑功能网络之一, 其状态会受到外部其他功能网络的影响. 这里我们首先以高级视觉皮层 (hV) 和听觉皮层 (Aud) 为例, 分析 DMN 与其他脑区激活行为的交互关系.

图 4(a) 给出了高级视觉皮层 (hV) 和听觉皮层 (Aud) 的 BOLD 时间序列, 若进行二值化, 用 1 和 0 对应激活与不激活, 则时间序列可用

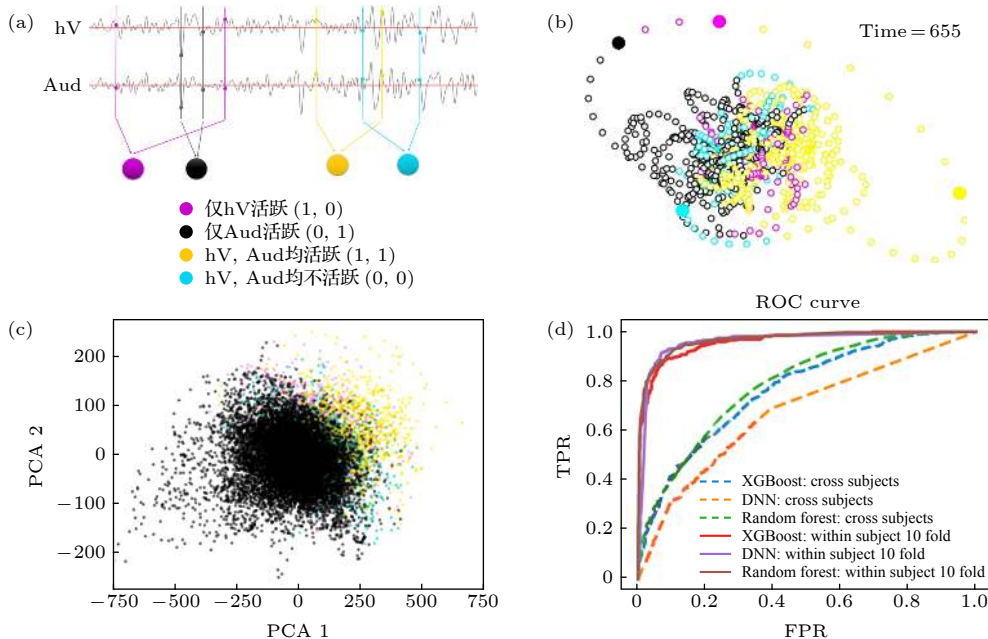


图 4 默认模式网络 (DMN) 状态与其他脑区激活行为的共变 (a) 对高级视觉皮层 (hV) 和听觉皮层 (Aud) 活动二值化, 得到 4 种激活状态的 (hV, Aud) 标签序列; (b) 根据 (hV, Aud) 标签序列着色的 DMN 状态轨道; (c) DMN 的 23 个脑区激活行为在 PCA1-PCA2 空间的投影; (d) 对 DMN 的激活行为数据按照同时刻 (hV, Aud) 标签进行 4 分类训练所得 ROC 曲线, 其中包括 XGBoost, DNN, Random Forest 三种方法及两种训练方式

Fig. 4. Covariation of the default mode network (DMN) states with activity mode in other brain regions: (a) Binarizing the activities of high-level visual cortex (hV) and auditory cortex (Aud) to obtain (hV, Aud) label sequence; (b) DMN state orbit colored according to (hV, Aud) label sequence; (c) projection of the activity states of the 23 brain regions in the PCA1-PCA2 space; (d) the ROC curve obtained by performing 4-classes training on the (hV, Aud) labels of DMN activity data by XGBoost, DNN, Random Forest with two training schemes.

4种 (hV, Aud) 激活状态表示, 即 (1, 0), (0, 1), (0, 0) 和 (1, 1), 图中用4种颜色示意. 为了分析 DMN 与 hV, Aud 的共变行为, 可根据同时刻的 (hV, Aud) 标签对 DMN 状态轨道进行着色, 如图 4(b), 并与图 3(b) 中根据吸引域标签染色的轨迹进行对比. 图 5(a) 和图 5(b) 分别给出了被试 1 和被试 2 根据吸引域标签进行着色的 DMN 状态轨道在 2 维空间 (由 PCA 前两个主成分对应的本征矢张成) 的投影, 图 5(c) 和图 5(d) 给出了被试 1 和被试 2 根据 (hV, Aud) 标签着色的同样的轨迹投影. 从图 5 可见不同的 (hV, Aud) 标签在 DMN 的状态空间中对对应特定的 DMN 子状态区域 (即吸引子及其吸引域), 说明 DMN 的状态与 (hV, Aud) 的状态有较好的对应关系. 我们通过 15 个被试数据的组分析对该现象进行验证, 也可见一致的现象, 见图 4(c), 其中二维空间是由 (通过对 15 个被试数据进行组 PCA 得到的) 前两个主成分 PCA1 和 PCA2 张成, 每个点表示某被试某一时刻的 DMN 状态投影. 所有被试的 DMN 均表现出了一致地与自身 (hV, Aud) 存在对应关系的激活行为, 这一规律没有被 DMN 轨道显著的个体间差异所影响.

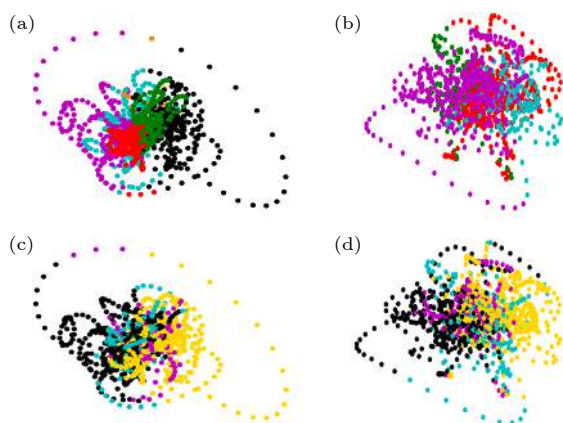


图 5 DMN 状态与视听觉皮层状态的对应现象 (a), (b) 根据吸引子的吸引域标签着色的 DMN 状态轨道; (c), (d) 根据 (hV, Aud) 标签序列着色的 DMN 状态轨道; (a), (c) 被试 1 的结果; (b)(d) 被试 2 的结果

Fig. 5. Correspondence between the states of DMN and (visual, auditory) cortexes: (a), (b) DMN state orbits colored according to the labels of attraction basin; (c), (d) DMN state orbits colored according to the labels of (hV, Aud) states. (a), (c) the results of subject 1; (b)(d) the results of subject 2.

以上是从状态空间的分布特征观察 DMN 与 (hV, Aud) 的状态对应关系, 我们还可通过

XGBoost, DNN(deep neural network), Random Forest 等机器学习算法对上述关系进行验证. 以 (hV, Aud) 的 4 个状态作为标签, 用 15 个被试 DMN 的 23 个脑区激活行为的 1200 个采样数据对模型进行 4 分类训练. 图 4(d) 给出了对应的 ROC 曲线, 横坐标为假阳性率 (FPR), 纵坐标为真阳性率 (TPR), 图中虚线表示通过跨被试的交叉验证, 即随机选择 13 个被试构成训练集, 剩余的 2 个被试作为测试集的结果. 图中实线表示通过十折交叉验证得到的结果, 具体处理方法是对某被试的 1200 个数据点随机拆分为训练集和测试集, 用十折交叉验证得到结果, 再对 15 个被试的 FPR 和 TPR 进行平均得到 ROC 曲线. 从 ROC 曲线可以看出: 1) 根据某被试 DMN 的状态能够得到对其自身 (hV, Aud) 状态的较为准确的推断, 即 DMN 与 hV, Aud 的状态存在密切的依赖关系; 2) 跨被试的验证结果劣于单个被试十折交叉验证结果, 是由于 DMN 轨迹的个体间差异等因素所致.

从上述 DMN 与 (hV, Aud) 的对应关系可知, 静息态下 DMN 在多吸引子间跳转的动力学过程与 DMN 外部脑区活动密切相关. 需要说明的是, 静息态并非对人脑状态的严格定义, 而是 BOLD 数据采集的实验条件, 在静息态下, 某些事件 (也可能是自发涨落) 导致视觉或听觉皮层的激活, 会伴随着 SN 等网络参与的对 DMN 特定形式的抑制作用.

2.2.2 静息态 DMN 内部功能网络分析

为了分析不同被试在静息态 DMN 内部功能网络 (这里包括关联网和因果性网络) 方面的共性特征, 我们计算了平均的结果, 得到了一些在静息态下 DMN 内部普遍存在的功能性结构特征.

2.2.2.1 静息态 DMN 内部的关联网

基于偏相关计算得到 DMN 的 23 个 ROI 的关系矩阵及网络 (见图 6(a) 和图 6(b)). 可以看到, 网络中强的正相关性边包括 vDMN (V1-V10), pDMN (P1-P4), dDMN 的主区域 (D1-D6) 以及左右海马旁回 (parahippocampal gyrus, 图中的 D8 和 D9) 各自的内部连边. 此外, 强的正相关性连边还包括 D2 (左角回, angular gyrus)-P3 (左下顶叶, inferior parietal lobule)、D4 (左背侧后扣带回, dorsal posterior cingulate gyrus)-P2 (右楔前

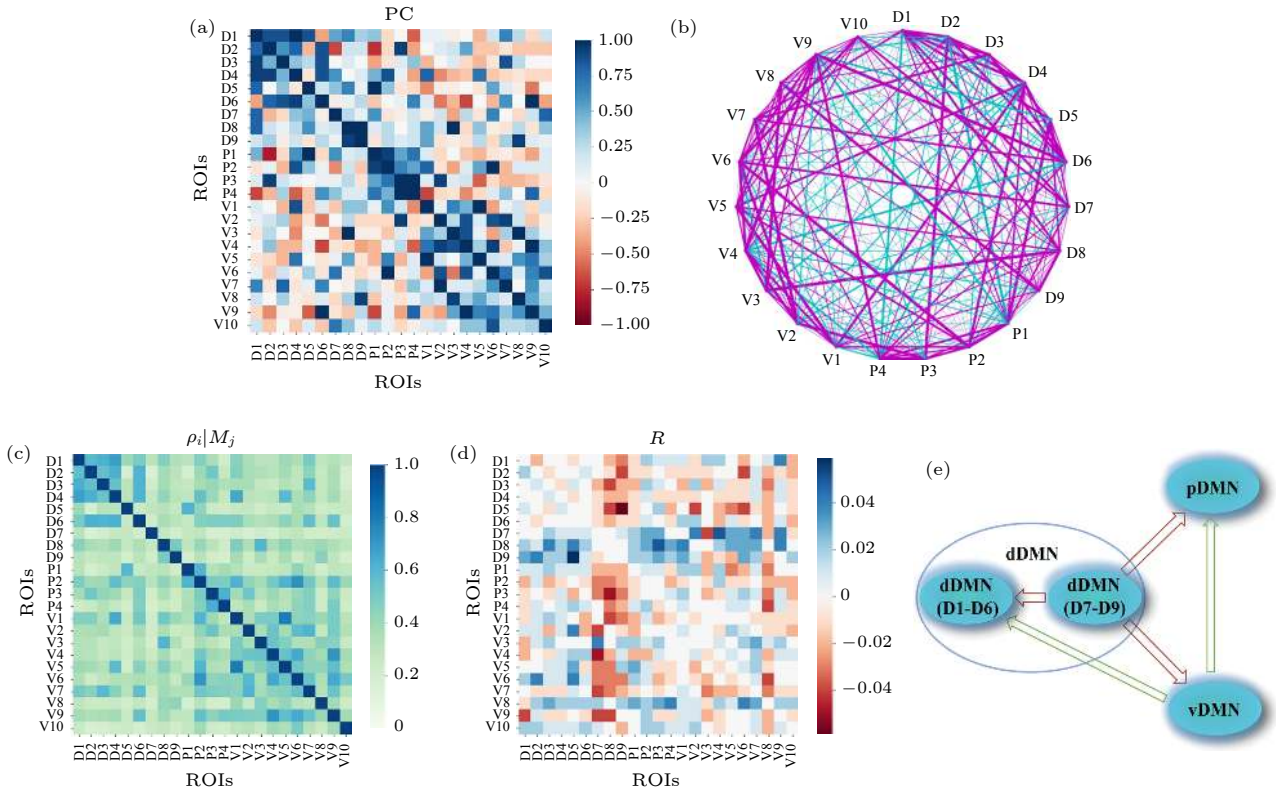


图 6 DMN 内部 23 个 ROI 间关系分析 (a), (b) 基于压缩感知的偏相关矩阵及网络结构, 网络中紫色线表示正相关, 青色线表示负相关, 线的粗细表示相关性强弱; (c) CCM 方法中得到的 ROI 之间的单向影响程度 $\rho_{i|M_j}$, 延迟坐标嵌入维数为 2, 延迟为 1, 下同; (d) CCM 方法得到的 ROI 间动力学因果性矩阵; (e) 对应的 ROI 间动力学影响的粗粒化结果

Fig. 6. Relationships of the 23 ROIs within DMN: (a) Partial correlation matrix calculated based on compressive sensing; (b) the corresponding network structure, which includes the positive (magenta links) and negative (cyan links) correlations. The width of the line indicates the correlation strength; (c) the directed dynamical influence from ROI i to j , $\rho_{i|M_j}$, obtained in the CCM method. The delay coordinate embedding dimension is 2 and the delay is 1; (d) the causality matrix obtained by CCM; (e) the coarse-grained result of the corresponding influence among ROIs within DMN.

叶, precuneus)、D4-V1 (左后扣带回, posterior cingulate)、D4-V5 (右后扣带回, posterior cingulate)、P2-V6 (左楔前叶, precuneus)、D5 (右腹侧扣带回前部, ventral anterior cingulate gyrus)-P1 (左腹后扣带回, ventral posterior cingulate gyrus)、P4 (右下顶叶, inferior parietal lobule)-V9 (右角回, angular gyrus) 等, 这些强的正相关点对多为空间近邻。

图 6(a) 和图 6(b) 中也存在强的反相关性边, 如 D2-P1, P1-V9, D1-P4, D6-V4. 对于反相关性边, 一种可能的原因是信息沿着正相关性边所连成的通路传递的过程中积累的延迟相位差导致的, 并且两节点的反相关程度与该两节点之间由正相关边所组成的最短路径长度有关^[30], 随着最短路径长度的增加, 反相关程度增强, 继续增加又会使之减弱 (相位引起的效应). 这里我们从 DMN 的 ROI 中也观察到了类似的情况, 如 D6-V9, V9-V4

均为正相关, V6-D4 为反相关; P1-P4, P4-V9 均为正相关, P1-V9 为反相关.

2.2.2.2 静息态 DMN 内部的因果网络

图 6(c) 为 CCM 方法中得到的 i 脑区对 j 脑区的影响程度 $\rho_{i|M_j}$, 该矩阵为近似对称的非对称矩阵. 可以看到, 其中较强的影响程度对应的脑区在偏相关分析中 (图 6(a)) 也被发现具有较强的关联性. 图 6(d) 给出了 CCM 方法得到的动力学因果性网络, 可以明显地看到 DMN 的子网络结构, 其中 vDMN, pDMN 相对分离, dDMN 内部 D1-D6 相对一致, D7 (丘脑)、D8 和 D9 (左右海马旁回) 则相对独立. CCM 方法得到的结果 R 可对应于 DMN 内部区域间的信息流, 见图 6(e). 在静息态下, 丘脑 (D7) 和左右海马旁回 (D8, D9) 处于 DMN 内部信息流的上游, dDMN 中的 D1—D6 和 pDMN 是下游子网络; vDMN 从丘脑和左右海马旁回接收信息, 并向 pDMN 和 dDMN 的主区

域 (D1—D6) 发送信息. 在 dDMN 中, D1(内侧额叶) 相对特殊, 它对 pDMN 和 vDMN 中的多个脑区都有较强的动力学影响; 在 vDMN 中, V3 和 V8(左右嘴峰, Culmen) 也具有广泛的对其他脑区的影响.

3 总结

研究表明, 大脑 DMN 在静息态下出现多种激活模式切换的行为, 但从复杂系统非平衡动力学的角度研究静息态 DMN 激活行为的工作还比较少见. 本文基于静息态 fMRI 数据分析了静息态 DMN 的动力学行为. 首先通过构造静息态 DMN 能量图景, 观察了不同被试在静息态的 DMN 状态非平衡演化过程. 为了理解该非平衡过程受到外部影响的机制, 我们以高级视觉皮层和听觉皮层为例, 通过 DMN 状态空间中的分布特征以及 XGBoost、深度神经网络等机器学习算法, 明确了外部脑区活动与 DMN 状态跳转的密切关系. 这为深入理解 DMN 的动力学行为启发了新的思路. 文中图 3 与图 4 的结果也验证了对 DMN 的传统认识, 即 DMN 作为维持觉醒功能的网络会因外部刺激而发生状态调节. 并且, 我们的结果从能量图景与外部脑区激活模式标签对照的新视角提供了静息态下 DMN 的动力学状态与外部脑区 (视、听觉皮层) 状态之间关联的证据.

另外, 本文采用偏相关、收敛交叉映射等方法计算了 DMN 内部 23 个 ROI 的功能网络, 包括关联性网络和因果性网络, 识别出了部分主要功能性连接, 以及 DMN 内部的子网络结构和动力学影响的核心区域 (见图 6). 虽然 DMN 的动力学行为具有较大的个体间差异, 但组平均的关联网络能够反映出比较稳定的强相关边; 因果性网络给出了 DMN 的子网络结构, 该结构可以得到关联性网络的佐证. 这些结果印证了前人对于 DMN 主要内部结构的研究结论 (例如 DMN 分为 dDMN, vDMN 和 pDMN 三大区域, 其各自具有密集的内部连接而彼此之间连接相对较少, 丘脑和海马旁回则是相对独立的功能区等). 同时, 也给出了一些新的启发, 例如根据因果性网络平均结果给出的 DMN 内部 ROI 间信息流, 我们可以得到这样的结论: 在静息态下, 1) 丘脑和海马旁回是 DMN 其他区域的驱动者, 即信息是从丘脑和海马旁回流向 DMN 其他

区域; 2) 嘴峰 (Culmen, 相应的 ROIs 为 vDMN3 和 vDMN8) 是 DMN 内很多其他区域的驱动者 (甚至处于丘脑和海马旁回的上游).

上述围绕 DMN 内部功能连接以及与外部脑区共变特征的探讨, 增加了我们对 DMN 动力学行为的理解. 但由于 DMN 的吸引子状态激活与视听觉、突显网络等功能网络活动的相互作用具体细节机制尚不完全清楚, 目前还无法把外部脑区对 DMN 活动非平衡过程的影响形式化为特定的方程, 这还有待神经影像学与复杂系统等领域的联合探索. 本文基于静息态 DMN 状态能量图景的动力学分析方法具有潜在的应用价值, 例如: 1) 已有研究发现静息态数据采集之前, 被试所处环境对其静息水平存在遗留影响 (例如, 被试刚刚经历了嘈杂环境或完成任务态测试, 与充分休息后所观察到的静息态存在一定差异), 那么, 特定被试在不同“静息水平”下的 DMN 状态分布及切换行为就可以根据状态能量图景中各个吸引子的能量大小以及吸引域大小进行分析. 例如, “弱静息”条件下 (高“温度”), DMN 可能更多地在具有较大吸引域的亚稳态附近漫游, 而“强静息”条件下 (低“温度”) 则陷入具有较低“能量”的亚稳态附近, 即使其吸引域较小; 2) 根据静息态 DMN 状态能量图景中各个吸引子相对位置以及吸引域构型特征, 结合 DMN 状态与外部脑区激活的对应关系, 预测被试在静息态下的特征“心理活动轨迹”.

感谢北京师范大学王文旭教授, 兰州大学俞连春、黄亮教授, 陕西师范大学王圣军教授的有益讨论, 以及西安交通大学黄文敏、曹玲灿、李蒙在数据处理方面的协助.

参考文献

- [1] Sarraf S, Sun J 2016 *arXiv: 1602.02225* [physics.med-ph]
- [2] Liu C, Zhou C, Wang J, Loparo K 2018 *IEEE Trans. Neural Syst. Rehabil. Eng.* **26** 1649
- [3] Lei Y, Song B, Chen L, Su J, Zhang X, Ni W, Yu Y, Xu B, Yu L, Gu Y, Mao Y 2018 *Brain Imaging Behav.* **11** 682
- [4] Chen J E, Glover G H, Greicius M D, Chang C 2017 *Hum. Brain. Mapp.* **38** 2454
- [5] Chen B, Li X 2017 *IEEE International Conference on Systems, Man, and Cybernetics (SMC)* Banff, Canada, October 1–4, 2017 p2820
- [6] Roberto T, Fox P Tomás P J 2008 *Cereb. Cortex* **18** 2553
- [7] Raichle M E, Macleod A M, Snyder A Z, Powers W J, Gusnard D A, Shulman G L 2001 *Proc. Natl. Acad. Sci. U.S.A.* **98** 676
- [8] Greicius M D, Ben K, Reiss A L, Vinod M 2003 *Proc. Natl.*

Acad. Sci. U.S.A. **100** 253

- [9] Greicius M D, Kaustubh S, Vinod M, Dougherty R F 2009 *Cereb. Cortex* **19** 72
- [10] Lin P, Yang Y, Jovicich J, Pisapia N D, Wang X, Zuo C S, Levitt J J 2016 *Brain Imaging Behav.* **10** 212
- [11] Gusnard D A, Raichle M E 2001 *Nat. Rev. Neurosci.* **2** 685
- [12] Li Y, Yao H, Lin P, Zheng L, Li C, Zhou B, Wang P, Zhang Z, Wang L, An N 2017 *Front. Ag. Neurosci.* **9** 259
- [13] Anticevic A, Cole M W, Murray J D, Corlett P R, Wang X J, Krystal J H 2012 *Trends Cogn. Sci.* **16** 584
- [14] Wang J, Wang Y, Wu X, Huang H, Jia Y, Zhong S, Wu X, Zhao L, He Y, Huang L, Huang R 2020 *Brain Imaging Behav.* **14** 186
- [15] Zhao Q, Swati Z N, Metmer H, Sang X, Lu J 2019 *Neurosci. Lett.* **701** 154
- [16] Greicius M D, Srivastava G, Reiss A L, Menon V 2004 *Proc. Natl. Acad. Sci. U.S.A.* **101** 4637
- [17] Fassbender C, Zhang H, Buzy W M, Cortes C R, Mizuiri D, Beckett L, Schweitzer J B 2009 *Brain Res.* **1273** 114
- [18] Uddin L Q, Kelly A M, Biswal B B, Margulies D S, Shehzad Z, Shaw D, Ghaffari M, Rotrosen J, Adler L A, Castellanos F X, Milham M P 2008 *J. Neurosci. Methods* **169** 249
- [19] Manoliu A, Riedl V, Zherdin A, Mühlau M, Schwerthöffer D, Scherr M, Peters H, Zimmer C, Förstl H, Bäuml J, Wohlschläger A M, Sorg C 2014 *Schizophrenia Bull.* **40** 428
- [20] Supekar K, Cai W, Krishnadas R, Palaniyappan L, Menon V 2019 *Biol. Psychiatry* **85** 60
- [21] Bonnelle V, Ham T E, Leech R, Kinnunen K M, Mehta M A, Greenwood R J, Sharp D J 2012 *Proc. Natl. Acad. Sci. U.S.A.* **109** 4690
- [22] Cui Y, Yu S, Zhang T, Zhang Y, Xia Y, Yao D, Guo D 2018 *Brain Res.* **1696** 71
- [23] Wang S J, Ouyang G, Guang J, Zhang M, Wong K M, Zhou C 2016 *Phys. Rev. Lett.* **116** 018101
- [24] Guo D, Guo F, Zhang Y, Li F, Xia Y, Xu P, Yao D 2018 *Front. Comput. Neurosci.* **12** 21
- [25] Watanabe T, Rees G 2017 *Nat. Commun.* **8** 1
- [26] Ashourvan A, Gu S, Mattar M G, Vettel J M, Bassett D S 2017 *Neuroimage* **157** 364
- [27] Lee H, Lee D S, Kang H, Kim B N, Chung M K 2011 *IEEE Trans. Med. Imaging* **30** 1154
- [28] Sugihara G, May R, Ye H, Hsieh C H, Deyle E, Fogarty M, Munch S 2012 *Science* **338** 496
- [29] Finn E S, Shen X, Scheinost D, Rosenberg M D, Huang J, Chun M M, Papademetris X, Constable R T 2015 *Nat. Neurosci.* **18** 1664
- [30] Qian J, Diez I, Ortiz-Terán L, Bonadio C, Liddell T, Goñi J, Sepulcre J 2018 *Front. Neurosci.* **12** 38

SPECIAL TOPIC—Statistical physics and complex systems

Dynamics of the default mode network in human brain*Yao Nan^{1) #} Su Chun-Wang^{2)3) #} Li You-Jun^{2) †} Wang Jue²⁾Zhou Chang-Song⁴⁾ Huang Zi-Gang^{2) ‡}1) (*School of Science, Xi'an University of Technology, Xi'an 710048, China*)2) (*Key Laboratory of Biomedical Information and Engineering of Ministry of Education, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an 710049, China*)3) (*School of Physical Science and Technology, Lanzhou University, Lanzhou 730000, China*)4) (*Center for Nonlinear Research, Institute of Computing and Theory, Department of Physics, Hong Kong Baptist University, Hong Kong, China*)

(Received 2 February 2020; revised manuscript received 22 March 2020)

Abstract

Brain is a typical complex system with characteristics such as self-adaptation, self-organization, and multistability. The activity of the default mode network (DMN), a crucial functional subnetwork of the human brain in resting state, obeys typical non-equilibrium statistical mechanical processes in which the system continually switches among multiple metastable states. Revealing the underlying dynamical mechanism of these processes has important scientific significance and clinical application prospects. In this paper, according to the blood oxygen level dependent (BOLD) signals obtained from functional magnetic resonance imaging (fMRI), we build an energy landscape, disconnectivity graph and transition network to explore the non-equilibrium processes of DMN switching among different attractors in resting state. Taking the activities of high-level visual and auditory cortices for examples, we verify the intimate relationship between the dynamics of DMN and the activity modes of these external brain regions, through comparing the distributions in state space and the algorithms such as XGBoost and deep neural networks. In addition, we analyze the interaction between various DMN regions in the resting state by using the techniques such as compressive-sensing-based partial correlation and convergence cross mapping. The results in this paper may present new insights into revealing the dynamics of the intrinsic non-equilibrium processes of brain in resting state, and putting forward clinically significant biomarkers for brain dysfunction from the viewpoint of dynamics.

Keywords: brain functional network, resting state, default mode network, energy landscape, state transition**PACS:** 02.50.-r, 64.60.aq, 87.18.-h**DOI:** 10.7498/aps.69.20200170

* Project supported by the National Natural Science Foundation of China (Nos. 11975178, 11647052), the Open Project of State Key Laboratory of Cognitive Neuroscience and Learning, China (Grant No. CNLYB1802), the Natural Science Basic Research Plan of Shaanxi Province, China (Grant Nos. 2018JQ1010, 2020JM-058, 2020JQ-096), the Scientific Research Program Funded by Shaanxi Provincial Education Department, China (Grant No. 17JK0553), and the Young Talent Fund of University Association for Science and Technology of Shaanxi Province, China (Grant No. 20170606).

These authors contributed equally.

† Corresponding author. E-mail: liyujun1@xjtu.edu.cn

‡ Corresponding author. E-mail: huangzg@xjtu.edu.cn

耦合相振子系统同步的序参量理论

郑志刚 翟云 王学彬 陈宏斌 徐灿

Synchronization of coupled phase oscillators: Order parameter theory

Zheng Zhi-Gang Zhai Yun Wang Xue-Bin Chen Hong-Bin Xu Can

引用信息 Citation: *Acta Physica Sinica*, 69, 080502 (2020) DOI: 10.7498/aps.69.20191968

在线阅读 View online: <https://doi.org/10.7498/aps.69.20191968>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

Pre-Btzing器复合体的从簇到峰放电的同步转迁及分岔机制

Synchronization transition from bursting to spiking and bifurcation mechanism of the pre-Btzing器 complex

物理学报. 2020, 69(4): 040501 <https://doi.org/10.7498/aps.69.20191509>

激光混沌并行串联同步及其在中继器保密通信系统中的应用

Chaotic laser parallel series synchronization and its repeater applications in secure communication

物理学报. 2019, 68(17): 170502 <https://doi.org/10.7498/aps.68.20190212>

多个耦合星型网络的同步优化

Optimizing synchronizability of multiplecoupled star networks

物理学报. 2019, 68(18): 180503 <https://doi.org/10.7498/aps.68.20190308>

弱Soret效应混合流体对流系统的分岔与非线性演化

Bifurcation and nonlinear evolution of convection in binary fluid mixtures with weak Soret effect

物理学报. 2020, 69(7): 074701 <https://doi.org/10.7498/aps.69.20191836>

兴奋性作用诱发神经簇放电个数不增反降的分岔机制

Bifurcation mechanism of not increase but decrease of spike number within a neural burst induced by excitatory effect

物理学报. 2018, 67(24): 240502 <https://doi.org/10.7498/aps.67.20181675>

专题：统计物理和复杂系统

耦合相振子系统同步的序参量理论*

郑志刚^{1)2)†} 翟云³⁾¹⁾²⁾ 王学彬¹⁾²⁾ 陈宏斌¹⁾²⁾ 徐灿^{1)2)‡}

1) (华侨大学系统科学研究所, 厦门 361021)

2) (华侨大学信息科学与工程学院, 厦门 361021)

3) (北京邮电大学理学院, 北京 100876)

(2019年12月26日收到; 2020年1月31日收到修改稿)

节律行为, 即系统行为呈现随时间的周期变化, 在我们的周围随处可见. 不同节律之间可以通过相互影响、相互作用产生自组织, 其中同步是最典型、最直接的有序行为, 它也是非线性波、斑图、集群行为等的物理内在机制. 不同的节律可以用具有不同频率的振子(极限环)来刻画, 它们之间的同步可以用耦合极限环系统的动力学来加以研究. 微观动力学表明, 随着耦合强度增强, 振子同步伴随着动力学状态空间降维到一个低维子空间, 该空间由序参量来描述. 序参量的涌现及其所描述的宏观动力学行为可借助于协同学与流形理论等降维思想来进行. 本文从统计物理学的角度讨论了耦合振子系统序参量涌现的几种降维方案, 并对它们进行了对比分析. 序参量理论可有效应用于耦合振子系统的同步自组织与相变现象的分析, 通过进一步研究序参量的动力学及其分岔行为, 可以对复杂系统的涌现动力学有更为深刻的理解.

关键词: 同步, 序参量, 涌现, 分岔

PACS: 05.45.-a, 05.45.Xt

DOI: 10.7498/aps.69.20191968

1 引言

节律, 即系统呈现的随时间周期变化, 广泛存在于我们生活的世界, 从钟摆、星体运动到生物钟, 都是不同系统典型的动力学行为. 复杂与非线性系统会表现出多种多样的自组织与群体性行为^[1-3], 而不同的节律之间也可以通过相互作用涌现出各种集体的自组织现象, 其中同步是最典型、最直接的有序行为, 它也是诸如非线性波、时空斑图乃至各种生物集群行为(例如鸟群、鱼群、蜂群、蚁群、人类社会等)的内在物理机制^[4,5].

作为一种最基本的协同现象, 同步研究可以追溯到1673年Huygens发现的关于两个相邻钟摆同步的讨论. 在后来的很多不同物理问题中, 人们

都发现了同步现象. 电子与无线电工程的发展很大程度上促进了关于同步的研究. 相比于现象观察及实验研究, 同步的理论研究相对要滞后得多, 直到20世纪初关于极限环的研究才得到了突破. 极限环是非线性耗散系统的一类典型时间振荡解和吸引子, 不同时间振荡之间可以通过相互作用产生新的协同. 由于大量看似不同的系统都具有可以用极限环描述的时间振荡, 人们逐渐认识到, 在各种不同同步行为的背后应该具有共同的物理机理^[5]. Wiener^[6]在其专著中关于“脑电波与自组织系统”一章指出, 脑电波的出现是来自于不同频率的锁定(pulling together of frequencies), 而且这种行为与其他诸如萤火虫同步闪光、蟋蟀同声鸣唱等现象有着共同的机制. 而如何在极限环基础上来研究驱动或相互作用的振子间的同步在理论上就成为

* 国家自然科学基金(批准号: 11875135, 11905068)和泉州市科技计划(批准号: 2018C085R)资助的课题.

† 通信作者. E-mail: zgzheng@hqu.edu.cn

‡ 通信作者. E-mail: xucan@hqu.edu.cn

重要课题.

Winfree^[7]于1967年提出了全局耦合振子模型. 他认识到, 在同步问题上, 极限环之间相互作用起作用的关键自由度是其相位, 因此只需要考虑耦合的相位振子模型即可揭示同步的动力学实质. Winfree模型的动力学方程可写为

$$\dot{\theta}_i = \omega_i + \left(\sum_{j=1}^N X(\theta_j) \right) Z(\theta_i), \quad (1)$$

这里用 $i = 1, 2, \dots, N$ 来标记不同振子, $\{\theta_i\}$ 代表振子的相位, $\{\omega_i\}$ 为振子的自然频率, 它们各不相同, 设符合某个统计分布函数 $g(\omega)$. Winfree发现, 自然频率分布较窄时振子之间会相互同步^[7]. 这样, 在众多不同物理背景下出现于不同体系中的很多现象就都可以利用耦合相振子系统的同步动力学来得到非常好的解释. 大量相互作用振子出现的整体同步行为是一种典型的集体涌现, 如何解析刻画同步涌现是一个挑战性课题.

Kuramoto^[8-10]考虑振子数目 $N \rightarrow \infty$ (热力学极限), 自然频率 $\{\omega_i\}$ 为单峰分布 $g(\omega)$, 振子间相互作用是耦合强度为 K 的全局性平均场形式, 并取最简单的相位差的正弦函数 $\sin(\Delta\theta)$ 来描述相互作用. 平均场耦合相振子模型可写为

$$\dot{\theta}_i = \omega_i + \frac{K}{N} \sum_{j=1}^N \sin(\theta_j - \theta_i). \quad (2)$$

该模型被后来研究者称为 Kuramoto 模型^[11]. 我们还引入振子的集体自然频率 $\bar{\omega} = \frac{1}{N} \sum_{i=1}^N \omega_i$. 一般情况下自然频率分布函数设为关于 $\bar{\omega}$ 的对称分布. 为讨论方便, 通常设 $\bar{\omega} = 0$, 分布对称性满足 $g(\omega) = g(-\omega)$.

尽管振子的自然频率各不相同, 有相互作用时 ($K \neq 0$) 各振子的实际振动频率都会相应地从 $K = 0$ 时的自然频率 $\{\omega_i\}$ 处发生偏离, 并随耦合强度变化而变化. 可定义振子的平均频率 Ω_i 为

$$\Omega_i = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \dot{\theta}_i(t) dt. \quad (3)$$

为描述振子的整体同步情况 (相干程度), Kuramoto 引入如下的序参量, 将其定义为所有振子相位复函数的平均场:

$$z = R \exp(i\Theta) = \frac{1}{N} \sum_{j=1}^N \exp(i\theta_j), \quad (4)$$

其中复序参量的模 $R = |z|$ 描述振子的相干性强弱, Θ 为任一任意相位.

在耦合强度很弱时, 只有自然频率极其靠近的振子才会同步, 但它们所占的比例几乎可以忽略, 大量振子的 Ω_i 都不等, 它们在任一时刻的相位都均匀分布于 $0-2\pi$ 之间, 如图 1(a) 所示, 此时 $R = 0$. 随着耦合强度的增加, 越来越多的振子会同步, 平均频率 Ω_i 相等, 这些同步的振子相位之间会靠近并保持固定相位关系, 振子不再均匀分布, 如图 1(b) 所示. 当所有 Ω_i 都相等时, R 就不为零, 表明此时振子之间可保持固定的相位关系. 存在一个临界的耦合强度 K_c , 当 $K \leq K_c$ 时 $R = 0$; 当 $K \geq K_c$ 时, $R \neq 0$. 在很强的耦合下, 振子相位会靠得很近, 形成整体的同步大集团, 如图 1(c) 所示. 大量耦合振子可以通过相互作用克服自然频率不同带来的无序而涌现出同步态实质上就是一种典型的非平衡相变. 在临界 K_c 处发生的相变在理论上可以处理. Kuramoto 利用统计物理学方法和自洽方程成功求解^[9,10], 理论上得到了临界耦合强度 K_c , 并得到在 K_c 附近序参量 R 的临界行为.

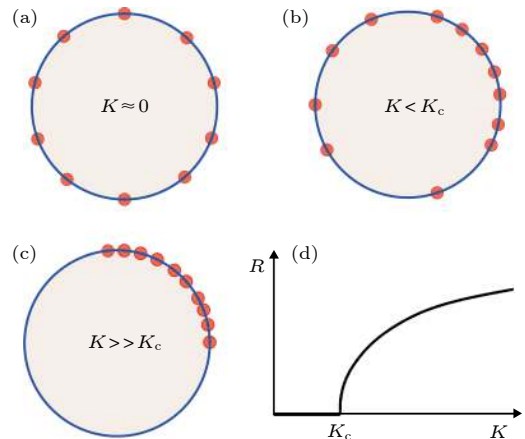


图 1 耦合振子同步示意图 (a) 在耦合强度很弱时, 大量振子不同步, 任一时刻相位均匀分布于 $0-2\pi$ 之间; (b) 随着耦合强度的增加, 越来越多的振子会同步, 振子不再均匀分布; (c) 在很强的耦合下, 振子相位会靠得很近, 形成整体的同步大集团; (d) 序参量随耦合强度的变化

Fig. 1. A schematic diagram of synchronization of coupled oscillators: (a) Most of the oscillators are asynchronous and evenly distributed along the circle; (b) with increasing the coupling, more and more oscillators are synchronized and are no longer evenly distributed; (c) under a strong coupling, oscillators form a single synchronous cluster, and the phases of oscillators are close to each other; (d) dependence of the order parameter on the coupling strength.

结合具体的物理背景, Kuramoto 模型的应用范围和领域大大拓展^[11-17]. 人们研究了不同耦合

函数^[18,19]、不同耦合网络拓扑结构^[14,15,20–25]、脉冲耦合^[26]等情况下的同步,探讨了同步涌现的玻璃态^[27]、非对称耦合及其行波和驻波波态^[28,29]、惯性效应^[30,31]、阻挫效应^[32–35]、时间延迟效应^[36]、外场驱动效应^[37]、时变耦合^[38]等与实际物理背景密切相关的问题.以 Kuramoto 模型为代表的耦合振子系统集体动力学问题时至今日一直是重要的研究热点.

对于大量振子同步的研究有几种不同层次和方法.利用非线性动力学^[39–41]和同步分岔树方法^[42–44]可以对耦合振子的微观动力学进行研究.当振子数目 $N \gg 1$ 时,微观动力学分析会变得很繁杂,可以利用统计物理和宏观方法开展研究.早期 Kuramoto 基于统计物理学方法和自洽方程理论成功地对同步相变进行了解析研究^[9].但自洽方程方法建立在同步态为定态的前提下,近年来人们发现了大量非定态集体动力学行为,因此有必要建立一套反映集体动态行为的统计与宏观理论^[45].20世纪90年代初,Watanabe 与 Strogatz 发现^[46],一类具有对称性的耦合振子系统的高维动力学可以通过引入 Möbius 变换(后称为 Watanabe-Strogatz 变换,简称 WS 变换)精确降维到三维空间,这意味着系统的部分可积性.Ott 与 Antonsen^[47,48]提出了序参量拟设理论,将高维微观动力学研究降维到二维序参量空间加以研究.人们随后意识到,OA 拟设下的二维动力学实际上是 WS 变换下的三维动力学的进一步降维^[49,50].

耦合振子系统同步的序参量动力学研究大致包括如下几个重要和基础问题.第一,序参量是表征整体行为的特征量,那么对于大自由度的耦合振子系统,序参量是如何涌现的?第二,序参量及其涌现的物理意义是什么?第三,如何用序参量刻画不同的同步或有序状态及其转变?第四,如果有不止一个序参量共存,这些序参量之间如何竞争?第五,对于实际系统,如何引入恰当的表征集体行为的序参量来刻画有序行为?如何从复杂系统的动态大数据中重构序参量动力学?本文将集中于上述的前三个基本问题,对大量振子组成的系统产生同步作为一种典型的涌现行为开展宏观和统计层面的研究,总结、回顾并对比几种不同的序参量理论框架^[51].

2 Kuramoto 自洽方程理论

首先来看 Kuramoto 如何通过引入序参量并利用求解自洽方程来解析处理同步问题.自洽方法的基础是通过假设系统存在一个不随时间改变的定态,在此状态下序参量为一个待定的定值,通过序参量的定义和系统定态的运动方程,得到待定的序参量的值,并在分析的过程中得到此序参量与相应定态的存在条件.自洽方法从其方法本身便限定了其适用范围,虽然只能用于对定态的分析,却可以不受具体动力学的限制,是振子系统分析中广泛使用的方法之一.

首先,假定振子数 N 足够多情况下,序参量(4)式与 N 无关且不随时间变化.考虑到 Kuramoto 模型(2)式相互作用的平均场形式,很容易可以将方程(2)重新写为

$$d\theta_i/dt = \omega_i + KR \sin(\Theta - \theta_i). \quad (5)$$

如果能定出 R ,则方程(5)完全可以求解,但这不是一件简单的事.一个可行的办法是带着未知 R 继续讨论,建立一个关于 R 的方程来将其求解.该方程即为自洽方程(self-consistent equation)^[8,51].下面的讨论即是围绕这一主题展开.

显然 $R = 0$ 对应于无相互作用的情形(均匀分布解),这个非相干态总是系统的一个解,但不总是稳定.当 $K > K_c$ 时均匀分布解失稳.另外一个解是 Ω_i 都相等时的解,此时振子之间可以保持固定的相位, $R \neq 0$,所有振子都以集体频率 $\bar{\omega}$ 转动, $\Theta = \bar{\omega}t$.此解在 $K \leq K_c$ 时不稳定.只要有 Ω_i 不相等, θ_i 就总是均匀分布于 $0-2\pi$ 之间.

引入变量 $\phi_i = \theta_i - \bar{\omega}t$, (5) 式可写为

$$d\phi_i/dt = \omega_i - \bar{\omega} - KR \sin \phi_i, \quad (6)$$

此方程正是过阻尼情形下的单摆方程.它有两个解.1) 同步解.当方程(6)描述的第 i 个振子满足 $|\omega_i - \bar{\omega}| \leq KR$ 时,该振子的相位 ϕ 就会保持定值 $\phi_i = \sin^{-1}[(\omega_i - \bar{\omega})/(KR)]$,这意味着所有满足该条件的振子都会以 $\bar{\omega}$ 频率运动,它们处于同步状态.2) 非同步解.当方程(6)描述的振子 i 满足 $|\omega_i - \bar{\omega}| > KR$ 时,则该振子的相位 ϕ 就会随时间变化,且凡是自然频率满足该条件的振子都处于非同步状态.

振子数 $N \rightarrow \infty$ 时, ϕ_i 在 $0-2\pi$ 之间会形成分布.设分布函数为 $P(\phi, \omega, t)$,它不仅依赖于 ϕ ,还

依赖于振子自然频率 ω . 平均场由分布函数可表为

$$Re^{i\Theta} = e^{i\bar{\omega}t} \int_0^{2\pi} \int_{-\infty}^{\infty} e^{i\phi} P(\phi, \omega, t) g(\omega) d\omega d\phi, \quad (7)$$

单振子分布函数满足连续性方程:

$$\frac{\partial P(\phi, \omega, t)}{\partial t} = -\frac{\partial}{\partial \phi} \{[\omega - \bar{\omega} - KR \sin \phi] P(\phi, \omega, t)\}. \quad (8)$$

考虑到振子具有自然频率分布, 如果只考察 ϕ 的统计分布, 则需要进一步对频率做积分 $P(\phi, t) = \int P(\omega, \phi, t) g(\omega) d\omega$. 这些振子相位的分布直接决定了相关的平均量, 例如序参量, 因而很重要的一点就是如何确定分布函数. 根据上面讨论的两类定态解, 可把分布 $P(\phi)$ 分解为同步与非同步两部分:

$$P(\phi) = P_s(\phi) + P_{as}(\phi). \quad (9)$$

同步振子的相位 ϕ 趋于不动点, 与时间无关, 因此 $P_s(\phi)$ 可由自然频率分布得到:

$$P_s(\phi) = g(\omega) \left| \frac{d\omega}{d\phi} \right| = KRg(\bar{\omega} + KR \sin \phi) \cos \phi, \\ \phi \in \left[-\frac{\pi}{2}, \frac{\pi}{2} \right]; \quad (10)$$

非同步振子的相位 ϕ_i 则随时间变化, 因为 ϕ_i 随时间变化是非均匀的, 单位时间内探测到 ϕ 在 $\phi \rightarrow \phi + d\phi$ 之间的概率反比于相速度 $|\dot{\phi}|$: $p(\phi, \omega) \propto |\dot{\phi}|^{-1}$. 把运动方程 (6) 式代入并归一化可得到

$$p(\phi, \omega) = \frac{\sqrt{(\omega - \bar{\omega})^2 - (KR)^2}}{2\pi |\omega - \bar{\omega} - KR \sin \phi|}, \quad (11)$$

因此

$$P_{as}(\phi) = \int_{|\omega - \bar{\omega}| > KR} g(\omega) p(\phi, \omega) d\omega. \quad (12)$$

令 $x = \omega - \bar{\omega}$, 并考虑到 $g(\bar{\omega} + x) = g(\bar{\omega} - x)$, 将 (11) 式代入 (12) 式积分可写为

$$P_{as}(\phi) = \int_{KR}^{\infty} \frac{g(\bar{\omega} + x) x \sqrt{x^2 - (KR)^2}}{\pi [x^2 - (KR \sin \phi)^2]} dx, \quad (13)$$

将前面引入的平均场用分布写出来, 并利用 (9) 式可得

$$Re^{i\Theta} = \int_{-\pi}^{\pi} e^{i(\phi + \bar{\omega}t)} P(\phi) d\phi \\ = \int_{-\pi}^{\pi} e^{i\phi + i\bar{\omega}t} [P_s(\phi) + P_{as}(\phi)] d\phi. \quad (14)$$

可以看出 (14) 式中 $P_{as}(\phi)$ 部分积分为零, 只有

$P_s(\phi)$ 对积分有贡献. 进一步将积分实部和虚部分离, 可得到 (注意 R 为实):

$$R = KR \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2 \phi g(\bar{\omega} + KR \sin \phi) d\phi, \quad (15)$$

$$0 = KR \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos \phi \sin \phi g(\bar{\omega} + KR \sin \phi) d\phi. \quad (16)$$

由 (16) 式可以确定 $\bar{\omega}$ (实际上在原方程中可通过对称性确定). (15) 式是一个自洽方程, 由此可以确定 R 和同步临界耦合强度 K_c . 当 $K \geq K_c$ 时, R 由零连续变为一个小量, 因此可把 $g(\omega)$ 在 $\bar{\omega}$ 附近展开:

$$g(\bar{\omega} + KR \sin \phi) \approx g(\bar{\omega}) + \frac{g''(\bar{\omega})}{2} (KR)^2 \sin^2 \phi + O(R^4), \quad (17)$$

其中 $g''(\bar{\omega}) = d^2g(\omega)/d\omega^2|_{\omega=\bar{\omega}}$. 代入 (17) 式可以得到

$$1 = \frac{\pi K}{2} g(\bar{\omega}) - \frac{1}{16} \pi K^3 R^3 g''(\bar{\omega}) + O(R^4). \quad (18)$$

当 $K \rightarrow K_c$ 时, 由于 R^2 项为高阶小量, $R \rightarrow 0$, 由此得到同步临界耦合强度为

$$K_c = 2/[\pi g(\bar{\omega})]. \quad (19)$$

将 (19) 式代入 (18) 式可以确定在临界点 K_c 附近 R 的行为:

$$R \approx \sqrt{\frac{8g(\bar{\omega})(K - K_c)}{g''(\bar{\omega})K^3}}. \quad (20)$$

例如对典型的自然频率符合洛伦兹分布 $g(\omega) = \{\pi[(\omega - \bar{\omega})^2 + \gamma^2]\}^{-1}\gamma$ 的情况, 临界耦合为 $K_c = 2\gamma$, 序参量 $R = \sqrt{(1 - 2\gamma/K)}$. 可以看出 R 在 K_c 附近为典型的连续相变 $R \propto (K - K_c)^{1/2}$, 如图 1(d) 所示. 需要指出的是, 这是一种非平衡相变, 是系统有序 (耦合) 压倒无序 (自然频率随机分布) 的结果.

3 序参量及其涌现的自组织原理

耦合振子同步微观动力学的研究表明, 大量耦合振子随着耦合强度的增加会经历由部分同步到整体同步的过渡, 相空间维数随同步进程而逐渐降低, 整体同步时, 系统在相空间的复杂运动会落到一个极低维的空间 [42-44]. 降维意味着系统发生同步的时候, 只需要少数变量即可刻画耦合振子系统的同步. 该结果为多振子体系同步的宏观及序参量描述提供了事实基础.

统计物理学建立了从微观到宏观之间的联系,

进一步通过统计定律来计算微观量的统计平均来得到宏观热力学量^[27]. 热力学和统计物理学的思想和方法为处理耦合振子同步转变问题提供了思路. 由大量相互作用振子组成的系统通常是非平衡系统, 同步是大量振子整体动力学从无序向有序的非平衡相变. 非平衡行为研究自 20 世纪中期发展起来的以耗散结构理论^[52,53]、协同学^[54,55]以及多学科分支形成的自组织理论等为描述众多的非平衡相变现象的共同本质提供了重要依据.

耗散结构理论认为, 处于非平衡状态的系统会在一定范围内维持原有热力学状态, 一直到远离平衡到一定临界值, 热力学分支失稳使系统进入到有序的结构分支. 非平衡相变现象在自然界各个领域有形形色色的表现, 物理表现十分不同, 而内含的数学实质是有明显的规律, 由此形成了以反映扩散方程等为核心的刻画宏观自组织过程的耗散结构理论^[53].

Haken 的协同学则关注于由大量自由度构成的复杂系统在外参量的驱动下和在子系统之间的相互作用下如何以自组织的方式在宏观尺度上形成空间、时间或功能有序结构的条件、特点及其演化规律^[54]. 协同学的基本原理是支配原理 (slaving principle), 它认为, 协同系统的状态由一组状态参量来描述, 这些状态参量弛豫时间尺度是不相同的, 慢变的线性不稳定模称为慢变量 (slow variable), 而快变的线性稳定模称为快变量 (fast variable). 当系统接近于发生显著质变的临界点时, 慢模数目会减少为只有一个或少数几个, 这些慢变量可以完全确定系统的宏观行为并表征系统的有序化程度, 故而称为序参量. 而为数众多的快模则由慢模/序参量所支配, 并可将其绝热消去, 由此可以建立少自由度的协同学基本方程. 在序参量方程基础上, 我们就可来研究协同系统的各种非平衡定态/非定态、稳定性及其非平衡相变^[55].

支配原理在数学上即绝热消去, 即可以对快速变化的变量进行平均加以消去, 只保留变化慢 (绝热) 的变量. 该方法在物理上有着深刻含义. 下面以如下的 n 维动力学系统为例来简单介绍一下支配原理:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}(\mathbf{x}), \quad (21)$$

其中 $\mathbf{x}(t) = (x_1(t), x_2(t), \dots, x_n(t))^T$ 为 n 维状态矢量, \mathbf{A} 为 $n \times n$ 常数矩阵, $\mathbf{B}(\mathbf{x})$ 包含 \mathbf{x} 的二次以

上代数式矢量. 设 $\mathbf{x} = 0$ 为方程的解. 设矩阵 \mathbf{A} 的本征值为 $\{\lambda_i, i = 1, 2, \dots, n\}$, 其中本征值按照其实部由大到小排列. 如果所有本征值实部均为负, 即 $\{\text{Re}\lambda_i < 0, i = 1, 2, \dots, n\}$, 那么 $\mathbf{x} = 0$ 为稳定解. 改变参数使 \mathbf{A} 的 m 个模失稳, 而其他模仍然保持稳定, 即假设该模对应的本征值 $\{0 < \text{Re}\lambda_k^u \ll 1, \text{Re}\lambda_i^s < 0, k = 1, 2, \dots, m, i = 1, 2, \dots, n - m\}$, 且 $\text{Re}\lambda_k^u \ll |\text{Re}\lambda_i^s|$. 通过引入一个 \mathbf{T} 矩阵进行如下的线性变换

$$\tilde{\mathbf{A}} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}, \quad (22)$$

将 \mathbf{A} 对角化

$$\tilde{\mathbf{A}} = \begin{pmatrix} \lambda^u & \mathbf{0} \\ \mathbf{0} & \lambda^s \end{pmatrix}, \quad (23)$$

这里 λ^u, λ^s 分别为 $m \times m$ 与 $(n - m) \times (n - m)$ 对角子矩阵. 则 (21) 式可改写为

$$\dot{\mathbf{u}} = \lambda^u \mathbf{u} + \mathbf{B}^u(\mathbf{u}, \mathbf{s}), \quad (24a)$$

$$\dot{\mathbf{s}} = \lambda^s \mathbf{s} + \mathbf{B}^s(\mathbf{u}, \mathbf{s}), \quad (24b)$$

这里 $\mathbf{x}(t) = (\mathbf{u}(t), \mathbf{s}(t))^T$, 其中 $\mathbf{u}(t) = (u_1(t), u_2(t), \dots, u_m(t))^T$, $\mathbf{s}(t) = (s_1(t), s_2(t), \dots, s_{n-m}(t))^T$, $\mathbf{B}_u = (B_1^u, B_2^u, \dots, B_m^u)^T$, $\mathbf{B}_s = (B_1^s, B_2^s, \dots, B_{n-m}^s)^T$. $\mathbf{s}(t)$ 为满足 (24b) 式的快变量, 而 $\mathbf{u}(t)$ 为满足 (24a) 式的慢变量. 按照支配原理绝热消去, 令 (24b) 式左边 $\dot{\mathbf{s}} = 0$, 由此 $n - 1$ 个方程可将 $n - m$ 维矢量 \mathbf{s} 作为 \mathbf{u} 的函数解出, 并代入 (24a) 式可得到 m 维的非线性方程

$$\dot{\mathbf{u}} = \lambda^u \mathbf{u} + \mathbf{B}^u(\mathbf{u}, \mathbf{s}(\mathbf{u})). \quad (25)$$

一般实际问题中首先失稳的模往往很少, 常常是一两个, 从 (24) 式到 (25) 式, 在理论上可以看作是从多变量方程到少数序参量方程的约化, 在动力学上也产生了极大简化, 这是支配原则在讨论相变点附近行为时给出的有益结果.

需要注意的是, 变量的快慢不能理解为其其他诸如振荡的快慢, 而是弛豫的快慢, 因此严格来说应该称为快/慢弛豫变量. 慢变模式有时可能本身就是高频振荡, 该模对于扰动响应的弛豫时间长短才是判断该模为快或慢的标准.

序参量有着几何上的意义, 支配原理也密切联系着中心流形定理. 中心流形定理是一种常用的几何降维方法, 它利用流形与对应子空间相切的特性, 求出系统在中心流形上的约化方程. 对于高维动力系统来说, 通过传统的分岔行为很难直接研究

其动力系统. 为了更好地抓住所要研究问题的本质, 一般采取中心流形定理等降维措施将其化为低维方程再进行研究.

4 序参量动力学理论

下面将从耦合振子微观动力学方程出发, 利用统计力学方法建立分布函数方程, 并通过统计平均引入各阶序参量, 建立相应的序参量方程, 并进一步讨论对方程的降维.

4.1 广义序参量

从一般形式的全局耦合振子系统运动方程出发:

$$\dot{\theta}_j(t) = F(\alpha, \theta_j, \beta, \gamma_j) + \xi_j(t), \quad (26)$$

其中 $j = 1, 2, \dots, N$, $\xi_j(t)$ 为作用于第 j 个振子上的随机噪声, 设为振子间无关联的高斯白噪声, 满足 $\langle \xi_i(t) \rangle = 0$, $\langle \xi_i(t) \xi_j(t') \rangle = 2D\delta_{ij}\delta(t - t')$. $\beta = \{\beta_1, \beta_2, \dots, \beta_m\}$ 为一组均匀控制参量, 这些参量对所有振子均相同, 如通常考虑振子间相互作用强度相同. $\gamma = \{\gamma_1, \gamma_2, \dots, \gamma_N\}$ 为一组非均匀控制参量, 它们对不同振子不一样. 例如振子的自然频率通常各不相同, 此时 $\{\gamma_i = \omega_i\}$.

定义一组广义序参量 $\alpha = \{\alpha_n\}$, 其分量为如下的 n 阶序参量:

$$\alpha_n = \frac{1}{N} \sum_{j=1}^N \exp(in\theta_j). \quad (27)$$

当 $n = 1$ 时, α_1 即是 Kuramoto 引入的相干因子. $n > 1$ 时 α_n 为高阶序参量 [56]. 在热力学极限 $N \rightarrow \infty$ 下, 可以引入振子相位的密度分布函数 $\rho(\gamma, \theta, t)$, $\rho(\gamma, \theta, t)d\theta$ 为一个振子在时刻 t 相位处于 $\theta \rightarrow \theta + d\theta$ 的概率或时刻 t 相位处于 $\theta \rightarrow \theta + d\theta$ 内的振子数密度. 它满足 (12) 式对应的 Fokker-Planck 方程:

$$\frac{\partial \rho}{\partial t} = -\frac{\partial(\rho v)}{\partial \theta} + D \frac{\partial^2 \rho}{\partial \theta^2}, \quad (28)$$

其中速度场 $v = F(\alpha, \theta, \beta, \gamma)$. 在没有外加噪声 ($D = 0$) 时, 分布函数方程退化为如下的连续性方程

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho v)}{\partial \theta} = 0. \quad (29)$$

在一般情况下振子是非全同的, 即分布函数 $\rho(\gamma, \theta, t)$ 与 γ 有关, 系统总的分布函数需要对所有非均匀参量求和, $\rho(\theta, t) = \sum_i \rho(\gamma_i, \theta, t)$. 例如, 如果 $\{\gamma_i = \omega_i\}$,

则分布函数为

$$\rho(\theta, t) = \sum_i \rho(\omega_i, \theta, t) = \int \rho(\omega, \theta, t) g(\omega) d\omega, \quad (30)$$

其中 $g(\omega)$ 为自然频率分布. $\rho(\theta, t)$ 包含了振子系统集体行为的所有信息. 在同步问题宏观层面, 我们最为关注序参量 $\alpha = \{\alpha_n\}$. 对于均匀系统即没有非均匀参量 γ , 利用分布函数, 序参量可表为如下积分:

$$\alpha_n = \int \exp(in\theta) \rho(\theta, t) d\theta. \quad (31)$$

如果系统有非均匀性, 则需要同时对非均匀参量 γ 求和. 例如, 如果非均匀参量是振子自然频率, 则有

$$\begin{aligned} \alpha_n &= \int \alpha_n(\omega, t) g(\omega) d\omega \\ &= \int \exp(in\theta) \rho(\omega, \theta, t) g(\omega) d\omega d\theta. \end{aligned} \quad (32)$$

可以看到, n 阶序参量 α_n 实际上就是 $\exp(in\theta)$ 的统计平均, 或称为 $\exp(i\theta)$ 的 n 阶矩. 由于宏观量的各阶矩描述与分布函数描述等价, 可由一方信息推知另外一方的信息. 还可以看到, 广义序参量实际上就是分布函数 $\rho(\theta, t)$ 的傅里叶变换系数.

4.2 Ott-Antonsen 拟设

下面先考虑全同振子系统的序参量运动方程. 利用序参量的定义 (27) 式, 对其进行时间求导, 并利用振子运动方程 (26) 式可以得到

$$\dot{\alpha}_n = \frac{in}{N} \sum_{j=1}^N \exp(in\theta_j) F(\alpha, \theta_j, \beta). \quad (33)$$

由于函数 $F(\alpha, \theta, \beta)$ 是循环变量相位 θ 的 2π 周期函数, 因此可做傅里叶展开:

$$F(\alpha, \theta_j, \beta) = \sum_{k=-\infty}^{\infty} f_k(\alpha, \beta) \exp(ik\theta_j), \quad (34)$$

将展开式代入运动方程 (33) 中可得序参量运动方程为

$$\dot{\alpha}_n = in \sum_{k=-\infty}^{\infty} f_k(\alpha, \beta) \alpha_{k+n}, \quad n = 1, 2, \dots \quad (35)$$

考虑最简单的耦合形式, 即 (34) 式的傅里叶分解中只需包含最低阶项:

$$\begin{aligned} F(\alpha, \theta) &= f_1(\alpha) \exp(i\theta) + f_{-1}(\alpha) \exp(-i\theta) \\ &\quad + f_0(\alpha), \end{aligned} \quad (36)$$

则序参量运动方程 (35) 简化为

$$\dot{\alpha}_n = in[f_1(\alpha)\alpha_{n+1} + \bar{f}_1(\alpha)\alpha_{n-1} + f_0(\alpha)\alpha_n], \quad (37)$$

其中 $n \geq 0$, $\alpha_{-n} = \bar{\alpha}_n$.

广义序参量可看成是相位振子的集体坐标变量, 这意味着相位运动方程 (26) 通过上述变换可以化为运动方程 (35) 和 (37), 它们都是一组耦合的序参量方程, 处理该序参量方程组的难度等价于相位运动方程 (26), 对其的简化需要新的条件.

方程 (37) 显然存在一个平庸的非相干解 $\alpha_n \equiv 0$, 对应于耦合振子的非同步态. 随着耦合强度的增加, 耦合振子会产生同步, 整体运动在相空间也会塌缩到一个低维空间中. 在广义序参量空间来看, 系统也必然会在低维空间运动. 根据协同学原理, 这些广义序参量中可能只有少数为慢变量, 其余为快变量, 其中慢变的广义序参量会成为系统状态的真实序参量. 由于快变量的变化都依赖于慢变量, 因此在发生同步转变的区域附近, 各阶序参量之间应存在一定的关系, 它们均依赖于慢变的序参量. 一种最简单的可能情形是所有的各阶序参量 α_n 都依赖于 α_1 , 不妨设为 $\alpha_n = G(\alpha_1, n)$. 将其代入运动方程 (37) 中并比较每一阶函数的傅里叶展开系数可以得到 [56]

$$\alpha_n = G(\alpha_1, n) = \alpha_1^n. \quad (38)$$

则 (37) 式的多个运动方程可简并为单一方程,

$$\dot{\alpha}_1 = i[f_1(\alpha_1)\alpha_1^2 + \bar{f}_1(\alpha_1) + f_0(\alpha_1)\alpha_1]. \quad (39)$$

(38) 式正是 Ott-Antonsen (OA) 拟设 (ansatz) [47,48].

4.3 Poisson 和不变子流形

OA 拟设有什么物理意义呢? 下面通过讨论分布函数来进行分析.

由于分布函数是 2π 循环变量相位的函数, 可将其进行傅里叶展开, 而展开系数即为广义序参量 α_n :

$$\rho(\theta, t) = \frac{1}{2\pi} \left[1 + \sum_{n=1}^{\infty} (\bar{\alpha}_n(t)e^{in\theta} + \alpha_n(t)e^{-in\theta}) \right]. \quad (40)$$

一般来说, 知道了各阶傅里叶系数即广义序参量 α_n , 就可以利用上述求和来得到分布函数, 通常这需要无穷阶的傅里叶系数.

OA 拟设认为相位分布函数 ρ 的傅里叶展开系数即广义序参量 α_n 相互之间并不独立, 且满足幂函数关系:

$$\alpha_n(t) = \alpha_1^n(t), \quad \bar{\alpha}_n(t) = \bar{\alpha}_1^n(t). \quad (41)$$

将其代入 (40) 式可以得到

$$\rho(\theta, t) = \frac{1}{2\pi} \left\{ 1 + \sum_{n=1}^{\infty} [\bar{\alpha}_1^n(t) \exp(in\theta) + \alpha_1^n(t) \exp(-in\theta)] \right\}, \quad (42)$$

(42) 式右边的求和为幂级数, 可以得到振子分布为泊松和形式

$$\rho(\theta, t) = \frac{1}{2\pi} \frac{1 - r^2}{1 - 2r \cos(\theta - \Theta) + r^2}, \quad (43)$$

其中 r 为序参量 $\alpha_1(t)$ 的幅度, Θ 为集体相位. 因而分布 $\rho(\theta, t)$ 完全由 $\alpha_1(t)$ 决定.

由于具有关系 (38) 式的解满足简并的运动方程 (39), 则在动力学演化过程中形式 (38) 一直得到满足. 振子的分布虽然随着系统的演化而改变, 但将始终具有泊松和分布的形式. 如果系统的初始相密度分布满足泊松分布, 那么不管在任何时刻系统的相密度分布将始终保持着这一性质. 序参量关系式 (38) 式及简并运动方程 (39) 被称为动力系统的 Poisson 和不变子流形. 这个不变子流形一个很重要的特点是虽然 $\alpha_1(t)$ 可以不含时, 也可以含时, 但分布形式随时间演化保持不变. 这一结果将 Kuramoto 自洽理论仅仅讨论定态的结果拓展至一般情形.

4.4 非全同振子系统的 OA 拟设

对于振子非全同的情形, 若振子自然频率各不相同, 即 $\{\gamma_i = \omega_i\}$, 设它们满足分布 $g(\omega)$. 那么当 $N \rightarrow \infty$ 时, 引入密度分布函数 $\rho(\omega, \theta, t)$, 广义序参量相应写为 (32) 式, 其中 $\alpha_n(\omega, t)$ 为 $\rho(\omega, \theta, t)$ 的 n 阶傅里叶展开式, 也可以理解为自然频率在 $\omega \rightarrow \omega + d\omega$ 之间的局域序参量. 利用连续性方程可以得到如下递归方程:

$$\dot{\alpha}_n = in \sum_{j=-\infty}^{\infty} f_j(\alpha, \beta, \omega) \alpha_{j+n}(\omega, t). \quad (44)$$

同样如果耦合函数只包含一阶傅里叶系数, 则可以引入如下局域序参量的 Ott-Antonsen 拟设:

$$\alpha_n(\omega, t) = \alpha_1^n(\omega, t) \quad (45)$$

是方程 (44) 的一组特解. 一阶序参量为

$$\alpha_1(t) = \int \alpha_1(\omega, t) g(\omega) d\omega. \quad (46)$$

当分布函数 $g(\omega)$ 的形式为 ω 的有理分式时, 可将 ω 从实轴延拓到复 ω 平面中去, 在不引起发散的情形下 ($|\alpha_1(t)| \leq 1$) 直接得到 $\alpha_1(t)$ 的演化方程.

5 微观动力学对称性与 Watanabe-Strogatz 变换

耦合相振子系统在振子数 $N \gg 1$ 时的微观动力学研究会变得很困难. OA 方法给出了一种有效的将高维耦合振子动力学降维的方案, 但高维系统降维到二维的序参量空间来研究是有前提和成立条件的.

OA 方法之所以成功, 其关键在于 Kuramoto 系统本身的微观动力学具有部分可积性. 这要从 1994 年 Watanabe 和 Strogatz 的工作谈起. 他们研究了全局耦合约瑟夫森结方程 (与 Kuramoto 模型系统同类). 研究发现, N 维方程中的每一条轨迹只能局限在一个三维子空间中, 这意味着原始的高维微观态可以通过一定方法降维至低维的宏观态^[46], 这就需要他们提出的后来被称为 Watanabe-Strogatz(WS) 变换的方法来实现. WS 变换的提出在一定程度上为人们寻求高维动力系统的低维解提供了方向, 但早期没有引起人们足够重视, 且数学与物理意义均不明确. 2009 年, 受 Ott 与 Antonsen 工作的启发, Marvel 等^[49,50] 成功将这类问题的解推广到一般形式, 给出了 WS 变换的数学意义, 并清晰地给出 OA 拟设的数学依据.

5.1 WS 变换

WS 变换实际上来自于复数 Möbius 变换^[57]. 定义复分数的 Möbius 变换 $F: \mathbb{C} \rightarrow \mathbb{C}$ 为

$$F(z) = \frac{az + b}{cz + d}, \quad a, b, c, d \in \mathbb{C}, \quad ad - bc \neq 0 \quad (47)$$

Möbius 变换有很多很好的性质, 其中之一就是存在且保持从直线到直线、从圆到圆的映射. 由所有满足上述条件的变换函数可以构成一个群, 称为 Möbius 变换群,

$$M(\mathbb{C}_\infty) = \left\{ F: \mathbb{C}_\infty \rightarrow \mathbb{C}_\infty \mid F(z) = \frac{az + b}{cz + d}, \quad ad - bc \neq 0 \right\}. \quad (48)$$

考虑一个子群, 该子群包含了那些将单位开圆盘上的复数一对一地映射到自身的所有分数线性变换.

将耦合函数 (36) 式对应的动力学方程重新写

为如下形式:

$$\dot{\theta}_j = f \exp(i\theta_j) + g + \bar{f} \exp(-i\theta_j), \quad (49)$$

其中 $j = 1, 2, \dots, N$, $f = f_1(\alpha)$ 为光滑复函数, \bar{f} 是其复共轭, $g = f_0(\alpha)$ 为实函数, 它们均不依赖于指标 j . 方程 (49) 定义了一个 N 维动力系统. Marvel 和 Strogatz^[49] 指出, 方程 (49) 的解满足如下 Möbius 群在复空间内单位圆上的含时 Möbius 变换 M_t :

$$\exp[i\theta_j(t)] = M_t[\exp(i\varphi_j)], \quad (50)$$

其中含时 Möbius 变换表达为

$$M_t(w) = \frac{\exp[i\psi(t)]w + \alpha(t)}{1 + \bar{\alpha}(t)\exp[i\psi(t)]w}, \quad (51)$$

这里 $\{\varphi_j, j = 1, 2, \dots, N\}$ 是系统的一组运动常数, $\psi(t)$ 是实参量函数, $\alpha(t)$ 是复函数, $|\alpha(t)| \leq 1$, $\bar{\alpha}(t)$ 是 $\alpha(t)$ 的复共轭. (50) 式和 (51) 式的变换就被称为 WS 变换. 对任意时刻 t 都对应一个变换 M_t , 这些变换构成的集合满足群的性质, 全体变换的集合 $\{M_t\}$ 构成一个代数系统即 Möbius 变换群. 系统 (49) 式的演化完全被变换 (50) 式和 (51) 式所支配.

5.2 三维动力学

方程 (49) 满足变换 (50) 式和 (51) 式, 说明了系统任一时刻的状态 $\{\theta_j(t)\}$ 可由一组运动不变量 $\{\varphi_j\}$ 在 WS 变换 $\{\alpha(t), \psi(t)\}$ 的作用下完全确定. 而运动常数或不变积分 $\{\varphi_j\}$ 的存在则反映了 WS 变换下系统的可积性. Möbius 群本质上是一个三参量 ($\psi(t), \text{Re}\alpha(t), \text{Im}\alpha(t)$) 的李群, 相轨迹 $\{\theta_j(t)\}$ 被 $\{\psi(t), \alpha(t)\}$ 以及 $\{\varphi_j\}$ 唯一地确定.

下面从代数方程出发去推导 $\psi(t), \alpha(t)$ 的运动方程. 根据方程 (50) 有

$$\theta_j(t) = -i \ln M_t(\exp[i\varphi_j(t)]), \quad (52)$$

则对时间的导数为

$$\begin{aligned} \dot{\theta}_j(t) &= -i \frac{1}{M_t[\exp(i\varphi_j)]} \frac{dM_t[\exp(i\varphi_j)]}{dt} \\ &= R \exp(i\varphi_j) + \frac{\dot{\psi} + i\dot{\alpha}\alpha - \alpha(i\dot{\alpha} - \bar{\alpha}\dot{\psi})}{1 - |\alpha|^2} \\ &\quad + \bar{R} \exp(-i\varphi_j), \end{aligned} \quad (53)$$

其中

$$R \equiv \frac{i\dot{\alpha} - \bar{\alpha}\dot{\psi}}{1 - |\alpha|^2}, \quad (54)$$

与 (49) 式对比可得

$$f = R = \frac{i\dot{\alpha} - \bar{\alpha}\dot{\psi}}{1 - |\alpha|^2}, \quad (55a)$$

$$g = \frac{\dot{\psi} + i\bar{\alpha}\dot{\alpha} - \alpha(i\dot{\alpha} - \bar{\alpha}\dot{\psi})}{1 - |\alpha|^2}. \quad (55b)$$

解出 $\dot{\alpha}$ 和 $\dot{\psi}$, 则有

$$\begin{aligned} \dot{\alpha} &= i(f\alpha^2 + g\alpha + \bar{f}), \\ \dot{\psi} &= f\alpha + g + \bar{f}\bar{\alpha}, \end{aligned} \quad (56)$$

因此通过 WS 变换, 可以将原始的 N 维相振子方程约化为三维的闭合方程, 而振子数目 N 可以是有限的, 也可以是无限大.

5.3 从 WS 变换到 OA 拟设

WS 变换是一个从 N 维相空间向三维相空间的严格动力学变换. 不同于统计方法例如基于序参量的 OA 拟设, WS 方法并不依赖于任何近似条件或者特定状态, 而是将整个系统的动力学完整而严格地投影到低维系统中 [56]. 当振子数 $N \rightarrow \infty$ 时, OA 拟设下单一复序参量 α_1 满足的方程为 (39) 式, 这是一个二维的闭合实方程. 可以看到, WS 变换下的三维方程 (56) 式中第一式即复序参量方程 (39), 因此一个自然的问题是, 这两者之间是否存在某种联系? 在什么情况下 WS 变换可以退化到 OA 拟设? 下面来分析一下 WS 变换和 OA 拟设这两种降维方案之间的区别和联系.

对于 N 振子系统 (49) 式, 通过 WS 变换可以将序参量 $\alpha_1(t)$ 重新表示为

$$\begin{aligned} \alpha_1(t) &= \frac{1}{N} \sum_{j=1}^N \exp[i\theta_j(t)] \\ &= \frac{1}{N} \sum_{j=1}^N \frac{\exp[i\psi(t)] \exp(i\varphi_j) + \alpha(t)}{1 + \bar{\alpha}(t) \exp[i\psi(t)] \exp(i\varphi_j)}, \end{aligned} \quad (57)$$

从一组不变量 $\{\varphi_j\}$ 出发, 若初始状态已知, 则系统的动力学情况可以由三维方程 (56) 描述. 取特定均匀分布 $\{\varphi_j = 2\pi(j-1)/N\}$, 则序参量 (57) 式可简化为

$$\alpha_1(t) = \alpha(t)(1 + I), \quad (58)$$

方程中

$$I = \frac{1 - |\alpha(t)|^{-2}}{1 \pm (\bar{\alpha}(t) \exp[i\psi(t)])^{-N}}, \quad (59)$$

其中“-”对应于 N 为偶数的情况, “+”对应于 N 为奇数的情况. 当 $N \gg 1, I \ll 1$, 序参量就可以取近似 $\alpha_1(t) \approx \alpha(t)$. 因而, 当运动常数 $\{\varphi_j\}$ 取均匀测度时, $\alpha_1(t)$ 与 $\alpha(t)$ 的方程完全一致, 方程 (56) 中 $\alpha(t)$ 和 $\psi(t)$ 的演化解耦, WS 变换的三维方程流形退化到二维 OA 流形 (39) 式.

需要指出的是, 当运动常数 $\{\varphi_j\}$ 不取均匀测度时, $\alpha(t)$ 和 $\psi(t)$ 的时间演化就会相互耦合而无法简单解耦, $\alpha(t)$ 的动力学行为将会受到 $\psi(t)$ 的影响, 此时系统的动力学在三维空间中进行, $\alpha(t)$ 和 $\psi(t)$ 的运动会变得非常复杂 [50,58]. 采用直角坐标, 设 $\alpha = x+iy, f = \text{Re}f+i\text{Im}f$, 则三维方程 (40) 化为

$$\begin{aligned} \dot{x} &= -uy + \text{Im}f(1 - x^2 - y^2), \\ \dot{y} &= ux + \text{Re}f(1 - x^2 - y^2), \\ \dot{\psi} &= u, \end{aligned} \quad (60)$$

这里 $u = 2x \text{Re} f + g - 2y \text{Im} f$.

图 2 给出了方程 (60) 的运动轨道在 α 平面的 Poincare 截面落点分布. 可以很清楚看到在 $|\alpha|$ 比较小的区域均为环面, α 的演化为定态或周期振荡. 但在 $|\alpha|$ 较大的范围, 运动轨道是不规则、混沌的 [58]. 在环面区域, 类似于哈密顿系统, 环面对应于低维的运动, OA 拟设可以成立. 在混沌区域, $\alpha(t)$ 和 $\psi(t)$ 的演化相互耦合, OA 拟设不适用.

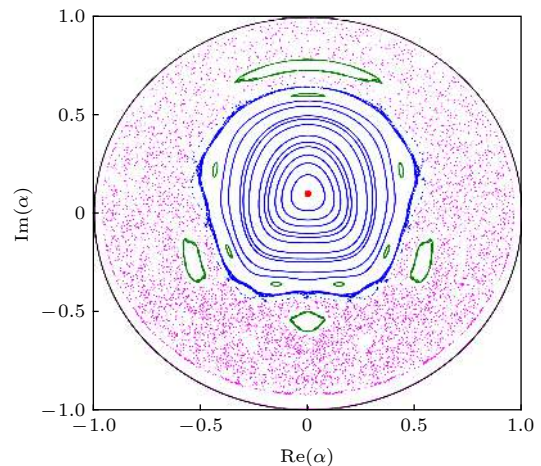


图 2 序参量 α 相空间的 Poincare 截面落点分布, 可以看到闭合环面和混沌散点

Fig. 2. The Poincaré section of the order parameter α in phase space, where one can find the closed tori and chaotic scattered points.

6 序参量理论在几个同步问题研究中的应用

6.1 经典 Kuramoto 模型的同步相变

对于经典 Kuramoto 模型, 即平均场耦合振子系统, Kuramoto 已成功地通过自洽方程理论得到了临界点及其临界行为. 下面利用 OA 拟设来研究该同步转变问题. 由于自然频率非均匀, 因此需用非均匀参数情况下的拟设. 利用序参量

$$z(t) = N^{-1} \sum_{j=1}^N \exp(i\theta_j) = \int \alpha_1(\omega, t) g(\omega) d\omega$$

定义式可知, 耦合函数中的分量分别为

$$f_1 \equiv -Kz/(2i), \quad \bar{f}_1 \equiv Kz/(2i), \quad f_0 \equiv \omega, \quad (61)$$

即 Kuramoto 模型耦合函数只包含到一阶傅里叶分量. 由 (39) 式可以得到

$$\dot{\alpha}_1(\omega, t) = [2i\omega\alpha_1(\omega, t) + Kz(t) - K\bar{z}(t)\alpha_1^2(\omega, t)]/2, \quad (62)$$

如果 $g(\omega)$ 为 Lorentz 分布

$$g(\omega) = 1/[\pi(\omega^2 + 1)], \quad (63)$$

由 (63) 式有

$$z(t) = \int_{-\infty}^{\infty} \frac{\alpha_1(\omega, t)d\omega}{\pi(\omega^2 + 1)}. \quad (64)$$

(64) 式的积分可以将 ω 延拓到复平面的上半平面进行, 利用留数定理可以得到 $z(t) = \alpha_1(\omega = i, t)$, 将其代入到 (63) 式和 (64) 式中有

$$\dot{z} = z(K - 2 - K|z|^2)/2, \quad (65)$$

此即序参量满足的动力学方程. 这个复方程可以写成振幅和幅角两个实方程. 可以看到, 该动力系统存在临界点 $K_c = 2$, 当 $K \leq K_c$ 时, 系统有唯一不动点 $z \equiv 0$, 此即非相干态; 当 $K \geq K_c$ 时, $z \equiv 0$ 解失稳, 系统分岔到新解

$$|z| = \sqrt{(K - K_c)/K}. \quad (66)$$

该解即同步解. 对比 Kuramoto 自洽方程得到的结果可见, 二者完全一致.

6.2 奇异态研究

奇异态 (chimera state) 是近年来发现的一类对称性破缺导致的时空斑图涌现行为, 它描述了结构全同的单元 (如相振子, 其振子的自然频率以及耦合方式都相同) 在非局域耦合下会产生相干

(coherence) 和非相干 (incoherence) 共存的态^[59,60]. 这种对称性自发破缺的现象在很多生物系统中可以看到, 例如海豚与其他海洋哺乳动物、迁徙的候鸟等都具有一类有趣的半脑睡眠现象^[61], 即它们可以在左 (右) 脑休息的时候让右 (左) 脑保持清醒的状态. 人们在 Kuramoto 模型以及其他很多振子系统特别是神经系统中均发现了奇异态^[62-67], 并在光学混沌实验^[68] 和化学 Belousov-Zhabotinsky (BZ) 反应实验^[69] 实现.

结合相振子系统, 我们考虑如下的空间一维 Ginzburg-Landau 方程

$$\frac{\partial \phi(x, t)}{\partial t} = \omega - \int_{\pi}^{\pi} G(\Delta x) \sin(\Delta \phi + \alpha) dx', \quad (67)$$

这里 $\phi(x, t)$ 为 t 时刻位于空间 x 处振子的相位, ω 为振子自然频率, 设为空间均匀且 $\omega = 0$. 积分代表空间不同位置振子间的耦合, $\Delta x = x - x'$, $\Delta \phi = \phi(x, t) - \phi(x', t)$. α 为振子间的相移 (阻挫), 设 $0 < \alpha < \pi/2$. 非负耦合核函数 $G(x) \geq 0$ 反映空间不同位置振子之间的非局域相互作用, 通常设为归一化且随 $|x|$ 增加而衰减的偶函数:

$$\int_0^{\infty} G(x) dx = 1, \quad G(-x) = G(x).$$

下面用 OA 拟设方法来研究系统 (67) 式的奇异态动力学^[62,70,71]. 设振子分布函数为 $f(x, \phi, t)$, 它满足连续性方程

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial \phi}(vf) = 0, \quad (68)$$

其中相速度为

$$v(x, t) = \omega - \int dx' G(\Delta x) \int_{\pi}^{\pi} d\phi' \sin(\Delta \phi + \alpha) f(x', \phi', t). \quad (69)$$

复序参量 $Z(x, t)$ 可表示为

$$\begin{aligned} Z &\equiv R(x, t) \exp[i\Phi(x, t)] \\ &= \int dx' G(x - x') \int_{\pi}^{\pi} d\phi \exp[i\phi(x')] f(x', \phi', t). \end{aligned} \quad (70)$$

利用序参量表达式 (70) 式, (69) 式的相速度 $v(x, t)$ 可表示为

$$\begin{aligned} v(x, t) &= \omega - \frac{1}{2i} (\bar{Z}(x, t) \exp[i(\phi + \alpha)] \\ &\quad - Z(x, t) \exp[-i(\phi + \alpha)]). \end{aligned} \quad (71)$$

考虑到相位 ϕ 的 2π 周期性, 可将分布函数表为相

位 ϕ 的傅里叶级数形式:

$$f(x, \phi, t) = \frac{1}{2\pi} \left\{ 1 + \sum_{n=1}^{\infty} [h_n(x, t) \exp(in\phi) + \bar{h}_n(x, t) \exp(-in\phi)] \right\}, \quad (72)$$

利用 OA 拟设^[47], 系数 h_n 相互之间不独立, 它们满足幂律关系:

$$h_n(x, t) = h^n(x, t), \quad \bar{h}_n(x, t) = \bar{h}^n(x, t), \quad (73)$$

这里 $h(x, t) = h_1(x, t)$. 将 (73) 式代入 (72) 式可得

$$f(x, \phi, t) = \frac{1}{2\pi} \left\{ 1 + \sum_{n=1}^{\infty} [h^n(x, t) \exp(in\phi) + \bar{h}^n(x, t) \exp(-in\phi)] \right\}. \quad (74)$$

将 (74) 式代入连续性方程 (68) 式, 并比较两边 $\exp(i\phi)$ 的不同阶系数可得

$$\frac{\partial h}{\partial t} = -i\omega h + \frac{1}{2} [\bar{Z} \exp(i\alpha) - Z \exp(-i\alpha) h^2]. \quad (75)$$

利用 (74) 式也可将序参量重新表达为

$$Z(x, t) = \int dx' G(x - x') \bar{h}(x', t). \quad (76)$$

利用 (75) 式和 (76) 式可得到序参量的演化动力学行为. 位于 x 处振子的相位长时分布函数为

$$f(\phi) \sim \sum_{n=-\infty}^{\infty} h^n(x) \exp(in\phi). \quad (77)$$

当 $|h(x, t)| < 1$ 时, 对 (77) 式求和并做归一化可得

$$f(\phi) = \frac{1 - |h|^2}{2\pi[1 - 2|h| \cos(\phi - \arg h) + |h|^2]}, \quad (78)$$

此处 \arg 代表复函数 h 的幅角. 上述分布函数为与 (43) 式形式相同的 Poisson 核函数, 其中心位于 $\arg h$ 处, $|h|$ 刻画分布的非均匀性. 当 $|h|=0$ 时, $f(\phi)$ 为均匀分布; 当 $0 < |h| < 1$ 时, $f(\phi)$ 为单峰分布; 当 $|h|=1$ 时, 由于分布 (78) 式分子分母均趋于零, 可求极限得 $f(\phi)$ 简并为 δ 函数 $f(\phi) = \delta(\phi - \arg h)$, 它代表所有振子的锁相.

系数 $h(x, t)$ 也可以看作是在空间 x 处 $\exp(i\phi)$ 的统计平均, 即

$$h(x, t) = \int_{-\pi}^{\pi} d\phi f(x, \phi, t) \exp(i\phi).$$

与 (70) 式定义的序参量 $Z(x, t)$ 相比较, $h(x, t)$ 没有对空间的积分, 即没有考虑非局域效应. 因此 $h(x, t)$ 量度的是空间 x 处附近的同步相干性, 是一

种局域序参量. 若对某一 x 处有 $|h(x, t)|=1$, 则在 $(x - \varepsilon, x + \varepsilon)$ ($\varepsilon \ll 1$) 范围内的振子会处于锁相 (相干态), 而当 $|h(x, t)| < 1$ 时, 则在 $(x - \varepsilon, x + \varepsilon)$ 范围内的振子处于非相干态, 以此可以给出两种态共存的奇异态.

在定态情况下, 进一步利用 (76) 式可以得到序参量 $R(x)$ 的自治方程, 进一步可以确定其空间分布. 当然, 也可用 Kuramoto 自治方程理论来进行讨论, 可以得到与上述 OA 方法在定态时一致的结果. 进一步的计算需要借助于数值方法, 详细讨论可见我们的专著^[51] 与综述文章^[59,60].

6.3 高阶序参量及其动力学

很长时间以来, 人们聚焦于 Kuramoto 模型中的简单耦合函数情形开展研究. 在多数情况下, 振子间相互作用并不是简单的一阶正弦耦合函数. 对于一般的耦合函数, 在第 4 节中已经看到, 不同阶序参量之间是复杂的相互依赖关系. Daido^[72] 早在 20 世纪 90 年代初就通过对一般耦合函数的 Fourier 展开提出了广义序参量, 但没有进一步讨论不同序参量之间的关系, 而是使用传统的 Kuramoto 自治方程方法进一步讨论. Pikovsky 和 Rosenblum^[73] 讨论了如下具有相同自然频率耦合振子的 Kuramoto-Daido 模型:

$$\dot{\theta}_k = \omega + \frac{1}{N} \sum_{j=1}^N h(\theta_j - \theta_k), \quad (79)$$

并引入了高阶序参量 $\{\alpha_n\}$. Pikovsky 在该工作中早于 Ott 等^[47] 提出了序参量之间的依赖关系 (拟设条件), 即高阶序参量 $\{\alpha_n, |n| \geq 2\}$ 均依赖于 $\alpha_{\pm 1}$. 把 (79) 式中的耦合函数写为 Fourier 级数形式, 则动力学方程可写为

$$\dot{\theta}_k = \omega + \sum_n h_n \alpha_n \exp(-in\theta_k), \quad (80)$$

其中 $\{\alpha_n\}$ 为 (27) 式定义的广义序参量. 考虑只有其中第 l 阶 Fourier 分量耦合的情形, 则动力学方程可写为

$$\dot{\theta}_k = \omega + \text{Im}[H(t) \exp(-il\theta_k)], \quad (81)$$

其中 $H(t)$ 依赖于序参量 α_l .

下面讨论利用 WS 变换来进行理论分析^[74]. 动力学方程 (81) 可以重新写为

$$\frac{d[\exp(i l \theta_k)]}{dt} = i l \exp(i l \theta_k) \omega + \frac{l}{2} [H - \bar{H} \exp(i 2 l \theta_k)]. \quad (82)$$

注意到 Möbius 变换不唯一, 而其逆变换唯一, 则利用 Möbius 逆变换

$$\exp(i l \theta_k) = \frac{\alpha(t) + \exp(i \varphi_k)}{1 + \bar{\alpha}(t) \exp(i \varphi_k)}, \quad (83)$$

(83) 式可变换为 $\{\varphi_k\}$, α 及其时间导数 $\{\dot{\varphi}_k\}$, $\dot{\alpha}$:

$$\begin{aligned} \dot{\alpha} &= l(i \omega \alpha + H/2 - \bar{H} \alpha^2/2), \\ \dot{\varphi}_k &= l\{\omega + \text{Im}[H(t)\bar{\alpha}]\}, \end{aligned} \quad (84)$$

可以看到 $\{\varphi_k\}$ 的方程右边不依赖于 k , 说明变换后的相角 $\{\varphi_k\}$ 均以相同角速度运动, 因此可引入新变量 $\psi(t) = \varphi_k(t) - \varphi_k^0$, 其中 $\{\varphi_k^0\}$ 为一组运动常数. 显然 $\dot{\psi} = \dot{\varphi}_k$, 说明 Möbius 变换映射可写为

$$M_l : \exp(i l \theta_k) = \frac{\alpha + \exp[i(\varphi_k^0 + \psi)]}{1 + \bar{\alpha} \exp[i(\varphi_k^0 + \psi)]}, \quad (85)$$

与一阶耦合函数相比, 这里形式上只是多出了 l 因子.

很多具体问题可以应用上述结果加以讨论. 例如近年来得到关注的三体相互作用平均场振子系统 [75,76]

$$\dot{\theta}_k = \omega + \frac{1}{N^2} \sum_{i,j=1}^N \sin(\theta_i + \theta_j - 2\theta_k), \quad (86)$$

该系统对应于方程 (81) 中 $l = 2$, $H(t) = \alpha_1^2$ 的情形, 因而可利用序参量方程 (84) 进行讨论.

再例如具有阻挫的二阶简谐耦合振子系统 [58]

$$\dot{\theta}_k = \omega + \frac{1}{N} \sum_{j=1}^N \sin[2(\theta_j - \theta_k + \gamma)], \quad (87)$$

其中 γ 为相移因子, 代表阻挫效应. 对于这种情况, $l = 2$, $H(t) = \exp(2i\gamma) \alpha_2$. 二阶或高阶简谐耦合函数在很多系统中都存在, 因此近年来的研究很多 [58,77-79], 也有相应的一些实验研究 [80-83], 这里不再展开.

7 结 论

综上所述, 大自由度复杂系统的同步是一类典型、基本的涌现行为. 这种涌现行为的出现是大量个体或自由度共同参与、通过相互作用自组织形成的整体有序. 有序意味着复杂系统低维宏观行为的出现, 它往往是以序参量的出现作为标志的, 可借

助于统计物理学思想和方法来加以研究. 就序参量的涌现机制来看, 协同学理论与方法揭示了大量自由度的复杂系统如何通过内部自组织与竞争产生出有序, 其中非常重要的理论任务就是对复杂系统通过集体变量的分析来甄别快慢变量, 并利用支配原理来进行降维, 得到序参量的动力学. 降维方法的精髓在于尺度的可分离性, 特别是时间尺度的快慢可区分性, 这种特性可以其他形式体现出来, 如拓扑流形中中心流形的出现、变量中守恒律的存在等等.

在复杂系统同步的序参量动力学理论方面, 除了 Kuramoto 的建立和求解序参量满足的自治方程方法以外, Ott 和 Antonsen 提出的拟设是近年来应用较为广泛的方案, 该方案首先引入集体坐标即各阶广义序参量, 然后通过各阶量之间的非独立性来达到降维的目标. 这里面就隐含了绝热消去快变量的思想. 需要指出的是, OA 拟设在很多情况下是一种推测, 且在很多情况下不只是一阶序参量是慢变量, 另外的高阶序参量也可能是慢变量, 这需要从原始的各阶序参量方程出发, 利用协同学基本原理来进行分析.

下面讨论一下对于实际系统同步作为整体动力学的研究分析. 利用序参量动力学对同步的研究既可以是理论建立模型的研究, 也可以是对实际问题利用的降维方法进行研究. 相比于基于模型的理论研究, 对于实际复杂系统有序行为的序参量动力学重构和研究显然有着更为重大的意义和价值. 对于很多实际中的复杂动态系统, 在宏观层面构建其序参量动力学是可能的, 但也富于挑战性. 序参量动力学对实际问题的研究是通过采集微观数据, 然后在此基础上进行宏观动力学重构, 因而是一个跨尺度动力学重构问题, 迄今为止这仍然是一个开放的领域, 与复杂网络重构研究具有同样重大的价值, 是对复杂系统动力学方面的重要探索范式, 值得大力开展研究. 就本文讨论的相位动力学及其同步而言, 其目前在应用方面的主战场之一是神经与脑科学 [84]. 我们不妨以此为例进行说明. 神经元的主体微观动力学是电信号的积累发放, 目前实验方面已经可以采集大量实时数据并进行分析整合. 对于脑电波及其各种动力学协同行为的研究, 可以先抽取不同部位采样的脉冲时间序列, 再定义对应的相位, 就可以建立相位的时间序列信息, 据此计算不同序参量, 通过分析序参量的低维行为进行讨论. 就目前的研究来看, 很重要的一个课题是序参

量动力学模型的重构, 这一问题至今无论在神经科学还是其他领域都还没有很大的突破, 但已有一些具有启发性的研究思路. 例如, Zhang 等^[85]和 Chen 等^[86]近几年提出了主超前相位方法、动力学权重因子方法以及网络动力学重构方法, 以此可以分析复杂系统的动力学并结合序参量层次的动力学描述, 完全有可能提供一个可行的图景. 这些工作还需要大力推进.

参考文献

- [1] Glass L, Mackay M C 1988 *From Clocks to Chaos: The Rhythms of Life* (Princeton: Princeton University Press) p10
- [2] Mikhailov A S, Calenbuhr V 2002 *From Cells to Societies: Models of Complex Coherent Action* (Berlin: Springer-Verlag) pp127–154
- [3] Zheng Z G 2004 *Spatiotemporal Dynamics and Cooperative Behaviors in Coupled Nonlinear Systems* (Beijing: Higher Education Press) pp53–85 (in Chinese) [郑志刚 2004 耦合非线性系统的时空动力学与合作行为 (北京: 高等教育出版社)第53—85页]
- [4] Pikovsky A, Rosenblum M, Kurths J 2001 *Synchronization, A Universal Concept in Nonlinear Sciences* (New York: Cambridge University Press) pp1–24
- [5] Strogatz S 2003 *Sync: The Emerging Science of Spontaneous Order* (New York: Hyperion) pp103–152
- [6] Wiener N 1965 *Cybernetics or Control and Communication in the Animal and the Machine* (Cambridge: MIT Press) p65
- [7] Winfree A T 1967 *J. Theor. Biol.* **16** 15
- [8] Kuramoto Y 1975 *Self-entrainment of a Population of Coupled Non-linear Oscillators, in: International Symposium on Mathematical Problems in Theoretical Physics* (Berlin: Springer) pp420–428
- [9] Kuramoto Y 1984 *Chemical Oscillations, Waves and Turbulence* (Berlin: Springer-Verlag) pp60–66
- [10] Kuramoto Y, Nishikawa I 1987 *J. Stat. Phys.* **49** 569
- [11] Acebrón J A, Bonilla L L, Vicente C J P, Ritort F, Spigler R 2005 *Rev. Mod. Phys.* **77** 137
- [12] Bocalletti S, Kurths J, Osipov G, Valladares D, Zhou C 2002 *Phys. Rep.* **336** 1
- [13] Arenas A, Díaz-Guilera A, Kurths J, Moreno Y, Zhou C 2008 *Phys. Rep.* **469** 93
- [14] Rodrigues F A, Peron T K DM, Ji P, Kurths J 2016 *Phys. Rep.* **610** 1
- [15] Osipov G V, Kurths J, Zhou C S 2007 *Synchronization in Oscillatory Networks, Springer Series in Synergetics* (Berlin: Springer-Verlag) pp3–10
- [16] Balanov A, Janson N, Postnov D, Sosnovtseva O 2009 *Synchronization: From Simple to Complex* (Berlin: Springer-Verlag) p21
- [17] Zheng Z G 2011 *Synchronization of Coupled Phase Oscillators* (New York: Nova Science Publishing House) pp293–327
- [18] Komarov M, Pikovsky A 2014 *Physica D* **289** 18
- [19] Komarov M, Pikovsky A 2013 *Phys. Rev. Lett.* **111** 204101
- [20] Hong H, Choi M Y, Kim B J 2002 *Phys. Rev. E* **65** 026139
- [21] Moreno Y, Pacheco A F 2004 *Europhys. Lett.* **68** 603
- [22] Gomez-Gardenes J, Gomez S, Arenas A, et al. 2011 *Phys. Rev. Lett.* **106** 12870
- [23] Peron T K D, Rodrigues F A 2012 *Phys. Rev. E* **86** 056108
- [24] Peron T, de Resende B M F, Mata A S, Rodrigues F A, Moreno Y 2019 *Phys. Rev. E* **100** 042302
- [25] Xu C, Gao J, Sun Y T, Huang X, Zheng Z G 2015 *Sci. Rep.* **5** 12039
- [26] O’Keeffe K P, Krapivsky P L, Strogatz S H 2015 *Phys. Rev. Lett.* **115** 064101
- [27] Bonilla L L, Perez Vicente C J, Rubi J M 1993 *J. Stat. Phys.* **70** 921
- [28] Hong H, Strogatz S H 2011 *Phys. Rev. Lett.* **106** 054102
- [29] Hong H, Strogatz S H 2012 *Phys. Rev. E* **85** 056210
- [30] Tanaka H, Lichtenberg A J, Oishi S 1997 *Phys. Rev. Lett.* **78** 2104
- [31] Gao J, Efstathiou K 2018 *Phys. Rev. E* **98** 042201
- [32] Omata S, Yamaguchi Y, Shimizu H 1988 *Physica D* **31** 397
- [33] Zheng Z G 2001 *Chin. Phys.* **10** 703
- [34] Zhang T X, Zheng Z G 2009 *Phys. B* **18** 1674
- [35] Omel’chenko E, Wolfrum M 2012 *Phys. Rev. Lett.* **109** 164101
- [36] Yeung M K S, Strogatz S H 1999 *Phys. Rev. Lett.* **82** 648
- [37] Childs L M, Strogatz S H 2008 *Chaos* **18** 043128
- [38] Petkoski S, Stefanovska A 2012 *Phys. Rev. E* **86** 046212
- [39] Cohen A H, Holmes P J, Rand R H 1982 *J. Math. Biol.* **13** 345
- [40] Ermentrout G B, Kopell N 1984 *SIAM J. Math. Anal.* **15** 215
- [41] Strogatz S H, Mirollo R E 1988 *Physica D* **31** 143
- [42] Zheng Z G, Hu G, Hu B 1998 *Phys. Rev. Lett.* **81** 5318
- [43] Zheng Z G, Hu B, Hu G 2000 *Phys. Rev. E* **62** 402
- [44] Hu B, Zheng Z G 2000 *Int. J. Bif. & Chaos* **10** 2399
- [45] Pikovsky A, Rosenblum M 2015 *Chaos* **25** 097616
- [46] Watanabe S, Strogatz S H 1994 *Phys. Rev. Lett.* **70** 2391
- [47] Ott E, Antonsen T M 2008 *Chaos* **18** 037113
- [48] Ott E, Antonsen T M 2009 *Chaos* **19** 023117
- [49] Marvel S A, Strogatz S H 2009 *Chaos* **19** 013132
- [50] Marvel S A, Mirollo R E, Strogatz S H 2009 *Chaos* **19** 043104
- [51] Zheng Z G 2019 *Emergence Dynamics in Complex Systems: From Synchronization to Collective Transport* (Beijing: Science Press) pp95–176 (in Chinese) [郑志刚 2019 复杂系统的涌现动力学: 从同步到集体运输 (北京: 科学出版社) 第95—176页]
- [52] Nicolis G, Prigogine I 1989 *Exploring Complexity: An Introduction* (New York: W. H. Freeman) pp21–25
- [53] Prigogine I, Nicolis G 1977 *Self-Organization in Non-Equilibrium Systems* (New Jersey: Wiley) pp63–222
- [54] Haken H 1983 *Synergetics: An Introduction* (3rd Ed.) (Berlin: Springer-Verlag) pp191–224
- [55] Haken H 1983 *Advanced Synergetics* (Berlin: Springer Verlag) pp187–221
- [56] Gao J, Xu C, Sun Y T, Zheng Z G 2016 *Sci. Rep.* **6** 30184
- [57] Tristan N 1997 *Visual Complex Analysis* (Oxford: Clarendon Press) p1
- [58] Xu C, Xiang H, Gao J, Zheng Z G 2016 *Sci. Rep.* **6** 31133
- [59] Yao N, Zheng Z G 2016 *Int. J. Mod. Phys. B* **30** 7 1630002
- [60] Zheng Z G, Zhai Y 2020 *Sci. Sin.-Phys. Mech. Astron.* **50** 010505 (in Chinese) [郑志刚, 翟云 2020 中国科学: 物理学 力学 天文学 **50** 010505]
- [61] Liu Z H 2018 *Chaotic Dynamics Foundation and Its Applications in Brain functions* (Beijing: Science Press) pp234–256 (in Chinese) [刘宗华 2018 混沌动力学基础及其在大脑功能方面的应用 (北京: 科学出版社) 第234—256页]
- [62] Abrams D M, Strogatz S H 2004 *Phys. Rev. Lett.* **93** 174102
- [63] Omel’chenko E, Maistrenko Y L, Tass P A 2008 *Phys. Rev. Lett.* **100** 044105
- [64] Sethia G C, Sen A 2014 *Phys. Rev. Lett.* **112** 144101
- [65] Zakhárova A, Kapeller M, Schol E 2014 *Phys. Rev. Lett.* **112** 154101
- [66] Omelchenko I, Maistrenko Y, Hovel P, et al. 2011 *Phys. Rev. Lett.* **106** 234102

- [67] Gu C, St-Yves G, Davidsen J 2013 *Phys. Rev. Lett.* **111** 134101
- [68] Hagerstrom A M, Murphy T E, Roy R, et al. 2012 *Nat. Phys.* **8** 658
- [69] Tinsley M R, Nkomo S, Showalter K 2012 *Nat. Phys.* **8** 662
- [70] Abrams D M, Strogatz S H 2006 *Int. J. Bif. & Chaos* **16** 21
- [71] Laing C R 2009 *Physica D* **238** 1569
- [72] Daido H 1992 *Prog. Theor. Phys.* **88** 1213
- [73] Pikovsky A, Rosenblum M 2008 *Phys. Rev. Lett.* **101** 264103
- [74] Gong C C 2019 *arXiv* 1909.07718
- [75] Komarov M, Pikovsky A 2015 *Phys. Rev. E* **92** 020901
- [76] Skardal P S, Arenas A 2019 *Phys. Rev. Lett.* **122** 248301
- [77] Skardal P, Ott E, Restrepo J G 2011 *Phys. Rev. E* **84** 036208
- [78] Yuan D, Cui H, Tian J, Xiao Y, Zhang Y 2016 *Commun. Nonlinear Sci. Numer. Simul.* **38** 23
- [79] Wang H, Han W, Yang J 2017 *Phys. Rev. E* **96** 022202
- [80] Czolczynski K, Perlikowski P, Stefanski A, Kapitaniak T 2013 *Commun. Nonlinear. Sci. Numer. Simul.* **18** 386
- [81] Goldobin E, Koelle D, Kleiner R, Kleiner R G 2011 *Phys. Rev. Lett.* **107** 227001
- [82] Goldobin E, Kleiner R, Koelle D, Mints R G 2013 *Phys. Rev. Lett.* **111** 057004
- [83] Kiss I Z, Zhai Y, Hudson J L 2005 *Phys. Rev. Lett.* **94** 248301
- [84] Huang X H, Zheng Z G, Hu G, Wu S, Rasch M J 2015 *New J. Phys.* **17** 035006
- [85] Zhang Z Y, Zheng Z G, Niu H J, Mi Y Y, Wu S, Hu G 2015 *Phys. Rev. E* **91** 012814
- [86] Chen Y, Wang S H, Zheng Z G, Zhang Z Y, Hu G 2016 *EPL* **113** 18005

SPECIAL TOPIC—Statistical physics and complex systems

Synchronization of coupled phase oscillators: Order parameter theory*

Zheng Zhi-Gang^{1)2)†} Zhai Yun³⁾¹⁾²⁾ Wang Xue-Bin¹⁾²⁾

Chen Hong-Bin¹⁾²⁾ Xu Can^{1)2)‡}

1) (*Institute of Systems Science, Huaqiao University, Xiamen 361021, China*)

2) (*College of Information Science and Engineering, Huaqiao University, Xiamen 361201, China*)

3) (*School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China*)

(Received 26 December 2019; revised manuscript received 31 January 2020)

Abstract

Rhythmic behaviors, i.e. temporally periodic oscillations in a system, can be ubiquitously found in nature. Interactions among various rhythms can lead to self-organized behaviors and synchronizations. This mechanism is also responsible for many phenomena such as nonlinear waves, spatiotemporal patterns, and collective behaviors in populations emerging in complex systems. Mathematically different oscillations are described by limit-cycle oscillators (pacemakers) with different intrinsic frequencies, and the synchrony of these units can be described by the dynamics of coupled oscillators. Studies of microscopic dynamics reveal that the emergence of synchronization manifests itself as the dimension reduction of phase space, indicating that synchrony can be considered as no-equilibrium phase transition and can be described in terms of order parameters. The emergence of order parameters can be theoretically explored based on the synergetic theory, central manifold theorem and statistical physics. In this paper, we discuss the order-parameter theory of synchronization in terms of statistical physics and set up the dynamical equations of order parameters. We also apply this theory to studying the nonlinear dynamics and bifurcation of order parameters in several typical coupled oscillator systems.

Keywords: synchronization, order parameter, emergence, bifurcation

PACS: 05.45.-a, 05.45.Xt

DOI: 10.7498/aps.69.20191968

* Project supported by the National Natural Science Foundation of China (Grant Nos. 11875135, 11905068) and the Science and Technology Program of Quanzhou, China (Grant No. 2018C085R).

† Corresponding author. E-mail: zgzheng@hqu.edu.cn

‡ Corresponding author. E-mail: xucan@hqu.edu.cn

极化电场对可激发介质中螺旋波的控制

潘军廷 何银杰 夏远勋 张宏

Control of spiral waves in excitable media under polarized electric fields

Pan Jun-Ting He Yin-Jie Xia Yuan-Xun Zhang Hong

引用信息 Citation: *Acta Physica Sinica*, 69, 080503 (2020) DOI: 10.7498/aps.69.20191934

在线阅读 View online: <https://doi.org/10.7498/aps.69.20191934>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

通过抑制波头旋转消除心脏中的螺旋波和时空混沌

Eliminating spiral wave and spatiotemporal chaos in cardiac tissues by suppressing the rotation of spiral wave tip

物理学报. 2018, 67(24): 248201 <https://doi.org/10.7498/aps.67.20181291>

用晚钠电流终止心脏中的螺旋波和时空混沌

Terminating spiral wave and spatiotemporal chaos in cardiac tissues by using late sodium current

物理学报. 2017, 66(13): 138201 <https://doi.org/10.7498/aps.66.138201>

具有早期后除极化现象的可激发系统中螺旋波破碎方式研究

Spiral wave breakup manner in the excitable system with early afterdepolarizations

物理学报. 2018, 67(9): 090501 <https://doi.org/10.7498/aps.67.20172505>

Hindmarsh-Rose神经元阵列自发产生螺旋波的研究

Spontaneous generation of spiral wave in the array of Hindmarsh-Rose neurons

物理学报. 2018, 67(3): 030502 <https://doi.org/10.7498/aps.67.20172140>

人类心室组织中波斑图引起的后除极化研究

Afterdepolarizations induced by wave pattern in human ventricular tissue

物理学报. 2017, 66(6): 068201 <https://doi.org/10.7498/aps.66.068201>

专题：统计物理和复杂系统

极化电场对可激发介质中螺旋波的控制*

潘军廷¹⁾ 何银杰²⁾ 夏远勋²⁾ 张宏^{2)†}

1) (浙江大学海洋学院, 舟山 316021)

2) (浙江大学物理学系, 浙江近代物理中心, 杭州 310027)

(2019年12月20日收到; 2020年1月17日收到修改稿)

螺旋波在不同的物理、化学和生物系统中普遍存在. 周期外场, 比如极化电场, 尤其是具有旋转对称性的圆极化电场可对螺旋波动力学产生重要影响. 本文综述了极化电场对可激发介质中螺旋波的控制, 包括共振漂移、同步、手征对称性破缺、多臂螺旋波的稳定、次激发介质中的螺旋波、三维回卷波湍流态的控制、心脏组织中螺旋波的去钉扎、心脏组织中螺旋波湍流态的控制等.

关键词: 极化电场, 可激发介质, 螺旋波, 控制

PACS: 05.45.-a, 05.65.+b, 05.45.Xt

DOI: 10.7498/aps.69.20191934

1 引言

许多生物膜、生理组织、化学反应系统都具有“可激发”的性质^[1]. 当此类系统的局部区域处于静息状态时, 对微扰是稳定的; 但对于较强的扰动将有一个快速的响应, 呈现激发状态, 最后回到静息状态. 螺旋波广泛地存在于各类可激发介质, 包括 Belousov-Zhabotinsky(BZ) 化学反应^[2]、一氧化碳在铂金表面的氧化反应^[3]、心脏组织^[4]等系统中. 心脏实验表明, 心律失常与螺旋波的自组织及螺旋波湍流态有密切关系^[4,5], 螺旋波的深入研究对心脏病研究将会产生重要影响. 另外, 与人们所熟悉的行波表现为平行的波前和波背完全一致地传播的情形不同, 螺旋波的波前和波背在顶点处融为一体, 形成具有奇异性结构的螺旋波端点. 此端点构成了螺旋波的组织中心, 是一个时空点拓扑缺陷(相位奇点), 研究螺旋波可为人们研究拓扑缺陷动力学提供一种有效的途径.

电场对螺旋波影响的实验和理论工作比较丰富, 包括直流电场和交流电场. 在 BZ 反应实验中, Steinbock 等^[6]发现直流电场可使螺旋波的漂移速度形成分别与电场方向平行和垂直的两个分量, 且其垂直分量的方向会因螺旋波的手征性不同而各异. Muñuzuri 等^[7]指出, 当交流电场频率成两倍于螺旋波频率时, 螺旋波会产生共振漂移. Zhang 等^[8]研究了直流、交流电场作用下的螺旋波漂移行为, 得到了螺旋波漂移速度的一个近似解析公式.

上述研究的外加电场, 只具有一个方向指向. 由于螺旋波绕其端点旋转, 具有旋转对称性, 人们希望外加的电场也具有旋转对称性. Chen 等^[9]通过把两个相互垂直且频率相同的交流电场叠加, 理论上提出了一个极化电场(图 1), 实现了外电场具有旋转对称性的特性. 通过调节两交流电场间的相位差, 可以得到不同模式的极化电场, 比如线极化、椭圆极化、圆极化电场等(图 2). 近来, Ji 等^[10]在 BZ 反应实验中实现了外加极化电场.

* 国家自然科学基金(批准号: 11675141)资助的课题.

† 通信作者. E-mail: hongzhang@zju.edu.cn

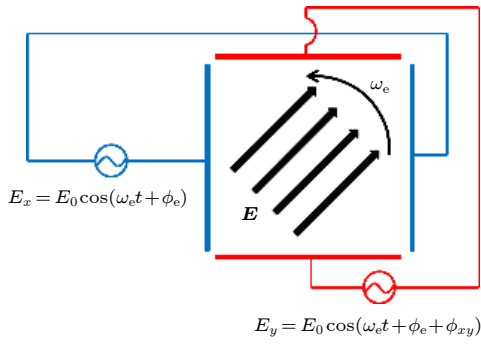


图 1 极化电场示意图^[11] E_x, E_y 表示两个相互垂直的交流电场, E_0, ω_e 分别是交流电场的振幅和频率, ϕ_e, ϕ_{xy} 分别是初相位和相位差

Fig. 1. Realization sketch of a polarized electric field^[11]: E_x, E_y are two ac electric fields perpendicular to each other, where E_0, ω_e are the amplitude and the frequency of the electric field, respectively, and ϕ_e, ϕ_{xy} are the initial phase and the phase difference, respectively.

描述螺旋波动力学的反应扩散方程一般可表述为^[12]

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{1}{\varepsilon} f(u, v) + D_u \nabla^2 u, \\ \frac{\partial v}{\partial t} &= g(u, v) + D_v \nabla^2 v, \end{aligned} \quad (1)$$

其中 u, v 是系统变量; f, g 代表反应函数 (f, g 的具体形式可参见文献 [12]); ε 描述介质的可激发性; D_u, D_v 是扩散系数. 系统 (1) 式可外加一极化电场 $E_x = E_0 \cos(\omega_e t + \phi_e), E_y = E_0 \cos(\omega_e t + \phi_e + \phi_{xy})$, 其中 E_x, E_y 表示两个相互垂直的交流电场.

如果研究的系统是化学系统, 比如 BZ 反应, 极化电场对系统的影响表现为^[13]: 方程 (1) 中, 第一个方程的右边加入 $M_u E_x \frac{\partial u}{\partial x} + M_u E_y \frac{\partial u}{\partial y}$, 第二

个方程的右边加入 $M_v E_x \frac{\partial v}{\partial x} + M_v E_y \frac{\partial v}{\partial y}$, 其中 M_u, M_v 表示离子迁移率. 本文第 2—7 节所研究的系统为化学系统.

如果研究的系统是心脏系统 (D_v 为零), 极化电场只对系统中的缺陷边界产生影响, 表现为在缺陷边界有一零流边界条件^[14]: $\mathbf{n} \cdot \nabla (u + \mathbf{E} \cdot \mathbf{r}) = 0$, 其中 \mathbf{n} 是缺陷边界处的法向矢量, \mathbf{r} 为缺陷边界处的点, u 为跨膜电压, $\mathbf{E} = (E_x, E_y)$ 为极化电场. 本文第 8 节和第 9 节所研究的系统为心脏系统.

2 螺旋波的共振漂移

极化电场下, Chen 等^[9] 研究得出, 当电场频率是螺旋波频率的 2 倍时, 螺旋波呈直线漂移, 即共振漂移. 当电场为圆极化电场, 且与螺旋波同向旋转时, 螺旋波的漂移速率最大; 反之, 当圆极化电场与螺旋波反向旋转时, 螺旋波不发生漂移. 另外, 通过改变极化电场的相位差, 可改变螺旋波的漂移速度.

近来, Li 等^[15] 基于响应函数理论^[16–18], 考虑方程 (1) 受弱电场扰动作用, 研究了弱极化电场下螺旋波的漂移行为. 研究解析给出了螺旋波的漂移速度公式, 且该公式适用于强激发及弱激发介质 (图 3). 当圆极化电场频率是螺旋波频率的 2 倍时, 研究得到顺时针旋转螺旋波的漂移速度公式为

$$\begin{aligned} |\dot{R}| &= 0.5\mu_2 E_0 \sqrt{2(1 + \sin \phi_{xy})}, \\ \Theta &= 2\Phi + \phi_e + 0.5\phi_{xy} + v_2 + 0.75\pi, \\ \phi_{xy} - 1.5\pi &\in (0, 2\pi), \end{aligned} \quad (2)$$

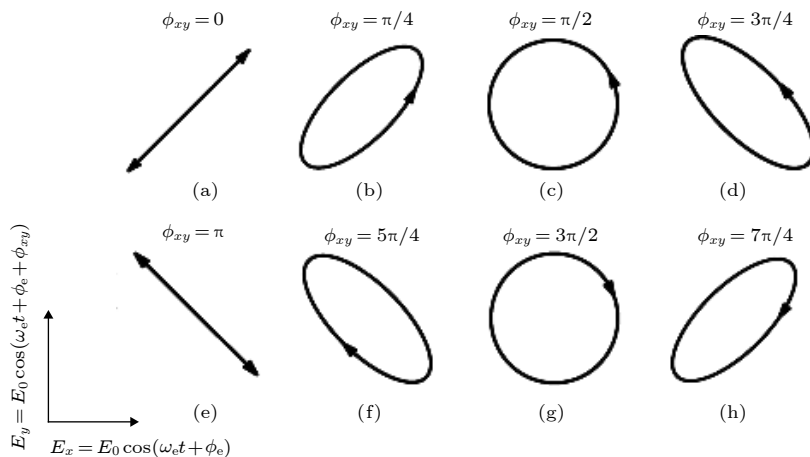


图 2 不同相位差的极化电场^[9]

Fig. 2. Polarized electric fields at different phase differences^[9].

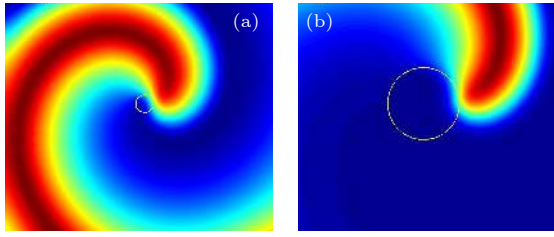


图3 无电场作用下, 顺时针旋转的螺旋波^[15] (a) 强激发介质; (b) 弱激发介质

Fig. 3. Clockwise (cw) rotating spiral waves without electric field^[15]: (a) Highly excitable medium; (b) weakly excitable medium.

其中 $|\dot{R}|$ 表示螺旋波的漂移速率; Θ 是漂移角; Φ 是螺旋波的初相位; μ_2, v_2 是漂移系数, 可由响应函数理论计算得到. 从方程(2)可看出, 当圆极化电场与螺旋波同向旋转时 ($\phi_{xy} = 0.5\pi$), 螺旋波的漂移速率 $|\dot{R}|$ 最大; 当圆极化电场与螺旋波反向旋转时 ($\phi_{xy} = 1.5\pi$), 螺旋波漂移速率 $|\dot{R}| = 0$. 当 Φ 或

ϕ_e 改变时, 漂移角 Θ 会均匀改变, 且 $\Delta\Theta = 2\Delta\Phi$ 及 $\Delta\Theta = \Delta\phi_e$. 当 ϕ_{xy} 改变时, 漂移的方向也会发生改变, 且 $\Delta\Theta = 0.5\Delta\phi_{xy}$. 研究表明, 理论结果与数值结果定量符合(图4).

3 螺旋波的同步

若周期外力使系统的振动模式受迫变为与外力驱动同频, 即出现了锁频现象. 锁频状态中, 外力振幅与频率差呈现出舌状关联, 即存在 Arnold tongue 区域^[19]. Chen 等^[20]数值研究了圆极化电场对螺旋波频率的影响, 发现当圆极化电场和螺旋波同向旋转, 且二者频率较为接近时, 原来处于严格旋转或漫游状态的螺旋波均可被圆极化电场同步(图5), 并与圆极化电场保持着恒定的旋转方向偏角. 锁频后, 圆极化电场强度与频率之间, 出现了 Arnold tongue 区域.

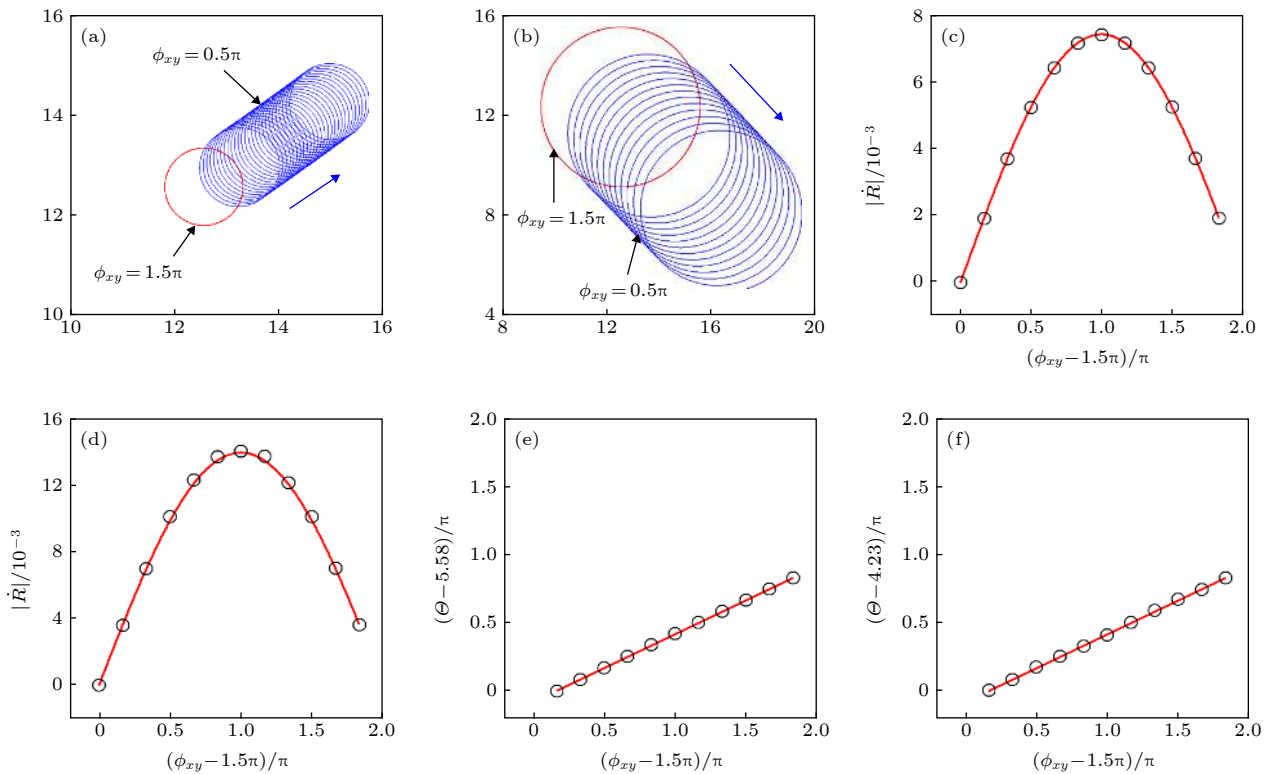


图4 极化电场作用下顺时针旋转螺旋波的漂移^[15] (a),(b) 顺时针 ($\phi_{xy} = 0.5\pi$)、逆时针 ($\phi_{xy} = 1.5\pi$) 旋转的圆极化电场作用下螺旋波的漂移, 其中, $\omega_e = 2\omega$, $\phi_e = 0$, $\Phi = 0$, ω 是螺旋波的频率; (c),(d) 漂移速率与相位差 ϕ_{xy} 的关系, 实线为理论结果, 圆圈为数值结果; (e),(f) 漂移角与相位差 ϕ_{xy} 的关系, 实线为理论结果, 圆圈为数值结果. 当漂移速率为 0 ($\phi_{xy} = 1.5\pi$), 无漂移角; (a),(c),(e) 强激发介质; (b),(d),(f) 弱激发介质

Fig. 4. Drifting behaviors of cw spirals under the influence of a polarized electric field^[15]: (a),(b) Drifting behaviors of spirals under the influence of a cw ($\phi_{xy} = 0.5\pi$) and a counterclockwise (ccw) ($\phi_{xy} = 1.5\pi$) circularly polarized electric fields (CPEFs) with $\omega_e = 2\omega$, $\phi_e = 0$, $\Phi = 0$, and ω being the frequency of the spiral waves; (c),(d) dependence of theoretical (lines) and numerical (circles) drift speeds on the phase difference ϕ_{xy} ; (e),(f) dependence of theoretical (lines) and numerical (circles) drift angles on the phase difference ϕ_{xy} . When the drift speed is 0 ($\phi_{xy} = 1.5\pi$), the drift angle cannot be defined. (a),(c),(e) Highly excitable medium; (b),(d),(f) Weakly excitable medium.

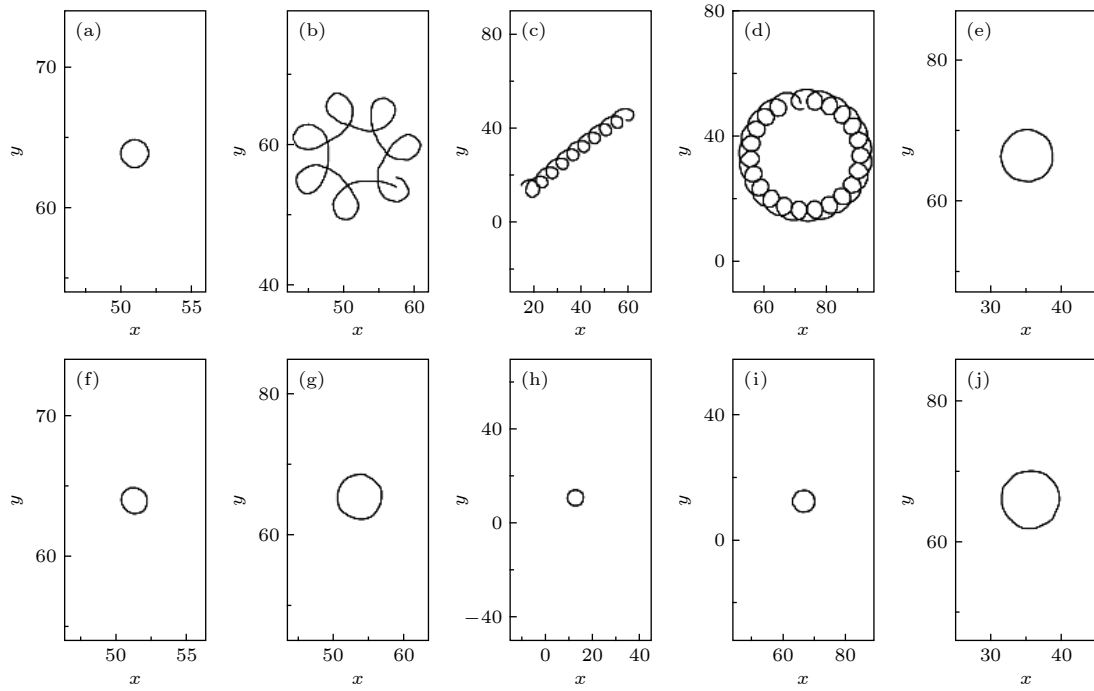


图 5 螺旋波端点轨迹^[20] (a)—(e) 无圆极化电场作用; (f)—(j) 有圆极化电场作用, 同一列上下两个图的介质尺寸相同
 Fig. 5. Trajectories of spiral tips without control (a)–(e) and under control (f)–(j) of CPEF^[20]. The size in the same column is identical.

近来, Li 等^[21] 基于响应函数理论, 对 Chen 等^[20] 的发现进行了理论研究, 解析给出了圆极化电场作用下, 螺旋波的稳定区域及同步条件, 从理论上得到了圆极化电场的 Arnold tongue 区域 (图 6), 理论和数值结果符合较好.

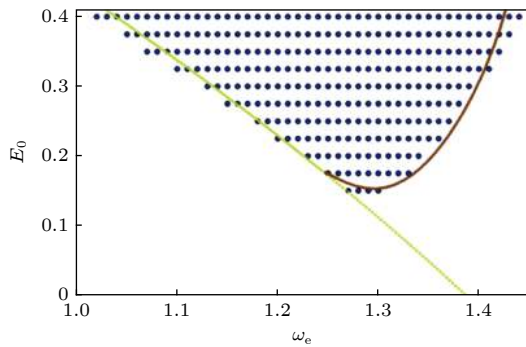


图 6 圆极化电场强度与频率之间的 Arnold tongue 区域^[21]
 实线为理论结果, 圆圈为数值结果

Fig. 6. Arnold tongue of the (ω_e, E_0) -plane of CPEF^[21]: Lines and circles denote the theoretical and the numerical results, respectively.

4 手征对称性破缺

螺旋波湍流态下, 系统中具有顺时针、逆时针旋转方向的螺旋波数目一般情况下一样多, 系统总体上呈现出手征对称性, 即零旋转手征对称性. 早年,

Nicolis 和 Prigogine^[22] 预言, 若反应扩散系统受手征性外场作用, 可能产生手征对称性破缺. 近来, Li 等^[23] 研究了圆极化电场对反应扩散系统中螺旋波湍流态手征性的影响, 发现外加的圆极化电场, 可使原先具有零旋转手征对称性的系统出现手征对称性破缺, 且使系统具有与圆极化电场的旋转手征性一致的手征性 (图 7). 研究利用漫游螺旋波, 进一步发现其原因在于, 外加的圆极化电场, 可使系统中与其同向旋转的螺旋波同步, 进而产生频移, 从而压制住具有另外旋转方向的螺旋波, 使得系统出现了手征对称性破缺 (图 8). 对于严格旋转的螺旋波, Li 等^[24] 理论分析和数值模拟发现, 与圆极化电场同向、反向旋转的螺旋波, 均可通过调节电场频率对其手征性进行控制.

5 多臂螺旋波的稳定

两个或多个相同手征性的螺旋波可以形成一个多臂螺旋波^[25], 其稳定性与它的周期密切相关^[26]. 比如, 当双臂螺旋波的周期小于一定临界值时, 两个螺旋波之间的排斥作用会增强, 使其衰退变成两个单臂螺旋波. Deng 等^[27] 研究发现, 与双臂螺旋波反向旋转的圆极化电场, 可使双臂螺旋波的周期

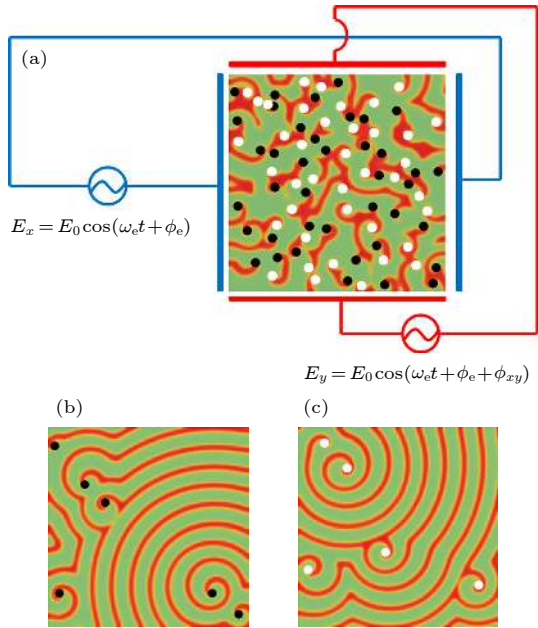


图 7 螺旋波湍流态中的手征对称性破缺^[23] (a) 螺旋波湍流态中逆时针(黑点)、顺时针(白点)旋转的螺旋波; (b) 在逆时针旋转的圆极化电场作用下, 系统中仅存留逆时针旋转的螺旋波; (c) 当(b)中的圆极化电场变成顺时针旋转后, 系统中仅存留顺时针旋转的螺旋波. $\omega_e = 1.25$, $E_0 = 0.20$

Fig. 7. Coherent state out of defect-mediated turbulence accompanied by chiral symmetry breaking^[23]: (a) An initial defect-mediated turbulence state consists of ccw spiral defects (black dots) and cw ones (white dots); (b) coherent state with only ccw spiral waves exists in the asymptotic state when the system is coupled to a ccw CPEF with $\omega_e = 1.25$ and $E_0 = 0.20$; (c) similar to (b) but with a cw CPEF, and in such a case, only cw spiral waves survive in the system.

增大, 进而减弱两个螺旋波之间的相互作用, 从而可使其处于稳定状态(图 9). 进一步研究表明, 外加的圆极化电场, 需要具有合适的强度和频率, 才可使双臂螺旋波稳定(图 10 给出了详细的相图).

6 次激发介质中的螺旋波

若介质的激发性只能维持平面波的传播, 但不能让螺旋波旋转起来, 则介质属于次激发介质^[28]. Jung 和 Mayer-Kress^[29,30] 提出, 噪声作用可助次激发介质产生螺旋波. Cai 等^[11] 发现, 圆极化电场可使次激发介质中原先不断收缩的半平面波不再收缩, 而是发生卷曲形成螺旋波(图 11). 所形成的螺旋波可处于严格旋转状态, 且与圆极化电场同步. 基于运动学关系^[31]

$$c_t = c_0 + c_B + c_E, \quad (3)$$

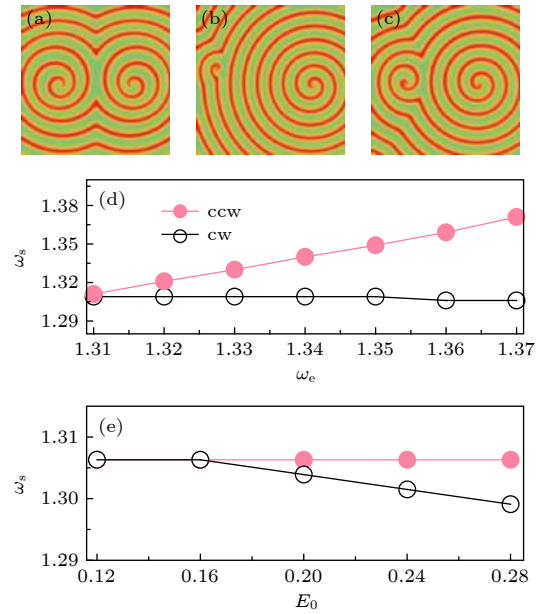


图 8 逆时针旋转圆极化电场作用下, 漫游螺旋波对的手征对称性破缺^[23] (a) $E_0 = 0$; (b) $E_0 = 0.10$, $\omega_e = 1.350 > \omega_0^{\text{ms}}$ ($\omega_0^{\text{ms}} \approx 1.309$), 其中 ω_0^{ms} 是无圆极化电场作用下, 漫游螺旋波的主频率; (c) $E_0 = 0.24$, $\omega_e = 1.307 < \omega_0^{\text{ms}}$; (d) 圆极化电场作用下, 逆时针旋转螺旋波频率 ω_s^{ccw} (实心圆圈) 及顺时针旋转螺旋波频率 ω_s^{cw} (空心圆圈) 与圆极化电场频率 ω_e 的关系, 其中 $E_0 = 0.10$; (e) 圆极化电场作用下, 逆时针旋转螺旋波频率 ω_s^{ccw} (实心圆圈) 及顺时针旋转螺旋波频率 ω_s^{cw} (空心圆圈) 与圆极化电场强度 E_0 的关系, 其中 $\omega_e = 1.307$

Fig. 8. Symmetry breaking of a meandering spiral pair under a ccw CPEF^[23]: (a) $E_0 = 0$; (b) $E_0 = 0.10$, $\omega_e = 1.350 > \omega_0^{\text{ms}}$ ($\omega_0^{\text{ms}} \approx 1.309$), where ω_0^{ms} is the principal frequency of the meandering spiral without the CPEF; (c) $E_0 = 0.24$, $\omega_e = 1.307 < \omega_0^{\text{ms}}$; (d) dependence of ω_s^{ccw} (the frequency of the ccw spiral wave) (full circles) and ω_s^{cw} (the frequency of the cw spiral wave) (open circles) on ω_e with $E_0 = 0.10$; (e) dependence of ω_s^{ccw} (full circles) and ω_s^{cw} (open circles) on E_0 with $\omega_e = 1.307$.

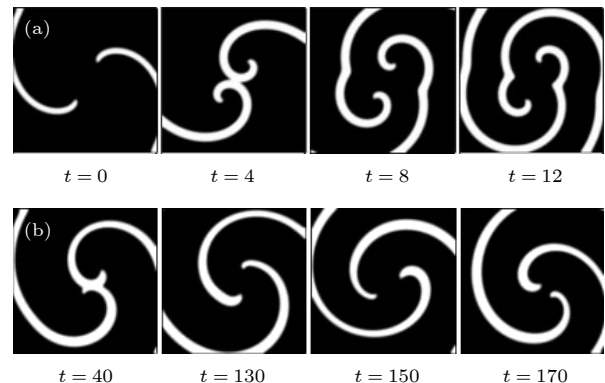


图 9 圆极化电场对双臂螺旋波的稳定作用^[27] (a) 无外场作用; (b) 有圆极化电场作用, $E_0 = 1.0$, $\omega_e = 1.22$

Fig. 9. Stabilization of two-armed spiral by CPEF^[27]: (a) Without external fields; (b) in the presence of a CPEF with $E_0 = 1.0$, $\omega_e = 1.22$.

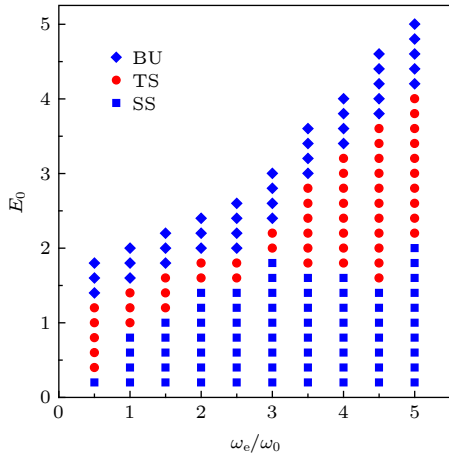


图 10 圆极化电场对双臂螺旋波作用的相图^[27] BU, TS 分别表示破碎、稳定区域; SS 表示电场过弱, 不足以稳定双臂螺旋波而使其衰减为单臂螺旋波的区域; $\omega_0 = 1.24$ 是单臂螺旋波的频率

Fig. 10. The phase diagram for the effects of CPEF on two-armed spiral^[27]: BU, TS denote the breakup and the stabilization regions, respectively, and SS means the region where the electric field is not strong enough to stabilize the two-armed spiral and it decays into two single-armed spirals. The frequency of the single-armed spiral $\omega_0 = 1.24$.

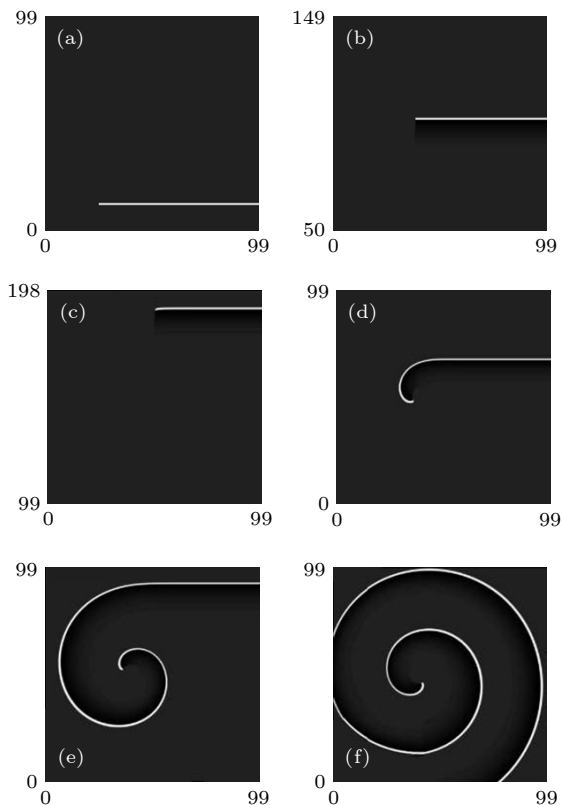


图 11 次激发介质中半平面波的演化^[11] (a)–(c) 无外场作用; (d)–(f) 有圆极化电场作用. $E_0 = 0.2$, $\omega_e = 0.2$
Fig. 11. The evolution of a broken plane wave in the subexcitable system without (a)–(c) and with (d)–(f) CPEFs^[11]. $E_0 = 0.2$, $\omega_e = 0.2$.

研究给出了圆极化电场使次激发介质产生螺旋波的定量解释. 方程 (3) 中, c_t 是半平面波端点的切向速度, c_0 表示平面波的速度. $c_B = c_0(B - B_c)/K$, 其中, $B = 4\sqrt{3}\epsilon/\Delta^3$ 表征介质的激发性, $\Delta = \delta^3 - 3\delta$ 由系统参数 δ 确定; $B_c = 0.535$ 是区分次激发介质 ($B > B_c$) 和激发介质 ($B < B_c$) 的临界值; $K \approx 0.63$ 是常数. $c_E = \gamma_{//}E_{//} + \gamma_{\perp}E_{\perp}$, 其中 $E_{//} = E_0 \cos \theta$, $E_{\perp} = E_0 \sin \theta$ 分别是与 c_t 平行和垂直的电场分量; $\gamma_{//} \approx -0.850$, $\gamma_{\perp} \approx 0.929$ 是常数; θ 是 c_t 与电场 E 之间的夹角 (图 12(a)). Hakim 和 Karma^[31] 提出, 方程 (3) 中, 若 $c_t < c_0$, 则介质可产生螺旋波. Cai 等^[11] 通过半解析分析, 得出与图 11 中的螺旋波对应的方程 (3) 中, $c_B + c_E < 0$ (图 12(b)), 即 $c_t < c_0$, 从而使得次激发介质中形成了螺旋波, 且数值与半解析结果取得了一致的结果 (图 12(c)). 值得提出的是, 图 11 中的系统参数 $B = 0.5515$, $c_0 \approx 0.8887$, 从而可有 $c_B > 0$. 因此, $c_B + c_E < 0$ 是圆极化电场作用的结果.

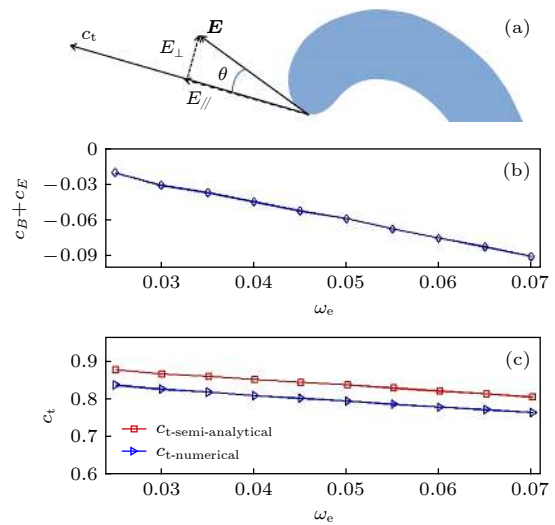


图 12 圆极化电场使次激发介质产生螺旋波的机制分析^[11], $E_0 = 0.1$ (a) 圆极化电场对螺旋波端点的作用示意图; (b) $c_B + c_E$ 随 ω_e 的变化关系; (c) c_t 的半解析解与数值解的比较

Fig. 12. The mechanism analyses for spiral waves sustained by CPEF in subexcitable media^[11], $E_0 = 0.1$: (a) The sketch of a spiral wave tip submitted to a CPEF; (b) results of $c_B + c_E$ varying with ω_e ; (c) the comparison of the semi-analytical c_t with the numerical c_t .

7 三维回卷波湍流态的控制

回卷波会因其奇异线张力为负而失稳成湍流

态^[32], 相反, 若奇异线张力为正, 则可使回卷波处于稳定状态. 在系统全空间加周期信号^[33], 或在系统的某一局部加一个较强的周期信号^[34], 这种张力为负引起的回卷波湍流态可以被控制到空间均匀状态. 最近, Li 等^[35]发现, 圆极化电场可抑制因张力为负引起的回卷波湍流态, 而使其从无序湍流态走向有序回卷波态(图 13). 其原因在于圆极化电场可使回卷波锁相, 进而将奇异线张力由负变为正, 以及具有较原先高些的频率. 研究进一步发现, 在圆极化电场频率与回卷波频率较为接近的范围

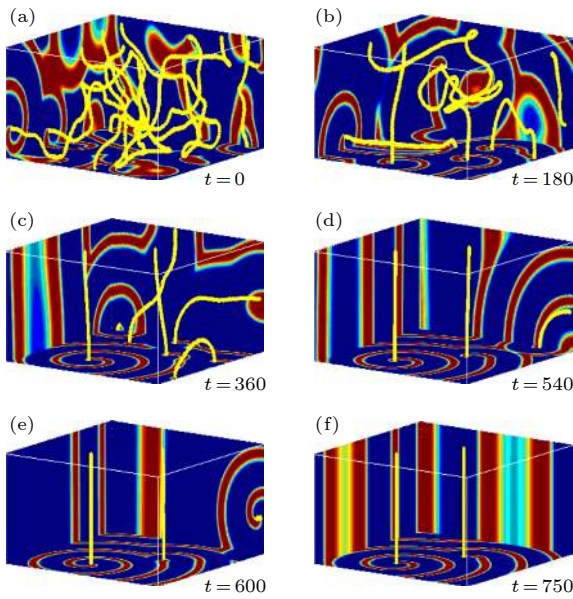


图 13 圆极化电场作用下, 回卷波湍流态从无序走向有序^[35] $E_0 = 0.4$, $\omega_e = \omega_0 = 1.2455$, 其中 ω_0 表示螺旋波的频率; $t = 0$, 施加圆极化电场作用; 黄线表示奇异线

Fig. 13. Ordering of scroll wave turbulence by switching on a ccw CPEF at $t = 0$ with $E_0 = 0.4$ and rotation frequency $\omega_e = 1.2455$ equal to the natural spiral wave frequency ω_0 ^[35]. Filaments are shown in yellow.

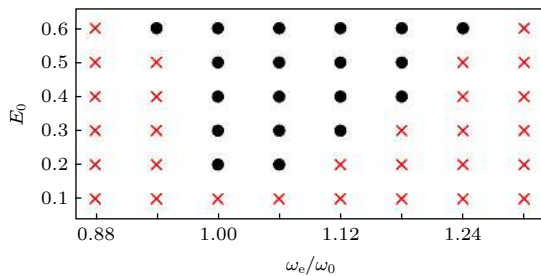


图 14 圆极化电场抑制(实心圆圈)回卷波湍流态的参数区域中, 电场强度 E_0 与归一化频率 ω_e/ω_0 的关系^[35], 交叉表示不能抑制回卷波湍流态

Fig. 14. Parameter region of scroll wave turbulence suppression (full circles) as a function of external field amplitude E_0 and normalized frequency ω_e/ω_0 ^[35]. Crosses denote failure of ordering turbulence.

内, 如此的回卷波湍流被抑制现象存在一定的范围(图 14). 另外, 基于响应函数理论, 研究发现圆极化电场下, 奇异线张力满足以下关系^[35]:

$$\Gamma_1 = \gamma_1 + g_{//} \frac{\Delta\omega}{M} + g_{\perp} \sqrt{E_0^2 - \left(\frac{\Delta\omega}{M}\right)^2}, \quad (4)$$

其中 Γ_1 , γ_1 分别表示回卷波在有、无外电场作用下的奇异线张力, E_0 为电场强度, $\Delta\omega$ 表征电场与回卷波之间的频差. M , $g_{//}$, g_{\perp} 三个参数可由理论计算得到. 数值计算与理论值取得一致的结果(图 15).

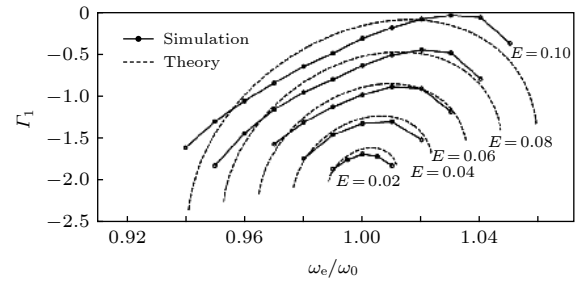


图 15 锁相回卷波的奇异线张力^[35]

Fig. 15. Filament tension of phase-locked scroll waves^[35].

8 心脏组织中螺旋波的去钉扎

心脏系统中, 螺旋波会引起心动过速, 若螺旋波被缺陷钉扎住而无法驱赶走, 则可能导致更严重的心脏疾病. 为了使螺旋波去钉扎, 人们研究发现, 匀强电场在缺陷边界处可产生激发波^[36–39], 进而可使螺旋波去钉扎. Feng 等^[40]发现, 相比于匀强电场, 圆极化电场作用下缺陷边界的膜电势分布呈动态的“太极图”状(匀强电场下, 呈静态的对称偶极子状), 且与圆极化电场同步旋转(图 16). 研究发现, 圆极化电场能有效地使螺旋波去钉扎(图 17). 由于圆极化电场作用产生的“太极图”式的膜电势分布, 即去极化及超极化区域可与电场动态同步旋转, 使得螺旋波的去钉扎较匀强电场更有效、所需的外加电压更低, 以及有更大的适用范围(图 18).

9 心脏组织中螺旋波湍流态的控制

心脏组织中的螺旋波湍流态, 会引起严重的心脏病, 比如心颤^[41]. 人们通常采用单个高压电击作用于心脏而使其恢复正常心电活动^[42–44], 但因其副作用太大^[43,44]而局限过大. 缺陷激发波法^[36–39]利用低电压匀强电场在缺陷边界产生激发波, 从而

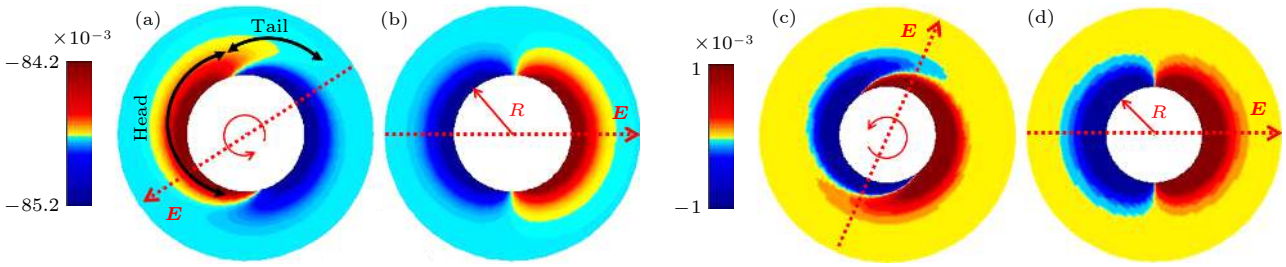


图 16 圆极化电场 (a)(c) 和匀强电场 (b)(d) 作用下的膜电势分布^[40] (a) Luo-Rudy 模型, $E_0 = 0.05 \text{ V/cm}$, $\omega_e = 0.2 \text{ rad/ms}$; (b) Luo-Rudy 模型, $E_0 = 0.05 \text{ V/cm}$; (c) Barkley 模型, $E_0 = 0.05$, $\omega_e = 4$; (d) Barkley 模型, $E_0 = 0.05$. R 表示缺陷半径, Luo-Rudy 模型中, $R = 0.32 \text{ cm}$; Barkley 模型中, $R = 3$. 红色点箭头表示电场方向, 红色曲线箭头表示圆极化电场逆时针旋转. 围绕缺陷的红色、蓝色分别表示去极化、超极化区域

Fig. 16. Distribution of the membrane potential induced by CPEF and uniform electric field (UEF)^[40]: (a) CPEF in Luo-Rudy model, $E_0 = 0.05 \text{ V/cm}$, $\omega_e = 0.2 \text{ rad/ms}$; (b) UEF in Luo-Rudy model, $E_0 = 0.05 \text{ V/cm}$; (c) CPEF in Barkley model, $E_0 = 0.05$, $\omega_e = 4$; (d) UEF in Barkley model, $E_0 = 0.05$. In Luo-Rudy model, the obstacle size $R = 0.32 \text{ cm}$, and in Barkley model, $R = 3$. The red dotted arrows represent the directions of electric fields. The red curved arrows mean CPEFs rotate counterclockwise. The red and the blue regions around obstacles demonstrate de-polarizations and hyper-polarizations, respectively.

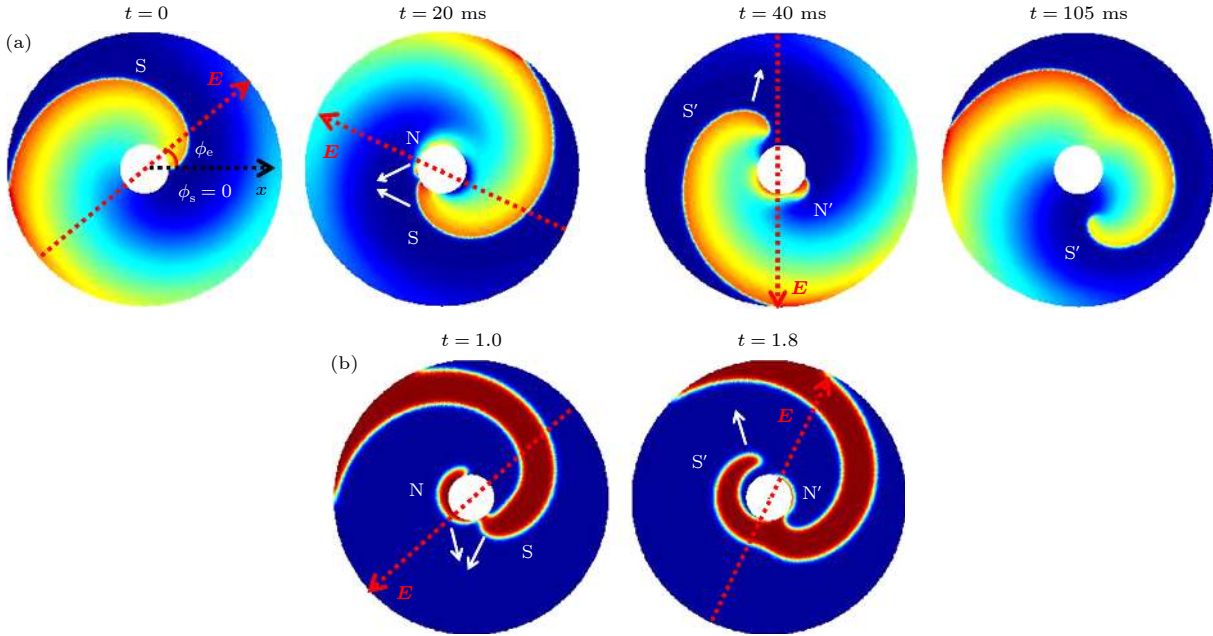


图 17 圆极化电场去除顺时针旋转的钉扎螺旋波^[40] (a) Luo-Rudy 模型, 螺旋波频率 $\omega_s = 0.136 \text{ rad/ms}$, 圆极化电场 $E_0 = 0.7 \text{ V/cm}$, $\omega_e = 0.1 \text{ rad/ms}$; 电场作用时间 $t = 0 - 46.2 \text{ ms}$, 圆极化电场相对于 x 轴的初相位为 ϕ_e ; 螺旋波相对于 x 轴的初相位 $\phi_s = 0$; (b) Barkley 模型, 螺旋波频率 $\omega_s = 1.024$, 圆极化电场 $E_0 = 1.8$, $\omega_e = 3.686$; 电场作用时间 $t = 0 - 6$. N 和 N' 表示不同时刻圆极化电场激发产生的激发波, S 和 S' 分别表示初始的钉扎螺旋波、去钉扎后的螺旋波. 白色箭头表示波的传播方向

Fig. 17. Unpinning the cw rotating anchored spiral by CPEF^[40]: (a) Luo-Rudy model, the frequency of spiral $\omega_s = 0.136 \text{ rad/ms}$; $E_0 = 0.7 \text{ V/cm}$, $\omega_e = 0.1 \text{ rad/ms}$; CPEF is applied from $t = 0$ to $t = 46.2 \text{ ms}$. ϕ_e is the initial phase of CPEF relative to x axis; ϕ_s is the initial phase of the anchored spiral front relative to x axis and sets as zero; (b) Barkley model, $\omega_s = 1.024$, $E_0 = 1.8$, $\omega_e = 3.686$; CPEF is applied from $t = 0$ to $t = 6$. N and N' represent different new waves nucleated by CPEF in different time. S and S' represent the initial anchored spiral and the new free spiral, respectively. White arrows are the propagation directions of waves.

可以去颤. 但该方法只能在缺陷邻近处有效而使其效率受限 (需要较多缺陷, 才能产生出较多的激发波去颤). Feng 等^[45] 发现, 圆极化电场在低电压下, 可使缺陷组织激发形成不断向外传播的圆形波, 且其频率可高于湍流波的频率 (图 19). 这些具有较

高频率的圆形波可抑制住心脏组织中的湍流态, 从而可以去颤 (图 20). 相对而言, 匀强电场虽也可使缺陷组织激发出圆形波, 但其频率几乎都是低于湍流波的频率 (图 21), 从而很难实现对湍流态的有效控制.

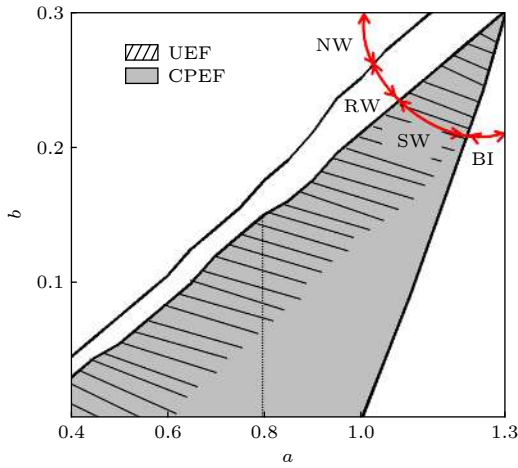


图 18 Barkley 模型中, 圆极化电场 (灰色区域) 和匀强电场 (阴影区域) 去除钉扎螺旋波的适用范围^[40] SW, NW, RW 和 BI 分别表示螺旋波、无波、收缩波和双稳区域; 圆极化电场 $E_0 = 1.8$, $\omega_e = 3.6 \omega_s$; 匀强电场 $E_0 = 7$

Fig. 18. Unpinning scope of CPEF (gray) and UEF (shaded) in Barkley model^[40]; SW, NW, RW and BI regions represent spiral waves, no wave, retracting waves and bi-stability, respectively; for CPEF, $E_0 = 1.8$, $\omega_e = 3.6 \omega_s$ and for UEF, $E_0 = 7$.

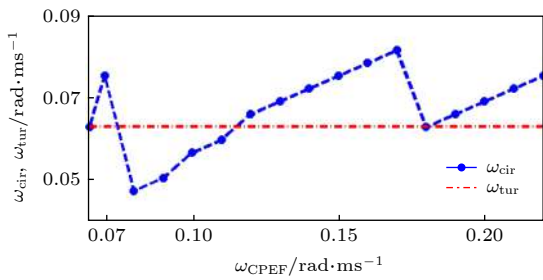


图 19 静息态下圆极化电场及其激发产生的圆形波列之间的频率关系^[45] $E_0 = 1.0 \text{ V/cm}$, $0.065 \text{ rad/ms} \leq \omega_{\text{CPEF}} \leq 0.22 \text{ rad/ms}$; ω_{cir} 是圆形波列的频率; ω_{tur} 是螺旋波湍流态的主频率

Fig. 19. The frequency relations between the circular wave train and CPEF in a two-dimensional quiescent medium^[45]: $E_0 = 1.0 \text{ V/cm}$, $0.065 \text{ rad/ms} \leq \omega_{\text{CPEF}} \leq 0.22 \text{ rad/ms}$; ω_{cir} is the frequency of the circular wave trains, and ω_{tur} is the dominant frequency of the spiral turbulence.

10 结 论

本文着重介绍了圆极化电场对可激发介质中螺旋波的控制. 由于具有旋转对称性, 圆极化电场可对螺旋波动力学产生独特而有效的影响, 并可基于运动学、响应函数等理论, 给出相应的理论解释.

近来, 人们通过对两对相互垂直的电极板通以交流电, 在 BZ 反应实验中实现了外加极化电场^[10]. 心脏实验方面, 人们一般把两对正负电极分别放置

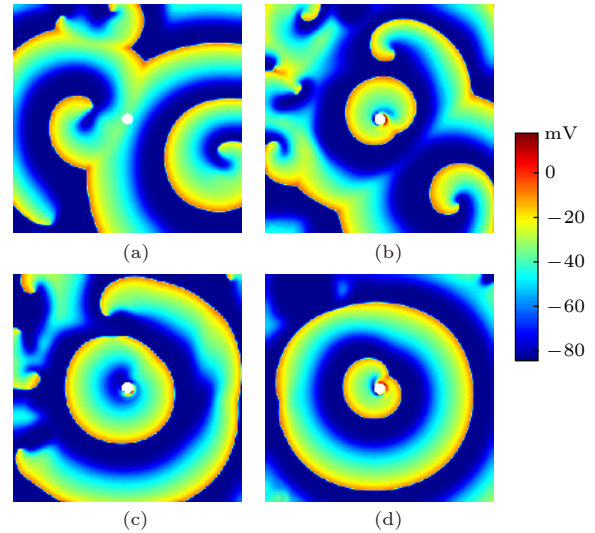


图 20 圆极化电场抑制螺旋波湍流态^[45] $E_0 = 1.0 \text{ V/cm}$, $\omega_{\text{CPEF}} = 0.14 \text{ rad/ms}$ (a) $t = 0$; (b) $t = 1000 \text{ ms}$; (c) $t = 1800 \text{ ms}$; (d) $t = 2800 \text{ ms}$

Fig. 20. Suppression of spiral turbulence by CPEF^[45]: $E_0 = 1.0 \text{ V/cm}$, $\omega_{\text{CPEF}} = 0.14 \text{ rad/ms}$: (a) $t = 0$; (b) $t = 1000 \text{ ms}$; (c) $t = 1800 \text{ ms}$; (d) $t = 2800 \text{ ms}$.

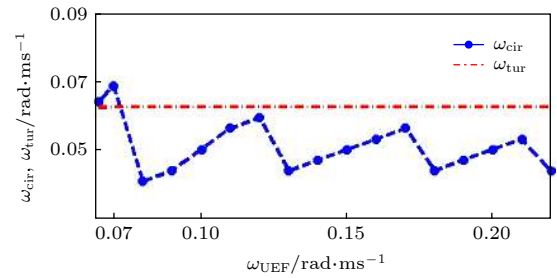


图 21 静息态下匀强电场及其激发产生的圆形波列之间的频率关系^[45] $E_0 = 1.0 \text{ V/cm}$, $0.065 \text{ rad/ms} \leq \omega_{\text{UEF}} \leq 0.22 \text{ rad/ms}$; 脉冲间隔为 10 ms ; ω_{cir} 是圆形波列的频率; ω_{tur} 是螺旋波湍流态的主频率

Fig. 21. The frequency relations between the circular wave train and UEF in a two-dimensional quiescent medium^[45]: $E_0 = 1.0 \text{ V/cm}$, $0.065 \text{ rad/ms} \leq \omega_{\text{UEF}} \leq 0.22 \text{ rad/ms}$; the pulse duration is 10 ms ; ω_{cir} is the frequency of the circular wave trains, and ω_{tur} is the dominant frequency of the spiral turbulence.

于心脏的四周, 然后通以直流电来对心脏产生作用 (文献 [38] 的图 5(D)). 很显然, 只要把其中的直流电换成交流电, 就可以类似于 BZ 反应实验一样, 在心脏实验中实现外加极化电场. 因此, 极化电场在心脏病研究方面, 也具有重要的意义. 希望本文介绍的数值及理论结果可以在实验上实现.

参考文献

[1] Winfree A T 2001 *The Geometry of Biological Time* (2nd ed.) (New York: Springer-Verlag) pp258–302

- [2] Winfree A T 1972 *Science* **175** 634
- [3] Jakubith S, Rotermund H H, Engel W, von Oertzen A, Ertl G 1990 *Phys. Rev. Lett.* **65** 3013
- [4] Gray R A, Pertsov A M, Jalife J 1998 *Nature* **392** 75
- [5] Christoph J, Chebbok M, Richter C, Schröder-Schetelig J, Bittihn P, Stein S, Uzelac I, Fenton F H, Hasenfuß G, Gilmour Jr R F, Luther S 2018 *Nature* **555** 667
- [6] Steinbock O, Schütze J, Müller S C 1992 *Phys. Rev. Lett.* **68** 248
- [7] Muñozuri A P, Gómez-Gesteira M, Pérez-Muñuzuri V, Krinsky V I, Pérez-Villar V 1994 *Phys. Rev. E* **50** 4258
- [8] Zhang H, Hu B, Hu G, Xiao J 2003 *J. Chem. Phys.* **119** 4468
- [9] Chen J X, Zhang H, Li Y Q 2006 *J. Chem. Phys.* **124** 014505
- [10] Ji L, Zhou Y, Li Q, Qiao C, Ouyang Q 2013 *Phys. Rev. E* **88** 042919
- [11] Cai M C, Pan J T, Zhang H 2012 *Phys. Rev. E* **86** 016208
- [12] Winfree A T 1991 *Chaos* **1** 303
- [13] Krinsky V, Hamm E, Voignier V 1996 *Phys. Rev. Lett.* **76** 3854
- [14] Pumir A, Nikolski V, Hörning M, Isomura A, Agladze K, Yoshikawa K, Gilmour R, Bodenschatz E, Krinsky V 2007 *Phys. Rev. Lett.* **99** 208101
- [15] Li T C, Gao X, Zheng F F, Pan D B, Zheng B, Zhang H 2017 *Sci. Rep.* **7** 8657
- [16] Biktashev V N, Holden A V 1995 *Chaos, Solitons Fractals* **5** 575
- [17] Biktasheva I V, Barkley D, Biktashev V N, Bordyugov G V, Foulkes A J 2009 *Phys. Rev. E* **79** 056702
- [18] Biktasheva I V, Barkley D, Biktashev V N, Foulkes A J 2010 *Phys. Rev. E* **81** 066202
- [19] Pikovsky A, Rosenblum M, Kurths J 2001 *Synchronization: A Universal Concept in Nonlinear Sciences* (Cambridge: Cambridge University Press) pp65–66
- [20] Chen J X, Zhang H, Li Y Q 2009 *J. Chem. Phys.* **130** 124510
- [21] Li T C, Li B W, Zheng B, Zhang H, Panfilov A, Dierckx H 2019 *New J. Phys.* **21** 043012
- [22] Nicolis G, Prigogine I 1981 *Proc. Natl. Acad. Sci. U.S.A.* **78** 659
- [23] Li B W, Deng L Y, Zhang H 2013 *Phys. Rev. E* **87** 042905
- [24] Li B W, Cai M C, Zhang H, Panfilov A V, Dierckx H 2014 *J. Chem. Phys.* **140** 184901
- [25] Agladze K I, Krinsky V I 1982 *Nature* **296** 424
- [26] Vasiev B, Siegert F, Weijer C 1997 *Phys. Rev. Lett.* **78** 2489
- [27] Deng L Y, Zhang H, Li Y Q 2009 *Phys. Rev. E* **79** 036107
- [28] Mikhailov A S, Zykov V S 1991 *Physica D* **52** 379
- [29] Jung P, Mayer-Kress G 1995 *Phys. Rev. Lett.* **74** 2130
- [30] Jung P, Mayer-Kress G 1995 *Chaos* **5** 458
- [31] Hakim V, Karma A 1999 *Phys. Rev. E* **60** 5073
- [32] Biktashev V N, Holden A V, Zhang H 1994 *Philos. Trans. R. Soc. London, Ser. A* **347** 611
- [33] Alonso S, Sagués F, Mikhailov A S 2003 *Science* **299** 1722
- [34] Zhang H, Cao Z J, Wu N J, Ying H P, Hu G 2005 *Phys. Rev. Lett.* **94** 188301
- [35] Li T C, Gao X, Zheng F F, Cai M C, Li B W, Zhang H, Dierckx H 2016 *Phys. Rev. E* **93** 012216
- [36] Ripplinger C M, Krinsky V I, Nikolski V P, Efimov I R 2006 *Am. J. Physiol. Heart Circ. Physiol.* **291** H184
- [37] Cysyk J, Tung L 2008 *Biophys. J.* **94** 1533
- [38] Fenton F H, Luther S, Cherry E M, Otani N F, Krinsky V, Pumir A, Bodenschatz E, Gilmour Jr R F 2009 *Circulation* **120** 467
- [39] Luther S, Fenton F H, Kornreich B G, Squires A, Bittihn P, Hornung D, Zabel M, Flanders J, Gladuli A, Campoy L, Cherry E M, Luther G, Hasenfuss G, Krinsky V I, Pumir A, Gilmour Jr R F, Bodenschatz E 2011 *Nature* **475** 235
- [40] Feng X, Gao X, Pan D B, Li B W, Zhang H 2014 *Sci. Rep.* **4** 4831
- [41] Jalife J 2000 *Annu. Rev. Physiol.* **62** 25
- [42] Koster R W, Dorian P, Chapman F W, Schmitt P W, O'Grady S G, Walker R G 2004 *Am. Heart J.* **147** e1
- [43] Babbs C F, Tacker W A, VanVleet J F, Bourland J D, Geddes L A 1980 *Am. Heart J.* **99** 734
- [44] Santini M, Pandozi C, Altamura G, Gentilucci G, Villani M, Scianaro M C, Castro A, Ammirati F, Magris B 1999 *J. Interv. Card. Electrophysiol.* **3** 45
- [45] Feng X, Gao X, Tang J M, Pan J T, Zhang H 2015 *Sci. Rep.* **5** 13349

SPECIAL TOPIC—Statistical physics and complex systems

Control of spiral waves in excitable media under polarized electric fields*

Pan Jun-Ting¹⁾ He Yin-Jie²⁾ Xia Yuan-Xun²⁾ Zhang Hong^{2)†}

1) (*Ocean College, Zhejiang University, Zhoushan 316021, China*)

2) (*Zhejiang Institute of Modern Physics, Department of Physics, Zhejiang University, Hangzhou 310027, China*)

(Received 20 December 2019; revised manuscript received 17 January 2020)

Abstract

Spiral waves are ubiquitous in diverse physical, chemical, and biological systems. Periodic external fields, such as polarized electric fields, especially circularly polarized electric fields which possess rotation symmetry may have significant effects on spiral wave dynamics. In this paper, control of spiral waves in excitable media under polarized electric fields is reviewed, including resonant drift, synchronization, chiral symmetry breaking, stabilization of multiarmed spiral waves, spiral waves in subexcitable media, control of scroll wave turbulence, unpinning of spiral waves in cardiac tissues, control of spiral wave turbulence in cardiac tissues, etc.

Keywords: polarized electric fields, excitable media, spiral waves, control

PACS: 05.45.-a, 05.65.+b, 05.45.Xt

DOI: [10.7498/aps.69.20191934](https://doi.org/10.7498/aps.69.20191934)

* Project supported by the National Natural Science Foundation of China (Grant No. 11675141).

† Corresponding author. E-mail: hongzhang@zju.edu.cn

卡西米尔力

苗兵

Casimir force

Miao Bing

引用信息 Citation: *Acta Physica Sinica*, 69, 080505 (2020) DOI: 10.7498/aps.69.20200450

在线阅读 View online: <https://doi.org/10.7498/aps.69.20200450>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

运用四元数分析椭球微粒所受的光阱力

Using quaternions to analyze the trapping force of an ellipsoidal bead

物理学报. 2017, 66(4): 048701 <https://doi.org/10.7498/aps.66.048701>

非球形效应对强声场中次Bjerknes力的影响

Influence of nonspherical effects on the secondary Bjerknes force in a strong acoustic field

物理学报. 2017, 66(1): 014302 <https://doi.org/10.7498/aps.66.014302>

基于多引导滤波的图像增强算法

Image enhancement based on multi-guided filtering

物理学报. 2018, 67(23): 238701 <https://doi.org/10.7498/aps.67.20181425>

雷暴云内电场力对起电和电荷结构的反馈作用

Feedback effect of electric field force on electrification and charge structure in thunderstorm

物理学报. 2018, 67(16): 169201 <https://doi.org/10.7498/aps.67.20180505>

融合散斑干涉技术的阵列式洛伦兹力微颗粒探测方法

Lorentz force particle analyzer with an array probe based on speckle pattern interferometry

物理学报. 2017, 66(20): 208102 <https://doi.org/10.7498/aps.66.208102>

专题：统计物理和复杂系统

卡西米尔力*

苗兵†

(中国科学院大学材料科学与光电技术学院, 北京 100049)

(2020年3月26日收到; 2020年4月9日收到修改稿)

量子电动力学中的卡西米尔力是真空零点能的体现. 广义的卡西米尔力则依赖于涨落介质的类型广泛地出现于物理中, 包括量子, 临界, 戈德斯通模, 以及非平衡卡西米尔力. 长程关联的涨落介质和约束是产生卡西米尔力的两个条件. 本文通过回顾卡西米尔物理的发展, 讨论了不同类型的卡西米尔力, 几种正规化方法, 并对卡西米尔物理的进一步发展做了展望.

关键词: 卡西米尔力, 涨落, 长程关联, 正规化**PACS:** 05.40.-a, 05.70.Jk, 11.10.-z**DOI:** 10.7498/aps.69.20200450

1 引言

在一个具有长程关联的涨落介质中引入外加约束, 由于约束对于介质涨落模式的修改, 因而在介质之间产生一种有效的相互作用力, 即卡西米尔力 (Casimir force)^[1]. 在物理中, 存在多种机制可产生长程关联的涨落介质. 例如, 在量子场论中, 规范对称性要求零质量波色子, 其传递的相互作用是长程的, 这样, 零质量波色子场是满足卡西米尔力产生要求的长程关联涨落介质, 典型的例子是量子电动力学中的电磁场 (光子场). 在统计物理中, 调整参数至连续相变的临界点时, 序参量的关联长度发散, 序参量临界涨落的关联函数是长程衰减的幂函数, 这样, 序参量场提供了产生卡西米尔力的长程关联涨落介质; 另一方面, 即使不在临界点, 对于具有连续对称的统计模型, 由戈德斯通定理 (Goldstone theorem) 可知, 破缺连续对称可以产生零质量的戈德斯通粒子, 因此戈德斯通场也提供了卡西米尔力产生的背景介质, 典型的例子如铁磁相变的 n -矢量模型、 H_2^+ 超流相变、液晶、薄膜或界

面、聚合物等. 在非平衡体系中, 动力学满足的流守恒可以在非平衡体系中产生长程关联, 从而可以产生卡西米尔力. 由此可见, 卡西米尔物理是一个涵盖广泛的研究题目^[2].

本文从卡西米尔力最初在量子电动力学中的提出, 到广义的卡西米尔力, 对卡西米尔效应进行考察, 最后对卡西米尔物理在未来的发展进行展望.

2 量子场论

2.1 发现

量子场论中的卡西米尔力最早是由荷兰物理学家 Hendrik Casimir 在 1948 年发现的^[3]. 由于量子力学的不确定性原理, 真空非空, 而是充满了量子涨落. 量子涨落导致非零的真空能量, 物理上称为真空零点能 (zero point energy of vacuum). 卡西米尔力是真空零点能存在的一个直接结果^[4].

Casimir 考虑在真空中引入两个平行的中性理想导电金属平板, 要求平板是无穷大和无穷薄的. 金属板的存在修改了真空中电磁场的涨落模式. Casimir 分别计算了自由真空的零点能和加板约束

* 国家自然科学基金 (批准号: 21774131, 21544007) 资助的课题.

† 通信作者. E-mail: bmiao@ucas.ac.cn

后真空的零点能,发现两者都有紫外发散,然而通过计算两者之差,Casimir得到了一个随着板间距离变化的有限能量,进一步计算该能量对板间距离的变化率,则得到两板之间的一个吸引力,即著名的卡西米尔力,公式如下:

$$P = -\frac{\pi^2 \hbar c}{240 a^4}, \quad (1)$$

(1)式是两板之间单位板面积的力,即卡西米尔压力.式中, a 是板间距离, \hbar 和 c 分别是普朗克常数和光速.由(1)式可以发现:(1)式中不出现金属板的细节性质,而是仅依赖于物理学基本常数 \hbar , c ,以及板间距离 a ,因此卡西米尔力具有普适性;对于 \hbar 和 c 的依赖说明卡西米尔力的产生是由于量子效应和相对论效应;卡西米尔力是随板间距离以幂率(-4)衰减的长程力;负号表明这里的卡西米尔力是吸引力.

由于卡西米尔力的普适性,可以基于简单的量纲分析(dimensional analysis)迅速得出(1)式的幂率关系.考虑两块平行的金属板(Casimir平板),由于卡西米尔力 P 归因于量子效应和相对论效应,因此 P 与 $\hbar c$ 成比例.由量纲 $[\hbar c] = E \cdot L$, $[P] = E/L^3$;本问题里唯一剩下的长度尺度是板间距离 a ;由量纲分析得出: $P \sim \hbar c/a^4$.可是,简单的量纲分析无法得出卡西米尔力普适常数的大小和符号,确定普适常数需要做具体的计算.在下文中我们将给出卡西米尔力基于真空零点能的计算.

2.2 与范德瓦耳斯力的关系

本节叙述卡西米尔力的研究背景,讲述其如何起源于对分子间范德瓦耳斯力(van der Waals force)在长程或者retarded情形下的计算^[5],以及卡西米尔力在后来的发展中远远超出了范德瓦耳斯力的范畴.

对胶体稳定性问题的研究启发了Casimir关于卡西米尔力的计算.胶体体系的不稳定性是由分子间相互作用力驱动的.理想气体的状态方程由于没有考虑分子间相互作用而不能描述气体的凝聚相变(气-液相变).荷兰物理学家van der Waals将理想气体方程修改为van der Waals方程,可以描述相变.在这个修正中,van der Waals引入了气体分子的尺寸(排除体积效应)和气体分子间相互作用力(范德瓦耳斯力).对于中性的气体分子,范德瓦耳斯力是一种短程的吸引力,它将驱动气体

状态的失稳而凝聚成液态.两个中性(可极化)分子间范德瓦耳斯力的量子力学计算是由London完成的,称为色散力(dispersion force),其原因是计算里需要考虑极化率(polarizability).

London的计算发现:两个中性分子间的色散力表现为一个以 r^{-6} 衰减的吸引相互作用势,其中 r 是分子间距.粗略来理解,指数为-6是由于:该力是分子的诱导偶极-诱导偶极相互作用,由于量子涨落,分子1产生诱导偶极,诱导偶极产生的诱导电场 $E \sim 1/r^3$,分子2在电场 E 作用下产生诱导偶极 $p \sim \alpha E \sim 1/r^3$,这里 α 是分子的极化率,相互作用能 $V \sim -p \cdot E \sim -1/r^6$.

Derjaguin-Landau-Verwey-Overbeek(DLVO)将London的色散力引入了他们的胶体稳定性理论.当胶体间由分子色散力带来的吸引大于由静电带来的排斥时,胶体体系将失去稳定性而发生凝聚.在胶体稳定性的研究中,Verwey和Overbeek发现,实验中分子之间的吸引力比 $1/r^6$ 衰减的更快,实际上应该是 $1/r^7$.Verwey和Overbeek因此评论说,这是因为在London理论里,没有考虑相对论效应,而胶体的间距(微米尺度)相对于微观的分子尺度很大,因此需要考虑相对论修正.

Casimir接受了这个建议,与Polder合作,重新研究了London理论,加入了相对论效应带来的所谓retardation效应,推导出了分子间的Casimir-Polder相互作用,是 r^{-7} 的色散吸引力.之后,Casimir认为Casimir-Polder分子间力的计算过于复杂,他希望简化计算.在与玻尔(Bohr)谈论时,玻尔评论这个力应该与零点能有关.受到启发后,Casimir完成了两宏观平板之间卡西米尔力的计算,通过计算真空零点能得到著名的卡西米尔力公式(1)式.

卡西米尔力被提出的时候,并没有引起很大的关注.这可能是因为当时人们已经知道了两个中性分子之间的色散吸引力,因此当中性分子变成两个宏观中性板时,其间会产生长程吸引力似乎并不值得过于惊讶.然而,我们需要指出,卡西米尔力为什么是吸引力,至今仍然没有完全令人满意的解释.在后来的发展中,人们发现取决于涨落介质的类型,边界条件,约束的性质和几何,卡西米尔力可以是排斥力.实际上,对于一个具体的体系,在未做计算前,人们不能判断卡西米尔力是吸引还是

排斥力. 因此, 在卡西米尔物理的发展中, 卡西米尔力虽起源于分子间范德瓦耳斯力的研究, 然而其发展已经远远超出了分子间力的研究范畴. 并且, 在讨论宏观或者介观物体之间的力时, 人们通常将卡西米尔力和范德瓦耳斯力进行区分: 范德瓦耳斯力是短距离, 近程的作用力, 依赖于作用物体的分子细节; 卡西米尔力是长程作用力, 长程特点将分子细节平均掉, 表现出普适性.

下面, 我们不做具体的量子力学计算, 而是利用标度讨论 (scaling argument) 得出两个中性分子之间的色散力, 包含 London 的不包含相对论效应的 non-retarded 情况和 Casimir-Polder 的考虑相对论修正的 retarded 情况 [4].

考虑两个体积分别为 V_1 和 V_2 的中性分子, 其电子运动的特征频率是 ω_0 . 基于标度分析, 间距为 r 的两分子间相互作用势能应采取如下形式:

$$V(r) \approx -\hbar\omega_0 \frac{V_1 V_2}{r^6} f\left(\frac{r}{c/\omega_0}\right), \quad (2)$$

其中, 负号是因为吸引力, r^{-6} 是基于两个体积乘积的量纲, $\hbar\omega_0$ 是体系的特征能量, c/ω_0 是体系的特征距离, $f(x)$ 是一个无量纲标度函数, 用于在 non-retarded 和 retarded 两种情形之间进行过渡.

I) 在短距离时, 即 $r < (c/\omega_0)$, 两分子之间的信号传递时间小于特征时间, 不需要考虑相对论效应, 取 $f(x) = \text{常数}$, 这样就得到 London 的结果: $V(r) \approx -\hbar\omega_0 (V_1 V_2 / r^6)$.

II) 在长距离时, 即 $r > (c/\omega_0)$, 两分子之间的信号传递时间大于特征时间, 需要考虑相对论效应. 此时相互作用势不再依赖于表征特征时间的 ω_0 , 为了抵消掉 ω_0 , 取 $f(x) = 1/x$, 这样就得到 Casimir-Polder 的结果: $V(r) \approx -\hbar c (V_1 V_2 / r^7)$, 其中, 光速 c 的出现体现了相对论效应.

3 计算

本节给出几个用零点能计算卡西米尔力的算例, 并且对其中处理发散的正规化方法 (regularization) 进行介绍.

3.1 一维空间标量场

首先考虑一个简单的情形: 一维空间的零质量标量场, 受限于两点之间, 取狄利克雷边界条件 (Dirichlet boundary condition) 引入约束, 点间距

为 a .

由量子场论, 量子化的标量场可看作为一组振动的谐振子. 由量子力学, 基态或真空态下, 频率为 ω 的谐振子具有零点能 $\hbar\omega/2$. 标量场的零点能为所有频率谐振子的零点能之和. 受限情形下, 由边界条件知道, 允许的频率为: $\omega_n = ck_n$, 这里, $k_n = (n\pi)/a$, $n = 1, 2, 3, \dots$

计算标量场真空能量:

$$E_0(a) = \frac{\hbar c}{2} \sum_{n=1}^{\infty} \frac{n\pi}{a}, \quad (3)$$

显然, 该无穷级数求和发散. 为了处理发散, 需要对 (3) 式进行正规化. 有多种正规化的方法, 这里, 我们采取以下两种方法.

3.1.1 衰减函数法

引入衰减函数 $f(n, x) = \exp(-k_n/A)$, 这里, $x = A^{-1} > 0$. 函数满足: $f(n, 0) = 1$; $f(n, x \neq 0) \rightarrow 0$, 当 $k_n \gg A$. 因此有:

$$\begin{aligned} E_0(a, x) &= \frac{\hbar c}{2} \sum_{n=1}^{\infty} \frac{n\pi}{a} e^{-x(n\pi/a)} = -\frac{\hbar c}{2} \frac{\partial}{\partial x} \sum_{n=1}^{\infty} e^{-x(n\pi/a)} \\ &= -\frac{\hbar c}{2} \frac{\partial}{\partial x} \left(\frac{e^{-x\pi/a}}{1 - e^{-x\pi/a}} \right) = \frac{\hbar c\pi}{2a} \frac{e^{x\pi/a}}{(e^{x\pi/a} - 1)^2} \\ &\approx \frac{\hbar c\pi}{2a} \left[\frac{1}{(x\pi/a)^2} - \frac{1}{12} + O(x^2) \right] \\ &= \frac{\hbar ca}{2\pi x^2} - \frac{\pi\hbar c}{24a} + O(x^2). \end{aligned} \quad (4)$$

对于 $x \neq 0$, (4) 式给出正规化的有限大小真空能量. 在正规化方法里, 令 $x \rightarrow 0$, 显然首项发散, 第二项是有限项, 而高阶 $O(x^2) \rightarrow 0$.

如何处理发散的首项呢? 需要减去标量场的自由真空能, 这是重整化操作 (renormalization). 下面计算自由真空能.

在自由情形下, 标量场有连续的涨落谱: $w(k) = \sqrt{c^2 k^2} = c|k|$. 真空能为

$$E_0^F = \frac{\hbar cL}{2} \int_{-\infty}^{\infty} \frac{dk}{2\pi} |k| = \frac{\hbar cL}{2\pi} \int_0^{\infty} k dk. \quad (5)$$

显然积分发散, 现引入衰减函数 $f(k, x) = \exp(-xk)$ 做正规化, 这里 $x > 0$. 有

$$E_0^F(x) = \frac{\hbar cL}{2\pi} \int_0^{\infty} dk k e^{-xk} = \frac{\hbar cL}{2\pi x^2} \Gamma(2) = \frac{\hbar cL}{2\pi x^2}, \quad (6)$$

其中, $\Gamma(x)$ 是伽马函数. 因此, 在有限区间内的自由真空能为

$$E_0^F(a, x) = \frac{E_0^F(x) a}{L} = \frac{\hbar c a}{2\pi x^2}, \quad (7)$$

可发现, 该结果正等于受限真空能 (4) 式的首项发散项. 现在计算正规化卡西米尔能量:

$$E_c(x, a) = E_0(a, x) - E_0^F(a, x) = -\frac{\pi\hbar c}{24a} + O(x^2), \quad (8)$$

取极限 $x \rightarrow 0$, 得到卡西米尔能量:

$$E_c(a) = -\frac{\pi\hbar c}{24a}, \quad (9)$$

计算卡西米尔力:

$$F = -\frac{\partial E_c(a)}{\partial a} = -\frac{\pi\hbar c}{24a^2}, \quad (10)$$

这是一个长程吸引力. 由该结果可以看出:

1) 结果仅依赖于基本常数 \hbar 和 c , 以及约束间距 a , 与微观细节无关, 因此卡西米尔力具有普适性;

2) 由于普适性, 标度关系 a^2 可由量纲分析直接得到, 因为 $[\hbar c] = E \cdot L$, $[F] = E/L$, a 是本问题的唯一长度尺度, 故有 $F \approx \hbar c/a^2$;

3) 常数项的大小和符号, 需要具体计算, 这里的结果是常数项为负, 故为吸引力, 我们指出, 依赖于场的类型、边界条件、空间维度等, 卡西米尔力可为排斥或吸引, 需要进行具体的计算.

3.1.2 zeta 函数正规化

做正规化:

$$\begin{aligned} E_0(a, s) &= \frac{\hbar c}{2} \sum_{n=1}^{\infty} \left(\frac{n\pi}{a}\right)^{1-2s} \\ &= \frac{\hbar c}{2} \left(\frac{\pi}{a}\right)^{1-2s} \sum_{n=1}^{\infty} n^{1-2s} \\ &= \frac{\hbar c}{2} \left(\frac{\pi}{a}\right)^{1-2s} \zeta(2s-1), \quad (11) \end{aligned}$$

(11) 式中使用了黎曼 zeta 函数的级数定义: $\zeta(z) = \sum_{n=1}^{\infty} n^{-z}$. 我们知道, 级数定义只是在 $\text{Re}(z) > 1$ 收敛, 即 (11) 式正规化能量的收敛要求是 $\text{Re}(s) > 1$. 然而, 黎曼 zeta 函数可以解析延拓至复平面, 是一个只在 $z = 1$ 有简单极点的亚纯函数 (meromorphic function).

在正规化方法中取 $s \rightarrow 0$. 利用解析延拓后黎曼 zeta 函数的性质有 $\zeta(-1) = -1/12$, 得到 $E_0(a) = \lim_{s \rightarrow 0} E_0(a, s) = -(\pi\hbar c)/(24a)$, 同 (9) 式通

过衰减函数正规化方法得到的结果一致. 可以说, zeta 函数 (不限于黎曼 zeta 函数) 正规化是一种最优雅的正正规化方法 [6].

3.2 三维空间电磁场

本节研究 Casimir 对于真空中受限电磁场的计算. 不同于 Casimir 采用的 Euler-Maclaurin 方法, 我们用 zeta 函数正规化方法迅速得到结果.

在量子场论中, 电磁场表现为一组谐振子自由度, 真空中充满了这些涨落模式. 在真空中放入两个 Casimir 平板, 板面积为 A , 板间距为 a . 板的存在将涨落模式修改为: $\omega_{\mathbf{k}_{\perp}, n} = c\sqrt{\mathbf{k}_{\perp}^2 + (n\pi/a)^2}$, 其中, $n = 0, \pm 1, \pm 2, \dots$. 注意, 不同于标量场的情况, 这里 n 需要同时取正负整数, 这是为了考虑电磁场涨落 (光子) 的两个极化自由度, 同时, 模式 $(\mathbf{k}_{\perp}, 0)$ 可以被激发, 即允许 $n = 0$.

计算受限的真空能:

$$\begin{aligned} E_0(a) &= \frac{A\hbar}{2} \int_{-\infty}^{\infty} \frac{d^2\mathbf{k}_{\perp}}{(2\pi)^2} \sum_{n=-\infty}^{\infty} \omega_{\mathbf{k}_{\perp}, n} \\ &= \frac{A\hbar}{2} \int_{-\infty}^{\infty} \frac{d^2\mathbf{k}_{\perp}}{(2\pi)^2} \left[2 \sum_{n=1}^{\infty} \omega_{\mathbf{k}_{\perp}, n} + \omega_{\mathbf{k}_{\perp}, 0} \right]. \quad (12) \end{aligned}$$

舍去与板间距 a 无关的项, 得到 Casimir 能量:

$$\begin{aligned} E_0^C(a) &= 2 \times \frac{A\hbar}{2} \int_{-\infty}^{\infty} \frac{d^2\mathbf{k}_{\perp}}{(2\pi)^2} \sum_{n=1}^{\infty} \omega_{\mathbf{k}_{\perp}, n} \\ &= 2 \times \frac{A\hbar c}{2} \int_{-\infty}^{\infty} \frac{d^2\mathbf{k}_{\perp}}{(2\pi)^2} \sum_{n=1}^{\infty} \left[\mathbf{k}_{\perp}^2 + \left(\frac{n\pi}{a}\right)^2 \right]^{1/2}, \quad (13) \end{aligned}$$

显然, (13) 式发散. 现在引入正规化因子 s , 推导正规化的 Casimir 能量如下:

$$\begin{aligned} E_0^C(a, s) &= A\hbar c \int_{-\infty}^{\infty} \frac{d^2\mathbf{k}_{\perp}}{(2\pi)^2} \sum_{n=1}^{\infty} \left[\mathbf{k}_{\perp}^2 + \left(\frac{n\pi}{a}\right)^2 \right]^{(1-2s)/2} \\ &= \frac{A\hbar c}{2\pi} \left(\frac{\pi}{a}\right)^{3-2s} \sum_{n=1}^{\infty} n^{3-2s} \int_0^{\infty} x dx [x^2 + 1]^{(1-2s)/2} \\ &= \frac{A\hbar c}{2\pi} \left(\frac{\pi}{a}\right)^{3-2s} \zeta(2s-3) \frac{1}{2s-3}, \quad (14) \end{aligned}$$

在 (14) 式的推导中, 应用了黎曼 zeta 函数的级数定义, 由级数定义的收敛性要求 $\text{Re}(s) > 2$. 并且在求无穷积分时假定了 $\text{Re}(s) > 3/2$.

现在在正规化能量 (14) 式中取极限 $s \rightarrow 0$, 利用黎曼 zeta 函数的解析延拓, 可得 $\zeta(-3) = 1/120$.

代入 (14) 式, 得到单位面积的 Casimir 能量:

$$\frac{E_0^C(a)}{A} = -\frac{\pi^2 \hbar c}{720a^3}, \quad (15)$$

则 Casimir 力为

$$P(a) = -\frac{1}{A} \frac{\partial E_0^C(a)}{\partial a} = -\frac{\pi^2 \hbar c}{240a^4}, \quad (16)$$

即 (1) 式. 这样就利用真空零点能推导出卡西米尔力, 得到普适常数的数值.

4 临界卡西米尔力

在连续相变的临界点, 序参量涨落的关联长度发散, 因此序参量场具有长程关联, 引入外加约束可以产生临界卡西米尔力. 1978 年, Fisher 和 de Gennes 考虑在两元混合液体中引入两个 Casimir 平板, 即两个平行的宏观平板, 调整参数至混合液体相分离的临界点, 此时序参量 (液体浓度) 具有长程关联的临界涨落, 关联长度 $\xi \rightarrow \infty$. 由于两板存在, 涨落谱被修改, Fisher 和 de Gennes^[7]认为两板间产生临界卡西米尔力.

临界卡西米尔力具有普适性, 我们再次用量纲分析研究该问题. 由于温度的存在, 取代能量, 需要研究体系的自由能 $F(a)$. 自由能是一个广延量, 因此与板面积 A 成比例. 序参量涨落是热涨落, 因此特征能量是热能 $k_B T$. 由于关联长度在临界点发散, 问题里只剩一个长度尺度, 即板间距 a . 由量纲分析, 立即可以写出:

$$F(a) = C k_B T \frac{A}{a^2}, \quad (17)$$

其中 C 是普适常数.

上述论证还可以按照统计物理中的 Blob 方法估算得出: 本问题的序参量涨落是热涨落, 因此可定义热 Blob, 其能量为特征能量 $k_B T$; 由于关联长度发散, 所以体系有唯一的特征尺寸 a , 故热 Blob 的体积是 a^3 ; 体系的 Blob 数目为 $g = (Aa)/a^3 = A/a^2$; 体系的自由能为 $F \approx g k_B T = k_B T A/a^2$.

和量子场论的情况一样, 普适常数 C 的数值需要通过具体的计算得出. 一般来说, 临界卡西米尔力普适常数的计算比第 3 节中量子电动力学卡西米尔力的计算要复杂, 这是因为电磁理论是一个线性理论, 而这里的相变问题则是一个非线性的场论. 计算过程需要使用临界现象理论里发展出的 $d = 4 - \epsilon$ 展开等技术. 在两维体系, 由于体系在临

界点不仅具有标度不变性, 并且进而有共形不变性, 因此可以用共形场论的方法计算普适常数. 另外我们指出, 依赖于两板上施加的边界条件类型, 临界卡西米尔力可以是吸引或排斥力, 即普适常数 C 可正可负^[8].

5 戈德斯通模

由戈德斯通定理可知, 体系在破缺连续对称性时, 将产生戈德斯通粒子 (Goldstone particle) 或戈德斯通模 (Goldstone mode), 这是一种零质量的模式, 因此是长程关联的. 扰动戈德斯通模也会产生卡西米尔力.

统计力学里, 戈德斯通模通常采用具有 $O(n)$ 对称的 n -矢量模型进行展示. 体系的序参量: $\psi = (\psi_1, \psi_2, \dots, \psi_n)$. 体系的哈密顿量对于序参量具有连续的旋转对称性. 经历相变后, 体系发生自发对称破缺, 即序参量破缺哈密顿量的连续旋转对称, 在 n -矢量空间中自发选出一个方向排列, 体系进入有序态. 序参量的涨落模被分成 1 个平行于有序方向的纵向模式 (软模) 和 $(n-1)$ 个垂直于有序方向的横向模式 (戈德斯通模). 由于戈德斯通模不需要耗能, 因此是长程关联的涨落. 需要指出, 纵向的软模导致了体系的有序, 其只在临界点质量为零, 软模涨落导致的正是前面讨论的临界卡西米尔力. 然而, 不限于临界点, 零质量的戈德斯通模存在于整个临界点以下.

下面以几个软物质物理中的例子来说明戈德斯通模诱导产生的卡西米尔力. 1) 液晶. 液晶体系通常由棒状分子组成, 这一分子特点在统计问题中引入了取向自由度, 即可以研究体系的连续旋转对称性. 如上所述, 这就为问题提供了戈德斯通模, 进而戈德斯通模可以诱导产生卡西米尔力. 2) 薄膜或界面. 由于薄膜或界面平面具有连续平移对称性, 因此破缺该对称性会导致戈德斯通模, 这里就是声子 (phonon), 声子导致薄膜或界面的毛细涨落 (capillary fluctuations). 在薄膜或界面上引入外界约束物, 将修改体系的涨落谱, 因而可以诱导产生卡西米尔力. 3) 聚合物. 聚合物统计力学里的一个重要结果是 de Gennes 认识到聚合物的排除体积问题 (对应于自回避无规行走) 可以用自旋 n -矢量模型的 $n \rightarrow 0$ 极限描述^[9]. 如上所述, n -矢量模型有 $(n-1)$ 个戈德斯通模, 若考虑在聚合物熔体

中引入两个 Casimir 平板, 计算戈德斯通模诱导的卡西米尔力:

$$P(a) = -\frac{(n-1)\partial F(a)}{A\partial a} \xrightarrow{n \rightarrow 0} -\left(-\frac{\partial F(a)}{A\partial a}\right), \quad (18)$$

因此, 聚合物涨落背景诱导的卡西米尔力与一般的卡西米尔力反号, 如果一般的卡西米尔力是吸引力, 则聚合物诱导卡西米尔力是排斥力^[10,11].

6 实验验证

自量子电动力学的卡西米尔力在 1948 年提出后, 引发了一系列的实验工作来验证卡西米尔力. 先计算一下卡西米尔力的大小, 由 (1) 式可写出: $P \approx -1.3 \times 10^{-5} \text{ atm} \times (1 \mu\text{m}/a)^4$ ($1 \text{ atm} = 101.325 \text{ kPa}$). 因此, 当板间距 $a = 1 \mu\text{m}$ 时, 卡西米尔力大约为 1.3 mPa , 这已经是一个宏观的数值了. 减小板间距离至微米以下, 卡西米尔力以 a^{-4} 方式增大.

经过这个估算, 实验上应该在微米尺度测量卡西米尔力. 早期有一些定性的探测卡西米尔力的实验工作^[12]和一些高精度的实验验证^[13,14]. 此后, 由于卡西米尔力在微米尺度是不能忽视的作用力, 因此出现了应用卡西米尔力的工作, 主要是在微米尺度的一些金属机器的制造上, 称为微米电力系统 (microelectromechanical systems, MEMS)^[15-17], 这些微米尺度机器的应用研究导致的两个重要的问题是: 1) 由于卡西米尔力是吸引力, 器件将被粘附在一起而无法正常工作, 那么, 如何实现卡西米尔排斥力; 2) 由于器件有不同的形状, 那么, 如何计算超越 Casimir 平板的不同形状器件之间的力.

临界卡西米尔力的重要验证实验包括研究液氮薄膜在经历超流相变时膜厚的变薄过程^[18]和研究在两元混合液体临界点附近胶体与表面的相互作用^[19]. 超流实验有意思的地方在于临界卡西米尔力和戈德斯通卡西米尔力将一起影响膜厚. 在超流相变临界点 (T_λ) 附近, 序参量涨落变成零质量的临界涨落, 该涨落在液氮薄膜的两个表面 (固体铜片表面和液气界面) 之间诱导临界卡西米尔吸引力, 导致膜厚降低. 在临界点 T_λ 以下, 临界涨落消失, 膜厚回升. 可是实验发现, 膜厚没有回到临界点以上的数值, 而是相对偏小, 原因有两个: 1) 超流相变的序参量 (位相角) 涨落导致戈德斯通模, 戈德斯通模诱导的卡西米尔吸引力使 T_λ 以下膜厚较临界点以上偏小; 2) 薄膜的下表面是氮的液气

界面, 在超流态, 黏度为零, 界面涨落的戈德斯通模 (毛细涨落) 诱导产生两表面间的卡西米尔吸引力, 进一步减小膜厚. 两种因素的联合作用解释了实验的观察结果^[1].

7 一些发展

在卡西米尔力的处理中, 有一些重要的假定. 1) 将外加约束与涨落介质的相互作用以施加的边界条件处理. 这种处理是基于体系中时间尺度的分离, 即外界约束运动的特征时间远远大于介质涨落的特征时间. 该处理本质上是在一个介质和约束的体系中, 将约束的运动作为慢变量而固定, 得到一个关于涨落介质的有效模型, 而约束慢变量则影响介质的涨落谱. 显然, 这种处理中, 只有约束对于介质的作用, 而没有介质对于约束的反作用. 2) 在 Casimir 平板的边界条件中, 平板无穷薄, 因此没有考虑介质的穿透效应.

在卡西米尔力理论的发展中, 针对不同的假定有后续的发展. Lifshitz^[20]和 Woods 等^[21]发展了 Lifshitz 理论, 该理论将卡西米尔力推广到实际材料体系, 利用涨落耗散定理推导了麦克斯韦能动量张量, 从而得出不同温度下两个平行的介电材料之间的相互作用力. Schwinger^[22]利用自己的量子场论 non operator 版本, 即源理论 (source theory), 重新推导了卡西米尔力, 在该推导中没有出现真空的零点能. 随着理论物理的发展, 人们认识到产生卡西米尔力的约束, 可以是来自外加边界条件, 也可以是在温度场论中通过松原理论 (Matsubara formulation) 将温度引入场论时带来的周期性条件约束. 在宇宙学中, 还可以是在处理非欧空间 (non Euclidean) 时由无边界空间的拓扑性质 (topology) 引入的“Identification”约束等. 另一方面, 在长程关联涨落介质的寻找上, 人们发现在非平衡统计体系里, 动力学守恒率可以使非平衡体系产生幂率衰减的长程关联行为, 这使得非平衡体系, 包括颗粒体系、活性物质等, 成为另一个研究卡西米尔力的重要情形^[23], 需要指出, 与平衡体系相比, 非平衡体系的卡西米尔力不具有普适性. 这些进展都是卡西米尔物理发展的精彩篇章.

8 结论和展望

本文讨论了广义的卡西米尔力. 通过考察具有

长程关联的不同涨落介质, 包括电磁场、临界场论、戈德斯通模, 以及非平衡体系, 可以看出, 卡西米尔物理涵盖了物理学里宽泛的研究方向. 在未来卡西米尔力的研究中, 从基础物理的角度, 通过研究约束的不同实现和处理方法, 以及研究不同类型的长程关联涨落场的产生, 卡西米尔力将会有有一个更为宽广的研究方向, 并有希望通过卡西米尔力的概念, 在不同的研究题目和方向之间建立联系. 从应用的角度, 量子电动力学里的卡西米尔力是量子效应的宏观反映, 在微米以下器件的设计中, 必须考虑卡西米尔力, 并通过对其进行调控实现不同的应用功能, 而临界、戈德斯通、非平衡等热卡西米尔力也可以用来调控统计体系中结构和序的形成.

参考文献

- [1] Kardar M, Golestanian R 1999 *Rev. Mod. Phys.* **71** 1233
- [2] Bordag M, Klimchitskaya G L, Mohideen U, Mostepanenko V M 2009 *Advances in the Casimir Effect* (New York: Oxford University Press)
- [3] Casimir H B G 1948 *Proc. K. Ned. Akad. Wet.* **51** 793
- [4] Milton K A 2001 *Casimir Effect: Physical Manifestations of Zero Point Energy* (Singapore: World Scientific Publishing Co. Pte. Ltd.)
- [5] Dalvit D, Milonni P, Roberts D, Rosa F 2011 *Casimir Physics* (Heidelberg: Springer) pp1–3
- [6] Elizalde E, Odintsov S D, Romeo A, Bytsenko A A, Zerbini S 1994 *Zeta Regularization Techniques with Applications* (Singapore: World Scientific Publishing Co. Pte. Ltd.)
- [7] Fisher M E, de Gennes P G 1978 *C. R. Seances Acad. Sci., Ser. B* **287** 207
- [8] Maciołek A, Dietrich S 2018 *Rev. Mod. Phys.* **90** 045001
- [9] de Gennes P G 1979 *Scaling Concepts in Polymer Physics* (London: Cornell University Press) pp271–281
- [10] Obukhov S P, Semenov A N 2005 *Phys. Rev. Lett.* **95** 038305
- [11] Semenov A N, Obukhov S P 2005 *J. Phys.: Condens. Matter* **17** S1747
- [12] Sparnaay M J 1958 *Physica* **24** 751
- [13] Lamoreaux S K 1997 *Phys. Rev. Lett.* **78** 5
- [14] Mohideen U, Roy A 1998 *Phys. Rev. Lett.* **81** 4549
- [15] Serry F M, Walliser D, Maclay G J 1998 *J. Appl. Phys.* **84** 2501
- [16] Buks E, Roukes M L 2001 *Phys. Rev. B* **63** 033402
- [17] Chan H B, Aksyuk V A, Kleiman R N, Bishop D J, Capasso F 2001 *Science* **291** 1941
- [18] Garcia R, Chan M H W 1999 *Phys. Rev. Lett.* **83** 1187
- [19] Hertlein C, Helden L, Gambassi A, Dietrich S, Bechinger C 2008 *Nature* **451** 172
- [20] Lifshitz E M 1956 *Sov. Phys. JETP* **2** 73
- [21] Woods L M, Dalvit D A R, Tkatchenko A, Rodriguez-Lopez P, Rodriguez A W, Podgornik R 2016 *Rev. Mod. Phys.* **88** 045003
- [22] Schwinger J 1975 *Lett. Math. Phys.* **1** 43
- [23] Aminov A, Kafri Y, Kardar M 2015 *Phys. Rev. Lett.* **114** 230602

SPECIAL TOPIC—Statistical physics and complex systems

Casimir force^{*}

Miao Bing[†]

(College of Materials Science and Opto-Electronic Technology, University of Chinese Academy of Sciences, Beijing 100049, China)

(Received 26 March 2020; revised manuscript received 9 April 2020)

Abstract

Casimir force in quantum electrodynamics is the representation of zero point energy of vacuum. Depending on the type of fluctuation medium, generalized Casimir force covers a wide spectrum of topics in physics, such as, quantum, critical, Goldstone mode, and non-equilibrium Casimir force. In general, long range correlated fluctuations and constraints are two conditions for generating the Casimir force. In this paper, through a survey of the development of Casimir physics, we discuss several types of Casimir forces and several regularization methods. We end the paper with an outlook for the further development of Casimir physics in the future.

Keywords: Casimir force, fluctuations, long range correlation, regularization

PACS: 05.40.-a, 05.70.Jk, 11.10.-z

DOI: 10.7498/aps.69.20200450

* Project supported by the National Natural Science Foundation of China (Grant Nos. 21774131, 21544007).

† Corresponding author. E-mail: bmiao@ucas.ac.cn

复杂势场量子弹球中疤痕态的量子化条件

李晓亮 陈宪章 刘郴荣 黄亮

Quantization condition of scarring states in complex soft-wall quantum billiards

Li Xiao-Liang Chen Xian-Zhang Liu Chen-Rong Huang Liang

引用信息 Citation: *Acta Physica Sinica*, 69, 080506 (2020) DOI: 10.7498/aps.69.20200360

在线阅读 View online: <https://doi.org/10.7498/aps.69.20200360>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

二维方形量子点体系等离激元的量子化

Quantization of plasmon in two-dimensional square quantum dot system

物理学报. 2017, 66(22): 227301 <https://doi.org/10.7498/aps.66.227301>

耗散耦合腔阵列耦合量子化腔场驱动三能级体系中的单光子输运

Single photon transport by a quantized cavity field driven cascade-type three-level atom in a dissipative coupled cavity array

物理学报. 2017, 66(5): 054204 <https://doi.org/10.7498/aps.66.054204>

复杂氧化物中电子相分离的量子调控

Quantum manipulation of electronic phase separation in complex oxides

物理学报. 2018, 67(22): 227502 <https://doi.org/10.7498/aps.67.20182007>

量子计算与量子模拟

Quantum computation and quantum simulation

物理学报. 2018, 67(12): 120301 <https://doi.org/10.7498/aps.67.20180710>

量子相干

Quantum coherence

物理学报. 2019, 68(3): 030304 <https://doi.org/10.7498/aps.68.20181779>

专题：统计物理和复杂系统

复杂势场量子弹球中疤痕态的量子化条件*

李晓亮 陈宪章 刘桀荣 黄亮†

(兰州大学物理科学与技术学院, 兰州 730000)

(2020年3月11日收到; 2020年4月3日收到修改稿)

量子疤痕是波函数在经典不稳定周期轨道周围反常凝聚的一种量子或波动现象. 人们对疤痕态的量子化条件进行了大量研究, 对深入理解半经典量子化起到了一定的促进作用. 之前大部分研究工作主要集中在硬墙量子弹球上, 即给定边界形状的无穷深量子势阱系统. 本文研究具有光滑复杂势场的二维量子弹球系统, 考察疤痕态的量子化条件及其重复出现的规律, 得到了与硬墙弹球不一样的结果, 对理解这类现象是一个有益的补充. 这些结果将有助于理解具有无规长程杂质分布的二维电子系统的态密度谱和输运行为.

关键词: 量子疤痕, 软墙量子弹球, 复杂势场量子弹球, 量子化条件**PACS:** 05.45.Mt, 03.65.Sq**DOI:** 10.7498/aps.69.20200360

1 背景

经典量子对应是量子力学一个重要的基本问题. 从玻尔的旧量子论, 到量子力学的建立, 一直到近代对量子力学基本原理的探讨, 都离不开这个问题. 众所周知, 对于经典系统, 有可积和不可积之分, 而对于实际系统来说, 绝大多数都是不可积的. 可积系统可以通过玻尔-索末菲量子化规则或后来发展的 Einstein-Brillouin-Keller (EBK) 量子化条件^[1]来量子化, 得到对应量子系统的能级. 而不可积的系统由于没有整体的作用量—角变量这些对偶量, 不能简单地应用量子化规则^[2,3]. 到上世纪六七十年代, 随着原子核物理、随机矩阵理论、经典混沌的发展, 人们重新考察经典不可积系统的量子化问题, 得到了系列结果, 特别是 Gutzwiller^[4]的求迹公式, 发现对于对应经典动力学为混沌的量子系统, 虽然经典不稳定周期轨道的测度是零, 但是这些轨道却决定了该量子系统能谱的涨落特征. 其中一类重要的模型是二维硬墙 (hard-wall) 量子弹球^[5–8], 即二维无限深势阱, 其经典极限对应着

动力学弹球——粒子在其中做自由运动, 在边界处发生镜像反射. 系统的动力学行为完全由边界形状决定. 当边界不规则时, 粒子的经典动力学行为往往是混沌的. 对于量子弹球系统, 其方程为单粒子薛定谔方程. 由于这一方程为线性方程, 排除了混沌发生的可能, 因此, 在半经典极限下如何发展出混沌就成了一个非常引人入胜的问题. 此外, 当改变边界使得经典动力学是可积或混沌时, 所对应的量子系统有没有什么特征? 有没有什么量子系统的指标可以明确说明其所对应的经典动力学是否混沌? 对这些问题的研究构成了量子混沌学, 简称为量子混沌^[9]. 经过半个多世纪的研究, 在很多方面, 如波函数统计特征、能谱统计、量子混沌散射、量子保真度、动力学演化特征等, 人们都得到了深入的理解^[5,8,10–15]. 注意在讲可积或混沌时, 均指的经典系统, 或量子系统在经典极限下的行为.

“量子疤痕”作为量子波函数在经典不稳定轨道上的反常凝聚, 是量子世界对经典轨道的一个敏锐的反常响应^[16–18]. 轨道为经典的概念, 波函数为量子的概念, 在一些情况下, 量子波函数在空间中

* 国家自然科学基金 (批准号: 11775101) 资助的课题.

† 通信作者. E-mail: huangl@lzu.edu.cn

会凝聚在经典的周期轨道上, 即粒子在这些经典轨道附近的概率要远大于其他地方. 对于经典混沌系统, 如运动场形或其他具有特定边界形状的弹球系统, 由于其周期轨道的不稳定性, 周期轨道在相空间的测度为零, 粒子保持在该轨道上运动的概率是零; 但对于量子系统, 在某些能级下, 粒子反而会以很大的概率“凝聚”在这些轨道上, 这一违反直观的现象就是量子疤痕. 当然, 对于经典可积系统或混合动力系统 (即相空间中既有混沌海又有稳定轨道所形成的 Kolmogorov-Arnold-Moser (KAM) 岛), 对应量子系统的本征波函数在稳定轨道上也有很强的凝聚, 而且比混沌轨道上的凝聚更强, 但由于在经典情形下粒子就能够自然地在这一轨道上运动, 所以并不奇怪. 量子疤痕作为联系经典轨道和量子现象的纽带, 受到了人们的广泛关注, 并给出了半经典的解释, 即量子波函数在该轨道上传播时相位的相干叠加产生的加强效应 [19–22].

之前关于量子疤痕的研究主要集中在硬墙量子弹球里, 关于软墙 (soft-wall) 量子弹球的研究相对较少. 软墙 (soft-wall) 和硬墙的主要区别在于弹球边界上的势场是光滑渐变的还是不连续的阶梯势. 本文将主要研究软墙量子弹球系统以及更加广泛的具有复杂光滑势场弹球系统的疤痕态及其所满足的量子化条件. 软墙会对经典动力学带来重要的影响. 已经发现, 即使对于经典可积的椭圆形硬墙弹球, 当边界足够软时, 系统变为混合型的, 在相空间中 KAM 岛周围出现了混沌海 [23]. 在软墙量子弹球方面, Luukko 等 [24] 和 Keski-Rahkonen 等 [25–27] 研究了二维异向谐振子, 在加入很多短程杂质破坏系统的可分离性后, 发现了与经典李萨如图形一致的李萨如疤痕态. 他们的工作详细讨论了频率比作为主要参数对系统行为的影响, 认为李萨如疤痕态的出现是量子力学对经典混沌的抑制作用, 但并没有讨论半经典量子化的条件. 本文将主要讨论疤痕态的半经典量子化条件, 并考虑更加复杂的光滑势场所带来的影响. 一般来讲, 疤痕态指凝聚在经典不稳定周期轨道上的波函数. 在此, 由于我们主要关心在这些轨道上波函数半经典量子化的条件, 我们将对稳定和不稳定轨道都做处理, 来考察他们半经典量子化的条件.

疤痕态出现的必要条件与 EBK 量子化条件类似, 即从一点出发, 沿轨道行进一周, 所累积的总的相位是 2π 的整数倍, 以保证波函数的单值性

和加强相干叠加 [14,28]. 只是对于较复杂的情况, 沿轨道行进时的相位贡献来源比较复杂. 关于相位的计算, 对于硬墙二维量子弹球, 人们已经比较清楚, 主要有在弹球区域内的自由运动项, 即 $\mathbf{p} \cdot d\mathbf{q}$ 项, 以及在边界处的反射系数项和 Maslov 指数项; 而对于二维光滑约束势场 (软墙) 中的量子弹球, 据我们所知, 由于所处理的情形的复杂性, 详细的疤痕态出现时能量所满足的量子化条件以及疤痕态重复出现的规律还没有给出. 虽然半经典量子化条件可以精确求出量子谐振子的本征能量, 但对于一般的系统, 比如受扰动的谐振子, 半经典量子化条件只在较高能级时才适用, 可以近似给出疤痕态出现时的能量值. 在这里, 我们发现, 对于较低的能级, 仍然能够得到疤痕态出现的能量值和半经典理论较好的对应.

改变能量或波矢, 当沿轨道一周累积的相位改变接近 2π 的整数倍时, 疤痕态能够重复出现, 并且重复出现的规律是一个重要指标, 依赖于系统态密度的信息. 对于同一类轨道, 由于在边界处获得的相位相同, 相位差主要由作用量 S 的改变贡献. 对于硬墙量子弹球, 在经典极限下, 由于轨道与能量无关, 动量大小与位置无关, 且动量与轨迹切向平行, 故 $S = \oint \mathbf{p} \cdot d\mathbf{q} = \hbar \oint \mathbf{k} \cdot d\mathbf{q} = \hbar \mathbf{k} \cdot L$, L 为轨道长度. 相邻疤痕态重复出现的必要条件即为 $\Delta k \cdot L = 2\pi$ 或 $\Delta k = 2\pi/L$, 即疤痕态在波矢空间是近似等间距出现的. 对于非相对论量子系统, 由于 $E - E_0 \propto k^2$, E_0 为能量的最低点, 即疤痕态在能量空间是按照 $\sqrt{E - E_0}$ 等间距出现的. 而对于极端相对论性系统, $E \propto k$, 此时在能量空间疤痕态则是按照 E 等间距出现的 [29,30]. 相对于硬墙弹球, 光滑势场量子弹球或软墙弹球的行为更加复杂, 此时周期轨道与能量有关, 哪怕对于同一类轨道, 当能量改变时, 轨道的长度和位置也会发生变化, 而且动量与位置有关, 作用量只能通过数值积分得到. 此外, 对于硬墙量子弹球系统, 每次在边界处的碰撞, 反射系数会贡献一个相位, 同时取决于边界形状, 如果碰撞点处边界为弧形, 还会有共轭点的贡献. 而对于光滑势系统, 除了端点, 在轨道发生偏转时没有相应的反射系数贡献的相位, 只有在轨道的端点处发生完全反射时才会有这个贡献, 同时对轨道共轭点的判断也更加困难. 特别地, 对于某些复杂势场的情况, 有可能存在一些临界点, 当

能量值跨过这些临界点时, 其经典周期轨道可能会出现定性上的改变^[31]. 我们将会碰到这种情况, 并考察当经典周期轨道发生定性的改变时, 比如从不稳定变为稳定, 其上疤痕态的半经典量子化条件的变化.

目前由于二维材料的广泛研究, 人们研究了由杂质原子所形成的岛上区域的电子态密度, 发现对于比较规则的系统, 态密度在随能量变化时会出现一些特征的峰, 与疤痕态所满足的量子化条件一致. 而对于一般情况, 当杂质较多时, 也会出现一些比较明显的峰, 这些峰的来源与特征是什么, 与经典轨道有没有关系, 还是一个没有解决的问题. 杂质所产生的势场往往能够形成大小不等的势谷, 电子局限在势谷中, 在一定的能量值时, 在特定的轨道上形成满足量子化条件的疤痕态, 显现出较高的态密度峰, 并且由于疤痕态的重复出现, 这些峰随着能量的改变也会重复出现, 只是由于势场的复杂性, 态密度峰的周期性会受一定的影响. 因此, 我们研究扰动下二维谐振子势中量子弹球的量子化条件, 以期对复杂约束势场中疤痕态出现的量子化条件以及实验上能够观察得到的态密度峰有进一步的理解.

2 模型

因为我们的目的是研究复杂势场下疤痕态出现的量子化条件, 因此将分两步进行, 一是轻微扰动下的二维量子谐振子, 这时有解析的分析, 可以验证我们的数值结果; 二是在大扰动下出现复杂经典动力学行为的光滑势场中的量子弹球. 在轻微扰动下, 我们将能够看到李萨如疤痕态, 即凝聚在经典李萨如图形(当 x, y 两个方向上的频率比为简单的整数比时所出现的图形)上的疤痕态, 并研究这些疤痕态出现所满足的量子化条件. 在大扰动下, 经典轨道和疤痕态的特征将会有较大的变化, 但是可以预期这些量子化条件仍然适用. 这样, 可以一般性地给出在二维复杂势场量子弹球系统中的量子化条件, 可以预期这些结果对于约束在复杂的随机长程势场中的电子仍然有效. 对于理想的二维量子谐振子, 其哈密顿为

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\mathbf{r}) = -\frac{\hbar^2 \nabla^2}{2m} + V(\mathbf{r}), \quad (1)$$

其中

$$V(\mathbf{r}) = \frac{1}{2}(\omega_x^2 x^2 + \omega_y^2 y^2). \quad (2)$$

系统的本征能量为 $E_{nm} = \hbar[(n+1/2)\omega_x + (m+1/2)\omega_y]$, 相应的本征态为

$$\psi_{n,m}(x, y) = N_{nm} H_n(\sqrt{\omega_x} x) H_m(\sqrt{\omega_y} y) \times e^{-\frac{1}{2}(\omega_x^2 x^2 + \omega_y^2 y^2)}, \quad (3)$$

其中 N_{nm} 是归一化常数, $\alpha_{x,y} = \sqrt{m\omega_{x,y}/\hbar}$, H_n 是 n 阶厄米多项式. 由于该系统可以分离为 x, y 两个方向的一维谐振子的直积, 本征波函数也具有这样的对称性, 一般不会出现凝聚在经典周期轨道上的情况, 因此不会出现与经典情况中的李萨如图形所对应的疤痕态. 文献^[26]通过在二维异向谐振子中加入大量有短程作用的杂质的方法来破坏这个对称性, 在这些杂质的诱导下, 该系统中出现了与经典李萨如图形一致的李萨如疤痕态.

为了方便, 在计算的时候我们重新标度各物理量. 令系统的特征长度为 L_0 , 则 $x = x' L_0$. 由于我们将采用有限差分的方式进行计算, 因此将弹球区域分成间距为 a 的正方形格子, 并令 $t = \hbar^2/(2ma^2)$, 将其作为能量的单位. 因 $\hbar\omega$ 亦是能量的单位, 故令 $\omega = \omega' t/\hbar$, 则 x', ω' 均为无量纲的数, 而 $m\omega_x^2 x^2/2 = m\omega_x'^2 t^2 x'^2 L_0^2/(2\hbar^2) = \omega_x'^2 [L_0^2/(2a^2)] x'^2 t/2$, 这样势能的单位亦为 t .

图1展示了本文所研究的轻微扰动和大扰动下的二维谐振子势, 其中 $\omega_x' L_0/(\sqrt{2}a) = 1$, $\omega_y' = 2\omega_x'$. 在图1(a)中, 我们在谐振子势(2)的基础上, 在 $(x_0/L_0, y_0/L_0) = (-0.278, -0.226)$ 处加了一个 $\delta(x-x_0, y-y_0)$ 的函数势, 用来破坏原来势场沿 x 轴、 y 轴的镜面对称以及沿 z 轴的旋转 180° 的对称性. 在图1所示的区域边界, 再加一个硬墙势, 注意这里也可以在这个区域边界之外加一个增长更快的连续势. 我们将主要分析低能级的态, 这时这个区域边界上及其外围的势场对结果影响很小. 图1(b)在图1(a)的基础上加一个高斯势, 即

$$V(\mathbf{r}) = \frac{1}{2}(\omega_x^2 x^2 + \omega_y^2 y^2) + U e^{-[(x-x_G)^2 + (y-y_G)^2]/(2\sigma^2)}, \quad (4)$$

其中 $U=1t$, $\sigma=0.2828$, $(x_G, y_G) = (-0.3441, 0.1226)$, 来进一步破坏系统的可积性. 这样整个势场形成左右两个谷, 中间一个峰, 其中右侧谷底的位置为 $(x_V/L_0, y_V/L_0) = (0.3574, -0.0179)$, 对应的势能为 $V_{\min} = 0.1053t$. 两个势谷之间存在两个鞍点, 对应

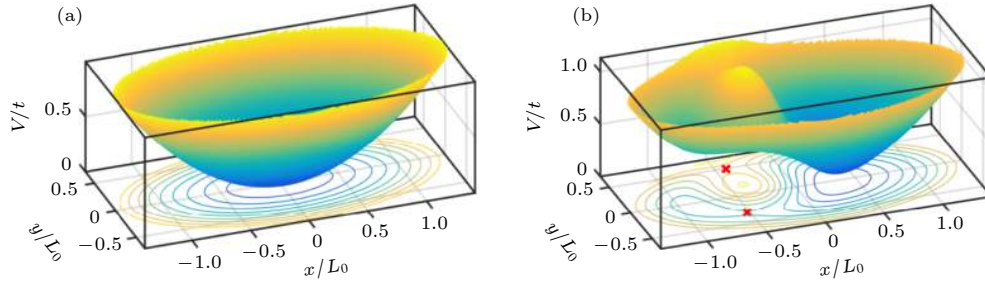


图 1 (a) 二维谐振子势 ((2) 式), 其中 L_0 是系统的特征尺度, $\omega'_x L_0 / (\sqrt{2}a) = 1$, $\omega'_y = 2\omega'_x$. 为了破坏对称性, 我们在 $(x_\delta/L_0, y_\delta/L_0) = (-0.278, -0.226)$ 处加了一个 $\delta(x - x_0, y - y_0)$ 函数势; (b) 在势场 (a) 的基础上, 加入了一个高斯势场 $V_G = Ue^{-[(x-x_G)^2 + (y-y_G)^2]/(2\sigma^2)}$, 其中 $U = 1t$, $\sigma/L_0 = 0.2828$, $(x_G/L_0, y_G/L_0) = (-0.3441, 0.1226)$. 这样整个势场形成左右两个谷, 一个峰, 其中右侧谷底的位置为 $(x_V/L_0, y_V/L_0) = (0.3574, -0.0179)$, 对应的势能为 $V_{\min} = 0.1053t$. 两个势谷之间存在两个鞍点, 如叉号所标示的位置, 对应的势能分别为 $0.591t$ 和 $0.976t$

Fig. 1. (a) Two-dimensional harmonic potential (Eq. (2)), where L_0 is the character scale of the system, $\omega'_x L_0 / (\sqrt{2}a) = 1$, and $\omega'_y = 2\omega'_x$. To break the discrete symmetry, we added a $\delta(x - x_0, y - y_0)$ potential at $(x_\delta/L_0, y_\delta/L_0) = (-0.278, -0.226)$; (b) On the potential given in Fig.(a), we added an additional Gaussian potential $V_G = Ue^{-[(x-x_G)^2 + (y-y_G)^2]/(2\sigma^2)}$, where $U = 1t$, $\sigma/L_0 = 0.2828$, and $(x_G/L_0, y_G/L_0) = (-0.3441, 0.1226)$. Thus the potential field forms two valleys and one peak, and the position of the bottom of the right valley is $(x_V/L_0, y_V/L_0) = (0.3574, -0.0179)$, with corresponding potential $V_{\min} = 0.1053t$. There are two saddle points between the two valleys, as marked by the crosses, with corresponding potential values $0.591t$ and $0.976t$.

的势能分别为 $0.591t$ 和 $0.976t$. 因此, 当粒子能量低于 $0.591t$ 时, 其经典轨迹只能局限在其中的一个势谷中. 这里需注意, 哪怕粒子只被囚禁在一个谷中, 其经典动力学仍然不一定是可积的, 因为软的势场往往能够将经典动力学复杂化^[23].

我们参照文献 [23] 计算了粒子约束在图 1(b) 势场中右侧谷时的庞加莱截面, 结果展示在图 2 中. 粒子在势场中运动时, 动量分解为平行于势能等值线的分量 $p_{//}$ 和垂直于势能等值线的分量 p_{\perp} . 庞加莱截面定义为 $p_{\perp} = 0$ 时的截面, 此时粒子轨迹与势能等值线相切. 这个切点相对于谷底的角度 θ 作为庞加莱截面的参数. 从这个庞加莱截面可以看出, 对于这个系统, 即使能量不太高时, 也有着复杂的结构, 破坏了系统整体的可积性. 但是系统有明确的 Kolmogorov–Arnold–Moser (KAM) 岛, 岛中间对应着稳定轨道. 利用 KAM 岛中心点的参数, 可以精确找到该轨道, 研究该轨道上波函数的量子化条件.

在求解该系统的本征能量和本征态时, 我们将 ∇^2 进行有限差分, 将弹球区域分成间距为 a 的正方形格子, 则得到系统在紧束缚近似下的哈密顿量

$$\hat{H} = 4t \sum_i |i\rangle\langle i| - t \sum_{i,j} |i\rangle\langle j| + \hat{V}, \quad (5)$$

其中 $t = \hbar^2/(2ma^2)$, 第一个求和对所有格点进行, 第二个求和对所有相邻点进行. 在格点表象中, 哈密顿矩阵的对角元为 $H_{ii} = 4t + V(x_i, y_i)$, 如果

格点 i 与格点 j 相邻, 则 $H_{ij} = -t$, 其余矩阵元为 0. 对于边界点, V 取一个很大的值. 在这里取为 $1000t$. 将该哈密顿矩阵对角化即可得到对应的本征能量和本征态. 由于方格子的色散关系为 $E(k_x, k_y) = 2t[1 - \cos(k_x a)] + 2t[1 - \cos(k_y a)]$, 只有当能量或波矢较小时才能近似为 $E = \hbar^2 k^2 / (2m)$, 其中 $k^2 = k_x^2 + k_y^2$. 因此我们将主要考虑低能级情况. 另一方面, 格点系统对原有二维谐振子系统的偏离在一定程度上也促使了李萨如疤痕态的出现.

本文考虑的均为 2 自由度系统. 对于这类系统, 如果轨道是稳定的, 考虑对轨道的变分, 那么在轨道的横截方向, 当有一个小扰动时, 这个小扰动会反复振荡. 考虑横截方向的一个整体模式 o , 假设其振荡频率为 ω_o , 那么这个模式在低能下 (小扰动下) 一般与谐振子的运动相似. 则对于量子化的系统, 它会在能量上贡献^[32]

$$E_o = (m + 1/2)\hbar\omega_o, \quad (6)$$

m 为在横截方向的量子数, 并约定当波函数在横截方向只有一个波包 (即基态) 时, 对应的 m 为 0. 这样总的能量为 $E = E_o + E_t$, 其中 E_t 为沿该轨道运动的能量. 而沿轨道运动的作用量为

$$S_t = \oint \mathbf{p} \cdot d\mathbf{q} = \oint \sqrt{2m(E_t - V(q_t))} dq_t. \quad (7)$$

对于软墙弹球, 当轨道有两个端点 (折返点) 时, 假设在端点处 V 随 q_t 为线性关系, 即 $\partial V / \partial q_t \neq 0$ 或 ∞ , 由于线性势的边界条件与一维

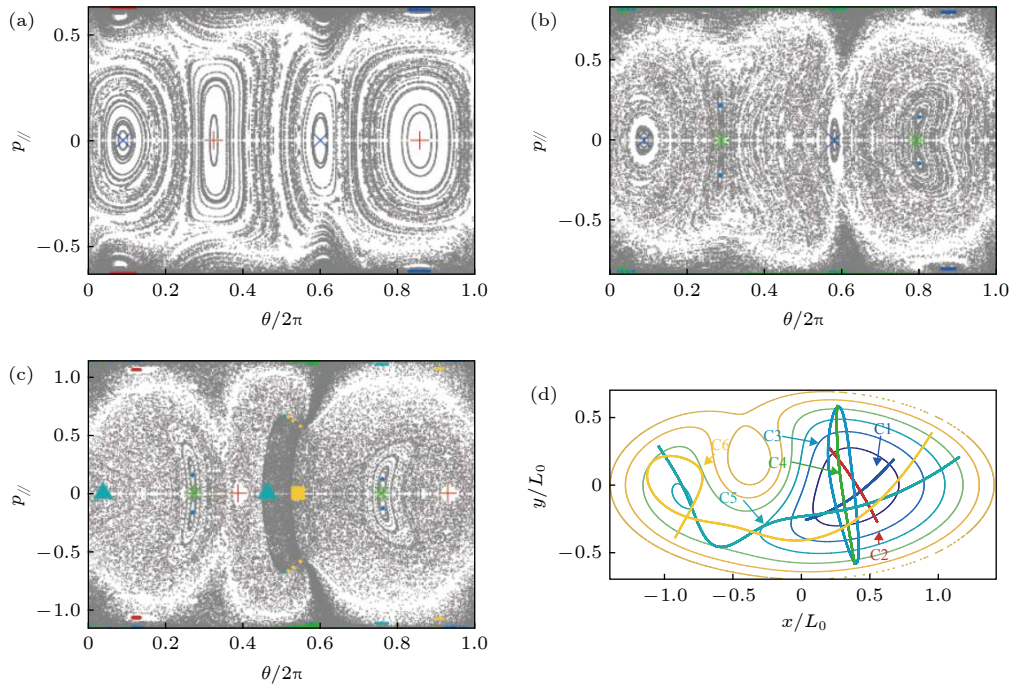


图2 图1(b)势场的庞加莱截面,即 $p_{\perp} = 0$ 时 $p_{//}$ 对于此时的位置相对于右侧谷底的角度 θ ((a)–(c))和后面所处理的6类轨道(d). (a)总能量 $E = 0.3t$; (b) $E = 0.45t$; (c) $E = 0.75t$

Fig. 2. The Poincaré section of the motion of a particle moving in the potential field given by Fig. 1(b), e.g., when $p_{\perp} = 0$, $p_{//}$ vs. the angle θ of this point with respect to the bottom of the right valley (x_v, y_v) ((a)–(c)). The total energy of the particle is $E = 0.3t$ (a), $E = 0.45t$ (b), and $E = 0.75t$ (c), respectively. (d) The six classes periodic orbits that will be discussed later.

情况类似,半经典处理会在波函数中给出 $\pi/4$ 的相位.沿轨道运行一周,抵近和远离每个端点,总共会贡献4倍,即 π 的相位^[14].因此量子化条件为

$$S_t = (n + 1/2)2\pi\hbar. \quad (8)$$

我们约定 n 的意义为,对于凝聚在周期轨道上的量子态,当沿轨道一周时粒子波函数有 $n + 1$ 个波长.这样的约定当系统退化为谐振子势时与一维谐振子的本征能量 $E_n = (n + 1/2)\hbar\omega$ 的约定一致.软墙弹球和硬墙弹球具有很大的区别,首先在折返点,硬墙弹球每次折返贡献 π 的相位,两个折返点共贡献 2π 的相位.此外,在轨道与硬墙边界碰撞发生反射时,由于在边界处波函数恒为0,每次反射会附加一个 π 的相位.而对于软墙弹球,这种反射不贡献额外的相位.

对于给定的疤痕态,通过数疤痕态的波包,可以确定量子数 n ,并由(8)式确定 S_t ,由(7)式可以反推出 E_t . E_o 由(6)式给出.这样,可以得到由半经典的量子化条件给出的本征能量 $E_{mn} = E_o(m) + E_t(n)$ ^[14,32].对于封闭的没有端点的周期轨道,则(8)式应为 $S_t = (n + 1)2\pi\hbar$.注意这里所讨论的是疤痕态出现的必要条件,即疤痕态的能量

需要近似满足在相应轨道上的量子化条件.但是这还不是充分条件,尤其是对于经典混沌的系统,疤痕态只是其中一类态,更多的是没有规律杂乱无章的态.即使满足量子化条件,在所预言的能量值附近,并不一定能够找到对应的疤痕态.因此,相较于在考虑沿周期轨道运动一周所产生的相位后,由半经典量子化给出相应的能量值并在该值附近寻找疤痕态,我们更倾向于反过来,由能量值根据半经典量子化公式得到相应的量子数 n ,并与根据波函数数出来的量子数 n 比较,看是否符合.

如果轨道是不稳定的,则量子化条件与前面类似,只是(6)式中的 $\hbar\omega_o$ 需更换为与轨道不稳定性有关的一个参数 u ^[14,33].

对于约束在复杂的随机长程势场中的电子,当局域在某个或某几个相邻的势谷中时,其所感受到的势场将在定性上与我们所考虑的大扰动下的二维谐振子系统类似,因此可以预期我们的结果对于理解这些系统中态密度的规律具有一定的帮助.

在此必须指出,上述半经典公式是Gutzwiller^[4], Miller^[32], Voros^[33]等在20世纪七八十年代提出来的,在顾雁老师^[14]的书里有详细的阐述.我们只是研究一个具体的软墙弹球的例子,

确定其中疤痕态的量子化条件, 以期对目前广泛研究的具有不规则杂质的二维电子系统的态密度谱峰有所理解. 我们的工作也发现, 这些半经典公式不是仅在短波长极限下成立, 而是对于一般的二维体系, 只要波长不是太长时就能成立, 而且在一个比较大的范围内都与量子计算结果一致.

3 结果

3.1 小扰动下二维谐振子系统

对于小扰动的情况, 我们固定 $\omega_x = 1$, 考察了不同的 ω_y . 对于不同的 ω_x 和 ω_y 的比例, 会得到不同的李萨如疤痕态. 图 3 展示了一些具有代表性的结果.

首先给出 bouncing ball 轨道的本征态的结果, 来检验分析的准确性. 对于图 3(a) 中那一类沿 x 方向的 bouncing ball 态, 它们也是原二维谐振子系统的本征态 ((3) 式), 对应着 $m = 0$ 的态. 此时 $E_{n,0} = \hbar[(n + 1/2)\omega_x + (1/2)\omega_y]$, 其轨道方向的能量为 $E_t(n) = E_{n,0} - (1/2)\omega_y$. 而在 x 方向, 对于经典的谐振子, 能量和作用量的关系为 $S_t = \oint \mathbf{p} \cdot d\mathbf{q} = \oint p_x \cdot dx = 2\pi E_t/\omega_x$. 对于该轨道, 沿轨道方向的量子化条件为 $S_t = 2\pi\hbar(n + 1/2)$. 这样找出所有沿 x 方向 bouncing ball 轨道的本征态出来, 从其本征能量 E , 可以得到 E_t , 并进而得到 S_t , 从而得到量子数 n . 注意这样得到的 n 不一定

是整数.

另一方面, 可以通过数波函数的波包的个数把沿轨道一周的波长的数目确定下来, n 即为波长数减 1. 通过对比这两种方式确定的 n , 可以判断能否通过数波函数波包的方式来得到该波函数相应的本征能量. 对于半经典公式的检验, 一般是反过来, 先确定横截方向的频率值, 在给定量子数下计算沿轨道的作用量, 再由作用量和能量的关系确定能量, 与计算得到的本征能量做对比. 这里我们的目的是给出疤痕态出现时所满足的必要条件, 即半经典量子化条件, 并考察半经典处理对于这些态的适用性问题, 由于疤痕态及其本征能量已经得到, 我们可以通过上述方式得到相应的轨道方向的量子数 n , 和数出来的波长数减 1 做对比. 由于整数 n 对应满足量子化条件, 所以从对 n 的偏离就可以很容易看出半经典描述对该组疤痕态的适用性. 对于复杂光滑势场中的疤痕态, 其横截方向量子数一般为 0, 在能量变化范围不大的时候, 横截方向量子数 n 和能量之间接近线性关系, 因此从 n 的差异也可以推得能量的差异.

纵向的 bouncing ball 态可以做类似的分析, 只需把 ω_x 和 ω_y 互换即可. 图 4 做了这样的分析, 可以看出, 不管对于横向还是纵向的 bouncing ball 态, 当能量较小时, 比如 $E/t < 0.4$, 两种方式得到的量子数符合得非常好. 能量较大时会有所偏离, 而且能量越大偏离越多. 这是由于 ∇^2 算符的有

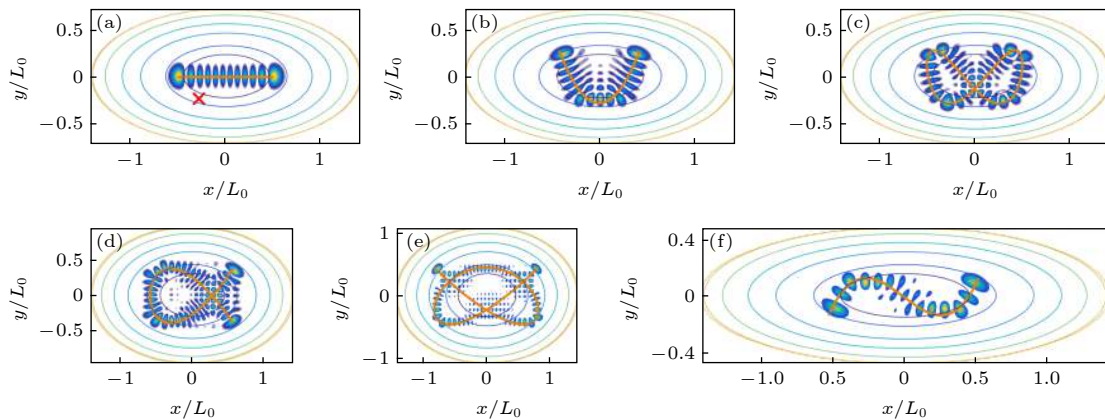


图 3 图 1(a) 势场下粒子的本征态. 图中所画为一些具有代表性的波函数的模方, 凝聚在李萨如轨道上. x 和 y 方向的频率比为 (a)–(c) 1:2, (d) 2:3, (e) 3:4, (f) 1:3. 对所有情况, $\omega'_x L_0/\sqrt{2}a = 1$, x/L_0 的范围为 $[-\sqrt{2}, \sqrt{2}]L_0$, 谐振子势在 $y = 0$ 边界上的值为 $1t$

Fig. 3. The representative eigen-wavefunctions of the billiard Fig. 1(a). Shown are the the square of the modulus of wavefunctions that are condensed on the Lissajous orbits. The ratio of the frequency in x and y directions are: (a)–(c) 1:2, (d) 2:3, (e) 3:4, (f) 1:3. For all case, $\omega'_x L_0/\sqrt{2}a = 1$, the range of x is $[-\sqrt{2}, \sqrt{2}]L_0$ and the value of the harmonic potential at the $y = 0$ boundary is $1t$.

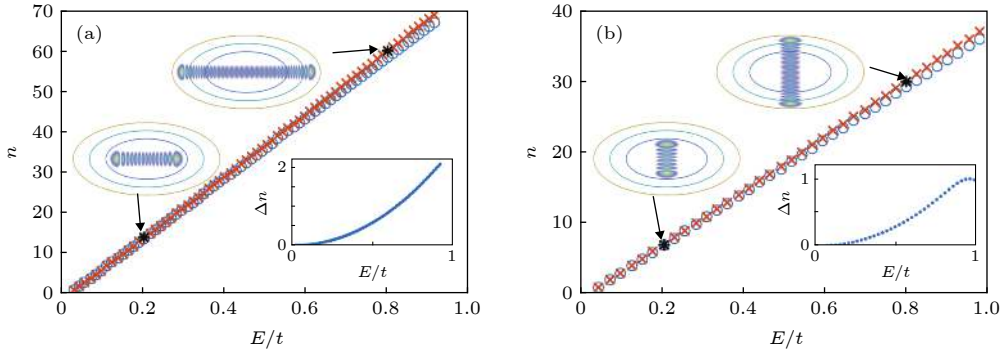


图 4 小扰动下二维谐振子势中的 bouncing ball 量子态的轨道方向量子数 n 对能量的依赖图。叉号为根据波函数数出来的波长数减 1, 圆圈为根据半经典公式得到的 n (a) 横向 bouncing ball 态; (b) 纵向 bouncing ball 态. 小图 Δn 为根据波函数数出来的结果和根据本征能量计算出来的结果的差

Fig. 4. The quantum numbers n along the trajectory vs. energy for bouncing ball states in the harmonic potential with a small perturbation. Crosses are the numbers of wavelenghts counted from the wavefunctions minus one, circles are derived from the semi-classical formulas: (a) Horizontal bouncing ball orbits; (b) vertical bouncing ball orbits. Insets show the difference Δn between these two methods.

限差近似导致的, 同时能量 E/t 在接近 1 时, 区域边界的硬墙势也会带来一定的影响.

对于图 3(b)—图 3(f) 中的轨道, 虽然比较复杂, 但仍然是一维轨道, 可以分为沿轨道方向和横截方向来处理. 横截方向在小振幅扰动下近似为谐振子, 对能量的贡献为 $E_o = (m + 1/2)\hbar\omega_o$, 其中 m 与横截方向的波包数有关, 对于图中所示各例, m 均为 0. 注意对于这些复杂的轨道, 与 bouncing ball 轨道不同, ω_o 的值是不知道的, 但对于能量相差不大的凝聚在同类型轨道上的态, 可以假设这个值是近似相等的, 所以是出现在模型里的一个待定系数, 这样可以通过联立不同的但凝聚在同类型轨道上的态把它确定下来. 沿轨道方向, 在图 3(b), 图 3(d)—图 3(f) 中有两个端点, 而且在端点处势场对轨道参数的依赖关系为线性, 因此在两个端点处会有额外的 $(1/2)2\pi\hbar$ 贡献, 半经典量子化条件为 $S_t = (n + 1/2)2\pi\hbar$, n 为沿轨道一周的波长数减 1. 相应态的总能量即为 $E = E_t + E_o$. 对于图 3(c) 中的轨道, 由于是自身封闭的轨道, 没有端点, 故 $S_t = (n + 1)2\pi\hbar$, n 为沿轨道一周的波长数减 1. 对于沿轨道方向作用量的计算, 考虑受轻微扰动的二维谐振子, 作用量对能量的依赖关系略微复杂, 依赖于具体的轨道类型, 但仍然可以依据二维谐振子的情况进行计算, 有

$$S_t = \oint \mathbf{p} \cdot d\mathbf{q} = \oint (p_x \cdot dx + p_y \cdot dy).$$

由于谐振子 x 方向的运动与 y 方向的运动解耦, 上述积分相对于对 x 及 y 方向两个独立的一维谐振子

进行积分, 有 $S_t = n_x 2\pi E_x / \omega_x + n_y 2\pi E_y / \omega_y$, 其中 n_x (n_y) 为沿该轨道运行一周, 粒子在 x (y) 方向来回振荡的次数, E_x (E_y) 为粒子沿轨道运动的能量在 x (y) 方向的分量, $E_x + E_y = E_t$.

对于图 3(b) 中的轨道, 有 $n_x = 1$, $n_y = 2$, 同时有 $\omega_y = 2\omega_x$, 因此 $S_t = 2\pi E_x / \omega_x + 2 \times 2\pi E_y / \omega_y = 2\pi E_x / \omega_x + 2\pi E_y / \omega_x = 2\pi E_t / \omega_x$. 由于该轨道有两个端点, 其沿轨道方向的量子化条件为 $S_t = (n + 1/2)2\pi\hbar$ 或 $(n + 1/2)\hbar\omega_x = E_t$. 这意味着如果把沿轨道的运动近似为一维谐振子 (一般来说不是), 其平均频率与 x 方向谐振运动的频率居然一样. 但是这一结果又可以理解, 因为从经典运动来讲, 沿该轨道运动的周期和频率确实是和 x 方向谐振运动的周期和频率是一致的. 图 5(a) 展示了对这一类李萨如态的分析结果. 通过拟合, 得到 $E_o = \hbar\omega_o/2 = 0.026t$. 可以看到, 与图 4 中 bouncing ball 轨道态的情况对比, 半经典量子化条件对于这个轨道符合得非常好, 而且 n 越小符合得越好. 我们找到的最低的能级对应的 n 为 6. 注意这里 n 大时的偏离应该归因于对 ∇^2 算符的有限差分近似, 而不是半经典方法的问题. 这个例子展示了对于这些复杂的轨道, 半经典处理在能量很小时也能得到很好的结果.

图 5(b) 展示了对另一类李萨如态的分析结果. 对应的这个轨道是一个封闭的轨道, 没有端点, 沿轨道方向的量子化条件为 $S_t = (n + 1)2\pi\hbar$, n 为沿轨道一周的波长数减 1. 这一轨道与图 5(a) 类似, 有 $n_x = 1$, $n_y = 2$, 同时有 $\omega_y = 2\omega_x$, 因此 $S_t =$

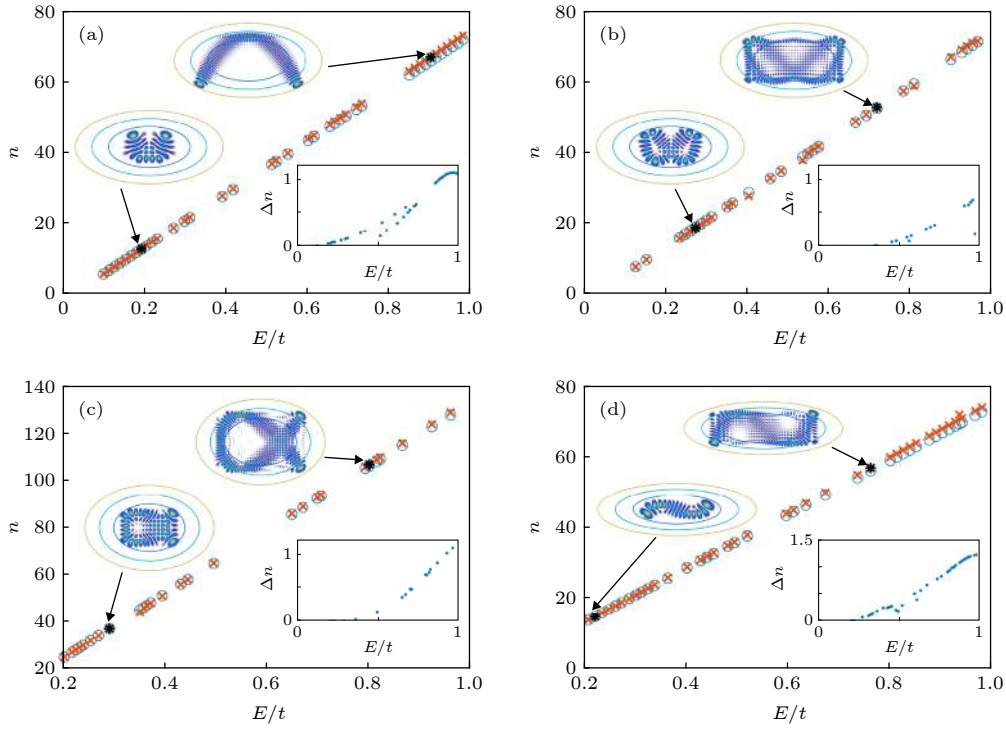


图 5 小扰动二维谐振子势场中 4 类李萨如态的量子化条件. 叉号为根据波函数数出来的波长数减 1, 圆圈为根据半经典公式得到的 n . 图中小图为根据波函数数出来的结果和根据本征能量计算出来的结果的差

Fig. 5. The quantization condition for the four types of scars for the harmonic potential with a small perturbation. Crosses are the numbers of wavelengths counted from the wavefunctions minus one, circles are the quantum numbers derived from the semiclassical formulas. Insets show the difference Δn between these two methods.

$2\pi E_t/\omega_x$. 对于该组数据的拟合, 得出 $E_o = \hbar\omega_o/2 = 0.0275t$. 对于这组轨道, 除了个别偏差较大的态, 半经典分析给出的结果与从波函数中数出的量子数 n 符合得很好, 并且由于该轨道在边界上没有折返点 (端点), 受最外层硬墙边界的影响较小, 与其他有端点的轨道相比, 当能量靠近边界能量 ($1t$) 时仍符合得较好.

图 5(c) 和图 5(d) 展示了另外两类有端点的李萨如态的分析结果, 基本与图 5(a) 相似. 略微有些差异的地方是对于图 5(c) 中的轨道, $n_x = 2$, $n_y = 3$, 有 $S_t = 2 \times 2\pi E_t/\omega_x$. 如果仍然把粒子沿周期轨道的运动套用谐振子的公式并假设平均频率为 ω_t , 则 $S_t = 2\pi E_t/\omega_t$, 有 $\omega_t = \omega_x/2$. 这一结果与经典粒子沿该轨道运动的情况一致, 不难看出, 当粒子运动一周时, 在 x 方向往返了 2 次, 与前面的 $n_x = 2$ 也一致.

3.2 大扰动下复杂势场系统

对于具有较大扰动的谐振子, 3.1 节中联系作用量与能量之间的公式不能继续使用, 作用量只能通过沿轨道的积分得到. 这时沿轨道的作用量 S_t 和

沿轨道运动的能量 E_t 仍然具有一个确定的关系, 这个关系不能显式写出, 只能数值得到, 而且对于不同类型的轨道会有较大的差别. 并且, 由于势函数复杂的形式, 系统动力学对能量的依赖关系非常敏感. 首先即使对于同种类型的轨道, 随着能量的改变, 其轨道的形状也可能发生改变. 图 6 展示了在右侧势谷中的两组 bouncing ball 轨道. 当能量较低时, 轨道接近直线, 而随着能量的增加, 其形状在势场的约束下发生了弯折, 如图 6 中的第一类 (C1) 轨道. 其次, 有些轨道可能形状变化不大, 但是随着能量的改变, 可能会出现从稳定到不稳定或者从不稳定到稳定的转变, 比如图 6 中的第二类 (C2) 轨道. 这组轨道在能量 E 小于 $0.35t$ 时稳定, 在庞加莱截面图上对应着明显的 KAM 岛, 超过 $0.35t$ 后该 KAM 岛分裂成上下两个小岛, 如图 2(b) 中 $\theta/(2\pi)$ 大约在 0.3 和 0.8 处的 KAM 岛所示, 而中心点为上下两个小 KAM 岛的交点, 是一个不稳定点. 而在能量超过 $0.6t$ 以后, 该轨道又成为稳定的, 对应着庞加莱截面上新出现的一个 KAM 岛, 如图 2(c) 中最右侧的 KAM 岛所示. 而第一类 (C1) 轨道在 $0.35t$ 时也失稳, 所对应的 KAM 岛分成了

左右两个小岛, 类似倍周期分叉, 但之后又成为稳定轨道, 一直到 $0.6t$ 附近重新失稳. 还有其他一些轨道, 在低能时没有, 只在能量比较高时才出现, 如后面的第三类 (C3) 轨道, 就是由 C2 分化而来的, 其所对应的不稳定轨道即为 C4. C3 和 C4 随着能量的增加一起演化, 逐渐偏离原来的 C2 轨道. 当能量为 $0.7t$ 时, 一个新的与原来 C2 对应的稳定轨道出现, 虽然所对应的 KAM 岛比较小, 但后续一直存在. 而 C4 这一组不稳定轨道, 只有当能量非常高时, 比如 $0.95t$, 才又变为稳定轨道了. 还有一些轨道, 比如连接左右两个势谷的轨道, 只有当能量高过他们之间的鞍点时 ($0.591t$) 才会出现, 如后面处理的 C5 和 C6 两组轨道. C5 对应有一个小的 KAM 岛. C6 轨道非常敏感, 只有少量参数值的时候轨道才稳定, 有一个很小的 KAM 岛 (在图中已经看不出来). 这些分析是通过考察确定这些轨道在庞加莱截面的位置以及庞加莱截面的结构随能量的变化而得出的. 对这些经典轨道有所了解之后, 我们就可以考察量子态在这些轨道上的凝聚及其量子化条件.

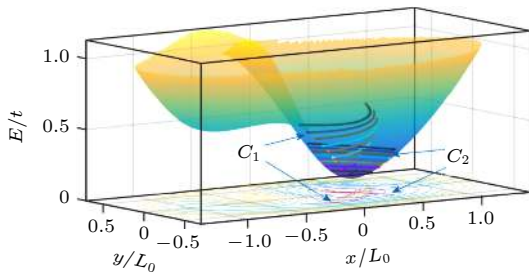


图 6 大扰动二维谐振子势场中的两类 bouncing ball 轨道以及它们在零能量面上的投影. 纵轴对应的是每个轨道的能量值. 为了便于辨认, 势函数及其等势线也一起画在了图上. 第一组 (C1) 轨道对应着图 2 中在 $p_{//} = 0$, $\theta/(2\pi) \approx 0.1$ 和 0.6 处两个最显著的 KAM 岛的中心轨道, 第二组 (C2) 轨道对应着图 2(a) 中在 $p_{//} = 0$, $\theta/(2\pi) \approx 0.33$ 和 0.86 处两个 KAM 岛的中心轨道

Fig. 6. Two types of bouncing ball orbits in the potential shown in Fig. 1(b) and their projections on the zero energy surface. The potential function and its equipotential lines are also plotted. The first class of orbits (C1) corresponds to the center point of the two most significant KAM islands for $p_{//} = 0$, $\theta/(2\pi) \approx 0.1$ and 0.6 in Fig. 2, and (C2) corresponds to the center points of the KAM islands for $p_{//} = 0$, $\theta/(2\pi) \approx 0.33$ and 0.86 in Fig. 2(a).

对于这些轨道, 由于经典的能量和作用量没有确定的函数关系, 只能通过数值积分计算. 另一方面, 由于很多轨道在给定的能量值时由于 KAM 岛太小或本身不稳定就没有 KAM 岛, 并不容易得

到, 这使得对于所有给定的疤痕态都根据其 E_t 来直接寻找经典轨道并积分求得作用量变得异常困难. 在此, 我们采取一个迂回的方法. 首先, 根据量子系统的疤痕态所凝聚的轨道以及庞加莱截面上 KAM 岛的信息, 在不同能量 E 下分别找到相应的轨道并计算其作用量 S , 对于每一类轨道, 我们在所涉及到的区间得到 10 个左右的数据点. 对于谐振子 S 正比于 E , 这里由于势场的不规则性以及轨道形状的变化, S 对 E 的依赖关系接近线性, 会有一些偏离. 利用这些数据点做二次多项式拟合, 得到拟合函数 $S = f_C(E)$, 下标 C 表示这个函数是依赖于轨道的. 这样, 对于凝聚在某一类轨道 C 上的疤痕态及其本征能量 $E_{n,m}$, 其中 n, m 分别为轨道方向和横截方向的量子数, 有 $E_{n,m} = E_t(n) + E_o(m) + V_{\min}$, 其中 $E_o(m) = (m + 1/2)\hbar\omega_o$, ω_o 为经典轨道横截模式的振荡频率, $V_{\min} = 0.1053t$ 为该势场的最低值. 这样 $E_t = E_{n,m} - E_o(m) - V_{\min}$, 其经典作用量可以通过 $S_t = f_C(E_t)$ 来得到. 由经典作用量, 可以通过 $S_t = (n + 1/2)2\pi\hbar$ 得到相应的量子数 n , 然后与从波函数图形上数出来的波长数减 1 相比, 来检验半经典量子化条件的适用性.

对于上一节中小扰动下的二维谐振子, 同一类轨道对应的这个频率近似为常数, 与能量无关. 这里将看到, 虽然对于复杂势场下同一类经典轨道可能存在重要差异, 但是其 ω_o 仍然近似保持恒定, 对能量依赖关系较小, 因此仍然能够近似成一个常数, 作为待定参数来处理. 由于横截方向的运动可以近似为简谐振子, ω_o 与横截方向波函数的宽度的平方具有反比关系, 从我们所考察的能量区间不同波函数横截方向宽度基本一致也可以佐证 ω_o 近似为一常数. 而 E_o 可以通过调整其值使得两者符合得最好来得到.

首先考察 C1 和 C2 这两类 bouncing ball 轨道. 这两类轨道是能量较低时最稳定的轨道, 具有最大的 KAM 岛. 因此, 它们所对应的疤痕态非常规律, 并且除了常见的 $m = 0$ 的态, 还有很好的 $m = 1$ 的态, 这样 $\hbar\omega_o$ 除了可以作为拟合参数得到, 还能通过两组本征能量 $E_{n,1}$ 和 $E_{n,0}$ 相减直接得到. 图 7 内的小图展示了这个结果. 我们发现, 对于 C1 和 C2 两类轨道, 两种方法得到的 ω_o 值一致, 并且在整个能量范围, ω_o 的变化幅度不超过 10%, 而且这部分能量本身就小, 因此 ω_o 可以作为常数来处理.

图 7 展示了 C1 和 C2 这两类 bouncing ball

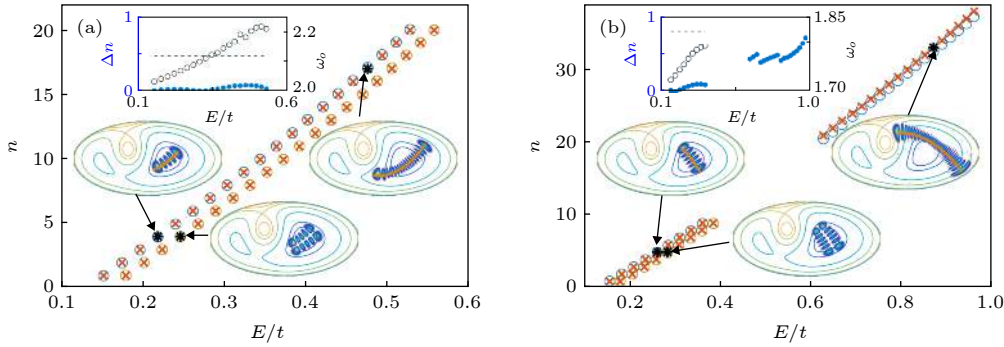


图 7 沿轨道方向量子数 n 与能量的依赖关系. 叉号为根据波函数数出来的波长数减 1, 圆圈为根据半经典公式得到的 n (a) 第一类 bouncing ball 轨道 (C1); (b) 第二类 bouncing ball 轨道 (C2). 每个图中横截量子数 $m = 0$ 为上面那组点, $m = 1$ 的为下面那组点. 对于 C2 轨道, 只有能量较低的时候有 $m = 1$ 的量子态, 能量较高时在计算中没有发现 $m = 1$ 的态. 图中小图展示了一些标准的疤痕态及其对应的经典轨道, 两种 n 的差值 (蓝色实心圆, 左侧坐标) 以及由 $E_{n,1} - E_{n,0}$ 计算出的 ω_0 值 (黑色空心圆, 右侧坐标), 其中横虚线为拟合得到的 ω_0 值 ($\omega_0 L_0 / (\sqrt{2}a)$), 对应的 $E_0 = \hbar\omega_0/2$ 分别为 $0.0141t$ 和 $0.0121t$

Fig. 7. The quantum numbers n along the trajectory vs. energy for bouncing ball states in the modified harmonic potential shown in Fig. 1(b) for C1 orbits (a) and C2 orbits (b). Crosses are the numbers of wavelenghtes counted from the wavefunctions minus one, circles are derived from the semiclassical formulas. In each panel, the upper set of points are for $m = 0$, and the lower set of points are for $m = 1$. For C2 orbits, only when energy is small there are $m = 1$ states. Insets show the difference Δn (solid circles, left coordinates) between these two methods, and ω_0 obtained from $E_{n,1} - E_{n,0}$ (empty circles, right coordinates), where the horizontal dashed line is the ω_0 obtained from fitting to the data, and the corresponding energies $E_0 = \hbar\omega_0/2$ are $0.0141t$ and $0.0121t$ for C1 and C2 orbits, respectively.

轨道半经典估计的沿轨道方向的量子数和从波函数直接数出来的量子数. 对于 C1 轨道两者符合得较好, 对于 C2 轨道, 在能量低时符合得较好, 在能量高时有系统性的偏差. 这可能是由于对 ω_0 的估计的问题, 因为这里我们用了统一的 ω_0 值, 但是实际上低能轨道和高能轨道的差异还是很大的.

图 8 展示了其他轨道的情况. 对于 C3, C4, C6 轨道, 半经典方法得到的量子数与从波函数中数出来的量子数符合得比较好, 而 C5 轨道两者符合得要差一些, 这与 C5 轨道对应的量子疤痕态比较模糊不易分辨有关. 这里注意, C4 在多数能量值下为不稳定轨道, 所以这里 E_0 并不对应着横截模式的振荡频率 ω_0 , 而是与轨道稳定性有关的一个参数^[14,33]. 此外, 当能量大于 $0.95t$ 时, 该轨道已经稳定, 但是对应的 KAM 岛很小. 此时疤痕态仍然满足相同的半经典公式, 并没有因为经典轨道稳定性发生变化而出现显著差别.

这些结果显示, 在这种复杂的势场下, 半经典量子化仍然能够对疤痕态对应的能级给出比较好的描述, 并且在低能下也符合得很好. 这是由于, 虽然势场本身及其经典动力学很复杂, 但是在势谷附近, 系统仍然具有两个较好的轨道方向, 使得系统沿每组轨道的运动近似为稳定的简谐运动, 在庞加莱截面上表现为四个较大的 KAM 岛 (每个稳定 bouncing ball 轨道有两个 KAM 岛), 因此量子化

条件能够比较好地适用. 对于我们的情况, 轨道 C1 和轨道 C2 在低能时基本互相垂直, 因此 C1 的 ω_t 应为 C2 的 ω_0 , C1 的 ω_0 应为 C2 的 ω_t . 图 7 展示 ω_0 , 在低能时 $\omega_0 L_0 / (\sqrt{2}a)$ 分别接近 2.03 和 1.72, 而由图中数据按 $S_t = 2\pi E_t / \omega_t$ 估计得到的 $\omega_t L_0 / (\sqrt{2}a)$ 在低能时分别为 1.69 和 2.00, 与上述分析一致. 当能量较高系统动力学较复杂时, 对应的 KAM 岛变小或消失, 部分轨道从稳定变到不稳定, 或者出现新的稳定轨道, 这时虽然每个轨道比较复杂, 但在一个比较小的能量区间, 轨道性质变化不大, 量子化条件对于这组轨道所对应的疤痕态仍能给出较好的描述. 但如图 7(b) 所示, 如果轨道已经存在一定的差异, 比如 KAM 岛的显著变化 (低能时 KAM 岛很大, 高能时很小), 那么 ω_0 应该分段取值. 可以预期, 对于图 7(b) 中能量较高的态, 重新标度 ω_0 将能够使得半经典得到的量子数整体上移, 与从波函数数出来的 n 符合得更好.

关于疤痕态重复出现的规律, 鉴于低能时在势谷中沿轨道方向的运动近似为谐振子运动, 这些疤痕态重复出现的规律与一维谐振子类似, 在能量上等间距出现, 或 n 与 E 是线性关系. 在数值上, 这一规律对于我们所处理的小扰动下二维谐振子的各组轨道 (图 4 和图 5) 以及复杂势场中的各组轨道 (图 7 和图 8) 均成立. 当然这一特征的根源是作用量与能量之间近似线性的关系. 此外, 并不需要作

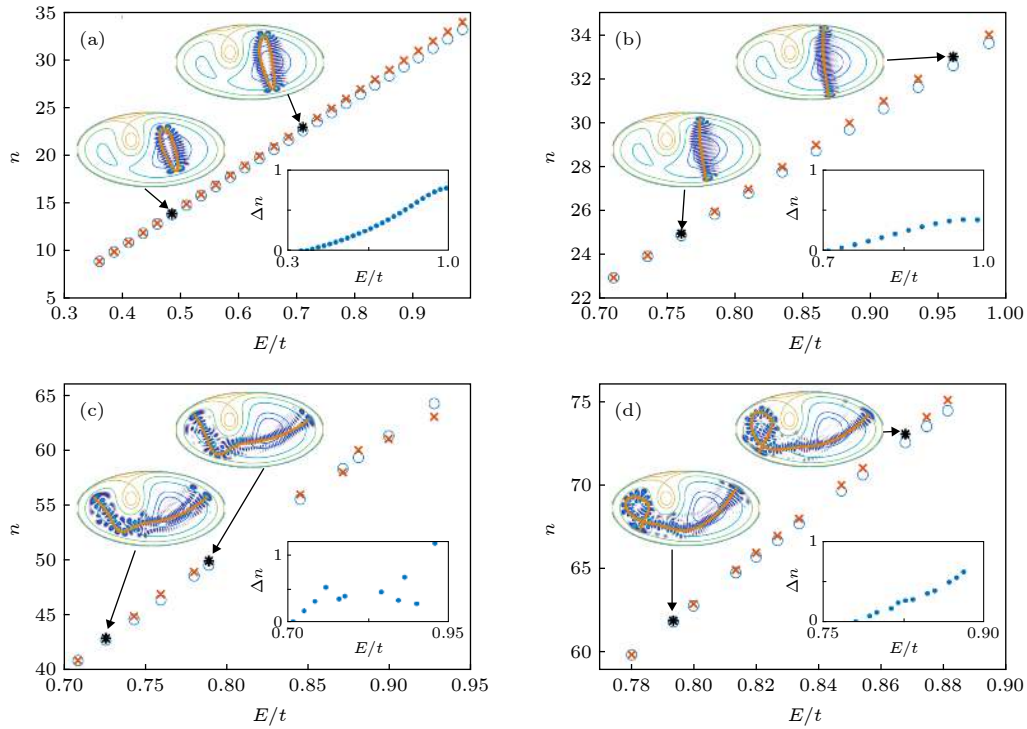


图 8 沿轨道方向量子数 n 与能量的依赖关系. 叉号为根据波函数数出来的波波长数减 1, 圆圈为根据半经典公式得到的 n . 这里所有态的横截量子数 m 均为 0. (a)–(d) 分别对应第三类 (C3)、第四类 (C4)、第五类 (C5)、第六类 (C6) 轨道, 对应的 E_o 分别为 $0.0220t$, $0.0013t$, $0.0280t$, 和 $0.0193t$. 图中小图展示了一些标准的疤痕态及其对应的经典轨道, 以及两种 n 的差值. C3 只在能量为 $0.35t$ 时 C2 失稳后才出现. C4 为 C2 的另外一支不稳定轨道, 只在能量超过 $0.95t$ 后才稳定. C5 和 C6 是连接两个势谷的轨道, 只在较高能级时出现

Fig. 8. The quantum numbers n along the trajectory vs energy. $m = 0$ for all cases. Crosses are the numbers of wavelenghtes counted from the wavefunctions minus one, circles are derived from the semiclassical formulas. (a)–(d) correspond to C3–C6 orbits, with $E_o = 0.0220t$, $0.0013t$, $0.0280t$ and $0.0193t$, respectively. Insets show some typical scarring states and the corresponding classical orbits, and the difference Δn between these two methods. Note that C3 orbits only appear for $E > 0.35t$ when C2 becomes unstable. C4 is the other unstable branch of C2, and becomes stable only for $E > 0.95t$. C5 and C6 are orbits connecting the two potential valleys, only appear when higher energy is high enough.

用量对能量有全局的线性关系, 由于疤痕态重复出现的规律只与作用量的改变量有关, 即 $\Delta S/h = 2\pi$ 或 $\Delta S = h$, 因此只要在一个能量区间在某一类周期轨道上的作用量线性依赖于能量, 那么在这个能量区间, 这一轨道上的疤痕态就会在能量上等间距出现. 注意由于作用量对能量的函数关系取决于轨道 (势场沿该轨道的 profile), 对于不同类型的轨道哪怕在相同的能量区间, 函数关系也可能不同. 所以在一类轨道上的疤痕态在能量上等间距出现并不会意味着其他轨道也是如此. 但是对于一般的复杂势场的低能情况, 当粒子局限在某个势谷中时, 由于沿轨道方向的运动近似为谐振子运动, 一般会有 n 与 E 的线性关系.

随机杂质系统. 当系统足够随机, 大部分态只提供杂乱的背景, 那么这组凝聚在经典轨道上的比较强的局域态将具有一定的主导性, 在态密度的谱中能够体现出来. 在分析其重复出现的规律时, 由

于是非相对论量子系统, 一般会假定按照 \sqrt{E} 的方式, 我们此次的结果表明, 在复杂势场下对于疤痕态重复出现的规律, 要根据其经典轨道做相应的分析, 找出作用量对能量的依赖关系, 才能判断疤痕态按照 E 出现的方式.

4 结 论

半经典量子化公式具有广泛的应用, 被人们用来分析量子系统中各种与经典轨道有关的物理量, 特别是在解释量子疤痕态及其出现的量子化条件时起到了重要的作用. 我们研究了软墙量子弹球以及复杂势场下的量子弹球, 考察其中疤痕态出现的量子化条件及其重复出现的规律, 发现半经典方法在处理复杂势场下的量子弹球时对疤痕态出现的量子化条件能够给出较好的预测, 而且对于同一类轨道, 甚至当轨道的稳定性发生改变时, 半经典预

测仍能与数据符合得较好. 此外, 一般认为半经典量子化公式对于较高能级适用得较好, 对于低能级符合得较差. 但是由于半经典量子化公式能够精确描述谐振子, 而我们在大扰动光滑势场中的系统虽然整体动力学结构非常复杂, 但在低能下当粒子主要约束在一个势谷中时, 经典庞加莱截面上有非常大的 KAM 岛, 粒子在 KAM 岛中心点所对应的轨道上运动时近似为一维简谐振动, 因此半经典公式能够较好地描述相应疤痕态的量子化条件. 当能量较高, 经典动力学更加复杂, 甚至经典粒子能够在两个势谷中穿梭时, 我们的结果显示半经典公式仍然能够较好地给出疤痕态的量子化条件.

关于疤痕态重复出现的规律, 在低能时, 由于粒子沿轨道运动类似一维谐振子, 疤痕态按照在能量 E 轴上等间距的方式出现, 并且对于我们所处理的系统, 在较高能级甚至轨道不稳定时仍是如此. 对比之前非相对论性硬墙量子弹球系统疤痕态按照 $\sqrt{E - E_0}$ 等间距出现的结论, 我们的结果是一个有益的补充, 使得人们在处理软墙系统时必须重新审视这些规律. 虽然我们期待这个结果对于无规的较复杂的势场中的量子系统也适用, 但是必须指出, 最终疤痕态在能量轴上重复出现的规律取决于作用量与能量的关系. 因为疤痕态在作用量轴上等间距重复出现, 因此对于一些特殊情况, 必须要分析在所研究的轨道上作用量与能量之间的关系. 在我们的系统中, 对于大多数轨道, 作用量均近似线性依赖于能量, 导致疤痕态在能量轴上等间距出现的结果.

由于疤痕态对应着较强的局域化的态, 尤其是稳定轨道上的局域态在开放系统中仍能存在^[34], 成为影响系统态密度及输运性质的重要因素^[35], 这些疤痕态重复出现的规律也会反映在态密度谱峰和输运特性上. 我们期待这些结果对于理解普遍存在的无规长程杂质下的二维电子的态密度及输运性质能够提供一定的帮助, 一方面可以解释实验观察到的现象, 另一方面也可以反推出势谷的信息, 有助于理解杂质的特性.

感谢北京大学汪知昌博士、江颖教授分享的未发表实验数据及讨论, 引起了我们对这一理论问题的研究.

参考文献

- [1] Keller B J 1958 *Ann. Phys.* **4** 180
 [2] Einstein A 1917 *Verh. Dtsch. Phys. Ges.* **19** 82

- [3] Stone A D 2005 *Phys. Today* **58** 37
 [4] Gutzwiller M C 1971 *J. Math. Phys.* **12** 343
 [5] Cvitanovic P, Artuso R, Mainieri R, Tanner G, Vattay G, Whelan N, Wirzba A 2005 *Chaos: Classical and Quantum* (Copenhagen: Niels Bohr Institute) pp143–145
 [6] Lichtenberg A J, Leiberman M A 1992 *Regular and Chaotic Dynamics* 2nd edition (New York: Springer-Verlag) pp7–60
 [7] Ott E 2002 *Chaos in Dynamical Systems* (2nd Ed.) (Cambridge: Cambridge University Press) pp421–450
 [8] Knauf A, Sinai Y G 1997 *Classical Nonintegrability, Quantum Chaos* (Birkhuaser: Springer-Verlag) pp41–47
 [9] Berry M V 1989 *Phys. Scr.* **40** 335
 [10] Stöckmann H J 2006 *Quantum Chaos: An Introduction* (New York: Cambridge University Press) pp296–338
 [11] Haake F 2010 *Quantum Signatures of Chaos* (3rd Ed.) (Berlin: Springer-Verlag) pp62–71
 [12] Gutzwiller M C 2013 *Chaos in Classical and Quantum Mechanics* (New York: Springer-Verlag) pp116–118
 [13] Xu G O 1995 *Quantum Chaotic Motions in Quantum Systems* (Shanghai: Shanghai Scientific and Technical Publishers) pp58–69 (in Chinese) [徐躬耦 1995 量子混沌运动 (上海: 上海科学技术出版社) 第58–69页]
 [14] Gu Y 1996 *Quantum Chaos* (Shanghai: Shanghai Scientific and Technological Education Publishing House) pp69–153 (in Chinese) [顾雁 1996 量子混沌 (上海: 上海科学技术出版社) 第69–153页]
 [15] Casati G and Chirikov B 2006 *Quantum Chaos: Between Order and Disorder* (New York: Cambridge University Press) pp317–385
 [16] McDonald S W, Kaufman A N 1979 *Phys. Rev. Lett.* **42** 1189
 [17] Heller E J 1984 *Phys. Rev. Lett.* **53** 1515
 [18] McDonald S W, Kaufman A N 1988 *Phys. Rev. A* **37** 3067
 [19] Bogomolny E B 1988 *Physica D* **31** 169
 [20] Berry M V 1989 *Proc. R. Soc. London, Ser. A* **423** 219
 [21] Agam O, Fishman S 1993 *J. Phys. A Math. Gen.* **26** 2113
 [22] Agam O, Fishman S 1994 *Phys. Rev. Lett.* **73** 806
 [23] Kroetz T, Oliveira H A, Portela J S E, Viana R L 2016 *Phys. Rev. E* **94** 022218
 [24] Luukko P J J, Drury B, Klaes A, Kaplan L, Heller E J, Räsänen E 2016 *Sci. Rep.* **6** 37656
 [25] Keski-Rahkonen J, Luukko P J J, Kaplan L, Heller E J, Räsänen E 2017 *Phys. Rev. B* **96** 094204
 [26] Keski-Rahkonen J, Ruhanen A, Heller E J, Räsänen E 2019 *Phys. Rev. Lett.* **123** 214101
 [27] Keski-Rahkonen J, Luukko P J J, Åberg S, Räsänen E 2019 *J. Phys. Condens. Matter* **31** 105301
 [28] Eckhardt B 1988 *Phys. Rep.* **163** 205
 [29] Huang L, Lai Y C, Ferry D K, Goodnick S M, Akis R 2009 *Phys. Rev. Lett.* **103** 054101
 [30] Xu H Y, Huang L, Lai Y C, Grebogi C 2013 *Phys. Rev. Lett.* **110** 064102
 [31] Arnold V I 2013 *Mathematical Methods of Classical Mechanics* (New York: Springer Science & Business Media) pp30–50
 [32] Miller W H 1975 *J. Chem. Phys.* **63** 996
 [33] Voros A 1988 *J. Phys. A: Math. Gen.* **21** 685
 [34] Huang L, Lai Y C, Luo H G, Grebogi C 2015 *AIP Adv.* **5** 017137
 [35] Zhang G Q, Chen X, Lin L, Peng H, Liu Z, Huang L, Kang N, Xu H Q 2020 *Phys. Rev. B* **101** 085404

SPECIAL TOPIC—Statistical physics and complex systems

Quantization condition of scarring states in complex soft-wall quantum billiards*

Li Xiao-Liang Chen Xian-Zhang Liu Chen-Rong Huang Liang[†]*(School of Physical Science and Technology, Lanzhou University, Lanzhou 730000, China)*

(Received 11 March 2020; revised manuscript received 3 April 2020)

Abstract

Quantum scar is an intriguing phenomenon in quantum or wave dynamics that the wavefunction takes an exceptionally large value around an unstable periodic orbit. It has attracted much attention and advances the understanding of the semiclassical quantization. Most of previous researches involving quantum scars focus on hard-wall quantum billiards. Here we investigate the quantum billiard with a smooth confinement potential which possesses complex classical dynamics. We demonstrate that the semiclassical quantization approach works well for both the stable and unstable classical periodic orbit, besides the fact that the shape of the orbits varies as the energy increases or even the stability switches. The recurrence rule of the quantum scars in this complex soft-wall billiard differs from that of the hard-wall nonrelativistic quantum billiard, such as being equally spaced in energy instead of being equally spaced in the square root of energy. These results implement the previous knowledge and may be used for understanding the measurements of density of states and transport properties in two-dimensional electron systems with random long-range impurities.

Keywords: quantum scar, soft-wall quantum billiard, complex smooth potential quantum billiard, quantization rule

PACS: 05.45.Mt, 03.65.Sq

DOI: [10.7498/aps.69.20200360](https://doi.org/10.7498/aps.69.20200360)

* Project supported by the National Natural Science Foundation of China (Grant No. 11775101).

[†] Corresponding author. E-mail: huangl@lzu.edu.cn

自驱动杆状粒子在半柔性弹性环中的集体行为

仲颖 施夏清

Collective behaviors of self-propelled rods under semi-flexible elastic confinement

Zhong Ying Shi Xia-Qing

引用信息 Citation: *Acta Physica Sinica*, 69, 080507 (2020) DOI: 10.7498/aps.69.20200561

在线阅读 View online: <https://doi.org/10.7498/aps.69.20200561>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

自驱动颗粒体系中的熵力

Entropy forces of nanoparticles in self-propelled systems

物理学报. 2017, 66(19): 190701 <https://doi.org/10.7498/aps.66.190701>

自驱动的Janus微球在具有不同障碍物的表面上的运动行为研究

The dynamics of self-propelled Janus microspheres near obstacles with different geometries

物理学报. 2019, 68(13): 134702 <https://doi.org/10.7498/aps.68.20190711>

双电层相互作用下主动粒子系统的压强

Pressure of active system under the electric double layer interaction

物理学报. 2019, 68(17): 170501 <https://doi.org/10.7498/aps.68.20190435>

柔性棘轮在活性粒子浴内的自发定向转动

Spontaneous rotation of ratchet wheel with soft boundary in active particle bath

物理学报. 2019, 68(16): 161101 <https://doi.org/10.7498/aps.68.20190425>

基于半解析自洽理论研究相对论激光脉冲驱动下阿秒X射线源的产生

Attosecond X-ray generation driven by the relativistic laser pulse based on the semi-analytical self-consistent theory

物理学报. 2017, 66(20): 204205 <https://doi.org/10.7498/aps.66.204205>

专题：统计物理和复杂系统

自驱动杆状粒子在半柔性弹性环中的集体行为*

仲颖 施夏清†

(苏州大学物理科学与技术学院, 软凝聚态物理及交叉研究中心, 苏州 215006)

(2020年4月16日收到; 2020年4月18日收到修改稿)

在生物体系的活性系统中, 杆状粒子在弹性半柔性边界中的受限行为极为常见. 本文研究了二维情况下, 自驱动杆状粒子受限在半柔性弹性环中的集体行为. 改变系统的粒子数及噪声强度, 系统显示明显的自驱吸附有序态、无序态及中间的过渡态. 通过表征弹性环内部粒子的径向极性大小和空间分布的非球度性对这些状态进行了刻画. 进一步对弹性环中心附近粒子密度的分析, 发现环中心气态粒子分布存在一个与边界高密度区域共存的饱和平台, 出现类似吸附转变的粒子分布. 在过渡区间, 体系内存在较大的涨落会导致弹性环出现异常形变. 非对称的粒子分布对弹性环整体的迁移具有重要贡献, 系统在过渡区间能获得相对较强的定向迁移.

关键词: 活性物质, 自驱动杆状粒子, 吸附相分离, 饱和密度**PACS:** 05.10.Gg, 05.20.-y, 87.16.Uv, 05.40.Jc**DOI:** 10.7498/aps.69.20200561

1 引言

活性物质通过消耗自身携带的或从环境中吸收的能量来实现各种力学运动, 是一类典型的具有多级结构的非平衡体系, 展现出丰富的集体动力学行为. 不同于平衡态, 活性物质常常具有很强的密度与取向序的耦合, 产生极为复杂的时空演化行为^[1-8]. 在一些大尺度的模型和理论研究中, 活性物质在取向上的对称性破缺, 会导致系统在空间上的密度不均匀分布, 比如自驱动粒子在周期性边界中会自发形成一些大的相分离结构^[9-12]. 除此之外, 活性物质也会自发演化出特异性的形态, 比如在细胞骨架纤维的实验中, 可以看到肌动蛋白丝和微管, 会形成一些涡旋、波形的图案分布^[13]. 这其中粒子的形状会对系统的集体行为产生较大影响, 从最简单的球形^[9,14,15], 到哑铃型双球^[16,17]、蠕虫形粒子^[18]以及杆状粒子^[11,12,19-21].

生命系统中常常存在非常明显的空间受限条件. 迁移的细胞个体内部就有复杂而有序的物质结构. 细胞膜下覆盖的肌动蛋白微丝与分子马达的动态结构对于细胞迁移具有至关重要的作用. 比如在肌动蛋白贡献细胞迁移^[22,23]的过程中, 具有活性物质特性的蛋白纤维在细胞膜附近生长并对其施加压力^[24-26], 所有过程都是在细胞膜内进行的. 所以在活性物质, 尤其是相关生命系统中, 边界束缚条件或者界面作用是个值得理论研究注意的问题.

自驱动粒子和界面的相互作用会对内部自驱动粒子的集体运动行为产生很大影响^[27]. 在类似于通道的开放的受限系统中, 自驱动粒子会在边界附近形成大量集聚, 并根据通道的具体特性形成一些对应分布^[17,28,29]. 在封闭系统内部, 自驱动粒子同样会在界面位置形成非常稳定的堆积分布^[30,31]. 目前, 对更有趣的软受限条件研究得比较少. 已有的一些研究发现, 在该条件下系统大尺度的粒子涨落会诱发柔性边界的明显形变, 反过来又会影响粒子的均匀分布^[32-34].

* 国家自然科学基金 (批准号: 11674236, 11922506) 资助的课题.

† 通信作者. E-mail: sxiaqing@gmail.com

本文研究了大量自驱动杆状粒子在柔性边界中的动力学行为. 简单的球形自驱动粒子模型, 往往不能很好地体现活性物质中个体的形状各向异性. 杆状粒子在非平衡受限时会有丰富的堆积行为^[35]. 我们的模型使用了弹性杆体系, 不仅可以实现柔性边界的弹性伸缩, 还可以在保证杆特性的前提下, 对杆的算法进行优化. 我们用弹性杆链接而成的柔性环作为约束边界, 统计了大数目自驱动杆状粒子受限在该柔性环中的分布情况和整体运动. 在该柔性环提供的弹性约束下, 其内部的自驱动杆状粒子同样也倾向于在环边界上聚集分布, 但在中心位置会保持一个相对稳定的气态密度分布. 大部分情况下这些自驱动杆状粒子都具有明显的对称性分布, 但在有序无序转变区域, 粒子大尺度的涨落与膜的大尺度形变两者相互影响, 会形成各向异性的粒子分布, 同时伴随较强的整体定向迁移. 可以看到, 这部分系统整体的运动以及形变, 一定条件下发生明显极化, 并伴随较大尺度的迁移, 与一些细胞迁移运动行为类似, 因而该系统的研究对调控这类集体运动状态行为有一定的参考作用.

2 模拟模型和方法

模拟一个二维平面下的系统, 如图 1(a) 所示, 该系统由两部分组成: 其一是 N_r 个原长为 L_r 的自驱动杆状粒子; 另一部分是一条柔性环, 由 N_l 个原长为 L_l 的杆首尾链接而成, 柔性环将所有自驱动杆状粒子约束在其内部.

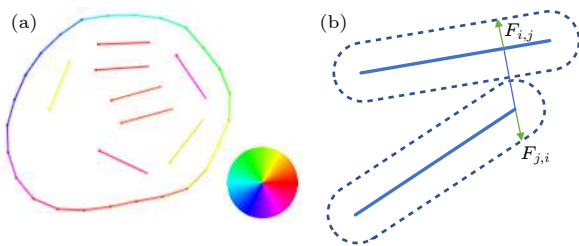


图 1 (a) 系统组成的示意图, 颜色代表杆身的取向; (b) 杆间碰撞受力示意图

Fig. 1. (a) The schematic diagram of this system, and the rods are colored according to their angle with respect to the radial direction; (b) the interaction between rods.

在该模型中的自驱动杆和环的结构单元都是采用相同构造的杆, 这种杆沿其径向两端是直径等宽的半圆形结构, 同时所有的杆都具有固定的杆宽 r_0 ($r_0 = 1$), 如图 1(b) 所示. 这样设定的杆可以比

较方便地计算杆间最短距离, 进而方便判断近邻杆间的碰撞及计算相互作用.

所有的这些杆状粒子, 杆身受外力不能弯曲. 但沿杆身方向是弹性的, 且只能沿径向弹性伸缩. 如果使用固定杆长的硬杆链接而成的环, 其中近邻的杆在受到碰撞时, 需要很高精度的迭代保证最近邻杆间的运动刚性. 而杆身弹性的设定可以有效简化这方面的计算, 通过杆自身的弹性约束保持环中杆间的连续. 此外我们选取较大的杆身弹性系数 k , 对于内部的自驱动杆状粒子, 可以视为是固定杆长的硬杆.

为方便表述, 约定所有杆两端的圆心位置分别为该杆的正负两端 (正负端形状完全对称), 杆的取向 θ 则定义为由其负端到正端的指向. 每根杆的位置信息可由其质心位置 \mathbf{r}_i , 取向 θ_i 和实际的杆长 l_i 确定. 相应的杆的正负两端的位置分别为 $\mathbf{r}_i^+ = \mathbf{r}_i + \frac{l_i}{2}\boldsymbol{\theta}_i$ 和 $\mathbf{r}_i^- = \mathbf{r}_i - \frac{l_i}{2}\boldsymbol{\theta}_i$. 特别地, 对于环中的相邻杆, 规定两者相连处的异号端点位置是始终重合的, 即环上杆 i 的“+”端端点位置, 与它顺时针方向的相邻杆 $i+1$ 的“-”端端点位置重合, 满足 $\mathbf{r}_i^+ \equiv \mathbf{r}_{i+1}^-$ 的关系^[36].

杆之间的相互作用是截断的简谐弹簧势

$$\phi(r) = \begin{cases} \phi_0(1 - r/r_0)^2, & r \leq r_0, \\ 0, & r > r_0, \end{cases} \quad (1)$$

其中 r 是两杆之间的最短距离, r_0 为固定的杆宽, ϕ_0 为势能的强度大小. 通过在 $r = r_0$ 处的截断, 体系为纯弹簧排斥势. 在体积排斥效应下, 发生碰撞的杆状粒子有平行排列倾向.

体系中所有的杆侧面都是光滑的, 即碰撞时不考虑杆之间的滑动摩擦. 在这种杆的模型中, 碰撞时杆所受的排斥力垂直于杆身或作用在端点位置, 如图 1(b) 所示. 杆间的碰撞所受的相互作用可以等效到端点位置处. 实际碰撞位置的排斥受力和端点上的等效受力, 其合力和力矩存在如下的等量关系:

$$\begin{aligned} \mathbf{F}_{i,j} &= \mathbf{F}_{i,j}^+ + \mathbf{F}_{i,j}^-, \\ (\lambda - \lambda_0)l\boldsymbol{\theta} \times \mathbf{F}_{i,j} &= (1 - \lambda_0)l\boldsymbol{\theta} \times \mathbf{F}_{i,j}^+ + \lambda_0(-l\boldsymbol{\theta}) \times \mathbf{F}_{i,j}^-, \end{aligned} \quad (2)$$

其中 $\mathbf{F}_{i,j}$ 是碰撞时 i 杆身上所受的与 j 杆相互作用的排斥力, $\mathbf{F}_{i,j}^+$ 和 $\mathbf{F}_{i,j}^-$ 分别对应杆正负两端位置的等效受力, λ 和 λ_0 分别是碰撞受力位置和杆质心,

沿杆取向 θ 在杆身上的相对位置, λ 在负端取 0 正端取 1, 杆质心的相对位置为 $\lambda_0 = 0.5$. 方程 (2) 联立可解得碰撞时杆正负两端的等效受力分别为:

$$\begin{aligned} \mathbf{F}_{i,j}^+ &= \lambda \mathbf{F}_{i,j}, \\ \mathbf{F}_{i,j}^- &= (1 - \lambda) \mathbf{F}_{i,j}, \end{aligned} \quad (3)$$

直接作用在端点处的碰撞受力同样符合上面的结果.

环内的自驱动粒子除了杆间的碰撞排斥作用外, 沿着其取向 θ 还受恒定大小的自驱动力 $s\theta$ 持续牵引, 从而具有自我推进的能力.

另外由于杆身是弹性的, 杆正负两端所受弹性力分别为 $\mathbf{F}_{e_i}^+ = -\frac{1}{2}k(l_i - L_\chi)\theta_i$ 和 $\mathbf{F}_{e_i}^- = \frac{1}{2}k(l_i - L_\chi)\theta_i$, 其中 k 为弹性系数, L_χ 为内部自驱动杆或柔性环中杆对应的原长. 杆正负端点位置的运动可以通过布朗动力学方程来表述:

$$\dot{\mathbf{r}}_i^\pm = \mu \left(F_p \theta_i + \mathbf{F}_{e_i}^\pm + \sum_{j \in \Omega} \mathbf{F}_{i,j}^\pm \right) + \xi_i(t), \quad (4)$$

其中 μ 是迁移率, 这里使用各向同性的迁移率. $j \in \Omega$ 是与当前杆发生碰撞的近邻杆. $\xi_i(t)$ 是高斯白噪声, 定义噪声强度为 η . 杆两端位置上的高斯白噪声可以给杆提供一个有效的角度上的扰动.

对于柔性环上的杆, 由于不存在自驱动力其动力学方程为

$$\dot{\mathbf{r}}_i^\pm = \mu \left(\mathbf{F}_{e_i}^\pm + \sum_{j \in \Omega'} \mathbf{F}_{i,j}^\pm \right). \quad (5)$$

这里也忽略了噪声的作用. 特别地, 柔性环上的相邻两杆之间不存在空间排斥作用, 即 $j \in \Omega'$ 是除去环上相邻两杆后与当前杆发生碰撞的近邻杆. 由于柔性环上相邻两杆端点重合, 环上杆 i 正端端点的实际位移 $\Delta \mathbf{r}_i^+$, 与它顺时针方向的相邻杆 $i+1$ 的负端端点实际位移 $\Delta \mathbf{r}_{i+1}^-$ 相同, 即 $\Delta \mathbf{r}_i^+ \equiv \Delta \mathbf{r}_{i+1}^- = (\dot{\mathbf{r}}_i^+ + \dot{\mathbf{r}}_{i+1}^-) \Delta t$.

在本文模拟中, 将原长 $L_r = 2$ 自驱动杆受限于半柔性环内. 半柔性环由 $N_l = 200$, 原长为 $L_l = 1$ 的杆首尾链接而成. 将系统整体放在足够大的二维周期性边界内 ($L_x = L_y = 200$). 选取足够小的时间步长 $dt = 0.0005\tau$, 保证数值稳定性及必要的精度, 其中 $\tau = 1$ 是模拟的单位时间量程. 根据动力学方程 (4) 式和 (5) 式对粒子的位置进行更新. 经过足够长时间演化后, 可对系统所处的稳定分布进行分析. 本文主要研究系统在自驱动杆的粒子数 N_r

和噪声强度 η 构成的二维参数空间中的统计动力学行为.

3 模拟结果和讨论

封闭空间中的自驱动粒子倾向于在边界附近聚集^[30-34]. 不同于球状粒子, 本文模型由杆状粒子组成. 由图 2 中的快照可以看出, 这样的自驱动杆在柔性环附近形成比较规则的极化液晶态排列. 由于杆粒子间向列型的相互作用, 整个柔性环内的自驱动杆状粒子基本是中心对称分布的, 此时用系统平均取向模的大小表示整体的极性程度, 柔性环内反向粒子的取向相互抵消, 系统的极性序接近于零. 但从快照上可以看出, 内部自驱动杆状粒子无序分布和有序聚集在环边界上是两种不同的分布, 而这两种分布的极性序都接近零无法很好地区分.

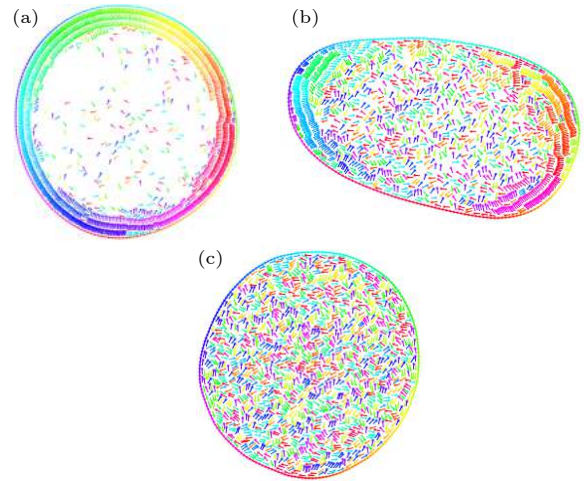


图 2 三种典型分布的快照, 自驱动杆粒子数 N_r 均为 1500, 噪声大小 η 分别为 0.10, 0.20 和 0.50, 依次对应 (a) 自驱吸附有序态、(b) 过渡态和 (c) 无序态. 粒子颜色代表取向, 同图 1

Fig. 2. The snapshots of three regions with fixed particle number $N_r = 1500$ for different noise levels, and, respectively, with (a) $\eta = 0.10$, self-propelled particle adsorbed ordered region, (b) $\eta = 0.20$ transient region, and (c) $\eta = 0.50$ disordered phase. The color represents the radial direction as Fig. 1.

为分析柔性环内自驱动杆状粒子角度上的分布, 定义一个径向极性序参量

$$S_p = \langle \theta_i \cdot \varphi_i \rangle, \quad (6)$$

其中 θ_i 是杆 i 的取向, φ_i 是杆 i 到环心位置相对位移的方向. 该极性序是粒子取向与其相对位移方向

两者夹角余弦值的平均, 反应了内部自驱动杆状粒子在沿环质心向外方向上的取向有序程度. 极性序 S_p 趋近 0 时, 内部自驱动杆的取向是各向均匀的; 而当极性序 S_p 趋近 1 时, 这些粒子基本都是背离环质心指向环外. 对于单个粒子, 其夹角的余弦值可以为负, 但由于柔性环边界会聚集内部的自驱动粒子, 系统整体平均后的极性序 S_p 基本都是正的, 且极性序 S_p 的值越大, 表明在该参数点下, 内部自驱动粒子在柔性环边界的聚集程度越高.

3.1 极性序相图

本文主要研究体系的密度和噪声对系统形态的影响. 系统改变自驱动杆的粒子数 N_r 和杆端的噪声强度 η , 测量各参数空间点的径向极性序 S_p 的值, 可以得到如图 3(a) 所示的相图. 明显地, 根据极性序 S_p 值的大小, 相图中从最左边极性序 S_p 接近 1 的有序相区域, 逐渐过渡到右侧极性序 S_p 接近 0 的无序区. 有序区主要集中在粒子数 N_r 较大, 噪声强度 η 较低的区域, 对应的快照如图 2(a) 所示. 大部分的自驱动杆状粒子都指向环外方向, 集中排列在柔性环上, 且可以构成完整的层状分布. 同时剩余的粒子在中心区域形成角度和位置都比较均匀地分布. 无序区域则主要分布在粒子数 N_r 较小或噪声强度 η 较大的区域, 如图 2(c) 所示, 其内部自驱动粒子的取向是无序的, 均匀分布在环内. 过渡区间主要分布在这两相之间的区域, 如图 2(b) 所示, 外层的自驱动杆无法形成完整的层状稳定排布, 而分别集中成反向的两个集团或异向的多个集团. 外侧的柔性环也因此有明显的变形, 中心区域同样存在一定密度的比较均匀的无序气态分布.

3.2 粒子分布特性

由三个相区不同的粒子分布可知, 内部自驱动粒子除了角度分布上具有向外的极性取向, 粒子本身的空间位置分布也存在各向异性的情况. 为分析粒子位置分布的不均匀性, 根据所有自驱动杆的位置信息定义体系分布非球度 Δ .

首先由所有杆的质心位置计算惯量张量 Q , 其元素分别是

$$Q_{mn} = \frac{1}{N} \sum_{i=1}^N R_m^{(i)} R_n^{(i)}, \quad (7)$$

其中 $R_m^{(i)}$ 是第 i 个粒子的相对位矢 $\mathbf{R}^{(i)}$ 在笛卡尔坐

标系第 m 维方向上的分量. 特别地, 相对位矢 $\mathbf{R}^{(i)}$ 以粒子的平均质心位置为原点, 故满足条件 $\sum_{i=1}^N \mathbf{R}^{(i)} = 0$. 该惯性张量有两个本征值 ε_1 和 ε_2 , 非球度 Δ 可以通过这两个本征值定义^[33]:

$$\Delta = \left\langle \frac{(\varepsilon_1 - \varepsilon_2)^2}{(\varepsilon_1 + \varepsilon_2)^2} \right\rangle_t, \quad (8)$$

$\langle \dots \rangle_t$ 表示系统稳定状态各时刻的时间平均. 显然非球度 $\Delta = 0$ 对应于粒子在各个方向上均匀分布, 非球度 Δ 的值越大, 表明粒子分布的各向异性越明显, 当非球度 $\Delta = 1$ 时粒子基本分布在一条直线上.

通过计算各参数点的非球度 Δ , 可以得到如图 3(b) 的热度图, 在有序相和无序相区域, 系统的非球度 Δ 都是接近 0, 这两个相区内粒子的位置分布都是各项同性的. 无序区由于粒子取向的无序性, 各向同性的粒子位置均匀分布在柔性环内; 而

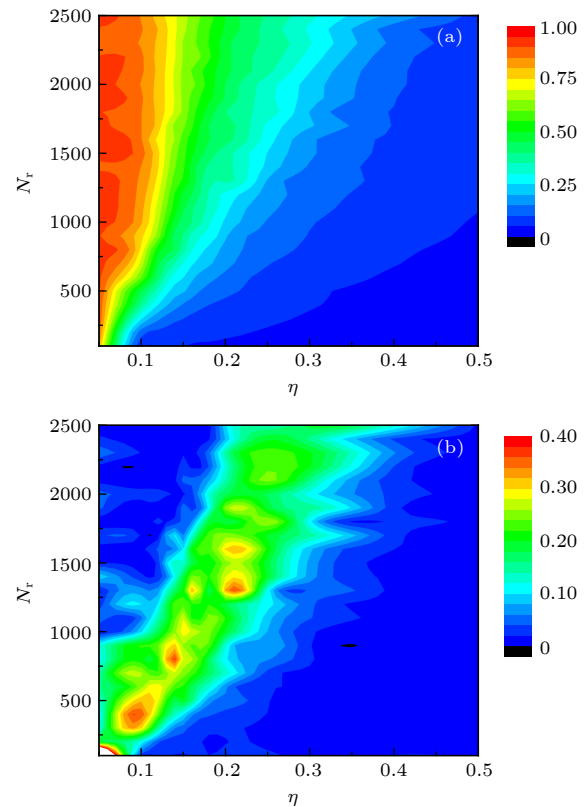


图 3 改变噪声强度 η 和弹性环中自驱动杆粒子数 N_r 得到的相图 (a) 比较径向极性序参 S_p 大小得到的热力图; (b) 比较非球度 Δ 大小得到的热力图, 其中转变区域具有极大值

Fig. 3. Phase diagrams for self-propelled rods in elastic-ring with varying the noise strength η and the number of self-propelled rods N_r , and the order parameter corresponding to (a) the radial polarity S_p and (b) the asphericity Δ . We have maximal asphericity Δ in the transition region.

有序区柔性环中心虽然也存在类似无序区的粒子均匀分布, 但它的极性序 S_p 主要由聚集在柔性环边界层状分布粒子贡献. 这部分粒子在边界高密度堆积形成稳定的层状液晶分布, 使系统整体处于类似汽液共存的动态平衡中. 而过渡区具有相对较大的非球度 Δ , 这是由于粒子数限制以及噪声影响, 外层的自驱动粒子很难在环边界处形成完整的层状结构. 在这个区间外层粒子由于角度上的偏离和数量上的减少, 以及转变区间涨落的增强, 对整体的径向极性序 S_p 贡献会有所降低.

内部自驱动粒子的位置分布除了在过渡相区会有明显的各向异性外, 也容易在柔性环边界位置聚集, 形成类似吸附相分离的密度分布差异. 我们将柔性环边界附近聚集形成的比较高密度分布的粒子排布区域划分为高密度态. 而中间比较无序分布的区域, 划分为低密度区. 由图 2 可以看到, 在有序区和过渡区 (图 2(a) 和图 2(b)), 内外侧粒子分布存在明显的密度差异, 而在无序区则不明显. 通过比较两种密度态的密度差异, 来表征粒子分布的相分离程度和粒子在环边界的聚集情况.

由于柔性环的形状易变, 以及杆状粒子自身的各向异性, 直接用杆的质心位置计算粒子密度不太合适. 我们根据杆的质心位置先得到每个粒子的泰森多边形, 计算各粒子相应的占据面积 S_i , 该面积对应每个粒子相对自由的运动范围. 然后定义粒子的数密度,

$$\psi = \frac{|\omega|}{\sum_{i \in \omega} S_i} = \frac{1}{\langle S_i \rangle}, \quad (9)$$

其中 ω 为符合高密度区或低密度区的粒子序号集合, $|\omega|$ 为集合 ω 中的元素个数. 相应的 $\psi_{in} = |\omega_{in}| / \sum_{i \in \omega_{in}} S_i$ 表示系统的中心粒子数密度, ω_{in} 是沿环边界等比例收缩后中心气态区域内的粒子集合. $\psi_{out} = |\omega_{out}| / \sum_{i \in \omega_{out}} S_i$ 是系统的外层粒子数密度, ω_{out} 即杆质心在沿环一圈宽度为杆长 L_r 的环状区间内的粒子集合. 由于内部自驱动粒子在环边界聚集排列, 或低粒子数 N_r 时与柔性环碰撞, 环边界附近 ω_{out} 区域长时间尺度上稳定有粒子存在, 而靠近环边界粒子的泰森多边形面积会明显小于内部粒子, 最后得到的 ψ_{out} 会稍大于 ψ_{in} .

为比较高密度两相的分离程度, 可以定义约化密度差为

$$P = \frac{\psi_{out} - \psi_{in}}{\psi_{out} + \psi_{in}}, \quad (10)$$

约化密度差 P 的值越大说明系统内外两侧粒子分布的密度差异越大, P 值越接近 0 则对应系统内部的粒子分布越均匀. 通过计算系统的约化密度差可以得到如图 4(a) 所示的热力图, 低噪声有序区域明显对应的区域约化密度差 P 值较高, 高低密度两相分离明显. 高噪声无序相区域的 P 值很低, 接近均匀分布. 这与图 2 快照中的粒子分布是一致的. 可以看到约化密度差 P 与极性序 S_p 有类似的分布, 随着噪声强度 η 的减弱, 系统从无序相进入有序相的过程中, 柔性环附近内部粒子的堆积程度也相应增强. 极性序 S_p 主要由环附近堆积的粒子贡献, 而自驱动粒子在环边界的堆积程度由约化密度差 P 刻画, 所以极性序 S_p 的大小一定程度上与约化密度差 P 正相关.

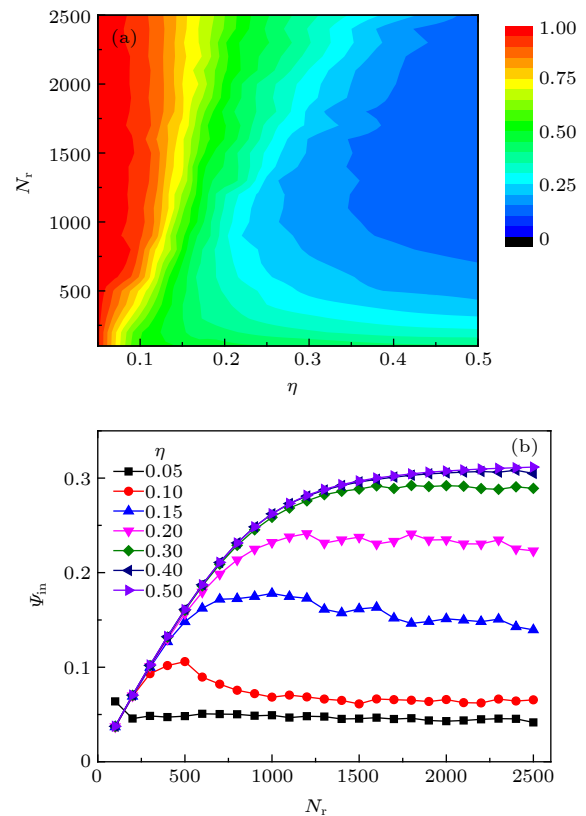


图 4 (a) 改变噪声强度 η 和自驱动杆粒子数 N_r , 比较约化密度差 P 得到的热力图; (b) 不同噪声强度 η 下, 弹性环中心附近粒子数密度随自驱动杆粒子数 N_r 的变化趋势

Fig. 4. (a) Phase diagram of the reduced density difference P for self-propelled rods with varying the noise strength η and the number of self-propelled rods N_r ; (b) density of central particles, ψ_{in} , versus the particle number N_r for different noise strength η .

注意到约化密度差的相图在高噪声强度 η 和低粒子数 N_r 区域, P 的值会有反向的变化. 此处的约化密度差 P 随粒子数 N_r 的增长而减小. 单独比较中心粒子数密度 ψ_{in} 的变化, 由图 4(b) 可以看到, 不同噪声下随粒子数 N_r 的增加 ψ_{in} 值最终都会有一个稳定的平台出现, 对应一个饱和密度的中心低密度态的存在. 该饱和密度与噪声强度有关, 噪声越强饱和密度越高. 而 ψ_{in} 在达到饱和密度前随粒子数 N_r 线性增加. 在相图右侧高噪声区域, 粒子数 N_r 比较小的情况下, 自驱动杆状粒子很难在环边界附近形成稳定的集聚, 此时外层区域 ω_{out} 内的粒子是与环边界碰撞的少数自驱动杆, 因而外层粒子数密度 ψ_{out} 比较稳定. 此时中心粒子数密度 ψ_{in} 未达到饱和, 增加粒子数 N_r 会减小内部粒子的泰森多边形面积, 即中心粒子数密度 ψ_{in} 会随粒子数 N_r 增加而迅速增大. 整体的约化密度差 P 主要受中心粒子数密度 ψ_{in} 的影响而减小. 而当中心粒子数密度 ψ_{in} 达到饱和后, 继续增加粒子数 N_r , 中心粒子数密度 ψ_{in} 基本保持稳定. 柔性环附近的粒子受挤压, 导致外层粒子数密度 ψ_{out} 小幅增大. 此时约化密度差 P 的值会相应有所增大. 所以约化密度差相图中 P 值变化的极值位置, 应该对应系统中心粒子数密度 ψ_{in} 刚达到饱和密度的参数点, 比较发现这两者基本是符合的.

在图 4(b) 中, 低噪声如 $\eta = 0.10$ 时, 系统的中心粒子数密度 ψ_{in} 会先有个小幅回落才能稳定在饱和密度. 这是因为低噪声区域, 粒子数 N_r 较小时, 内部的杆状粒子在自驱动作用下容易聚集在环边界位置, 但未能形成完整的层状排列. 噪声和粒子间的碰撞都参与到两侧的粒子交换中. 当粒子数 N_r 增大到形成完整序列时, 只有内侧层状排列的粒子在噪声作用下, 与环中心区域的低密度态的粒子发生交换. 此时的中心粒子数密度 ψ_{in} 会有所降低. 随噪声强度的增加, 粒子碰撞贡献的交换逐渐减少, 该密度回落过程也逐渐减弱, ψ_{in} 达到饱和平台的曲线也愈加平滑. 而当噪声较大时, 如噪声强度 $\eta = 0.4, 0.5$ 对应的 ψ_{in} 曲线几乎重合, 且没有明显水平的饱和平台. 此时由于噪声太大, 粒子很难在环边界形成有效的堆积, 主要依靠粒子间排斥将柔性环撑开. 且随粒子数增大, 柔性环拉伸后对内部粒子的压力也相应增大, 中心粒子数密度 ψ_{in} 会随粒子数变化而有一个小幅的增长.

3.3 整体迁移

由于内部自驱动粒子分布的不均匀性, 系统整体会有迁移运动, 为分析系统整体的动力学行为, 我们测量了柔性环质心的均方位移,

$$\text{MSD}(t) = \langle (\mathbf{R}(t'+t) - \mathbf{R}(t'))^2 \rangle, \quad (11)$$

其中 $\mathbf{R}(t)$ 是柔性环质心的位置. 如图 5(a) 所示, 我们测量了足够长时间尺度下, 三个相区各自的均方位移及其斜率 (补充材料 [movie1.mov](#), [movie2.mov](#), [movie3.mov](#), 分别对应 $\eta = 0.10, 0.20, 0.50$), 其中有序相和过渡相的均方位移斜率均接近 2. 无序相在长时间尺度时的均方位移斜率则接近 1. 系统在有序相和过渡相区域几乎都是在

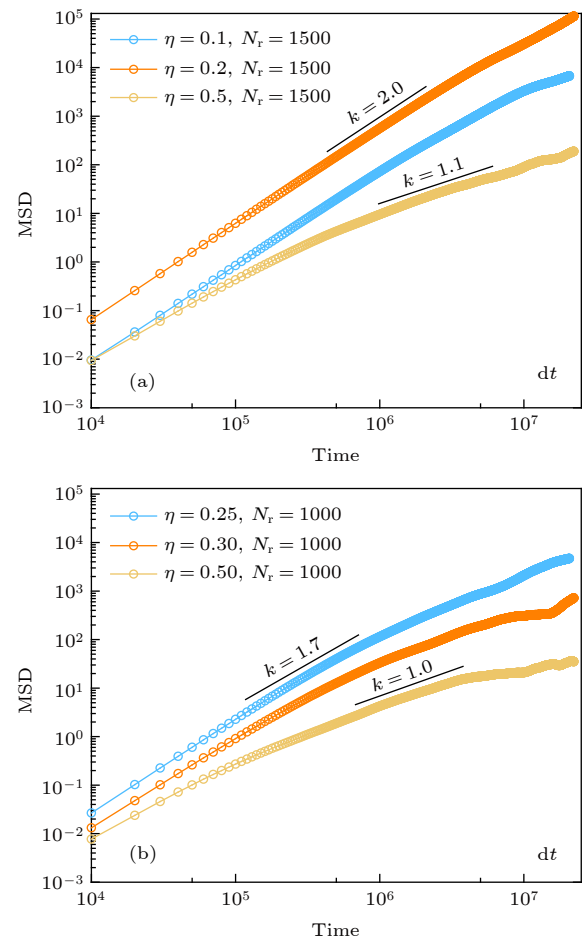


图 5 弹性环及杆状粒子质心均方位移随时间的变化 (a) 粒子数 N_r 为 1500 时, 噪声大小 η 为 0.10, 0.20 和 0.50 所在三个区区域的比较; (b) 粒子数 $N_r = 1000$, η 为 0.25, 0.30 和 0.50 下无序态时的对比

Fig. 5. Mean-squared displacement (MSD) for the center of mass of particle and elastic ring: (a) Noise levels $\eta = 0.10$, $\eta = 0.20$, and $\eta = 0.50$ for $N_r = 1500$; (b) noise levels for $\eta = 0.25$, $\eta = 0.30$, and $\eta = 0.50$ with particle number $N_r = 1000$ in the disordered regime.

做整体的定向迁移运动,而在无序区则接近随机游走.由于中心低密度态的粒子数比较少,且在角度上均匀分布,其对系统整体的运动影响很弱,系统的整体移动主要由沿环附近堆积排列的粒子贡献.当有序相和过渡相在形成稳定分布后,其沿环附近堆积的粒子不会有很大的变化.由于很难实现完全的对称,未被完全相互抵消的自驱动贡献,会导致系统整体沿某一方向定向迁移.从图 5(a)可以看出,由于有序相具有更高的对称性,多余的自驱动贡献比过渡相要低很多,其均方位移的尺度也要低一个数量级左右.而无序态的粒子分布和角度上的分布都很均匀,所以其瞬时的多余自驱动贡献在长时间尺度上是类似高斯白噪声的分布,其整体的运动也接近随机游走.

另外还比较了无序相区域不同噪声下的均方位移曲线,如图 5(b)所示,随着噪声强度 η 的增加,系统自身对称性更高,其均方位移尺度的数量级会相应降低,同时其长时间的扩散系数也相应减小,接近无规的随机运动.在无序相中,系统没有稳定的外层堆积分布,其整体的迁移主要来自噪声和极性涨落的竞争.当噪声较弱时系统的整体运动趋向于定向迁移,而噪声很强时系统会有类似随机游走的运动行为.值得指出的是质心的运动有两种可能,一种是整个弹性环形变导致质心运动而整体可能并没有发生明显的迁移,另一种是系统整体的迁移.在目前的体系中,系统的形变带来的影响远远小于整体迁移的作用,形变导致的质心迁移对均方位移的贡献是可以忽略的.

4 结 论

本文研究了在二维条件下,自驱动杆状粒子受限在一条可伸缩的柔性环中的动力学行为.通过随机动力学模拟,发现根据内部自驱动粒子整体的径向极性,可以从低噪声的有序区过渡到高噪声低密度时的无序区.同时粒子的空间分布也会产生变化.最主要的区别在于柔性环附近内部粒子的堆积方式:有序区内部自驱动粒子可以在环附近形成完整的层状排布,而无序区则由于高噪声影响,自驱动粒子无法在相应区域形成稳定的高密堆积.特别地,我们发现无序和有序转变的过渡区间环上有稳定的粒子堆积,但由于粒子数限制无法形成完整的层状排列,分散的堆积集团会导致柔性环明显形变.

除此之外,根据约化密度的异常转变,发现中心粒子数密度存在一个与噪声相关的饱和密度,该密度对应环边界附近位置开始形成稳定的粒子堆积.

系统整体的运动与这些边界集聚的粒子分布有关.在有序区和过渡区域中,环心的运动接近定向迁移.但由于有序相环边界位置粒子的层状分布更均匀,其整体定向迁移的强度会非常小.而当噪声增大,环边界附近稳定的粒子堆积逐渐减少,系统整体的运动也逐渐趋向无规则的随机游走.在我们的模型中,柔性环在过渡区会有明显形变,且此时系统整体的定向迁移最为显著.研究这类受限条件下的活性物质体系的动力学行为,对探讨细胞形变迁移方面的机制具有一定的参考意义^[37].目前在一些人工受限体系,比如将分子马达与杆状的微丝、微管系统受限到液滴表面,研究其中活性液晶态的动力学演化也正受到越来越多人的关注^[38,39].

参考文献

- [1] Cavagna A, Cimarelli A, Giardina I, Parisi G, Santagati R, Stefanini F, Viale M 2010 *Proc. Natl. Acad. Sci. U.S.A.* **107** 11865
- [2] Ballerini M, Cabibbo N, Candelier R, Cavagna A, Cisbani E, Giardina I, Lecomte V, Orlandi A, Parisi G, Procaccini A, Viale M, Zdravkovic V 2008 *Proc. Natl. Acad. Sci. U.S.A.* **105** 1232
- [3] Makris N C, Rattil P, Symonds D T, Jagannathan S, Lee S, Nero R W 2006 *Science* **311** 660
- [4] Katz Y, Tunstrom K, Ioannou C C, Huepe C, Couzin I D 2011 *Proc. Natl. Acad. Sci. U.S.A.* **108** 18720
- [5] Ginelli F, Peruani F, Pillot M H, Chate H, Theraulaz G, Bon R 2015 *Proc. Natl. Acad. Sci. U.S.A.* **112** 12729
- [6] J.-C. Tsai, Fangfu Ye, Juan Rodriguez, J. P. Gollub, Lubensky T C 2005 *Phys. Rev. Lett.* **94** 214301
- [7] Zhang H P, Be'er A, Florin E L, Swinney H L 2010 *Proc. Natl. Acad. Sci. U.S.A.* **107** 13626
- [8] Chen L M 2016 *Acta Phys. Sin.* **65** 186401 (in Chinese) [陈雷鸣 2016 物理学报 **65** 186401]
- [9] Fily Y, Marchetti M C 2012 *Phys. Rev. Lett.* **108** 235702
- [10] Farrell F D, Marchetti M C, Marenduzzo D, Tailleur J 2012 *Phys. Rev. Lett.* **108** 248101
- [11] Weitz S, Deutsch A, Peruani F 2015 *Phys. Rev. E* **92** 012322
- [12] Abkenar M, Marx K, Auth T, Gompper G 2013 *Phys. Rev. E* **88** 062314
- [13] Surrey T, Nedelec F, Leibler S, Karsenti E 2001 *Science* **292** 1167
- [14] Deseigne J, Dauchot O, Chate H 2010 *Phys. Rev. Lett.* **105** 098001
- [15] Bialke J, Speck T, Lowen H 2012 *Phys. Rev. Lett.* **108** 168301
- [16] Wysocki A, Elgeti J, Gompper G 2015 *Phys. Rev. E* **91** 050302
- [17] Costanzo A, Di Leonardo R, Ruocco G, Angelani L 2012 *J. Phys. Condens. Matter* **24** 065101

- [18] Isele-Holder R E, Elgeti J, Gompper G 2015 *Soft Matter* **11** 7181
- [19] Ginelli F, Peruani F, Bar M, Chate H 2010 *Phys. Rev. Lett.* **104** 184502
- [20] Gao T, Blackwell R, Glaser M A, Betterton M D, Shelley M J 2015 *Phys. Rev. Lett.* **114** 048101
- [21] Wensink H H, Lowen H 2012 *J. Phys. Condens. Matter* **24** 464130
- [22] Ofer N, Mogilner A, Keren K 2011 *Proc. Natl. Acad. Sci. U.S.A.* **108** 20394
- [23] Prass M, Jacobson K, Mogilner A, Radmacher M 2006 *J. Cell Biol.* **174** 767
- [24] Weichsel J, Schwarz U S 2013 *New J. Phys.* **15** 035006
- [25] Theriot J A, Mitchison T J, Tilney L G, Portnoy D A 1992 *Nature* **357** 257
- [26] Pantaloni D 2001 *Science* **292** 1502
- [27] Lushi E, Wioland H, Goldstein R E 2014 *Proc. Natl. Acad. Sci. U.S.A.* **111** 9733
- [28] Wensink H H, Lowen H 2008 *Phys. Rev. E* **78** 031409
- [29] Lee C F 2013 *New J. Phys.* **15** 055007
- [30] Yang X, Manning M L, Marchetti M C 2014 *Soft Matter* **10** 6477
- [31] Abaurrea Velasco C, Dehghani Ghahnaviyeh S, Nejat Pishkenari H, Auth T, Gompper G 2017 *Soft Matter* **13** 5865
- [32] Tian W D, Gu Y, Guo Y K, Chen K 2017 *Chin. Phys. B* **26** 100502
- [33] Paoluzzi M, Di Leonardo R, Marchetti M C, Angelani L 2016 *Sci. Rep.* **6** 34146
- [34] Spellings M, Engel M, Klotsa D, Sabrina S, Drews A M, Nguyen N H, Bishop K J, Glotzer S C 2015 *Proc. Natl. Acad. Sci. U.S.A.* **112** E4642
- [35] Zhang R F, Ren C L, Feng J W, Ma Y Q 2019 *Sci. China-Phys. Mech. Astron.* **62** 117012
- [36] Pesek J, Baerts P, Smeets B, Maes C, Ramon H 2016 *Soft Matter* **12** 3360
- [37] Lober J, Ziebert F, Aranson I S 2014 *Soft Matter* **10** 1365
- [38] Opathalage A, Norton M M, Juniper M P N, et al. 2019 *PNAS* **116** 4788
- [39] Keber F C, Loiseau E, Sanchez T, et al. 2014 *Science* **345** 1135

SPECIAL TOPIC—Statistical physics and complex systems

Collective behaviors of self-propelled rods under semi-flexible elastic confinement*

Zhong Ying Shi Xia-Qing[†]

(Center for Soft Condensed Matter Physics and Interdisciplinary Research, School of Physical

Science and Technology, Soochow University, Suzhou 215006, China)

(Received 16 April 2020; revised manuscript received 18 April 2020)

Abstract

In biological active systems there commonly exist active rod-like particles under elastic confinement. Here in this work, we study the collective behavior of self-propelled rods confined in an elastic semi-flexible ring. By changing the density of particles and noise level in the system, It is clearly shown that the system has an ordered absorbing phase-separated state of self-propelled rods and the transition to a disordered state as well. The radial polar order parameter and asphericity parameter are characterized to distinguish these states. The results show that the gas density near the central region of the elastic confinement has a saturated gas density that co-exists with the absorbed liquid crystal state at the elastic boundary. In the crossover region, the system suffers an abnormal fluctuation that drives the deformation of the elastic ring. The non-symmetric distribution of particles in the transition region contributes significantly to the collective translocation of the elastic ring.

Keywords: active matter, self-propelled rod, absorption phase separation, saturation density

PACS: 05.10.Gg, 05.20.-y, 87.16.Uv, 05.40.Jc

DOI: 10.7498/aps.69.20200561

* Project supported by the National Natural Science Foundation of China (Grant Nos. 11674236, 11922506).

[†] Corresponding author. E-mail: sxiaqing@gmail.com

行人跟踪算法及应用综述

曹自强 赛斌 吕欣

Review of pedestrian tracking: Algorithms and applications

Cao Zi-Qiang Sai Bin Lu Xin

引用信息 Citation: *Acta Physica Sinica*, 69, 084203 (2020) DOI: 10.7498/aps.69.20191721

在线阅读 View online: <https://doi.org/10.7498/aps.69.20191721>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

基于元胞传输模型的楼梯区域行人运动

Study of pedestrian flow on stairs with a cellular transmission model

物理学报. 2019, 68(2): 020501 <https://doi.org/10.7498/aps.68.20180912>

基于演化博弈论的行人与机动车冲突演化机理研究

Evolution mechanism of conflict between pedestrian and vehicle based on evolutionary game theory

物理学报. 2018, 67(19): 190201 <https://doi.org/10.7498/aps.67.20180534>

人脑默认模式网络的动力学行为

Dynamics of the default mode network in human brain

物理学报. 2020, 69(8): 080203 <https://doi.org/10.7498/aps.69.20200170>

通道中行人-机动车相互作用机理的建模和模拟

Modeling and simulation on interaction between pedestrians and a vehicle in a channel

物理学报. 2018, 67(24): 240503 <https://doi.org/10.7498/aps.67.20181499>

考虑在能见度受限下行人跟随行为特性的建模与模拟

Modeling and simulation of following behaviors of pedestrians under limited visibility

物理学报. 2019, 68(24): 240504 <https://doi.org/10.7498/aps.68.20190707>

专题：统计物理和复杂系统

行人跟踪算法及应用综述*

曹自强 赛斌 吕欣†

(国防科技大学系统工程学院, 长沙 410073)

(2019年11月11日收到; 2019年12月18日收到修改稿)

行人跟踪是计算机视觉领域中研究的热点和难点, 通过对视频资料中行人的跟踪, 可以提取出行人的运动轨迹, 进而分析个体或群体的行为规律. 本文首先对行人跟踪与行人检测问题之间的差别进行了阐述, 其次从传统跟踪算法和基于深度学习的跟踪算法两个方面分别综述了相关算法与技术, 并对经典的行人动力学模型进行了介绍, 最终对行人跟踪在智能监控、拥堵人群分析、异常行为检测等场景的应用进行了系统讲解. 在深度学习浪潮席卷计算机视觉领域的背景下, 行人跟踪领域的研究取得了飞跃式发展, 随着深度学习算法在计算机视觉领域的应用日益成熟, 利用这一工具提取和量化个体和群体的行为模式, 进而对大规模人群行为开展精确、实时的分析成为了该领域的发展趋势.

关键词: 行人跟踪, 轨迹提取, 计算机视觉, 行人动力学**PACS:** 42.30.Tz, 05.45.TP, 89.75.-k**DOI:** 10.7498/aps.69.20191721

1 引言

近年来, 深度学习的浪潮席卷计算机视觉领域, 这不仅提高了通用物体的检测性能, 也极大地促进了行人检测的发展, 为行人跟踪领域的研究奠定了良好的基础^[1]. 行人检测的主要任务是判断图片或者视频中是否有行人, 如果有, 则用框图把行人标记出来^[2], 不用考虑前后两帧中行人的匹配问题. 行人跟踪与行人检测不同, 需要利用数据关联技术关联前后两帧中相似度最大的行人, 以达到对视频中的行人持续跟踪^[3]的目的, 从而得到行人运动的速度、轨迹和方向等信息^[4], 并将其进一步应用到个人或大规模群体行为的研究领域中去^[5,6]. 行人跟踪是计算机视觉应用中的一项基本任务, 虽然已有大量文献提出了各种算法, 但由于行人跟踪问题比较复杂, 不仅需要考虑到拍摄的角度、光照的变化^[7-9], 还需要考虑新目标出现, 旧目标消失, 以

及当跟丢目标再次出现时, 如何进行再识别^[10]等问题, 这使得健壮的行人跟踪算法仍然是一个巨大的挑战.

随着深度学习技术在计算机视觉领域的广泛应用, 用深度学习的方法来研究行人跟踪问题俨然成为了学术界的主流^[11]. 虽然已有相关的文献综述对行人跟踪领域中的算法进行总结, 但这些综述大多不够新颖, 所提到的算法依旧是传统的目标跟踪算法, 没有将最新的深度学习算法包含进来. 为了弥补已有文献的不足, 同时使得广大科研工作者掌握行人跟踪领域的最新发展趋势, 本文首先将行人跟踪领域的算法按照传统跟踪算法和深度学习跟踪算法的分类方法进行了系统介绍, 并选取相应的指标评估性能, 然后介绍几种经典的人类行为动力学模型, 回顾人类行为动力学领域的发展历程, 最后围绕新技术条件下的视频监控、拥堵人群分析、异常行为监测等典型应用场景进行了系统地阐述.

* 国家自然科学基金 (批准号: 82041020, 71771213, 91846301, 71790615, 71901067) 和湖南省科技计划项目 (批准号: 2017RS3040, 2018JJ1034) 资助的课题.

† 通信作者. E-mail: xin.lu@flowminder.org

2 传统跟踪算法

2.1 卡尔曼滤波算法

1960年, Kalman^[12]为了解决离散数据的线性滤波问题,提出了卡尔曼滤波算法,该算法后来被扩展到目标跟踪领域^[13],其核心思想是利用上一时刻目标状态的预测值和当前时刻目标状态的测量值得到当前时刻目标状态的最优估计,并把当前时刻得到的最优估计作为下一时刻目标的预测值进行迭代运算,如此循环往复,逼近目标的真实值^[14].该算法的创新点在于同时考虑了在预测过程和测量过程中的误差,并且认为这两种误差独立存在,不受测量数据的影响.

该算法包括包括预测阶段和更新阶段两部分.在预测阶段,利用目标上一时刻的预测值 $\hat{\mathbf{x}}_{k-1}^-$ 预测当前状态 $\hat{\mathbf{x}}_k^-$,并对误差协方差矩阵 \mathbf{P}_k^- 进行估计;在更新阶段,卡尔曼滤波器用加权的测量结果来矫正预测结果^[15].卡尔曼滤波的两个阶段如表1所列,各参数的含义如表2所列.

表1 卡尔曼滤波的预测阶段和更新阶段
Table 1. Prediction and update process of Kalman filtering.

预测阶段	更新阶段
$\hat{\mathbf{x}}_k^- = \mathbf{A}\hat{\mathbf{x}}_{k-1}^- + \mathbf{B}\mathbf{U}_{k-1}$	$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}^T (\mathbf{H}\mathbf{P}_k^- \mathbf{H}^T + \mathbf{R})^{-1}$
$\mathbf{P}_k^- = \mathbf{A}\mathbf{P}_{k-1} \mathbf{A}^T + \mathbf{Q}$	$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H}\hat{\mathbf{x}}_k^-)$
	$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}) \mathbf{P}_k^-$

表2 卡尔曼滤波公式中的参数及含义

Table 2. Parameters and meanings in the Kalman filter formula.

参数	含义
$\hat{\mathbf{x}}_k^-$	目标在 k 时刻的先验状态估计值,包括目标的位置、速度等参数,一般是 n 维向量
$\hat{\mathbf{x}}_k$	目标在 k 时刻的后验状态估计值,是对 $\hat{\mathbf{x}}_k^-$ 应用卡尔曼滤波更新后的值
$\hat{\mathbf{x}}_{k-1}$	目标在 $k-1$ 时刻的后验状态估计值
\mathbf{A}	状态转移矩阵,一般是 $n \times n$ 阶的方阵
\mathbf{B}	控制矩阵,一般为0
\mathbf{U}_{k-1}	外部控制量,一般也为0
\mathbf{P}_k^-	k 时刻的先验误差协方差矩阵,需要事先给定一个初始值,以后的值可以由卡尔曼滤波递归得到
\mathbf{P}_k	k 时刻的后验误差协方差矩阵,是对 \mathbf{P}_k^- 的修正
\mathbf{K}_k	卡尔曼增益
\mathbf{y}_k	测量值,一般只能测量目标的位置,是 m 维向量
\mathbf{Q}	系统噪声协方差矩阵,是一个需要调节的参数,一般假定它是一个固定的值,在实验中需要通过不断调节 \mathbf{Q} 值,来寻找滤波器的最优值
\mathbf{R}	观测噪声协方差矩阵,和测量仪器有关,在实验中要不断尝试来确定最优的 \mathbf{R} 值
\mathbf{H}	观测矩阵,是 $m \times n$ 阶矩阵,用于将 m 维的测量值 \mathbf{y}_k 转换为与预测值 $\hat{\mathbf{x}}_k$ 相同的 n 维向量

将卡尔曼滤波算法应用到行人跟踪领域时,一般令 $\hat{\mathbf{x}}_k^- = [d_x(k), d_y(k), v_x(k), v_y(k)]^T$, 其中 $d_x(k)$, $d_y(k)$ 分别表示 k 时刻目标中心点的 x 坐标和 y 坐标, $v_x(k)$, $v_y(k)$ 分别表示 k 时刻目标中心点沿 x 轴, y 轴的分速度^[16],因为正常行人在行走过程中不会突然加速或者减速,所以视频中相邻帧的行人之间的运动可以近似看作匀速运动,假设相邻两帧之间的时间间隔为 Δt ,则行人在 k 时刻的运动方程为:

$$\begin{cases} d_x(k) = d_x(k-1) + \Delta t \cdot v_x(k-1), \\ d_y(k) = d_y(k-1) + \Delta t \cdot v_y(k-1), \\ v_x(k) = v_x(k-1), \\ v_y(k) = v_y(k-1). \end{cases} \quad (1)$$

利用该运动方程,再结合卡尔曼滤波算法中两个阶段的相关公式,就可以预测目标在下一时刻的状态,具体操作流程为:

首先将 $\hat{\mathbf{x}}_k^- = [d_x(k), d_y(k), v_x(k), v_y(k)]^T$ 代入(1)式,则有

$$\hat{\mathbf{x}}_k^- = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 1 & 0 & \Delta t & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \hat{\mathbf{x}}_{k-1}^-$$

由此可得状态转移矩阵

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

其中外部控制量 \mathbf{U}_{k-1} 和控制矩阵 \mathbf{B} 均为零,之后

分别为误差协方差矩阵、系统噪声协方差矩阵赋予初始值, 一般令

$$P_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

根据以上参数便可求得预测阶段的 \hat{x}_k^- 和 P_k^- ;

然后用 $\mathbf{y}_k = [y_x(k), y_y(k)]$ 表示测量的目标中心点坐标, 因为 $\mathbf{y}_k = \mathbf{H} \times \mathbf{x}_k$, 所以状态转移矩阵

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix};$$

最后设置观测噪声协方差矩阵的初始值 $\mathbf{R} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, 再结合预测阶段得到的 \hat{x}_k^- 和 P_k^- 就可以算出目标在 k 时刻的后验状态估计值 \hat{x}_k 和后验协方差矩阵 P_k .

利用该算法, 有了系统前一时刻状态就能估计出系统下一时刻的状态, 同时还可利用当前的观测值修正系统状态, 使得估计的均方误差最小^[17].

卡尔曼滤波算法经常与其他算法结合共同完成行人跟踪任务, 很少单独使用. 例如李娟等^[18]用卡尔曼滤波算法原理对视频中的行人进行跟踪, 他们首先采用混合高斯模型得到运动行人的前景图像, 然后利用 HSV(hue, saturation, value) 颜色空间模型和基于形态学的目标重构方法消除运动阴影, 最后用卡尔曼滤波预测行人的位置并得到了行人的运动轨迹. 石龙伟^[19]将卡尔曼滤波与光流法结合起来, 先用光流法对视频进行预处理, 然后根据光流法获取的目标位置等信息用卡尔曼滤波实现对行人的有效跟踪. 王宏选^[20]在动态行人跟踪 TLD (tracking-learning-detection, 跟踪-学习-检测) 算法中引入了卡尔曼滤波器预测行人下一帧可能出现的区域, 以缩小检测范围, 提高检测速度, 改善行人之间因存在遮挡而导致跟踪丢失的问题.

卡尔曼滤波算法比较简单, 容易实现, 适合跟踪简单场景中的行人. 但其易受光照变化的影响, 而且当行人突然改变行走的方向或者速度时, 跟踪效果较差.

2.2 多假设跟踪算法

1979年, Reid^[21]提出了多假设跟踪算法, 该算法的最初目的是解决雷达信号的自动跟踪问题, 后来 Kim 等^[22,23]对多假设跟踪算法进行了改进, 将其扩展到目标跟踪领域. 多假设跟踪算法本质上

是基于卡尔曼滤波算法在多目标跟踪问题中的扩展, 其中假设是指聚簇内一组目标和量测的分配互联关系^[24]. 多假设跟踪是一种延时决策算法, 在数据关联发生冲突时, 会形成多种假设, 直到获取到新的信息再做决定, 主要包括数据聚簇、假设生成、计算假设得分、假设删除四部分^[25], 其中假设生成和假设删除是该算法的核心, 该算法的流程如下.

Step 1 数据聚簇, 将新接收的量测点迹与以前的假设进行关联.

Step 2 将所有可能的航迹生成假设并保存, 生成的假设用下面的公式表示:

$$\mathbf{Z}(k) \triangleq \{Z_m(k), m = 1, 2, \dots, M_k\}, \quad (2)$$

$$\mathbf{Z}^k \triangleq \{\mathbf{Z}(1), \mathbf{Z}(2), \dots, \mathbf{Z}(k)\}, \quad (3)$$

$$\Omega^k(k) \triangleq \{\Omega_i^k, i = 1, 2, \dots, I_k\}, \quad (4)$$

其中 $\mathbf{Z}(k)$ 表示 k 时刻的量测集合; \mathbf{Z}^k 表示 k 时刻的累积量测集合; Ω^k 表示 k 时刻关联假设的集合; M_k 是可用量测个数; Ω_i^k 表示先验假设; $Z_m(k)$ 的来源可能是原有目标的继续、新目标的量测、虚警等. 如果量测是原有目标的继续, 则它符合原有航迹的高斯分布, 否则量测是一个均匀分布的噪声; 如果是新目标的量测、虚警, 则出现当前关联的可能性可以通过泊松分布和二项分布的乘积表示.

Step 3 计算假设概率:

$$P_i^k = \frac{1}{c} P_D^{(N_{DT})} (1 - P_D)^{(N_{TGT} - N_{DT})} \beta_{FT}^{N_{FT}} \beta_{NT}^{N_{NT}} \times \left[\prod_{m=1}^{N_{DT}} N(Z_m - H\bar{x}, B) P_g^{k-1} \right], \quad (5)$$

其中 P_i^k 表示假设概率; c 表示归一化因子; P_D 表示检测概率; N_{DT} 表示与先前目标相关的量测数量; N_{FT} 表示与错误目标相关的量测数量; N_{NT} 表示与新目标相关的量测数量; N_{TGT} 表示先前已知目标数; β_{FT} 表示错误目标的密度; β_{NT} 表示已检测到的先前未知目标的密度.

Step 4 假设删除, 因为假设的积累会占据大量的内存, 增加运算量, 不利于实时跟踪, 所以对假设进行剔除. 目前有两种删除假设的方法, 分别是零扫描法和多扫描法.

零扫描法 首先使用零扫描滤波器处理每个数据集, 然后仅保留概率最大的那个假设. 另外一种改进的方法是不仅选择最大似然假设, 而且增加

卡尔曼滤波器中的协方差以解释误相关的可能性.

多扫描法 使用多扫描算法处理数据集之后仍存在若干假设, 然后再次修剪所有不太可能的假设, 但保持所有概率高于指定的阈值的假设.

多假设跟踪算法保留了假设的大量历史信息, 确保了跟踪效果的稳定性, 但同时由于这些历史信息占据了过多的存储空间, 使得该算法的计算量大, 实时性差^[26].

2.3 粒子滤波算法

针对卡尔曼滤波需要目标的状态变量满足高斯分布的缺点, Breitenstein 等^[27]提出了一种基于粒子滤波框架的多行人跟踪检测算法, 该算法是卡尔曼滤波算法的一般化方法. 卡尔曼滤波建立在线性的状态空间和高斯分布的噪声上, 而粒子滤波的状态空间模型可以是非线性的, 且噪声分布可以是任何型式, 是一种通过非参数化的蒙特卡罗方法来实现递推的贝叶斯滤波, 粒子滤波的基本原理是通过先验概率和当前观测值估计后验概率^[28], 该算法分为两个步骤.

首先进行数据关联, 用匹配算法最多将一次检测分配给至多一个目标, 再用匹配函数 $s(tr, d)$ 评估检测 d 与跟踪器 tr 的每个粒子 p 之间的距离, 并用为 tr 训练的分类器 $ctr(d)$ 对 d 进行评估:

$$s(tr, d) = g(tr, d) \cdot \left[ctr(d) + \alpha \cdot \sum_{p \in tr} p_N(d-p) \right], \quad (6)$$

其中 $p_N(d-p)$ 表示评估 d 和 p 之间距离的正态分布; $g(tr, d)$ 是门控函数, 代表检测相对于目标的速度和运动方向的位置.

其次计算跟踪器 tr 的粒子 p 的权重 $w_{(tr,p)}$:

$$w_{(tr,p)} = p(y_t | x_t^i) = \beta \cdot I(tr) \cdot p_N(p - d^*) + \gamma \cdot d_c(p) \cdot p_o(tr) + \eta \cdot ctr(p), \quad (7)$$

其中参数 β, γ, η 是实验设定的; $I(tr)$ 是指示函数, 如果检测与跟踪器关联, 则返回 1, 否则返回 0; $d_c(p)$ 表示置信密度; $p_o(tr)$ 是加权函数.

该算法不依赖于背景建模, 可以在复杂的遮挡场景中对大量动态移动的人进行鲁棒跟踪, 是完全的二维操作 (不需要摄像机或地面平面标定), 在行人跟踪的实验中可以很好地复现行人的运动模式^[29]. 针对现实世界中行人身体之间的遮挡问题, Xu 等^[30]用粒子滤波跟踪行人的头部, 并用基于颜色直方图

和方向梯度直方图的方法对头部外观模型进行更新, 有效地减少了由于遮挡问题而造成的行人标号变化频繁的问题. 该算法在 UT-Interaction 数据集的测试结果中, 行人的身份标号仅变化了 4 次.

2.4 基于马尔科夫决策的多目标跟踪算法

2015 年, Xiang 等^[31]提出了一种基于马尔科夫决策过程的在线多目标跟踪框架, 将多目标跟踪问题视作一个马尔科夫决策过程来处理. 马尔科夫决策过程由一个元组 $(S, A, T(\cdot), R(\cdot))$ 组成^[32], 其中 S 表示目标所处的状态, A 表示目标可以执行的动作, $T(\cdot)$ 表示状态转移函数, $R(\cdot)$ 表示奖励方程, 行人跟踪问题的马尔科夫决策过程如图 1 所示.

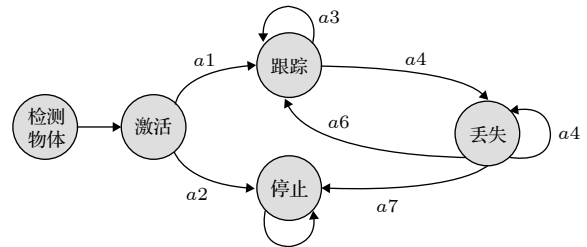


图 1 马尔科夫决策过程流程图^[31]
Fig. 1. Markov decision process flow chart^[31].

该算法将行人跟踪中的行人划分为激活、跟踪、丢失、停止四种状态. 其处理流程如下.

Step 1 行人被检测器检测到后首先进入激活状态, 然后用一个训练好的 SVM 分类器判断行人是进入跟踪状态还是停止状态, 其分类结果用一个 5 维的特征向量 $\Phi_{Active}(s)$ 表示. SVM 分类器是从训练视频序列中选出训练样本, 并将行人的 2 维坐标、高度、宽度以及检测得分归一化为一个 5 维的标准化向量训练得到的^[31], 其奖励函数为

$$R_{Active}(s, a) = y(a)(\mathbf{w}_{Active}^T \Phi_{Active}(s) + b_{active}). \quad (8)$$

如果转移到跟踪状态, 则 $y(a) = 1$; 如果转移到停止状态, 则 $y(a) = -1$; $(\mathbf{w}_{Active}^T, b_{active})$ 用于定义 SVM 的超平面.

Step 2 在跟踪状态下的奖励函数:

$$R_{Tracked}(s, a) = \begin{cases} y(a), & e_{medFB} < e_0 \text{ 且 } o_{mean} > o_0, \\ -y(a), & \text{其他,} \end{cases} \quad (9)$$

其中 e_{medFB} 表示预测误差的中间值, 如果预测误差 e_{medFB} 太大则跟踪失效; o_{mean} 表示前后两帧边界框重叠区域的平均值, 只有当 o_{mean} 在阈值 o_0 以上时

才被认为正确检测到目标; 因此当且仅当 e_{medFB} 小于阈值 e_0 和 o_{mean} 大于阈值 o_0 才表示跟踪有效, $y(a) = +1$; 否则进入丢失状态, $y(a) = -1$.

Step 3 丢失状态下的奖励函数:

$$R_{\text{Loss}}(s, a) = y(a)(\mathbf{w}^T \phi(t, d_k) + b). \quad (10)$$

如果转入跟踪状态, 则 $y(a) = +1$; 如果进入停止状态, 则 $y(a) = -1$; $\mathbf{w}^T \phi(t, d_k)$ 是捕捉目标和检测之间相似性的特征向量.

基于马尔科夫决策的多目标跟踪算法的跟踪效果比较好, 文献 [31] 将其在 MOT Benchmark 上进行了测试, 其多目标跟踪准确率可达 30.3%, 多目标跟踪精度可达 70.3%, 但是在行人长时间遮挡后容易发生误判.

2.5 相关滤波算法

相关滤波最初是表示信号处理领域中两个信号之间相似度的概念, 两个信号之间的相似度越高, 它们就越相关. 2019 年, Bolme 等 [33] 首次将相关滤波算法运用到目标跟踪领域, 其核心思想是利用误差平方和最小的滤波器 (minimum output sum of squared error, MOSSE) 训练图像, 使得图像的平方和误差最小, 从而建立跟踪目标的外观模型. 该算法的处理流程如下:

Step 1 首先训练相关滤波器, 最小化实际输出 $F_i H^*$ 与期望输出 G_i 之间的平方和误差:

$$\min_H \sum_i |F_i \odot H^* - G_i|^2; \quad (11)$$

Step 2 然后用训练好的相关滤波器 H^* 与输入图像 F 做相关操作, 求其响应 G :

$$G = F \odot H^*; \quad (12)$$

Step 3 最后用 PSR 作为响应 G 峰值强度的度量, 只有当 PSR 大于某个阈值时才会跟新目标的位置, 否则执行 Step 1, 一般 PSR 小于 7 表示跟踪失败,

$$\text{PSR} = \frac{g_{\text{max}} - \mu_{s1}}{\sigma_{s1}}. \quad (13)$$

该算法运用傅里叶变换操作, 极大地提高了运算速度, 使得该算法具有很好的实时性. 但由于没有考虑尺度的变化, 导致算法的鲁棒性比较差. 针对以上问题, Henriques 等 [34] 提出了 KCF 算法, 利用 HOG (histograms of oriented gradients) 特征代替 MOSSE 中使用的原像素, 增强了滤波器对目标

和环境的判别能力. 此外在 MOSSE 线性回归模型的基础上加入了正则项 $\lambda \|\mathbf{w}\|^2$, 建立了线性岭回归模型 [35,36]:

$$\min_w \sum_{i=1}^n (f(x_i) - y_i)^2 + \lambda \|\mathbf{w}\|^2, \quad (14)$$

其中 $\lambda \|\mathbf{w}\|^2$ 是 L_2 正则项, 用于解决过拟合问题; 对于非线性问题, 运用高斯核函数 $\phi(a, b) = \exp\left(-\frac{1}{\sigma^2} \|a - b\|^2\right)$ 将其转换为线性问题, 此时目标函数形式为:

$$f(x) = \sum_{i=1}^n a_i k(z, x_i). \quad (15)$$

KCF 算法利用循环矩阵和核函数大大提高了跟踪的速度和精度. 但其跟踪模型仅使用了 HOG 特征, 在特征信息出现模糊时容易导致跟踪失败, 而且学习率是固定不变的, 使得跟踪模型易受到周围环境的污染, 不适合长时间跟踪 [37].

3 深度学习算法

深度学习是包含多级非线性变换的层级机器学习方法, 深层次神经网络是其主要形式. 神经网络中层与层之间的神经元连接模式受启发于生物神经网络 [38]. 深度学习算法与传统算法相比, 不需要手动选择特征, 具备良好的特征提取能力. 但由于深度学习算法需要大量的数据以及高性能的计算机来训练数据, 而之前的计算机性能不能满足深度学习算法的需求, 也没有大量的数据用来训练 [39], 因而深度学习算法沉寂了相当长的一段时间 [40]. 近年来, 随着大数据时代的到来、计算机性能的提升, 深度学习开始广泛应用于计算机视觉的各个领域, 深度学习算法在行人跟踪领域应用的主流思路是 tracking-by-detection [41], 即首先用深度学习模型提取目标行人的特征, 检测出视频中的行人所在位置, 然后用多目标跟踪器对目标进行持续跟踪. 本文以卷积神经网络为例介绍深度学习算法在行人跟踪上的应用.

3.1 卷积神经网络

卷积神经网络 (convolutional neural network, CNN) 是一种典型的深度学习模型. LeCun [42] 最早提出了 CNN 的概念. 2012 年, Krizhevskyd

等^[43]首次将深度卷积神经网络应用到图像分类领域,其设计的 AlexNet 网络模型赢得了 ImageNet 图像分类比赛的冠军,成功地把深度卷积神经网络引入了计算机视觉领域.此后 ImageNet 比赛的冠军均是采用深度卷积神经网络的方法完成的. CNN 的基本结构包括输入层、卷积层、池化层、全连接层及输出层^[44],如图 2 所示.

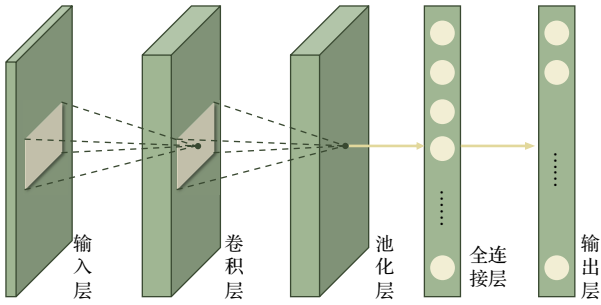


图 2 CNN 基本结构图

Fig. 2. CNN basic structure diagram.

图 2 中只绘制出了一个卷积层和池化层,然而在实际的网络中经常有若干个卷积层和池化层交替连接.在卷积层中通常使用一个大小为 $f \times f$ 的滤波器执行卷积操作来提取图像中的特征,网络前面浅层的卷积层用来提取图像的低级特征,后面更深层的卷积层用来提取图像的高级特征,全连接层将提取到的图片的特征归一化为一维的特征向量,输出层在分类问题中用来输出每个类别的概率.在处理行人检测问题时,卷积神经网络模型按照检测的步骤可以划分为 One-stage 和 Two-stage 两类. Two-stage 把检测行人分为两个阶段,首先产生行

人候选区域 (region proposals), 然后对候选区域进行分类,其典型代表是 RCNN, SPP-Net, Faster-RCNN 等模型,特点是准确率较高,但检测速度慢. One-stage 可以直接生成行人的类别概率和位置坐标,其典型代表是 YOLO 系列模型以及 CornerNet, CenterNet 等模型,其特点是运算速度较快,但准确率一般较低.

3.2 RCNN 网络模型

2014 年, Girshick 等^[45]提出 RCNN (regions with CNN features) 模型进行目标检测,将卷积神经网络引入目标检测领域. RCNN 就是在目标候选区内用 CNN 的方法来提取特征,处理流程如图 3 所示.

该模型首先使用选择性搜索 (selective search) 算法^[46]在输入图片上生成 2000 个左右的目标候选区域,然后对这些候选区域进行归一化操作,并将归一化后的候选区域送到 AlexNet 卷积网络提取特征.在 AlexNet 网络中有 5 个卷积层可以提取特征,经过一轮训练后,每个候选区域都能够得到一个 4096 维的特征向量,然后将提取到的特征传入 SVM 分类器中进行分类,最后使用卷积层的输出训练一个回归器 (dx, dy, dw, dh) 对候选区域进行微调,使其接近真实标注的区域.

RCNN 的出现使得科研工作者不用再手工设计大量的人工特征,而且准确率与传统检测算法相比有了很大的提高.但由于 RCNN 要求图片输入到卷积层的尺寸大小是固定的,因此需要对原始图片做尺寸变换,这会让图片产生形变,损失一部分

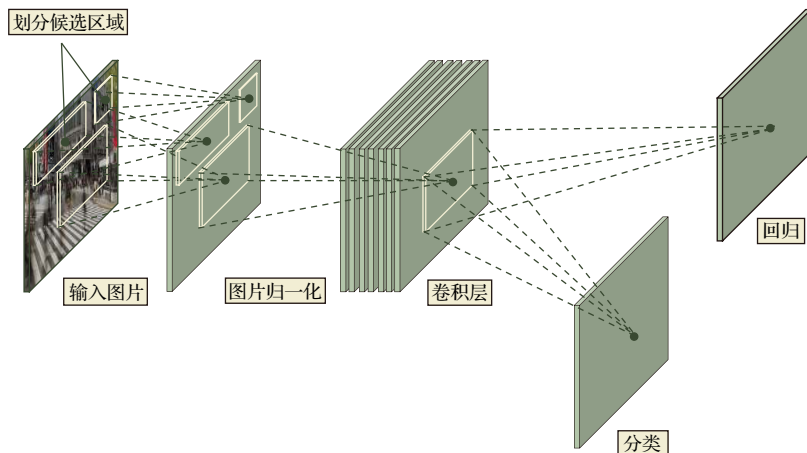


图 3 RCNN 算法流程图^[45]

Fig. 3. RCNN algorithm flowchart^[45].

特征, 降低了检测的准确率. 并且每次检测都要先生成 2000 个候选区域, 运算较为耗时, 导致 RCNN 在 VOC2007 数据集上检测一张图片大约需要 47 s^[47], 不能满足实时性要求.

3.3 SPP-Net 网络模型

2014 年, He 等^[48]发现感兴趣区域 (region of interest, ROI) 的特征都可以与特征图上相应位置的特征一一对应, 于是提出了 SPP-Net 网络模型. 该模型一次检测只需一次卷积运算, 这使得检测速度得到了极大提升, 其检测速度大约是 RCNN 的 100 倍. SPP-Net 网络模型的结构如图 4 所示.

SPP-Net 网络模型首先对输入图片使用选择性搜索 (selective search) 算法生成 2000 个左右的目标候选区域, 并将每个候选区域的大小划分为 4×4 , 2×2 , 1×1 的块, 然后用金字塔池化 (spatial

pyramid pooling, SPP) 层进行池化操作, 得到维度为 $(4 \times 4 + 2 \times 2 + 1 \times 1) \times 256$ 的特征向量, 最后将特征向量作为全连接层的输入, 在输出层输出. SPP-Net 网络模型的核心是在卷积层后加了空间金字塔池化层, 该层可以生成固定大小的图片, 不用对图像进行裁剪, 减少了特征的损失, 而且在整个过程中仅对图片做一次卷积特征提取, 极大地提高了通用物体的目标检测速度. 但其需要将数据分为多个训练阶段, 步骤较为复杂^[49].

3.4 Faster-RCNN 模型

虽然 RCNN 网络模型和 SPP-Net 网络模型检测目标的准确度较高, 但是它们在检测之前均要先生成 2000 个左右的候选区域, 这增加了目标检测的时间. Faster-RCNN 模型^[47]的最大贡献在于废除了选择性搜索算法, 利用区域提议网络

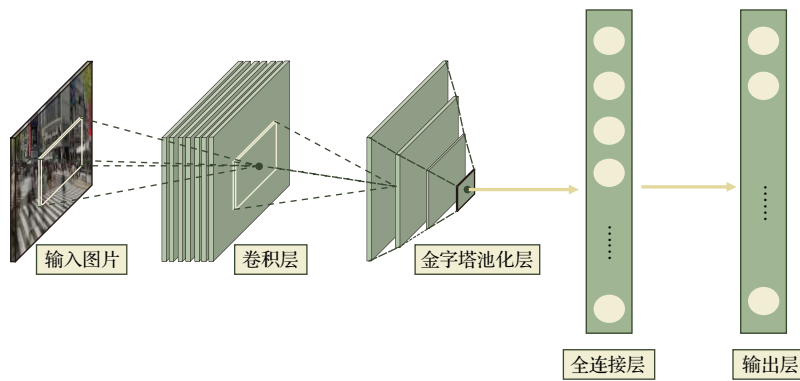


图 4 SPP-Net 结构图^[48]

Fig. 4. SPP-Net structure diagram^[48].

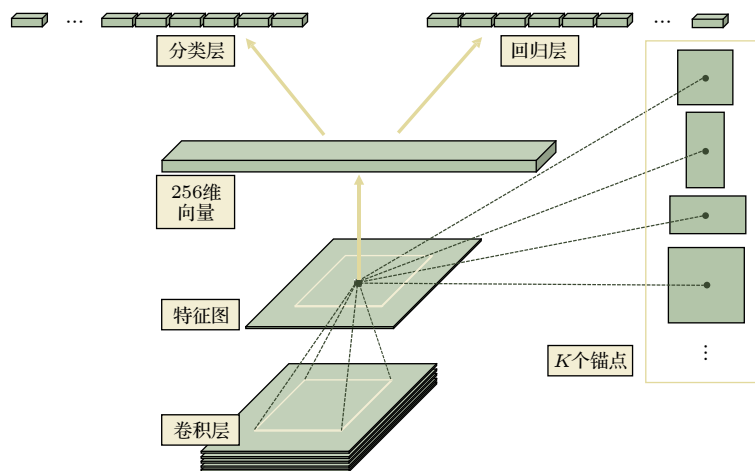


图 5 RPN 结构图^[47]

Fig. 5. RPN structure diagram^[47].

(region proposal network, RPN) 生成候选区域, 并通过共享卷积运算提取图片特征, 极大地降低了运算量^[50], 提高了检测速度. RPN 网络模型的结构如图 5 所示.

RPN 网络首先在 n 维特征图的每个像素上生成 k 个不同尺寸的锚框, 并给每个锚框分配一个二进制的标签 (是否是目标), 若锚框与实际目标的重叠区域的面积大于 0.7 倍总面积则被标记为正标签, 若锚框与实际目标的的重叠区域的面积小于 0.3 倍总面积则被标记为负标签. 然后用一个大小为 $s \times s$ 的滑动窗口生成一个 n 维的特征, 最后连接到分类层和回归层^[51], 判断是否存在目标并记录目标位置.

Faster-RCNN 采用 RPN 网络降低了生成候选框所需的时间, 并将 Softmax 分类器和回归器结合, 不用分别训练网络模型, 检测效果和速度均优于 RCNN 和 SPP-Net 网络, 作者在 PASCAL VOC 2007 数据集中测试的平均精度为 0.73. 但由于 RPN 网络中采用滑动窗口遍历卷积特征图, 因此也增加了时间的开销.

3.5 YOLO 系列模型

以上所介绍的目标检测算法都是先划分目标候选区域然后再预测目标类别, 而 YOLO^[52] 将目标区域候选区域的划分与类别的预测当作一个回归问题来处理, 直接在图片上输出多个目标的位置和类别, 在保证较高准确率的前提下实现目标的快速检测, 更能满足现实需求^[53]. YOLO 总共经历了 YOLO, YOLOV2, YOLOV3 三个版本, 下面分别对其进行介绍.

YOLO 采用改进的 InceptionV1^[54] 模型提取图片的特征, 因为 InceptionV1 模型要求输入图片的尺寸大小是 448×448 , 因此首先要将输入图片的尺寸调整为 448×448 , 其次将调整过尺寸的整张图片作为卷积网络的输入, 并用大小为 $S \times S$ 的网格对原始图片划分, 此时图片中物体的中心点就会落在某个网格单元内, 则对应的网格单元就负责检测该物体. 每个网格单元预测 B 个候选框和候选框内的置信度得分. 每个候选框中包含 x, y, w, h 和置信度 5 个信息^[55], 其中 (x, y) 表示预测边框的中心点坐标, (w, h) 表示预测边框的宽和高, 但需要注意的是中心点坐标的数值是相对于小网络边框而言的, 宽和高的数值是相对于整张图片而言

的. 如果网格单元不包含物体, 则置信度为 0, 否则置信度的计算公式为

$$\text{pr}(\text{object}) \cdot \text{IOU}_{\text{pred}}^{\text{truth}}, \quad (16)$$

得到置信度之后, 每个网格单元会给出 C 个类别的条件概率, 然后用 (17) 式计算各个网格单元内所有类别的概率.

$$\begin{aligned} \text{pr}(\text{Class}_i) &= \text{pr}(\text{Class}_i | \text{object}) \cdot \text{pr}(\text{object}) \cdot \text{IOU}_{\text{pred}}^{\text{truth}} \\ &= \text{pr}(\text{Class}_i) \cdot \text{IOU}_{\text{pred}}^{\text{truth}}. \end{aligned} \quad (17)$$

最后这些类别预测信息和置信度得分被编码到 $S \times S \times (B \times 5 + C)$ 大小的向量中作为 YOLO 输出层的输出. 在预测时每个网格单元会依据类别概率生成一个候选框, 但是大物体会生成多个候选框, 作者利用非极大值抑制算法选择交并比 (intersection over union, IoU) 得分最高的候选框, 并去除冗余窗口, 优化检测结果. YOLO 的检测速度很快, 可以达到每秒 21 帧, 但是精度不高, 平均精度只有 0.66, 容易漏检小物体. 针对以上问题, 作者在 2017 年提出了 YOLOV2 模型^[56], 在 YOLO 模型的基础上做了 5 个方面的改进. 首先在每一个卷积层后面都增加了批标准化 (batch normalization, BN) 操作, 对数据做归一化预处理, 加快了收敛速度. 其次将输出层的全连接层替换为卷积层, 由此可以微调图片的输入尺寸, 使网络适应不同尺寸的输入. 然后引入了 Faster-RCNN 中候选区域框的概念, 并采用 K-均值聚类方法调整候选区域框的尺寸, 使其更好地适应目标的尺寸. 然后在模型中添加转移层, 将浅层特征图连接到深层特征图上, 有利于检测小目标. 最后 YOLOV2 不再让每一个小网格预测目标类别, 而把这一任务交给候选区域框. 这 5 个方面的改进使得 YOLOV2 在 PASCAL VOC 数据集上的检测速度达到每秒 40 帧, 平均精度为 0.786.

2018 年, 作者又对 YOLOV2 在速度和精度上进行了改进, 提出了 YOLOV3 模型^[57]. 该模型采用具有 Darknet-53 网络来做特征提取, 是 YOLOV3 的精度得以提升的关键因素. 为了进一步加强对小物体的检测能力, YOLOV3 利用多尺度特征对目标进行检测, 在论文中作者采用了大小为 13×13 , 26×26 和 52×52 三个不同尺度的特征. 最后在分类时用 Logistic 回归替代 YOLOV2 的 Softmax 回归, 以便对多标签任务分类. YOLOV3 检测一

张尺寸为 320×320 的图片所消耗的时间为 22 ms, 平均精度为 28.2%.

3.6 CornerNet 网络模型

以 YOLO 系列为代表的 One-stage 深度学习网络模型和 Two-stage 深度学习网络模型均属于 anchor-base 模型, 需要使用不同大小、不同高宽比的 anchor 作为检测目标的候选区域. anchor 的优点是将目标检测问题转化为目标与 anchor 的匹配问题, 不必用目标检测算法遍历图片, 极大地缩短了检测目标的时间, 使得以 YOLO 为代表的 One-stage 模型可以和 Two-stage 模型竞争. 但使用 anchor 的深度学习模型也存在着两个主要缺点, 在使用 anchor 时不仅需要预先生成大量的 anchor 以便和图片中的目标重叠, 这导致只有少量的 anchor 与目标重叠, 造成了正负样本不均匀的问题 [58], 而且这些 anchor 包含很多超参数, 比如 anchor 的数量、尺寸等, 使得训练过程变得复杂. 针对 anchor-base 模型存在的缺点, Law 和 Deng [59] 提出一种新的 One-stage 模型——CornerNet 模型, 该模型利用一对关键点——物体边界框的左上角点和右下角点来检测物体从而取代 anchor, CornerNet 网络模型的结构如图 6 所示.

CornerNet 网络模型首先用 hourglass 网络提取图片特征, hourglass 网络先将特征图下采样到一个很小的尺度, 之后再进行上采样还原特征图的尺度, 这样可以获取不同尺度下图片所包含的信息, 然后在 hourglass 网络之后连接两个预测模块, 这两个预测模块分别预测边界框的左上角点和右下角点, 最后在每个预测模块内部经过 Corner pooling 操作后生成 Heatmaps, Embeddings 和 Offsets.

Heatmaps 的作用是预测左上角点和右下角点

的位置, 其预测角点的损失函数为

$$L_{\text{det}} = \frac{1}{N} \sum_{c=1}^C \sum_{i=1}^H \sum_{j=1}^W \begin{cases} (1-p_{cij})^\alpha \log(p_{cij}), & y_{cij} = 1, \\ (1-y_{cij})^\beta (p_{cij})^\alpha \log(1-p_{cij}), & \text{其他}, \end{cases} \quad (18)$$

其中, c 表示类别, heatmaps 的尺寸是 $H \times W$, p_{cij} 为 heatmaps 中 c 类物体在位置 (i, j) 得分, y_{cij} 表示对应位置的 groundtruth, N 是目标数量, α, β 是超参数, 作者在实验中设置 $\alpha = 2, \beta = 4$; Embeddings 的作用是匹配同一个边界框的左上角点和右下角点, 其匹配原理是, 如果左上角点和右下角点来自同一个边界框, 则它们的 Embeddings 之间的距离应该比较小, 反之它们的 Embeddings 之间的距离应该比较大. Embeddings 通过两个损失函数来表示这两种距离:

$$L_{\text{pull}} = \frac{1}{N} \sum_{k=1}^N \left[(e_{t_k} - e_k)^2 + (e_{b_k} - e_k)^2 \right], \quad (19)$$

$$L_{\text{push}} = \frac{1}{N(N-1)} \sum_{k=1}^N \sum_{\substack{j=1 \\ j \neq k}}^N \max(0, \Delta - |e_k - e_j|), \quad (20)$$

其中 e_{t_k} 是目标 k 的左上角点, e_{b_k} 是目标 k 的右下角点, L_{pull} 表示属于同一个边界框的左上角点和右下角点之间的距离, L_{push} 表示属于不同边界框的左上角点和右下角点之间的距离, e_k 表示 e_{b_k} 和 e_{t_k} 的平均值, Δ 在实验中设置为 1; Offset 的作用是微调预测出的边界框, 因为对图片进行全卷积操作之后, 输出的图片尺寸会很小, 因此, 当将位置信息从热力图映射到输入图片时会存在精度损失, 这部分损失用 o_k 表示:

$$o_k = \left(\frac{x_k}{n} - \left\lfloor \frac{x_k}{n} \right\rfloor, \frac{y_k}{n} - \left\lfloor \frac{y_k}{n} \right\rfloor \right), \quad (21)$$

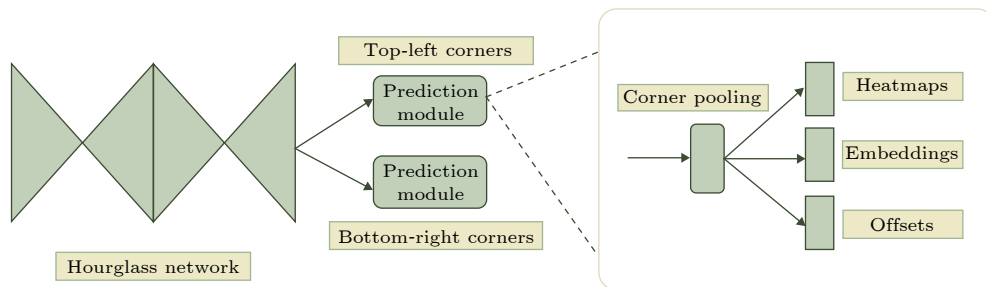


图 6 CornerNet 结构图 [59]

Fig. 6. CornerNet structure diagram [59].

其中 (x_k, y_k) 表示角点 k 的 (x, y) 坐标, 得到 o_k 之后, 利用 smooth L1 损失函数监督学习该参数:

$$L_{\text{off}} = \frac{1}{N} \sum_{k=1}^N \text{SmoothL1Loss}(o_k, \hat{o}_k). \quad (22)$$

由于左上角点和右下角点不在物体内部, 因此 Law 和 Deng^[59] 提出 corner pooling 来确定左上角点和右下角点, corner pooling 的原理是, 利用图片中的上边界和左边界的信息确定左上角点, 利用图片中的下边界和右边界确定右下角点. CornerNet 在 MSCOCO 数据集上测试的平均精度为 42.1%, 超过了绝大部分 One-stage 模型 在 MSCOCO 数据集中的平均精度.

3.7 CenterNet 网络模型

虽然 CornerNet 网络模型利用一对边角点取代 anchor 提高了物体检测的精度, 由于 CornerNet 网络模型中的边角点不在物体内部, 因此 CornerNet 网络模型无法感知物体内部的信息, 这其实也是大部分 One-stage 模型普遍存在的问题, 而 Two-stage 模型可以感知物体内部的信息, 因此 Two-stage 模型的准确率一般比 One-stage 模型的准确率高. 针对 CornerNet 模型无法感知物体内部信息的缺点, Zhou 等^[60] 提出 CenterNet 网络模型, 利用关键点估计找到物体的中心点并返回目标的尺寸、3D 位置、方向、甚至姿态等其他属性, 充分利用了物体内部的信息. Zhou 等^[60] 在 COCO 数据集上对 CenterNet 网络模型的速度和精度进行了测试, CenterNet 网络模型在 Resnet-18 网络下取得了每秒 142 帧的检测速度和 28.1% 的检测精度, 在 DLA-34 网络下取得了每秒 52 帧的检测速度和 37.4% 的检测精度, 在 Hourglass-104 网络下取得了每秒 1.4 帧的检测速度和 45.1% 的检测精度, 其精度可以媲美 Two-stage 网络, 实现了速度和精度的完美权衡. 该模型的核心思想是, 将图片输入到全卷积网络中生成一个热力图, 其中热力图的峰值对应目标的中心点, 每个峰值点的图像特征还可以预测边界框的宽和高, 返回目标的其他属性.

CenterNet 网络在训练过程中采用标准的密度监督学习训练网络, 首先用预先标注好的目标真实中心点坐标作为标签来预测目标的中心点坐标, 目标真实中心点坐标的计算公式为

$$p = \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2} \right), \quad (23)$$

其中 (x_1, x_2, y_1, y_2) 表示目标边界框的坐标, 计算出真实坐标 p 之后, 用下采样后的 $\tilde{p} = \lfloor p/R \rfloor$ 替代 p , 其中 R 是下采样因子, 然后采用高斯核 $Y_{xyc} = \exp\left(-\frac{(x - \tilde{p}_x)^2 + (y - \tilde{p}_y)^2}{2\sigma_p^2}\right)$ 将关键点分布到特征图上, 其中 σ_p 是目标尺寸自适应标准差, 并用如下所示的损失函数使得预测的目标中心点坐标与真实值之间的距离最小:

$$L_k = \frac{-1}{N} \sum_{xyc} \begin{cases} (1 - \hat{Y}_{xyc})^\alpha \log(\hat{Y}_{xyc}), & Y_{xyc} = 1, \\ (1 - Y_{xyc})^\beta (\hat{Y}_{xyc})^\alpha, & \text{otherwise,} \\ \log(1 - \hat{Y}_{xyc}), & \text{otherwise,} \end{cases} \quad (24)$$

其中 α, β 是损失函数的超参数, 在实验中作者设 $\alpha = 2, \beta = 4$; N 是图片中关键点的数量.

由于在训练过程中, 用下采样因子 R 对图片进行了下采样, 把特征图重新映射到原始图片时会存在误差, 因此用 local offset 补偿损失, 并用 L1 Loss 训练偏置值 L_{off} :

$$L_{\text{off}} = \frac{1}{N} \sum_p \left| \hat{O}_{\tilde{p}} - \left(\frac{p}{R} - \tilde{p} \right) \right|, \quad (25)$$

然后再利用得到的中心点坐标对每个目标的尺寸进行回归, 预测边界框的尺寸信息:

$$L_{\text{size}} = \frac{1}{N} \sum_{k=1}^N |\hat{S}_{p_k} - s_k|, \quad (26)$$

其中 $s_k = (x_2^{(k)} - x_1^{(k)}, y_2^{(k)} - y_1^{(k)})$ 是标准边界框的大小. 最后整体的损失函数为 L_k, L_{size} 与 L_{off} 三者的和, 而且每个损失都有相应的权重.

$$L_{\text{det}} = L_k + \lambda_{\text{size}} L_{\text{size}} + \lambda_{\text{off}} L_{\text{off}}, \quad (27)$$

其中, $\lambda_{\text{size}} = 0.1, \lambda_{\text{off}} = 1$; 这样用一个网络就可以得到目标中心点的预测值、偏置和尺寸.

由于 CenterNet 较为简单, 同时兼顾了速度和精度, 因此可以考虑将其应用到算力比较小的嵌入式平台中. 但 CenterNet 网络也存在一定的缺点, 由于它只检测物体的中心点, 因此当多个物体的中心点重叠时, CenterNet 只能检测出一个中心点, 会出现漏检的情况. 针对 CenterNet 的这一缺点, Duan 等^[61] 用中心点、左上角点和右下角点三个关键点检测物体, 提高了物体的检测精度, 在 MSCOCO 数据集中的检测精度达到了 47.0%, 但是

检测速度比较慢, 检测一张图片需要 340 ms^[61].

相较于传统方法, 深度学习能够从训练数据中抽取更加抽象的特征, 而且随着训练数据的增加, 模型的效果也显著增加. 但深度学习模型计算量大, 难以调参, 对设备的运算能力要求比较高, 大多数深度学习模型很难满足实时性要求. 因此在实际应用中一般会对深度模型的结构做轻量化处理, 牺牲一定的准确度来换取实时性, 或者提高硬件的运算能力.

3.8 多目标跟踪评价指标

由于深度学习算法的评价指标平均精度 (mean average precision, mAP) 属于目标检测领域的指标, 只能用于衡量检测目标的准确性, 不能用来衡量多目标算法的跟踪性能, 因此为了比较以上跟踪算法的性能, 需要选择相应的多目标跟踪评价指标对其进行衡量. 文献 [62] 最早提出了多目标跟踪准确度 MOTA、多目标跟踪精度 MOTP 两种评价指标, 此外在 MOT Challenge^[63] 多目标跟踪评价平台上也提供了部分评价指标, 如跟踪轨迹

大致完整 (大于 80%) 的比率 MT、虚警数 FP、丢失数 FN 以及轨迹误配数 IDS. 其中 MOTA 是最重要的一个指标, 用来度量算法能否准确确定目标个数.

$$MOTA = 1 - \frac{\sum_i (fp_t + m_t + mme_t)}{\sum_t g_t}, \quad (28)$$

其中 fp_t , m_t , mme_t 分别表示在第 t 帧时的误判数, 丢失数, 误配数, g_t 表示第 t 帧时跟踪的目标数. MOTA 的取值范围是 $(-\infty, 1]$, 仅当没有错误的时候取 1.

MOTP 用来度量算法能否准确确定目标的位置,

$$MOTP = \frac{\sum_{t,i} d_t^i}{\sum_t g_t}, \quad (29)$$

其中 d_t^i 表示目标 i 的预测位置与真实位置的距离; g_t 表示第 t 帧时跟踪的目标数.

本文从文献中查找了相关算法的性能指标并以表格的形式列举出来, 其中 “↑” 表示数值越大, 性能越好. “↓” 表示数值越小, 性能一越好. “—” 表示没有找到相关指标, 所以表中并未列出. 不同算法之间的性能对比如表 3 所列.

表 3 不同算法之间的性能对比
Table 3. Performance comparison between different algorithms.

算法	MOTA ↑	MOTP ↑	MT ↑	ML ↓	IDS ↓	数据集	类别
卡尔曼滤波 ^[64]	85.00%	—	—	—	—	MIT Traffic video dataset	传统跟踪算法
多假设跟踪算法 ^[21]	29.10%	71.70%	12.10%	53.30%	476	MOT Benchmark	传统跟踪算法
粒子滤波算法 ^[27]	—	—	80.80%	0.70%	10	CAVIAR dataset	传统跟踪算法
基于马尔科夫决策的多目标跟踪算法 ^[31]	30.30%	71.30%	13.00%	38.40%	680	MOT Benchmark	传统跟踪算法
相关滤波算法 ^[65]	83.40%	73.50%	—	—	—	Urban Tracker dataset	传统跟踪算法
基于Faster-RCNN的跟踪算法 ^[66]	38.50%	72.60%	8.70%	37.40%	586	MOT 15 Benchmark	深度学习跟踪算法
基于YOLOV3的跟踪算法 ^[67]	60.50%	79.30%	30.20%	19.60%	1129	MOT 16 Benchmark	深度学习跟踪算法

4 行人动力学模型

利用以上部分介绍的算法和模型, 便可以得到行人 k 的边界框坐标 $(x_{k1}, y_{k1}, x_{k2}, y_{k2})$, 对相应的坐标求平均值便可得到行人 k 的中心点坐标 (\bar{x}_k, \bar{y}_k) , 然后进一步用时间间隔 Δt 记录行人的中心点坐标, 最后用相邻两次的坐标之差与 Δt 相除便可求得行人的速度信息. 在得到行人的运动轨迹、速度等参数之后, 便可对移动人群的运动模式进行分析, 从中挖掘出群体行为的潜在规律^[68]. 已

有大量文献建立了各种模型分析行人的动力学行为^[69,70], 本文介绍三种典型的流体动力学模型、社会力模型、启发式模型, 以及结合了多种模型的集成模型.

4.1 流体动力学模型

针对人群行为分析问题, Henderson^[71-74] 首先将气体动力学和流体动力学模型应用到行人群体中. 他通过测量各种人群的速度频率分布发现行人在经过十字路口时大部分人为了躲避来往车辆会降低行走速度, 而少部分人会加快行走速度. 进一

步的研究还表明经过十字路口时女性的速度比男性的速度要低. 因此作者用性别和遇到的车辆数量对行人群体进行划分, 对女性和男性分别用二维气体的麦克斯韦-玻尔兹曼速度分布进行描述, 得到了很好的拟合效果, 但其存在一个动量守恒和能量守恒的假设条件.

后来, Helbing^[75] 舍弃了动量守恒和能量守恒的假设, 用一个改进的类玻尔兹曼的气体动力学模型来描述不同的行人群体. 该模型首先将行人按照行走方向的不同划分为不同的模式 μ , 其次给不同模式的 μ 设置三个变量 x, v_μ^0, v_μ , 其中 x 表示模式 μ 所处的位置, v_μ^0 表示模式 μ 的理想速度, v_μ 表示模式 μ 的实际速度, 然后利用上述三个变量建立密度方程 $\hat{\rho}_\mu(\mathbf{x}, \mathbf{v}_\mu, \mathbf{v}_\mu^0, t) = \frac{N_\mu(u(\mathbf{x}) \times v(\mathbf{u}_\mu), t)}{A \cdot V}$, 密度 $\hat{\rho}_\mu$ 与行人达到理想速度 v_μ^0 的趋势^[76,77]、行人间的相互作用、运动模式的改变、每单位时间内区域密度的增加或减少四个因素有关, 下面具体介绍这四个因素对密度 $\hat{\rho}_\mu$ 的影响.

1) 行人达到理想速度 v_μ^0 的趋势使得其密度 $\hat{\rho}_\mu$ 接近平衡密度 $\hat{\rho}_\mu^0$:

$$\hat{\rho}_\mu^0(\mathbf{x}, \mathbf{v}_\mu, \mathbf{v}_\mu^0, t) := \delta(v_\mu, -v_\mu^0) \rho_\mu^0(\mathbf{x}, \mathbf{v}_\mu^0, t), \quad (30)$$

ρ_μ^0 表示理想速度为 v_μ 而实际速度为 v_μ 的行人密度, $\delta(\cdot)$ 是狄拉克函数.

2) 行人之间的相互作用通过影响行人的行进速度进而影响行人密度, 这一因素的影响可以用类玻尔兹曼混沌假设^[78,79]建模:

$$\begin{aligned} & \hat{\sigma}_{\mu v}(\mathbf{u}_\mu^1, \mathbf{u}_v^1; \mathbf{u}_\mu^2, \mathbf{u}_v^2; \mathbf{x}, t) \\ &= \sigma_{\mu v}(\mathbf{v}_\mu^1, \mathbf{v}_v^1; \mathbf{v}_\mu^2, \mathbf{v}_v^2; \mathbf{x}, t) \delta(\mathbf{v}_\mu^{0,2} - \mathbf{v}_\mu^{0,1}) \\ & \quad \times \delta(\mathbf{v}_v^{0,2} - \mathbf{v}_v^{0,1}), \end{aligned} \quad (31)$$

$\hat{\sigma}_{\mu v}(\mathbf{u}_\mu^1, \mathbf{u}_v^1; \mathbf{u}_\mu^2, \mathbf{u}_v^2; \mathbf{x}, t)$ 表示模式 μ 和模式 v 的行人将其状态从 $(\mathbf{u}_\mu^1, \mathbf{u}_v^1)$ 更改为 $(\mathbf{u}_\mu^2, \mathbf{u}_v^2)$ 的相对速度.

3) 若行人的运动模式在运动过程中发生变化, 则有

$$\begin{aligned} & \hat{C}_{\mu v}(\mathbf{x}, \mathbf{x}_\mu, t) \\ &= \int \hat{\sigma}_u^{\nu\mu}(\mathbf{u}_v; \mathbf{u}_\mu; \mathbf{x}, t) \rho_\nu^0(\mathbf{x}, \mathbf{u}_v, t) d^4 \mathbf{u}_v \\ & \quad - \int \hat{\sigma}_u^{\mu\nu}(\mathbf{u}_v; \mathbf{u}_\mu; \mathbf{x}, t) \rho_\mu^0(\mathbf{x}, \mathbf{u}_v, t) d^4 \mathbf{u}_v. \end{aligned} \quad (32)$$

4) 行人走出和进入某个区域对该区域在单位时间内密度的影响, 可以用公式表示为

$$\hat{q}_\mu(\mathbf{x}, \mathbf{v}_\mu, \mathbf{v}_\mu^0, t) := \hat{q}_\mu^+(\mathbf{x}, \mathbf{v}_\mu, \mathbf{v}_\mu^0, t) - \hat{q}_\mu^-(\mathbf{x}, \mathbf{v}_\mu, \mathbf{v}_\mu^0, t). \quad (33)$$

最后利用密度方程推导出关于行人的空间密度 $\langle \hat{\rho}_\mu \rangle$, 平均速度 $\langle v_\mu \rangle$ 和速度方差 $\langle (\sigma_{\mu u, i})^2 \rangle$ 的流体动力学方程.

流体动力学模型从宏观层面来描述行人, 注重整个人群中密度和速度的平均值, 可以很好地把握行人的整体行为特征, 在解释行人拥堵、人群分流等方面得到了广泛的应用^[80]. 但是其忽视了人群中个体与个体之间的相互作用, 不能从微观层面刻画每个行人的行为特征, 也无法解释个体行为对群体行为的影响^[81].

4.2 社会力模型

针对流体动力学模型无法从微观层面上描述个体行为的问题, Johansson 等^[82] 提出了社会力模型. 社会力模型中的社会力由自驱动力、排斥力和吸引力三部分组成.

自驱动力的作用是驱使行人向目标方向前进. 行人正常行走时, 其实际速度 \mathbf{v}_α 等于理想速度 \mathbf{v}_α^0 . 如果受到障碍物的干扰, 行人会调整自己的速度, 但在自驱动力的作用下会产生一个指向 \mathbf{v}_α^0 的加速度, 使得实际速度 \mathbf{v}_α 向理想速度 \mathbf{v}_α^0 靠近, 自驱动力的作用可以表示为

$$\mathbf{F}_\alpha^0(\mathbf{v}_\alpha, v_\alpha^0 \mathbf{e}_\alpha) := \frac{1}{\tau_\alpha} (v_\alpha^0 \mathbf{e}_\alpha - \mathbf{v}_\alpha), \quad (34)$$

其中 \mathbf{F}_α^0 表示自驱动力, v_α^0 表示理想速度的大小, \mathbf{e}_α 表示理想速度的方向, τ_α 表示从实际速度变为理想速度所需的时间.

行人之间的排斥力用于减小其他行人对个体私人领域侵犯的影响. 当个人的私人领域受到侵犯时, 通常会感到不舒服, 从而产生一个排斥力与其他行人保持一定距离. 排斥力的大小取决于行人的密度和理想速度 v^0 , 行人 α 在接近陌生人 β 时, 可以用下面的方程表示:

$$\mathbf{f}_{\alpha\beta}(\mathbf{r}_{\alpha\beta}) := -\nabla_{\mathbf{r}_{\alpha\beta}} V_{\alpha\beta}[b(\mathbf{r}_{\alpha\beta})], \quad (35)$$

其中 $V_{\alpha\beta}[b(\mathbf{r}_{\alpha\beta})]$ 是 b 的单调递减函数, 且其椭圆形等值线指向运动方向, b 是椭圆的半短轴.

$$2b := \sqrt{(\|\mathbf{r}_{\alpha\beta}\| + \|\mathbf{r}_{\alpha\beta} - v_\beta \Delta t \mathbf{e}_\beta\|)^2 - (v_\beta \Delta t)^2}, \quad (36)$$

其中 $\mathbf{r}_{\alpha\beta} = \mathbf{r}_\alpha - \mathbf{r}_\beta$ 是行人 β 步距的数量级.

此外, 当行人靠近障碍物时也会感觉到不舒

服,行人与障碍物也会保持一定的距离以免自己受伤.这种与障碍物之间的排斥力可以描述为

$$\mathbf{F}_{\alpha\beta}(\mathbf{r}_{\alpha\beta}) := -\nabla_{\mathbf{r}_{\alpha\beta}} U_{\alpha\beta}(\|\mathbf{r}_{\alpha\beta}\|). \quad (37)$$

同样,斥力 $U_{\alpha\beta}(\|\mathbf{r}_{\alpha\beta}\|)$ 单调递减.

行人有时也会被其他行人(朋友)或者障碍物所吸引.在地点 \mathbf{r}_i 的吸引力 $f_{\alpha i}$ 可以用单调递增的吸引潜力 $W_{\alpha i}(\|\mathbf{r}_{\alpha i}\|, t)$ 表示:

$$\begin{aligned} f_{\alpha i}(\|\mathbf{r}_{\alpha i}\|, t) &:= -\nabla_{\mathbf{r}_{\alpha i}} W_{\alpha i}(\|\mathbf{r}_{\alpha i}\|, t), \\ \mathbf{r}_{\alpha i} &= \mathbf{r}_{\alpha} - \mathbf{r}_i. \end{aligned} \quad (38)$$

但是吸引力 $f_{\alpha i}(\|\mathbf{r}_{\alpha i}\|)$ 会随着时间不断减小,因为对行人或者障碍物的兴趣会随着时间不断降低.同时,这种吸引力效应是形成人群的原因.

上面所提到的力都是行走在运动方向上所能感知到的.除此之外,还应考虑在行人身后的弱影响 c , $0 < c < 1$.由此引入方向相关权重:

$$w(\mathbf{e}, \mathbf{f}) := \begin{cases} 1, & \mathbf{e} \cdot \mathbf{f} \geq \|\mathbf{f}\| \cos \varphi, \\ c, & \text{其他.} \end{cases} \quad (39)$$

综上,一个行人行为的排斥力和吸引力可以表示为:

$$\begin{aligned} \mathbf{F}_{\alpha\beta}(\mathbf{e}_{\alpha}, \mathbf{r}_{\alpha} - \mathbf{r}_{\beta}) &:= w(\mathbf{e}_{\alpha} - \mathbf{f}_{\alpha\beta}) \mathbf{f}_{\alpha\beta}(\mathbf{r}_{\alpha} - \mathbf{r}_{\beta}), \\ \mathbf{F}_{\alpha i}(\mathbf{e}_{\alpha}, \mathbf{r}_{\alpha} - \mathbf{r}_i) &:= w(\mathbf{e}_{\alpha} - \mathbf{f}_{\alpha i}) \mathbf{f}_{\alpha i}(\mathbf{r}_{\alpha} - \mathbf{r}_i, t); \end{aligned} \quad (40)$$

行人所受的总社会力为

$$\begin{aligned} \mathbf{F}_{\alpha}(t) &:= \mathbf{F}_{\alpha}^0(\mathbf{v}_{\alpha}, v_{\alpha}^0 \mathbf{e}_{\alpha}) \\ &+ \sum_{\beta} \mathbf{F}_{\alpha\beta}(\mathbf{e}_{\alpha}, \mathbf{r}_{\alpha} - \mathbf{r}_{\beta}) \\ &+ \sum_B \mathbf{F}_{\alpha B}(\mathbf{e}_{\alpha}, \mathbf{r}_{\alpha} - \mathbf{r}_B^{\alpha}) \\ &+ \sum_i \mathbf{F}_{\alpha i}(\mathbf{e}_{\alpha}, \mathbf{r}_{\alpha} - \mathbf{r}_i, t); \end{aligned} \quad (41)$$

社会力模型为

$$\frac{d\mathbf{w}_{\alpha}}{dt} := \mathbf{F}_{\alpha}(t) + \text{fluctuations}, \quad (42)$$

其中 *fluctuations* 表示波动,代表偶然或者故意偏离了常规的运动模式.社会力模型不仅能解释微观层面上行人行为模式的改变,比如遇到障碍物行人会减慢速度等,而且可以从个体行为模式的改变推导出整个人群中压力、密度等宏观参数的变化.

4.3 行为启发式模型

社会力模型虽然能够解释一部分人群行为,但

是在实际的应用过程中会产生比较复杂的数学公式,参数很难校准^[83].因此 Moussaïd 等^[84]提出一种简单的行为启发式模型来捕捉人群行为中的潜在规律.该模型认为行人行为模式的改变是通过两种简单的认知过程完成的,并加入了行人的视觉信息^[85-87].此外还考虑了极度拥挤情况下行人之间无意的碰撞行为.该模型主要包括视觉信息复现、认知过程、碰撞效应三部分.

首先对现实世界中行人所感受到的视觉信息进行复,视觉信息主要包括视野范围和视野范围内最近的障碍物到行人的距离两部分.其中视野范围用区间 $[-\phi, \phi]$ 表示,其含义是以行人行走视线 H_i 所在方向为基准,向视线 H_i 左右两边最大各倾斜 ϕ 度.视野范围内最近一个障碍物到行人 i 的距离用 f 表示,如果行人 i 以速度 v_i^0 在向目标方向前进的过程中不会与障碍物发生碰撞,则令 f 等于无穷大.其次构造两个认知过程来模拟行人对视觉信息的处理,第一个认知过程用来确定行人在遇到障碍物之后所选择的行走方向,第二个认知过程用来确定行人遇到障碍物之后应该调整为多大的速度.

经验数据表明^[87],行人在避开障碍物的前提下并不愿意在目的路线上偏离太多.因此,第一个认知过程是“行人在保证不与障碍物发生碰撞的前提下,选择一条到目的地 O_i 的最短路径”,可以表示为:

$$\begin{aligned} d(\alpha) &= d_{\max}^2 + f(\alpha)^2 - 2d_{\max}f(\alpha)\cos(\alpha_0 - \alpha), \\ \alpha_{\text{des}} &= \min d(\alpha), \end{aligned} \quad (43)$$

其中 α_{des} 表示行人遇到障碍物后所选择的最短路径方向, α_0 表示目的地的方向.

行人在发现障碍物之后,需要一个缓冲时间 τ 来防止与障碍物发生意外碰撞.由此第二个启发式是“行人在所选择的行走方向上与障碍物保持一定的距离”,可以表示为

$$v_{\text{des}}(t) = \min(v_i^0, d_h/\tau) \quad (44)$$

其中 $v_{\text{des}}(t)$ 表示实际速度, v_i^0 表示理想速度, d_h 表示最近的障碍物和行人的距离, τ 表示缓冲时间.

上述两部分可以很好地模拟出稀疏场景下行人的行为,但在拥堵场景下需要考虑行人行人之间无意的碰撞.

$$\mathbf{f}_{ij} = kg(r_i + r_j - d_{ij})\mathbf{n}_{ij}, \quad (45)$$

其中 \mathbf{n}_{ij} 是从行人 i 指向行人 j 的归一化向量; d_{ij}

表示行人之间的距离; r_i, r_j 分别表示行人 i, j 的半径; kg 是系数; f_{ij} 表示行人 i, j 之间的相互作用力.

启发式模型相比于流体动力学模型和社会力模型, 更加真实地模拟出了人们在遇到地震、恐怖袭击、火灾等紧急场景下的群体行为, 且该模型对个体的运动轨迹和集体的运动模式的预测与大量的经验和实验数据相一致.

4.4 集成模型

随着研究的深入, 人们发现单一的模型都存在一定的缺点, 很难准确描述人类行为. 比如启发式模型虽然简单高效, 但它无法像社会力模型那样描述来自其他行人或者障碍物的排斥力. 同样社会力模型中也不能运用启发式模型中的视觉信息来调节行走方向. 于是 Porter 等 [88] 提出了 IM 模型框架, 该模型集成了社会力模型、行为启发式模型以及材料科学理论, 充分发挥了每个模型的优势.

作者首先用社会力模型中的公式 (24) 式描述行人遇到障碍物后实际速度与理想速度的不同, 用 (25) 式和 (27) 式描述行人之间、行人与障碍物之间的排斥力. 其次用行为启发式模型确定行人在遇到障碍物时所选择的行走方向, 公式如下:

$$e_\alpha(t) = d_{\max}^2 + r(e)^2 - 2d_{\max}r(e)\cos(e_0 - e), \quad (46)$$

其中 $e_\alpha(t)$ 表示行人 α 的目标方向, d_{\max} 表示行人 α 视线的最远距离, $r(e)$ 表示距离行人最近的一个障碍物的距离, e_0 表示目的地的方向, e 表示视野范围内的方向.

作者最后将材料科学中“只考虑直接相邻的分子, 就可以很好地模拟分子间的相互作用”的理论应用到行人之间的交互中, 只考虑在行人视野中的多个行人的影响, 这样就不需要考虑周围所有的行人, 大大简化了运算量.

5 典型系统及应用

5.1 智能监控领域

现在市场上安装的监控摄像机需要有专门的人员不断地监察影像以应付可疑事件, 但实际上这些摄像头拍摄的视频很少有人一直监察或者根本没有人监察, 导致摄像机没有起到应有的作用 [89]. 因此, 如果计算机能自动监视它便可起到预警的作用, 帮助人们提前发现异常情况.

在拥挤的环境 (比如机场、车站、地铁等) 中经常发生行人丢失物品的事件, 一旦丢失物品, 就需要视频调度员从大量的视频中寻找, 耗时且滞后. 为了解决这一问题, Ferrando 等 [90] 建立了一种从室内场景中检测丢失物体的系统. 该系统包括目标分割、目标的识别和跟踪以及动作决策三个模块. 该系统首先将目标按照是行人还是物体、是动态还是静态划分为 4 类, 然后对动态的行人和物体进行持续跟踪, 最终物体处于静止状态时则被认定为丢失, 及时发出警报, 来协助工作人员进行处理. 但该系统仅适合在室内场合使用, 并且行人之间的遮挡时间过长会导致误报, 这时就需要人工来判别.

类似的系统还包括国内清华大学刘晓东等 [91] 开发的一套集运动目标检测、目标跟踪、目标分类于一体的智能监控系统, 湖南大学万琴和王耀南 [92] 提出的一种针对固定监控场景的运动检测与目标跟踪方法. 在国外, Nikouei 等 [93] 把智能监控作为一种边缘网络服务, 提出了一种轻量级 CNN 算法, 拥有更高的运算速度和较少的内存使用, 提高了行人检测的实时性. Gajjar 等 [94] 用 K -均值算法来跟踪监控视频中的行人, 首先在视频中记录与行人位置相关的 HOG 特征向量, 然后用 K -均值算法聚类得到了行人的轨迹.

5.2 拥堵人群分析

人群踩踏是拥堵人群中最具灾难性的事件之一 [95]. 当聚集在一个地点的人群密度过高时, 人与人之间不可避免地会发生身体接触, 此时一个行人对其周围人的作用力会像水纹一样不断向外传播, 并与其他各个方向不同大小的力相互叠加起来共同作用在人群中. 这些合力将人们在人潮中推来推去, 同时由于人群密度的增加, 人群之间的温度也随之升高. 拥挤的空间加上闷热的环境会让人产生头晕、乏力等症状. 如果此时有人不幸跌倒, 那么在多米诺骨牌效应的作用下会引发一系列连锁反应——周围的人也相继跌倒 [96], 从而引发大规模的踩踏伤亡事件. 拥堵人群的安全问题在应急管理、消防安全、建筑设计等领域都有着重要意义 [97].

然而行人跟踪领域长期缺乏带注释的视频数据集, 很难用计算机视觉的方法对其进行分析. 直到最近 Zawbaa 等 [98] 利用 HUER 数据集, 对 6 个不同的朝圣地点进行了建模, 可以实现在不同的朝圣视频场景中对仪式地点进行分类. 该系统包括预

处理、分割、特征提取和位置分类四个阶段. 将视频帧作为输入输入到 k 近邻 (KNN)、人工神经网络 (ANN) 和支持向量机 (SVM) 分类器中. 该系统普遍提高了六个朝觐仪式的识别准确率, 其准确率超过 90%. 虽然这个系统在识别朝觐仪式上有更好的准确性, 但还不足以对踩踏事件作出预警.

针对以上问题, Helbing 等^[99]通过对拥堵人群中密度、速度、压力等参数的分析, 发现当人群中的压力是导致踩踏事件的关键因素, 一旦压力超过 0.02 则踩踏事件不可避免, 由此提供了一个踩踏事件的预警机制. 在实际生活中, 踩踏事件多由突发事件引起, Zhao 等^[100]通过对由突发事件的研究发现, 在突发事件发生之后人们会争先恐后地往出口方向逃脱, 这造成出口区域附近人员密度极高, 人与人之间的拥挤程度加深, 减缓了人们的逃离速度. 针对这一现象, Zhao 等^[100]在出口处设置类似面板形状的障碍物引导行人分流, 降低出口区域的人群密度, 并通过实验表明该方法可以提高疏散效率, 减少人员伤亡. 因此对拥堵人群设计有效的疏散方法, 有助于减少人群中由于恐慌和从众行为所造成的经济损失和人员伤亡, 对改进应对突发事件的策略具有十分重要的意义^[101].

5.3 异常行为分析

在公共安全领域, 需要及时发现危险分子的异常行为, 确保人民群众的生命和财产安全. Yogameena 和 Nagananthini^[102]利用投影和骨架化方法对个体的正常和异常行为进行分类, 该系统主要包括运动检测与跟踪、行为分析, 它可以检测到人类走路、跑步、打架、弯腰等异常行为, 准确率高. 但是适合稀疏人群, 在拥堵人群中效果不佳.

随着我国老龄化人口越来越严重, 如何照顾好老人、防止老人摔倒是很重要的问题. Miaou 等^[103]认为约 70% 的意外摔倒是可以预防的, 并提出了一种结合个人信息 (如年龄、性别、体重) 的全摄像头来检测老年人是否摔倒. 摄像机的图像实时传输到一台服务器上, 对前景中感兴趣的对象进行背景减除, 然后系统使用连接组件标记来获取每个对象的面积、高度和宽度, 设置一个简单的判定跌落阈值来确定一个人是否跌落. 此外, 马里兰大学的计算机视觉实验室利用对灰度图像中的人体建模, 同时根据手、腿及头部等部位的动作的分析与跟踪, 能够实现多个行人进行检测与跟踪^[104]. Kocabas

等^[105]提出了一种多人姿态估计框架, 该框架将多任务模型与残差网络相结合, 可以联合处理人体检测、关键点检测和姿态估计问题, 作者在 COCO 关键点数据集上测试的检测速度是 23 帧每秒.

6 结 论

行人跟踪是计算机视觉领域中的难点和热点问题, 同时也是人类行为动力学中一个难点问题. 虽然之前的工作提出了大量再现人群行为的模型, 但由于缺乏对现有模型进行验证或校准的公开数据集, 使得这些模型之间没有一个很好的评判标准. 而且目前关于大规模人群行为的分析大都停留在统计和宏观层面上, 如基于手机数据对受地震、极端气候影响的人口迁移模式^[106-108]以及对个人旅行模式的时空分布的研究^[109,110]; 基于社交网络软件对自然灾害发生前后社交网络的结构及其演化的研究^[111], 对男同这一特殊群体的行为研究^[112]; 基于传感器的应急管理救援研究等^[113], 但微观层面上的研究极度缺乏. 近年来, 随着深度学习技术在计算机视觉领域的兴起, 已有部分文献运用深度学习技术分析人群中的异常行为、估计行人的运动姿态, 且取得了良好的效果. 这使得人们应用开源分析平台建立大规模标准数据集成为了可能, 并在微观和中观尺度下对大规模人群行为分析提供了一个新思路.

参考文献

- [1] Wang C, Sun X, Li H 2019 *J. Phys.* **1176** 032028
- [2] Li H X 2018 *M.S. Thesis* (Hefei: University of Science and Technology of China) (in Chinese) [李海翔 2018 硕士学位论文 (合肥: 中国科学技术大学)]
- [3] Hang Z Z 2011 *M.S. Thesis* (Changsha: National University of Defense Technology) (in Chinese) [黄忠主 2011 硕士学位论文 (长沙: 国防科学技术大学)]
- [4] Li X, Hu W, Shen C, Zhang Z, Dick A, Hengel A 2013 *ACM Trans. Intell. Syst. Technol.* **4** 58
- [5] Yan X Y 2011 *J. Univ. Electron. Sci. Technol. China* **40** 168 (in Chinese) [闫小勇 2011 电子科技大学学报 **40** 168]
- [6] Han X P, Wang B H, Zhou T 2010 *Complex Syst. Complex. Sci.* **07** 132 (in Chinese) [韩筱璞, 汪秉宏, 周涛 2010 复杂系统与复杂性科学 **07** 132]
- [7] Zhang S, Yao H, Sun X, Lu X 2013 *Pattern Recogn.* **46** 1772
- [8] Zhang K, Song H 2013 *Pattern Recogn.* **46** 397
- [9] Zhang S, Wang J, Wang Z, Gong Y, Liu Y 2015 *Pattern Recogn.* **48** 580
- [10] Wei R 2014 *M.S. Thesis* (Harbin: Harbin Engineering University) (in Chinese) [魏然 2014 硕士学位论文 (哈尔滨: 哈尔滨工程大学)]
- [11] Brunetti A, Buongiorno D, Trotta G F, Bevilacqua V 2018

- Neurocomputing* **300** 17
- [12] Kalman R E 1960 *J. Basic. Eng.* **82** 35
- [13] Comaniciu D, Ramesh V, Meer P 2003 *IEEE Trans. Pattern Anal. Mach. Intell.* **5** 564
- [14] Bishop G, Welch G 2001 *Proceedings of SIGGRAPH 2001* Los Angeles, August 12–17, 2001 p41
- [15] Chui C K, Chen G 2017 *Kalman Filtering* (New York: Springer) pp19–26
- [16] Huang S, Hong J 2011 *International Conference on Consumer Electronics, Communications and Networks* Xianning, China, March 11–13, 2011 p1423
- [17] Wang H 2018 *Comput. Know. Tech.* **14** 0194 (in Chinese) [王慧 2018 电脑知识与技术 **14** 0194]
- [18] Li J, Shao C F, Yang L Y, Li Q 2009 *J. Transp. Syst. Eng. Inf. Tech.* **9** 0148 (in Chinese) [李娟, 邵春福, 杨励雅, 李琦 2009 交通运输系统工程与信息 **9** 0148]
- [19] Shi L W 2017 *M.S. Thesis* (Chongqing: Chongqing University of Posts and Telecommunications) (in Chinese) [石龙伟 2017 硕士学位论文(重庆: 重庆邮电大学)]
- [20] Wang X H 2017 *M.S. Thesis* (Xian: Xidian University) (in Chinese) [王宏选 2017 硕士学位论文(西安: 西安电子科技大学)]
- [21] Reid D 1979 *IEEE Trans. Autom. Control* **24** 843
- [22] Kim C, Li F, Ciptadi A, Rehg J M 2015 *IEEE International Conference on Computer Vision* Santiago, Chile, December 13–16, 2015 p4696
- [23] Kim C, Li F, Rehg J M 2018 *European Conference on Computer Vision* Munich, Germany, September 8–14, 2018 p200
- [24] Zhai T H 2010 *Inf. Res.* **36** 25 (in Chinese) [翟海涛 2010 信息化研究 **36** 25]
- [25] Finn L, Kingston P 2019 *Proceedings of the IEEE Aerospace Big Sky*, Montana, March 2–9, 2019 p1
- [26] Yilmaz A, Javed O, Shah M 2006 *ACM Comput. Surv.* **38** 13
- [27] Breitenstein M D, Reichlin F, Leibe B, Koller-Meier E, Van Gool L 2010 *IEEE Trans. Pattern Anal. Mach. Intell.* **33** 1820
- [28] Gordon N J, Salmond D J, Smith A F 1993 *IEE Proc. F.* **140** 107
- [29] Breitenstein M D, Reichlin F, Leibe B, Koller-Meier E, Van Gool L 2009 *IEEE International Conference on Computer Vision* Kyoto, Japan, September 29–October 2, 2009 p1515
- [30] Xu R, Guan Y, Huang Y 2015 *Multimed. Tools. Appl.* **74** 729
- [31] Xiang Y, Alahi A, Savarese S 2015 *IEEE International Conference on Computer Vision* Santiago, Chile, December 13–16, 2015 p4705
- [32] White C 2001 *Markov Decision Processes* (New York: Springer) pp32–40
- [33] Bolme D S, Beveridge J R, Draper B A, Lui Y M 2010 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* San Francisco, CA, June 13–18, 2010 p2544
- [34] Henriques J F, Caseiro R, Martins P, Batista J 2015 *IEEE Trans. Pattern Anal. Mach. Intell.* **37** 583
- [35] Henriques J F, Caseiro R, Martins P, Batista J 2012 *European Conference on Computer Vision* Firenze, Italy, October 7–13, 2012 p702
- [36] Wang S 2017 *M.S. Thesis* (Hangzhou: Zhejiang University) (in Chinese) [王松 2017 硕士学位论文(杭州: 浙江大学)]
- [37] Deng X F, Peng X Y, Zhang J L, Xu Z Y 2019 *Semiconduct. Optoelectron.* **40** 742 (in Chinese) [邓雪菲, 彭先容, 张建林, 徐智勇 2019 光电技术及应用 **40** 742]
- [38] Zhou F Y, Jin L P, Dong J 2019 *Chin. J. Comput.* **40** 1229 (in Chinese) [周飞燕, 金林鹏, 董军 2019 计算机学报 **40** 1229]
- [39] Marcus G 2018 arXiv: 1801.00631[cs]
- [40] Goodfellow I, Bengio Y, Courville A 2016 *Deep Learning* (London: MIT press) pp10–15
- [41] Luo W, Xing J, Milan A, Zhang X, Liu W, Zhao X, Kim T K 2014 arXiv: 1409.7618 [cs]
- [42] LeCun Y 1989 *Connectionism in Perspective* (North Holland: Citeseer) pp23–25
- [43] Krizhevsky A, Sutskever I, Hinton G E 2012 *Advances in Neural Information Processing Systems* Lake Tahoe, Nevada, December 3–6, 2012 p1097
- [44] LeCun Y, Bottou L, Bengio Y, Haffner P 1998 *Proceedings of the IEEE* Leuven, Belgium, May 20–20, 1998 p2278
- [45] Girshick R, Donahue J, Darrell T, Malik J 2014 *IEEE Conference on Computer Vision and Pattern Recognition* Columbus, Ohio, June 24–27, 2014 p580
- [46] Uijlings J R, Van De Sande K E, Gevers T, Smeulders A W 2013 *Int. J. Comput. Vis.* **104** 154
- [47] Ren S, He K, Girshick R, Sun J 2015 *IEEE Trans. Pattern Anal. Mach. Intell.* **39** 1137
- [48] He K, Zhang X, Ren S, Sun J 2014 *IEEE Trans. Pattern Anal. Mach. Intell.* **37** 1904
- [49] Li Y, Hu J, Ji B 2019 *Journal of Physics: Conference Series* p022119
- [50] Chen Y J 2019 *M.S. Thesis* (Harbin: Harbin University of Science and Technology) (in Chinese) [陈怡佳 2019 硕士学位论文(哈尔滨: 哈尔滨理工大学)]
- [51] Nie W C 2019 *M.S. Thesis* (Harbin: Harbin Engineering University) (in Chinese) [聂文昌 2019 硕士学位论文(哈尔滨: 哈尔滨工程大学)]
- [52] Redmon J, Divvala S, Girshick R, Farhadi A 2016 *IEEE Conference on Computer Vision and Pattern Recognition* Las Vegas, Nevada, June 26–July 1, 2016 p779
- [53] Mittal N, Akarsh V, Kapoor S 2019 *Int. J. Sci. Res. Eng. Trends* **5** 562
- [54] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A 2015 *IEEE Conference on Computer Vision and Pattern Recognition* Boston, Massachusetts, June 8–10, 2015 p1
- [55] Yang J Y 2019 *M.S. Thesis* (Chengdu: University of Electronic Science and Technology of China) (in Chinese) [杨眷玉 2016 硕士学位论文(成都: 电子科技大学)]
- [56] Redmon J, Farhadi A 2017 *IEEE Conference on Computer Vision and Pattern Recognition* Honolulu, HI, July 21–26, 2017 p7263
- [57] Redmon J, Farhadi A 2018 arXiv: 1804.02767[cs]
- [58] Lin T Y, Goyal P, Girshick R, He K, Dollár P 2017 *IEEE International Conference on Computer Vision* Venice, Italy, October 22–29, 2017 p2980
- [59] Law H, Deng J 2018 *European Conference on Computer Vision* Munich, Germany, September 8–14, 2018 p734
- [60] Zhou X, Wang D, Krähenbühl P 2019 arXiv: 1904.07850[cs]
- [61] Duan K, Bai S, Xie L, Qi H, Huang Q, Tian Q 2019 *IEEE International Conference on Computer Vision* Seoul, Korea, October 27–November 2, 2019 p6569
- [62] Bernardin K, Stiefelhagen R 2008 *Eurasip J. Image Vide.* **2008** 246309
- [63] Leal-Taixé L, Milan A, Reid I, Roth S, Schindler K 2015 arXiv: 1504.01942[cs]
- [64] Shantaiya S, Verma K, Mehta K 2015 *Eur. J. Adv. Eng. Tech.* **2** 34
- [65] Yang Y, Bilodeau G A 2017 *Proceedings of the Computer*

- and Robot Vision Edmonton, Canada, May 16–19, 2017 p209
- [66] Chen L, Ai H, Shang C, Zhuang Z, Bai B 2017 *IEEE International Conference on Image* Beijing, China, September 17–20, 2017 p645
- [67] Yi Z, Shen Y, Zhao Q 2019 *Optik* **194** 163124
- [68] Zhou T, Han P X, Yan Y X, Yang Z M, Zhao Z D, Wang B H 2013 *J. Univ. Electron. Sci. Technol. China* **42** 481 (in Chinese) [周涛, 韩筱璞, 闫小勇, 杨紫陌, 赵志丹, 汪秉宏 2013 *电子科技大学学报* **42** 481]
- [69] Barbosa H, Barthelemy M, Ghoshal G, James C R, Lenormand M, Louail T, Menezes R, Ramasco J J, Simini F, Tomasini M 2018 *Phys. Rep.* **734** 1
- [70] Yan X Y, Wang W X, Gao Z Y, Lai Y C 2017 *Nat. Commun.* **8** 1639
- [71] Henderson L F 1974 *Transp. Res.* **8** 509
- [72] Henderson L 1971 *Nature* **229** 381
- [73] Henderson L, Lyons D 1972 *Nature* **240** 353
- [74] Henderson L, Jenkins D 1974 *Transp. Res.* **8** 71
- [75] Helbing D 1998 *Complex Syst.* **6** 391
- [76] Alberti E, Belli G 1978 *Transp. Res.* **12** 33
- [77] Helbing D 1991 *Behav. Sci.* **36** 298
- [78] Keizer J 2012 *Statistical Thermodynamics of Nonequilibrium Processes* (New York: Springer) pp22–34
- [79] Helbing D 1992 *Physica A* **181** 29
- [80] Haase K, Kasper M, Koch M, Müller S 2019 *Oper. Res.* **67** 376
- [81] Dong H, Zhou M, Wang Q, Yang X, Wang F Y 2019 *IEEE Trans. Intell. Transp. Syst.* DOI: [10.1109/TITS.2019.2915014](https://doi.org/10.1109/TITS.2019.2915014)
- [82] Johansson A, Helbing D, Shukla P K 2007 *Adv. Complex Syst.* **10** 271
- [83] Frank G A, Dorso C O 2011 *Physica A* **390** 2135
- [84] Moussaïd M, Helbing D, Theraulaz G 2011 *Proc. Natl. Acad. Sci. U.S.A.* **108** 6884
- [85] Gibson J J 1958 *Br. J. Psychol.* **49** 182
- [86] Batty M 1997 *Nature* **388** 19
- [87] Turner A, Penn A 2002 *Environ. Plann. B Plann. Des.* **29** 473
- [88] Porter E, Hamdar S H, Daamen W 2018 *Transp.* **14** 361
- [89] Collins R T, Lipton A J, Kanade T 2000 *IEEE Trans. Pattern Anal. Mach. Intell.* **22** 745
- [90] Ferrando S, Gera G, Regazzoni C 2006 *IEEE International Conference on Video and Signal Based Surveillance* Sydney, Australia, November 22–24, 2006 p21
- [91] Liu X D, Su G D, Zhou Q, Tian C 2019 *J. Image Graph.* **5** 1024 (in Chinese) [刘晓冬, 苏光大, 周全, 田超 2019 *中国图像图形学报* **5** 1024]
- [92] Wang Q, Wang N Y 2007 *Appl. Res. Comput.* **1** 199 (in Chinese) [万琴, 王耀南 2007 *计算机应用研究* **1** 199]
- [93] Nikouei S Y, Chen Y, Faughnan T R 2018 *Proceedings of the IEEE/ACM Symposium on Edge Computing* Bellevue, WA, October 25–27, 2018 p336
- [94] Gajjar V, Gurnani A, Khandhediya Y 2017 *IEEE International Conference on Computer Vision* Venice, Italy, October 22–29, 2017 p2805
- [95] Helbing D, Frankas I, Vicsek T 2000 *Nature* **407** 487
- [96] Helbing D 2013 *Nature* **497** 51
- [97] Haghani M, Sarvi M 2017 *Transp. Res. B Meth.* **107** 253
- [98] Zawbaa H M, Aly S A, Gutub A A 2012 arXiv: 1209.3433[cs]
- [99] Helbing D, Johansson A, Zein H, Abideen A 2007 *Phys. Rev. E* **75** 046109
- [100] Zhao Y, Li M, Xin L, Tian L, Yu Z, Kai H, Wang Y, Li T 2017 *Physica A* **465** 175
- [101] Wang B H, Zhou T, Shi D M 2016 *Mod. Phys.* **28** 50 (in Chinese) [汪秉宏, 周涛, 史冬梅 2016 *现代物理知识* **28** 50]
- [102] Yogameena B, Nagananthini C 2017 *Int. J. Disast. Risk. Re.* **22** 95
- [103] Miaou S G, Sung P H, Huang C Y 2006 *Proceedings of the Distributed Diagnosis and Home Healthcare* Arlington, VA, April 2–4, 2006 p39
- [104] Zhu J, Javed O, Liu J, Qian Y, Sawhney H 2014 *IEEE Conference on Computer Vision and Pattern Recognition* Columbus, Ohio, Jun 24–27, 2014 p3510
- [105] Kocabas M, Karagoz S, Akbas E 2018 *European Conference on Computer Vision* Munich, Germany, September 8–14, 2018 p417
- [106] Lu X, Wrathall D J, Sundsøy P R, Nadiruzzaman M, Wetter E, Iqbal A, Qureshi T, Tatem A J, Canright G S, Engo-Monsen K 2016 *Clim. Change* **138** 505
- [107] Lu X, Wrathall D J, Sundsøy P R, Nadiruzzaman M, Wetter E, Iqbal A, Qureshi T, Tatem A, Canright G, Engo-Monsen K 2016 *Glob. Environ. Change* **38** 1
- [108] Lu X, Bengtsson L, Holme P 2012 *Proc. Natl. Acad. Sci. U.S.A.* **109** 11576
- [109] Lu X, Wetter E, Bharti N, Tatem A J, Bengtsson L 2013 *Sci. Rep.* **3** 2923
- [110] Gonzalez M C, Hidalgo C A, Barabasi A L 2008 *Nature* **453** 779
- [111] Lu X, Brelford C 2014 *Sci. Rep.* **4** 6773
- [112] Huang G, Cai M, Lu X 2019 *Inter. J. Environ. Res. Pub. Heal.* **16** 3597
- [113] Lu X 2018 *Commun. CCF* **14** 56 (in Chinese) [吕欣 2018 *中国计算机学会通讯* **14** 56]

SPECIAL TOPIC—Statistical physics and complex systems

Review of pedestrian tracking: Algorithms and applications^{*}

Cao Zi-Qiang Sai Bin Lu Xin[†]

(*College of Systems Engineering, National University of Defense Technology, Changsha 410073, China*)

(Received 11 November 2019; revised manuscript received 18 December 2019)

Abstract

Pedestrian tracking is a hotspot and a difficult topic in computer vision research. Through the tracking of pedestrians in video materials, trajectories can be extracted to support the analysis of individual or collected behavior dynamics. In this review, we first discuss the difference between pedestrian tracking and pedestrian detection. Then we summarize the development of traditional tracking algorithms and deep learning-based tracking algorithms, and introduce classic pedestrian dynamic models. In the end, typical applications, including intelligent monitoring, congestion analysis, and anomaly detection are introduced systematically. With the rising use of big data and deep learning techniques in the area of computer vision, the research on pedestrian tracking has made a leap forward, which can support more accurate, timely extraction of behavior patterns and then to facilitate large-scale dynamic analysis of individual or crowd behavior.

Keywords: pedestrian tracking, trajectory extraction, computer vision, human behavioral dynamics

PACS: 42.30.Tz, 05.45.TP, 89.75.-k

DOI: [10.7498/aps.69.20191721](https://doi.org/10.7498/aps.69.20191721)

^{*} Project supported by the National Natural Science Foundation of China (Grant Nos. 82041020, 71771213, 91846301, 71790615, 71901067) and the Science and Technology Program of Hunan, China (Grant Nos. 2017RS3040, 2018JJ1034).

[†] Corresponding author. E-mail: xin.lu@flowminder.org

基于液滴界面不稳定性的表面粗糙聚合物微球的制备及其细胞捕获应用

王月桐 商珞然 赵远锦

Surface-textured polymer microspheres generated through interfacial instabilities of microfluidic droplets for cell capture

Wang Yue-Tong Shang Luo-Ran Zhao Yuan-Jin

引用信息 Citation: *Acta Physica Sinica*, 69, 084701 (2020) DOI: 10.7498/aps.69.20200362

在线阅读 View online: <https://doi.org/10.7498/aps.69.20200362>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

电场对协流式微流控装置中乳液液滴生成行为的调控机理

Mechanism for regulation and control of emulsion droplet generation in co-flow microfluidic device via electric field

物理学报. 2018, 67(17): 176801 <https://doi.org/10.7498/aps.67.20180616>

经典瑞利-泰勒不稳定性界面变形演化的改进型薄层模型

Improved thin layer model of classical Rayleigh-Taylor instability for the deformation of interface

物理学报. 2018, 67(9): 094701 <https://doi.org/10.7498/aps.67.20172613>

爆轰加载下高纯铜界面Rayleigh-Taylor不稳定性实验研究

Experimental research on Rayleigh-Taylor instability of oxygen-free high conductivity copper under explosive loading

物理学报. 2020, 69(9): 094701 <https://doi.org/10.7498/aps.69.20191999>

可压流体Rayleigh-Taylor不稳定性的离散Boltzmann模拟

Discrete Boltzmann simulation of Rayleigh-Taylor instability in compressible flows

物理学报. 2018, 67(8): 080501 <https://doi.org/10.7498/aps.67.20171952>

声场中空化气泡的耦合振动及形状不稳定性的研究

Coupled oscillation and shape instability of bubbles in acoustic field

物理学报. 2018, 67(3): 034301 <https://doi.org/10.7498/aps.67.20171573>

专题：统计物理和复杂系统

基于液滴界面不稳定性的表面粗糙聚合物微球的制备及其细胞捕获应用*

王月桐¹⁾ 商珞然^{2)†} 赵远锦^{1)‡}

1) (东南大学生物科学与医学工程学院, 生物电子学国家重点实验室, 南京 210096)

2) (复旦大学生物医学研究院, 复旦大学附属中山医院徐汇医院, 上海 200032)

(2020年3月12日收到; 2020年3月25日收到修改稿)

具有不同组成和形态的聚合物颗粒近来受到越来越多的关注, 它们的表面粗糙度显著影响着其理化性能, 尤其在调节生物材料与生物系统间的相互作用中发挥着重要作用. 本文设计了一种具有表面可调褶皱结构的聚苯乙烯微球. 首先通过微流控装置产生尺寸均一的含有疏水聚合物和助表面活性剂的液滴. 在有机溶剂的挥发过程中, 不断收缩的液滴出现界面不稳定现象. 表面面积自发增大, 固化后得到表面具有褶皱的微球. 研究表明, 调节助表面活性剂的浓度以及溶剂挥发速率均可以有效调控微球表面粗糙程度. 循环肿瘤细胞捕获实验表明, 这种褶皱结构能明显增强细胞黏附力, 提高细胞捕获量. 以上这些特征表明这种表面褶皱微球将在生物医学分析领域具备良好的应用前景.

关键词: 界面不稳定性, 微流控, 表面褶皱, 微球, 细胞捕获**PACS:** 47.20.Ib, 61.30.Pq, 83.10.Tv, 87.17.Uv**DOI:** 10.7498/aps.69.20200362

1 引言

具有不同组成和形态的聚合物颗粒近来受到越来越多的关注. 基于高分子材料的理化特性利用各种加工方法可以制备出不同尺寸和结构的聚合物微粒, 例如光滑微球、表面粗糙微球、多孔微球、片状微粒、红细胞状微球等^[1,2]. 它们的光学、电学性质等可以通过结构、尺寸和组成进行调节, 从而在生物材料、药理学、化工等领域有重要应用^[3-5]. 各向异性的表面微结构是这类聚合物颗粒最突出的特征之一. 表面形貌显著影响其理化性能, 例如催化性能和效率、自组装结构、血小板黏附特性、

生物化学反应程度等^[6,7]. 更有趣的是, 表面粗糙的粒子在调节生物材料和生物体系之间的相互作用方面起着至关重要的作用. 具有表面粗糙结构的微粒由于与某些具有特殊形状和表面结构的细胞及微生物在结构上有相似之处, 所以会对细胞的黏附、迁移、增长和组织愈合等进行调控从而行使特定的生物学功能^[8,9]. 因此, 调控粒子的表面微观结构在设计功能性生物材料方面提供了新的思路.

目前, 有很多方法可以制备聚合物微球, 主要有聚合法 (如乳液聚合、悬浮聚合、沉淀聚合)、乳液-溶剂挥发法 (如相分离、界面失稳)、微加工法 (如软印刷、微流控技术) 等^[10-12]. 通过调控实验参数, 可以得到各种尺寸、形状可调的微粒. 通常, 由

* 国家自然科学基金 (批准号: 61927805, 51522302)、国家自然科学基金委员会-中国工程物理研究院联合基金 (批准号: U1530260)、江苏省自然科学基金 (批准号: BE2018707) 和复旦大学和中山医院徐汇医院启动研究基金 (批准号: JIH1340032) 资助的课题.

† 通信作者. E-mail: luoranshang@fudan.edu.cn

‡ 通信作者. E-mail: yjzhao@seu.edu.cn

于在颗粒和介质之间的界面自由能趋于最小化,在异相体系(例如乳液、悬浮液和分散液等)中聚合制备的聚合物颗粒具有球形形貌^[13,14].对于如何突破球形界面的限制产生表面粗糙的聚合物微粒这一问题,研究人员已经做了许多尝试.其中,通过引入表面活性剂可以对乳液液滴界面现象进行调节.表面活性剂是一类两亲性物质,在乳液界面起到降低界面张力的作用,从而维持乳液稳定存在^[15,16].在溶剂挥发过程中会导致界面曲率增加、界面面积减小、表面活性剂浓度增加等一系列变化.当乳液液滴界面张力不断降低至一定程度会导致界面面积自发增大,从而使界面发生扰动,这种现象称之为乳液液滴界面不稳定^[17,18].利用此机制制备具有复杂表面微结构的聚合物微球操作简单、可调控性强、各向异性明显;但由于常规的搅拌等方法制备的乳液液滴单分散性较差,因此常常导致生成的微球尺寸差异较大,且该方法仅局限于两亲性嵌段共聚物的体系,而对于疏水性聚合物(均聚物或者共聚物)鲜有报道,这限制了该方法的适用范围.微流控方法可以较好地解决乳液的单分散性问题,微流控技术是指在微尺度(几十至几百微米)范围下对微量流体(10^{-9} — 10^{-18} L)进行系统地整合和操控的技术^[19–21].在材料制备方面,微流控技术举足轻重.通过多样化的材料选择和精巧的流体通道设计,可以对所生成的材料的形貌以及结构进行精确的控制,因此其在复杂形貌及多功能材料制备方面显示出巨大的应用前景^[22,23].基于微流控技术制备的微粒具有粒径均一、尺寸可控、结构复杂、应用广泛等明显优势,但其缺点是需要与其他技术相结合才能对聚合物微粒的复杂结构进行调控.因此,利用微流控技术大量生产具有可控表面粗糙度的功能化聚合物颗粒仍然是一个较大的挑战.

本文将微流控技术与液滴界面不稳定机制相结合,构建了尺寸均一、表面粗糙结构可调的聚苯乙烯微球.如图1所示,首先基于毛细管微流控芯片制备出单分散的水包油乳液液滴,进而研究了疏水的均聚物聚苯乙烯(Polystyrene, PS)在表面活性剂十二烷基硫酸钠(Sodium dodecyl sulfate, SDS)和助表面活性剂 n -十六烷醇(n -hexadecanol, HD)的共同作用下使乳液液滴发生的界面不稳定的过程.接着,通过调控溶剂挥发速率及助表面活性剂的浓度等实验参数实现对微球表面粗糙度的调控.由于可操控表面微结构微球在细胞培养、捕

获等方面具有的良好应用前景,我们将制备得到的微球用于循环肿瘤细胞(circulating tumor cells, CTC)的捕获研究,考察其对提高捕获效率的作用效果.

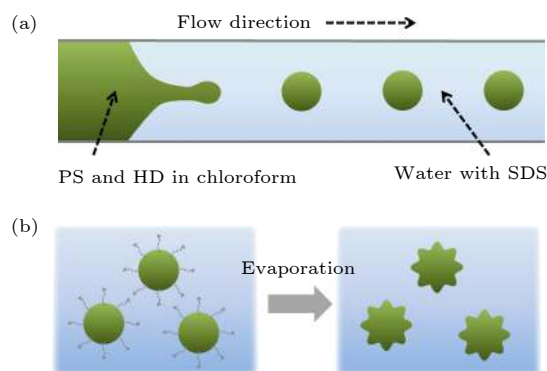


图1 表面粗糙聚苯乙烯微球的制备示意图 (a) 均匀乳液液滴的微流控生成示意图; (b) 具有表面褶皱的固体微球的形成过程示意图,在收集皿中,液滴内氯仿的挥发触发了界面不稳定性现象,随之产生表面的褶皱

Fig. 1. Schematic illustration of the fabrication of poly-styrene microspheres with rough surface: (a) Schematic representation of generating uniform emulsion droplets from microfluidic devices; (b) schematic diagram of the formation process of solidified microspheres with surface wrinkles. In the collection dish, the volatilization of chloroform from droplets triggers the interface instability phenomenon, then the surface wrinkles are generated.

2 实验

2.1 微流控芯片制备乳液液滴

本实验中首先通过玻璃毛细管微流控装置制备单乳液液滴模版,玻璃毛细管的组装按照典型的协流式单乳液通道结构来进行.分散相(内相)溶液由10 mg/mL的聚苯乙烯/氯仿溶液、不同浓度的 n -十六烷醇/氯仿溶液组成;连续相(外相)溶液由3 mg/mL的SDS/甘油溶液组成,选择十二烷基硫酸钠固体粉末,溶于30 vol.%甘油/水溶液中;收集液的组分同连续相溶液.

为产生单分散性的水包油乳液液滴,将一定量的内外相溶液抽取到不同规格的玻璃注射器中并将其分别安放在蠕动泵上,设定内外相流速,启动蠕动泵工作.通过改变内外相流速对生成乳液液滴的尺寸和产率进行调控.

2.2 表面褶皱聚苯乙烯微球的生成与调控

将上述生成的液滴收集在不同体积的SDS/甘

油溶液中,可在显微镜下观察到随着有机溶剂在水相中的扩散、蒸发,乳液液滴界面不稳定性现象被触发,液滴逐渐固化并在表层形成不同程度的褶皱.悬浮在溶液中的微粒首先在去离子水中透析 7 天以去除表面的甘油、SDS 和残余的氯仿,然后在乙醇中透析 7 天以去除 HD.

微球表面褶皱程度的调控可主要从以下两个因素的调节来实现: 1) 在 0—3 mg/mL 范围内改变分散相中 HD 的浓度,可得到从表面平滑到高度褶皱的聚合物微粒; 2) 在 0.5—10.0 mm 范围内改变收集液 SDS/甘油溶液的高度,从而影响有机溶剂的蒸发速率,造成表面形态变化.为了使微球适应于更广泛应用的功能,将上述具有褶皱结构的微球浸泡在四氧化三铁纳米粒子分散液中并保持 3 h 震荡.由于吸附有磁性纳米粒子的聚苯乙烯微球在水中的沉降速度较快,因此采用静置法使微球沉至底部,移除上方溶液后添加超纯水多次洗涤.洗涤完全后,可通过外加磁铁实现微球的快速分离.

2.3 褶皱微球用于循环肿瘤细胞 (CTC) 捕获实验

将干燥的聚苯乙烯褶皱微球样品放入真空等离子清洗机反应腔内,等离子体处理 5 min,使其惰性的表面活化,便于后续抗体的接枝和特异性固定.由于在等离子体处理过程中需要开启真空泵,这可能会导致一部分磁性粒子的损耗,但通过增加磁性纳米粒子的初始吸附数量、合理控制等离子体处理时间等措施,能够减弱这种损耗对整个褶皱微球磁性的影响.接着,将抗上皮细胞黏附分子 (epithelial cell adhesion molecule, EpCAM) 特异性抗体稳定、均一地修饰在聚苯乙烯微球表面.随后,采用外周血 MCF-7 肿瘤细胞作为靶细胞,将褶皱微球添加到 MCF-7 细胞悬液中.为了便于观察和计数,细胞通过钙黄绿素染色.

3 结果与讨论

微流控通道内液滴的生成是在流场作用下流体界面发生形变导致的界面不稳定性增强的结果.无论是协流式、T 形,还是流动聚焦式通道,当分散相和连续相流体相遇时,分散相流体被拉伸.最终,在自由表面失稳条件下,分散相流体断裂而形成液滴,如图 2(a) 所示.通常来讲,液滴的断裂是

黏性力和界面张力相抗衡的结果,前者使界面发生形变而后者抵抗形变.另外,液滴在断裂时对交叉口的阻隔作用会产生一个瞬时增强的上游压力,从而带动下一个液滴的断裂.此外,通道的浸润性也会影响液滴的生成,因为当连续相流体能够优先浸润通道内壁时,连续相流体会在内壁上形成薄膜,从而阻止分散相流体与管壁的直接接触,这样就可以避免三相接触线对整个体系的影响.因此在本工作中,通过选择合理的通道尺寸、流速条件,以及对连续相流体通道进行疏水修饰,保证了液滴能够稳定地生成.如图 2(b) 所示,体式显微镜下观察收集到的乳液液滴直径在 300 μm 左右,尺寸均一,分散度比较好.通过调节内外相流速能够对生成乳液液滴的尺寸和产率进行调控,如图 2(d) 和图 2(e),内相流速越小,外相流速越大,液滴尺寸越小,相应的间隔越大,即生成液滴越稀疏,产率变低.

将液滴收集在含有外相溶液的玻璃皿中,可以明显观察到随着氯仿的蒸发,乳液液滴不断缩小,同时产生“涟漪状”物质,最后聚苯乙烯团聚成微球的形状,呈现黑色,通过光学显微镜追踪此过程如图 3 所示.在有机溶剂氯仿的挥发过程中,乳液液滴变小导致了界面曲率增加、界面面积减小、表面活性剂浓度增加等一系列物理变化.当界面张力降低到一定程度时,液滴表面面积自发增大 (类似于花朵绽放的过程),即触发了乳液液滴界面不稳定性机制.另外,在氯仿挥发过程中,PS 与 HD 之间因为互不相溶而发生相分离,形成非常微小的液滴.随着整个乳液液滴体积收缩,微小液滴逐渐融合,PS 微球慢慢固化过程中小液滴占据的位置成为凹陷孔隙,进一步增加了微球表面的粗糙度.待氯仿彻底挥发后,该结构被固定,最终得到表面具有凹陷孔隙分布和均匀凸出褶皱结构的微球.由于液滴是基于微流控芯片产生的,具有均一的尺寸分布,因此同等条件下经过氯仿挥发后的固化微球也具有较为良好的单分散性,微球直径在 80 μm 左右,如图 2(c) 所示.

我们探究了内相溶液中 HD 的浓度和收集液高度这两个主要因素对微球表面褶皱度的影响.当 HD 的浓度从 0 逐步增加到 3 mg/mL 时,微球表面褶皱度发生了明显的变化.内相溶液中不含 HD,即浓度为 0 mg/mL 时,所得微球表面非常平滑,几乎没有褶皱,其光镜和电镜的表征图像如图 4(a) 和图 4(d) 所示.当 HD 浓度提高到 0.5 mg/mL

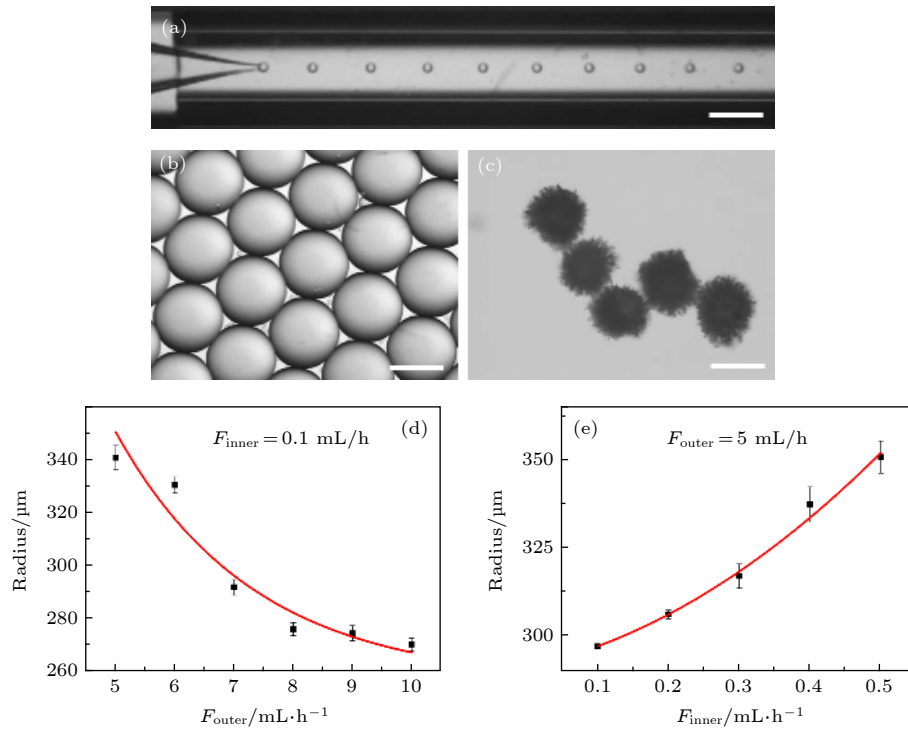


图 2 (a) 微流控通道内液滴的实时生成图像, 比例尺为 $600 \mu\text{m}$; (b), (c) 液滴及其相应 PS 微球的光镜图片, 比例尺分别为 200 和 $80 \mu\text{m}$; (d), (e) 液滴直径和内外相流速之间的曲线关系图: (d) 当内相流速保持在 0.1 mL/h 不变时, 液滴直径与外相流速间呈负相关; (e) 当外相流速保持在 5 mL/h 不变时, 直径与内相流速呈正相关

Fig. 2. (a) Real-time images of droplets generated in the microfluidic device, the scale bar is $600 \mu\text{m}$; (b), (c) microscope photographs of the (b) droplets and (c) the resultant PS microparticles. The scale bars are $200 \mu\text{m}$ in (b) and $80 \mu\text{m}$ in (c), respectively; (d), (e) Relationships of the droplets radius with (d) the same inner flow rate and (e) the same outer flow rate, respectively. Error bars represent standard deviations.

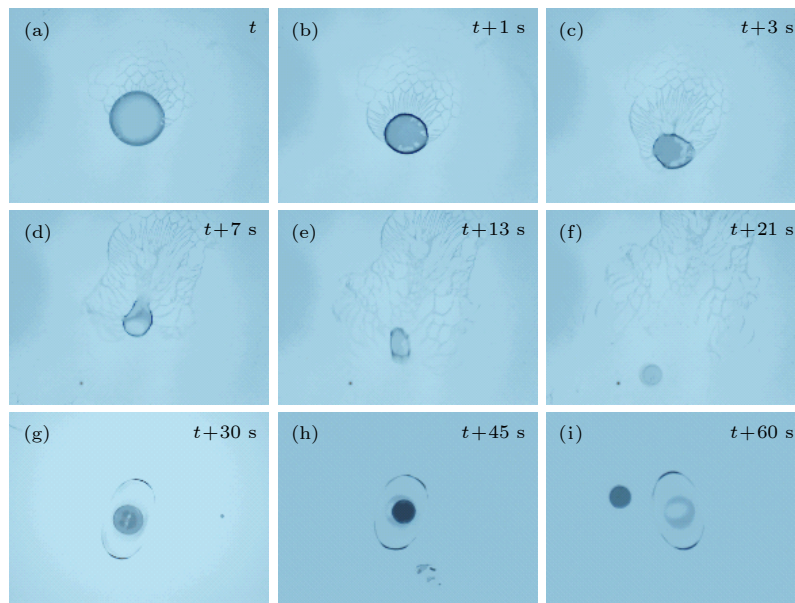


图 3 体式显微镜下观察到的具有褶皱表面的微球的生成过程. 随着氯仿的蒸发, 乳液液滴不断缩小, 同时产生“涟漪状”物质, 最后 PS 玻璃化形成固化微球, 此过程在常温下约持续 1 min . 图 (a)—(f) 的比例尺为 $300 \mu\text{m}$, 图 (g)—(i) 的比例尺为 $250 \mu\text{m}$

Fig. 3. Formation process of microspheres with wrinkled surface observed under the microscope. With the evaporation of chloroform, the droplets continue to shrink, and meanwhile, the "ripple-like" substance is produced. Finally, the PS vitrifies to form solidified microspheres. This process lasts about 1 minute at the room temperature. The scale bars are $300 \mu\text{m}$ in (a)–(f) and $250 \mu\text{m}$ in (g)–(i), respectively.

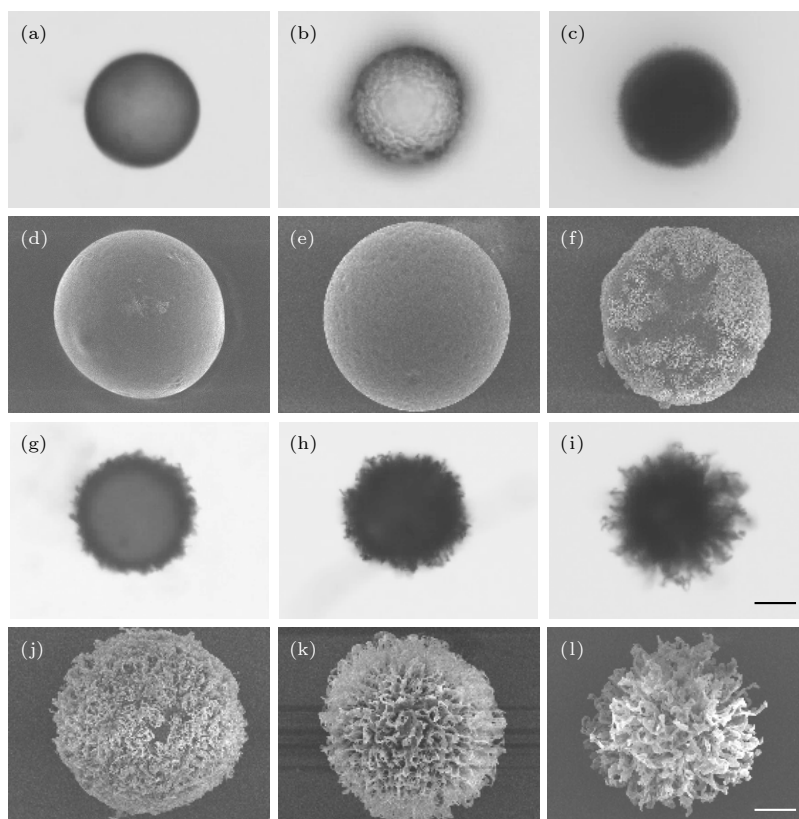


图 4 具有不同表面褶皱度的聚苯乙烯微球的光镜和电镜表征. 这些微球是由含有相同浓度 PS 和不同浓度 HD 的液滴模板固化得到, 其对应的 HD 浓度如下: (a), (d) 0 mg/mL; (b), (e) 0.5 mg/mL; (c), (f) 1 mg/mL; (g), (j) 1.5 mg/mL; (h), (k) 2 mg/mL; (i), (l) 3 mg/mL. 光镜图片 ((a)—(c), (g)—(i)) 和电镜图片 ((d)—(f), (j)—(l)) 的比例尺分别为 45 和 30 μm

Fig. 4. The optical microscopy and SEM images of the PS microparticles with different surface-roughness. They are obtained from microfluidic droplet templates containing 10 mg/mL PS and varied concentrations of HD: (a), (d) $c = 0$ mg/mL; (b), (e) $c = 0.5$ mg/mL; (c), (f) $c = 1$ mg/mL; (g), (j) $c = 1.5$ mg/mL; (h), (k) $c = 2$ mg/mL; and (i), (l) $c = 3$ mg/mL. The scale bars are 45 μm in ((a)—(c), (g)—(i)) and 30 μm in ((d)—(f), (j)—(l)), respectively.

时表面出现局部不平滑的区域 (图 4(b) 和图 4(e)), 具体表现为微球表面形成了凹陷孔隙的分布. 如图 4(c), 图 4(f) — 图 4(h), 图 4(j), 和图 4(k) 显示, 从 1 mg/mL 开始微球表面出现均匀褶皱, 在 1—3 mg/mL 之间, 随着 HD 浓度的增加, 微球表面伸出的粗糙结构长度不断增加. 当浓度达到 3 mg/mL 时, 球体核心变得非常小, 被大幅度褶皱覆盖. 表面活性剂 HD 的浓度对表面形貌的影响可归结于 HD 分子在油/水界面处的累积和重排, 且进一步渗透到 SDS 单层中诱导 SDS 的组装, 两个过程都共同导致了乳液液滴界面张力的降低, 从而引起界面不稳定现象的发生, 产生了伸展出表面的褶皱.

另外, 当收集液 SDS/甘油的液面高度在 0.5—10.0 mm 范围内变化时, 保持容器底面直径固定, 随着收集液体积增加, 褶皱度逐渐降低, 这与 HD 的浓度变化的影响相反. 由此可见, 水层高

度明显影响有机溶剂的挥发速率, 从而可以通过调节挥发速率的快慢制备各种形态的聚合物粗糙微球. 综合来看, 乳液液滴挥发触发的界面失稳过程和液滴的相变是两个竞争性的过程, 当溶剂挥发非常快时, 乳液液滴中的界面张力降低占主导作用, 首先发生了界面不稳定现象, 然后固化形成粗糙微球; 当溶剂挥发非常慢时, 液滴在没有达到很低的界面张力时就已经固化, 形成了较光滑的微球.

由于微球表面褶皱之间的间隙较小, 磁性纳米粒子能够被吸附在间隙中并保持稳定. 宏观磁吸实验可以证明, 微球整体可以在磁场的作用下进行可控运动. 将一批表面载有磁性纳米粒子的褶皱微球装在透明玻璃瓶内, 在永磁铁的作用下, 可以观察到微球被吸引到靠近磁铁的瓶壁一侧, 如图 5 所示. 因此可以证明, 所制备的微球具有磁控运动特性, 这为后续的手机捕获和富集带来了极大便利.

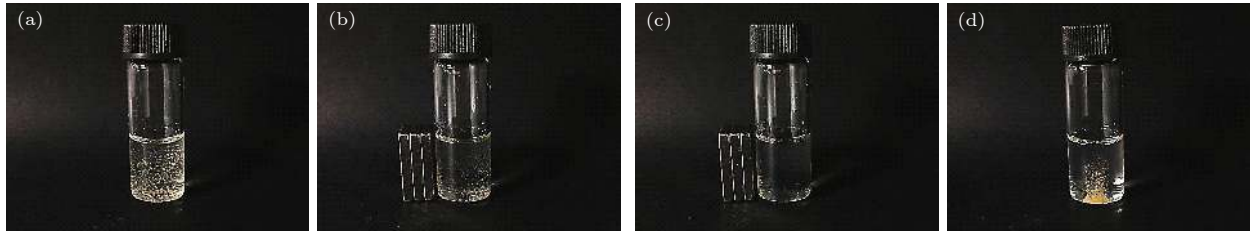


图 5 表面覆盖有四氧化三铁纳米粒子的聚苯乙烯褶皱微球被赋予较强的磁响应性. 在永磁铁的引导下, 微球被吸引到靠近磁铁的瓶壁一侧. 图中比例尺为 1.5 cm

Fig. 5. Strong magnetic responsiveness of the highly textured PS microparticles coated with Fe_3O_4 nanoparticles. Under the guidance of the permanent magnet, the microspheres are attracted to the side of the bottle wall near the magnet. The scale bar is 1.5 cm.

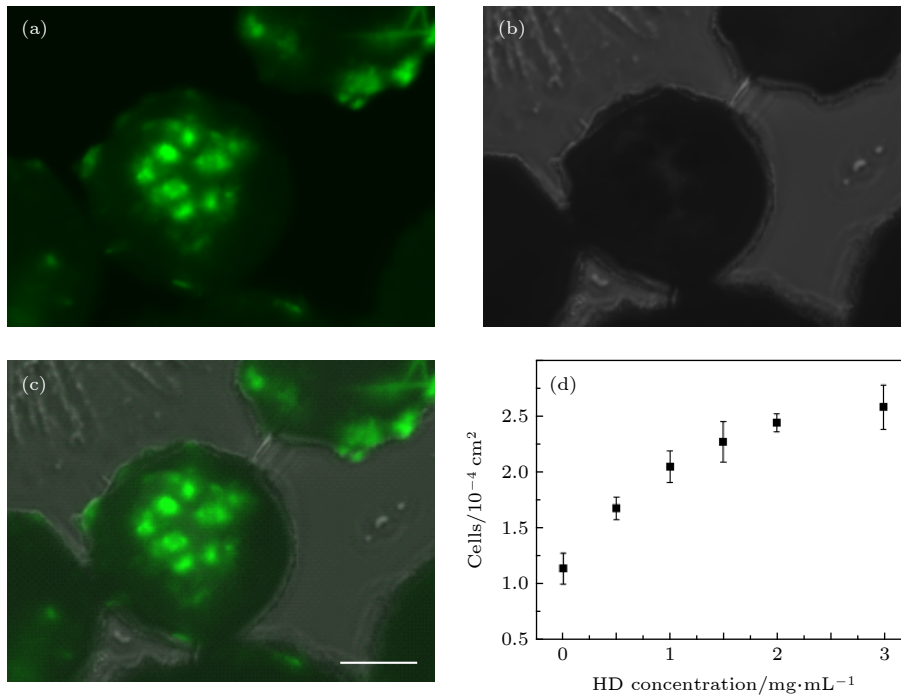


图 6 被捕获在微球表面的循环肿瘤细胞的共聚焦激光扫描显微镜图像 (a) 横截面荧光图像; (b) 明场图像; (c) (a) 和 (b) 合并后的图像; (d) 细胞捕获数量与微球表面褶皱度的关系

Fig. 6. The CLSM images of CTC captured onto the surface of the PS microparticles: (a) The cross-section fluorescent images; (b) the white-light images, and (c) the merged images. The scale bar is 30 μm ; (d) the relationships of the number of captured cells with the extents of surface wrinkles.

由于粗糙表面有利于细胞黏附、传播和生长, 我们探究了这种表面褶皱微球的肿瘤循环细胞捕获应用. 长期以来, 临床在癌症的诊断与治疗中, 通常是利用组织细胞活检来进行确诊并跟踪治疗效果, 这种方法不仅会给患者带来身体上的创伤, 而且价格昂贵. 诊断癌症更好的方法是对游离在血液中的肿瘤细胞, 即循环肿瘤细胞进行检测. CTC 被认为是癌症转移的一个重要因素, 但是它在血液循环系统中的含量极少, 一般 1 mL 血液中只有 1—10 个, 因此急需高效的捕获和富集载体^[24,25]. 我们将表面修饰特异性抗体的褶皱微球加入到细

胞悬液中并进行染色观察, 如图 6(a)—图 6(c), 共聚焦激光扫描显微镜 (confocal laser scanning microscope, CLSM) 图像显示钙黄绿素染色的 MCF-7 细胞附着在褶皱表面, 且发出明亮的绿色荧光. 为了获得更高的捕获效率, 我们还探索了细胞数量与褶皱程度之间的关系. 为此, 将具有相同浓度抗体, 但褶皱度不同的微球与相同浓度的细胞一起孵育, 根据 CLSM 图像计算并分析捕获的细胞. 图 6(d) 显示在光滑聚苯乙烯微球表面捕获的细胞数量较少, 而随着褶皱程度的增加, 被捕获的数量相应增多.

4 结 论

提出了一种具有表面粗糙微结构的聚苯乙烯微球的可控制备方法. 采用微流控技术制备出单分散的乳液液滴, 原位观测乳液液滴的界面不稳定现象及微球的形成过程, 并对其形成机制进行了分析, 提出了界面不稳定现象发生的机理. 表面活性剂 SDS 及助表面活性剂 HD 的加入使得溶剂挥发过程中界面张力降低而触发乳液液滴界面不稳定, 最终得到表面粗糙结构的聚合物微球. 改变 HD 的浓度以及有机溶剂的挥发速率可以有效调控微球表面的粗糙程度. 制备的粗糙微球还可以吸附磁性纳米粒子、包覆荧光染料等, 且其粗糙褶皱结构有利于增加与细胞间的黏附作用, 因此我们进一步探究了微球在捕获外周血中 CTC 的作用效果. 我们发现, 微球捕获富集细胞的能力与其表面褶皱程度呈现正相关, 因此能够通过调控粗糙度实现 CTC 捕获效率的精确控制. 此外, 这种用于制备具有高度褶皱表面颗粒的通用方法还可以扩展到其他类型的疏水性聚合物, 其将在药物递送、生化分析、组织工程、催化涂覆和装置制造领域中有重要的潜在应用.

参考文献

- [1] Jo Y K, Lee D 2019 *Small* **16** 1903736
- [2] Wang H, Liu Y X, Chen Z Y, Sun L Y, Zhao Y J 2020 *Sci. Adv.* **6** eaay1438
- [3] Shang L R, Cheng Y, Zhao Y J 2017 *Chem. Rev.* **117** 7964
- [4] Iwai Y, Uchida Y, Nishiyama N 2016 *Adv. Opt. Mater.* **4** 1961
- [5] Visaveliya N, Köhler J M 2015 *J. Mater. Chem. C* **3** 844
- [6] Zhang X X, Zhao Y J 2019 *Sci. Bull.* **64** 1472
- [7] Zhang H, Liu Y X, Wang J, Shao C M, Zhao Y J 2019 *Sci. China Chem.* **62** 87
- [8] Koh L B, Rodriguez I, Venkatraman S S 2010 *Biomaterials* **31** 1533
- [9] Scopelliti P E, Borgonovo A, Indrieri M 2010 *PloS One* **5** e11862
- [10] Campos E, Branquinho J, Carreira A S, Carvalho A, Coimbra P, Ferreira P, Gil M H 2013 *Eur. Polym. J.* **49** 2005
- [11] Wang Y T, Shang L R, Chen G P, Shao C M, Liu Y X, Lu P H, Rong F, Zhao Y J 2018 *Appl. Mater. Today* **13** 303
- [12] Wang B J, Pepijn P, Wang H Z, Bai Z H, Wang H L, Luque R, Xuan J 2017 *Chem. Soc. Rev.* **46** 855
- [13] Zhao C X, Middelberg A P J 2013 *RSC Adv.* **3** 21227
- [14] Hao N, Nie Y, Zhang J X J 2018 *Int. Mater. Rev.* **63** 461
- [15] Liu S Q, Deng R H, Li W K, Zhu J T 2012 *Adv. Funct. Mater.* **22** 1692
- [16] Hussain M, Xie J, Hou Z 2017 *ACS Appl. Mater. Interfaces* **9** 14391
- [17] Zhao K G, Xue C, Wang L F, Ye W H, Wu J F, Ding Y K, Zhang W Y, He X T 2018 *Acta Phys. Sin.* **67** 214209 (in Chinese) [赵凯歌, 薛创, 王立锋, 叶文华, 吴俊峰, 丁永坤, 张维岩, 贺贤士 2018 *物理学报* **67** 214209]
- [18] Xia T J, Dong Y Q, Cao Y G 2018 *Acta Phys. Sin.* **62** 214702 (in Chinese) [夏同军, 董永强, 曹义刚 2018 *物理学报* **62** 214702]
- [19] Li D Y, Zhang Y X, Ou Y X, Pu X Y 2019 *Acta Phys. Sin.* **68** 054203 (in Chinese) [李东阳, 张远宪, 欧永雄, 普小云 2019 *物理学报* **68** 054203]
- [20] Chen Q, Qi X B, Chen S F, Liu M F, Pan D W, Li B, Zhang Z W 2017 *Acta Phys. Sin.* **66** 046801 (in Chinese) [陈强, 漆小波, 陈素芬, 刘梅芳, 潘大伟, 李波, 张占文 2017 *物理学报* **66** 046801]
- [21] Yu Y R, Guo J H, Sun L Y, Zhang X X, Zhao Y J 2019 *Research* **2019** 1
- [22] Macosko E Z, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas A R, Kamitaki N, Martersteck E M 2015 *Cell* **161** 1202
- [23] Zhu P, Wang L 2017 *Lab Chip* **17** 34
- [24] Li Y Q, Chandran B K, Lim C T 2015 *Adv. Sci.* **2** 1500118
- [25] Luan C X, Wang H, Han Q, Ma X Y, Zhang D G, Xu Y S, Chen B A, Li M L, Zhao Y J 2018 *ACS Appl. Mater. Interfaces* **10** 21206

SPECIAL TOPIC—Statistical physics and complex systems

Surface-textured polymer microspheres generated through interfacial instabilities of microfluidic droplets for cell capture*

Wang Yue-Tong¹⁾ Shang Luo-Ran^{2)†} Zhao Yuan-Jin^{1)‡}

1) (*State Key Laboratory of Bioelectronics, School of Biological Science and Medical Engineering, Southeast University, Nanjing 210096, China*)

2) (*Zhongshan-Xuhui Hospital, Institutes of Biomedical Sciences, Fudan University, Shanghai 200032, China*)

(Received 12 March 2020; revised manuscript received 25 March 2020)

Abstract

Polymer microparticles with various compositions and morphologies have recently received much attention. Their surface-roughness significantly affects the physical and chemical properties, which especially counts in regulating the interaction between biological materials and living systems. In this paper, we design a polystyrene microsphere with controllable surface textures. At first, a microfluidic device is used to generate droplets with uniform size containing the hydrophobic polymer and a co-surfactant. During the volatilization of the organic solvent, the shrinking droplets appear to be unstable at the interface. Thus, the surface area increases spontaneously, and microspheres with wrinkles on the surface are obtained after being solidified. The results show that tuning the concentration of the co-surfactant and the rate of solvent evaporation can effectively regulate the surface roughness of the microspheres. Circulating tumor cell capture experiments reveal that this textured structure can facilitate the cell adhesion and increase the number of the captured cells. These features indicate that the coarse microspheres possess a promising application prospect in the field of biomedical analysis.

Keywords: interfacial instability, microfluidics, surface-texture, microspheres, cell capture

PACS: 47.20.Ib, 61.30.Pq, 83.10.Tv, 87.17.Uv

DOI: 10.7498/aps.69.20200362

* Project supported by the National Natural Science Foundation of China (Grant Nos. 61927805, 51522302), the NSAF Foundation of China (Grant No. U1530260), the Natural Science Foundation of Jiangsu Province, China (Grant No. BE2018707), and the Launching Research Funds from Fudan University and Zhongshan-Xuhui Hospital, China (Grant No. JIH1340032).

† Corresponding author. E-mail: luoranshang@fudan.edu.cn

‡ Corresponding author. E-mail: yjzhao@seu.edu.cn

基于香蕉形液晶分子自组装的纳米螺旋丝有机凝胶及其流变特性

王行政 杨晨静 蔡历恒 陈东

The rheology property of organogels based on 3D helical nanofilament bnetworks self-assembled by bent-core liquid crystals

Wang Xing-Zheng Yang Chen-Jing Cai Li-Heng Chen Dong

引用信息 Citation: *Acta Physica Sinica*, 69, 086102 (2020) DOI: 10.7498/aps.69.20200332

在线阅读 View online: <https://doi.org/10.7498/aps.69.20200332>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

基于液晶/聚合物光栅的高转化效率有机半导体激光器

Enhancement of conversion efficiency for an organic semiconductor laser based on a holographic polymer dispersed liquid crystal
物理学报. 2017, 66(24): 244204 <https://doi.org/10.7498/aps.66.244204>

利用银纳米立方增强效率的多层溶液加工白光有机发光二极管

Efficiency improvement in solution-processed multilayered phosphorescent white organic light emitting diodes by silica coated silver nanocubes
物理学报. 2020, 69(4): 047801 <https://doi.org/10.7498/aps.69.20191526>

金纳米四面体增强有机太阳能电池光吸收及光伏性能研究

Enhanced light absorption and device performances of organic photovoltaic devices with Au tetrahedra nanoparticles
物理学报. 2018, 67(24): 247201 <https://doi.org/10.7498/aps.67.20181502>

碳纳米管/二硒化钼有机玻璃的非线性吸收、非线性散射和光限幅特性

Nonlinear absorption, nonlinear scattering, and optical limiting properties of carbon nanotube/molybdenum diselenide organic glass
物理学报. 2018, 67(23): 236201 <https://doi.org/10.7498/aps.67.20181583>

以石墨烯为电极的有机噻吩分子整流器的设计及电输运特性研究

Design and electronic transport properties of organic thiophene molecular rectifier with the graphene electrodes
物理学报. 2017, 66(9): 098501 <https://doi.org/10.7498/aps.66.098501>

专题：统计物理和复杂系统

基于香蕉形液晶分子自组装的纳米螺旋丝 有机凝胶及其流变特性*

王行政^{1) #} 杨晨静^{1) #} 蔡历恒^{2) †} 陈东^{1) ‡}

1) (浙江大学能源工程学院化工机械研究所, 杭州 310027)

2) (弗吉尼亚大学材料科学工程学院, 夏洛茨维尔 22904, 美国)

(2020年3月4日收到; 2020年3月22日收到修改稿)

在香蕉形液晶分子 B4 相态中, 非手性香蕉形液晶分子自组装形成层状结构, 分子在层内倾斜, 形成层手性和自发极化, 并且造成层内不匹配, 最终形成纳米螺旋丝. 本文设计了 NOBOW/十六烷混合体系, 在高温时, 香蕉形液晶分子溶解于十六烷, 在低温时, 香蕉形液晶分子自组装形成纳米螺旋丝, 并最终形成三维网络, 变成有机凝胶. 为深入理解纳米螺旋丝有机凝胶的特性, 拓展其在软物质领域的应用, 本文通过流变实验对该有机凝胶的黏弹性质进行了系统研究. 实验表明纳米螺旋丝有机凝胶与传统凝胶不同, 纳米螺旋丝有机凝胶可以随温度变化形成凝胶-流体的可逆变化, 并且通过测量 NOBOW/十六烷混合体系在不同液晶分子浓度、温度、应变大小和应变速率下的流变特征, 揭示了该有机凝胶的流变特性与纳米螺旋丝的性质密切相关.

关键词: 香蕉形液晶, 螺旋纳米丝, 有机凝胶, 黏弹性

PACS: 61.30.-v, 77.84.Nh

DOI: 10.7498/aps.69.20200332

1 引言

液晶在显示、光学器件、功能材料等领域具有广泛的应用^[1-7]. 传统液晶分子一般为棍状, 主要由棍状的刚性部分和柔性碳链组成. 而香蕉形液晶分子主要由弯曲状的刚性部分和连接于其两端的柔性碳链组成. 由于香蕉形液晶分子特殊的几何形状, 可以自组装形成多种新颖的相态结构, 表现出分子弯曲、层内倾斜和分子层手性的相互耦合作用^[8,9]. 人们对香蕉形液晶分子的重点研究开始于 1996 年, 当时人们首次在非手性香蕉形液晶分子中发现了铁电性^[10]. 在此之前, 铁电性只存在于手

性棍状液晶分子中. 与棍状液晶不同, 香蕉形液晶分子本身是非手性的, 但是当其自组装形成层状结构, 分子在层内倾斜, 破坏了体系的镜像对称, 形成层手性和自发极化^[11]. 香蕉形液晶分子与棍状液晶分子的另一个重要区别是香蕉形液晶分子的层状结构存在层内不匹配. 当香蕉形液晶分子在层内倾斜, 其上下两臂的倾斜方向接近相互垂直, 因此上下两臂的扩张方向也相互垂直, 造成层内的不匹配. 为了消除这种层内不匹配, 香蕉形液晶分子的层状结构倾向于形成马鞍状弹性形变^[10,12,13].

在 20 多年的研究里, 人们先后在香蕉形液晶分子中发现 B1, B2, B3, B4, B5, B6, B7, B8 等多个新型相态^[14]. 其中在 B4 相态中, 香蕉形液晶分

* 国家自然科学基金 (批准号: 11704331, 21878258) 和稀土资源利用国家重点实验室开放课题基金 (批准号: RERU2019008) 资助的课题.

共同第一作者.

† 通信作者. E-mail: liheng.cai@virginia.edu

‡ 通信作者. E-mail: chen_dong@zju.edu.cn

子的层手性进一步形成微观的纳米螺旋丝手性结构和宏观的左旋和右旋手性区域. 已有研究表明纳米螺旋丝主要通过成核和保持手性生长, 最终形成宏观的手性区域, 即同一手性区域由同一手性纳米螺旋丝组成, 而左旋和右旋手性区域出现的概率是一样的^[15]. 这是首个由非手性分子形成手性结构的相态, 纳米螺旋丝也因其特有的手性结构引起了人们的广泛研究^[16-21], 其中包括纳米螺旋丝结构、非线性光学、手性限域效应等^[17,22]. 此前纳米螺旋丝与棍状液晶混合体系的研究表明, 在混合体系中, 香蕉形液晶分子可以通过相分离, 自组装形成纳米螺旋丝三维网络, 进而影响填充于其间的棍状液晶排列, 形成非对映区域, 在手性分离、非对称合成等领域展现出潜在的应用前景^[23]. 然而由纳米螺旋丝形成的有机凝胶尚缺乏深入的研究, 对其深入的研究可以拓展其在软物质领域的应用^[24].

本文设计了 NOBOW/十六烷混合体系, 制备由纳米螺旋丝形成的有机凝胶, 并通过有机凝胶流变性质的研究, 其中包括可逆的凝胶-流体相变特征和体系储能模量和损耗模量随液晶分子浓度、温度、应变大小和应变速率的变化特征, 进而揭示纳米螺旋丝性质对有机凝胶性质的影响规律, 为其在手性分离等领域的应用奠定基础.

2 结果与讨论

NOBOW 和十六烷的化学结构如图 1(a) 和图 1(b) 所示. NOBOW 是典型的香蕉形液晶分子, 由弯曲的刚性核心和连接于其两端的柔性碳链组成, 具有 Iso (175 °C) B2 (147 °C) B3 (143 °C) B4 相变行为. 当温度从高温逐渐降低, 在焓 (弯曲刚性核心趋向于有序排列) 和熵 (柔性碳链趋向于无序排列) 的共同驱动下, 香蕉形液晶分子自组装形成层状结构, 分子在层内倾斜, 形成层手性和自发极化, 即 B2 相态, 如图 1(c) 和图 1(d) 所示. 此时, 极化方向 p 同时垂直于指向矢 n 和层法线 z , 即垂直于指向矢 n 和层法线 z 所在的平面. 由于香蕉形液晶分子上下两臂的倾斜方向相互垂直, 对应的扩张方向也相互垂直, 造成层内不匹配, 如图 1(e) 所示. 为了消除这种层内不匹配, 香蕉形液晶分子的层状结构会自发形成马鞍状弹性形变. 当这种形变足够强时, 液晶分子可以自组装形成双连续相, 即 DC 相态 (如图 1(f) 所示), 或纳米螺旋丝相, 即

B4 相态 (如图 1(g) 所示). 而 NOBOW 液晶分子则形成纳米螺旋丝相.

在 NOBOW 和十六烷混合体系中, 十六烷作为 NOBOW 的有机溶剂, 体系随温度变化可以形成可逆的凝胶-流体转变, 如图 2(a) 所示. 在高温时, NOBOW 溶解于十六烷, 形成无序相. 因为十六烷的存在, NOBOW 的相变温度降低, 在冷却过程中, NOBOW 液晶分子直接从无序相自组装形成纳米螺旋丝, 并且在生长过程中, 通过分叉形成三维网络^[23], 体系进而形成有机凝胶. 当温度升高时, 纳米螺旋丝融化, 凝胶再次转变为流体, 形成凝胶-流体的可逆变化. 有机凝胶中, 香蕉形液晶分子自组装形成的纳米螺旋丝可以直接通过速冻断层电子显微镜 (FFTEM) 和透射电子显微镜 (TEM) 观察到, 如图 2(b) 和图 2(c) 所示. 已有 NOBOW/SCB(SCB, 一种棍状液晶材料) 混合体系的研究表明, 纳米螺旋丝具有基本固定的宽度 (~37 nm) 和螺距 (~220 nm), 不依赖于 NOBOW 的浓度和体系的温度^[15], 这主要是因为 NOBOW 分子自组装形成纳米螺旋丝是一个相分离过程. 在 NOBOW/十六烷混合体系中, 纳米螺旋丝的宽度约为 37 nm, 螺距约为 230 nm, 基本保持不变.

如前所述, 为降低层内不匹配, 香蕉形液晶分子的层状结构会自发形成马鞍状弹性形变, 最终形成纳米螺旋丝, 其弹性形变自由能可以表达为 $f_E = \bar{K}(\sigma - \sigma_0)^2$, 其中 \bar{K} 是弗兰克 (Frank) 高斯弹性形变常数, $\sigma_0 = 1/R_p = G/\bar{K}$ 是最优曲率. 但是在层间距保持不变的条件下, 纳米螺旋丝不能处处满足最优曲率, 其曲率从螺旋轴中心往两边逐渐减小, 可表达为 $\sigma = q/(1 + q^2 s^2)$, 其中 $p = 2\pi/q$ 是纳米螺旋丝的螺距, s 是距纳米螺旋丝中心轴的距离. 因此, 纳米螺旋丝内部仍然存在层内不匹配, 该层内不匹配将造成应力的存在. 其应力大小与偏离最优曲率的大小成正比, 所以应力从螺旋轴中心往两边逐渐增加, 对应颜色逐渐加深, 如图 2(c) 中的插图所示. 这与传统高分子形成的有机凝胶有很大的区别, 也促使我们进一步研究纳米螺旋丝有机凝胶的流变性质.

纳米螺旋丝有机凝胶的储能模量 (G') 和损耗模量 (G'') 由马尔文旋转流变仪测量所得. 首先, 调节流变仪温度从 140 °C 下降到 20 °C, 每下降 5 °C, 测量一次 NOBOW/十六烷混合体系在 0.5% 应变和 1 Hz 应变速率条件下的储能模量和损耗模量.

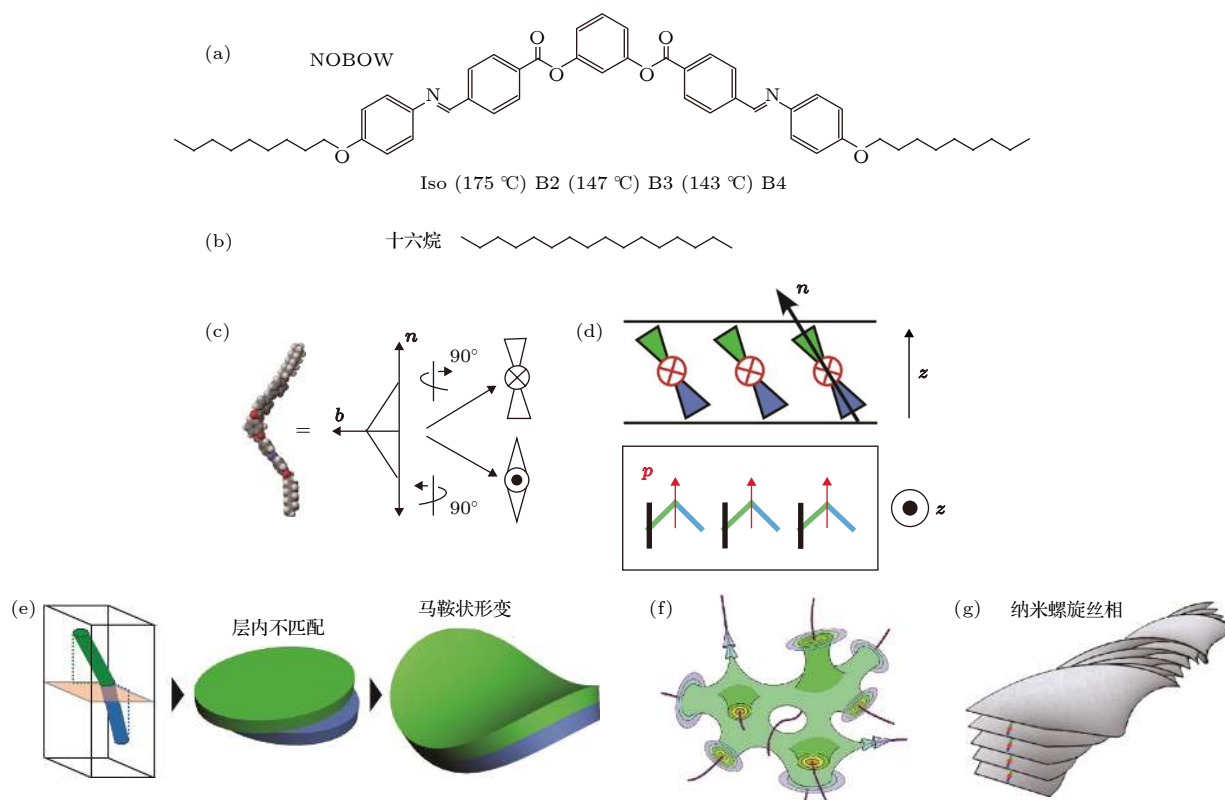


图 1 香蕉形液晶分子的自组装结构 (a) NOBOW 和 (b) 十六烷的化学结构; (c) 香蕉形液晶分子的指向矢 \mathbf{n} 沿长轴方向, 极化 \mathbf{p} 沿弓形方向; (d) 非手性香蕉形液晶分子自组装形成层状结构, 分子在层内倾斜, 形成层手性和极化, 极化方向 \mathbf{p} 同时垂直于指向矢 \mathbf{n} 和层法线 \mathbf{z} , 即垂直于指向矢 \mathbf{n} 和层法线 \mathbf{z} 所在的平面; (e) 香蕉形液晶分子上下两臂的倾斜方向相互垂直, 对应的扩张方向也相互垂直, 造成层内不匹配. 为了消除这种层内不匹配, 香蕉形液晶分子的层状结构会自发形成马鞍状弹性形变. 当马鞍状弹性形变足够大时, 液晶分子可以自组装形成 (f) 双连续相或 (g) 纳米螺旋丝相

Fig. 1. Self-assembly of bent-core liquid crystal. Chemical structures of (a) NOBOW and (b) hexadecane molecules; (c) the director, \mathbf{n} , which is along the molecular long axis, and the polarization, \mathbf{p} , which is along the bow direction; (d) smectic layers of tilted bent-core molecules. The molecules are tilted from the layer normal \mathbf{z} , and the macroscopic polarization \mathbf{p} , is orthogonal to both \mathbf{n} and \mathbf{z} ; (e) the tilt directions of the top and bottom molecular arms of bent-core liquid crystals are essentially orthogonal to each other. If each molecular arm is regarded as an elastic slab, the two elastic slabs dilate along their local molecular tilt directions, which are orthogonal to each other, thus resulting in internal structural mismatch. The internal structural mismatch could be released by the formation of saddle-splay deformations; (f) bicontinuous smectic layer of the dark conglomerate (DC) phase, in which smectic layers organize into disordered focal conics; (g) helical nanofilaments of the B4 phase, in which smectic layers of finite width form uniformly twisted ribbons.

随后, 在 1 Hz 应变速率条件下, 分别测量 NOBOW/十六烷混合体系在 20 °C, 60 °C 和 90 °C 的储能模量和损耗模量随应变大小变化的规律. 最后, 在 0.5% 应变条件下, 分别测量 NOBOW/十六烷混合体系在 20, 90, 100 和 120 °C 的储能模量和损耗模量随应变速率变化的规律.

在 $c = 0.25\%$ NOBOW/十六烷混合物体系中, 由于香蕉形液晶分子浓度太低, 纳米螺旋丝仅形成团簇, 而无法形成相互连通的三维网络, 如图 3(a) 插图所示. 该体系的储能模量 (G') 和损耗模量 (G'') 均很小, 且损耗模量略大于储能模量, 表现为典型的流体^[25,26], 如图 3(a) 所示. 当 NOBOW

的浓度增加到 $c = 0.5 \text{ wt.}\%$ 时, 在高温流体状态下, 体系储能模量和损耗模量均很小, 与溶剂十六烷本身相近. 随着温度从高温逐渐降低到 99 °C 时, 香蕉形液晶分子自组装形成纳米螺旋丝网络, 体系成为凝胶, 其储能模量开始显著大于损耗模量, 表现出典型的凝胶特性, 如图 3(b) 所示. 随着温度进一步下降, 体系储能模量逐渐增加. 已有 NMR 对纳米螺旋丝相态的研究表明, 随着温度降低, 液晶分子柔性碳链的热运动逐渐减小, 分子排列更有序^[27], 这很可能是纳米螺旋丝刚性随温度降低逐渐增强的原因. $c = 1 \text{ wt.}\%$ NOBOW/十六烷混合体系储能模量 (G') 和损耗模量 (G'') 随温

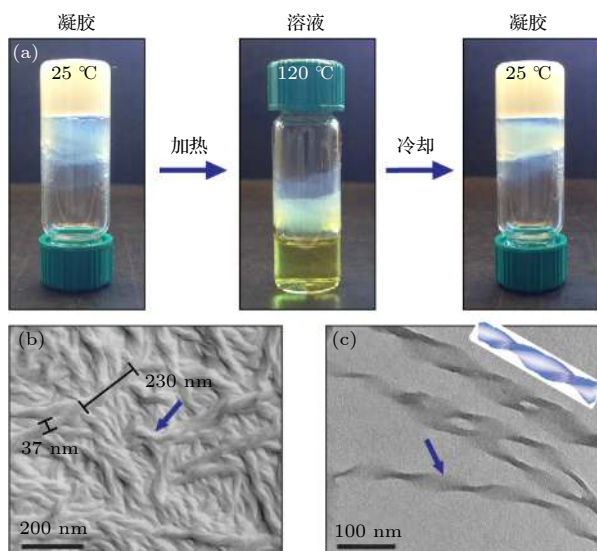


图 2 基于香蕉形液晶分子自组装形成纳米螺旋丝网络的有机凝胶 (a) 在 120 °C 高温下, 1 wt.% NOBOW 香蕉形液晶分子可以溶解于十六烷, 体系呈透明流体; 在室温 25 °C 下, 香蕉形液晶分子自组装形成纳米螺旋丝网络, 体系变成有机凝胶. 该流体-凝胶转变可通过加热和冷却实现可逆变化; (b) 有机凝胶中纳米螺旋丝网络的速冻断层电子显微镜 (FFTEM) 图像; (c) 纳米螺旋丝的透射电子显微镜 (TEM) 图像, 插图展示了纳米螺旋丝的应力分布, 应力从螺旋轴中心往两边逐渐增加, 对应颜色逐渐加深

Fig. 2. Organogels formed by self-assembled 3D helical nanofilament networks of bent-core liquid crystals: (a) At 25 °C, bent-core liquid crystal molecules self-assemble into helical nanofilament networks and the 1 wt.% NOBOW/hexadecane mixture forms organogels. At 120 °C, NOBOW dissolves in the hexadecane solvent and the system appears as transparent fluid, showing reversible gel-sol transitions upon heating and cooling; (b) freeze-fracture transmission electron microscopy (FFTEM) image of helical nanofilament networks in the organogels; (c) transmission electron microscopy (TEM) image of individual helical nano-filaments. The inset shows the distribution of internal stress in the helical nanofilaments. The internal stress gradually increases from the helical axis towards the edges, accompanied by the color increase.

度的变化行为与 $c = 0.5$ wt.% 体系类似, 但是由于香蕉形液晶分子浓度的增加, 流体-凝胶转变温度提高到 107 °C, 储能模量大小和储能模量随温度降低的增加速率也相应增加. 例如, 在 $T = 20$ °C 时, 纳米螺旋丝的储能模量从 $G' \approx 550$ Pa 增加到 $G' \approx 1200$ Pa, 差不多翻倍.

为了进一步研究纳米螺旋丝三维网络的力学性能, 我们在不同的条件下进行了一系列的流变实验. 对于 $c = 0.25$ wt.% NOBOW/十六烷混合体系, 其储能模量 (G') 和损耗模量 (G'') 都很小, 呈现流体特性, 如图 4(a) 所示. 对于 $c = 0.5$ wt.% 和

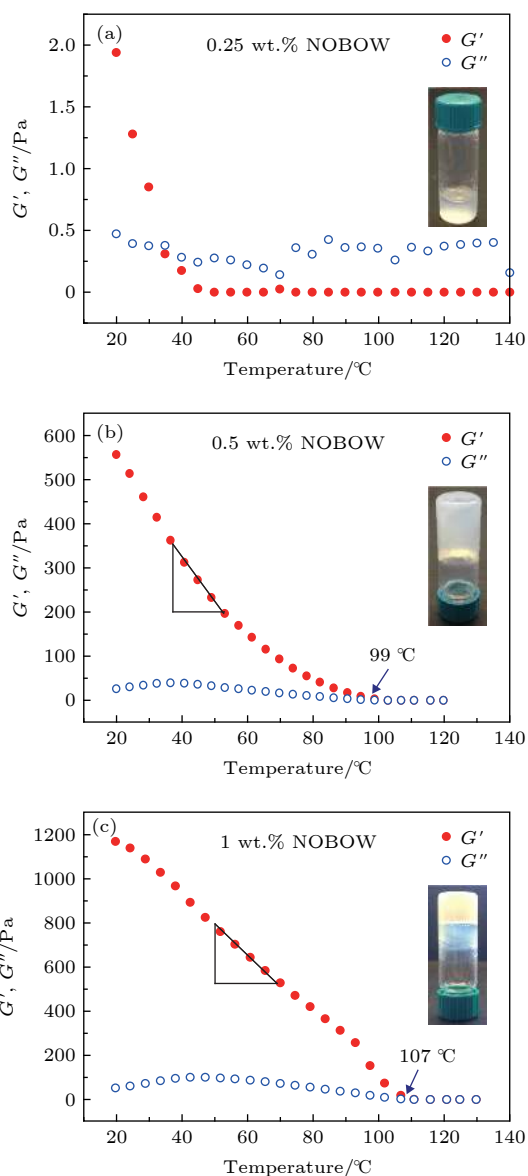


图 3 NOBOW/十六烷混合体系储能模量 (G') 和损耗模量 (G'') 随温度变化的关系 (a) 当 NOBOW 的质量分数为 0.25 wt.% 时, 体系不能形成凝胶, 损耗模量略大于储能模量; (b), (c) 当 NOBOW 的质量分数分别为 0.5 wt.% 和 1 wt.% 且温度分别低于 99 和 107 °C 时, 香蕉形液晶分子自组装形成纳米螺旋丝网络, 体系储能模量显著升高. 0.5 wt.% 体系储能模量增加的速率为 10.3 Pa/°C, 而 1 wt.% 体系储能模量增加的速率为 14.2 Pa/°C. 实验中, 应变保持 0.5%, 应变速率保持 1 Hz

Fig. 3. Temperature sweeps of the elastic (G' , solid symbols) and viscous moduli (G'' , open symbols) of NOBOW/hexadecane mixtures: (a) 0.25 wt.% NOBOW/hexadecane mixture could not form organogels and its G'' is larger than G' ; (b), (c) 0.5 wt.% and 1 wt.% NOBOW/hexadecane mixtures could self-assemble into helical nanofilament networks, forming organogels. The increase rate of the elastic modulus of 0.5 wt.% NOBOW/hexadecane is 10.3 Pa/°C and that of 1 wt.% NOBOW/hexadecane is 14.2 Pa/°C. The measurements are taken at a fixed strain of 0.5% and a fixed strain rate of 1 Hz.

$c = 1$ wt.% NOBOW/十六烷混合体系, 在 $T = 20$ °C, 应变 $\gamma > 1\%$ 时, 其储能模量随着应变增加逐渐减小, 表明有机凝胶中的纳米螺旋丝不能抵抗相应的应变, 开始被破坏 (如图 4(b) 和图 4(c) 所示); 当应变 $\gamma \approx 10\%$ 时 (如图 4(b) 和图 4(c) 中箭头所示), 储能模量 (G') 开始小于损耗模量 (G''),

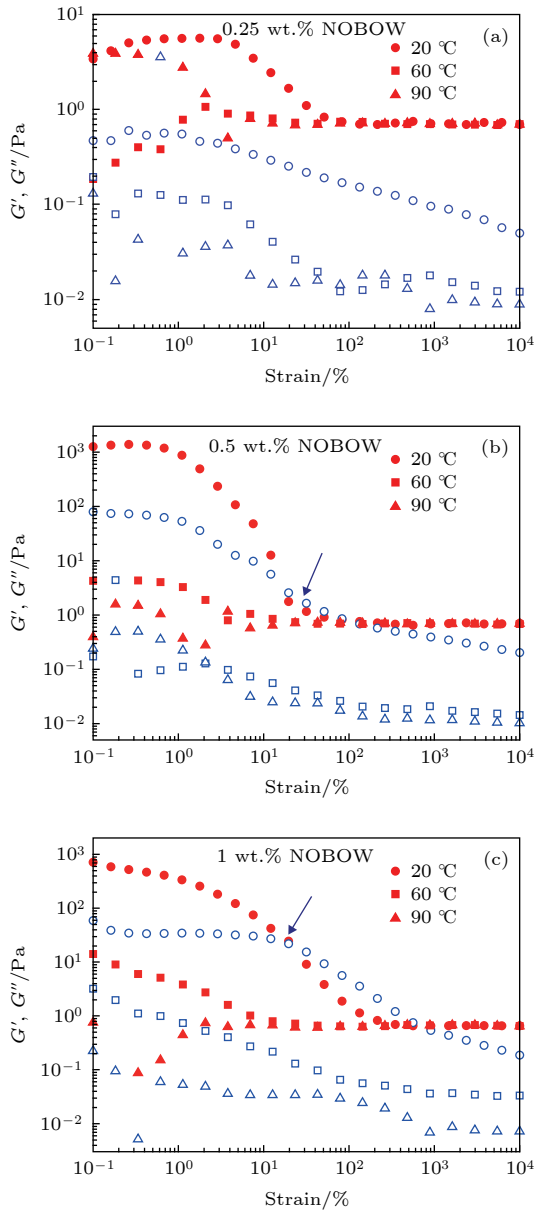


图 4 不同温度下, NOBOW/十六烷混合体系储能模量 (G') 和损耗模量 (G'') 随应变大小变化的规律 (a) 0.25 wt.%, (b) 0.5 wt.% 和 (c) 1 wt.% NOBOW/十六烷混合体系储能模量 (G') 和损耗模量 (G'') 分别在 20, 60 和 90 °C 随应变大小变化的关系. 实验中, 应变速率保持 1 Hz

Fig. 4. Strain sweeps of the elastic (G' , solid symbols) and viscous moduli (G'' , open symbols) of (a) 0.25 wt.%, (b) 0.5 wt.% and (c) 1 wt.% NOBOW/hexadecane mixtures measured at different temperatures of 20 °C, 60 °C and 90 °C. The measurements are taken at a fixed strain rate of 1 Hz.

表明有机凝胶基本被破坏, 体系开始表现出流体的基本特性. 同时由图 4 可见, 不同温度下, 在储能模量减小的时候, 没有出现损耗模量的峰值, 这和传统凝胶体系很不一样, 传统凝胶体系在储能模量减小的时候, 一般会出现损耗模量的峰值^[23], 这很可能是因为纳米螺旋丝偏刚性, 不能拉伸或弯折.

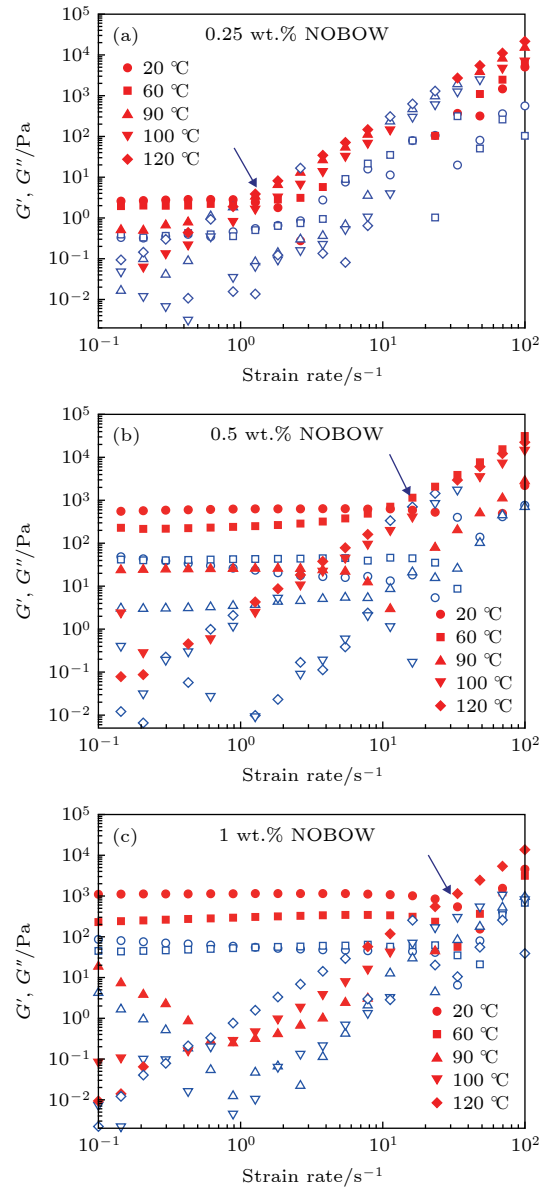


图 5 不同温度下, NOBOW/十六烷混合体系储能模量 (G') 和损耗模量 (G'') 随应变速率变化的规律 (a) 0.25 wt.%, (b) 0.5 wt.% 和 (c) 1 wt.% NOBOW/十六烷混合体系储能模量 (G') 和损耗模量 (G'') 分别在 20, 90, 100 和 120 °C 随应变速率变化的关系. 实验中, 应变大小保持 0.5%

Fig. 5. Strain rate sweeps of the elastic (G' , solid symbols) and viscous moduli (G'' , open symbols) of (a) 0.25 wt.%, (b) 0.5 wt.% and (c) 1 wt.% NOBOW/hexadecane mixtures measured at different temperatures of 20, 90, 100 and 120 °C. The measurements are taken at a fixed strain of 0.5%.

因为凝胶体系在 20 °C 时已经被破坏,所以在随后的 60 °C 和 90 °C 测试中,基本表现为流体的特性.

为了不破坏纳米螺旋丝的网络结构,测试有机凝胶流变随应变速率的变化关系,我们保持应变大小为 0.5%,然后逐渐增加应变速率,如图 5(a)—图 5(c) 所示.对于 $c=0.25$ wt.%, $c=0.5$ wt.% 和 $c=1$ wt.% 的 NOBOW/十六烷混合体系,当应变速率较小时,体系储能模量 (G') 和损耗模量 (G'') 不随应变速率增加而增加,而是保持不变,说明纳米螺旋丝具有一定刚性,能对应变速率范围内的变化做出快速响应.当应变速率足够大(如图 5 中箭头所示, $\dot{\gamma} \approx 1$ s⁻¹, $\dot{\gamma} \approx 10$ s⁻¹ 和 $\dot{\gamma} \approx 40$ s⁻¹ 分别对应 $c=0.25$ wt.%, $c=0.5$ wt.% 和 $c=1$ wt.% 的混合体系),体系储能模量 (G') 和损耗模量 (G'') 随着应变速率增加而逐渐增加,此时纳米螺旋丝已经不能响应应变的快速变化.

3 结 论

在 NOBOW/十六烷混合体系中,在较低 NOBOW 浓度 ($c=0.5$ wt.%) 下,纳米螺旋丝可以自组装生成三维网络,进而形成有机凝胶.该有机凝胶与传统凝胶不同,可以随温度发生凝胶-流体的可逆变化.当温度较高,体系处于流体状态时,体系储能模量和损耗模量均较小,与溶剂本身相近.当温度较低,香蕉形液晶分子自组装形成纳米螺旋丝时,体系储能模量随温度降低而逐渐增加,这主要是因为随着温度降低,液晶分子尤其是其柔性碳链的热运动逐渐减小,分子排列更有序,纳米螺旋丝刚度增大.因为纳米螺旋丝是由香蕉形液晶分子在层内不匹配驱动下自组装形成,具有内应力和一定刚性,能对低频应变做出快速响应,但容易被较大应力破坏.纳米螺旋丝力学性能的深入研究有助于我们更好地理解纳米螺旋丝三维网络有机凝胶的结构特性,拓展其在软物质领域的应用.

参考文献

[1] Qiao X X, Zhang X J, Tian Y, Meng Y G, Wen S Z 2013 *Acta Phys. Sin.* **62** 176101 (in Chinese) [乔小溪, 张向军, 田煜,

孟永钢, 温诗铸 2013 *物理学报* **62** 176101]

[2] Shen Z X, Tang M J, Chen P, Zhou S H, Ge S J, Duan W, Wei T, Liang X, Hu W, Lu Y Q 2020 *Adv. Opt. Mater.* **8** 1902124

[3] Qin J, Wang X Q, Yuan C, Zheng Z, Shen D 2019 *Liq. Cryst.* **47** 255

[4] Liu L, Wang M, Guo L X, Sun Y, Zhang X Q, Lin B P, Yang H 2018 *Macromolecules* **51** 4516

[5] Wang L, Chen D, Gutierrez Cuevas K G, Bisoyi H K, Fan J, Zola R S, Li G, Urbas A M, Bunning T J, Weitz D A 2017 *Mater. Horiz.* **4** 1190

[6] Ye F F, Mukhopadhyay R, Stenull O, Lubensky T C 2007 *Phys. Rev. Lett.* **98** 147801.1

[7] Ye F F, Lubensky T C 2009 *JPCB* **113** 3853

[8] Hough L E, Jung H T, Kruerke D, Heberling M S, Nakata M, Jones C D, Chen D, Link D R, Zasadzinski J, Heppke G 2009 *Science* **325** 456

[9] Dierking I 2010 *Angew. Chem. Int. Ed.* **49** 29

[10] Sekine T, Niori T, Sone M, Watanabe J, Takezoe H 1997 *Jpn. J. Appl. Phys.* **36** 6455

[11] Chen D, Madenian J E, Shao R, Dong K Y, Wang H, Korblova E, Walba D M, Glaser M A, Clark N A 2011 *J. Am. Chem. Soc.* **133** 12656

[12] Coleman D A, Fernsler J, Chattham N, Nakata M, Takanishi Y, Korblova E, Link D R, Shao R F, Jang W G, MacLennan J E 2003 *Science* **301** 1204

[13] Thisayukta J, Takezoe H, Watanabe J 2001 *Jpn. J. Appl. Phys.* **40** 3277

[14] Tschierske C, Dantlgraber G 2003 *Pramana* **61** 455

[15] Kondepudi D K, Kaufman R J, Singh N 1990 *Science* **250** 975

[16] Zep A, Salamonczyk M, Vaupotič N, Pocięcha D, Gorecka E 2013 *Chem. Commun.* **49** 3119

[17] Chen D, Zhu C, Wang H, MacLennan J E, Glaser M A, Korblova E, Walba D M, Rego J A, Soto-Bustamante E A, Clark N A 2013 *Soft Matter* **9** 462

[18] Gleeson H F, Liu H, Kaur S, Srigengan S, Görtz V, Mandle R, Lydon J E 2018 *Soft Matter* **14** 9159

[19] Matraszek J, Topnani N, Vaupotic N, Takezoe H, Mieczkowski J, Pocięcha D, Gorecka E 2015 *Angew. Chem. Int. Ed. Engl.* **128** 3529

[20] Shen Z, Jiang Y, Wang T, Liu M 2015 *J. Am. Chem. Soc.* **137** 16109

[21] Shen Z, Wang T, Shi L, Tang Z, Liu M 2015 *Chem. Sci.* **6** 4267

[22] Zheludev N I 2010 *Science* **328** 582

[23] Chen D, Tuchband M R, Horanyi B, Korblova E, Walba D M, Glaser M A, MacLennan J E, Clark N A 2015 *Nat. Commun.* **6** 1

[24] Liu L W, Wang Z W, Zhou L W, Wang Z J, Gao G J, Liu X J 2000 *Acta Phys. Sin.* **49** 1887 (in Chinese) [刘立伟, 王作维, 周鲁卫, 王冶金, 高广君, 刘晓君 2000 *物理学报* **49** 1887]

[25] Macosko C W 1994 *RHEOLOGY Principles, Measurements and Applications* (Canada: Wiley-VCH) pp65–106

[26] Mason T G, Weitz D A 1995 *Phys. Rev. Lett.* **74** 1250

[27] Chen D, Zhu C, Shoemaker R K, Korblova E, Walba D M, Glaser M A, MacLennan J E, Clark N A 2010 *Langmuir* **26** 15541

SPECIAL TOPIC—Statistical physics and complex systems

The rheology property of organogels based on 3D helical nanofilament bnetworks self-assembled by bent-core liquid crystals*

Wang Xing-Zheng^{1) #} Yang Chen-Jing^{1) #} Cai Li-Heng^{2) †} Chen Dong^{1) ‡}1) (*Institute of Process Equipment, College of Energy Engineering, Zhejiang University, Hangzhou 310027, China*)2) (*College of Materials Science and Engineering, University of Virginia, Charlottesville 22904, United States of America*)

(Received 4 March 2020; revised manuscript received 22 March 2020)

Abstract

In the B4 phase of bent-core liquid crystals, smectic layers of tilted achiral bent-core molecules are chiral and polar, which, driven by intra-layer structural mismatch, eventually twist into helical nanofilaments. We design a NOBOW/hexadecane organogel system, which is different from traditional organogel system, and the studied organogels show reversible gel-liquid transitions under temperature cycles. At high temperature, the NOBOW molecules dissolve in hexadecane and the storage modulus and viscous modulus show typical liquid characteristics. At low temperature, the mobility of NOBOW molecules decreases and the storage modulus of the organogels increases as the temperature decreases. We conduct a rheology experiment to systematically investigate the viscoelasticity of the organogel to understand the property of the organogel and develop the application in soft matter. The viscoelastic studies of the organogels reveal that the helical nanofilaments are internally strained and their 3D networks are relatively stiff, which provides an in-depth insight into the properties of the organogels and paves the way for their applications in soft matter.

Keywords: bent-core liquid crystal, helical nanofilament, organogel, viscoelastic property

PACS: 61.30.-v, 77.84.Nh

DOI: 10.7498/aps.69.20200332

* Project supported by the National Natural Science Foundation of China (Grant Nos. 11704331, 21878258) and the Open Funds of the State Key Laboratory of Rare Earth Resource Utilization, China (Grant No. RERU2019008).

These authors contributed equally.

† Corresponding author. E-mail: liheng.cai@virginia.edu

‡ Corresponding author. E-mail: chen_dong@zju.edu.cn

复杂网络链路可预测性：基于特征谱视角

谭索怡 祁明泽 吴俊 吕欣

Link predictability of complex network from spectrum perspective

Tan Suo-Yi Qi Ming-Ze Wu Jun Lu Xin

引用信息 Citation: *Acta Physica Sinica*, 69, 088901 (2020) DOI: 10.7498/aps.69.20191817

在线阅读 View online: <https://doi.org/10.7498/aps.69.20191817>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

基于复杂网络动力学模型的无向加权网络节点重要性评估

Evaluation methods of node importance in undirected weighted networks based on complex network dynamics models

物理学报. 2018, 67(9): 098901 <https://doi.org/10.7498/aps.67.20172295>

复杂网络谱粗粒化方法的改进算法

Improved algorithm of spectral coarse graining method of complex network

物理学报. 2017, 66(6): 060502 <https://doi.org/10.7498/aps.66.060502>

知识图谱复杂网络特性的实证研究与分析

Empirical study of knowledge network based on complex network theory

物理学报. 2019, 68(12): 128902 <https://doi.org/10.7498/aps.68.20190106>

基于相对距离的复杂网络谱粗粒化方法

A spectral coarse graining algorithm based on relative distance

物理学报. 2019, 68(10): 100501 <https://doi.org/10.7498/aps.68.20181848>

基于领域相似度的复杂网络节点重要度评估算法

Node importance measurement based on neighborhood similarity in complex network

物理学报. 2017, 66(3): 038902 <https://doi.org/10.7498/aps.66.038902>

专题：统计物理和复杂系统

复杂网络链路可预测性：基于特征谱视角*

谭索怡¹⁾ 祁明泽²⁾ 吴俊^{3)†} 吕欣^{1)‡}

1) (国防科技大学系统工程学院, 长沙 410073)

2) (国防科技大学文理学院, 长沙 410073)

3) (北京师范大学复杂系统国际科学中心, 珠海 519087)

(2019年11月3日收到; 2020年2月27日收到修改稿)

近年来链路预测的理论和实证研究发展迅速, 大部分工作关注于提出更精确的预测算法. 事实上, 链路预测的前提是网络的结构本身能够被预测, 这种“可被预测的程度”可以看作是网络自身的基本属性. 本文拟从特征谱的视角去解释网络的链路可预测性, 并刻画网络的拓扑结构信息, 通过对网络特征谱进行分析, 构造了复杂网络链路可预测性评价指标. 通过该指标计算和分析不同网络的链路可预测性, 能够在选择算法前获取目标网络能够被预测的难易程度, 解决到底是网络本身难以预测还是预测算法不合适的问题, 为复杂网络与链路预测算法的选择和匹配问题提供帮助.

关键词：可预测性, 链路预测, 特征谱, 复杂网络

PACS: 89.75.Hc, 89.75.Fb, 64.60.aq

DOI: 10.7498/aps.69.20191817

1 引言

近年来, 复杂网络研究迅速发展, 其学科分支在包括数学、统计物理、生物医学、化学、计算机等领域掀起研究热潮^[1-5]. 在现代信息科学领域中, 链路预测作为将复杂网络与信息科学连接起来的重要桥梁, 关心的是信息科学中最基本的问题——缺失信息的预测和还原问题^[6,7]. 即在一个网络中, 如何基于已知连边信息, 刻画网络的相似性, 进而重现因为数据缺失尚未观察到的连边, 或者预测未来网络演化过程中将要出现的连边.

目前链路预测相关理论方法研究主要围绕基于马尔科夫链、最大似然估计、概率模型、网络结构相似性等数学领域和统计物理的观点和方法展开. 早期的链路预测领域普遍关注的是马尔科夫链

和机器学习, 主要存在着计算复杂度较高, 参数设置不具有普适性等问题^[8]. 也有学者提出从似然分析的角度构建链路预测框架, 比较经典的有层次结构模型^[9]和随机分块模型^[10]. Pan等^[11]提出的闭路模型, 拥有比前两者更好的预测精度. 似然分析的优点在于能够从理论上帮助我们理解网络结构特征, 然而受限其自身理论的复杂性, 这类方法不是应用性很强的方法, 即使构思巧妙, 在处理大规模网络时也会显得吃力. 最早由 Taskar等^[12]提出的概率模型是数据挖掘领域的传统模型, 该模型在预测时同时运用了网络的结构信息和节点的属性信息, 概率模型拥有较高的预测精确度, 但是同时产生的高计算复杂度以及其参数设置存在非普适性, 都限制了该类方法的应用范围. 得益于 David 和 Kleinberg^[13]在 2007 年有关链路预测结构相似性的论文, 基于网络结构相似性的链路预测问题在

* 国家自然科学基金 (批准号: 82041020, 71771213, 71901067, 71871217) 和湖南省科技计划项目 (批准号: 2017RS3040, 2018JJ1034, 2019JJ20019) 资助的课题.

† 通信作者. E-mail: wujunpla@hotmail.com

‡ 通信作者. E-mail: xin.lu@flowminder.org

近年受到越来越多的关注. Zhou 等^[14]把链路预测问题和评价指标都进行了简化,很多研究人员开始利用同样的数据和指标分析链路预测问题.基于相似性的算法,作为最简单的链路预测算法框架,其中一系列算法复杂性低但预测精度不错的局部相似性指标的提出,大幅度增加了链路预测在超大规模网络中的可应用性.

利用链路预测算法精准地预测网络的未知结构有着广泛的应用前景.例如,在军事对抗中,通常只能侦测到敌方作战网络的部分结构信息,如果我们能够获得更多更准确的信息,就可以制定一定的优先级规则或重要性标准来选择性地攻击网络中的关键节点或连边^[15,16];在生物实验中,研究人员需要通过大量的实验研究去推断探索细胞组分内部的交互作用,一个具有指导作用的预测结果能有效降低实验成本并帮助人类理解生物网络连边演化机制的规律^[17,18];社交网络中,读懂用户的兴趣偏好和喜怒哀乐,对企业的发展事至关重要,一个好的“猜你要关注”推荐能够牢牢地黏住老用户、吸引新用户^[19,20].此外,一个优秀的链路预测算法往往蕴含着一种可能的网络演化机制^[21–24].遗憾的是,除非站在上帝视角,否则没有人能判断一个链路预测算法是否足够精准.如果网络的节点对之间随机连接,任何算法可能都会无功而返,难以做出有效预测;相反,面对一个有特定的连边演化机制,非常规则的网络,一个足够优秀的方法能够实现精度很高的预测.此外,即使是同一个网络,不同链路预测算法的准确性也不尽相同,这种精度值只能相对地反映出网络对于某种特定预测算法的预测精度,算法不同,精度也随之改变,并不能刻画网络自身的固有的链路可预测性,很多时候,我们都面临着是预测算法不合适还是网络本身就难以预测这样一个网络与算法的选择和匹配问题.

显然,网络中待预测的连边集合与网络中不存在的连边集合交集为空集,无论预测的准确性和效率如何,理论上我们总可以通过无限加边命中所有待预测的连接.然而这种上界是没有价值的,不考虑成本的加边会带来巨大的成本消耗和结构噪音,这样的情况显然偏离了链路预测的初衷.如果能够获悉一个网络的链路信息能够多大程度被预测出来,就能够提供一个导向,确定当前算法是否接近或者已经达到目标网络的可预测上限.因此,刻画网络多大程度上能够被预测是链路预测中首先需

要解决的问题,这个问题在相关文献中被称为复杂网络的链路可预测性问题.

近年来链路预测的理论和实证研究发展迅速,但绝大部分研究的目的都是希望提出更准确的预测算法^[25,26],关于复杂网络链路可预测性的研究起步较晚,相关成果少见报道.许小可等^[27]最早从理论上比较了各种算法的优劣,分析多个网络演化过程中形成链接的两个节点之间的拓扑距离分布,阐明了传统基于共同邻居相似性指标可有效进行链路预测的机理,从理论上分析了9种基于共同邻居相似性算法的预测上限. Lü等^[28]提出结构一致性的概念,认为网络“可被预测的程度”,是网络的一种重要固有属性.通过对已知网络进行扰动,刻画重构的邻接矩阵和真实邻接矩阵的差异.如果丢失的连边没有显著改变网络的结构,那么这个网络是可预测的,即网络的结构一致性越强,网络可预测性越好.熵被广泛用来测量物理系统中的无序度, Yin等^[29]设计了基于证据推理(Dempster-Shafer theory)的链路预测算法,从香农信息熵的视角出发,分析了网络链路信息的可预测性.

本文拟从特征谱的视角去理解网络拓扑信息,并刻画网络的链路可预测性.首先基于特征谱理论给出复杂网络链路可预测性的数学描述,提出可预测性指标.在此基础上,通过计算和分析不同实证网络的链路可预测性,验证该指标的有效性.

2 链路预测问题描述

在本文的研究中,主要讨论无向无权网络.令 $G(V, E)$ 表示无向无权网络, V 表示节点, E 表示连边.令 $U = N(N-1)/2$ 表示连边的全集.对于网络中未连边的节点对 (v_i, v_j) , 可以通过某种预测算法得到其得分矩阵,将所有未连边列表中节点对的得分降序排列,排在前面的节点对之间产生链接的可能性大.

在网络进行链路预测之前,我们并不知道网络缺失的部分和未来演化中可能出现的连边连接情况,因此,在实验中,将网络中已有的连边集合 E 划分为训练集 E^T 和测试集 E^P .显然, $E = E^T \cup E^P$, $E^T \cap E^P = \emptyset$.链路预测算法通过学习训练集 E^T 中的相似性进行预测,并通过测试集 E^P 检测算法预测效果,测试集中存在的预测连边越多,算法的准确性越高.其中,数据集的划分存在多种方式,

为排除其干扰, 本文所有实验中均采用随机抽样法. 常见的算法评价指标有 AUC (area under the receiver operating characteristic curve)^[30], 精确度 (precision)^[31] 和排序分 (ranking score)^[32], 本文选用 precision 对预测结果进行评价. precision 定义为在前 L 条边的预测中, 正确预测连边的比例. 如果有 m 条边被正确预测, 则 precision 的定义为

$$\text{precision} = m/L. \quad (1)$$

为了更好地解释链路预测问题, 图 1 给出了一个简

单的例子. 图 1(a) 为 8 个节点和 13 条连边的完全信息网络. 我们采取随机抽样的方法选择 3 条连边作为测试集, 如图 1(b) 中黄色连边所示, 显然, 训练集包含 10 条连边. 由于 8 个节点的全连通网络共有 28 条连边, 则未连边的数目为 $28 - 10 = 18$. 选择一种链路预测算法, 对 18 条未知连边进行打分, 并将得分按从大到小排序, 精确度越高的算法能更多地将测试集中的 3 条连边 $\{e_{15}, e_{17}, e_{34}\}$ 排在其余 15 条不存在边的前面.

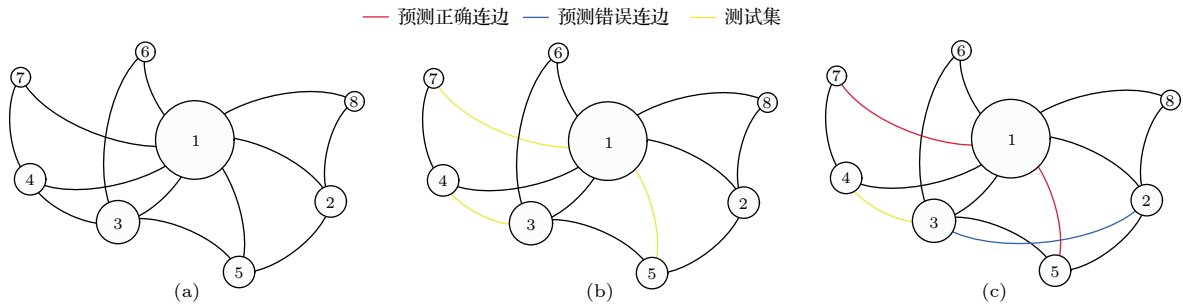


图 1 链路预测问题示意图

Fig. 1. Illustration of link prediction problem.

在这个例子中我们选择资源分配算法^[14]进行链路预测, 选取该算法认为存在可能性最高的 3 条连边添加到网络中, 如图 1(c) 所示, 红色连边表示正确预测, 蓝色连边表示错误预测, 可以看到, 算法正确预测了连边 e_{15} 和 e_{17} , 未能正确预测出节点 3 和节点 4 之间的连边 e_{34} 而是错误的认为连边 e_{23} 存在的可能性更高, 易计算得到, 此次预测精度 (precision) 为 $2/3$.

3 基于特征谱的复杂网络链路可预测性

3.1 复杂网络的特征谱

复杂网络的特征谱是代数图论的基本研究课题, 经过多年的研究, 如文献^[33]所述, 已有成熟的理论体系和丰富的研究成果. 网络的特征谱提供了包含网络功能和动力学行为在内的大量信息, 可以被形容为网络的“指纹”, 即网络与其特征谱是一一对应的, 不同类别的网络有着完全不同的特征谱. 因此, 通过分析和识别特征谱, 我们就能够锁定目标网络. 进一步, 特征谱不仅是网络的“指纹”, 还是网络的“脉象”. 通过分析特征谱这一网络“脉象”, 可以得到大量的网络结构信息. 例如, 通过拉

普拉斯矩阵 (Laplace matrix) 的最大特征根我们可以估计网络的度序列; 分析特征谱还可以挖掘网络社区结构; 网络的中心性和二部分性也可从特征谱得出^[34–37]. 最近有研究表明, 网络的特征值谱还可以表现网络结构和动力学 (例如神经与激发序列) 的层次性^[38].

令 G 表示无向无权图, $\mathbf{A}(G) = (a_{ij})_{N \times N}$ 表示 G 的邻接矩阵, 其中若节点对 v_i 与 v_j 之间有连边, 则 $a_{ij} = 1$, 否则 $a_{ij} = 0$. 令 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$ 为 $\mathbf{A}(G)$ 的特征根, 则称集合 $\{\lambda_i\}$ 为 G 的特征谱 (spectrum). 定义 d_i 为节点 v_i 的度. G 的拉普拉斯矩阵 (Laplace matrix) 可用数学公式表示为 $\mathbf{L}(G) = \mathbf{D}(G) - \mathbf{A}(G)$, 式中, $\mathbf{D}(G) = \text{diag}\{d_i\}$ 表示节点度的对角矩阵, 显然, $\mathbf{L}(G)$ 是对称半正定矩阵. 令 $\mu_1 \geq \mu_2 \geq \dots \geq \mu_N$ 表示 $\mathbf{L}(G)$ 的特征根, 则称集合 $\{\mu_i\}$ 为图 G 的拉普拉斯特征谱.

3.2 基于特征谱视角的网络链路可预测性

近年来, 很多统计物理领域的学者基于特征谱研究了图的沟通性 (communicability)^[34] 和可扩展性 (good expansion, GE)^[39]. 图的沟通性指网络中不同节点之间进行交流或传递信息的能力, 而可扩展性指那些既稀疏同时又高度连通的节点间的沟

通能力. 实际上, 统计物理角度的沟通和扩张, 在网络信息的视角中, 可以理解为网络结构某种程度上的演化和发展. 链路预测, 作为网络信息挖掘的技术手段, 一个很重要的功能便是预测缺失连边和未来可能存在的连边. 可以说, 链路预测算法与网络连边形成机制相辅相成, 好的链路预测算法本身就给出了很多网络演化可能机制的暗示; 反之网络的链路可预测性也可以理解为网络连边演化机制的另一种表现形式. 因此, 我们可以认为, 沟通性和可扩张性这两个指标所刻画的拓扑信息从某种

程度上来说和网络的链路可预测性是相似的, 即具备较好的链路可预测性的网络, 一般也具有较好的沟通性和可扩张性.

已有研究表明, 可扩张性好的网络同时也表现出良好的沟通性, 且这些网络特征谱的最大特征根 λ_1 远大于次大特征根 λ_2 , 即 $\lambda_1 \gg \lambda_2$. 我们在之前的工作中^[40]研究了无标度网络特征谱, 同样发现不同参数的无标度网络中存在着不同程度的 $\lambda_1 \gg \lambda_2$, 即存在谱隙 (spectrum gap) 现象, 如图 2 所示. 因此, 如果能够定量地刻画特征谱中 λ_1 和其他特征根之间的差距, 就能够像中医把脉一样, 定量刻画网络的链路可预测性.

3.3 可预测性的数学表达式

在各种各样衡量网络结构属性的指标中, 文献^[41]提出的子图中心性是基于网络特征谱的指标. 其认为闭环回路的路径长度越小, 回路信息交流越便利, 节点之间的联系越紧密, 对节点的中心性贡献越大. 节点 i 的子图中心性可以定义为

$$SC(i) = \sum_{j=1}^N (\xi_j^i)^2 e^{\lambda_j}, \quad (2)$$

其中 λ_j ($j = 1, 2, \dots, n$) 是邻接矩阵 \mathbf{A} 第 j 个特征向量的特征值, ξ_j 是 λ_j 所对应的特征向量, ξ_j^i 是邻接矩阵第 j 个特征向量的第 i 个组分, 例如, λ_1 和 ξ_1 分别是邻接矩阵 \mathbf{A} 的最大特征值及其对应的特征向量. 对于 (2) 式而言, 显然, $SC(i)$ 包含了从节点 i 出发, 偶数长度和奇数长度的所有的闭途径. 因此, (2) 式也可以表示为

$$SC(i) = \sum_{j=1}^N (\xi_j^i)^2 \cosh(\lambda_j) + \sum_{j=1}^N (\xi_j^i)^2 \sinh(\lambda_j) = SC_{\text{even}}(i) + SC_{\text{odd}}(i). \quad (3)$$

显然, 偶数长度的闭途径更多的是一些无环的轨迹, 更多地出现在二部分图中; 而奇数长度的闭途径则不包含这部分无效的路径. 本文的研究对象是简单无权图, 因此, 奇数长度的闭途径更适合于用来描述网络中节点与其邻居间的拓扑结构关系. 我们可以将 $SC_{\text{odd}}(i)$ 写成如下形式^[39]:

$$SC_{\text{odd}}(i) = (\xi_1^i)^2 \sinh(\lambda_1) + \sum_{j=2}^N (\xi_j^i)^2 \sinh(\lambda_j), \quad (4)$$

其中 λ_1 是网络的主特征值, ξ_1^i 是主特征向量的第 i 个组分. 当网络存在一个巨大的谱隙 (spectral

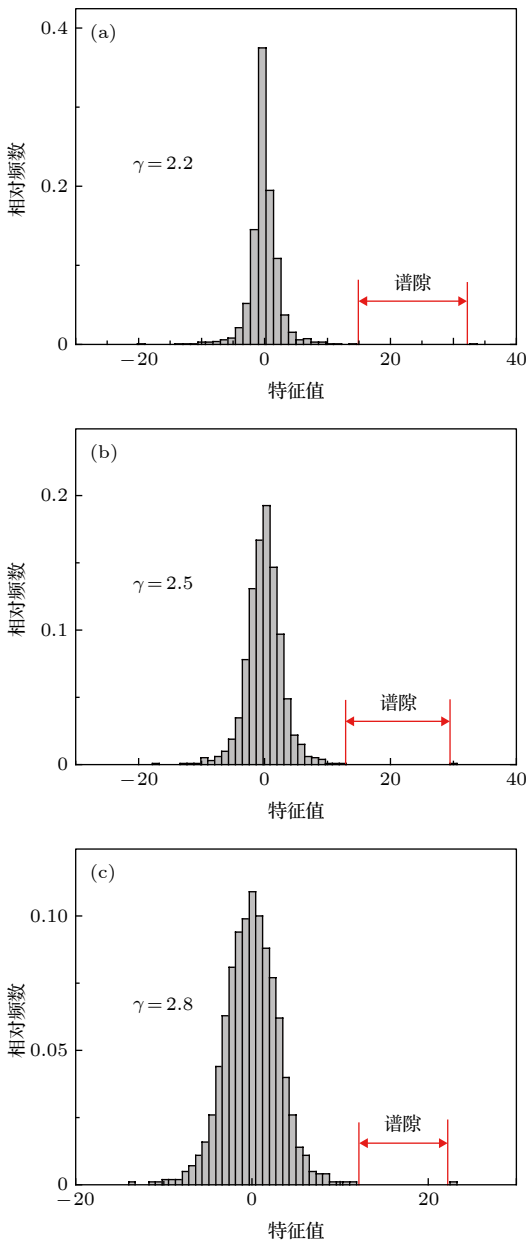


图 2 无标度网络特征谱直方图

Fig. 2. The histograms of eigenvalues of random scale-free networks.

gap) 时, 有 $\lambda_1 \gg \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_N$. 因此, 在这种情况下,

$$(\xi_1^i)^2 \sinh(\lambda_1) \gg \sum_{j=2}^N (\xi_j^i)^2 \sinh(\lambda_j), \quad (5)$$

则

$$\text{SC}_{\text{odd}}(i) \approx (\xi_1^i)^2 \sinh(\lambda_1). \quad (6)$$

也就是说, 要判断网络特征谱中 λ_1 和 λ_2 之间是否有足够大的谱隙. 需要检测

$$\xi_1^i \propto \sqrt{\text{SC}_{\text{odd}}(i)} - \sqrt{\sinh(\lambda_1)}. \quad (7)$$

令 $A = [\sinh(\lambda_1)]^{-0.5}$, $\eta = 0.5$, (7) 式可以表示为 $\xi_1^i \propto A [\sqrt{\text{SC}_{\text{odd}}(i)}]^\eta$. 显然, ξ_1^i 和 $\text{SC}_{\text{odd}}(i)$ 之间存在着线性关系. 因此, 我们可以在双对数形式下将 (7) 式改写成:

$$\log \xi_1^i = 0.5 \log \text{SC}_{\text{odd}}(i) - 0.5 \log \sinh(\lambda_1). \quad (8)$$

通过测量不同情况与理想情况的偏差 $\Delta \log \xi_1^i$, 我们可以判断网络是否具有良好的可预测性.

$$\Delta \log \xi_1^i = \log \frac{\xi_1^i}{\xi_1^{\text{ideal}}(i)} = \log \left\{ \frac{(\xi_1^i)^2 \sinh(\lambda_1)}{\text{SC}_{\text{odd}}(i)} \right\}^{0.5}, \quad (9)$$

当 $\Delta \log \xi_1^i \approx 0$ 时, 网络具有良好的可预测性. 由于 $\Delta \log \xi_1^i$ 是一个一维数组, 直接比较 $\Delta \log \xi_1^i$ 与 0 的关系并不容易, 因此, 我们构建一个可预测性的数学表达式 p 去测度 $\Delta \log \xi_1^i$ 多大程度接近于 0. 其数学表达式如下:

$$p = \exp \left[- \sum \sqrt{\left(\frac{\text{SC}_{\text{odd}}(i) - (\xi_1^i)^2 \sinh(\lambda_1)}{\text{NSC}_{\text{odd}}(i)} \right)^2} \right], \quad (10)$$

易知, 可预测性 p 的值域为 $[0, 1]$, 如果偏差趋近于 0, 那么 p 趋近于 1, 表示网络的链路可预测性很好; 反之, 若网络 p 值较小, 表示网络的可预测性差.

4 实验结果分析

4.1 模型网络的可预测性分析

相比于随机网络, BA 无标度网络具有节点生长和边的偏好链接 (preferential attachment) 两种明确的生成机制, 即新加入的节点更倾向于与那些具有较大连接度的节点相连. 一个新节点与一个已经存在的节点 v_i 相连接的概率 Π_i 与节点的度 d_i 成正比:

$$\Pi_i = \frac{d_i}{\sum_j d_j}, \quad (11)$$

这意味着, 如果我们的指标能够有效刻画网络的可预测性, 则 p 值会随着网络演化机制的变化而改变. 为全面比较网络演化机制对 p 刻画可预测性能力的影响, 我们基于 (11) 式, 加入参数 α , 调控 BA 模型中偏好链接机制的强度. 构造连接概率 Π_i' :

$$\Pi_i' = \frac{d_i^\alpha}{\sum_j d_j^\alpha}. \quad (12)$$

当 $\alpha = 0$ 时, 网络生成仅由生长机制决定 (在此情况下老的节点仍有更高的概率获得更多连接); 当 $\alpha = 1$ 时, 网络生成过程具有显著的偏好链接特征. 图 3 展示了当网络平均度为 4 时, 不同节点规模下可预测性 p 值随参数 α 的变化, 结果表明, 随着 α 的增加, 优先链接特性逐渐增强, 我们的指标能够捕捉到网络逐渐明显的优先链接特性, 网络的可预测性越来越好. 并且在连边机制固定的情况下, 可预测性随网络规模的变化不大, 证明 p 指标具有良好的稳定性.

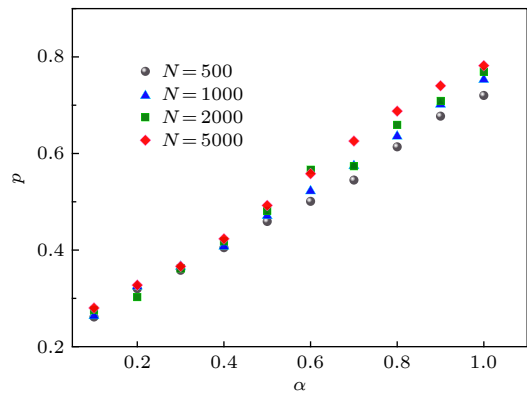


图 3 不同节点规模下, 模型网络可预测性 p 随 α 的变化
Fig. 3. The link predictability of model network versus α with various N .

同时, 为了清晰地对比各种链路预测算法在这两类模型网络中的表现, 我们生成一个节点数为 1000, 平均度为 6 的 BA 无标度网络和与之同样规模的随机网络进行对比实验, 实验结果如图 4 和表 1 所示. 表 1 给出了 12 种基于相似性的链路预测算法在这两个模型网络中的精确度 (precision). 包括 6 种基于节点局部信息的相似性指标: 共同邻居 (CN), Adamic-Adar(AA), 资源分配 (RA), 偏

好连接 (PA), Individual Attraction(IA) 和 CAR 指标; 3 种基于路径的相似性指标: 局部路径 (LP), Katz 和 LHN-II 指标; 3 种基于随机游走的相似性指标: 平均通勤时间 (ACT), 重启的随机游走 (RWR), 局部随机游走指标 (LRW), 具体的算法原理参见文献 [25]. 实验通过随机抽样的方式, 将训练集和测试集按照 9:1 的比例进行划分, 我们固定预测连边的比例, 令 L 等于测试集中的连边数量.

可以看到, 由于 BA 无标度网络中, 新的节点进入网络后会选择网络中已存在的大度节点产生链接. 网络具有固定的网络连边演化机制, 连边都是按照优先链接产生, 因此, 网络具有很好的可预测性. 在图 4(c) 中表现为 ξ_1^i 和 $SC_{\text{odd}}(i)$ 在一条直线上, 体现出 (7) 式表示的线性关系; 反观同样规模的随机网络, 网络中的连边以固定概率随机产生, 不根据任何演化机制和节点属性, 很难基于某一演化机制去预测连边是否存在, 网络可预测性较差. 在图 4(a) 中表现为 ξ_1^i 和 $SC_{\text{odd}}(i)$ 并不具有很强的线性关系. 图 4(b) 的雷达图直观地展示了链路预测算法在两类模型网络中的表现, 结果表明, 各类算法在 BA 无标度网络中的精确度显著优于随机网络, 这与我们对这两类网络可预测性的判断是一致的. 观察各个算法的表现, 在 BA 无标度网络中, PA 指标表现最为出色, 这是因为 PA 算法的思想来源于优先链接的方法, 即连边存在的可能性大小正比于两端度值的积. 因此, PA 算法对于相似性的定义更贴近于 BA 无标度网络的连边演化机制, 故在这类网络中有着优异的表现. 纵使 BA 无标度网络具有优秀的可预测性, LHN-II 算法在网络中的表现却很差. 这是因为 LHN-II 算法是基于一般等价 (regular equivalence) 的思想, 其相似性的定义更多地取决于节点连接的节点之间的相似性, 即使节点对之间不存在共同邻居. 然而, 节点的属性如果不是特殊背景的网络或者有特定的标准往往是很难去量化的, 因此, 虽然无标度网络有着高的可预测性, 但 LHN-II 算法却不是适用于该网络的合适的链路预测算法. 上述结果初步表明, 通过计算链路可预测性 p 的值, 能够回答到底是不可预测的网络还是不合适的算法这个问题, 从而为决策者筛选算法提供指导意见. 上述模型网络只是真实网络一种演化机制的抽象, 真实网络在生长演化过程

中往往表现出如集聚性、社团性、无标度性等多种复杂的结构特征, 为进一步证明指标的有效性, 我们在更多真实网络中进行了实验.

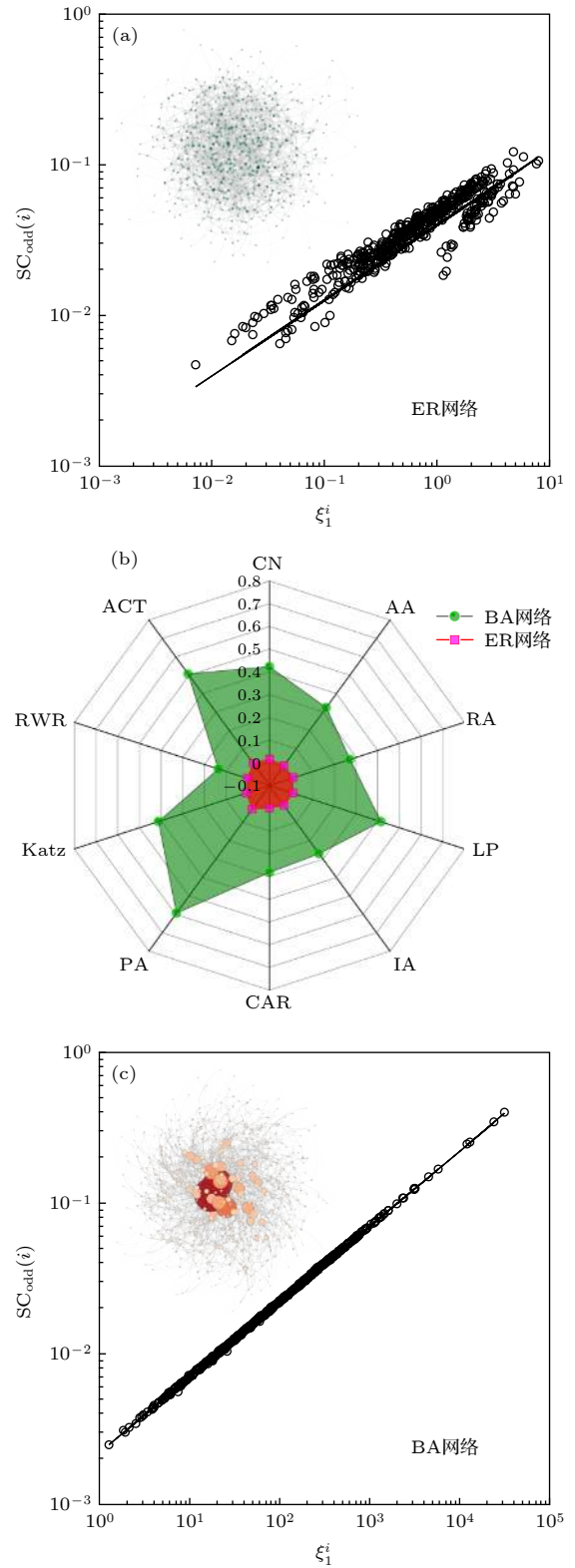


图 4 无标度网络和随机网络可预测性示意图

Fig. 4. The link predictability of BA scale-free network and random graph.

表 1 链路预测算法在模型网络中的表现
Table 1. Performance of link prediction algorithms in model networks.

网络	p	CN	AA	RA	LP	IA	CAR	PA	Katz	RWR	ACT	LRW	LHN-II
BA网络	0.975	0.423	0.324	0.272	0.415	0.271	0.283	0.594	0.412	0.136	0.507	0.085	0.003
随机网络	0.543	0.015	0.008	0.009	0.008	0.009	0	0.030	0.008	0.002	0.020	0	0.001

4.2 真实网络的链路可预测性分析

本节进一步考察可预测性指标在真实世界网络中的表现. 我们选取了各个不同领域的 15 个真实网络作为实验网络. 网络拓扑属性如表 2 所列, V 和 E 表示网络中的节点和边, $\langle k \rangle$ 表示平均度, C 表示集聚系数, r 表示同配系数, $\langle l \rangle$ 表示平均最短路径长度.

在实验中, 我们采用随机抽样的方法, 将训练集和测试集按照 9:1 的比例进行划分, 即测试集包含 10% 的真实连边. 针对预测结果采取 precision 衡量算法的表现, 由于真实网络规模不同, 在实验时固定预测连边的比例, 令 L 等于测试集中的连边数量.

表 3 为链路预测算法在真实网络中的最大精确度 (precision) 测试结果, 每个链路预测算法在每个网络运行 100 次取平均. 从网络间纵向比较来看, 算法在可预测性高的网络上的预测精度要明显高于可预测低的网络, 如图 5 大图所示, 不同颜色的圆代表不同的网络, 圆的大小与网络可预测性 p 成正比. 横坐标表示网络可预测性的值, 纵坐标表

示链路预测算法的最大 precision 值, 纵坐标越大, 算法的最大精度越高. 可以看到, 那些可预测性好的网络对应的 precision 值也相对较高, 如图中右

表 2 不同领域真实网络拓扑属性
Table 2. Basic statistics of real networks.

网络	V	E	r	$\langle k \rangle$	$\langle l \rangle$	C
C_elegans	297	2148	-0.163	14.47	2.46	0.308
Windsurfers	43	336	-0.147	15.63	1.70	0.564
Adolescent health	2539	12969	0.251	10.22	4.52	0.142
Jazz	198	2742	0.020	27.69	2.21	0.520
USAirport	1574	28236	-0.113	35.87	3.14	0.384
Metabolic	453	4596	-0.226	20.29	2.64	0.124
Yeast	2375	11693	0.454	9.85	5.10	0.388
US powergrid	4941	6594	0.003	2.67	20.09	0.103
Physicians	241	1098	-0.056	9.11	3.02	0.552
Air Traffic Control	1226	2615	-0.015	4.27	6.10	0.063
Contiguous USA	49	107	0.233	4.37	4.26	0.406
Email	1133	5451	0.078	9.62	3.65	0.166
King James Bible	1773	9131	-0.048	10.30	3.38	0.163
Protein Stelzl	1706	6207	-0.191	7.28	5.09	0.006
Router	5022	6258	-0.138	2.49	6.45	0.033

表 3 链路预测算法在真实网络中的 precision 值
Table 3. The precision of link prediction algorithms in real networks.

网络	p	CN	AA	RA	LP	IA	CAR	PA	Katz	RWR	ACT	LRW	SRW
C_elegans	0.999	0.100	0.107	0.105	0.101	0.108	0.094	0.058	0.101	0.105	0.055	0.110	0.108
Windsurfers	0.999	0.379	0.396	0.413	0.370	0.393	0.381	0.214	0.369	0.360	0.247	0.402	0.426
Adolescent health	0.422	0.103	0.103	0.088	0.089	0.101	0.094	0.003	0.088	0.053	0.008	0.042	0.047
Jazz	1.000	0.502	0.523	0.542	0.489	0.535	0.517	0.133	0.489	0.352	0.168	0.342	0.393
USAirport	0.998	0.333	0.336	0.364	0.332	0.332	0.330	0.280	0.332	0.087	0.294	0.076	0.080
Metabolic	0.999	0.137	0.195	0.269	0.141	0.168	0.132	0.104	0.141	0.196	0.092	0.214	0.215
Yeast	0.998	0.154	0.177	0.267	0.158	0.161	0.148	0.094	0.174	0.073	0.211	0.045	0.059
US powergrid	0.362	0.054	0.032	0.028	0.058	0.047	0.037	0.000	0.057	0.016	0.034	0.015	0.018
Physicians	0.368	0.119	0.126	0.121	0.117	0.122	0.106	0.014	0.117	0.119	0.015	0.132	0.127
Air Traffic Control	0.480	0.036	0.024	0.018	0.037	0.021	0.025	0.007	0.037	0.002	0.015	0.002	0.002
Contiguous USA	0.540	0.096	0.130	0.132	0.005	0.000	0.000	0.012	0.004	0.067	0.053	0.133	0.121
Email	0.950	0.144	0.158	0.143	0.142	0.159	0.145	0.018	0.141	0.065	0.024	0.052	0.051
King James Bible	0.960	0.167	0.270	0.446	0.163	0.256	0.176	0.078	0.163	0.186	0.069	0.197	0.224
Protein Stelzl	0.441	0.001	0.002	0.001	0.001	0.002	0.006	0.014	0.001	0.006	0.013	0.006	0.006
Router	0.511	0.051	0.029	0.020	0.056	0.031	0.055	0.022	0.055	0.006	0.164	0.005	0.005

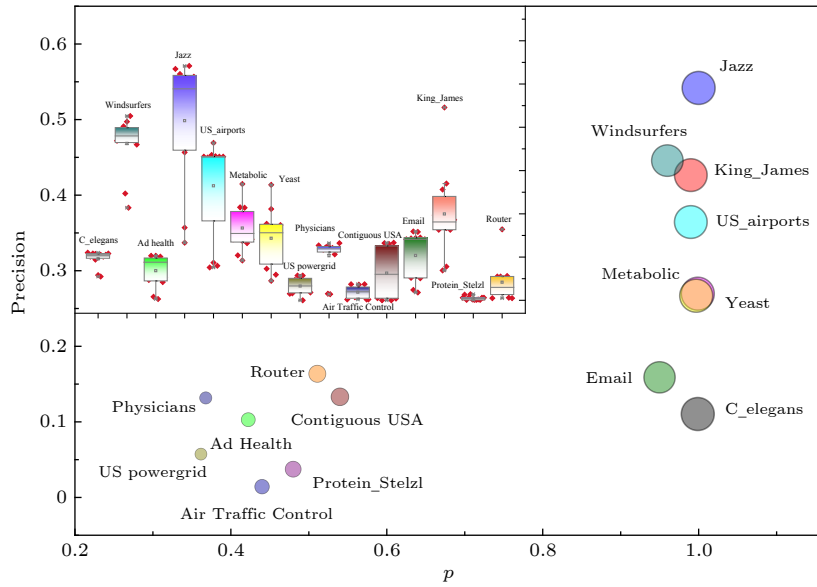


图 5 预测算法在真实网络上的表现

Fig. 5. The performance of prediction algorithms in real networks.

侧较大的圆所示. 相比之下, 图中左下方网络的 precision 值较低, 其对应的网络可预测性值也较差. 值得注意的是, *C_elegans* 网络的可预测值很高, 然而实验选取的多种基于结构相似性的链路算法在网络中的 precision 值普遍不高, 我们进一步计算了各类算法在该网络中的 AUC 值, 结果表明, 最大 AUC 值同样来自于 LRW 指标, 达到了 0.907, 各类算法的平均 AUC 值为 0.848. 因为 AUC 值是基于整个边列表, 而这里的 L 是基于前 10% 的边, 说明这些算法对于该网络的正确预测大部分来自于边列表的后半段, 测试的算法在该网络上的预测效率均较低, 对于该网络有待挖掘能够快速找到更多正确连边的预测算法. 图 5 里的箱线图则显示了对每个网络预测算法之间的比较结果. 我们发现, 除了少数网络各个预测算法的表现基本维持在同一水准外, 大多数预测算法在网络中的效果差别很大, 这也体现出网络和算法选择与匹配问题的重要性. 例如, 在 Jazz 网络中, CN, AA, RA 等指标的 precision 都达到了 0.5 左右, 然而, 优先链接指标 PA 的精确度却很差, 仅为 0.133. 该结果说明, 对于 Jazz 网络来说, 基于共同邻居的算法更能捕获爵士音乐家间的合作关系, 且这种关系不是以优先链接的形式展开的, 所以由优先链接思想演化而来, 依靠节点度乘积来刻画相似性的 PA 指标与 Jazz 网络并不匹配. 事实上, 不仅是 Jazz 网络, PA 算法在很多真实网络中的 precision 值都低于

其他算法, 这说明真实网络在演化生长的过程中往往表现出集聚性、社团性、无标度性、小世界性等多种结构特征, PA 算法虽然能够在无标度网络中表现突出, 其相对单一的相似性刻画思想难以胜任网络结构特征比较复杂的真实网络. 类似的情况还出现在 Router 网络中, 该网络中大多数算法的 precision 值都低于 0.05, 只有 ACT 算法的 precision 达到了 0.164, 这说明 ACT 的算法思想能够更好地匹配该网络更多的结构特征. 结合表 3 和箱线图进行讨论分析, 能够为网络与算法间的匹配和选择问题提供帮助.

5 结论与展望

在网络科学和信息科学领域中, 我们常常会遇到信息缺失的情况. 链路预测作为数据挖掘领域重要的研究方向之一, 是一个长期存在的挑战和难题. 近年来有关链路预测理论和实证的研究发展迅速, 大量研究工作的重心都放在提出算法本身上, 各种各样精确度越来越高的算法层出不穷. 但是, 大量实验结果表明, 在同一网络中不同算法的精度有好有坏. 因此, 到底是不可预测的网络, 还是不合适的链路预测方法, 是一个很有挑战性问题.

本文从统计物理中图的可扩张性得到启发, 提出了一种基于特征谱的链路可预测性度量指标. 通过对网络特征谱的分析, 构造了一个指标来评价网络缺失链路的“可被预测的程度”. 模型网络和大量

真实网络中的实验结果证明, 该指标能够有效地刻画网络的链路可预测性, 且能够就链路预测算法选择提供建议. 例如, 随机网络的可预测性较差, 而无标度网络的可预测性较好. 然而, 虽然各类算法在无标度网络中的精确度明显优于随机网络, LHN-II 算法却不是适用于该网络的合适的链路预测算法; 在实证网络中, Jazz 网络具有较好的可预测性, 一些基于共同邻居的相似性指标如 CN, AA, RA 表现较好, 然而优先链接指标 PA 的精确度却很差. 这是因为网络中音乐家之间的合作关系不是以优先链接的形式展开的, 依靠节点度乘积来刻画相似性的 PA 算法不是适合 Jazz 网络的链路预测算法. 事实上, 我们认为网络的可被预测的程度差并不绝对, 可预测性差的网络也许只是没能遇到理解它结构特征的链路预测算法. 一个好的链路预测算法背后往往有一套贴近网络生长演化的连边机制, 同样道理, 一个重要的机制往往能够提取出一种精确的链路预测算法. 链路预测的研究与网络的结构和演化密切相关, 即算法与网络连边形成机制相辅相成, 是互通的, 网络的链路可预测性即是网络连边演化机制的另一种表现形式. 我们把基于特征谱视角计算可预测性, 获取网络能够被预测的难易程度的工作视作基础, 在下一步的研究中, 拟通过对具有典型演化机制的网络进行分析, 说明预测背后的主要机制以及预测正确或者错误的原因, 去探索一些因果关系. 同时, 考虑基于特征谱挖掘和学习不同网络的拓扑结构信息, 对网络拓扑结构进行标记分类, 针对不同类型网络的连边机制, 提出与之相匹配的链路预测算法.

参考文献

- [1] Albert R, Jeong H, Barabási A L 2000 *Nature* **406** 378
- [2] Albert R, Barabási A L 2002 *Rev. Mod. Phys.* **74** 47
- [3] Newman M E J 2003 *SIAM Rev.* **45** 167
- [4] Wang X F 2002 *Int. J. Bifurcat. Chaos* **12** 885
- [5] Hou L L, Lao S Y, Xiao Y D, Bai L 2015 *Acta Phys. Sin.* **64** 188901 (in Chinese) [侯绿林, 老松杨, 肖延东, 白亮 2015 物理学报 **64** 188901]
- [6] Lü L L 2010 *J. Univ. Electron. Sci. Technol. China* **39** 651 (in Chinese) [吕琳媛 2010 电子科技大学学报 **39** 651]
- [7] Lü L L, Zhou T 2013 *Link Prediction* (Beijing: Higher Education Press) p41 (in Chinese) [吕琳媛, 周涛 2013 链路预测 (北京: 高等教育出版社) 第 41 页]
- [8] Sarukkai R R 2010 *Comput. Networking* **33** 377
- [9] Clauset A, Moore C, Newman M E J 2008 *Nature* **453** 98
- [10] Guimerá R, Marta S P 2009 *Proc. Natl. Acad. Sci. U.S.A.* **106** 22073
- [11] Pan L M, Zhou T, Lü L Y, Hu C K 2016 *Sci. Rep.* **6** 22955
- [12] Taskar B, Wong M F, Abbeel P, Koller D 2003 *Proceedings of the 16th International Conference on Neural Information Processing Systems* (Cambridge: MIT Press) pp659-666
- [13] David L N, Kleinberg J 2007 *J. Am. Soc. Inf. Sci. Technol.* **58** 1019
- [14] Zhou T, Lü L Y, Zhang Y C 2009 *Eur. Phys. J. B* **71** 623
- [15] Xu X K, Fang J Q 2010 *Complex Syst. Complex Sci.* **7** 116 (in Chinese) [许小小, 方锦清 2010 复杂系统与复杂性科学 **7** 116]
- [16] Tan S Y, Wu J, Lü L Y, Li M J, Lu X 2016 *Sci. Rep.* **6** 22916
- [17] Amaral L A N 2008 *Proc. Natl. Acad. Sci. U.S.A.* **105** 6795
- [18] Menche J, Sharma A, Kitsak M, Ghiassian S D, Vidal M, Loscalzo J, Barabási A L 2015 *Science* **347** 1257601
- [19] Lü L Y, Medo M, Yeung C H, Zhang Y C, Zhang Z K, Zhou T 2012 *Phys. Rep.* **519** 1
- [20] Zhou Y X, Lü L Y 2012 *J. Univ. Electron. Sci. Technol. China* **41** 163 (in Chinese) [朱郁筱, 吕琳媛 2012 电子科技大学学报 **41** 163]
- [21] Liu H K, Lü L Y, Zhou T 2011 *Scientia Sinica: Phys. Mech. Astron.* **41** 816 (in Chinese) [刘宏鲲, 吕琳媛, 周涛 2011 中国科学: 物理学 力学 天文学 **41** 816]
- [22] Zhang Q M, Lü L Y, Wang W Q, X Y, Zhou T 2013 *PLoS One* **8** 1
- [23] Wang W Q, Zhang Q M, Zhou T 2012 *EPL* **98** 28004
- [24] Zhang Q M, Xu X K, Zhu Y X, Zhou T 2015 *Sci. Rep.* **5** 10350
- [25] Lü L Y, Zhou T 2011 *Physica A* **390** 1150
- [26] Yu H, Liu Z, Li Y J, Yi C 2016 *Acta Phys. Sin.* **65** 020501 (in Chinese) [于会, 刘尊, 李勇军, 尹超 2016 物理学报 **65** 020501]
- [27] Xu X K, Xu S, Zhu Y X, Zhang Q M 2014 *Complex Syst. Complex Sci.* **11** 41 (in Chinese) [许小小, 许爽, 朱郁筱, 张千明 2014 复杂系统与复杂性科学 **11** 41]
- [28] Lü L Y, Pan L M, Zhou T, Zhang Y C, Stanley H E 2015 *Proc. Natl. Acad. Sci. USA* **112** 2325
- [29] Yin L K, Zheng H Y, Bian T, Deng Y 2014 *Physica A* **482** 699712
- [30] Hanley J A, McNeil B J 1982 *Radiology* **143** 29
- [31] Herlocker J L, Konstan J A, Terveen L G, Riedl J T 2004 *ACM Trans. Inf. Syst.* **22** 5
- [32] Zhou T, Ren J, Matúš M, Zhang Y C 2007 *Phys. Rev. E* **76** 046115
- [33] Farkas I J, Derényi I, Barabási A L, Vicsek T 2001 *Phys. Rev. E* **64** 026704
- [34] Estrada E, Hatano N, Benzi M 2012 *Phys. Rep.* **514** 89
- [35] Newman M E J 2006 *Phys. Rev. E* **74** 036104
- [36] Kousik D, Sovan S, Madhumangal P 2018 *Soc. Netw. Anal. Min.* **8** 1
- [37] Zhang J H, Shen Y Z, Li Y Y, Sun J, Li X X 2017 *Acta Phys. Sin.* **66** 188901 (in Chinese) [张金浩, 申玉卓, 李艳雨, 孙娟, 李晓霞 2017 物理学报 **66** 188901]
- [38] Wang R, Lin P, Liu M X, Wu Y, Zhou T, Zhou C S 2019 *Phys. Rev. Lett.* **123** 038301
- [39] Estrada E 2006 *EPL* **73** 649
- [40] Tan S Y, Wu J, Li M J, Lu X 2016 *EPL* **114** 58002
- [41] Estrada E, Hatano N 2007 *Chem. Phys. Lett.* **439** 247

Link predictability of complex network from spectrum perspective^{*}

Tan Suo-Yi¹⁾ Qi Ming-Ze²⁾ Wu Jun^{3)†} Lu Xin^{1)‡}¹⁾ (*College of Systems Engineering, National University of Defense Technology, Changsha 410073, China*)²⁾ (*College of Liberal Arts and Sciences, National University of Defense Technology, Changsha 410073, China*)³⁾ (*International Academic Center of Complex Systems, Beijing Normal University, Zhuhai 519087, China*)

(Received 3 November 2019; revised manuscript received 27 February 2020)

Abstract

Link prediction in complex networks has attracted much attention in recent years and most of work focuses on proposing more accurate prediction algorithms. In fact, “how difficultly the target network can be predicted” can be regarded as an important attribute of the network itself. In this paper it is intended to explain and characterize the link predictability of the network from the perspective of spectrum. By analyzing the characteristic spectrum of the network, we propose the network link predictability index. Through calculating the index, it is possible to learn how difficultly the target network can be predicted before choosing algorithm, and to solve the problem whether the network is unpredictable or the algorithm is inappropriate. The results are useful for the selecting and matching the complex network and link prediction algorithms.

Keywords: link predictability, link prediction, spectrum theory, complex network**PACS:** 89.75.Hc, 89.75.Fb, 64.60.aq**DOI:** [10.7498/aps.69.20191817](https://doi.org/10.7498/aps.69.20191817)

^{*} Project supported by the National Natural Science Foundation of China (Grant Nos. 882041020, 71771213, 71901067, 71871217) and the Science and Technology Plan Project of Hunan Province, China (Grant Nos. 2017RS3040, 2018JJ1034, 2019JJ20019).

[†] Corresponding author. E-mail: wujunpla@hotmail.com

[‡] Corresponding author. E-mail: xin.lu@flowminder.org

复杂网络上的部分同步化：奇异态、遥同步与集团同步

王振华 刘宗华

Partial synchronization in complex networks: Chimera state, remote synchronization, and cluster synchronization

Wang Zhen-Hua Liu Zong-Hua

引用信息 Citation: *Acta Physica Sinica*, 69, 088902 (2020) DOI: 10.7498/aps.69.20191973

在线阅读 View online: <https://doi.org/10.7498/aps.69.20191973>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

复杂网络谱粗粒化方法的改进算法

Improved algorithm of spectral coarse graining method of complex network

物理学报. 2017, 66(6): 060502 <https://doi.org/10.7498/aps.66.060502>

基于相对距离的复杂网络谱粗粒化方法

A spectral coarse graining algorithm based on relative distance

物理学报. 2019, 68(10): 100501 <https://doi.org/10.7498/aps.68.20181848>

知识图谱复杂网络特性的实证研究与分析

Empirical study of knowledge network based on complex network theory

物理学报. 2019, 68(12): 128902 <https://doi.org/10.7498/aps.68.20190106>

基于复杂网络动力学模型的无向加权网络节点重要性评估

Evaluation methods of node importance in undirected weighted networks based on complex network dynamics models

物理学报. 2018, 67(9): 098901 <https://doi.org/10.7498/aps.67.20172295>

复杂网络链路可预测性: 基于特征谱视角

The link predictability of complex network from a spectrum perspective

物理学报. 2020, 69(8): 1 <https://doi.org/10.7498/aps.69.20191817>

专题：统计物理和复杂系统

复杂网络上的部分同步化：奇异态、遥同步与集团同步*

王振华 刘宗华†

(华东师范大学物理与电子科学学院, 上海 200241)

(2019年12月26日收到; 2020年2月14日收到修改稿)

近年来, 复杂网络上耦合振子的部分同步化引起了人们极大的关注, 其潜在或背后的原因是部分同步化斑图在大脑网络中广泛存在, 并很可能与大脑的认知或记忆等功能有密切的联系. 本文对这些进展进行简单的总结与归纳, 并按照学者们研究的不同侧重点, 将其分成三方面来进行介绍, 即奇异态、遥同步与集团同步化. 着重强调这三种情形各自出现的条件、常用的研究模型、检测的方法以及侧重解释的生物现象等方面, 并对它们三者之间的相互关系及今后的研究方向做一些简单的探讨.

关键词: 复杂网络, 奇异态, 遥同步, 集团同步**PACS:** 89.75.Hc, 05.45.Xt, 68.18.Jk**DOI:** 10.7498/aps.69.20191973

1 引言

混沌通常被认为是有害的现象, 直到 1990 年美国海军实验室的 Pecora 和 Carroll^[1] 发现两个全同的混沌系统实际上是可以同步化的, 从而导致了混沌系统的一个极其重要的应用——保密通信^[2]. 源于这个应用, 两个耦合系统间的混沌同步化受到了极大的关注, 并发现了一系列的同步化方式, 如完全同步、相同步、延迟同步及广义同步等. 随后, 这些研究被推广到了多个耦合系统的情形, 包括耦合映像格子的时空同步化等. 这些同步化的一个共同特征是耦合的各子系统间行为一致, 是作为一个整体来表现同步特征的.

实际上, 除了这些整体同步化外, 还存在着以非整体行为的方式体现的同步化. 比如, 在猫的视觉皮层中, 分布在不同空间位置的神经元与它们的振荡响应 (40—60 Hz) 是同步的^[3]. 然而, 这个现象直至十年之后, 也就是 2000 年左右, 才引起了

人们的广泛关注, 被称为部分同步化. 部分同步主要讨论两个振子以上但总振子数目不是太大的情形. 此时只有部分态变量是同步的, 而其他态变量并不与它们同步, 因此系统的轨道将收敛到一个由同步变量定义的不变线性子空间^[4–7]. 比如考虑由三个振子 x, y, z 组成的耦合混沌系统, 部分同步可以发生在振子 x 与 y 之间或 x 与 z 之间, 而第三个振子并不与这两者同步. 部分同步化已获得了实验验证^[8,9].

上述研究, 无论是同步还是部分同步, 包含的子系统数目都不大, 通常只有 2—6 个子系统. 另外, 这些耦合的连接方式是均匀的, 比如最近邻连接、规则连接或者全连等, 而不涉及复杂的网络结构, 因此在理解大脑的各种同步斑图方面帮助不大. 最近, 部分同步的研究有了长足的进展, 不仅仅是子系统数目扩大到成百上千的规模, 而且连接的方式也推广到多种多样^[10,11], 从分组连接到复杂网络连接, 甚至到了多层网络连接^[12,13]. 基于这些研究, 人们发现复杂网络上的部分同步可以展现出

* 国家自然科学基金 (批准号: 11675056, 11835003) 资助的课题.

† 通信作者. E-mail: zhliu@phy.ecnu.edu.cn

复杂多变、多姿多彩的斑图,这就对大脑斑图的多样性给出了一种可能的机制.我们知道,大脑皮层的大尺度同步^[14]与网络神经元子集同步放电^[15,16]均与认知过程密切相关,且额叶脑区内任务相关的 α 波同步与顶底加工(top-down processing)的需求也高度相关^[17].另外,神经活动的同步斑图还与大脑无序紧密相关.比如,癫痫发作时,特定脑区会变得高度同步,其他部分则不同步^[18].而在失去了部分脑细胞导致的帕金森氏病中,同步活动将会在特定的脑区缺失^[19].因此,复杂网络上部分同步化的研究为理解大脑的高级功能,如认知与记忆等打开了一扇新的窗口,亟待进一步深入下去.

由于复杂网络上部分同步的复杂性,历史上,不同的研究小组从不同的角度得到了不同的发现,并依据具体出现条件的不同而将其分别命名为奇异态、遥同步及集团同步等.奇异态是最先命名的,描述的是同步振子与无序振子的共存现象,可对半脑睡眠与首晚效应等生物现象做出合理的解释.遥同步是最后命名的,描述的是不直接相连的振子之间通过中心节点发生同步但却不与中心节点上的振子同步的现象,可用于帮助理解大脑分布式计算的内在机制.而集团同步化居中,描述的是复杂网络上的部分振子因具有某种对称性而形成同步集团的现象,可帮助理解大脑同步斑图的多样性.虽然这三种部分同步化状态发现的初始时间略有不同,但对它们研究的进展并没有绝对的先后,而是交替前行、相互促进的.实际上,这三种部分同步化状态并不是完全独立的,甚至有时描述的是同一种现象或斑图.本文将对这些进展进行综述,并对它们间的相互关系进行简单的讨论.

2 奇异态

奇异态描述的是同步与无序这一对矛盾体共存的特殊情形,是Abrams与Strogatz^[20]于2004年命名的.其英文名字为Chimera state,代表希腊神话中的一种狮头、羊身、蛇尾的喷火怪物.这个名字取得非常好,不仅仅在于它把这个领域做“热”了,更重要的是它描述了一大类生物系统半脑睡眠现象的微观机制.来自于实验探测的EEG数据表明,在一些鸟类与海洋哺乳动物的睡眠过程中,它们的一半大脑是同步的,而另一半大脑是非

同步的,这就使得一半大脑比另一半具有更强的警觉性,从而减少睡眠时所受到的外部攻击,称为“半脑睡眠”^[21,22].众所周知,人类早期住在山洞而没有房子保护时也需要提防野兽的攻击,因而这种警觉性或许到目前应该还有某些残留痕迹.这种推断直到2016年才被实验证实,称为“首晚效应”^[23].说的是当人们出差到异地时,第一晚的睡眠通常不踏实,而是表现为左边半脑觉醒的次数要远大于右边半脑,但第二晚以后会恢复正常.

另一方面,奇异态的研究也拓展并丰富了非线性动力学及耦合振子理论.长期以来,人们一直相信同步与非同步的共存是限制在非均匀振子网络中的,其中具有相似频率的振子可以发生同步,而那些偏离平均频率较大的振子则保持不相关的漂移.因此,当发现全同振子在相同的耦合方式下也可以分离成为同步与漂移的两组,即奇异态时,就从根本上令人感到吃惊,因为全同振子通常是作为整体,表现为要么同步要么非同步的.这样,作为一种新的动力学状态的奇异态就实质性地拓展了我们对从同步到混乱的转变或相反的转变的认识.

事实上,奇异态这种现象在其命名的两年前就由Kuramoto与Battogtokh^[24]在Ginzburg-Landau方程中注意到了,但当时并没有引起人们足够的重视,这也从一个侧面说明了命名的重要性.在这个发现中,一个极其重要的因素是,不同于以往的全局耦合或最近邻耦合,甚至也不同于随机耦合,作者采用了一种随距离衰减的指数耦合,称为非局域耦合(nonlocal coupling).虽然这种非局域耦合没有局域或全局耦合那么受到重视,但它其实也有广泛的应用,比如在约瑟夫逊结、化学振子、蜗牛壳模式的神经网络及眼优势条纹中等. Abrams与Strogatz^[20]继承了这个非局域耦合的思路,但代之以Kuramoto相振子模型.此外,他们还将指数衰减的耦合核改为余弦耦合核,使得模型可以解析求解.

研究奇异态最常用到的是Kuramoto相振子模型,这个模型因只含振荡的相位特征却没有涉及振荡的振幅而被认为是描述自然界各种振荡行为最简单的模型. Abrams与Strogatz^[20]考虑的模型就是这样一个相振子环,其方程为

$$\frac{\partial \phi}{\partial t} = \omega - \int_{-\pi}^{\pi} G(x-x') \sin[\phi(x,t) - \phi(x',t) + \alpha] dx', \quad (1)$$

这里 $\phi(x, t)$ 为振子在位置 x 时间 t 时的相, 角度 $0 \leq \alpha \leq \pi/2$ 为可调参数. 耦合核 $G(x - x')$ 假定取如下形式:

$$G(x) = \frac{1}{2\pi}(1 + A \cos x), \quad (2)$$

其中 $0 \leq A \leq 1$. 引入序参量来对局部振子的有序与无序进行描述. 令 Ω 代表旋转框架的角频率, $\theta = \phi - \Omega t$ 代表振子相对于这个框架的相, 于是有

$$R(x, t)e^{i\Theta(x, t)} = \int_{-\pi}^{\pi} G(x - x')e^{i\theta(x', t)} dx', \quad (3)$$

其中 R 为序参量, Θ 为局部平均相. 此时, 方程 (1) 变为

$$\frac{\partial \theta}{\partial t} = \omega - \Omega - R \sin[\theta - \Theta + \alpha]. \quad (4)$$

稳态时, R 与 Θ 将独立于时间, 而只依赖于空间. 作参数变换: $\beta = \pi/2 - \alpha$ 与 $\Delta = \omega - \Omega$. 图 1(a) 为方程 (4) 的快照, 其中靠近 $x = \pm\pi$ 的振子是锁相与相关的, 即它们以相同的瞬时频率运动并处于近似的同相. 同时, 图 1(a) 中间的分散振子是漂移与无序的. 图 1(b) 则给出了对应图 1(a) 的序参量曲线, 证实同步态具有较高的 R 值, 而无序态具有较小的 R 值. 图 1(c) 则为对应的局部平均相 $\Theta(x)$.

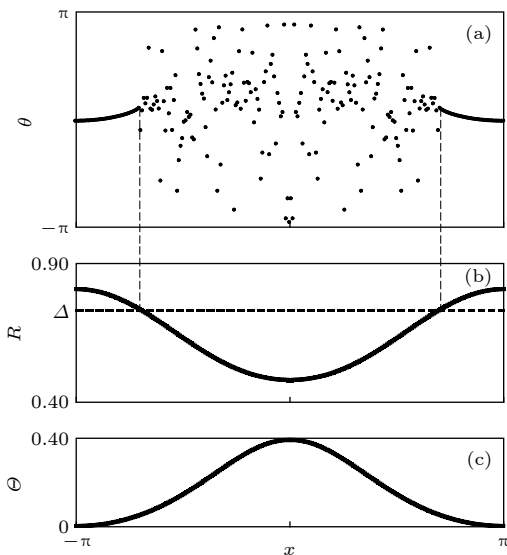


图 1 (a) 奇异态的斑图, 参数: $A = 0.995$, $\beta = 0.18$, $N = 256$ 个振子; (b) 来自于方程 (3) 的局部相关度 $R(x)$, 锁相的振子满足 $R(x) \geq \Delta$; (c) 局部平均相 $\Theta(x)$. 改编自文献 [20]
 Fig. 1. (a) Phase pattern for a chimera state, parameters: $A = 0.995$, $\beta = 0.18$, $N = 256$ oscillators; (b) local phase coherence $R(x)$, computed from Eq. (3), locked oscillators satisfy $R(x) \geq \Delta$; (c) local average phase $\Theta(x)$. Figure adapted from Ref. [20].

随后, Abrams 等 [25] 于 2008 年提出了一个仅包含两组相互作用的振子集团且精确可解的奇异态模型. 这个模型人为地将振子分成大小相等的两组, 各组内与组间的耦合方式完全一样, 通过初值条件选取的不同, 使得它们一个处于同步态, 而另一个处于无序态. 通过 Ott 与 Antonsen [26] 的降维方法, 他们得到了关于奇异态稳定性、动力学及分岔的精确解. 其模型如下 [25]:

$$\frac{d\theta_i^\sigma}{dt} = \omega + \sum_{\sigma'=1}^2 \frac{K_{\sigma\sigma'}}{N_{\sigma'}} \sum_{j=1}^{N_{\sigma'}} \sin(\theta_j^{\sigma'} - \theta_i^\sigma - \alpha), \quad (5)$$

其中 $\sigma = 1, 2$, N_σ 为第 σ 组内的振子数目. 因假定全同振子, 因此频率 ω 与相延迟 α 对所有振子都是一样的. 组内的耦合强度设定为 $K_{11} = K_{22} = \mu > 0$, 组间的耦合强度则取为 $K_{12} = K_{21} = \nu > 0$, 且 $\mu > \nu$, 即组内的耦合大于组间的耦合. 这有点类似于上面的随距离衰减的非局域耦合. 对时间做标度变换, 可设定 $\mu + \nu = 1$. 进一步为方便寻找奇异态存在的区域, 定义参数 $A = \mu - \nu$ 与 $\beta = \pi/2 - \alpha$.

对于随机选择的初值, 方程 (5) 很容易走向同步态. 但当初值设置在奇异态附近时, 比如一组为近似同步的初值, 另一组为随机初值, 则能观察到奇异态 [25]. 图 2 为奇异态的动力学, 只画了非同步那一组的序参量, 其满足 $r(t) = |\langle e^{i\theta_j(t)} \rangle|$. 从图 2 (a)—图 2 (c) 可见, 对应不同的参数值 A , 序参量分别呈现常数值、小振动及大振荡行为, 分别表示稳定的奇异态、呼吸奇异态及长周期呼吸子.

当取其他的 β 时, 也可能出现稳定的奇异态与呼吸奇异态, 但所需的耦合强度差 A 会有所不同. 借助于 OA 降维方法, Abrams 等 [25] 得到了依赖于参数 β 与 A 的解析解, 并给出了各稳定区域间的边界曲线. 图 3 是相应的相图.

上述奇异态一旦形成, 便不再改变, 因此是稳定的奇异态. 为了实现海豚与鸟类在睡眠中出现的交替奇异态, Ma 等 [27] 引入了一个周期性的延迟信号来描述变化的外部环境. 具体做法是将方程 (5) 改写为

$$\frac{d\theta_i^\sigma}{dt} = \omega_i + \sum_{\sigma'=1}^2 \frac{K_{\sigma\sigma'}}{N_{\sigma'}} \sum_{j=1}^{N_{\sigma'}} \sin(\theta_j^{\sigma'}(t) - \theta_i^\sigma(t) - \alpha) + A \sin[\Omega(t - \tau_\sigma)], \quad (6)$$

其中 A 与 Ω 分别为外部信号的振幅与频率, τ_σ 代

表系统对环境的响应时间. 考虑到睡觉半脑与清醒半脑对外界应该有不同的响应, 可令 $\tau_1 = 0$ 与 $\tau_2 \neq 0$. 耦合参数取为 $\mu = 5$ 与 $\nu = -4$. 负耦合 $\nu = -4$ 来自于组间的相互作用, 不同于组内的相互

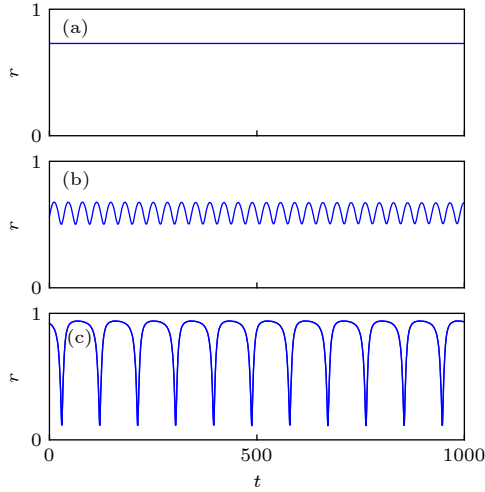


图 2 序参量 r 随时间的变化, 其中参数 $N_1 = N_2 = 128$, $\beta = 0.1$ (a) $A = 0.20$, 稳定奇异态; (b) $A = 0.28$, 呼吸奇异态; (c) $A = 0.35$, 长周期呼吸子. 初值条件取为靠近奇异态, 结果为丢掉了 2000 个暂态时间单位后的情形. 改编自文献 [25]

Fig. 2. Order parameter r versus time. In all three panels, $N_1 = N_2 = 128$ and $\beta = 0.1$: (a) $A = 0.20$, stable chimera; (b) $A = 0.28$, breathing chimera; (c) $A = 0.35$, long period breather. Numerical integration began from an initial condition close to the chimera state, and plots shown begin after allowing a transient time of 2000 units. Figure adapted from Ref. [25].

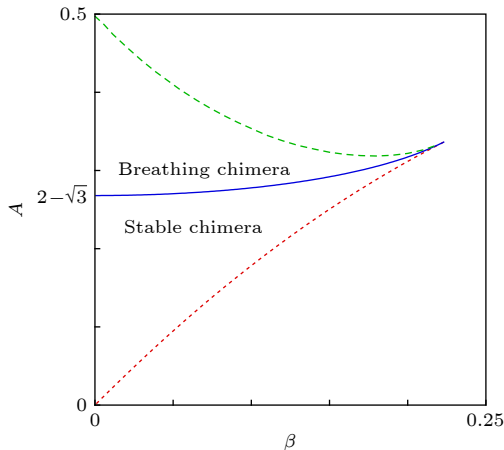


图 3 奇异态的稳定性相图 分岔曲线: 鞍-节分岔 (点划线)、超临界 Hopf 分岔 (实线), 两者都是解析发现的; 同宿分岔 (虚线), 数值发现的. 改编自文献 [25]

Fig. 3. Stability diagram for chimera states. Bifurcation curves: saddle-node (dotted line) and supercritical Hopf (solid line), both found analytically; homoclinic (dashed line), found numerically. Figure adapted from Ref. [25].

作用, 其依据是视交叉上核中左右振荡组之间是反相同步 [28]. 另外, 由于真实大脑中神经元的非全同, 可令 ω_i 随机均匀地分布在 $[1 - \delta, 1 + \delta]$, 其中 δ 代表振子的多样性. Ma 等 [27] 发现, 对一个固定的 δ , 存在一个阈值 A_0 , 以便 $A \geq A_0$ 时可以出现交替的奇异态. 比如当 $\delta = 0.1$ 时, 有 $A_0 = 0.2$, 且随着 δ 的增加, A_0 会逐渐降低. 图 4 给出了 $N_1 = N_2 = 128$ 与 $\delta = 0.1$ 时的结果, 其中图 4(a) 代表 $A = 1$, $\Omega = 1$ 与 $\tau_2 = 1.5$ 的情形; 图 4(b) 为 $A = 1$, $\Omega = 0.5$ 与 $\tau_2 = 1.5$ 的情形; 图 4(c) 为 $A = 1.5$, $\Omega = 1$ 与 $\tau_2 = 1.5$ 的情形; 图 4(d) 为 $A = 1$, $\Omega = 1$ 与 $\tau_2 = 3.0$ 的情形. 从图 4 可见, 交替奇异态存在于所有的四种情形中, 表明交替奇异态对信号参数具有鲁棒性.

当网络结构不是全连网, 而是二维或三维规则格子甚至复杂网络时, 奇异态会表现出多种多样性, 其斑图丰富多彩 [29–33]. Zhu 等 [34] 研究了复杂网络情形, 发现此时的斑图会比较混乱, 但如果按振子的有效频率进行重新排序, 则仍可见清晰的同步与无序共存. 详细的信息可见综述文献 [10, 11, 35].

有鉴于此, 为了定量刻画奇异态, Kemeth 等 [36] 于 2016 年引入了一个特征分类方法来测量奇异态. 其基本思想是测量空间关联, 即对于具有局部或非局部耦合拓扑的系统, 采用局部曲率来测量空间关联. 据此, 一维系统时局域曲率可由二阶导数测量, 一维以上的系统则由拉普拉斯算子来完成. 具体步骤如下.

i) 对于包含空间数据 f 的每一个时间步或快照, 我们用分立的拉普拉斯算子 D 来计算每个节点的局部曲率. 以一维情形为例, 对时间 t 时的快照, 我们计算

$$\hat{D}f = f(x + \Delta x, t) - 2f(x, t) + f(x - \Delta x, t). \quad (7)$$

以图 5(a) 所示的快照为例, 按方程 (7) 计算出的 $|\hat{D}f|$ 如图 5(b) 所示, 其中 D_m 代表 \hat{D} 的最大值. 对相振子系统, 可将这个算子用于复平面的数据, 即相振子处于一个常数振幅 A 的圆环上. D_m 代表特定振子的曲率, 其两个邻居都在相平面中转移 180° , 即两个邻居振子在圆环上都处于相对立的位置. 在连续极限下, D_m 收敛到 $4A$. 在同步化区域, $\lim_{N \rightarrow \infty} |\hat{D}| = 0$. 因此, 通过这个变换, 同步化区域被投影到 x -轴, 而在非相关区域 $|\hat{D}|$ 有限且显示明显的涨落.

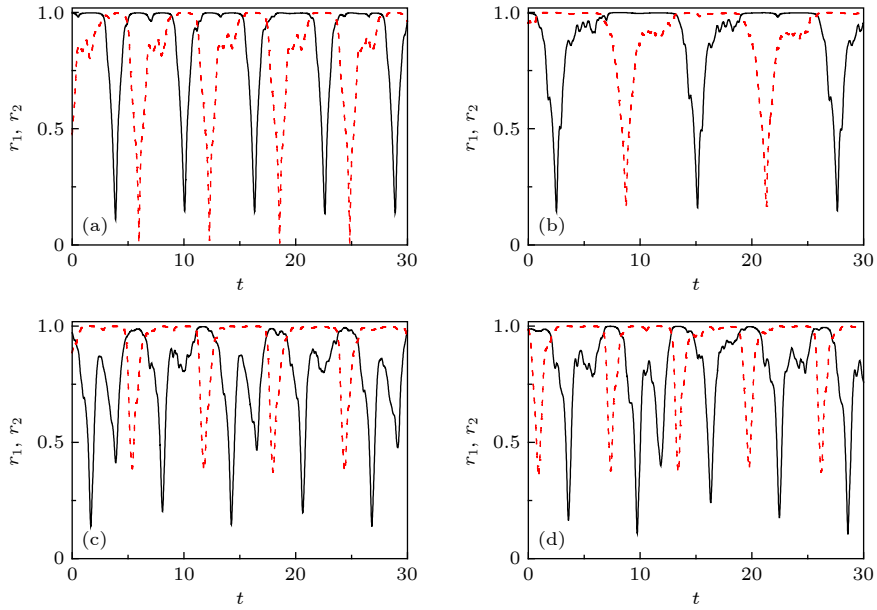


图 4 外部信号诱导的交替奇异态 (a) $A = 1, \Omega = 1$ 与 $\tau_2 = 1.5$; (b) $A = 1, \Omega = 0.5$ 与 $\tau_2 = 1.5$; (c) $A = 1.5, \Omega = 1$ 与 $\tau_2 = 1.5$; (d) $A = 1, \Omega = 1$ 与 $\tau_2 = 3.0$. 改编自文献 [27]

Fig. 4. Alternating chimera states induced by external signals: (a) $A = 1, \Omega = 1$ and $\tau_2 = 1.5$; (b) $A = 1, \Omega = 0.5$ and $\tau_2 = 1.5$; (c) $A = 1.5, \Omega = 1$ and $\tau_2 = 1.5$; (d) $A = 1, \Omega = 1$ and $\tau_2 = 3.0$. Figure adapted from Ref. [27].

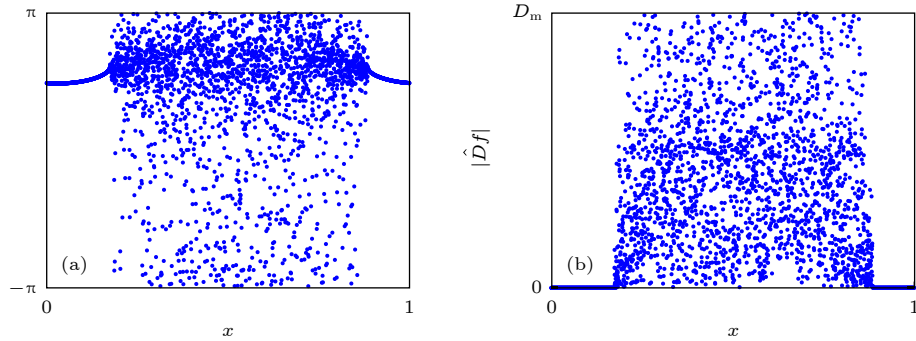


图 5 (a) 来自于 Kuramoto 模型的快照; (b) 将拉普拉斯算子用于 (a) 中的数据后得到的局部曲率的绝对值. 改编自文献 [36]

Fig. 5. (a) Snapshot of the Kuramoto model; (b) absolute value of the local curvature obtained by applying the discrete Laplace operator on the data set shown in (a). Figure adapted from Ref. [36].

ii) 计算 $|\hat{D}|$ 的归一化几率密度函数 $g(|\hat{D}| = 0)$ 代表每个时间步实现中空间关联区域所占的相对大小. 对于完全同步的系统有 $g(|\hat{D}| = 0) = 1$, 而完全无序的系统则有 $g(|\hat{D}| = 0) = 0$. 当 $g(|\hat{D}|)$ 取 0—1 之间的值时, 就表示同步与无序的共存.

iii) 考虑到数值模拟中的误差涨落, Kemeth 等 [36] 建议只要局部曲率的绝对值小于最大曲率的百分之一就应该被当作相关态, 其他的为非相关态. 为此, 他们引入关联测度

$$g_0(t) = \int_0^\delta g(t, |\hat{D}|) d|\hat{D}|, \quad (8)$$

其中 $\delta = 0.01D_m$. 实际计算中可用 $g_0(t)$ 的时间平均值来衡量奇异态.

奇异态概念最直接的应用就是神经系统, 特别是大脑网络, 目前这方面的研究很丰富 [37–44]. 不同于一维的相振子模型, 神经元系统通常由二维或三维模型来表示, 此时的相互作用由耦合矩阵表示. 研究中比较常用到的是 FitzHugh-Nagumo (FHN) 神经元模型. Omelchenko 等 [37] 为 FHN 模型引入了一种旋转耦合, 并发现了多奇异态. 他们将 N 个神经元放置在一个圆环上, 每个神经元与其最近邻的 $2R$ 个神经元发生耦合. 其动力学方程如下:

$$\begin{aligned} \epsilon \frac{du_k}{dt} &= u_k - \frac{u_k^3}{3} - v_k + \frac{\sigma}{2R} \sum_{j=k-R}^{k+R} [b_{uu}(u_j - u_k) \\ &\quad + b_{uv}(v_j - v_k)], \\ \frac{dv_k}{dt} &= u_k + a_k + \frac{\sigma}{2R} \sum_{j=k-R}^{k+R} [b_{vu}(u_j - u_k) \\ &\quad + b_{vv}(v_j - v_k)], \end{aligned} \quad (9)$$

其中 u_k 与 v_k 分别表示快变量与慢变量, ϵ 为一个小小参数, 这里取为 0.05. 依赖于阈值 a_k , 每个单独的 FHN 振子显示振荡 ($|a_k| < 1$) 或可激 ($|a_k| > 1$) 行为, 这里取 $a_k \in (-1, 1)$. 为简单起见, 他们将耦合取为旋转矩阵

$$B = \begin{pmatrix} b_{uu} & b_{uv} \\ b_{vu} & b_{vv} \end{pmatrix} = \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix}, \quad (10)$$

此耦合依赖于相参数 $\phi \in [-\pi, \pi]$.

令 $r = R/N$, 则参数 r 与 σ 就将决定系统的行为. 图 6(a)—图 6(d) 给出了参数 r 与 σ 分别取 $(0.33, 0.1)$, $(0.33, 0.19)$, $(0.33, 0.23)$ 及 $(0.33, 0.28)$ 时变量 u_k 的快照. 为方便比较, 同时给出了每个振子的平均相速度 $\omega_k = 2\pi M_k / \Delta T$, 其中 M_k 为时间间隔 ΔT 内完成的圈数. 常数 ω_k 代表锁相区域. 从

图 6(a)—图 6(d) 可见, 随着耦合强度 σ 的增加, 奇异态从单一的非相关区域变成了两个非相关区域, 即观察到了多奇异态^[37]. 类似地, 图 6(e)—图 6(h) 给出了从单一的非相关区域变成三个非相关区域的情形.

在大脑网络方面的应用主要是将网络结构考虑成不同大小的真实脑皮层网^[13,45–47], 然后研究这些特定网络结构下的奇异态. 值得一提的是, Kang 等^[13] 发现了对应“首晚效应”或半脑睡眠的奇异态. 他们使用的脑皮层网络数据含有 998 个节点、17865 条边^[48,49], 但其中含有 9 个孤立的没有边的节点. 去掉这些孤立节点后, 网络只有 989 个节点, 其中右半脑含有 496 个节点 ($i = 1–496$), 左半脑含有 493 个节点 ($i = 497–989$). 考虑到左右半脑只能通过胼胝体连接, 脑皮层网实际上是个双层网络, 它的 17865 条边可分成右半脑内的 8037 条边、左半脑内的 7773 条边以及连接左右半脑的 2055 条边.

考虑到通过胼胝体的连边与半脑内的连边具有不同的信号传播速度, Kang 等^[13] 提出了一个双层网模型 (A, B 层) 并让层内与层间的耦合强度不同, 分别由 λ_{in} 与 λ_{out} 表示. A 层的方程为

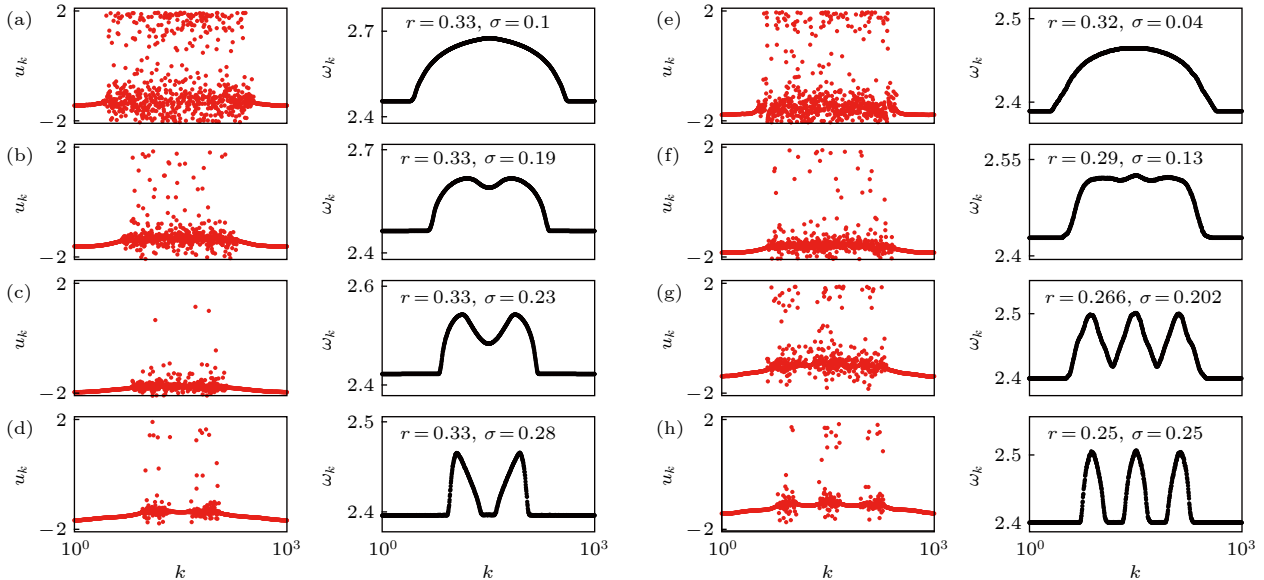


图 6 单一的非相关区域变成多个非相关区域的情形 (a)—(d) 变成两个非相关区域; (e)—(h) 变成三个非相关区域. 在每个子图中, 左列代表变量 u_k 的快照, 右列为对应的平均相速度. 改编自文献^[37]

Fig. 6. Transition from a classical chimera state with one incoherent domain to multichimera states with two (a)–(d), and three (e)–(h) incoherent domains. In each panel the left column shows snapshot of variables u_k , and the right column shows the corresponding mean phase velocities. Figure adapted from Ref. ^[37].

$$\begin{aligned}
 \epsilon \dot{u}_i^a &= u_i^a - \frac{(u_i^a)^3}{3} - v_i^a \\
 &+ \frac{\lambda_{\text{in}}}{k_{\text{in},i}^a} \sum_{j=1}^N \mathbf{A}_{ij} [d_{uu}(u_j^a - u_i^a) + d_{uv}(v_j^a - v_i^a)] \\
 &+ \frac{\lambda_{\text{out}}}{k_{\text{out},i}^a} \sum_{j=1}^N (\mathbf{AB})_{ij} [d_{uu}(u_j^b - u_i^a) + d_{uv}(v_j^b - v_i^a)], \\
 \dot{v}_i^a &= u_i^a + a + \frac{\lambda_{\text{in}}}{k_{\text{in},i}^a} \sum_{j=1}^N \mathbf{A}_{ij} [d_{vu}(u_j^a - u_i^a) + d_{vv}(v_j^a - v_i^a)] \\
 &+ \frac{\lambda_{\text{out}}}{k_{\text{out},i}^a} \sum_{j=1}^N (\mathbf{AB})_{ij} [d_{vu}(u_j^b - u_i^a) + d_{vv}(v_j^b - v_i^a)],
 \end{aligned} \tag{11}$$

其中 $k_{\text{in},i}^a$ 与 $k_{\text{out},i}^a$ 分别代表层内与层间的节点度, \mathbf{A}_{ij} 与 $(\mathbf{AB})_{ij}$ 代表层内与层间的耦合矩阵, $\epsilon = 0.05$, $a = 0.5$. 旋转矩阵 \mathbf{D} 取方程 (10) 一样的形式. 相应地, B 层的方程为

$$\begin{aligned}
 \epsilon \dot{u}_i^b &= u_i^b - \frac{(u_i^b)^3}{3} - v_i^b \\
 &+ \frac{\lambda_{\text{in}}}{k_{\text{in},i}^b} \sum_{j=1}^N \mathbf{B}_{ij} [d_{uu}(u_j^b - u_i^b) + d_{uv}(v_j^b - v_i^b)] \\
 &+ \frac{\lambda_{\text{out}}}{k_{\text{out},i}^b} \sum_{j=1}^N (\mathbf{AB})_{ij} [d_{uu}(u_j^a - u_i^b) + d_{uv}(v_j^a - v_i^b)], \\
 \dot{v}_i^b &= u_i^b + a + \frac{\lambda_{\text{in}}}{k_{\text{in},i}^b} \sum_{j=1}^N \mathbf{B}_{ij} [d_{vu}(u_j^b - u_i^b) + d_{vv}(v_j^b - v_i^b)] \\
 &+ \frac{\lambda_{\text{out}}}{k_{\text{out},i}^b} \sum_{j=1}^N (\mathbf{AB})_{ij} [d_{vu}(u_j^a - u_i^b) + d_{vv}(v_j^a - v_i^b)].
 \end{aligned} \tag{12}$$

这个双层网模型可展示不同的状态, 包括同步化、无序及奇异态. 图 7 给出了四个典型的状态, 其中上层为网络 A 、下层为网络 B 、子图为它们对应的动力学. 非常有趣的是, 图 7(c) 与图 7(g) 显示出网络 A 无序而网络 B 同步, 为典型的半脑睡眠现象, 从而再现了“首晚效应”.

除了这些数值模拟与理论研究外, 奇异态已被数个实验证实 [50–54]. 比如 Martens 等 [52] 简单地使用了两个摆座、一根弹簧和一些全同的节拍器实现了奇异态. 随着弹簧弹力系数的增大, 一个摆座上的节拍器全同步, 另一个摆座的节拍器会出现各种不同的非协同行为. 该实验无需精妙耦合计算和时间延迟, 且能容易地拓展至其他物理模型.

Gambuzza 等 [53] 则使用电路实验模拟了 FHN 模型的奇异态, 发现了一种静止态和同步态混合的奇异态.

3 遥同步

大量的证据表明, 人脑的信息处理是分布式的, 并由具有社区结构的复杂网络承担 [55]. 一个脑区的信息可为其他脑区用来进行认知计算, 这种跨多个脑区的处理过程称为信息的分散与整合 [56]. 在这个分散与整合的过程中, 有一种特殊的信息传递方式—遥同步, 指不直接相连的两个节点通过连接一个公共的邻居而相互间发生同步, 但却不与那个公共邻居同步 [57]. 其证据是在采集的多电极大脑数据中, 无论是 β 带 (14–30 Hz) 还是 γ 带 (40–80 Hz) 都有各脑区间神经活动的长程同步. 由于这些远距离脑区间轴向传导延迟可达到几十毫秒, 这种远距离的同步放电很可能是中继振子介导的 [58]. 事实上, 中继神经元广泛存在于我们的神经系统中, 比如丘脑、脊髓和外侧膝状体. 例如, 丘脑就负责视觉信号从视网膜到视觉皮质的转播. 目前, 遥同步已在部分实验中得到证实 [59].

遥同步是 Bergner 等 [57] 在研究星形网的相同步时引入的. 他们考虑的模型如下:

$$\frac{du_n}{dt} = (\alpha + i\omega_n - |u_n|^2)u_n + \frac{\kappa}{d_n^{\text{in}}} \sum_{m=1}^N g_{nm}(u_m - u_n), \tag{13}$$

其中 u_n 是复数, α 为霍普夫分岔参数, ω_n 为振子未耦合时的本征频率, κ 为耦合强度, d_n^{in} 为节点的入度, (g_{nm}) 为邻接矩阵. 图 8 为他们研究遥同步时的星形网示意图.

以 $N = 5$ 为例, Bergner 等 [57] 取中心节点的频率 $\omega_1 = 2.5$, 4 个叶子节点的频率为 $\{\omega_n\}_{n=2}^5 = \{0.975, 0.992, 1.008, 1.025\}$, 即中心节点的频率远大于叶子节点的频率. 令 $u_n = x_n + iy_n$, 并引入相位 $\phi_n = \tan^{-1}(y_n/x_n)$, 则可定义

$$r_{nm} = |\langle e^{i(\phi_n(t) - \phi_m(t))} \rangle_t|, \tag{14}$$

其中 $\langle \cdot \rangle_t$ 代表时间平均, r_{nm} 为节点 n 与 m 之间的局部序参量. 为了检验遥同步, 可引入两个序参量. 一个用来测量中心节点与叶子节点间的关联

$$r^{\text{direct}} = \frac{1}{N-1} \sum_{n=2}^N r_{1n},$$

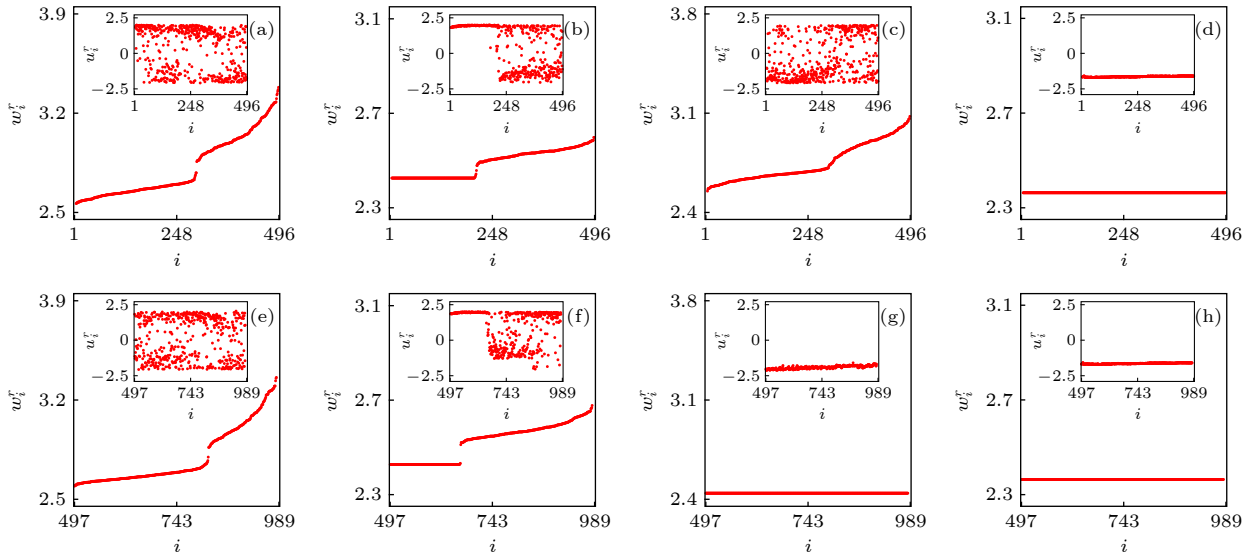


图 7 大脑皮层网中的四个典型的行为. 参数为 $\alpha = \pi/2 - 0.1$, 上层为网络 A、下层为网络 B、子图为它们对应的动力学 u_i (a) 与 (e) 的耦合参数为 ($\lambda_{in} = 0.1, \lambda_{out} = 0.3$); (b) 与 (f) 的耦合参数为 ($\lambda_{in} = 0.1, \lambda_{out} = 1.8$); (c) 与 (g) 的耦合参数为 ($\lambda_{in} = 0.4, \lambda_{out} = 3.5$); (d) 与 (h) 的耦合参数为 ($\lambda_{in} = 4.0, \lambda_{out} = 3.5$). 改编自文献 [13]

Fig. 7. Four typical behaviors in the cerebral cortex with $\alpha = \pi/2 - 0.1$ where the up and down panels represent the two hemispheres, respectively, and the insets are their corresponding dynamics of u_i at a moment t . The parameters are $\lambda_{in} = 0.1$ and $\lambda_{out} = 0.3$ in panels (a) and (e) of disorder; $\lambda_{in} = 0.1$ and $\lambda_{out} = 1.8$ in panels (b) and (f) of chimera state; $\lambda_{in} = 0.4$ and $\lambda_{out} = 3.5$ in panels (c) and (g) of an emergent state conceptually similar to the state of unihemispheric sleep; and $\lambda_{in} = 4.0$ and $\lambda_{out} = 3.5$ in panels (d) and (h) of synchronization. Figure adapted from Ref. [13].

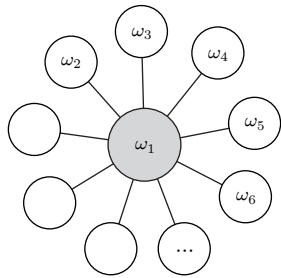


图 8 用于遥同步的星形网示意图. 改编自文献 [57]

Fig. 8. Graphic visualization of a hub network motif (star motif). Figure adapted from Ref. [57].

另一个则用来测量叶子节点间的关联

$$r^{\text{indirect}} = \frac{2}{(N-1)(N-2)} \sum_{n=2, m>n}^N r_{nm},$$

图 9 给出了这两个序参量对耦合强度 κ 的依赖性, 可见叶子节点间关联的增加要远快于中心节点与叶子节点间的关联. 叶子节点在 κ 大约为 0.47 时达到完全同步, 而中心节点与叶子节点间的完全同步要推迟到 $\kappa \approx 0.74$.

为了明白遥同步的机制, Bergner 等 [57] 将模型进一步简化为只有三个节点的情形: 两个叶子节点 1 与 2, 由中心节点 3 间接连接. 为了实现节点 1 与 2 的相互同步, 节点 1 的活动需要传递给节

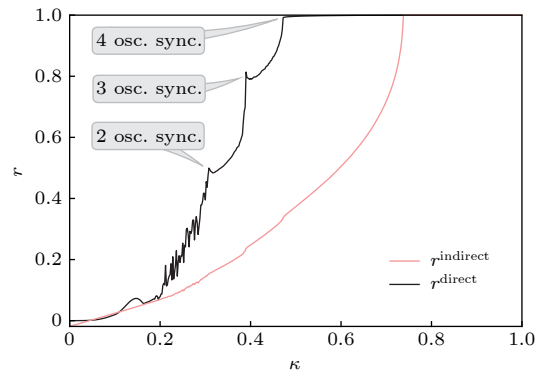


图 9 相同步的过程. 相同步的建立清晰可见, 其中的三个标注分别表明两个、三个与四个叶子节点间的同步化. 改编自文献 [57]

Fig. 9. Transition to PS for the hub motif. From the plot the onset of RS is clearly visible. The three annotations indicate synchronization between two, three, and four peripheral oscillators, respectively. Figure adapted from Ref. [57].

点 2, 反之亦然. 这就意味着节点 3 的动力学必须保证节点 1 与 2 的活动的传递尽可能不被改变, 为此需要满足两个条件: 首先, 为了不让中心节点与叶子节点同步, 节点 3 的吸引子的平均时间尺度应该与节点 1 与 2 的吸引子有很大的不同. 此外, 节点 1 与 2 的区别不应太大, 以便它们能够通过弱相互作用同步. 其次, 为了得到经过节点 3 的传递,

节点 3 的扰动必须不能衰减得太快. 将中心节点在其极限环 ($|u_0|^2 = \alpha$) 附近作线性化并保持其他两个叶子节点非接触, 可得如下方程:

$$\begin{aligned} \frac{du_h}{dt} &= (-2\alpha + i\omega_h)u_h + \frac{1}{2}\kappa(u_1 + u_2 - 2u_h), \\ \frac{du_{1,2}}{dt} &= (\alpha + i\omega_{1,2} - |u_{1,2}|^2)u_{1,2} + \kappa(u_h - u_{1,2}). \end{aligned} \quad (15)$$

这是两个 Stuart-Landau 振子通过一个线性过滤器耦合的方程. 通过拉普拉斯变换可将这个过滤器的传递函数写为

$$H_h(\omega) = \frac{\kappa}{i\omega + 2\alpha + \kappa - i\omega_h}, \quad (16)$$

于是, 中心节点可被一个有效耦合取代并有

$$\frac{du_{1,2}}{dt} = (\alpha + i\omega_{1,2} - \kappa - |u_{1,2}|^2)u_{1,2} + \kappa H_h(\omega_{2,1})u_{2,1}. \quad (17)$$

从这个角度, 我们可以解释两个主要的特点: 首先, 由于 $\lim_{\alpha \rightarrow \infty} H_h = 0$, 叶子节点间的有效耦合强度下降为零, 从而有效解耦而不能同步. 其次, 通过检验方程 (16), 可看清楚耦合强度对中心节点的频率的依赖性. 中心节点越快, 则有效耦合强度越低.

叶子节点实际上是对称的, 因此 Nicosia 等 [60] 研究了网络对称性对出现遥同步的影响. 他们考虑的是带相参数的全同的 Kuramoto 相振子, 其相参数使得相连的振子保持特定的相差, 从而阻碍完全同步. 发现耦合网络的对称性起着重要作用, 具有相同对称性的两个节点可以同步化, 即使它们相距遥远. 这种遥同步是由网络的对称性导致的. Zhang 等 [61] 则进一步研究了这种对称性与遥同步的关系, 将星形网的遥同步推广到了网络的情形, 称为非相干介导远程同步. 具体地, 将 N 个节点组织成三组 A, B 与 C , 其中 A 与 B 相连, B 与 C 相连, 但 A 与 C 不相连, 见图 10 所示. 假定 B 组至少有两个节点, 且每个组内的节点与连边形成一个相连的子网. 他们发现那两个非连续的 A 与 C 同步而中间 B 组与它们无关, 且镜像对称性是这种遥同步发生的机制, 比如 A 组中的节点 1 与 C 组中的节点 N 就具有镜像对称性而会完全同步.

基于大脑中的遥同步现象, Kang 等 [62] 讨论了第 2 节中那个真实脑皮层网络上遥同步的可能性, 发现对全同振子很难观察到遥同步. 但若考虑大脑中信号速度的有限而加进延迟时, 遥同步则成为可能. 具体地, 考虑真实脑皮层网络上每个节点为

Stuart-Landau 振子, 满足如下方程:

$$\dot{u}_j = (\alpha + i\omega - |u_j|^2)u_j + \varepsilon \sum_{k=1}^N \mathbf{W}_{jk}(u_k(t-\tau) - u_j(t)), \quad (18)$$

其中 $i = \sqrt{-1}$, $j = 1, 2, \dots, N = 989$, \mathbf{W}_{jk} 为权重链接矩阵 [48,49]; $\sqrt{\alpha}$ 与 ω 分别为未耦合时振子的振幅与自然频率; ε 为耦合强度; τ 为时间延迟. 在数值模拟中, 将方程 (18) 改写为

$$\begin{aligned} \dot{x}_i &= (\alpha - x_i^2 - y_i^2)x_i - \omega y_i \\ &+ \varepsilon \sum_{j=1}^N \mathbf{W}_{ij}(x_j(t-\tau) - x_i(t)), \\ \dot{y}_i &= \omega x_i + (\alpha - x_i^2 - y_i^2)y_i \\ &+ \varepsilon \sum_{j=1}^N \mathbf{W}_{ij}(y_j(t-\tau) - y_i(t)), \end{aligned} \quad (19)$$

其中 x_i 与 y_i 分别代表 u_i 的实部与虚部. 这里固定 $\alpha = 1.0$ 与 $\omega = 2.0$. 图 11 给出了 $\tau = 0.5$ 与 $\varepsilon = 0.1$ 时的六个典型的遥同步斑图, 相似的斑图也可在其他的参数 τ 与 ε 处发现. 图 11 的第二列与第三列清楚地表明每一个斑图都有两个或更多的中心节点, 与第一列的斑图只有一个中心节点形成了鲜明的对比.

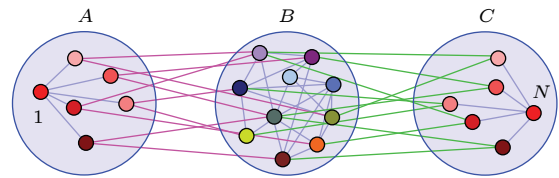


图 10 A 与 C 组之间的遥同步由 B 组的介导来完成. 节点的颜色代表了它们的状态, 可见节点 1 与 N 是完全同步的, 而 B 组中节点的动力学是无关联的. 改编自文献 [61]

Fig. 10. Remote synchronization between node groups A and C mediated by incoherence in group B . The colors of the nodes schematically represent their states, indicating that nodes 1 and N are identically synchronized, while the dynamics of the nodes in B are incoherent. Figure adapted from Ref. [61].

图 11 的发现也不同于文献 [57] 中的遥同步斑图, 因此有必要进一步挖掘. 图 11 的第二列与第三列的一个公共特征是它们的两个中心节点是由一些公共叶子节点相连的. 由此, Kang 等 [62] 提出了一个新的遥同步框架, 见图 12, 其中红、蓝、粉红数字分别代表中心节点、叶子节点与公共叶子节点. 这个模型揭示, 当公共叶子节点的耦合强度较小时, 只能形成单个中心节点的遥同步. 但当公共

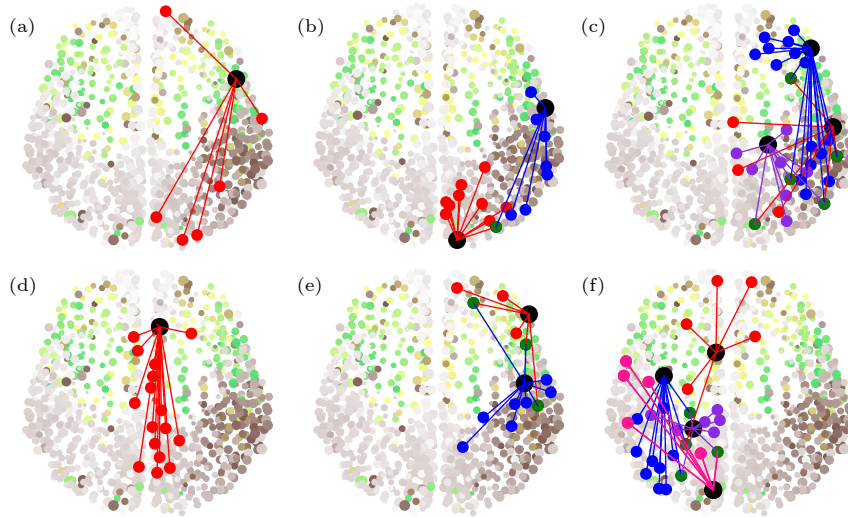


图 11 参数取 $\tau = 0.5$ 与 $\varepsilon = 0.1$ 时的六个典型的遥同步斑图. 每个斑图都是按如下条件挑选的: (i) 在中心节点与叶子节点间没有同步化; (ii) 所有的叶子节点均相互同步化. 改编自文献 [62]

Fig. 11. Six typical patterns of RS for $\tau = 0.5$ and $\varepsilon = 0.1$. Each pattern is chosen by the conditions: (i) There is no synchronization between the hub and its peripheral nodes; (ii) all the peripheral nodes are synchronized each other. Figure adapted from Ref. [62].

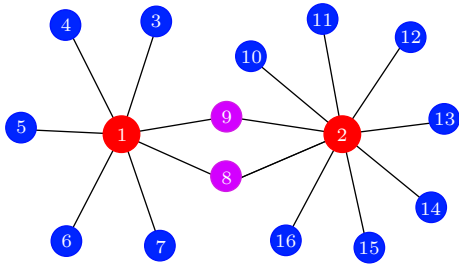


图 12 具有两个中心节点的遥同步新框架示意图, 其中红、蓝、粉红数字分别代表中心节点、叶子节点与公共叶子节点. 改编自文献 [62]

Fig. 12. A schematic figure of the new framework of RS with two hubs, where the nodes with red, blue and pink numbers represent the hub, leaf and common leaf nodes, respectively. Figure adapted from Ref. [62].

叶子节点的耦合强度较大时, 两个遥同步集团可以合并成一个较大的遥同步集团, 此时那两个中心节点同步成为一个同步集团, 而所有的叶子节点成为另一个同步集团. 详情见原文, 这里不展开讨论了.

此外, 还有其他的一些关于遥同步的研究, 比如 Gambuzza 等 [63] 研究了多层网上的遥同步, Punetha 等 [64] 研究了二分网上带时间延迟的遥同步, Leyva 等 [65] 研究了多层对称网络上的遥同步等.

4 集团同步

集团同步表示网络中特定集团内的振子将同步到相同的轨道而不同集团之间的行为则互不相

同. 这种同步集团可以出现在成群的动物中, 其网络为邻居间的视觉连接, 或者由局部通讯网连接的成群的无人驾驶汽车. 集团同步也可以出现在电力网中, 这将是出问题的征兆, 即失去全局同步. 最近的研究表明, 网络拓扑的对称性与同步动力学集团的形成有重要关系 [66–70], 并已为较多的实验证实 [71–74], 王新刚 [70] 对此有一个较全面的综述.

为了说明耦合振子网络中的对称性与集团的概念, 图 13(a) 展示了一个四节点网络, 其中振子是全同的、耦合是双向的 [73]. 这个网络总共有六个对称. 图 13(b) 给出了图 13(a) 中交换节点 1 与 2 的一个反射, 网络结构保持不变. 图 13 (c) 展示了旋转 120° 的网络, 其与原始结构 (图 13(a)) 不可区分. 这些对称性可由邻接矩阵 \mathbf{A} 的对称性来表示, 其元素为 $A_{ij} = 1$ 如果节点 i 与节点 j 相连, 否则 $A_{ij} = 0$. 网络的动力学可写为

$$\dot{\mathbf{x}}_i = \mathbf{F}(\mathbf{x}_i) + \sigma \sum_j A_{ij} \mathbf{H}(\mathbf{x}_j), \quad (20)$$

其中 $i = 1, \dots, N$, \mathbf{x}_i 为振子 i 的动力学矢量, \mathbf{F} 为振子的矢量场, \mathbf{H} 为耦合函数. 这个动力学方程也可写成拉普拉斯耦合的形式

$$\dot{\mathbf{x}}_i = \mathbf{F}(\mathbf{x}_i) + \sigma \sum_j L_{ij} \mathbf{H}(\mathbf{x}_j), \quad (21)$$

其中拉普拉斯矩阵 $\mathbf{L} = \{L_{ij}\}$, $L_{ij} = A_{ij} - \delta_{ij} \sum_j A_{ij}$. 此时, 矩阵 \mathbf{L} 每行之和等于零, 即节点 i 的输入被对角自耦合平衡了.

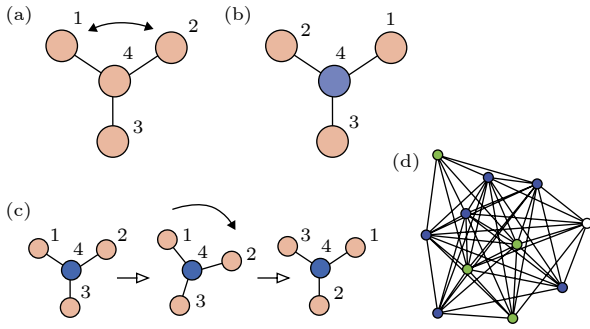


图 13 网络中对称性的例子 (a) 四个全同振子通过三根连线耦合的网络; (b) 一个反射操作后的同一网络; (c) 一个旋转操作后的同一网络; (d) 一个 11 个节点的网络有三个集团 (蓝、绿、白). 改编自文献 [73]

Fig. 13. Examples of symmetries in networks: (a) A network of four identical oscillators coupled through three identical links; (b) the same network after a reflection operation; (c) the same network after a rotation operation; (d) an 11-node network showing three clusters (blue, green, and white). Figure adapted from Ref. [73].

在图 13(a) 的网络中, 节点 1, 2 与 3 有相同的运动方程, 因此如果它们从相同的初始条件出发, 就将被永久地保持同步化. 节点 4 不能与任何其他

节点置换且也不与它们同步, 因此在图 13(b) 与图 13(c) 中将其标为不同的颜色. 这是网络中对称性与动力学间的亲密关系, 它将网络分成了两个集团 1, 2, 3 与 4. 文献 [73] 讨论了对于给定的网络, 如何发现所有可能的集团同步斑图. 其主要结论是邻接矩阵的对称性对应动力学的对称性, 当它们从一个同步态出发, 可置换的节点集构成的集团就将保持同步, 即每个集团的同步态是动力学流不变的. 图 14 给出了所有允许的斑图, 其中属于相同集团的节点标成了同一种颜色.

文献 [68] 指出, 基于集团的坐标变换可以更简单且更快地计算网络的对称性. 图 15 是关于 24 个节点的网络上的 9 个对称性集团, 其中每个集团中的节点具有相同的稳定性, 而不同的集团则稳定性不同, 且这些同步集团可唯一确定. 此外, Dahms^[66] 等考虑了时间延迟对全同振子网络上集体行为的影响, 发现延迟耦合确实可以诱导出不同的同步集团, 且允许集团同步的耦合矩阵在本征值谱上也显示了非常类似的对称性.

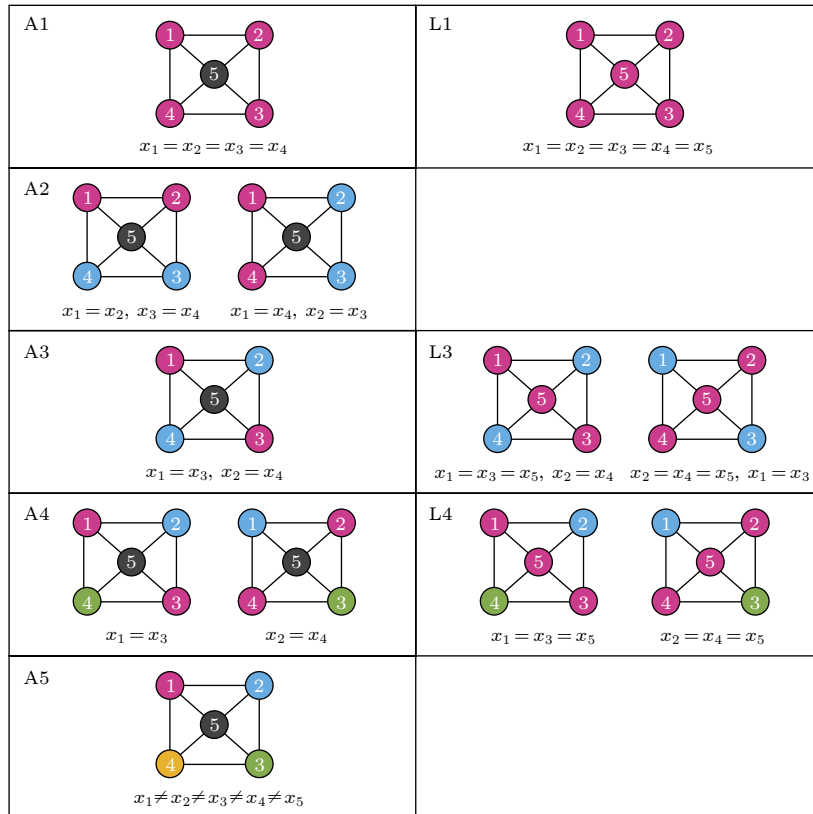


图 14 五个节点的网络中的集团斑图 左边: 当网络连接为邻接矩阵 (方程 (20)) 时所有可能的斑图; 右边: 当网络连接为拉普拉斯矩阵 (方程 (21)) 时额外的斑图. 改编自文献 [73]

Fig. 14. Patterns of clusters in a five-node network. Left: All possible patterns displayed when the network connectivity is given by the adjacency matrix (Eq. (20)); right: Additional patterns displayed when the network connectivity is given by the Laplacian matrix (Eq. (21)). Figure adapted from Ref. [73].

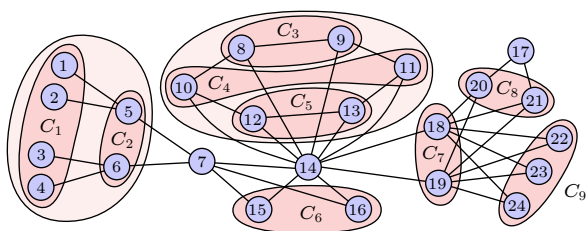


图 15 24 个节点构成的网络中对称集团的分组. 改编自文献 [68]

Fig. 15. Grouping of symmetry clusters in a CS pattern for a 24-node network. Figure adapted from Ref. [68].

5 讨论与展望

至此, 我们完成了对奇异态、遥同步与集团同步这三种部分同步态研究进展的介绍. 从这些研究可见, 它们三者之间既有联系, 也有区别. 首先, 它们三者都是介于完全同步与完全无序之间的中间态, 承担着从无序到同步的过渡任务. 但这种中间态并不是通常那种大家逐渐靠近、序参量逐渐增大的二级相变过程, 而是在位形空间中形成两种性质截然不同的状态且相互共存, 这就比较神奇, 从而成为了它的反直观之处, 同时也成就了它用于解释大脑斑图的特质. 这个特质在奇异态中表现为相互连接的同步组与非同步组的共存, 也可以是分开的多个同步组与非同步组的共存; 在遥同步中表现为不直接相连的分散振子成为了同步态, 而与它们相连的振子, 尤其是中心节点, 却不与它们同步; 在集团同步中表现为具有网络对称性的节点达成了同步, 而不同对称性的集团则成为不同的同步集团, 同步的振子之间可以存在连接, 也可以没有直接连接. 这些不同的表观形式其实可以看成是同一种现象的不同表述, 就像量子力学中不同的表象一样, 供我们在不同的环境下有选择地使用.

除了在表观形式上不相同之外, 它们三者出现的条件也不相同. 奇异态最初要求的条件是非局域耦合, 由衰减核描述. 然后这个条件推广为初始条件的不同, 但非局域耦合的要求适当放宽, 不再要求各个振子的耦合对象均有些不同, 而改为各组内个体耦合的对象完全相同, 但不同组间个体耦合的对象不相同. 再然后进一步推广为延迟系统, 此时所有个体耦合的对象可以完全相同. 遥同步主要讨论的是朗道振子系统, 本质原因是朗道振子可作为线性滤波器来传递不直接相连的节点间的耦合信息, 从而使中心节点周围的叶子节点达到同步. 遥

同步最初是在星形网上进行讨论的, 现已拓展至复杂网络乃至真实的脑网络. 集团同步最初要求的是网络位置完全对称, 这就导致对称性相同的节点具有完全一样的动力学方程, 从而可达到完全同步的动力学行为. 后来这个条件推广为输入量相同即可, 因为所谓对称性保证的实际上是耦合贡献的等同.

事实上, 这三种部分同步并不是完全独立的, 而是都可看成某种局部对称性的体现. 例如奇异态中同步组内的振子实际上具有相同的对称性, 遥同步中围绕中心节点的叶子也具有完全对称的地位, 集团同步就更不用说了, 它的定义就源自对称性. 按照这个局部对称性的特点, 我们其实可以抛开直接相连或不直接相连的具体细节, 而将这三者统称为局部同步态的涌现. 在这个意义上, 就不难理解在某些系统中可以同时出现多种部分同步态. 比如图 11 中的遥同步振子与其他的非同步振子一起就构成了奇异态, 图 12 中的遥同步振子因结构对称也可当作集团同步. 同理, 对于图 15 中的那些同步集团中的振子, 如果它们不直接相连的话, 就是遥同步了, 比如 C_2 中的振子 5 与 6. 这种同时具有多个属性的部分同步在大脑皮层网络中非常普遍, 或者说大脑皮层网就是一个典型的呈现各种各样部分同步的系统.

这三种部分同步态的研究意义就在于它们可以被很好地用来解释大脑高级功能的物理机制, 包括感知、学习与记忆等. 现在已经知道大脑神经网络具有典型的复杂网络特征, 包括度分布的异质性、小世界特征及社团结构等. 正是由于这些特征, 才保证了大脑认知活动期间广泛的斑图, 从而保障了大脑高级功能的实施. 当大脑接收到不同的信号或刺激时, 相应的动力学斑图就会被启动, 或者实现从一个斑图到另一个的转换. 具体表现就是我们观察到的对应奇异态的首晚效应、对应遥同步的分布式计算及对应集团同步的斑图多样性等.

虽然这三种部分同步的研究取得了不错的结果, 在大脑机制的阐述方面也取得了重要的进展, 但对大脑及大脑运行机制的研究依然是个长期而艰巨的任务, 还有许多开放的问题等着我们进一步研究. 下面是根据我们自己的研究经验建议的几个开放问题.

1) 如何将这三种部分同步得到的现有结论充分地用于探索大脑功能的机制, 比如认知、记忆与信号传播等方面.

2) 现有研究大多局限于单层网络. 由于大脑的社区结构特征, 多层复杂网络上的部分同步更值得期待. 文献 [13] 是这样的一个例子, 但更进一步、更深入的探讨还很缺乏.

3) 如何能跳开结构层面的分析, 从更深的层面来揭示出一些更本质的东西, 比如从本征值谱分析的角度来进行分析等. 这方面的一个先期探索是关于本征模与模块化结构间的内在联系 [75].

4) 如何根据这三种部分同步的特征来控制各种神经疾病如帕金森氏病、癫痫、阿尔茨海默病、精神分裂症与脑肿瘤等.

5) 除了这三种典型的部分同步外, 是否还存在其他的部分同步态或者根本就不能归结为部分同步但却是脑功能所必须的状态? 如果有的话, 如何去揭示它们?

我们期待部分同步化的研究在大脑神经元网络的机制与应用方面发挥越来越重要的作用. 希望本综述起到抛砖引玉的作用, 激励出更多、更好的工作.

参考文献

- [1] Pecora L M, Carroll T L 1990 *Phys. Rev. Lett.* **64** 821
- [2] Liu Z H 2018 *Chaotic Dynamics Foundation and Its Applications in Brain Functions* (Beijing: Science Press) (in Chinese) [刘宗华 2018 混沌动力学基础及其在大脑功能方面的应用 (北京: 科学出版社)]
- [3] Gray C M, Konig P, Engel A K, Singer W 1989 *Nature* **338** 334
- [4] Hasler M, Maistrenko Y, Popovych O 1998 *Phys. Rev. E* **58** 6843
- [5] Hu G, Zhang Y, Cerdeira H A, Chen S 2000 *Phys. Rev. Lett.* **85** 3377
- [6] Zhang Y, Hu G, Cerdeira H A, Chen S, Braun T, Yao Y 2001 *Phys. Rev. E* **63** 026211
- [7] Fu C, Lin W, Huang L, Wang X 2014 *Phys. Rev. E* **89** 052908
- [8] Heisler I A, Braun T, Zhang Y, Hu G, Cerdeira H A 2003 *Chaos* **13** 185
- [9] Kumar P, Verma D K, Parmananda P 2017 *Phys. Lett. A* **381** 2337
- [10] Panaggio M J, Abrams D M 2015 *Nonlinearity* **28** R67
- [11] Majhi., Bera B K, Ghosh D, Perc M 2019 *Phys. Life Rev.* **28** 100
- [12] Majhi S, Perc M, Ghosh D 2016 *Sci. Rep.* **6** 39033
- [13] Kang L, Tian C, Huo S, Liu Z 2019 *Sci. Rep.* **9** 14389
- [14] Roelfsema P R, Engel A K 1997 *Nature* **385** 157
- [15] Vogels T P, Abbott L F 2005 *J. Neurosci.* **25** 10786
- [16] Diesmann M, Gewaltig M, Aertsen A 1999 *Nature* **402** 529
- [17] Benedek M, Bergner S, Konen T, Fink A, Neubauer A C 2011 *Neuropsychologia* **49** 3505
- [18] Ayala G F, Dichter M, Gummit R J, Matsumoto H, Spencer W A 1973 *Brain Res.* **52** 1
- [19] Levy R, Hutchison W D, Lozano A M, Dostrovsky J O 2000 *J. Neurosci.* **20** 7766
- [20] Abrams D M, Strogatz S H 2004 *Phys. Rev. Lett.* **93** 174102
- [21] Rattenborg N C, Amlaner C J, Lima S L 2000 *Neurosci. Biobehav. Rev.* **24** 817
- [22] Mathews C G, Lesku J A, Lima S L, Amlaner C J 2006 *Ethology* **112** 286
- [23] Tamaki M, Bang J W, Watanabe T, Sasaki Y 2016 *Curr. Biol.* **26** 1190
- [24] Kuramoto Y, Battogtokh D 2002 *Nonlinear Phenom. Complex Syst.* **5** 380
- [25] Abrams D M, Mirollo R, Strogatz S H, Wiley D A 2008 *Phys. Rev. Lett.* **101** 084103
- [26] Ott E, Antonsen T M 2008 *Chaos* **18** 037113
- [27] Ma R, Wang J, Liu Z 2010 *Europhys. Lett.* **91** 40006
- [28] Indic P, Schwartz W J, Paydarfar D 2008 *J. R. Soc. Interface* **5** 873
- [29] Omelchenko O E, Wolfrum M, Yanchuk S, Maistrenko Y L, Sudakov O 2012 *Phys. Rev. E* **85** 036210
- [30] Maistrenko Y, Sudakov O, Osiv O, Maistrenko V 2015 *New J. Phys.* **17** 073037
- [31] Yeldesbay A, Pikovsky A, Rosenblum M 2014 *Phys. Rev. Lett.* **112** 144103
- [32] Sethia G C, Sen A, Atay F M 2008 *Phys. Rev. Lett.* **100** 144102
- [33] Tian C, Cao L, Bi H, Xu K, Liu Z 2018 *Nonlinear Dyn.* **93** 1695
- [34] Zhu Y, Zheng Z, Yang J 2014 *Phys. Rev. E* **89** 022914
- [35] Zheng Z G, Zhai Y 2020 *SCIENTIA SINICA Physica, Mechanica & Astronomica* **50** 010505 (in Chinese) [郑志刚, 翟云 2020 中国科学: 物理学 力学 天文学 **50** 010505]
- [36] Kemeeth F P, Haugland S W, Schmidt L, Kevrekidis I G, Krischer K 2016 *Chaos* **26** 094815
- [37] Omelchenko I, Omelchenko E, Hovel P, Scholl E 2013 *Phys. Rev. Lett.* **110** 224101
- [38] Hizanidis J, Kanas V G, Bezerianos A, Bountis T 2014 *Int. J. Bifurcation Chaos* **24** 1450030
- [39] Sakaguchi H 2006 *Phys. Rev. E* **73** 031907
- [40] Olmi S, Politi A, Torcini A 2010 *Europhys. Lett.* **92** 60007
- [41] Santos M S, et al. 2017 *Chaos, Solitons Fractals* **101** 86
- [42] Hizanidis J, Kouvaris N E, Zamora-Lopez G, Diaz-Guilera A, Antonopoulos C G 2016 *Sci. Rep.* **6** 19845
- [43] Tian C, Bi H, Zhang X, Guan S, Liu Z 2017 *Phys. Rev. E* **96** 052209
- [44] Tian C, Zhang X, Wang Z, Liu Z 2017 *Front. Phys.* **12** 128904
- [45] Chouzouris T, Omelchenko I, Zakharova A, et al. 2018 *Chaos* **28** 045112
- [46] Bansal K, Garcia J O, Tompson S H, Verstynen T, Vettel J M, Muldoon S F 2019 *Sci. Adv.* **5** eaau8535
- [47] Huo S, Tian C, Kang L, Liu Z 2019 *Nonlinear Dyn.* **96** 75
- [48] Hagmann P, et al. 2008 *PLoS Biol.* **6** 1479
- [49] Honey C J, et al. 2009 *Proc. Natl. Acad. Sci. U.S.A.* **106** 2035
- [50] Tinsley M R, Nkomo S, Showalter K 2012 *Nat. Phys.* **8** 662
- [51] Nkomo S, Tinsley M R, Showalter K 2013 *Phys. Rev. Lett.* **110** 244102
- [52] Martens E A, Thutupalli S, Fourriere A, et al. 2013 *Proc. Natl. Acad. Sci. U.S.A.* **110** 10563
- [53] Gambuzza L V, Buscarino A, Chossari S, et al. 2014 *Phys. Rev. E* **90** 032905
- [54] Totz J F, Rode J, Tinsley M R, Showalter K, Engel H 2018 *Nat. Phys.* **14** 282
- [55] Ito T, Kulkarni K R, Schultz D H, Mill R D, Chen R H, Solomyak L I, Cole M W 2017 *Nat. Commun.* **8** 1027

- [56] Sporns O 2013 *Curr. Opin. Neurobiol.* **23** 162
- [57] Bergner A, Frasca M, Sciuto G, Buscarino A, Ngamga E J, Fortuna L, Kurths J 2012 *Phys. Rev. E* **85** 026208
- [58] Vicente R, Gollo L L, Mirasso C R, Fischer I, Pipa G 2008 *Proc. Natl. Acad. Sci. U.S.A.* **105** 17157
- [59] Minati L 2015 *Chaos* **25** 123107
- [60] Nicosia V, Valencia M, Chavez M, Diaz-Guilera A, Latora V 2013 *Phys. Rev. Lett.* **110** 174102
- [61] Zhang L, Motter A E, Nishikawa T 2017 *Phys. Rev. Lett.* **118** 174102
- [62] Kang L, Wang Z, Huo S, Tian C, Liu Z 2020 *Nonlinear Dyn.* **99** 1577
- [63] Gambuzza L V, Frasca M, Gomez-Gardenes J 2015 *Europhys. Lett.* **110** 20010
- [64] Punetha N, Ujjwal S R, Atay F M, Ramaswamy R 2015 *Phys. Rev. E* **91** 022922
- [65] Leyva I, Sendina-Nadal I, Sevilla-Escoboza R, Vera-Avila V P, Chholak P, Boccaletti S 2018 *Sci. Rep.* **8** 8629
- [66] Dahms T, Lehnert J, Scholl E 2012 *Phys. Rev. E* **86** 016202
- [67] Jalan S, Singh A, Acharyya S, Kurths J 2015 *Phys. Rev. E* **91** 022901
- [68] Cho Y S, Nishikawa T, Motter A E 2017 *Phys. Rev. Lett.* **119** 084101
- [69] Siddique A B, Pecora L, Hart J D, Sorrentino F 2018 *Phys. Rev. E* **97** 042217
- [70] Wang X G 2020 *SCIENTIA SINICA Physica, Mechanica & Astronomica* **50** 010503 (in Chinese) [王新刚 2020 中国科学: 物理学 力学 天文学 **50** 010503]
- [71] Williams C R S, Murphy T E, Roy R, Sorrentino F, Dahms T, Scholl E 2013 *Phys. Rev. Lett.* **110** 064104
- [72] Pecora L M, Sorrentino F, Hagerstrom A M, Murphy T E, Roy R 2014 *Nat. Commun.* **5** 4079
- [73] Sorrentino F, Pecora L M, Hagerstrom M, Murphy T E, Roy R 2016 *Sci. Adv.* **2** e1501737
- [74] Cao B, Wang Y F, Wang L, Yu Y Z, Wang X G 2018 *Front. Phys.* **13** 130505
- [75] Wang R, Lin P, Liu M, Wu Y, Zhou T, Zhou C 2019 *Phys. Rev. Lett.* **123** 038301

SPECIAL TOPIC—Statistical physics and complex systems

Partial synchronization in complex networks: Chimera state, remote synchronization, and cluster synchronization^{*}

Wang Zhen-Hua Liu Zong-Hua[†]

(School of Physics and Electronic Science, East China Normal University, Shanghai 200241, China)

(Received 26 December 2019; revised manuscript received 14 February 2020)

Abstract

In recent years, the study of partial synchronization of coupled oscillators in complex networks has attracted great attention. The underlying reason is both the extensive existence of the patterns of partial synchronization in brain network and their close relationship to brain functions of cognition and memory. In this paper, we briefly review the research progress in this field. According to the researches by different groups, we classify them as three types, i.e. chimera state, remote synchronization, and clustering synchronization. We mainly discuss the conditions of these three states, as well as their models, detections, and their applications in biology. We discuss the relationship among the three types of states and give some outlooks for future studies.

Keywords: complex network, chimera state, remote synchronization, cluster synchronization

PACS: 89.75.Hc, 05.45.Xt, 68.18.Jk

DOI: 10.7498/aps.69.20191973

^{*} Project supported by the National Natural Science Foundation of China (Grant Nos. 11675056, 11835003).

[†] Corresponding author. E-mail: zhliu@phy.ecnu.edu.cn

社会引力定律追根溯源

闫小勇

Exploring the roots of social gravity law

Yan Xiao-Yong

引用信息 Citation: *Acta Physica Sinica*, 69, 088903 (2020) DOI: 10.7498/aps.69.20191686

在线阅读 View online: <https://doi.org/10.7498/aps.69.20191686>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

个体行为与社会环境耦合演化的舆论生成模型

Opinion formation model with co-evolution of individual behavior and social environment

物理学报. 2019, 68(11): 118902 <https://doi.org/10.7498/aps.68.20182254>

引力彩虹时空中Kerr黑洞的熵谱和面积谱

Entropy spectrum and area spectrum of Kerr black hole in gravity's rianbow

物理学报. 2018, 67(6): 060401 <https://doi.org/10.7498/aps.67.20172374>

近距离牛顿反平方定律实验检验进展

Recent progress in testing Newtonian inverse square law at short range

物理学报. 2018, 67(16): 160401 <https://doi.org/10.7498/aps.67.20180636>

基于电荷守恒定律的航天器内带电三维仿真简化模型

A three-dimensional simplified simulation model based on charge conservation law for internal charging in spacecraft

物理学报. 2019, 68(19): 195201 <https://doi.org/10.7498/aps.68.20190631>

一种基于社交影响力和平均场理论的信息传播动力学模型

An information diffusion dynamic model based on social influence and mean-field theory

物理学报. 2017, 66(3): 030501 <https://doi.org/10.7498/aps.66.030501>

专题：统计物理和复杂系统

社会引力定律追根溯源*

闫小勇^{1)2)†}

1) (北京交通大学, 交通系统科学与工程研究院, 北京 100044)

2) (电子科技大学, 复杂性实验室, 成都 611731)

(2019年11月4日收到; 2019年11月19日收到修改稿)

在交通出行、人口迁移、商品贸易、信息流通、社会交往、科研合作等大量人、物、信息的空间流动现象中, 都存在类似万有引力定律的规律, 即两地之间的某种流动量与两地活力的乘积成正比、与两地距离的幂成反比. 类比万有引力定律建立的引力模型也在交通出行分布预测、人口迁移量预测、地区间贸易量预测等诸多方面获得了广泛应用. 但复杂的社会系统中为何会有这样简单的引力定律存在? 这是个非常有趣也有价值的问题. 本文对从统计物理学、微观经济学和博弈论等不同视角探索社会引力定律根源的研究进行了综述.

关键词: 社会引力定律, 最大熵原理, 离散选择模型, 拥挤博弈

PACS: 89.75.-k, 89.65.-s, 89.70.Cf, 02.50.Le

DOI: 10.7498/aps.69.20191686

1 引言

预测地点间人、物、信息的流动是社会学、经济学、人口学、交通科学、网络科学等诸多学科长期以来的一个重要研究主题^[1,2]. 一百多年来, 研究者们陆续提出了多种预测地点间流动量的模型(在经济学中被称为空间交互模型^[3]), 其中最有影响力的是引力模型^[4] (gravity model, 也译为“重力模型”). 引力模型在许多方面都获得了广泛应用. 例如在人口学中, 引力模型被用来预测地区间的人口迁移量^[5]; 在经济学中, 引力模型被用来预测国家间的商品贸易量^[6]; 在交通科学中, 引力模型被用来预测地点间的交通出行量^[7]; 在网络科学中, 引力模型被用来评估网络节点的相似性^[8]、影响力^[9,10]等. 引力模型之所以受到如此青睐, 是由于在许多空间流动现象中都存在类似万有引力定律^[11]的规律, 即两地之间的某种流动量正比于两地“活力”(多用地点人口数量、经济发展水平、进出流动量等

表示)的乘积, 反比于两地之间空间距离的幂函数. 这种规律被称为社会引力定律^[12].

早在1846年, Desart^[13]就在比利时铁路客运数据中发现, 两个车站间的客运量正比于两个车站所在地的人口数乘积, 反比于两车站间距离的2.25次幂. 这可能是关于发现社会引力定律的最早记载^[14], 其发现时间比Carey^[15]在人口迁移现象中发现社会引力定律的时间(1858年)还要早12年. 后来, Ravenstein^[16]从19世纪70—80年代英国人口普查数据中也发现两地人口迁移量与两地人口乘积成正比、与两地距离成反比的现象. 而Reilly^[17]在1929年的著作中研究零售市场问题时也发现, 零售中心从其周围某个城镇吸引到的顾客数量与该零售中心的规模成正比, 与两地间的距离平方成反比. 此后的几十年中, 研究者在交通^[18]、人口^[19]、经济^[20]等领域又陆续发现了许多服从社会引力定律的现象.

近年来, 随着现代电子与信息技术的不断发展, 有越来越多的手段(如IC卡、GPS、手机、社

* 国家自然科学基金(批准号: 71822102, 71621001, 71671015, 61304177)资助的课题.

† 通信作者. E-mail: yanxy@bjtu.edu.cn

交网站等)可长期记录人、物、信息在空间中的流动数据^[21,22].通过对这些大数据进行统计分析,统计物理与复杂系统领域的学者在很多复杂系统中都发现了符合社会引力定律的现象.例如, Viboud 等^[23]通过分析美国人口普查数据,发现两郡之间的通勤出行量与郡人口、郡间距离的关系符合社会引力定律; Jung 等^[24]通过分析韩国高速公路收费数据,发现两城市间的高速公路交通量与城市人口、城市间距离的关系符合社会引力定律; Krings 等^[25]分析了比利时 250 万名匿名手机用户在 571 个城镇之间的手机通讯数据,发现城镇间通讯量与城镇人口、城镇间距离的关系符合社会引力定律; Balcan 等^[26]通过分析国际航空运输协会的航空客运数据,发现全球 29 个国家主要机场之间的航空运输量与机场服务区域人口、机场间距离的关系符合社会引力定律; Kaluza 等^[27]分析了全球船舶港口自动识别系统数据,发现全球 951 个港口间的海运量与港口进出量、港口间距离的关系符合社会引力定律; Pan 等^[28]通过分析美国科学信息研究所 2003—2010 年的论文数据,发现全球 18199 个城市之间的科研合作强度与城市整体科研水平、城市间距离的关系符合社会引力定律; Goh 等^[29,30]分析了韩国首尔市地铁网络和常规公交网络中的刷卡数据,发现车站间客流与车站上下车客流、车站间距离的关系符合社会引力定律; Levy 和 Goldenberg^[31]分析了 Facebook 等四类社交网络数据,发现两人交往概率与两人距离之间的关系符合社会引力定律.诸如此类的研究还有很多.可以预见,社会引力定律还会在更多的复杂系统中被发现.

从以上研究回顾中可以看出,交通出行、人口迁移、商品贸易、信息流通、社会交往、科研合作等大量空间交互现象都符合社会引力定律.但在诸多复杂的社会系统中为何会有如此简单的社会引力定律存在?这是一个困扰人类上百年的问题,也吸引了很多不同领域的学者对其追根溯源^[32,33].从 20 世纪 60 年代起,统计物理领域的学者用最大熵原理^[34]、经济学领域的学者用效用理论^[35,36]相继对社会引力定律进行了解释.最近,复杂系统领域的研究人员又从博弈论角度探索了社会引力定律的根源^[37].本文将对这些从不同视角探索社会引力定律根源的研究进行综述.

2 社会引力定律的统计物理解释

2.1 最大熵原理导出引力模型

Wilson^[34]以交通系统中地点间的出行分布问题为背景,用统计物理学中的最大熵原理^[38]对社会引力定律给出了最早的理论解释.在交通系统的出行需求预测工作中,常采用的模型框架是“四阶段法”,即依次进行出行生成预测、出行分布预测、交通方式划分和交通分配^[7].其中,在出行生成预测阶段可得到地点 i 的出行发生量 O_i 和出行吸引量 D_i .而在出行分布预测阶段,则要求预测出的地点间出行分布量 T_{ij} 既满足发生量约束 $\sum_j T_{ij} = O_i$, 又满足吸引量约束 $\sum_i T_{ij} = D_j$. Wilson 用最大熵原理导出了满足以上两类约束的引力模型:在交通系统总出行量为 T 的情况下,系统的宏观状态是任意两个地点 i, j 之间的出行量为 T_{ij} , 则系统的微观状态数为

$$\Omega = \frac{T!}{\prod_i \prod_j T_{ij}!}. \quad (1)$$

Wilson 为地点间出行量 T_{ij} 又增加了一个出行成本约束

$$\sum_i \sum_j T_{ij} c_{ij} = C, \quad (2)$$

其中 c_{ij} 为从地点 i 到 j 的出行成本, C 为系统可支配的总成本.那么,系统最可能出现的出行分布就是以下最大熵模型的解:

$$\begin{aligned} \max \ln \Omega &= \ln \frac{T!}{\prod_i \prod_j T_{ij}!}, \\ \text{s.t.} \quad &\sum_j T_{ij} = O_i, \\ &\sum_i T_{ij} = D_i, \\ &\sum_i \sum_j T_{ij} c_{ij} = C. \end{aligned} \quad (3)$$

用拉格朗日乘法可求得该模型的解为

$$T_{ij} = a_i O_i b_j D_j e^{-\gamma c_{ij}}, \quad (4)$$

其中 γ 是模型参数, $a_i = 1 / \sum_j b_j D_j e^{-\gamma c_{ij}}$ 和 $b_j = 1 / \sum_i a_i O_i e^{-\gamma c_{ij}}$ 是两组相互依赖的迭代因子.这

就是带有负指数成本函数的双约束引力模型 [7].

2.2 成本与距离的对数关系

经典引力模型中的距离函数是幂律函数 [4]. 但 (4) 式中用最大熵原理导出的引力模型, 其成本函数是负指数函数 $e^{-\gamma c_{ij}}$, 这与经典引力模型中使用的距离幂函数不同. 简单地看, 直接将 $\gamma c_{ij} = \beta \ln d_{ij}$ 这种成本与距离之间的对数关系代入 (4) 式, 即可得到经典的双约束引力模型 $T_{ij} = a_i O_i b_j D_j d_{ij}^{-\beta}$. 但为什么成本与距离之间具有这种对数关系, 之前并未得到很好的解答 [39].

Yan 等 [40] 为社会引力定律中距离的幂函数形式提供了一种解释. 他们认为, 交通系统中的出行成本 c 主要由出行时间 t 和货币费用 m 两部分组成, 可以表示为二者的加权和形式 $c \approx \eta t + \mu m$ [41]. 其中, 货币费用通常与出行距离 d 具有近似线性关系 $m \approx \nu d$ [41]. 他们通过分析出行日志数据 [42] 中出行时间与距离的关系, 发现出行时间与距离也具有近似对数关系 $t \approx \phi \ln d + \psi$ [43]. 这种对数关系源于人们在不同出行距离所采用交通方式的速度差异 [44], 例如人们在进行几百米的出行时往往是步行或骑自行车, 而几十千米时就要使用公交、小汽车等交通方式, 当几百千米时就要乘坐更快速的火车或飞机了, 这使得出行时间和距离之间呈现非线性关系. 综上, 可得到出行成本与距离之间的关系为 $c = \eta \phi \ln d + \eta \psi + \mu \nu d$. 进而结合最大熵原理, 还可导出群体出行距离的截尾幂律分布 $P(d) \sim d^{-\beta} e^{-d/\kappa}$. 这一研究不仅能解释社会引力定律中距离函数的幂律形式, 还解释了一大批出行距离分布的实证统计结果 [45–50].

3 社会引力定律的微观经济学解释

Wilson 用最大熵原理为社会引力定律提供了一个非常合理的宏观解释. 但是, 最大熵原理仅能给出系统最可能的宏观分布状态, 并不考虑系统中个体选择出行目的地的微观决策过程. 从出行目的地选择决策行为的角度来看, 社会引力定律的微观底层机制仍未得到满意解答 [32,33]. 而经济学家则很早就开始用效用理论来研究个体选择出行目的地的微观决策行为 [35,36]. 早期的研究使用确定效用理论来解释社会引力定律 [35], 而影响更为广泛的研究则是基于随机效用理论的离散选择模型 [36]. 离

散选择模型是微观经济学中非常重要的描述个体选择行为的方法, 在交通系统出行选择问题上已获得非常广泛的应用 [51,52]. 本节将介绍用离散选择模型对个体出行目的地选择行为进行建模以及导出引力模型的方法.

3.1 目的地离散选择模型

离散选择模型被用来描述、解释和预测决策者对多个离散的选项进行选择的行为. 一次离散选择行为通常包含以下要素 [51].

1) 决策者, 即做出选择行为的主体. 在出行目的地选择行为中, 决策者就是指位于某起点选择目的地的一个出行者.

2) 选项, 即供决策者选择的多项事物. 在出行目的地选择行为中, 选项是指供出行者选择的目的地.

3) 选项属性. 决策者选择选项时会考虑诸如价格、质量等因素, 每一种因素称为一个属性. 在出行目的地选择行为中, 选项属性一般包括各地点的活力、到各地点的距离或出行成本等.

4) 决策准则, 即决策者在做出选择时的行为准则. 最常用的是效用最大化准则 [53], 即决策者在所有选项中选择效用最高的选项. 此处的效用是指决策者选择某个选项所能获得的满意程度. 在出行目的地选择问题中, 可用下式来表达这一准则:

$$U_{ij} > U_{ik}, \forall k \in K - \{j\}, \quad (5)$$

其中 U_{ij} 是位于地点 i 的出行者所选择的目的地 j 的效用, U_{ik} 则是其他任一目的地的效用, K 是所有目的地的集合.

(5) 式的效用值 U_{ik} 本质上是指目的地 k 所具有的真实效用. 根据效用最大化准则, 决策者会选择所有选项中真实效用最大的一个选项. 但实际选择问题中, 决策者有可能并不是总选择一个固定的选项, 而是以不同的概率选择不同的选项. 这可能是决策者受到一些内部或外部因素的影响, 对选项效用的认知发生了变化. 但研究者并不能直接观测真实效用变化, 只能观测选项的确定效用值 [54]. 在离散选择模型中, 这种概率选择的情况是用随机效用理论来处理的. 此时目的地 k 的真实效用表示为

$$U_{ik} = V_{ik} + \varepsilon_{ik}, \quad (6)$$

其中 V_{ik} 是研究者直接观测到的选项 k 的确定效用, ε_{ik} 是描述 V_{ik} 与真实效用 U_{ik} 偏差程度的随机项.

3.2 离散选择模型导出引力模型

根据 (5) 式和 (6) 式可知, 在出行目的地选择问题中, 所处 i 地点的决策者选择目的地 j 能获得的效用是 $U_{ij} = V_{ij} + \varepsilon_{ij}$, 则目的地 j 被 i 点决策者选择的概率是

$$P_{ij} = \Pr(V_{ij} + \varepsilon_{ij} > V_{ik} + \varepsilon_{ik}, \forall k \in K - \{j\}), \quad (7)$$

其中 K 是所有备选目的地集合 (不包括起点 i).

如果 (7) 式中所有随机项 ε 都服从相互独立同分布的标准 Gumbel 分布 $F(\varepsilon) = e^{-e^{-\varepsilon}}$, 那么决策者选择目的地 j 的概率就是

$$\begin{aligned} P_{ij} &= \Pr(\varepsilon_{ik} < V_{ij} - V_{ik} + \varepsilon_{ij}, \forall k \in K - \{j\}) \\ &= \Pr(\varepsilon_{ij} = x) \cdot \prod_{k \neq j} \Pr(\varepsilon_{ik} < V_{ij} - V_{ik} + x) \\ &= \int_{-\infty}^{\infty} \frac{dF(x)}{dx} F(V_{ij} - V_{ik} + x) dx \\ &= \int_{-\infty}^{\infty} e^{-x} e^{-e^{-x}} \prod_{k \neq j} e^{-e^{-(V_{ij} - V_{ik} + x)}} dx \\ &= \int_{-\infty}^{\infty} e^{-e^{-x} \sum_k e^{-(V_{ij} - V_{ik})}} e^{-x} dx \\ &= \frac{e^{V_{ij}}}{\sum_k e^{V_{ik}}}. \end{aligned} \quad (8)$$

这就是 Logit 模型 [54], 是微观经济学中应用非常广泛的离散选择模型.

由 Logit 模型可以很容易地导出引力模型: 假设目的地 j 对 i 点出行者的确定效用主要受目的地活力 A_j 和地点 i 到 j 的距离 d_{ij} 影响, 即

$$V_{ij} = \ln A_j - \beta \ln d_{ij} + \gamma. \quad (9)$$

如果所有出行者之间的选择决策行为无差异, 那么从点 i 到 j 的出行量 T_{ij} 就是从 i 出发的总出行量 O_i 乘以出行者们选择 j 点的概率, 即

$$T_{ij} = O_i P_{ij} = O_i \frac{e^{V_{ij}}}{\sum_k e^{V_{ik}}} = O_i \frac{A_j \cdot d_{ij}^{-\beta}}{\sum_k A_k \cdot d_{ik}^{-\beta}}, \quad (10)$$

这就是经典的单约束引力模型 [7].

3.3 收益与活力的对数关系

从 (9) 式中可以看出, 目的地的效用 V_{ij} 实际上是目的地活力为出行者带来的收益 $\ln A_j$ 与出行者到目的地的成本 $\beta \ln d_{ij}$ 的差值. 2.2 节已对出行成本与距离的对数关系进行了解释, 本节将解释为

何使用地点活力的对数来表达地点会带给出行者收益.

根据心理物理学中著名的 Weber-Fechner 定律 [55], 人的直观感觉差异 dp 正比于某种物理刺激强度 W 的相对变化量 dW/W , 即 $dp = \kappa dW/W$, 其中 κ 是个常数. 据此可导出人的直观感觉程度 p 与物理刺激强度之间的对数关系 $p = \kappa \ln(W/W_0)$, 其中 W_0 可以解释为刺激阈值. Weber-Fechner 定律在行为经济学中被广泛应用于收益函数 [55], 即认为选项带给决策者的收益是决策者对选项属性的实际度量指标 (如商品的价格、反映地点活力的人口数等) 的直观感觉. 在本文即将介绍的目的地选择博弈研究工作中, Weber-Fechner 定律导出的收益活力对数关系也会被再次使用.

4 社会引力定律的博弈论解释

基于随机效用理论的离散选择模型从个体选择决策角度为社会引力定律提供了新的解释, 但这类模型并未考虑实际出行目的地选择过程中个体之间的相互作用. 在包括交通系统在内的社会经济复杂系统中, 个体间的相互作用是一个非常普遍的现象, 而博弈论则是社会学、经济学以至军事、政治等社会科学学科研究个体相互作用的一个重要科学工具 [56]. 最近, Yan 和 Zhou [37] 从博弈论的角度解释了社会引力定律的可能根源. 他们将出行者选择目的地的过程刻画为一种拥挤博弈 [57,58], 体现了出行者个体之间的相互作用. 本节将介绍这种目的地选择博弈模型的基本框架及其导出引力模型的方法.

4.1 目的地选择博弈模型

出行目的地选择问题可以看作是多个参与者进行的博弈. 每个参与者在面对多个可供选择的目的地时, 总会选择带给自己收益 (或称效用) 最大的目的地. 位于地点 i 的个体选择目的地 j 获得的收益 U_{ij} 由两部分组成.

1) 出行成本 $C_{ij} + g(T_{ij})$, 即负收益, 与 i 到 j 的固定出行成本 C_{ij} 直接相关, 并且随两地间出行量 T_{ij} 的增加而增加. 其中 $g(T_{ij})$ 是一个增函数, 体现了路途上的拥挤效应.

2) 目的地带给选择者的收益 $h(A_j) - f(D_j)$, 取决于目的地的活力 A_j 和选择该目的地的人数

$D_j = \sum_i T_{ij}$. 其中 $f(D_j)$ 是一个增函数, 体现了目的地的拥挤效应, 即目的地的收益会随着选择人数的增加而下降. 这种拥挤效应具有现实依据: 例如在城市出行中, 对于非通勤出行 (例如购物、娱乐等) 来说, 选择某一目的地的人数增加后可能导致环境不舒适或可获取资源的减少, 从而降低选择者的收益.

综上, i 地点的出行者选择目的地 j 的收益可表示为

$$U_{ij} = h(A_j) - f(D_j) - C_{ij} - g(T_{ij}). \quad (11)$$

这种模型被命名为目的地选择博弈 (destination choice game, DCG) 模型 [37]. 在 DCG 模型中, 当个体信息完备并总是选择使自己收益最大化的目的地时, 相同起点出行的所有个体都具有相等的收益, 没有人能通过单方面改变选择而增加自己的收益.

实际应用中, 需要先把 (11) 式中的各项成本和收益函数具体化. 根据出行成本与距离之间的对数关系 (见 2.2 节), 可令 $C_{ij} \propto \ln d_{ij}$. 根据 Weber-Fechner 定律 [55] (见 3.3 节), 则可将目的地活力收益函数 $h(A_j)$ 、目的地拥挤成本函数 $f(D_j)$ 和路途拥挤函数 $g(T_{ij})$ 均表示为对数函数. 此时 (11) 式可写为

$$U_{ij} = \alpha \ln A_j - \beta \ln d_{ij} - \gamma \ln D_j - \ln T_{ij}, \quad (12)$$

其中 α , β 和 γ 是三个非负参数. 通过使用多种真实人类移动数据进行的模型测试结果 [37] 显示, DCG 模型相对于传统的引力模型 [4]、介入机会模型 [59] 和新型辐射模型 [60]、人口权重机会模型 [61,62], 具有更高的预测精度.

4.2 拥挤博弈模型导出引力模型

若忽略 DCG 模型收益函数中的目的地拥挤成本, 则可得到一个简化的目的地选择博弈 (degenerated destination choice game, DDCG) 模型, 其收益函数如下:

$$U_{ij} = \alpha \ln A_j - \beta \ln d_{ij} - \ln T_{ij}, \quad (13)$$

这是一个典型的拥挤博弈模型 [57,58]. 拥挤博弈可表示为一个四元组: $(N, R, (\Psi_k)_{k \in N}, (w_j)_{j \in R})$ [58,63], 其中 $N = \{1, 2, \dots, n\}$ 是所有参与博弈的个体 (在 DDCG 中就是某个出发地点 i 上数量为 O_i 的所有出行者), $R = \{1, 2, \dots, m\}$ 是备选资源集合 (在 DDCG 中就是备选目的地), $(\Psi_k) \subseteq 2^R$ 是参与者

k 的策略空间, w_j 是备选资源 j 的收益函数 (在 DDCG 中就是目的地 j 对 i 起点出行者的收益函数 U_{ij}). 所有参与者的策略集合 $S = S_1, \dots, S_n$ 就是拥挤博弈的一个状态, 其中参与者 k 的策略 $S_k \in \Psi_k$. 资源 j 的拥挤程度 $n_j(S)$ (在 DDCG 中就是目的地 j 的路途拥挤程度) 则表示在状态 S 下选择该资源的参与者数量 (在 DDCG 中就是从起点 i 到目的地 j 的出行量 T_{ij}). 据此可知, DDCG 中每个出行者获取的收益不仅取决于其所选择的目的地的活力与距离, 还受选择同样目的地的其他出行者数量的影响, 即个体之间的相互作用. 根据拥挤博弈理论中的势函数 [58,63] 定义, 可知 DDCG 的势函数为

$$\begin{aligned} \phi &= \sum_{j \in R} \sum_{k=1}^{T_{ij}(S)} U_{ij}(k) \\ &= \sum_{j \in R} \sum_{k=1}^{T_{ij}(S)} (\alpha \ln A_j - \beta \ln d_{ij} - \ln k). \end{aligned} \quad (14)$$

最大化此势函数的策略集合, 就是拥挤博弈的纳什均衡解 [57,63].

为求解 DDCG 的纳什均衡解, 可将 T_{ij} 视为一个连续变量, 并令势函数最大化, 即可得到一个最优化模型

$$\begin{aligned} \max \phi(S) &= \sum_j \int_0^{T_{ij}} (\alpha \ln A_j - \beta \ln d_{ij} - \ln x) dx, \\ \text{s.t.} \quad &\sum_j T_{ij} = O_i. \end{aligned} \quad (15)$$

用拉格朗日乘子法可求得该模型的解为

$$T_{ij} = O_i \frac{A_j^\alpha d_{ij}^{-\beta}}{\sum_j A_j^\alpha d_{ij}^{-\beta}}, \quad (16)$$

是一个双参数的单约束引力模型. 若令参数 $\alpha = 1$, 则可转换为标准的单约束引力模型, 见 (10) 式.

4.3 DDCG 模型与 Logit 模型的对比

DDCG 模型与 3.2 节中介绍的离散选择 Logit 模型都体现了个体选择目的地的微观决策行为, 且都假设个体总是追求效用最大. 二者导出的引力模型在形式上也没有本质区别, 见 (10) 式和 (16) 式. 但二者的底层机制并不相同: Logit 模型假设个体对效用的认知具有随机性, 但并未考虑复杂系统中广泛存在的个体相互作用 [56]; 而 DDCG 模型则考

考虑了群体中个体间的相互作用.

著名的红蓝巴士问题^[64]可以用来说明这两个模型的机制差异. 在红蓝巴士问题中, 出行者可选择的交通方式有小汽车与蓝巴士两种. 简单起见, 假设这两种交通方式的确定效用是相等的, 根据 Logit 模型可知二者的被选概率为 $P_{\text{car}} = P_{\text{blue}} = 1/2$. 现在把一半蓝巴士的颜色改为红色, 并把红蓝巴士看成是两种交通方式, 那么两者的确定效用也必然相等. 按照 Logit 模型的结果, 此时出行者选择这三种交通方式的概率为 $P_{\text{car}} = P_{\text{blue}} = P_{\text{red}} = 1/3$. 然而, 从常识推断, 将半数巴士涂红仅会影响原先选择蓝巴士的人, 而小汽车的被选概率仍为 $P_{\text{car}} = 1/2$, 红、蓝巴士则各为 $P_{\text{red}} = P_{\text{blue}} = 1/4$. 但 Logit 模型却高估了巴士被选择的概率, 低估了小汽车被选择的概率, 形成了典型的悖论^[64].

在考虑了个体相互作用的 DDCG 模型中则不存在这种悖论: 假设总共有 b 辆巴士, 每辆巴士的固定收益值为 A_{bus} , 拥挤成本 (体现了个体之间相互作用的结果) 为 $\ln(ax)$, 其中 x 代表乘坐这辆巴士的人数. 那么当把所有巴士看成同一交通方式时, 其收益值为 $U_{\text{bus}} = A_{\text{bus}} - \ln(ax/b)$; 而把红蓝巴士看作两种交通方式时, 二者的收益值则均为 $U_{\text{red}} = U_{\text{blue}} = A_{\text{bus}} - \ln(2ax/b)$. 根据 DDCG 模型的均衡条件, 如果在 m 个人中选择巴士和小汽车的比例各为 $1/2$, 说明 $U_{\text{car}}(m/2) = U_{\text{bus}}(m/2) = A_{\text{bus}} - \ln[am/(2b)]$. 那么把红蓝巴士看作两种方式时, 一定有 $U_{\text{car}}(m/2) = U_{\text{red}}(m/4) = U_{\text{blue}}(m/4)$, 即选择红蓝巴士的比例各为 $1/4$, 与实际相符.

当然, 实际中的出行者对目的地效用 (即收益) 的感知可能会有一定的随机性. 此时可对 DDCG 模型进行进一步的扩展: 假设出行者对目的地收益的理解与可观测的目的地固定收益之间存在服从独立同 Gumbel 分布的随机性偏差, 根据 Fisk^[65] 的证明, 此时的均衡解等价于如下最优化模型的解

$$\begin{aligned} \min Z(x) &= \sum_j \int_0^{T_{ij}} (-\alpha \ln A_j + \beta \ln d_{ij} + \ln x) dx \\ &\quad + \frac{1}{\theta} \sum_j T_{ij} \ln T_{ij}, \\ \text{s.t.} \quad &\sum_j T_{ij} = O_i, \end{aligned} \tag{17}$$

其解为

$$T_{ij} = O_i \frac{A_j^{\frac{\theta\alpha}{1+\theta}} d_{ij}^{-\frac{\theta\beta}{1+\theta}}}{\sum_j A_j^{\frac{\theta\alpha}{1+\theta}} d_{ij}^{-\frac{\theta\beta}{1+\theta}}}, \tag{18}$$

式中的参数 θ 体现了随机性程度对目的地选择结果的影响: 当 $\theta \rightarrow \infty$ 时该模型退化为 DDCG 模型, 均衡结果与 (16) 式一致; 当 $\theta = 0$ 时, 个体将完全随机地选择目的地, 各目的地被选择的概率相等. 这一结果说明, DDCG 模型可以扩展到个体对目的地收益感知有随机性的情形. 但是, 基于随机效用理论的离散选择模型却无法退化到目的地收益固定的情形: 如果离散选择模型不假设个体对收益感知具有随机性, 那么所有个体都会选择固定收益最高的同一目的地. 这是因为离散选择模型并未考虑实际社会系统中普遍存在的个体相互作用, 即群体拥挤问题.

5 结 论

社会引力定律是诸多社会复杂系统中广泛存在的普适规律, 受到社会学、经济学、地理学、交通科学、统计物理与复杂系统等学科学者持久的关注^[1,4,6,7,32,33,39]. 本文综述了解释社会引力定律存在根源的三类主要理论, 既包括传统的最大熵原理和效用理论, 也包括最近提出的目的地选择拥挤博弈. 其中, 最大熵原理为社会引力定律建立了一个统计物理的理论基础, 但其仅能给出系统最可能的宏观分布状态, 却无法反映系统中个体的微观决策行为. 从个体决策行为角度建立的随机效用离散选择模型为社会引力定律提供了微观经济学解释, 但其并未考虑在实际社会系统中普遍存在的个体相互作用. 而目的地选择博弈则为社会引力定律建立了一个更简洁的理论框架: 它不需像最大熵模型一样事先指定出行总成本作为额外的约束条件, 也不需像随机效用理论一样作出个体对效用感知具有随机性的假设, 除“个体追求收益最大化”这条经济学公理之外, 目的地选择博弈的核心假设就只有“群体拥挤影响收益”这一体现个体相互作用的基本事实. 与现有对社会引力定律的其他理论解释相比, 从博弈论角度给出的解释更符合“奥卡姆剃刀”原理^[66], 有助于我们理解很多复杂社会系统中由个体相互作用而涌现出的群体空间交互模式.

本文介绍的这些对社会引力定律的理论解释都是以交通出行问题为背景的, 离散选择模型和目

的地选择博弈模型分析的核心都是出行者如何选择目的地. 但这种个体选择交互对象的行为并不仅仅存在于交通系统中, 在空间交互模式符合社会引力定律的系统中几乎都存在. 例如, 人口迁移是选地点做居住地, 社会交往是选人做交流对象, 科研合作则是选研究者做合作伙伴等. 在这些系统中, 个体多会倾向于选择活力相对高、距离相对近的对象. 但往往选择同一对象的人越多, 该对象能带来的收益就会越低: 在交通系统中拥挤会导致出行成本上升, 在商品贸易中竞争对手增加会导致商品价格下降, 在科研合作中选择某人的合作者增多会导致双方的合作强度降低, 等等. 而这些不同系统中的空间交互对象选择行为, 都能被本文所介绍的目的地选择博弈等模型来刻画. 总之, 对社会引力定律追根溯源, 不仅有助于深入理解交通出行、人口迁移、商品贸易、信息流通、社会交往、科研合作等空间交互现象的底层机制, 对于更好地预测、引导甚至控制各种复杂社会系统中人、物、信息的流动, 都具有广阔的应用前景.

感谢合作者周涛、汪秉宏、韩筱璞对本文中相关研究成果所做出的贡献.

参考文献

- [1] Yan X Y 2019 *Beyond Gravity Law* (Beijing: China Science Press) piii (in Chinese) [闫小勇 2019 超越引力定律 (北京: 科学出版社) 第iii页]
- [2] Ding Y M, Yang C P 2012 *Acta Phys. Sin.* **61** 238901 (in Chinese) [丁益民, 杨昌平 2012 物理学报 **61** 238901]
- [3] Fotheringham A S, O'Kelly M E 1989 *Spatial Interaction Models: Formulations and Applications* (Dordrecht: Kluwer Academic Publishers) p1
- [4] Sen A, Smith T 1995 *Gravity Models of Spatial Interaction Behavior* (Berlin: Springer) p1
- [5] Karemera D, Oguledo V I, Davis B 2000 *Appl. Econ.* **32** 1745
- [6] Porojan A 2001 *Open Econ. Rev.* **12** 265
- [7] de Dios Ortuzar J, Willumsen L G 2011 *Modelling Transport* (West Sussex: John Wiley & Sons) p182
- [8] Liu J H, Zhang Z K, Chen L, Liu C, Yang C, Wang X 2014 *PloS ONE* **9** e91070
- [9] Ma L L, Ma C, Zhang H F, Wang B H 2015 *Physica A* **451** 205
- [10] Li Z, Ren T, Ma X Q, Liu S M, Zhang Y X, Zhou T 2019 *Sci. Rep.* **9** 8387
- [11] Newton I 1729 *Mathematical Principles of Natural Philosophy* (London: Benjamin Motte Publisher) p5
- [12] Dodd S C 1950 *Am. Sociol. Rev.* **15** 245
- [13] Desart H G 1846 *Chemin de fer Direct de Bruxelles vers Gand, par Alost, en Communication avec les Stations Diverses* (Bruxelles: Devroye) p16
- [14] Odlyzko A 2015 *Economia Hist. Meth. Philos.* **5** 157
- [15] Carey H C 1858 *Principles of Social Science* (Philadelphia: Lippincott) p28
- [16] Ravenstein E G 1889 *J. R. Stat. Soc.* **48** 167
- [17] Reilly W J 1929 *Methods for the Study of Retail Relationships* (Texas: University of Texas Bulletin) p2944
- [18] Zipf G K 1946 *Am. Sociol. Rev.* **11** 677
- [19] Tobler W 1995 *Urban Geogr.* **16** 327
- [20] Fagiolo G 2010 *J. Econ. Interact. Coor.* **5** 1
- [21] Zhou T, Han X P, Yan X Y, Yang Z M, Zhao Z D, Wang B H 2013 *J. Univ. Electron. Sci. Technol. China* **42** 481 (in Chinese) [周涛, 韩筱璞, 闫小勇, 杨紫陌, 赵志丹, 汪秉宏 2013 电子科技大学学报 **42** 481]
- [22] Xu Z X, Wang Y, Si H B, Feng Z M 2011 *Acta Phys. Sin.* **60** 040501 (in Chinese) [徐赞新, 王钺, 司洪波, 冯振明 2011 物理学报 **60** 040501]
- [23] Viboud C, Bjørnstad O N, Smith D L, Simonsen L, Miller M A, Grenfell B T 2006 *Science* **312** 447
- [24] Jung W S, Wang F, Stanley H E 2008 *Europhys. Lett.* **81** 48005
- [25] Krings G, Calabrese F, Ratti C, Blondel V D 2009 *J. Stat. Mech.* **2009** L07003
- [26] Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco J J, Vespignani A 2009 *Proc. Natl. Acad. Sci. U.S.A.* **106** 21484
- [27] Kaluza P, Kölsch A, Gastner M T, Blasius B 2010 *J. R. Soc. Interface* **7** 1093
- [28] Pan R K, Kaski K, Fortunato S 2012 *Sci. Rep.* **2** 902
- [29] Goh S, Lee K, Park J S, Choi M Y 2012 *Phys. Rev. E* **86** 026102
- [30] Goh S, Lee K, Choi M Y, Fortin J Y 2014 *PloS ONE* **9** e89980
- [31] Levy M, Goldenberg J 2014 *Physica A* **393** 418
- [32] Sheppard E S 1978 *Geogr. Anal.* **10** 386
- [33] Hua C I, Porell F 1979 *Int. Regional Sci. Rev.* **4** 97
- [34] Wilson A G 1967 *Transp. Res.* **1** 253
- [35] Niedercorn J H, Bechdolt Jr B V 1969 *J. Regional Sci.* **9** 273
- [36] Domencich T A, Mcfadden D 1975 *Urban Travel Demand: A Behavioral Analysis* (Amsterdam: North-Holland) p1
- [37] Yan X Y, Zhou T 2019 *Sci. Rep.* **9** 9466
- [38] Pressé S, Ghosh K, Lee J, Dill K A 2013 *Rev. Mod. Phys.* **85** 1115
- [39] Barthélemy M 2011 *Phys. Rep.* **499** 1
- [40] Yan X Y, Han X P, Wang B H, Zhou T 2013 *Sci. Rep.* **3** 2678
- [41] Hensher D A, Button K J 2000 *Handbook of Transport Modelling* (Oxford: Pergamon) p20
- [42] Chalasani V S, Engebretsen O, Denstadli J M, Axhausen K W 2005 *J. Transp. Stat.* **2** 1
- [43] Rietveld P, Zwart B, Van Wee B, van den Hoorn T 1999 *Ann. Regional Sci.* **33** 269
- [44] Brakman S, Garretsen H 2005 *Location and Competition* (London: Routledge) p205
- [45] Brockmann D, Hufnagel L, Geisel T 2006 *Nature* **439** 462
- [46] González M C, Hidalgo C A, Barabási A L 2008 *Nature* **453** 779
- [47] Han X P, Hao Q, Wang B H, Zhou T 2011 *Phys. Rev. E* **83** 036117
- [48] Liang X, Zheng X, Lü W, Zhu T, Xu K 2012 *Physica A* **391** 2153
- [49] Gallotti R, Bazzani A, Rambaldi S 2012 *Int. J. Mod. Phys. C* **23** 1250061
- [50] Huang F H, Peng J, You M Y 2016 *Acta Phys. Sin.* **65** 228901 (in Chinese) [黄飞虎, 彭舰, 由明阳 2016 物理学报 **65** 228901]

- [51] Ben-Akiva M E, Lerman S R 1985 *Discrete Choice Analysis: Theory and Application to Travel Demand* (Cambridge: MIT Press) p1
- [52] McFadden D 2007 *Transp. Policy* **14** 269
- [53] Herrstein R J, Loewenstein G F, Prelec D, Vaughan Jr W 1993 *Journal of Behav. Decis. Making* **6** 149
- [54] Train K E 2009 *Discrete Choice Methods with Simulation* (Cambridge: Cambridge University Press) p38
- [55] Takemura K 2014 *Behavioral Decision Theory: Psychological and Mathematical Descriptions of Human Choice Behavior* (Berlin: Springer) p52
- [56] He D R, Liu Z H, Wang B H 2009 *Complex Systems and Complex Networks* (Beijing: Higher Education Press) p59 (in Chinese) [何大韧, 刘宗华, 汪秉宏 2009 复杂系统与复杂网络 (北京: 高等教育出版社) 第59页]
- [57] Rosenthal R W 1973 *Int. J Game Theory* **2** 65
- [58] Monderer D, Shapley L S 1996 *Games Econ. Behav.* **14** 124
- [59] Stouffer S A 1940 *Am. Sociol. Rev.* **5** 845
- [60] Simini F, González M C, Maritan A, Barabási A L 2012 *Nature* **484** 96
- [61] Yan X Y, Zhao C, Fan Y, Di Z R, Wang W X 2014 *J R Soc. Interface* **11** 20140834
- [62] Yan X Y, Wang W X, Gao Z Y, Lai Y C 2017 *Nat. Commun.* **8** 1639
- [63] Voorneveld M, Borm P, Van Meegen F, Tijs S, Facchini G 1999 *Int. Game Theory Rev.* **1** 283
- [64] McFadden D 1980 *J. Bus.* **1980** S13
- [65] Fisk C 1980 *Transp. Res. B* **14** 243
- [66] Burgess J 1998 *Occam's Razor and Scientific Method* (Oxford: Clarendon Press) p195

SPECIAL TOPIC—Statistical physics and complex systems

Exploring the roots of social gravity law^{*}

Yan Xiao-Yong^{1)2)†}

1) (*Institute of Transportation System Science and Engineering, Beijing Jiaotong University, Beijing 100044, China*)

2) (*Complex Laboratory, University of Electronic Science and Technology of China, Chengdu 611731, China*)

(Received 4 November 2019; revised manuscript received 19 November 2019)

Abstract

Many spatial mobility of people, goods and information, such as human travel, population migration, commodity trade, information communication, social interaction and scientific cooperation, follow a law similar to Newton's law of universal gravitation. This law, named social gravity law, is that the flow between two locations is directly proportional to the product of the vitality of these two locations, and inversely proportional to a power function of their distance. The gravity model established by analogy with the gravity law has also been widely used to predict trip distribution, population migration, interregional trade flows, etc. But why do many complex social systems have such a simple law? It is an interesting and valuable issue. This paper reviews the research on exploring the roots of the social gravity law from various perspectives, including statistical physics, microeconomics, and game theory.

Keywords: social gravity law, principle of maximum entropy, discrete choice model, congestion game

PACS: 89.75.-k, 89.65.-s, 89.70.Cf, 02.50.Le

DOI: 10.7498/aps.69.20191686

* Project supported by the National Natural Science Foundation of China (Grant Nos. 71822102, 71621001, 71671015, 61304177).

† Corresponding author. E-mail: yanxy@bjtu.edu.cn

多层网络级联失效的预防和恢复策略概述

蒋文君 刘润然 范天龙 刘霜霜 吕琳媛

Overview of precaution and recovery strategies for cascading failures in multilayer networks

Jiang Wen-Jun Liu Run-Ran Fan Tian-Long Liu Shuang-Shuang Lü Lin-Yuan

引用信息 Citation: *Acta Physica Sinica*, 69, 088904 (2020) DOI: 10.7498/aps.69.20192000

在线阅读 View online: <https://doi.org/10.7498/aps.69.20192000>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

相依网络上基于相连边的择优恢复算法

Enhancing resilience of interdependent networks against cascading failures under preferential recovery strategies

物理学报. 2018, 67(8): 088901 <https://doi.org/10.7498/aps.67.20172526>

复杂网络电输运性能与通信序列熵之间的关联

Correlation between the electrical transport performance and the communicability sequence entropy in complex networks

物理学报. 2019, 68(11): 118901 <https://doi.org/10.7498/aps.68.20190230>

复杂网络链路可预测性: 基于特征谱视角

The link predictability of complex network from a spectrum perspective

物理学报. 2020, 69(8): 1 <https://doi.org/10.7498/aps.69.20191817>

基于复杂网络动力学模型的无向加权网络节点重要性评估

Evaluation methods of node importance in undirected weighted networks based on complex network dynamics models

物理学报. 2018, 67(9): 098901 <https://doi.org/10.7498/aps.67.20172295>

基于领域相似度的复杂网络节点重要度评估算法

Node importance measurement based on neighborhood similarity in complex network

物理学报. 2017, 66(3): 038902 <https://doi.org/10.7498/aps.66.038902>

专题：统计物理和复杂系统

多层网络级联失效的预防和恢复策略概述*

蒋文君¹⁾ 刘润然^{1)†} 范天龙²⁾¹⁾ 刘霜霜¹⁾ 吕琳媛^{2)1)‡}

1) (杭州师范大学, 阿里巴巴复杂科学研究中心, 杭州 311121)

2) (电子科技大学, 基础与前沿研究院, 成都 611731)

(2019年12月31日收到; 2020年2月1日收到修改稿)

现实生活中, 与国计民生密切相关的基础设施网络大多不是独立存在的, 而是彼此之间相互联系或依赖的, 于是用于研究这些系统的多层网络模型随之产生. 多层网络中的节点在失效或者遭受攻击后会因“层内”和“层间”的相互作用而产生级联效应, 从而使得失效能够在网络层内和层间反复传播并使得失效规模逐步放大. 因此, 多层网络比单个网络更加脆弱. 多层网络级联失效产生的影响和损失往往是非常巨大的, 所以对多层网络级联失效的预防和恢复的研究具有重大意义. 就多层网络级联失效的预防而言, 主要包含故障检测, 保护重要节点, 改变网络耦合机制和节点备份等策略. 就多层网络发生级联失效后的恢复策略而言, 主要包含共同边界节点恢复、空闲连边恢复、加边恢复、重要节点优先恢复、更改拓扑结构、局域攻击修复、自适应边修复等策略.

关键词: 复杂网络, 级联失效, 预防策略, 恢复策略

PACS: 89.75.-k, 89.90.+n

DOI: 10.7498/aps.69.20192000

1 引言

在复杂网络研究的早期, 单个网络上的多个动力学特性均受到了广泛的关注, 如疾病传播^[1,2]、网络同步^[3,4]、级联失效^[5]和网络控制^[6,7]等. 而随着研究的深入, 人们发现很多现实网络都不是孤立存在的, 比如电力网络和通讯网络之间存在相互依赖关系. 这些网络会因与其他网络之间的依赖关系, 在面临蓄意攻击或随机故障时比孤立网络更加脆弱^[8]. 现实中存在互相依赖和联系的复杂系统非常多, 如黑客或者病毒的攻击会使得因特网出现故障甚至瘫痪, 从而导致银行金融系统、电力网络、交通网络和物流信息网络等一系列关键基础设施的数据采集与监视控制系统 (supervisory control and data acquisition, SCADA) 无法正常工作, 进

而导致这些系统的瘫痪和崩溃. 比如电力网络中的发电站需要铁路网络等为其运送燃料和物资的补给, 而铁路网络也需要通过电力网络和通信网络提供支撑和控制. 图 1^[9]总结了电力基础设施网络和其他基础设施之间的依赖关系. 这些基础设施中的一个或某几个一旦出现故障或受到攻击, 其影响都会快速地扩散到其他相关网络中, 引发一系列迭代级联事故, 从而将损害扩大到更广范围. 发生在 2003 年意大利停电事故^[10]和 2005 年 8 月印度尼西亚大停电事故均凸显了这种大规模的耦合网络故障对社会生产生活甚至国家安全带来的巨大风险. Tootaghaj 等^[11]搜集了全球近年来比较重大的停电事故, 如表 1 所列. 为了避免和减少级联失效对基础设施所带来的损害, 2016 年, 我国通过了《网络安全法》, 构建起以信息共享为基础, 事前预防、事中控制、事后恢复与惩治的关键信息基础

* 国家自然科学基金 (批准号: 61773148, 61673150, 11622538) 和“沿创, 研途人生”科研类项目 (批准号: 2018YCZD09) 资助的课题.

† 通信作者. E-mail: runranliu@163.com

‡ 通信作者. E-mail: linyuan.lv@uestc.edu.cn

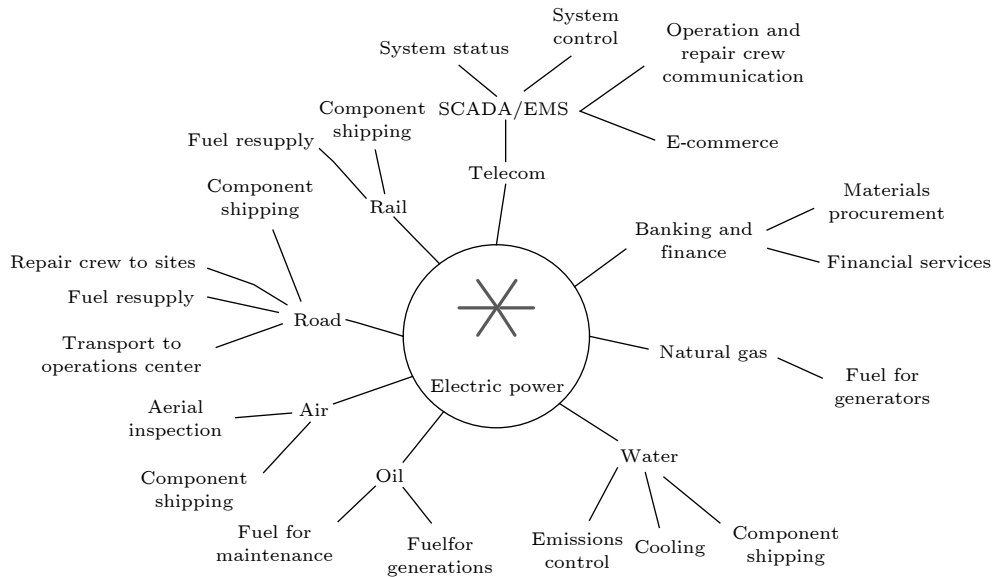


图 1 电力基础设施依赖关系^[9]

Fig. 1. Dependencies among power infrastructures^[9].

设施保护体系^[12]. 美国在这方面也甚为关注, 自克林顿政府以来就出台大量的相关法律文件: 第 63 号总统令和《国土安全法》等, 扩展对关键基础设施的保护范围^[13].

表 1 重大停电事故数据^[11]

Table 1. Data on major power outages^[11].

地点	日期	受灾人数/百万
巴西	1999/3/11	97
印度	2001/1/2	230
美国, 加拿大	2003/8/14-15	55
意大利, 瑞士	2003/9/28	55
印度尼西亚	2005/8/18	100
巴西, 巴拉圭	2009/11/10-11	87
土耳其	2015/3/31	70
印度	2012/7/30-31	620
孟加拉	2014/11/1/	150
肯尼亚	2016/6/7	44

如何才能避免相互依赖系统级联效应的发生呢? 在这些系统发生级联失效后, 如何在系统完全崩溃之前修复并减小级联失效所带来的损失呢^[14]? 常用策略和方案通常分为两种类型: 1) 通过故障检测和保护关键节点等预防措施来减少故障发生的可能性; 2) 当故障发生时采用合适的恢复策略, 对故障进行恢复.

接下来, 本文第 2 节介绍相依网络级联失效的经典模型. 第 3, 4 节则分别梳理了预防多层网络

级联失效的方法和级联失效发生后的网络恢复策略, 而第 5 节是讨论部分.

2 相依网络级联失效模型

虽然很多人都注意到了网络之间的耦合关系, 也认为多个基础设施网络之间往往不是单一存在的, 但是这一问题的研究一直未能形成一个清晰的框架. 直到 2010 年, Buldyrev 等^[10]在 *Nature* 杂志上提出了由两个一对一相互依赖的网络构成的双层网络级联失效模型, 并建立了相关的理论分析方法, 发现相依网络在遭受攻击后的破碎形式为一阶不连续相变, 这与单层网络的二阶连续相变有着本质的不同. 这一结果意味着级联过程对原有的多层网络动力学具有深刻的影响. 但 Buldyrev 等提出的模型较为严格, 具有度分布 $p_A(k)$ 的网络 A 和度分布 $p_B(k)$ 的网络 B 需要有相同的节点数 N , 并且两个网络中的所有节点之间要建立一对一的完全依赖关系, 也就是说, 属于网络 A 中的节点 a_i 唯一依赖网络 B 中的节点 b_j , 而网络 B 中的 b_j 也必定唯一依赖网络 A 中的 a_i , 并且当网络 A 中的节点失效后, 网络 B 中与之对应的节点也会立刻失效, 反之亦然. 其级联失效的过程如下: 从网络 A 中随机删除一定比例的节点后, 网络 A 中新产生的孤立节点会失效, 随后网络 B 中与网络 A 中的所有失效节点有依赖关系的节点也会失效, 而网

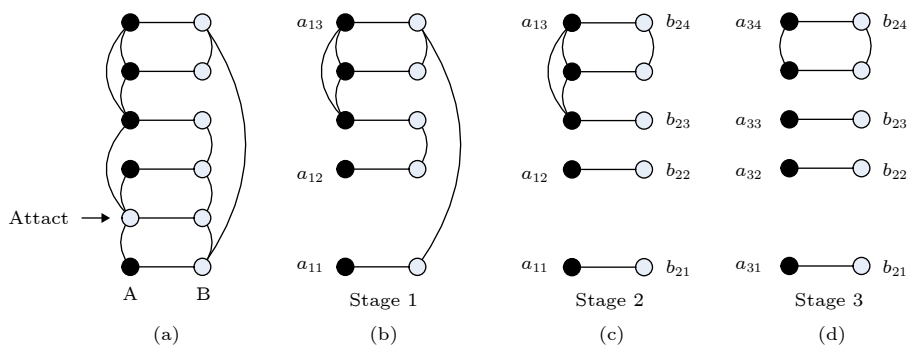


图 2 级联失效迭代过程的建模^[10] (a) 网络在初始状态下遭到攻击; (b), (c) 和 (d) 网络在遭受攻击后网络级联失效的不同阶段, 并最终达到了稳态, 级联过程结束

Fig. 2. Modeling of cascading failure iterative processes^[10]: (a) The network is attacked in the initial state; (b), (c), and (d) are the cascading failure processes of the network due to the dependencies between dependent networks after the attack, respectively. Eventually reached a steady state.

络 B 中失效的节点又会反馈到网络 A 中, 如此反复迭代, 直到不再有新的节点失效, 网络达到稳态, 级联失效过程结束. 图 2 展示了两个网络的级联失效过程.

经典相依网络级联失效模型的提出引发了更多学者开展对多层网络的动力学研究. 在两层网络的基础上, Gao 等^[15,16] 进一步提出了多个网络存在相互依赖的“网络的网络 (network of network, NON)”模型. 随着对相依网络级联失效模型研究的推进, 关于网络鲁棒性的研究^[17] 以及多层网络之间的动力学研究也吸引了愈来愈多的学者并取得丰硕成果. Baxter 等^[18] 阐明了一阶相变的本质其实是混合相变, 当临界节点形成的临界簇发散的时候, 一旦其中的“基石”节点 (keystone vertex) 失效, 整个临界簇就会发生雪崩现象, 导致多层网络互连巨分量出现不连续的跳跃. 多层网络的结构特性也对网络的鲁棒性有着重要的影响, 如簇结构和度关联等. Faqeeh 等^[19] 在具有模块的网络中发现了不止一个渗流簇 (coexist percolation cluster, CPC), 对理解网络的鲁棒性以及传染病模型中的疾病爆发具有重要意义. 在网络的同配性研究方面^[20,21], 文献^[20] 比较了同配性对单层网络和多层网络上动力学过程的不同影响, 发现在多层网络上, 增加层间同配程度可以提高信息扩散效率; 而增加层间异配程度则可以分散链路通信负载, 增强网络鲁棒性. 此外还有一些学者研究了空间地理效应^[22,23] 和有向网络^[24–27] 等因素对网络的鲁棒性所产生的影响. 通过考虑多层网络层内节点的不同耦合方式, 一些文献也研究了多层网络上的扩展渗

流模型, 例如 k -core 渗流^[28]、靴攀渗流 (bootstrap percolation)^[29,30] 和弱渗流^[31] 等, 极大地丰富了相依网络级联失效方面的研究.

3 预防策略

预防策略是指在多层网络发生故障前采取的预防应对措施, 主要包括故障检测、预先保护重要节点、节点备份以及改变耦合机制等. 通过采取恰当的预防策略可以减少网络关键节点发生失效的概率并有效防御恶意攻击, 从而极大降低由于故障和攻击所带来的社会经济损失. 例如, 安装杀毒软件可以大概率防止电脑被计算机病毒感染和损坏. 相比事后的修复和采取应对措施, 恰当的预防策略相对成本低且效用大.

3.1 故障检测

故障检测是指定期查找设备或系统是否出现故障并对这些故障进行修复的过程. 很多时候这些故障可能不会立即产生可以觉察的危害和损失, 但如果不加以排除, 则会在关键时刻或者随着时间积累对系统产生严重损害, 所以应该定期加以排查. 这里将各种基础设施抽象成了网络, 但在现实生活中不同的系统根据其自身特点有不同侧重点和相应的检测方法, 不能一概而论. 比如, 对于电力网络来说需要检测的可能会发生的故障包括发电机组故障、母线故障、输电线路故障和变电所故障等, 而对于计算机网络来说其故障检测则包括硬件故障、软件故障、人为故障和病毒故障等. 通过故障

检测可以提前排查系统中存在的问题,减少其对系统造成的危害.故障检测是事前采取针对性措施,减少和防范故障发生的策略.

3.2 预先保护重要节点

重要节点一般是指少量对网络结构或功能非常重要,且其影响可以快速地波及到网络中大部分节点的节点^[32],关于重要节点的衡量方法有很多^[33,34].通过保护重要节点,可以大大提高复杂网络的鲁棒性^[35-37].

文献^[38]研究在相互依赖的网络上,分别以其中一个网络中的大度节点或者小度节点为攻击目标时网络鲁棒性的变化,他们发现即使是在较低的攻击概率下,在相互依赖的无标度(scale-free)网络^[39]上采取保护大度节点的策略后,网络仍是十分脆弱的. Du等^[40]发现具有较大数量的相连边(同一层网络节点之间的连边)和相依边(不同层网络具有相互依赖关系节点之间的连边)的节点很重要,所以不仅要保护层内或层间连接程度较高的节点,而且要保护层内和层间连边数量之和较大的节点,以此来增加系统的鲁棒性.

另一方面,应该优先对那些能够使得网络快速瓦解的节点进行保护. Osat等^[41]将单层网络的最优渗流推广到多层网络上,并通过最优渗流找到那些被删除后网络不会再出现 $N^{1/2}$ 规模的簇的最小节点集.最优渗流问题的解决方案在网络鲁棒性研究中具有直接的适用性,是瓦解网络最简易的方法. Baxter等^[18]发现多层网络中存在能够导致网络临界簇雪崩的基石节点,并且发现在网络瓦解过程中巨分支崩溃的方式是不连续的混合相变,这与单个网络中平滑的连续相变存在明显差异.除此之外,文献^[42]定义了一种节点的通用性(versatility)属性,来刻画那些在多种不同的动力学过程中都扮演重要角色的节点,并基于此提出了多种中心性指标来识别这类节点,如 Eigenvector versatility 和 PageRank versatility 等.如果我们优先对上述文献中的这些重要节点采取保护措施,就可以有效减缓或者抑制网络的破碎.此外,还有一些其他指标也可以作为选取重要节点进行保护的依据,如度中心性^[43,44]、介数中心性^[45]、 k -壳分解^[46]、半局部中心性^[47]、PageRank^[48]、LeaderRank^[49]、圈比^[34]等.

另有文献^[17]将两层网络看成一个整体,将单层网络上一些衡量节点重要性的指标运用在了双

层网络(或更多层的网络)上,从而找到在多层网络中需要保护的重要节点,提高了相依网络的鲁棒性,主要包括以下方法.

1) T-度中心性保护策略.将单层网络中的度中心性概念推广到两层网络上,得 T-度(two-layer-degree)保护策略.在 T-度中心性保护策略中,将不同层中的节点同等看待,计算每个节点在各自层内的度值大小,节点的重要性按它的 T-度从大到小依次递减.例如,图 3(a)给出了一个由网络 A 和网络 B 组成的相依网络.传统的方法是选择网络 A 中的节点 1 和节点 2(度值分别为 5 和 4),使它们和其依赖节点在故障发生时能够正常工作.而实验表明,在保护节点比例不变的情况下,保护网络 A 中的节点 1 和网络 B 中的节点 1(度值分别为 5 和 5)效果更好,这正是因为它们的 T-度最大.

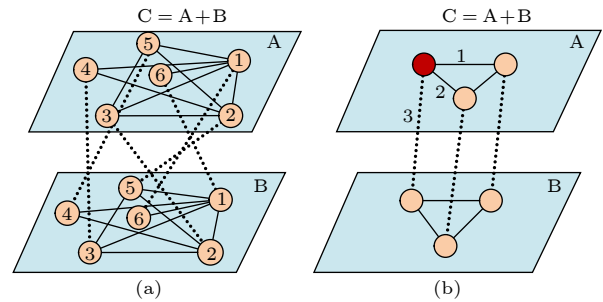


图 3 基于相依网络的保护节点模型^[17]

Fig. 3. Nodes protection model based on interdependent networks^[17].

2) T-介数中心性保护策略.给定两层网络 $G = (C, (L_A, L_B, L_{AB}))$,其中 L_A 表示包含 n_A 个节点的网络 A 中的边, L_B 表示包含 n_B 个节点的网络 B 的边, L_{AB} 则表示网络 A 和网络 B 之间的相依边.将整个相依网络看成一个由 $(n_A + n_B)$ 个节点构成的网络, L_A, L_B 和 L_{AB} 全都同等视为网络 C 中的正常连边.如图 3(b) 中的深色节点,它具有三条连边.对整个网络 C 计算所有节点的介数中心性 $C(i)$,即

$$C(i) = \sum_{i \neq j, i \neq q, j \neq q} \frac{\delta_{jq}(i)}{\delta_{jq}}, \quad (1)$$

其中 δ_{jq} 为从节点 j 到 q 的所有最短路径的数目, $\delta_{jq}(i)$ 表示从节点 j 到节点 q 的 δ_{jq} 条最短路径中经过节点 i 的数目.选择介数中心性最高的部分节点进行保护,称为 T-介数(two-layer-betweenness)

保护策略. 这些既包含相连边又包含相依边的交叉路径在信息传递和故障传播中发挥着重要作用, 但在传统的研究中却被忽视.

3) T-社团 (two-layer-comm) 保护策略^[50]. ① 同上, 将两个相依的网络 A 和网络 B 看作一个网络 C, 初始时假设每个节点本身就是一个社团, 当出现最大的模块度增量^[51]后, 合并相应的社团 i 和 j . 当达到了局部最大模块度时, 此步骤停止. ② 将此时的每个社区继续当作一个“节点”, 重复步骤直到模块度停止变化. 节点 i 的模块度增益 ΔQ 为

$$\Delta Q = \left[\frac{\sum_{\text{in}} + k_{i,\text{in}}}{2m} - \left(\frac{\sum_{\text{tot}} k_i}{2m} \right)^2 \right] - \left[\frac{\sum_{\text{in}}}{2m} - \left(\frac{\sum_{\text{tot}}}{2m} \right)^2 - \left(\frac{k_i}{2m} \right)^2 \right], \quad (2)$$

其中 \sum_{in} 表示属于社团内的连边权重之和, \sum_{tot} 表示社团内的节点所产生的连边权重之和. k_i 是节点 i 的度, $k_{i,\text{in}}$ 表示社团中节点 i 和社团中的其他节点的连边的权重之和. m 表示整个网络中的边的权重之和. 在整个过程中, 假设所有边的权重都是 1. T-社团 (two-layer-comm) 保护策略, 就是优先保护模块度最大的节点, 使得当它们或者它们的依赖节点失效时, 这些节点依然可以正常运行, 从而增加了网络的鲁棒性.

3.3 优化层间耦合机制

不同的网络拓扑结构使多层网络的鲁棒性之间存在较大差异. Reis 等^[52] 针对随机连接的相依模型网络鲁棒性低, 而自然界中的真实相依网络鲁棒性却相对较高的问题进行研究, 发现多层网络的鲁棒性由每一层的内部结构和层间的节点连接模式共同决定. 他们指出一个网络的中心节点 (hub nodes) 与另一个网络中心节点之间存在度同配相关性的相依网络在随机故障中具有更好的鲁棒性. Parshani 等^[35] 提出了部分依赖模型, 通过降低相互依赖节点的数量, 降低级联失效的危害. Liu 等^[31] 提出的弱依赖模型则降低了相互依赖网络之间的依赖程度, 这也使得网络的破碎形式从一阶相变变成二阶相变, 增强了网络的鲁棒性. 文献^[53] 则提出多层网络间非对称性的依赖, 来控制网络层间的依赖程度, 从而增加网络的鲁棒性. Hu 等^[54] 分析了相依网络结构相似性对级联故障带来的影响, 他

们发现增加结构相似性会减弱级联故障的程度. 文献^[55] 也指出网络间的相似性越高, 则发生节点随机失效时系统的鲁棒性也越高. 文献^[56] 提出网络同配性研究可以为提高网络鲁棒性带来新的启发, 从而提高关键基础设施的保护水平. Radicchi 等^[57] 将具有相互依赖关系的节点在一个失效后其余节点也会失效这一规则更改为: 当一个节点有至少两个副本节点 (多层网络的不同层中具有相互依赖关系的节点) 存活时, 这个节点就不会失效. 在多层 (> 2) 网络中建立冗余的相互依赖关系可以提高整个系统的鲁棒性. 除此之外还可以通过减弱网络层间的相关性^[58]、设置加强节点^[59] 等策略抑制系统的级联效应.

3.4 节点备份

节点备份也是一个有效预防级联效应的措施, 它是指对多层网络中的少数重要节点预先从结构或功能方面设计并添加它们的备份, 万一这些节点日后失效, 这些备用节点能够立即启用代替失效的节点, 维持网络功能正常运行^[60]. Valdez 等^[61] 认为对一些节点备份后, 即使在缺少其他网络支持的情况下, 这些节点仍然能保留功能, 从而增加了系统的鲁棒性. Quattrociocchi 等^[62] 通过引入节点的自愈 (self-healing) 机制, 即增加网络固有的冗余度来增强网络的鲁棒性. Schneider 等^[37] 选择最少的自治节点^[63,64] (通过 k -shell, 介数中心性等方法进行筛选节点) 进行备份, 以避免网络在遭受攻击时发生突然瓦解和破碎.

然而节点备份的策略也有一些不足之处. 文献^[60] 提出在自修复网络中通过相互复制进行自我修复是一把“双刃剑”. 除此之外, 在现实应用中, 节点备份策略和网络冗余设计还需要增加网络的设计和维护成本 (有些情况下节点的备份还面临技术难题), 付出一定的时间和经济代价, 造成一定的浪费, 因此在实际应用中往往需要考虑这一策略的代价和效果的平衡.

4 恢复策略

恢复策略并不是重新设计或者构建一个网络, 而是在正发生级联失效的网络上同步进行补救和修复, 使其级联过程减缓甚至停止, 并逐步恢复原有功能的办法^[65]. Schneider 等^[66] 认为对于给定度分布的网络, 在抵御恶意攻击中最有效的网络结构

仍然是未知的;而对于给定连边数量的网络,鲁棒性最高的结构是所有节点度都相同的网络.他们在欧洲电力系统、互联网以及复杂网络模型上对此进行了仿真,结果表明,网络结构的很小变化(低成本)就可以显著提高不同网络的鲁棒性,并保持其功能不变.该研究结果不仅对提高现有基础设施的鲁棒性有重要意义,而且对设计经济可靠的网络系统也有一定的参考价值. Di Muro 等^[14]提出的通过寻找共同边界节点的恢复策略(详见 4.1 节)和 La Rocca 等^[67]在 2018 年提出的空闲连边策略(详见 4.2 节),其相同点都是从网络的巨分量入手来对抗级联效应.此外还有一些策略关注到了相依网络中节点的两种不同属性的边(相依边和相连边^[68,69]),他们认为对于来自相依网络的节点,它的重要性与其相依边和相连边的数量有关.文献^[68]研究了如何在合理分配有限成本的情况下来添加连接边和依赖边.文献^[69]则是通过衡量节点的相依边和相连边的数量,来确定优先恢复的节点.对于多层网络, Berezin 等^[70]发现局部攻击引起的危害比起同等情况下的随机攻击更加严重.文献^[71]对于局部攻击产生的故障提出在故障节点存活邻居中选择两个低度值的节点进行加边的修复方法.

4.1 边界节点恢复策略

Muro 等^[14]提出相依网络恢复策略,旨在对未被级联失效波及的剩余网络进行保护.这一策略使得级联失效过程和恢复过程动态交替进行,其核心是找到两个相依网络中的共同边界节点.共同边界节点是指两个网络中距离各自巨分支距离为 1 的一对失效的相互依赖节点.图 4 中节点 1 和 2 即为共同边界节点.初始网络 A 发生了故障,网络 B 中所对应的节点也会失效,在网络 B 将故障传递回网络 A 之前,恢复机制会介入并找出当前的共同边界节点,每轮恢复阶段以概率 γ 对共同边界节点进行恢复,从而尽可能地遏制级联失效在相依网络上的传播. Muro 等发现,最终网络有以下三种情况,第一是系统不被修复也不会崩溃;第二是部分节点在这一过程中失效,但恢复策略避免了系统的崩溃;第三是恢复过程也不能阻断级联失效过程,最终系统崩溃.

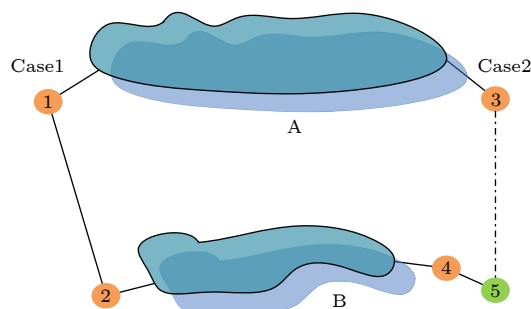


图 4 故障恢复策略图解^[14]. 网络 A 和网络 B 的巨分支如图所示. 情况 1: 两个通过相依边连接的失效节点(节点 1 和节点 2)分别距离其巨分支的距离 $l = 1$, 然后以恢复概率 γ 进行修复; 情况 2: 如果两个相互依赖的故障节点(节点 3 和节点 5)中至少有一个与其巨分支的距离大于 1, 则不符合恢复的条件, 所以放弃恢复这一对节点

Fig. 4. Illustration of failure recovery strategy^[14]. The giant components of network A and network B are shown in the figure. Case 1: Two failed nodes (nodes 1 and 2) connected by dependent edges are respectively at a distance of $l = 1$ from their maximal cluster, and then repaired with recovery probability γ . Case 2: If at least one of the two interdependent failed nodes (nodes 3 and 5) is more than 1 away from its maximal cluster, the recovery condition is not met, so the pair of nodes is abandoned to be restored.

选择共同边界节点进行恢复有以下两个原因: 第一, 当故障发生时, 通常都是优先抢修正常区域周边的基础设施; 第二, 如果候选恢复目标不是共同边界节点, 其对应的相依节点若是脱离巨分支的节点, 这个节点就会因其依赖节点的失效而失效, 那么对该节点的修复就没有意义.

吴佳键等^[69]对此策略进行了一些修改, 他们认为用恢复概率 γ 来随机选择恢复节点不是最优方案. 于是提出利用共同边界节点在巨分支内外的连接边数计算和定义边界节点的重要性, 也就是基于相连边的择优恢复算法 (preferential recovery based on connectivity link, PRCL). 实验显示, PRCL 算法的恢复策略更好, 可以识别出恢复过程中更重要的边界节点.

4.2 空闲连边恢复策略

在 Buldyrev 提出的相依网络级联失效模型的基础上, La Rocca 等^[67]2018 年提出在两个相依网络中对相较而言恢复代价更低的那个网络进行恢复的策略. 这里假设网络 B 为符合条件的网络, 在步骤 $n = 0$ 时, 从网络 A 中移走 $1 - p$ 比例的节点, 得到网络 A 的巨分支. 因为网络 A 与网络 B 的节点一对一依赖, 所以可以得到网络 B 此时的巨分

支. 以概率 γ 同时恢复网络 B 中某个有限簇 (其规模不小于 2) 中的两个节点与巨分支之间的连边. 但如果该有限簇只有单个节点, 那么就以相同方式恢复其与巨分支之间的一条连边. 需要注意的是所有可能被恢复的有限簇中的节点必须有空闲连边. 空闲连边是一种虚拟连边, 指的是那些在级联失效过程中断开的连边. 当一条连边断开以后, 则它两端的节点各自得到一条空闲连边, 如图 5(a) 中的虚线边所示.

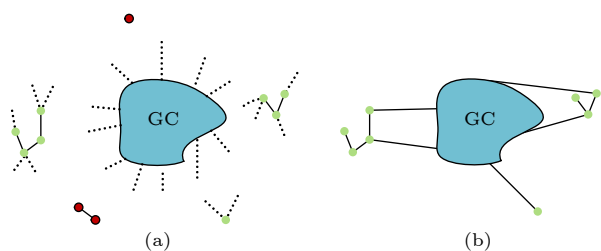


图 5 网络 B 中恢复策略的实现示意图^[67] (a) GC 表示网络巨分支, 虚线表示空闲连边, 带有空闲连边的簇表示可修复的簇, 没有空闲连边的簇表示无法进行恢复的簇; (b) 网络 B 完成重连后的巨分支

Fig. 5. Schematic diagram of the implementation of recovery strategy in network B^[67]: (a) GC represents the giant component of the network, the dashed lines indicate idle connected edges, clusters with free connected edges represent repairable clusters, and clusters without free connected edges represent clusters that cannot be recovered; (b) the giant component of network B after reconnection.

之所以选择一个有限簇里面的两个节点与巨分支相连接, 是为了减少这个有限簇再次脱离巨分支的概率. 增加有限簇与巨分支相连的节点数量虽然可以提高网络的鲁棒性, 但是在现实应用中也会增加成本. 阶段 $n = 0$ 结束后, 以概率 $1 - \gamma$ (γ 是恢复概率) 删掉网络 B 中没有被恢复的有限簇, 如图 5(a) 中没有空闲连边的深色节点. 至此第一轮网络 B 恢复过程结束. 当网络 B 将这个结果反馈给网络 A 的时候, 与网络 B 中失效节点相连接的网络 A 中的节点就会失效, 然后又反馈到网络 B, 与网络 A 中失效节点相依赖的网络 B 中的节点失效, 对此时的网络 B 开启新一轮的恢复过程. 以此类推, 两个网络就这样迭代下去, 直至两个网络构成的系统达到稳态.

La Rocca 等^[67] 指出随着 γ 的增加, 临界阈值 p_c 降低, 网络破碎的形式也会从一阶相变转变为二阶相变, 这就使得两个网络在发生网络崩溃前可以克服更多节点的失效. 这也是为什么 La Rocca 等

认为将此恢复策略应用在两个相依网络中较为脆弱的那个网络, 会使整个系统的抗毁性提高. 根据不同的恢复概率 γ 和网络的初始保留概率 p , 就可以知道网络是能被恢复的, 还是不能避免它最终的崩溃.

4.3 加边恢复策略

加边恢复策略是以相依网络级联失效模型^[15]为基础通过进行一系列的加边, 来增加网络鲁棒性的恢复策略^[72]. 过程如下: 在晶格网络 A 中, 对于一个被移除的节点, 将其邻居中没有被移除的具有功能性的两个节点以概率 w 相连. 网络 B 中与网络 A 中的失效节点具有依赖关系的节点也会失效, 因此网络 B 也需要实施上述的恢复过程, 即失效节点的两个未失效且不直连的邻居以概率 w 进行连接. 其实恢复步骤是将每个失效节点的所有邻居对当作候选者以概率 w 连接, 节点失效, 找出邻居对连接, 再有节点失效……, 整个过程待两个相依网络达到动态平衡, 即没有连边或者节点再失效后结束. 随着时间的推移, 修复连接可能会极大地改变拓扑结构, 新建立连接的两个节点之间在原始晶格网络上的距离可能越来越大.

4.4 优先恢复重要节点策略

Gong 等^[73] 提出在级联失效后, 优先恢复重要节点的策略. 通过在三个不同类型的耦合网络 (随机网络-随机网络 (ER-ER)^[74], 随机网络-无标度网络 (ER-SF)^[73], 电力网络-无标度网络 (power-SF)^[73] 上分别应用 6 种不同的重要节点识别指标 (随机、度中心性、介数中心性、PageRank、LeaderRank) 来确定优先恢复的节点, 结果发现只需恢复网络中 5% 的重要节点就可以显著恢复网络功能, 尤其是介数中心性指标效果最优. 而基于度中心性指标和 PageRank 指标的恢复策略的优点在于其较低的计算复杂度, 可用于具有数百万节点的大规模相依网络.

相依网络的恢复过程如下 (例如按照随机选择): 对于一个已经级联失效的网络, 初始时假设图中的节点均为失效节点. 如图 6(a) 所示, 假设恢复网络 C 中的节点 1, 2, 3 和 4, 那么网络 D 中与网络 C 相依赖的节点 5, 6, 7 和 8 被触发而恢复正常 (图 6(b)). 然后, 由于网络 C 中已恢复节点 4 和它在 D 网络中的依赖节点 8 不在各自网络的巨分

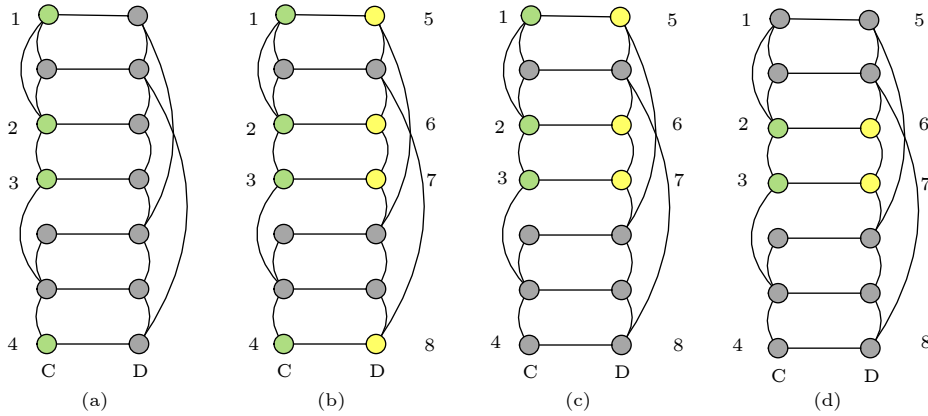


图 6 恢复模型

Fig. 6. Schematic diagram of recovery model.

支中, 所以它们会再次失效 (图 6(c)). 同理, 网络 D 中的节点 5 由于孤立失效, 而导致与其依赖的 C 网络中的节点 1 也再次失效 (图 6(d)). 此时网络达到稳态, 网络中的节点 2, 3, 6 和 7 就是本次恢复过程执行后最终被真正恢复的节点.

4.5 更改拓扑结构策略

Schneider 等 [66] 提出了一种有效恢复电力网络故障的方法, 他们发现在不增加连边数量的情况下, 只要对给定网络的拓扑结构进行相对较小的修改, 就有可能大大降低恶意攻击的危害. 这一结论在两个真实网络, 即欧洲电网和互联网中得到了验证. 这一发现, 一方面可以指导现有网络通过结构的优化来提升鲁棒性, 另一方面也可以用来设计未来的基础设施, 使其具有更好的鲁棒性. 我们通常是用临界阈值来衡量网络的鲁棒性. 而这种方法忽略了如果网络受到了攻击但是并没有崩溃的情况. 所以他们引进了一个独特的测量鲁棒性的方法,

$$R = \frac{1}{N} \sum_{Q=1}^N S(Q), \quad (3)$$

其中 N 是网络中的节点数, $S(Q)$ 是在删除了 Q 个节点之后网络巨分量中节点的数目.

4.6 局域攻击修复策略

局域攻击也叫局部攻击, 指的是网络位于某个地理空间范围内的节点受到了攻击. 在现实生活中, 局域攻击比随机攻击更为普遍, 如军事打击、自然和人为的灾害等 [75]. 文献 [71] 提出了优先最小度修复策略 (the healing strategy by prioritizing minimum degrees, HPMD), 空间相依网络出现局

部攻击时可以采用此策略. 失效网络的模型依托于文献 [76], 优先最小度策略是将一个失效节点的两个度值最低的邻居相连进行恢复, 对比度中心性、随机选择和局部中心性, 此方法更优.

4.7 自适应边恢复策略

Liu 等 [77] 提出在多层网络中添加自适应边的恢复策略. 为了增加多层网络的鲁棒性, 抵御大规模节点失效而导致的网络崩溃, 在多层网络中, 网络 A 定义为控制层网络, 而网络 B, C, ... 是非控制层 (不能人为干预), 当网络 A 中的节点 a_i 脱离其巨分支时, 我们规定节点 a_i 会随机产生 M 条边连接在网络 A 中的其他节点上, 也就是说产生的 M 条自适应边中只要有一条连接在了网络 A 的巨分支上, 节点 a_i 就会从失效状态恢复成具有正常功能的状态 (在此过程中假设与 a_i 相依赖的其他网络层中的节点均没有失效, 都具有正常功能). 根据经典的相依网络级联失效模型, 我们知道当 a_i 脱离巨分支时, 其他网络层中与之相互依赖的节点也要失效, 而在节点 a_i 脱离时, 产生的 M 条自适应边会很大程度上保证这个节点被修复, 这意味着其他层中与节点 a_i 具有依赖关系的节点会因为这个自适应边的加入而以很大概率避免了脱离其各自网络的巨分支. 所以对其中一个网络层的自适应扰动不仅可以增强控制层网络自身的鲁棒性, 还可以增强其他互连网络层的鲁棒性.

5 讨论

多层网络鲁棒性是当前复杂网络和复杂系统研究的核心问题之一. 基于渗流理论的研究发现由

于网络之间的联系和依赖, 多层网络往往是非常脆弱的. 这一结果为一些基础设施出现突发大规模级联失效给出了理论解释. 但是还有一些基础设施系统却非常稳定, 大规模的失效现象很少出现. 因此, 为了理解基础设施系统的鲁棒性和脆弱性, 有关多层网络鲁棒性第一个方面的研究是对具有不同耦合机制、拓扑结构的多层网络进行建模, 并研究这些因素对多层网络鲁棒性的影响以及网络在遭受攻击时破碎的机理. 第二个方面是如何设计有效的预防措施或节点恢复策略来降低级联失效对多层网络的损害. 这两个方面的研究相辅相成, 第一个方面的研究为第二方面的研究提供了基础理论和思路. 本文所介绍的多层网络级联失效的预防策略大多基于第一个方面的研究成果, 由渗流理论可知度值较小的节点很容易因网络中其他节点的删除而失效, 而度值较大的节点失效的时候会产生较大的破坏性, 因此, 当跨网络层的节点随机耦合时多层网络会比较脆弱, 这为调整耦合机制提供了重要思路. 此外, 节点的保护策略也同样基于对多层网络破碎机理的研究, 例如理论研究发现多层网络临界簇中的“基石节点”是至关重要的, 这为保护多层网络中的重要节点提供了重要思路.

多层网络级联失效的抑制策略研究同样建立在多层网络破碎机理的理解之上, 如边界节点恢复模型、加边恢复策略等. 无论是对节点的恢复, 还是对网络进行加边, 都需要一定的代价. 如何达到效用和代价的最优? 为什么需要恢复边界节点? 哪些节点需要优先加边恢复? 回答这些问题同样需要理解多层网络的破碎规律和特点. 对于不同的多层网络发生级联失效的时候, 该采用什么样的决策方法来选用恢复策略呢? 目前来说, 回答这一问题尚有比较大的挑战, 但可以肯定的是恢复策略的选用需要结合具体的情况, 如多层网络拓扑结构特性、耦合机制、网络损害规模和实际需求, 以及恢复的代价限制和速度要求等. 随着对多层网络级联效应研究的深入, 相信对这一问题的研究会不断取得突破, 而且会有更多的、更加贴合现实情景的预防和恢复策略被提出.

参考文献

- [1] Eubank S, Guclu H, Kumar V A, Marathe M V, Srinivasan A, Toroczkai Z, Wang N J N 2004 *Nature* **429** 180
- [2] Keeling M J, Eames K T 2005 *J. R. Soc. Interface* **2** 295
- [3] Pecora L M, Carroll T L 1990 *Phys. Rev. Lett.* **64** 821
- [4] Yu W, Chen G, Lü J J A 2009 *Automatica* **45** 429
- [5] Qi X, Yang G, Liu L 2020 *Physica A* **539** 122870
- [6] Liu Y Y, Slotine J J, Barabási A L 2011 *Nature* **473** 167
- [7] Wang X F, Chen G 2002 *Physica A* **310** 521
- [8] Rueda D F, Calle E 2017 *Int. J. Crit. Infrastruct. Prot.* **16** 3
- [9] Rinaldi S M, Peerenboom J P, Kelly T K 2001 *IEEE Control Syst. Mag.* **21** 11
- [10] Buldyrev S V, Parshani R, Paul G, Stanley H E, Havlin S 2010 *Nature* **464** 1025
- [11] Tootaghaj D Z, Bartolini N, Khamfroush H, La Porta T 2007 *IEEE 36th Symposium on Reliable Distributed Systems (SRDS)* pp54–63
- [12] Cui C C <http://mini.eastday.com/a/180419160201052.html> [2018-04-19] (in chinese) (in Chinese) [崔聪聪 <http://mini.eastday.com/a/180419160201052.html> [2018-04-19]]
- [13] China Daily http://www.xinhuanet.com/world/2015-07/24/c_128056543.htm [2015-07-24] (in Chinese) [中国日报网 http://www.xinhuanet.com/world/2015-07/24/c_128056543.htm [2015-07-24]]
- [14] Di Muro M A, La Rocca C E, Stanley H, Havlin S, Braunstein L A 2016 *Sci. Rep.* **6** 22834
- [15] Gao J, Buldyrev S V, Stanley H E, Havlin S 2012 *Nat. Phys.* **8** 40
- [16] Gao J, Li D, Havlin S 2014 *Natl. Sci. Rev.* **1** 346
- [17] Gong M, Wang Y, Wang S, Liu W 2017 *Sci. Rep.* **7** 12753
- [18] Baxter G, Dorogovtsev S, Goltsev A, Mendes J 2012 *Phys. Rev. Lett.* **109** 248701
- [19] Faqeeh A, Melnik S, Colomer-de-Simón P, Gleeson J P 2016 *Phys. Rev. E* **93** 062308
- [20] Murakami M, Ishikura S, Kominami D 2017 *Appl. Netw. Sci.* **2** 6
- [21] Malgorzata T, Keith B, Martin R, Ananthram S, Raissa M D 2019 *Phys. Rev. E* **99** 032308
- [22] Shekhtman L M, Berezin Y, Danziger M M, Havlin S 2014 *Phys. Rev. E* **90** 012809
- [23] Zhao J, Li D, Sanhedrai H, Cohen R, Havlin S 2016 *Nat. Commun.* **7** 10094
- [24] Dorogovtsev S N, Mendes J F F, Samukhin A N 2001 *Phys. Rev. E* **64** 025101
- [25] Liu X, Stanley H E, Gao J 2016 *Proc. Natl. Acad. Sci.* **113** 1138
- [26] Azimi Tafreshi N, Dorogovtsev S N, Mendes J F 2014 *Phys. Rev. E* **90** 052809
- [27] van der Hoorn P, Litvak N 2015 *Phys. Rev. E* **92** 022803
- [28] Klimek P, Thurner S, Hanel R 2009 *J. Theor. Biol.* **256** 142
- [29] Baxter G J, Dorogovtsev S N, Goltsev A V, Mendes J F 2010 *Phys. Rev. E* **82** 011103
- [30] Parisi G, Sallitto M 2015 *EPL* **109** 36001
- [31] Liu R R, Eisenberg D A, Seager T P, Lai Y C 2018 *Sci. Rep.* **8** 2111
- [32] Albert R, Jeong H, Barabási A L 2000 *Nature* **406** 378
- [33] Lü L, Chen D, Ren X L, Zhang Q M, Zhang Y C, Zhou T 2016 *Phys. Rep.* **650** 1
- [34] Fan T, Lü L, Shi D, Zhou T 2020 arXiv: 2001.08541 [physics.soc-ph]
- [35] Parshani R, Buldyrev S V, Havlin S 2010 *Phys. Rev. Lett.* **105** 048701
- [36] Fan T L, Zhu Y Y, Wu L L, Ren X L, Lü L Y 2017 *JEST* **46** 766 (in Chinese) [范天龙, 朱燕燕, 吴蕾蕾, 任晓龙, 吕琳媛 2017 *电子科技大学学报* **46** 766]
- [37] Schneider C M, Yazdani N, Araújo N A, Havlin S, Herrmann H J 2013 *Sci. Rep.* **3** 1969

- [38] Huang X, Gao J, Buldyrev S V, Havlin S, Stanley H E 2011 *Phys. Rev. E* **83** 065101
- [39] Barabási A L, Albert R 1999 *Science* **286** 509
- [40] Du R, Dong G, Tian L, Liu R 2016 *Physica A* **450** 687
- [41] Osat S, Faqeeh A, Radicchi F 2017 *Nat. Commun.* **8** 1540
- [42] De Domenico M, Solé-Ribalta A, Omodei E, Gómez S, Arenas A 2015 *Nat. Commun.* **6** 6868
- [43] Bonacich P 1972 *J. Math. Sociol.* **2** 113
- [44] Freeman L C 1978 *Soc. Networks* **1** 215
- [45] Freeman L C 1977 *Soc. Networks* **40** 35
- [46] Kitsak M, Gallos L K, Havlin S, Liljeros F, Muchnik L, Stanley H E, Makse H A 2010 *Nat. Phys.* **6** 888
- [47] Chen D, Lü L, Shang M S, Zhang Y C, Zhou T 2012 *Physica A* **391** 1777
- [48] Brin S, Page L 1998 *Comput. Networks* **30** 107
- [49] Lü L, Zhang Y C, Yeung C H, Zhou T 2011 *PLoS one* **6** e21202
- [50] Blondel V D, Guillaume J L, Lambiotte R, Lefebvre E 2008 *J. Stat. Mech.:Theory Exp.* **2008** P10008
- [51] Dugué N, Perez A 2015 *HAL* Id: hal-01231784
- [52] Reis S D, Hu Y, Babino A, Andrade Jr J S, Canals S, Sigman M, Makse H A 2014 *Nat. Phys.* **10** 762
- [53] Liu R R, Jia C X, Lai Y C 2019 *Phys. Rev. E* **100** 052306
- [54] Hu Y, Zhou D, Zhang R, Han Z, Rozenblat C, Havlin S 2013 *Phys. Rev. E* **88** 052805
- [55] Parshani R, Rozenblat C, Ietri D, Ducruet C, Havlin S 2011 *EPL* **92** 68002
- [56] Zhou D, Stanley H E, D'Agostino G, Scala A 2012 *Phys. Rev. E* **86** 066103
- [57] Radicchi F, Bianconi G 2017 *Phys. Rev. X* **7** 011013
- [58] Min B, Do Yi S, Lee K M, Goh K I 2014 *Phys. Rev. E* **89** 042811
- [59] Yuan X, Hu Y, Stanley H E, Havlin S 2017 *Proc. Natl. Acad. Sci.* **114** 3311
- [60] Ishida Y 2005 *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems* Melbourne, VIC, Australia, September 14–16, 2005 p86
- [61] Valdez L D, Macri P A, Braunstein L 2014 *J. Phys. A: Math. Theor.* **47** 055002
- [62] Quattrociochi W, Caldarelli G, Scala A 2014 *Plos one* **9** e87986
- [63] Nair D T, Malhotra M 2011 arXiv:1107.1956 v1 [cs.IR]
- [64] Mitchell J C, Teague V 2002 *International Symposium on Software Security* Nara, Japan, October 3–4, 2002 p58
- [65] Ishida Y, Mori T 2005 *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems* Melbourne, VIC, Australia, September 14–16, 2005 p79
- [66] Schneider C M, Moreira A A, Andrade J S, Havlin S, Herrmann H J 2011 *Proc. Natl. Acad. Sci.* **108** 3838
- [67] La Rocca C E, Stanley H E, Braunstein L A 2018 *Physica A* **508** 577
- [68] Cui P, Zhu P, Wang K, Xun P, Xia Z 2018 *Physica A* **497** 185
- [69] Wu J J, Gong K, Wang C, Wang L 2018 *Acta Phys. Sin.* **67** 88901 (in Chinese) [吴佳键, 龚凯, 王聪, 王磊 2018 物理学报 **67** 88901]
- [70] Berezin Y, Bashan A, Danziger M M, Li D, Havlin S 2015 *Sci. Rep.* **5** 8934
- [71] Gong K, Wu J J, Liu Y, Li Q, Liu R R, Tang M 2019 *Complexity* **2019** 10
- [72] Stippinger M, Kertész J J 2014 *Physica A* **416** 481
- [73] Gong M, Ma L, Cai Q, Jiao L 2015 *Sci. Rep.* **5** 8439
- [74] Erdős P, Rényi A 1959 *Publ. Math. Debrecen* **4** 3286
- [75] Shao S, Huang X, Stanley H E, Havlin S 2015 *New J. Phys.* **17** 023049
- [76] Li W, Bashan A, Buldyrev S V, Stanley H E, Havlin S 2012 *Phys. Rev. Lett.* **108** 228702
- [77] Liu R R, Jia C X, Lai Y C 2019 *New J. Phys.* **21** 045002

SPECIAL TOPIC—Statistical physics and complex systems

Overview of precaution and recovery strategies for cascading failures in multilayer networks^{*}

Jiang Wen-Jun¹⁾ Liu Run-Ran^{1)†} Fan Tian-Long²⁾¹⁾Liu Shuang-Shuang¹⁾ Lü Lin-Yuan^{2)1)‡}¹⁾ (*Alibaba Research Center for Complexity Sciences, Hangzhou Normal University, Hangzhou 311121, China*)²⁾ (*Institute of Fundamental and Frontier Science, University of Electronic Science and Technology of China, Chengdu 611731, China*)

(Received 31 December 2019; revised manuscript received 1 February 2020)

Abstract

In real life, most of the infrastructure networks closely related to the national economy and people's livelihood do not exist independently, but are interconnected with or dependent on each other, so the multilayer network model is proposed to study the independent complex systems and infrastructures. When the nodes in the multilayer network suffer initial failure or attack, the cascade occurs due to the interaction between the “intra-layer” and “inter-layer”, and the failure can propagate in the network layer and across the layers iteratively, so that the scale of the failures is enlarged gradually. As a result, many multilayer networks are more fragile than single networks. The cascading failure of multilayer network usually brings very serious catastrophes to our society. So, conducting the research on preventing the multilayer network from cascading failure and recovering is of great significance. As far as the prevention of cascading failure is concerned, what are mainly included are the strategies such as the fault detection, the protection of important nodes, the optimization of the coupling method of networks, and the backup of nodes. As for the recovery of multi-layer network, included mainly are the strategies such as common boundary node recovery, the idle connected link recovery, the link addition, the priority recovery of important nodes, the topology perturbation, and the repairing of localized attack and adaptive link.

Keywords: complex network, cascade failure, precaution strategy, recovery strategy**PACS:** 89.75.-k, 89.90.+n**DOI:** 10.7498/aps.69.20192000

^{*} Project supported by the National Natural Science Foundation of China (Grant Nos. 61773148, 61673150, 11622538) and the Research Project “The Life Journey of Innovation and Research”, China (Grant No. 2018YCZD09).

[†] Corresponding author. E-mail: runranliu@163.com

[‡] Corresponding author. E-mail: linyuan.lv@uestc.edu.cn

相对性区域创新指数与经济周期挖掘

方学进 崔俊英 胡淡淡 韩筱璞

Relativistic regional innovation index and novel business cycle

Fang Xue-Jin Cui Jun-Ying Hu Dan-Dan Han Xiao-Pu

引用信息 Citation: *Acta Physica Sinica*, 69, 088905 (2020) DOI: 10.7498/aps.69.20191970

在线阅读 View online: <https://doi.org/10.7498/aps.69.20191970>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

复杂系统重构

Complex system reconstruction

物理学报. 2020, 69(8): 088906 <https://doi.org/10.7498/aps.69.20200001>

周期波动性对金融市场稳定性的影响

Influence of periodic volatility on the stability of financial market

物理学报. 2017, 66(4): 040501 <https://doi.org/10.7498/aps.66.040501>

基于近似熵的斯隆数字化巡天中类星体光变复杂性分析

Analysis on complexity of optical variability based on approximate entropy in Sloan digital sky survey quasars

物理学报. 2019, 68(14): 149801 <https://doi.org/10.7498/aps.68.20182071>

不同硅晶面指数上的类倒金字塔结构研究与分析

Different silicon crystal face index of inverted pyramid structure

物理学报. 2018, 67(22): 226801 <https://doi.org/10.7498/aps.67.20181275>

圆锥料仓内颗粒周期性脉动特征研究

Cyclical pulsation properties of particles in cone silo

物理学报. 2017, 66(1): 014501 <https://doi.org/10.7498/aps.66.014501>

专题：统计物理和复杂系统

相对性区域创新指数与经济周期挖掘*

方学进 崔俊英 胡淡淡 韩筱璞†

(杭州师范大学, 阿里巴巴复杂科学研究中心, 杭州 311121)

(2019年12月16日收到; 2020年3月16日收到修改稿)

提出了一类新的相对性区域创新指数, 并采用世界专利申请数据对其进行了具体计算. 基于区域创新同经济发展水平之间的超线性关系, 该指数消除了经济发展水平对创新能力的影响, 可以实现对不同发展水平的经济体之间进行有效的创新能力横纵对比. 该创新指数尽管极其简单, 却揭示出一系列迥异于传统认知的现象, 例如中国大陆地区的技术创新能力在1980年代就已经位居世界前列. 采用该指数, 不但可以在较高水平上解释世界各国的经济增长, 还发现它同经济增长率之间的相关性存在一个20年的经济周期. 这些结果显示, 该指数作为一个单一性指标, 以极小的数据依赖就实现了较高程度的解释性, 不但重新定位了世界各经济体的创新能力, 对深入理解创新同经济发展之间的关系提供了新的角度, 而且暗示着这类相对性经济指标的发展潜力与应用空间.

关键词: 区域创新指数, 经济周期, 相对性经济指标, 经济复杂性**PACS:** 89.65.Gh, 89.75.-k, 89.65.-s**DOI:** 10.7498/aps.69.20191970

1 引言

创新是社会经济发展的关键驱动力之一. 对于一个经济体, 其社会创新能力正日益成为获取经济竞争优势的决定性因素, 对其经济发展有着长期而深刻的影响^[1-3]. 为了能够准确地评估一个经济体的社会创新能力, 研究者充分挖掘了各类社会创新相关因素, 从不同视角构建了大量的评价指标, 并基于这些指标来进行经济形态与产业政策等方面的分析^[4-9]. 一般而言, 根据指标所依据的因素结构, 这些区域创新指标通常可分为单一性指标和综合性指标两大类. 单一性指标一般是用来衡量有关社会创新的某一单项维度的发展程度, 例如研究与发展类(R&D)指标^[10]、专利类指标^[11,12]、科技论文类指标^[13,14]、区域创新效率指标^[15]. 例如自然指数(nature index), 它依据各国的学术组织在82种

顶级学术期刊上的论文发表信息所构建, 是一个典型的科技论文类单一性指标, 在衡量基础科研产出方面有着重要价值^[16,17]. 综合性指标往往包含了涉及多个维度的衡量项目, 可以从多个角度对地区创新水平进行综合性的判断. 例如, 由四川省社会科学院和中国科学院成都文献情报中心发布的中国区域创新指数, 所包含的评价维度涉及创新环境、创新投入和创新产出三个方面^[18]; 又如由世界知识产权组织所发布的全球创新指数(global innovation index, GII), 综合了80项指标, 可以对各国的创新能力、创新环境、创新效率等多个方面进行相当全面的评估和比较^[8,9].

需要指出的是, 以上这些区域创新指标中, 绝大多数属于绝对性指标, 即它们所衡量的是一个地区在社会创新方面的产出或投入的绝对量, 或是人均绝对量, 或绝对性的投入产出比. 这种绝对性指标虽然可以为度量社会创新相关的各类因素提供

* 国家自然科学基金(批准号: 61873081, 11622538, 61673150, 61673151)、浙江省自然科学基金(批准号: LGF18F030007, LR16A050001)和浙江省新苗人才计划项目(批准号: 2019R426070)资助的课题.

† 通信作者. E-mail: xp@hznu.edu.cn

直接计量,但是它们一般不能直接反映各个经济体之间巨大的发展不平衡所带来的结构性差异.换句话说,对于一个经济体,在不同的经济发展水平与外部环境下,往往有着与之相适应的不同的经济结构、创新水平等;而使用绝对性指标对处在不同发展水平的经济体进行直接对比,则会掩盖这种依赖于经济发展程度的结构性差异.

为了克服绝对性指标的这一缺陷,一种思路是,首先为处在不同经济发展水平下的各个经济体分别寻找到各自的参照,用各个经济体相对它们各自的参照的差异作为指标的构建基础.这一类指标则可称之为相对性指标.本文将基于这一思路,采用世界专利申请数据,提出一种新型的相对性指标来衡量各经济体的技术创新能力,并通过分析该指标同世界经济增长之间的关系来论证其有效性.

2 数据来源

本文所使用的数据是自1980年至2017年世界各经济体在各年份的人均国民生产总值(GDP)、专利申请数和人口数,以及世界总GDP、世界专利申请总数和世界总人口,均从世界银行网站(<https://www.worldbank.org/>)获取得到.该数据共包含264个国家、地区和组织,其中国家实体或地区共217个.在这217个经济体中,同时包含人均GDP信息和专利申请信息者,共148个.由于1985年前的各国信息缺失严重,因此在实际分析中我们使用的是1985年至2017年的数据.这148个经济体自1985年后的GDP和专利申请数就是我们构建区域创新指数的数据支撑.由于在这期间仍然有部分经济体的部分年份信息缺失,在实际计算中涉及每一年份时,所统计的范围是这148个经济体中所有具有该年份信息者.

3 区域创新指数的构建

在提出本文所述的区域创新指数之前,首先观察世界各国的专利申请数同其经济发展水平的关系.我们采用人均GDP作为各国经济发展水平的测度.由于汇率的变化,不同年份的人均GDP之间进行直接对比往往会引入较大误差,因此,在我们的后续分析中所采用的是“相对人均GDP”,其定义是:一个经济体在某一年份的相对人均GDP,

是该经济体的人均GDP相对该年份世界人均GDP的比值.其计算公式为

$$G_Y^i = \frac{\text{GDP}_Y^i}{\langle \text{GDP}_Y \rangle}, \quad (1)$$

其中 G_Y^i 为经济体*i*的在年份*Y*的相对人均GDP, GDP_Y^i 为经济体*i*在年份*Y*的人均GDP,尖括号表示在年份*Y*的世界人均GDP均值.我们也采用了“相对人均专利申请数”作为专利申请的主要分析对象,即该年份该经济体的人均专利申请数与世界人均专利申请数的比值:

$$n_Y^i = \frac{N_Y^i}{\langle N_Y \rangle}, \quad (2)$$

其中 n_Y^i 为经济体*i*在年份*Y*的相对人均专利申请数, N_Y^i 为经济体*i*在年份*Y*的人均专利申请数,尖括号表示在年份*Y*的人均专利申请数的世界均值.通过这两个相对性的人均指标,世界各经济体在不同年份的人均GDP和人均专利数都被纳入了一个可有效比较的范畴中.

图1显示了各个经济体自1985年以来在由相对人均GDP的对数值 g (例如对经济体*i*,有 $g^i = \log_{10}(G^i)$)和相对人均专利申请数的对数值 n_p (对经济体*i*,有 $n_p^i = \log_{10}(n^i)$)所构成的空间中的变化曲线.可以看到,发达经济体同发展中经济体之间有着较为明显的分离,其分离区域大致如图1点划线所示.中国大陆地区的曲线在 g 和 n_p 两者都呈现出稳定而快速的增长趋势,而韩国是极少数已经实现了从发展中经济体区域向发达经济体区域大幅度跨越的经济体之一.

进一步,通过对全部148个经济体的所有数据点进行线性回归,得到拟合直线,如图1中虚线所示,其斜率为1.12.该拟合直线所表征的是,在不同的相对经济水平下,世界平均意义上的相对人均专利申请数.其斜率大于1,表示相对人均专利申请数的增长快于相对人均GDP的增长,满足近似为幂次形式的超线性关系^[19-21].该超线性关系说明一个经济体的经济发展水平是决定其创新水平的最核心因素.因此,如果采用建立在绝对数量意义上的创新指标,对处在不同经济发展水平的经济体的创新能力进行直接比较,那么发达经济体的指标往往会明显占据相当大的优势,从而可能掩盖一些发展中经济体内部所隐藏的有利于社会创新的因素或效应.因此,为了有效地进行这一跨越经济发展程度的比较,首先需要排除其经济发展水平的影

响,为不同经济发展水平的经济体寻找到各自的创新水平基准.图1中的拟合直线,已经给出了不同的相对人均GDP水平下的相对人均专利申请数量的期望值,它所反映的是在不同经济发展水平下一个经济体在世界平均意义上的创新能力.或者说,假如一个经济体的创新能力始终保持在其当前经济发展水平下的平均值,那么伴随其经济发展,它在图1中的轨迹将沿着这条拟合直线增长.因此,这条拟合直线实际上给出了创新能力在同等经济发展水平的经济体中处于中游的经济体的相对人均专利申请数.

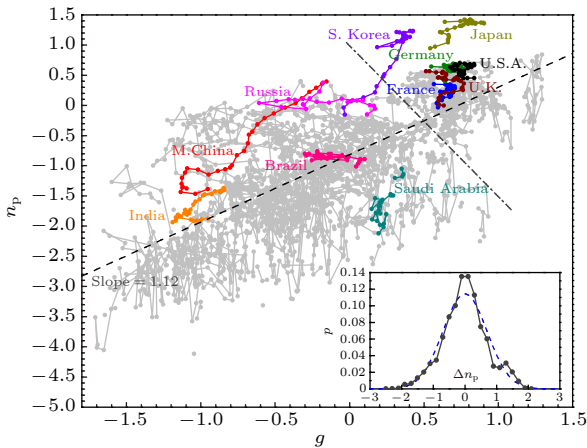


图1 世界各经济体在由相对人均GDP的对数 g 和相对人均专利申请数的对数 n_p (两者均以10为底数) 所构成的空间中的变化轨迹.彩色曲线为11个代表性经济体在该空间的轨迹,灰色曲线为其他经济体.灰色虚线为拟合直线 $n_p = 1.12g - 0.82$.灰色点划线大致区分了发达经济体的轨迹所在区域和发展中经济体所在区域,右上方主要为发达经济体,左下方主要为发展中经济体.插图显示了各个数据点相对拟合直线的离差 Δn_p 的概率分布,蓝色虚线为其高斯函数拟合

Fig. 1. The trajectories of economies from 1985 to 2017 in the space of the logarithmic relative GDP per capita (g) and the logarithmic relative number of patent applications per capita (n_p). The colored curves and gray curves represent the trajectory of 11 representative economies and the remain economies, respectively. The gray dashed line is the fitting function $n_p = 1.12g - 0.82$ of all data points. The gray dot dash line roughly distinguishes between the developed economies and the developing economies. Developed economies are mainly in the upper right area, while developing economies are in the lower left. The inset plots the distribution of the deviation Δn_p of each data point from the fitting line, in which the blue dashed line is its Gaussian fitting.

我们把该拟合直线作为各经济体进行对比的基准,把各个数据点相对该拟合直线的离差 Δn_p 作为不同经济发展水平的经济体之间进行创

新能力对比的依据,从而基本消除了经济发展水平的不同所带来的影响.我们观察了各个数据点相对该拟合直线的离差 Δn_p 的分布.如图1插图所示,该离差分布近似为期望值为0的正态分布.

我们基于该离差定义出一个经济体在某一年份的区域创新指数 I ,其计算方法如下:在图1中,若某经济体(例如经济体 i)在某个年份所对应的数据点坐标为 (g^i, n_p^i) ,它相对拟合直线的离差为 $\Delta n_p^i = n_p^i - h(g^i)$,其中 $h(g^i) = 1.12g^i - 0.82$ 为拟合直线方程,则该经济体在该年份的区域创新指数为 $I^i = P(\Delta n_p < \Delta n_p^i)$,即在所有经济体所有年份的数据点的离差值中,小于该数据点的离差的比例.需要注意,这一比例的计算并非进行同等经济发展水平的经济体之间的对比,而是涉及所有纳入统计的经济体和所有纳入统计的年份,也就是说图1中的所有数据点.这是因为,经济发展水平的不同所带来的影响已经基本通过拟合直线所剔除,不同经济发展水平的经济体相对拟合线的离差之间已经具备了可比较性.这种基于离差的比较,其意义在于:假如某个发展中经济体具有较高的正向离差,则说明它的绝对意义的技术创新能力相对同等经济水平的经济体已经处在较高位置;而且,如果它保持其离差不变,沿着平行于拟合直线的方向发展,那么当它经济发展达到发达经济体水平时,其绝对意义上的技术创新能力也将超过大部分发达经济体.

区域创新指数 I 是一个介于0到1的值,其表示的意义是:一个经济体的相对人均专利申请数,相对其当前经济发展水平下的期望数的偏离量的累积概率.如果该指数为0.5,表示其相对专利申请数量等于其经济发展水平下的期望数量,处在世界平均水平;如果该指数趋向于1,则表示该国的相对人均专利申请数远远高于期望数量,意味着极高的创新水平;若该指数趋向于0,则说明其相对人均专利申请数量严重低于期望.这里需要强调,该指数是建立在相对意义上的区域创新指数,它的关注点并非一个经济体的创新的绝对量或者人均量,而是关注这个经济体的创新能否超出其所处的经济发展水平下的世界期望.由于经济发展水平是决定区域创新能力的最主要因素,如果某经济体的创新能够获得超过其经济发展水平下的世界期望,则说明该经济体极可能存在某种能够有效促进创新的机制或环境.因此,该指数实际上反映的是经

济体中那些有利于技术创新的结构性因素的影响强弱.

图 2 显示了各经济体的区域创新指数 I 随年份的变化. 我们重点观察了 11 个代表性经济体, 包括美国、日本、德国、英国、法国等发达经济体, 中国、俄罗斯、巴西、印度等“金砖国家”, “亚洲四小龙”之一的韩国, 以及石油输出国沙特阿拉伯. 可以看出, 在 20 世纪 80 年代中国大陆地区的 I 值就已经超过 0.8, 位居世界前列, 而自 2016 年起更是高居世界首位, 意味着中国大陆地区事实上早已经步入创新型国家行列.

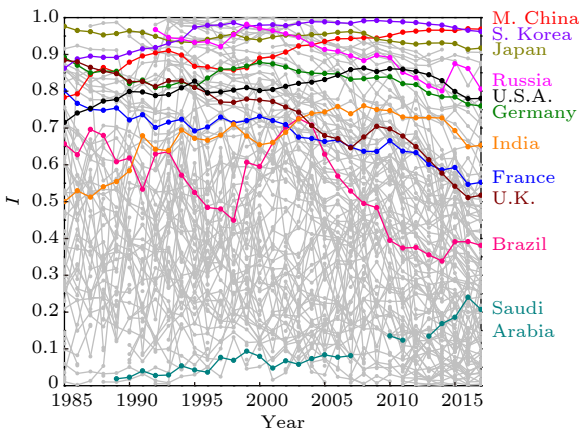


图 2 各经济体的区域创新指数 I 随年份的变化, 彩色线为 11 个代表性经济体, 灰色线为其他经济体

Fig. 2. The change of the regional innovation index I of each economy with the years. The colored curves and gray curves represent 11 representative economies and the remain economies, respectively.

除中国大陆地区之外, 日本和韩国也一直保持很高的创新水平. 德国和美国等国则位居第二梯队, 也一直稳定保持在一个较高的水平, 其 I 值起伏于 0.8 附近. 然而, 虽然同为发达经济体, 英国和法国的 I 值则处于持续下降中, 2016 年后已经下降到 0.5—0.6 附近, 仅居世界中等水平, 甚至低于印度.

从图 2 中也可以看出, 虽然和中国同为“金砖”国家, 但是印度和巴西的创新指数仅仅处在中等水平, 其中印度的 I 值在 0.7 上下浮动, 巴西则在 0.4—0.7 之间剧烈波动. 而以俄罗斯为代表的原苏联国家, 在 20 世纪 90 年代其 I 值普遍处于较高水平, 但在随后的 20 多年中持续性下降. 沙特阿拉伯的 I 值则长期处于很低的水平, 但在最近十年中出现了明显的提升.

通过对各代表性经济体的趋势的分析可以看出, 该区域创新指数 I 所显示的各经济体的技术创新能力截然不同于基于各类绝对性创新指标所构建的传统认知^[22]. 以中国大陆地区为代表的一批发展中经济体, 在绝对性创新指标之下一般并不处于领先地位, 而在该指数下则水平很高. 同时, 以英国、法国为代表的一些在传统认知中有着高度创新能力的发达国家, 在该指标下却仅仅处在世界中游. 附表 A1 中完整显示了各个经济体在 2016 年的指数 I 值. 而图 3 则对比了各个经济体的指数 I 值与该经济体的全球创新指数 (GII)^[8,9]. 不难发现 GI 作为代表性的绝对性区域创新指数, 它同经济发展水平的密切依赖; 而指数 I 则同 GI 的结果大相径庭, 同经济发展水平基本无关.

4 区域创新同经济增长的关系中的周期性

由于观察到一些长期保持了经济高速发展的经济体其指数 I 往往较高, 我们进一步通过挖掘该指数 I 同经济增长的关系, 来验证这一指数的有效性.

考虑到创新对经济发展的影响往往具有长期性, 我们首先观察了各经济体的指数 I 的 20 年均值 (从 1998 年到 2017 年), 同该时间段的相对人均 GDP 的平均增长率 $\langle \Delta g \rangle$ 的相关性. 这里 (也包括此后的全部计算), 对某经济体 i , 在从年份 Y 开始的 m 年的期间, 相对人均 GDP 的 m 年平均增长率 $\langle \Delta g \rangle^i$ 的计算方式是

$$\langle \Delta g \rangle_{Y \rightarrow Y+m-1}^i = \sqrt[m]{\frac{G_{Y+m-1}^i}{G_Y^i}} - 1, \quad (3)$$

其中 G_{Y-1}^i 是该经济体在年份 $Y-1$ 的相对人均 GDP 值, 同样 G_{Y+m-1}^i 对应年份 $Y+m-1$. 从 1998 年到 2017 年的 20 年期间, 各经济体 I 的 20 年均值 $\langle I \rangle$ 同各经济体相对人均 GDP 的 20 年平均增长率 $\langle \Delta g \rangle$ 的相关性如图 4 所示. 其中, 直接计算 $\langle I \rangle$ 同 $\langle \Delta g \rangle$ 之间的相关性, 所得的 Pearson 相关性系数 r 为 0.220 (图 4(a) 的插入图, 图中 P 为相关性的显著性值); 进一步计算 $\langle I \rangle^\beta$ 同 $\langle \Delta g \rangle$ 之间的相关性, 所得的最高 Pearson 相关性系数为 0.294, 此时对应指数 β 值为 3.80, 如图 4(a) 所示.

在此基础上, 由于经济增长率同时同经济发展水平本身存在依赖性, 我们进一步通过各经济体的

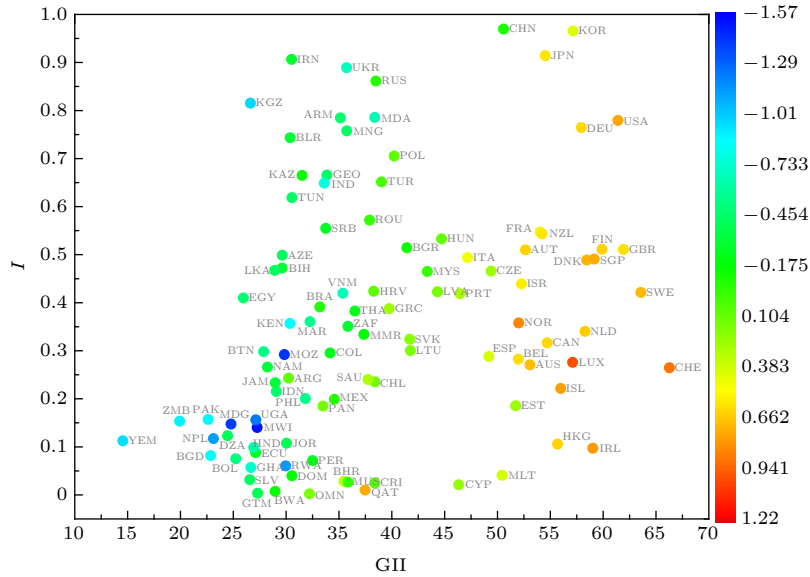


图 3 各经济体在 2016 年的区域创新指数 I 与该年份的全球创新指数 (GII) 的关系. 数据点的颜色表示该年份各经济体的相对人均 GDP 的对数值 g

Fig. 3. The regional innovation index I vs. global innovation index (GII) for each economy at 2016. The color of each data point shows the logarithmic relative GDP per capita (g) of each economy.

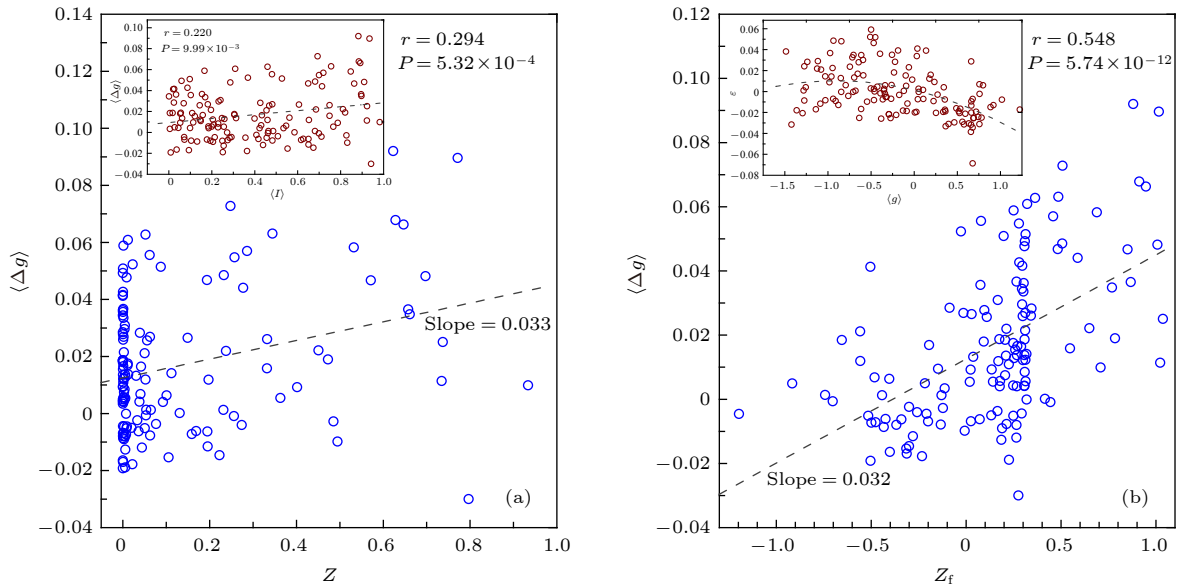


图 4 20 年时间段 (1998 年至 2017 年) 内各经济体平均区域创新指数 $\langle I \rangle$ 与相对人均 GDP 的平均增长率 $\langle \Delta g \rangle$ 之间的相关性 (a) $Z = \langle I \rangle^\beta$, 其中 $\beta = 3.80$ 为相关性最强时所对应 β 值, 直线为拟合直线; 插图显示为 $\langle I \rangle$ 同 $\langle \Delta g \rangle$ 之间的相关性 (即设定 $\beta = 1.0$ 时); (b) 通过相对人均 GDP 进行修正后的相关性, $Z_f = \langle I \rangle^\beta + (a \langle g \rangle^2 + b \langle g \rangle + c)/k$, 其中 $\langle g \rangle$ 为各经济体的相对人均 GDP 的对数值 g 的 20 年均值, 其中 $\beta = 3.80$, $a = -0.011$, $b = -0.020$, $c = 0.0013$, 而 $k = 0.033$ 为图 (a) 的拟合直线斜率; 插图显示了修正函数 $f(\langle g \rangle) = a \langle g \rangle^2 + b \langle g \rangle + c$ 的获得, 即对图 (a) 的回归残差 ε 同 $\langle g \rangle$ 的关系进行拟合所得

Fig. 4. The correlations between the average regional innovation index of each country $\langle I \rangle$ and the average growth rate of relative per capita GDP $\langle \Delta g \rangle$ in the period from 1998 to 2017: (a) $Z = \langle I \rangle^\beta$, where $\beta = 3.80$ corresponding to the strongest correlation between $\langle \Delta g \rangle$ and Z , and the dashed line is the fitting line. The inset of panel (a) shows the correlation between $\langle I \rangle$ and $\langle \Delta g \rangle$ (setting $\beta = 1.0$); (b) the correlation between $\langle \Delta g \rangle$ and the corrected prediction value Z_f of each economy, where $Z_f = \langle I \rangle^\beta + (a \langle g \rangle^2 + b \langle g \rangle + c)/k$, $\langle g \rangle$ is the 20-year average of the logarithmic relative GDP per capita g of each economy, and $\beta = 3.80$, $a = -0.011$, $b = -0.020$, $c = 0.0013$, and $k = 0.033$ is the slope of the fitting line in Fig.(a). The dashed line in the inset of Fig. (b) shows the correction function $f(\langle g \rangle) = a \langle g \rangle^2 + b \langle g \rangle + c$, which is obtained by the fitting for the correlation between ε and $\langle g \rangle$, where ε is the regression residuals in the linear regression shown in Fig. (a).

相对人均 GDP 进行修正. 修正的方法是, 首先对 $\langle I \rangle^\beta$ 同 $\langle \Delta g \rangle$ 的关系进行线性回归, 观察各经济体的数据点相对拟合直线的回归残差 ε 同各经济体相对人均 GDP 的对数值 g 的 20 年均值 $\langle g \rangle$ 的关系, 如图 4(b) 插图所示, 该离差分布近似可用二次函数 $f(\langle g \rangle) = a \langle g \rangle^2 + b \langle g \rangle + c$ 拟合, 表示经济发展水平较高和较低的经济体都容易出现相对较低的增长率. 随后构建新的预测指标 $Z_f = \langle I \rangle^\beta + f(\langle g \rangle)/k$, 其中 k 为 $\langle I \rangle^\beta$ 同 $\langle \Delta g \rangle$ 的拟合直线斜率. 如图 4(b) 所示, 预测指标 Z_f 同 $\langle \Delta g \rangle$ 之间的相关性高达 0.548, 说明仅仅通过指数 I 并结合各经济体的相对人均 GDP, 就已经可以在较高程度上解释各经

济体长时期的经济发展速度同其创新能力之间的关系, 表明了指数 I 在预测经济发展速度方面的有效性. 同时, 该正相关特性也说明, 从长时期来看, 创新能力较强的经济体往往具有更快的经济发展速度.

进一步, 我们挖掘指数 I 和各经济体经济短期发展速度的关系. 首先设置一个长度为 m 年的滑动窗口, 计算各经济体的创新指数在该窗口期内的平均值 $\langle I \rangle$ 与相对人均 GDP 的年度增长率在该窗口期内的平均值 $\langle \Delta g \rangle$ 之间的 Pearson 相关性系数 r_I . 图 5 显示了滑动窗口长度 m 为 1 年、3 年、5 年时该相关性随年份的变化, 其中年份标定为每个滑

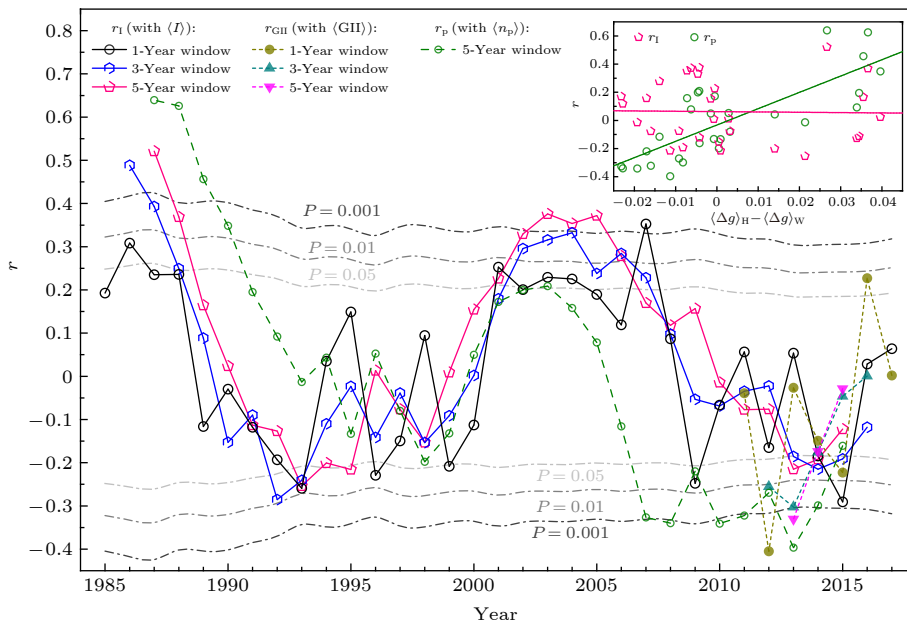


图 5 以 1 年期、3 年期和 5 年期为滑动窗口长度, 各类指标在滑动窗口期内各经济体的均值同相对人均 GDP 增长率的平均值 $\langle \Delta g \rangle$ 之间的相关性随年份的变化. 黑色、蓝色和粉色实线及其空心数据点对应指标为创新指数 (该相关性表示为 r_I); 其中不同灰度的虚线标志出相关性 r_I 在最低限度情况下 (即有效数据点最少的情况, 对应滑动窗口长度为 1 年时) 的不同显著性水平的边界, 浅灰、中灰和深灰虚线分别对应 $P = 0.05, 0.01, 0.001$ 的 r_I 值. 深黄、深青色、品红虚线及其实心数据点对应指标为全球创新指数 GII (该相关性表示为 r_{GII}). 橄榄绿色虚线及其空心数据点对应指标为相对人均专利申请数的对数值 n_p (该相关性表示为 r_p , 只显示了滑动窗口为 5 年期的情况). 插图显示的是, 采用 5 年期滑动窗口, 高收入经济体的相对人均 GDP 的平均增长率 $\langle \Delta g \rangle_H$ 与所有经济体的相对人均 GDP 增长率的均值 $\langle \Delta g \rangle_W$ 的差值 $(\langle \Delta g \rangle_H - \langle \Delta g \rangle_W)$, 同相关性 r_I (粉色点) 和相关性 r_p (橄榄绿色点) 的相关性; 实线分别为同色数据点的拟合直线

Fig. 5. Designing the moving window length of 1 year, 3 years and 5 years, for given index, the correlation between the average value of the index of each economy and the average growth rate $\langle \Delta g \rangle$ of the relative GDP per capita within the moving window are shown by curves and data points. The black, blue and pink lines and hollow data points show correlation r_I , corresponding to the index I . The different gray dashed lines show the thresholds of the correlation r_I for different level of significance in the case with the minimum data points (corresponding to the case with 1-year moving window length), and the light gray, medium gray and dark gray dashed lines correspond to the significance $P = 0.05, 0.01$ and 0.001 , respectively. The dark yellow, dark cyan, magenta dashed lines and solid data points show correlation r_{GII} , corresponding to global innovation index (GII). The olive dashed line and hollow data points show correlation r_p , corresponding to the index of the logarithmic relative number of patent applications per capita (n_p) (5-year-moving-window only). The inset shows the correlations between $(\langle \Delta g \rangle_H - \langle \Delta g \rangle_W)$ and r_I , and the correlation between $(\langle \Delta g \rangle_H - \langle \Delta g \rangle_W)$ and r_p , where $\langle \Delta g \rangle_H$ and $\langle \Delta g \rangle_W$ is the average growth rate of the relative GDP per capita within the moving window for high-income economies and all economies, respectively, and the solid lines respectively are the fitting curve for the data points with the same color.

动窗口期的中间年份. 对于不同长度的滑动窗口, 该相关性随年份的变化都呈现出较为明显的长周期波动现象, 其周期约为 20 年. 以 5 年滑动窗口的情况为例, 在该周期性波动中, 峰值处的相关性系数在 0.4 左右, 显著性 P 值可以低于 0.001, 一般强于谷值处 (谷值处相关性系数在 -0.2 左右, P 值在 0.01—0.05 之间), 如图 5 所示. 这也使得在前述的 20 年时间段内 $\langle I \rangle$ 同 $\langle \Delta g \rangle$ 之间的相关性仍然呈现为正.

图 5 还显示了各经济体自 2011 年到 2017 年的全球创新指数 GII 的窗口期均值同 $\langle \Delta g \rangle$ 的相关性. 为了同指数 I 作进一步的对比, 图 5 还展示了一个假定的绝对性指标的结果, 即直接把相对人均专利申请数的对数 n_p 作为指标, 各经济体的 n_p 在各窗口期的均值同 $\langle \Delta g \rangle$ 的 Pearson 相关性系数 r_p 随年份的变化 (图 5 橄榄绿色虚线). 可以看出, r_p 也存在波动, 但总体呈下降趋势. 根据图 1 中发达经济体与发展中经济体的大致分区, 我们把相对人均 GDP 的对数 g 大于 0.5 的经济体视作高收入经济体, 计算了每一年份高收入经济体的相对人均 GDP 增长率的均值 $\langle \Delta g \rangle_H$, 以及该年份所有经济体的相对人均 GDP 增长率的均值 $\langle \Delta g \rangle_W$, 观察其差值 $(\langle \Delta g \rangle_H - \langle \Delta g \rangle_W)$ 同 r_p 和 r_I 的关系. 如图 5 插入图所示, r_p 强烈正相关于该差值 (Pearson 相关性系数为 0.781, 显著性 P 值为 5.85×10^{-7}), 而 r_I 和该差值基本没有相关性 (Pearson 相关性系数为 0.020, P 值为 0.917). 这说明, n_p 对经济增长的解释性强烈依赖于高收入经济体的增长率, 一旦高收入经济体的增长放缓, r_p 就呈现为负值, 因此作为绝对性指标的 n_p 并不能真正反映创新对经济增长的作用; 然而, 作为相对性指标, 本文所提出的指数 I 的相关性 r_I 仅是平稳波动, 几乎不存在这种依赖性. 这一结果也暗示着, 如果某个创新指标呈现出同经济发展水平的强烈相关, 那么类似 n_p 的这种解释性依赖的问题它同样是难以避免的; 而相对性指标则可以有效避开此问题.

进一步, 为了寻找 r_I 所体现的这种周期性的成因, 我们采用 5 年滑动窗口, 选取了处于谷值的 1994 年和 2014 年, 以及处于峰值的 2004 年, 来对比其相关关系. 2004 年和 1994 年的 $\langle I \rangle$ 与 $\langle \Delta g \rangle$ 的关系对比如图 6(a) 所示, 可以观察到一些具有较高 $\langle I \rangle$ 值的经济体, 其相对人均 GDP 的年度增长率均值 $\langle \Delta g \rangle$ 在这期间出现了较大幅度的提升, 直接改变了 $\langle I \rangle$ 与 $\langle \Delta g \rangle$ 的相关性的方向. 类似的现象

也在 2014 年和 2004 年的对比中被观察到, 只是其经济增长率变化方向同 2004 年和 1994 年的对比是相反的, 如图 6(b) 所示.

通过观察各经济体在这期间的 $\langle I \rangle$ 的改变量和 $\langle \Delta g \rangle$ 的改变量, 发现在相邻谷峰处经济增长率发生较大变化的经济体, 主要是在前一谷值或峰值处出现极端性经济增长率的经济体. 例如, 一些在 2004 年窗口内具有很高的经济增长率的经济体, 在 2014 年窗口内经济增长率骤然降低 (图 6(a) 和图 6(b)). 进一步观察还发现, 在 1994 年—2004 年的谷峰转变阶段, 及 2004 年—2014 年的峰谷转变中, 发生经济增长率骤然变化的经济体有着较高的重合度, 而且多为高 $\langle I \rangle$ 值经济体. 例如, 在 2004 年窗口的 $\langle \Delta g \rangle$ 相对 1994 年窗口的增加幅度最大的 20 个经济体中, 同时位列 2014 年窗口的 $\langle \Delta g \rangle$ 相对 2004 年窗口的下降幅度最大的 20 个经济体者, 共有 12 个经济体, 而其中在 1994 年窗口和 2004 年窗口的 $\langle I \rangle$ 值均超过 0.80 的经济体就有 8 个 (阿塞拜疆、哈萨克斯坦、俄罗斯联邦、乌克兰、白俄罗斯、亚美尼亚、罗马尼亚、保加利亚), 除了俄罗斯联邦, 它们大多属于中小规模的经济体, 并主要集中在东欧、中亚等地区. 这一现象在图 6(c) 中得到了确认: 该图显示了 1994 年至 2004 年的谷峰变换中各经济体的相对人均 GDP 年度增长率均值 $\langle \Delta g \rangle$ 的改变量 $(\langle \Delta g \rangle_{2004} - \langle \Delta g \rangle_{1994})$ 同 2004 年至 2014 年的峰谷变换中的改变量 $(\langle \Delta g \rangle_{2014} - \langle \Delta g \rangle_{2004})$ 的关系, 呈现出较为强烈的负相关特性 (其相关性 $r_g = -0.725$), 而且变化最为剧烈的经济体 (如图 6(c) 右下角各经济体) 大多具有相当高的 $\langle I \rangle$ 值. 这些现象表示, 这种周期性的驱动力, 主要来自于一些创新能力较高但经济发展速度很不稳定的中小规模经济体的强烈的同步性经济波动.

5 结 论

综上所述, 本文所提出的区域创新指数 I 是一种相对性的指标, 它从专利申请角度对世界经济体的技术创新能力进行了度量. 这里首先需要说明的是, 本文对指数 I 的计算全部基于国家或地区层面, 即每个 I 值所表征的“区域”指的是国家或地区经济实体. 但是, 指数 I 本身并不局限于这一层面, 其计算方法可以拓展到对任意区域层面的技术创新能力的衡量, 例如省份、城市等. 但是, 在拓展到

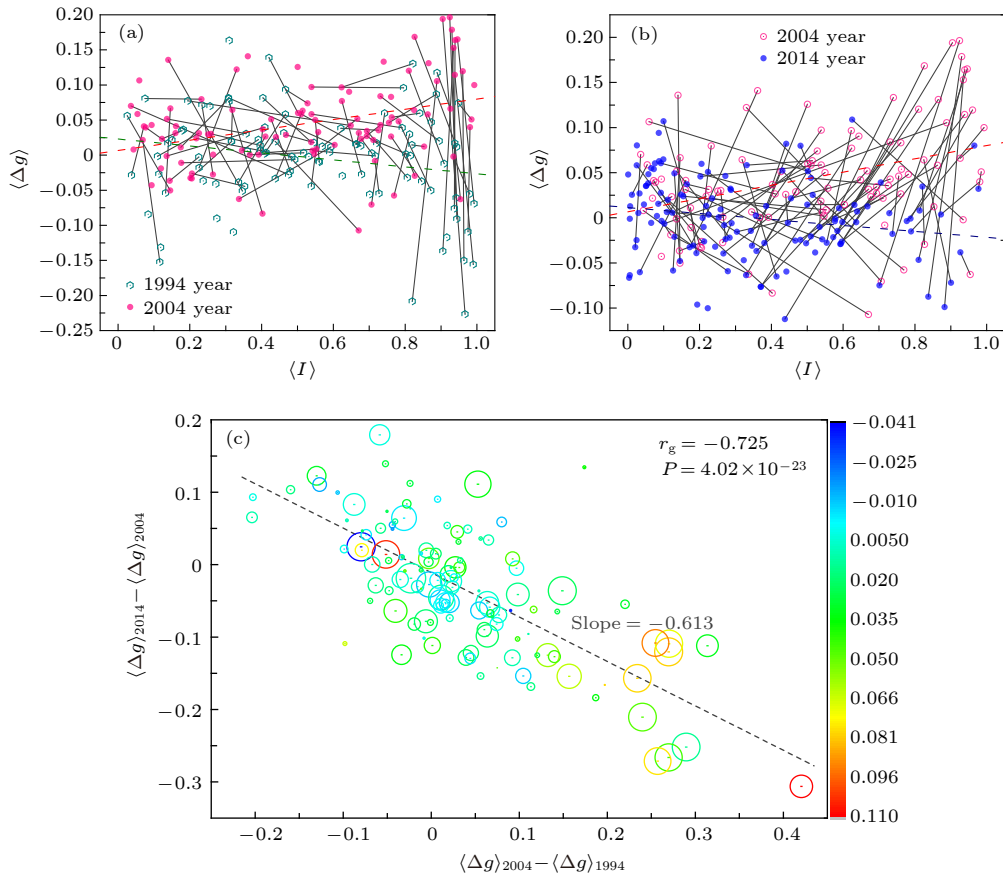


图 6 (a) 和 (b) 分别对比了在谷-峰变换和峰-谷变换前后的两个典型年份的 5 年滑动窗口内各经济体的创新指数 $\langle I \rangle$ 的平均值与相对人均 GDP 增长率 $\langle \Delta g \rangle$ 的平均值之间的相关性。(a) 青色六角圈对应 1994 年 (r_I 值谷值), 桃红色圆点对应 2004 年 (r_I 值峰值), 同一经济体由灰色线连接, 绿色虚线和桃红色虚线分别为 1994 年和 2004 年数据点的拟合直线, 斜率分别为 -0.050 和 0.073 ; (b) 桃红色圈和蓝色圆点分别对应 2004 年 (r_I 值峰值) 和 2014 年 (r_I 值谷值) 的数据点, 同一经济体由灰色线连接, 桃红色虚线 (斜率 0.073) 和蓝色虚线 (斜率 -0.033) 分别为 2004 年和 2014 年数据点的拟合直线; (c) 1994 年至 2004 年的谷-峰变换中各国的 $\langle \Delta g \rangle$ 改变量 ($\langle \Delta g \rangle_{2004} - \langle \Delta g \rangle_{1994}$) 同 2004 年至 2014 年的峰-谷变换中各经济体的 $\langle \Delta g \rangle$ 改变量 ($\langle \Delta g \rangle_{2014} - \langle \Delta g \rangle_{2004}$) 的关系, 其中各数据点的直径正比于该经济体自 1995 至 2014 年间的创新指数 I 的 20 年平均值, 颜色对应于该期间各经济体的相对人均 GDP 增长率 Δg 的 20 年平均值, 虚线为拟合直线

Fig. 6. (a) and (b) respectively compare the correlations between the average value of the index I of each economy $\langle I \rangle$ and the average growth rate $\langle \Delta g \rangle$ of the relative GDP per capita at the 5-year moving windows before and after the transition from bottom on r_I wave to peak and the one from peak to bottom. Fig. (a) shows the comparison between 1994 (the cyan hexagons, at the bottom) and 2004 (the pink dots, at the peak), where the data points of the same economy are connected by gray lines, and the green dashed line and the pink dashed line respectively show the linear fittings of 1994 (with a slope of -0.050) and the one of 2004 (with a slope of 0.073); Fig. (b) shows the comparison between 2004 (the pink circles, at the peak) and 2014 (the blue dots, at the valley), where the data points of the same economy are connected by gray lines, and the pink dashed line (with a slope 0.073) and the blue dashed line (with a slope of -0.033) show the linear fittings of 2004 and 2014, respectively; Panel (c) plots the relationship between the change ($\langle \Delta g \rangle_{2004} - \langle \Delta g \rangle_{1994}$) in the valley-peak transition and the change ($\langle \Delta g \rangle_{2014} - \langle \Delta g \rangle_{2004}$) in the peak-valley transition, where the diameter of each circle is proportional to the 20 year average $\langle I \rangle$ of the economy's index, and the color corresponds to the 20-year average growth rate $\langle \Delta g \rangle$ of the economy's relative GDP per capita.

其他区域层面时, 为了保证该指数的统一性, 计算基准需要固定在图 1 所示的拟合直线上, 即依据国家或地区经济实体层面所得到的相对人均专利申请数和相对人均 GDP 之间关系的拟合线; 也就是说, 不论进行什么层面的计算, 所参考的都是通过世界拟合线所得的期望值。

另一点需要强调的是, 由于地区经济发展水平

是影响创新产出绝对量的首要因素, 而指数 I 基本排除了地区经济发展水平的影响, 所表示的既非一个经济体的创新产出总量, 也非人均意义上的绝对总量, 而是其创新产出同其经济发展水平下的期望的相对量, 即一个地区在多大程度上能够超越其经济发展水平的制约而去促进其创新, 因此, 指数 I 虽然不论在形式上还是在数据基础上都是一个纯

粹的宏观指标,但它所揭示的则是区域经济中所隐藏的可促进创新的结构性的强弱。

该指数 I 作为一个单一性的宏观指标,其数据依赖性极低,在仅仅包含专利申请信息而尚未引入更多信息的情况下,就已经展示出对地区经济发展的较高水平的解释性,有助于挖掘技术创新对经济增长的长期促进作用.而且,该指数还揭示出若干新的发现,它们迥异于传统的基于绝对性指标的认知,主要包括以下两点.

第一,该指数重构了世界各经济体的技术创新能力排序,指出中国大陆地区最迟在 20 世纪 80 年代后期,技术创新就已经远远超越了经济发展水平下的期望而位居世界前列,并且长期处在稳步提高中,自 2016 年后更是冠居世界首位,大大超过了众多传统认知中的创新强国.这一发现重新定位了中国大陆地区的技术创新水平,暗示着中国大陆经济有着强大的内生驱动力,早已形成稳定的创新驱动机制,为从社会创新角度来理解中国大陆地区经济的高速增长提供了全新的支撑.

第二,通过该指数同经济增长的相关性分析,我们发现技术创新同经济增长之间可能存在一个长度约为 20 年的经济周期.进一步的挖掘发现,该经济周期主要受到一些创新能力较高但经济发展速度不稳定的经济体的影响.这里也引发一系列

新的有待进一步研究的问题,特别是在社会创新同其经济波动的关联性方面^[23–25].

这些发现说明了这一简单指数在衡量区域技术创新能力、挖掘创新和经济增长的关系方面的有效性,同时也暗示着,这一类以揭示经济体内部结构性因素为核心目标的相对性经济指标在区域经济分析方面可能有着很大的应用潜力.此外还需指出的是,该指数的构建考虑了创新与经济增长的超线性关系,这一点也是复杂性科学视角的体现.近年来,基于复杂性科学视角的新型宏观经济指标研究层出不穷^[26],例如经济复杂性指标和适应度指标等^[27–30].相比传统的经济指标,这类新型指标的构建一般建立在对经济系统的结构性分析基础之上,往往贯通从微观到宏观的不同层面,可以更为直接和精准地反映出经济体的结构性调整,在挖掘经济结构、解释和预测经济增长方面体现出相当的优势^[31,32].在后继研究中,将进一步结合经济大数据,构建出可以贯通微观与宏观的更具有解释性的创新指标.

感谢北京师范大学王有贵教授、电子科技大学吕琳媛教授的有益讨论.

附录

表 A1 148 个经济体在 2016 年的区域创新指数 I 值
Table A1 The index I of 148 economies at 2016.

排序	经济体名称	英文名称	2016年指数I	排序	经济体名称	英文名称	2016年指数I
1	中国大陆地区	Mainland China	0.969	75	墨西哥	Mexico	0.199
2	韩国	Republic of Korea	0.965	76	爱沙尼亚	Estonia	0.185
3	日本	Japan	0.914	77	巴拿马	Panama	0.185
4	伊朗伊斯兰共和国	Islamic Republic of Iran	0.906	78	巴基斯坦	Pakistan	0.157
5	乌克兰	Ukraine	0.889	79	乌干达	Uganda	0.156
6	俄罗斯联邦	The Russian Federation	0.861	80	赞比亚	Zambia	0.153
7	吉尔吉斯斯坦	Kyrgyzstan	0.815	81	马达加斯加	Madagascar	0.147
8	摩尔多瓦	Moldova	0.786	82	马拉维	Malawi	0.140
9	亚美尼亚	Armenia	0.784	83	摩纳哥	Monaco	0.131
10	美国	USA	0.779	84	阿尔及利亚	Algeria	0.123
11	德国	Germany	0.764	85	尼泊尔	Nepal	0.117
12	蒙古	Mongolia	0.758	86	也门共和国	Republic of Yemen	0.112
13	白俄罗斯	Belarus	0.743	87	约旦	Jordan	0.108
14	波兰	Poland	0.705	88	中国香港特别行政区	Hong Kong of China	0.106
15	格鲁吉亚	Georgia	0.666	89	洪都拉斯	Honduras	0.099
16	哈萨克斯坦	Kazakhstan	0.665	90	爱尔兰	Ireland	0.097
17	土耳其	Turkey	0.651	91	津巴布韦	Zimbabwe	0.089

表 A1 (续) 148 个经济体在 2016 年的区域创新指数 I 值
Table A1 (continued). The index I of 148 economies at 2016.

排序	经济体名称	英文名称	2016年指数 I	排序	经济体名称	英文名称	2016年指数 I
18	印度	India	0.649	92	厄瓜多尔	Ecuador	0.088
19	突尼斯	Tunisia	0.619	93	孟加拉国	Bangladesh	0.082
20	罗马尼亚	Romania	0.572	94	玻利维亚	Bolivia	0.075
21	乌兹别克斯坦	Uzbekistan	0.556	95	秘鲁	Peru	0.071
22	塞尔维亚	Serbia	0.555	96	古巴	Cuba	0.068
23	法国	France	0.547	97	卢旺达	Rwanda	0.060
24	新西兰	new Zealand	0.543	98	加纳	Ghana	0.057
25	匈牙利	Hungary	0.533	99	马耳他	Malta	0.041
26	保加利亚	Bulgaria	0.514	100	多米尼加共和国	Dominican Republic	0.040
27	芬兰	Finland	0.511	101	巴哈马	Bahamas	0.039
28	英国	Britain	0.511	102	萨尔瓦多	El Salvador	0.032
29	奥地利	Austria	0.510	103	巴林	Bahrain	0.028
30	阿塞拜疆	Azerbaijan	0.499	104	毛里求斯	Mauritius	0.026
31	意大利	Italy	0.494	105	哥斯达黎加	Costa Rica	0.025
32	新加坡	Singapore	0.491	106	塞浦路斯	Cyprus	0.021
33	丹麦	Denmark	0.489	107	特立尼达和多巴哥	Trinidad and Tobago	0.019
34	波斯尼亚和黑塞哥维那	Bosnia and Herzegovina	0.472	108	卡塔尔	Qatar	0.010
35	斯里兰卡	Sri Lanka	0.467	109	博茨瓦纳	Botswana	0.007
36	捷克共和国	Czech Republic	0.466	110	危地马拉	Guatemala	0.003
37	马来西亚	Malaysia	0.465	111	阿曼	Oman	0.002
38	苏丹	Sudan	0.441	—	阿鲁巴	Aruba	—
39	以色列	Israel	0.439	—	安哥拉	Angola	—
40	克罗地亚	Croatia	0.424	—	阿尔巴尼亚	Albania	—
41	拉脱维亚	Latvia	0.423	—	阿拉伯联合酋长国	United Arab Emirates	—
42	瑞典	Sweden	0.421	—	布隆迪	Burundi	—
43	越南	Vietnam	0.420	—	布基纳法索	Burkina Faso	—
44	葡萄牙	Portugal	0.419	—	伯利兹	Belize	—
45	阿拉伯埃及共和国	Arab Republic of Egypt	0.410	—	巴巴多斯	Barbados	—
46	巴西	Brazil	0.391	—	文莱达鲁萨兰国	Brunei Darussalam	—
47	希腊	Greece	0.387	—	科特迪瓦	Ivory Coast	—
48	泰国	Thailand	0.383	—	刚果(金)	The Democratic Republic of Congo	—
49	摩洛哥	Morocco	0.361	—	刚果(布)	The Republic of Congo	—
50	挪威	Norway	0.358	—	吉布提	Djibouti	—
51	肯尼亚	Kenya	0.357	—	埃塞俄比亚	Ethiopia	—
52	南非	South Africa	0.351	—	斐济	Fiji	—
53	荷兰	Netherlands	0.340	—	圭亚那	Guyana	—
54	黑山	Montenegro	0.334	—	海地	Haiti	—
55	斯洛伐克共和国	Slovak Republic	0.324	—	伊拉克	Iraq	—
56	加拿大	Canada	0.316	—	柬埔寨	Cambodia	—
57	圣马力诺	San Marino	0.305	—	黎巴嫩	Lebanon	—
58	立陶宛	Lithuania	0.300	—	利比亚	Libya	—
59	不丹	Bhutan	0.298	—	莱索托	Lesotho	—
60	哥伦比亚	Colombia	0.295	—	中国澳门特别行政区	Macau of China	—

表 A1 (续) 148 个经济体在 2016 年的区域创新指数 I 值
Table A1 (continued). The index I of 148 economies at 2016.

排序	经济体名称	英文名称	2016年指数 I	排序	经济体名称	英文名称	2016年指数 I
61	莫桑比克	Mozambique	0.292	—	北马其顿	North Macedonia	—
62	西班牙	Spain	0.288	—	马里	Mali	—
63	比利时	Belgium	0.283	—	尼日利亚	Nigeria	—
64	卢森堡	Luxembourg	0.276	—	尼加拉瓜	Nicaragua	—
65	澳大利亚	Australia	0.271	—	巴布亚新几内亚	Papua New Guinea	—
66	纳米比亚	Namibia	0.266	—	巴拉圭	Paraguay	—
67	瑞士	Switzerland	0.265	—	斯洛文尼亚	Slovenia	—
68	阿根廷	Argentina	0.243	—	阿拉伯叙利亚共和国	Syrian Arab Republic	—
69	沙特阿拉伯	Saudi Arabia	0.240	—	塔吉克斯坦	Tajikistan	—
70	智利	Chile	0.235	—	土库曼斯坦	Turkmenistan	—
71	牙买加	Jamaica	0.233	—	坦桑尼亚	Tanzania	—
72	冰岛	Iceland	0.221	—	乌拉圭	Uruguay	—
73	印度尼西亚	Indonesia	0.215	—	委内瑞拉玻利瓦尔共和国	Bolivarian Republic of Venezuela	—
74	菲律宾	Philippines	0.200	—	萨摩亚	Samoa	—

注: “—”说明该年份的该经济体的数据缺失, 相应也没有其排序序号.

参考文献

- [1] Hansen T, Winther L 2011 *Eur. Urban Reg. Stud.* **18** 321
- [2] Fagerberg J E, Srholec M 2008 *Res. Policy* **37** 1417
- [3] Wignaraja G 2012 *J. Asian Econ.* **23** 224
- [4] Mohnen P, Dagenais M 2002 *Towards an Innovation Intensity Index: The Case of CIS 1 in Denmark and Ireland* In: *Innovation and Firm Performance* (London: Kleinknecht A, Mohnen P, eds.) pp3–30
- [5] Slaper T F, Hart N R, Hall T J, Thompson M F 2011 *Econ. Dev. Q.* **25** 36
- [6] Clayton T, Borgo M D, Haskel J 2009 <https://ssrn.com/abstract=1345684>
- [7] Żelazny R, Pietrucha J 2017 *Q. J. Econ. Econ. Policy* **12** 43
- [8] Crespo N F, Crespo C F 2016 *J. Bus. Res.* **69** 5265
- [9] Sohn S Y, Kim D H, Jeon S Y 2016 *Technol. Anal. Strateg. Manag.* **28** 492
- [10] McGrath M E, Romeri M N 1994 *J. Prod. Innov. Manag.* **11** 213
- [11] Narin F, Olivastro D 1988 *Technology Indicators Based On Patents And Patent Citations* In: *Handbook of Quantitative Studies of Science and Technology* (Amsterdam: Van Raan A F J, eds.) pp465–507
- [12] Guan J C, Gao X 2009 *J. Associat. Inf. Sci. Technol.* **60** 35
- [13] Csajbók E, Berhidi A, Vasas L, et al. 2007 *Scientometrics* **73** 91
- [14] Mester G 2016 *Interdiscipl. Descript. Compl. Syst.* **14** 1
- [15] Gao H 2014 *Internat. J. Adv. Manag. Sci.* **3** 128
- [16] May M, Brody H 2015 *Nature* **522** S1
- [17] Haunschild R, Bornmann L 2015 *Scientometrics* **102** 1829
- [18] Report of China Regional Innovation Index Research Group 2019 *Western China 2019 19* (in Chinese) [中国区域创新指数报告课题组 2019 中国西部 2019 19]
- [19] Bettencourt L M A, Lobo J, Helbing D, Kühnert C, West G B 2007 *Proc. Natl. Acad. Sci. U. S. A.* **104** 7301
- [20] Bettencourt L M A, West G B 2010 *Nature* **467** 912
- [21] Arbesman S, Kleinberg J M, Strogatz S H 2009 *Phys. Rev. E* **79** 16115
- [22] Wunsch-Vincent S, Lanvin B, Dutta S 2015 *eSocialSciences Working Paper* id: 7491
- [23] Hirooka M 2003 *J. Evol. Econ.* **13** 549
- [24] Groot B de, Franses P H 2009 *Technol. Forecast. Social Chang.* **76** 1021
- [25] Hausman A, Johnston W J 2014 *J. Bus. Res.* **67** 2720
- [26] Gao J, Zhang Y C, Zhou T 2019 *Phys. Rep.* **817** 1
- [27] Hidalgo C A, Klinger B, Barabási A L, Hausmann R 2007 *Science* **317** 482
- [28] Hidalgo C A, Hausmann R 2009 *Proc. Natl. Acad. Sci. U.S.A.* **106** 10570
- [29] Caldarelli G, Cristelli M, Gabrielli A, Pietronero L, Scala A, Tacchella A 2012 *PLoS ONE* **7** e47278
- [30] Cristelli M, Gabrielli A, Tacchella A, Caldarelli G, Pietronero L 2013 *PLoS ONE* **8** e70726
- [31] Neffke F, Henning M, Boschma R 2011 *Econ. Geogr.* **87** 237
- [32] Tacchella A, Mazzilli D, Pietronero L 2018 *Nat. Phys.* **14** 861

SPECIAL TOPIC—Statistical physics and complex systems

Relativistic regional innovation index and novel business cycle^{*}

Fang Xue-Jin Cui Jun-Ying Hu Dan-Dan Han Xiao-Pu[†]*(Alibaba Research Center for Complexity Sciences, Hangzhou Normal University, Hangzhou 311121, China)*

(Received 16 December 2019; revised manuscript received 16 March 2020)

Abstract

In this paper, we propose a new type of relativistic regional innovation index by using the international patent application data. Based on the super-linear relationship between regional innovation and economic development, the new index can eliminate the influence of economic development level on innovation capabilities, and can effectively achieve the comparison of innovation capabilities among economies at different economic development levels. This new index is quite simple, and points out a series of new findings that are sharply different from the traditional cognitive phenomena, e.g. the index shows that the technological innovation capabilities of mainland China are among the highest in the world in 1980s. Moreover, the use of this new index not only can efficiently explain the economic growth of countries in the world at a higher level, but also find that there is a novel 20-year business cycle in the correlation between the index and economic growth rate. These results show that the index, as a simple single indicator, can achieve a higher degree of explanatory ability with minimal data dependence. This new index not only repositions the innovation capacity of world's economies, but also provides a new insight into an in-depth understanding of the relationship between innovation and economic development, and implies the development potential and application space such a kind of relativistic economic indicator.

Keywords: regional innovation index, business cycle, relativistic economic indicator, complexity in economy**PACS:** 89.65.Gh, 89.75.-k, 89.65.-s**DOI:** 10.7498/aps.69.20191970

^{*} Project supported by the National Natural Science Foundation of China (Grant Nos. 61873081, 11622538, 61673150, 61673151), the Natural Science Foundation of Zhejiang Province, China (Grant Nos. LGF18F030007, LR16A050001), and the Ximmiao Talent Program of Zhejiang Province, China (Grant No. 2019R426070).

[†] Corresponding author. E-mail: xp@hznu.edu.cn

复杂系统重构

张海峰 王文旭

Complex system reconstruction

Zhang Hai-Feng Wang Wen-Xu

引用信息 Citation: *Acta Physica Sinica*, 69, 088906 (2020) DOI: 10.7498/aps.69.20200001

在线阅读 View online: <https://doi.org/10.7498/aps.69.20200001>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

一种基于分支过程的信息流行度动力学模型

Modeling information popularity dynamics based on branching process

物理学报. 2019, 68(7): 078901 <https://doi.org/10.7498/aps.68.20181948>

基于指数加权-核在线序列极限学习机的混沌系统动态重构研究

Dynamic reconstruction of chaotic system based on exponential weighted online sequential extreme learning machine with kernel

物理学报. 2019, 68(10): 100503 <https://doi.org/10.7498/aps.68.20190156>

复杂网络链路可预测性: 基于特征谱视角

The link predictability of complex network from a spectrum perspective

物理学报. 2020, 69(8): 1 <https://doi.org/10.7498/aps.69.20191817>

基于广义似然比判决的混沌信号重构方法

Reconstruction algorithm of chaotic signal based on generalized likelihood ratio threshold-decision

物理学报. 2017, 66(4): 040503 <https://doi.org/10.7498/aps.66.040503>

基于瀑布型多重网格加速的复指数波前复原算法

Complex exponential reconstruction algorithm accelerated by cascadic multigrid method

物理学报. 2019, 68(10): 104209 <https://doi.org/10.7498/aps.68.20182137>

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.

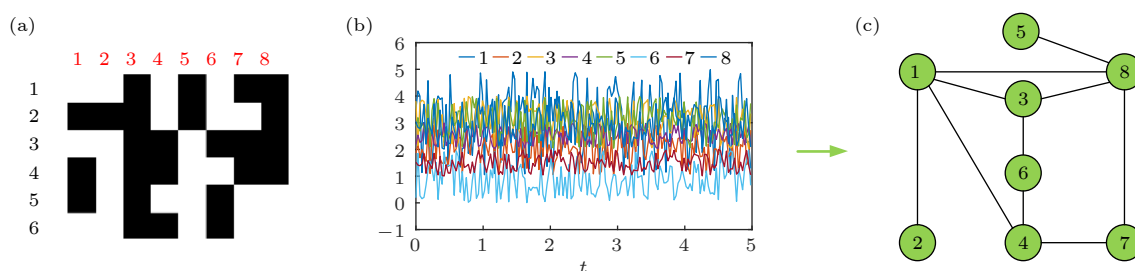


图 1 网络重构示意图 (a) 通过离散的数据; (b) 连续的数据; (c) 推断网络结构

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 $\mathbf{X} \in \mathbb{R}^N$:

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

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} :

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

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

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

(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

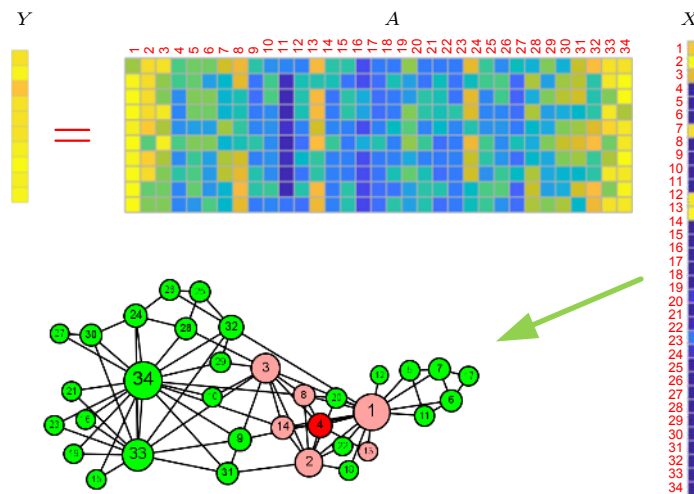


图 2 基于压缩感知方法重构 Karate 网络中 4 号节点的邻居 (重构方法见 2.4 节)

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.

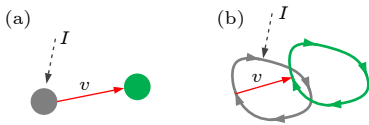


图 3 驱动-响应实验示意图. 对稳态系统施加 (稳态是一个稳定点 (a), 或者一个周期轨道 (b)) 一个持续驱动 I , 系统达到另外一个稳态. 两个稳态的差异 v 包含了网络的拓扑结构

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

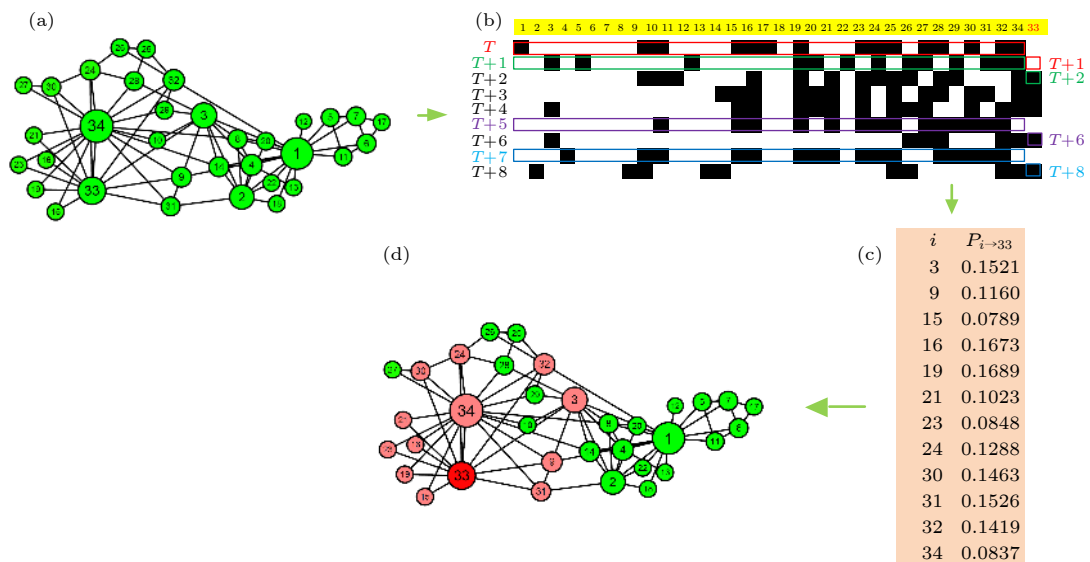


图 4 EM 算法推断 Karate 网络 33 号节点的结构 (a) 网络结构; (b) 二进制数据; (c) EM 算法推断出节点 33 的结构; (d) 真实网络 33 号节点的结构

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

电力电子化电力系统稳定的问题及挑战:以暂态稳定比较为例

杨子千 马锐 程时杰 占萌

Problems and challenges of power–electronic–based power system stability: A case study of transient stability comparison

Yang Zi-Qian Ma Rui Cheng Shi-Jie Zhan Meng

引用信息 Citation: *Acta Physica Sinica*, 69, 088907 (2020) DOI: 10.7498/aps.69.20191954

在线阅读 View online: <https://doi.org/10.7498/aps.69.20191954>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

高光束质量、高功率稳定性激光器的设计及实验研究

Design and experimental study on high quality beam and high stability power of laser

物理学报. 2018, 67(7): 074201 <https://doi.org/10.7498/aps.67.20171932>

复杂系统重构

Complex system reconstruction

物理学报. 2020, 69(8): 088906 <https://doi.org/10.7498/aps.69.20200001>

石墨炔衍生物结构稳定性和电子结构的第一性原理研究

First–principle study of structure stability and electronic structures of graphyne derivatives

物理学报. 2017, 66(10): 107102 <https://doi.org/10.7498/aps.66.107102>

腔光力系统制备微波非经典态研究进展

Research progress in non–classical microwave states preparation based on cavity optomechanical system

物理学报. 2020, 69(5): 054203 <https://doi.org/10.7498/aps.69.20191735>

负参数空间分数阶Chua系统的动力学行为及实验验证

Dynamics analysis and circuit implementation of fractional–order Chua's system with negative parameters

物理学报. 2017, 66(23): 230502 <https://doi.org/10.7498/aps.66.230502>

专题：统计物理和复杂系统

电力电子化电力系统稳定的问题及挑战： 以暂态稳定比较为例*

杨子千 马锐 程时杰 占萌†

(华中科技大学电气与电子工程学院, 强电磁工程与新技术国家重点实验室, 电力安全与高效湖北省重点实验室, 武汉 430074)

(2019年12月24日收到; 2020年3月19日收到修改稿)

随着电力电子技术的进步和环境保护对清洁能源的要求, 以同步发电机为主的传统电力系统正向着多样化电力电子装备为主的电力系统转变, 由此电力系统正面临着百年来未有之大变局. 近年来, 国内外不断报道出以电力电子装备为主的新能源基地和传统高压直流等机理不明的电力事故, 严重威胁了电力系统安全稳定运行. 针对上述问题, 本文首先介绍传统电力系统暂态稳定分析的主要方法, 接着分析了典型故障场景下简单电力电子化电力系统的动力学行为, 并建立了同时包含电力电子设备与传统同步机的多机耦合系统模型, 最后总结了电力电子化电力系统的非线性、多时标、复杂性的本质特点, 归纳其暂态稳定的基本问题与挑战以及对未来研究方向的展望, 希望引起复杂系统和统计物理背景的研究人员的广泛兴趣.

关键词: 电力电子化电力系统动态, 暂态稳定, 并网功率变换器, 复杂系统动态**PACS:** 89.75.-k, 05.45.-a, 84.70.+p, 84.30.Jc**DOI:** 10.7498/aps.69.20191954

1 引言

随着能源生产和消费转型, 电力系统正从依赖化石能源的电力系统向着高比例可再生能源的新一代电力系统转变. 电力电子设备正在取代以同步发电机和异步电动机为主的电磁变换装备. 如图 1 所示, 在电能生产环节, 风力和光伏发电等清洁能源正大规模接入电网; 在电能传输环节, 高压直流输电系统和柔性交流输电系统 (FACTS) 的发展可实现跨区域远距离输电; 在电力负荷侧, 交流传动变频负荷满足用户多样化需求, 并提高了用电效率. 源网荷多方面的电力电子化已成为现代电力系统发展的重要趋势, 在一次装备方面电力系统正经历伟大的历史性变革, 这是电力系统百年来的巨大

变化^[1-5].

电力电子设备为电力系统带来高效的同时, 也因其复杂性给电力系统的安全稳定运行带来了巨大的挑战. 作为国际上最近报道的一次大电网事故, 2019年8月9日英国 Little Wymondley 一条线路发生单相短路故障引发了伦敦在内的英格兰和威尔士大部分地区停电, 约 110 万人受到停电影响. 英国国家电网 (NationalgridESO) 的报告指出, 风电、直流输电等无惯性的电力电子并网电源高占比是本次事故的深层次原因^[6]. 当前电力电子化电力系统受到短路故障等各种大扰动后暂态行为机理尚未明确, 传统继电保护装置无法快速抑制事故影响的扩散, 电网时刻面临失稳风险. 基于传统电力系统的基本认识和基础理论将无法适应全新的装备, 寻找电力电子化电力系统较直观的物理图像和机理分析亟需解决.

* 国家重点研发计划 (批准号: 2017YFB0902000)、国家电网公司科技项目 (批准号: SGXJ0000KXJS1700841) 和国家自然科学基金国际 (地区) 合作与交流项目 (批准号: 11861131011) 资助的课题.

† 通信作者. E-mail: zhanmeng@hust.edu.cn

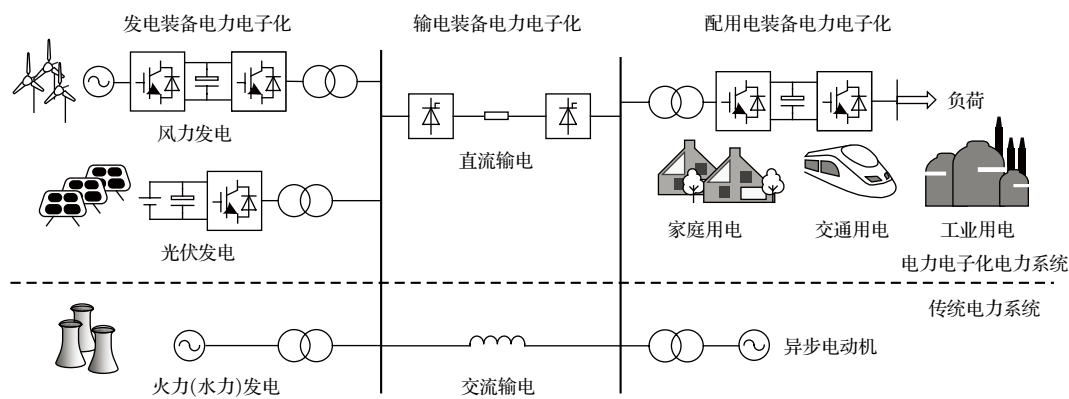


图 1 电力电子化电力系统示意图

Fig. 1. Schematic diagram of power electronic dominated power systems.

电力系统受到各种扰动后保持安全稳定运行是电网面临的最核心的问题之一^[7,8]. 工程中一般根据电力系统的特点对电力系统稳定性分析进行分类和简化. 对于某一特定的运行状态, 将遭受一个微小扰动后系统能否恢复扰动前状态称为小干扰稳定性分析. 其本质为研究系统运行点(不动点)附近的局部稳定性, 一般将系统的非线性方程组在工作点处进行线性化, 然后用线性系统理论方法, 如特征值, 加以分析. 与之相对应, 对某一稳定运行状态施加特定的大的扰动, 研究扰动后系统是否可以达到一个可以接受的(新的或旧的)运行状态称为暂态稳定分析, 或者称为大扰动稳定分析. 其本质为系统全局稳定性, 需要确定故障后运行点吸引子稳定边界与故障恢复时刻的状态之间的关系, 在分析的过程中需保留原始的非线性关系.

在传统电力系统动态研究中, 不同元件及控制器耦合相互作用引发振荡问题的研究通常属于小干扰稳定分析范畴, 用线性系统理论加以研究. 人们根据不同的振荡机理和数学模型, 总结出低频振荡和次同步振荡等不同问题, 并形成相应的分析方法和控制策略. 对于暂态稳定性分析, 同步发电机等电磁变换装备在一次调频时间尺度内的暂态特性主要由转子机电动态所决定. 由此同步发电机转子在不平衡功率(即: 原动机的输入机械功率与同步机的电磁输出功率)下的转子运动被认为是暂态稳定分析的核心内容. 针对单同步机连接到无穷大母线系统, 可以用等面积定则来分析^[7,8], 并通过加速和减速面积来提供基本物理图像. 进一步对于多同步机耦合系统, 学术界曾发展出多种分析方法, 主要有基于李雅普诺夫能量函数的直接法和基于多机空间映射的扩展等面积法^[9-14]. 传统电力系统

稳定分析经过五、六十年的发展, 已经形成了较为严格的数学理论, 并为电网部门在线运行和事故预判方面提供帮助.

与传统电力系统小干扰稳定性分析相对应, 电力电子化电力系统同样对非线性系统动力方程组在工作点处线性化, 通过研究线性系统的稳定性来判断系统稳定性. 其中包括状态空间法、阻抗法、幅相运动方程法等. 其中, 状态空间分析法(又称模态分析)是线性控制系统中常用的分析方法^[15]. 这种方法能精确分析振荡现象, 并适用于复杂多机系统, 但较缺乏对振荡机理的解释, 而且当系统结构发生改变(如投入或切除一条线路)时, 都需重新列写状态方程. 尤其重要的, 实际电力电子装备通常由不同的厂商生产, 内部控制方式和控制参数都不可能精确得到, 利用状态空间分析法来分析不太现实. 为克服这些不足, 工程上大量有实际应用的是基于复频率域分析的阻抗法^[16-18]. 其将电力电子设备黑箱化, 以电压和电流作为设备输入和输出变量, 以阻抗传递函数矩阵的形式来看待设备, 然后根据广义奈奎斯特判据或简单的波特图来判断整个系统的稳定性. 然而当前研究人员对于不同坐标系下的多种阻抗分析方法、交流运行点的处理、设备非线性系统线性化、如何推广用于多个电力电子设备系统、分析结果的准确性等也仍有诸多激烈的争论. 此外, 幅相运动方程法^[19,20]则推广传统的转子运动方程形式, 认为多样化电力电子设备也可统一表现为不平衡瞬时功率驱动下的内电势运动, 采用输入有功和无功功率与输出内电势幅值和相位的形式. 由于在电气领域, 基于非正弦时变电压和电流信号的功率理论还远未完备, 从功率的转换、储存、耗散等角度来分析大系统也仍然面

临巨大的挑战. 后两种频域分析方法都是对不同设备和传输线网络分开来建模, 方便网络结构发生改变情况下的分析, 但是应用到多机系统仍可能面临传递函数阶数过高和广义奈奎斯特判据判断不方便等问题. 近期我们的文章具体比较了状态空间分析、阻抗法和幅相运动方程法三种方法对系统小扰动稳定的分析结果, 用数学的相似矩阵变换、数值符号运算和实际算例证明了其结果的一致性^[21].

对于更加复杂的电力电子化电力系统暂态稳定问题, 由于控制器的切换与限幅、坐标变换等非线性环节、多时间尺度序贯控制方式间的复杂相互作用、机电与电磁暂态相互作用、复杂保护与控制主导的暂态控制方式等, 这些都使得传统暂态稳定分析理论难以适应当前电力电子化电力系统动态分析. 据作者们所知, 现阶段学术界对暂态稳定的研究主要针对特定简单场景下的零散研究, 仍然缺乏直观实用的方法来解释系统受扰后的物理过程^[22–32]. 如文献^[23,24]参照等面积定则, 从低维动力系统的角度分别分析了装备在功率同步控制和下垂控制的 VSC 系统的暂态稳定性. 文献^[26]发现了电压源型变换器 (VSC) 在遭受电压突然跌落故障后存在一种突变分岔现象. 文献^[29–31]总结了电力电子设备接入电网对同步机稳定的影响规律和对振荡的影响. 文献^[32]仍然借鉴传统的转子运动方程和等面积定则, 具体针对 VSC 内部的锁相环 (PLL) 动态, 得到了暂态稳定分析的一些简单结果.

2017 年国家电力调度控制中心主任李明节等^[33]总结了新能源并网系统引发的复杂振荡问题及其对策, 并归纳了振荡的一些新特点. 如可出现不同的振荡频率, 甚至是超同步振荡 (超出工频的振荡); 振荡频率可漂移; 振荡可发生在电网不同位置; 振荡特性 (频率、阻尼及稳定性) 受多变流器 (群) 聚合效应以及电网诸多参数、乃至风光等外部条件的影响, 具有影响因素复杂、大范围时变等特征; 振荡极易进入控制限幅, 导致振荡往往始于小信号负阻尼失稳, 而终于非线性自持续振荡; 振荡在机理上可能涉及到多变流器间及其与大电网之间的动态相互作用等. 由此可见, 为解决电力电子化电力系统复杂振荡问题, 我们可能必须对付和处理系统的固有非线性问题. 另外, 2015 年 9 月第 8 届 FEPPCON VIII 国际会议 (The Future of Electronic Power Processing and Conversion) 也

明确指出: 由于电力系统中的变流器越来越多, 二者之间的交互作用越来越复杂, 对传统电网运行特性的改造也越来越明显. 如何分析、设计、控制、集成, 才能确保电力电子化的供电系统仍然能够维持安全、稳定、高效地长期运行? 这是摆在电力电子、电力系统等学科研究人员面前的世纪难题.

在另一方面, 物理、数学和自动控制背景的研究人员近些年也积极进入到复杂电力网络领域之中, 在电力系统与复杂系统交叉方面开展了一系列卓有成效的研究, 如德国 Menck 等^[34]和 Schultz 等^[35]、德国 Rohden 等^[36]和 Schafer 等^[37]、美国 Motter 等^[38]和 Yang 等^[39]、瑞士 Dorfler 等^[40]和 Grob 等^[41]、英国 Carareto 等^[42]. 但是他们多以智能电网同步稳定问题为背景, 研究复杂网络上相互耦合的二阶 Kuramoto 相振子的同步化过程, 较多地从网络动力学的角度研究智能电网的网络拓扑结构如何影响整个系统的同步能力等问题. 国际上著名的复杂系统圣地圣塔菲研究所也于 2012 年召开学术沙龙: Power Grids as Complex Networks, 召集了智能电网和复杂系统两方面的专家, 探讨这两个领域共同关心的核心问题, 以促进复杂系统和电力系统的相互融合和进步.

针对上述研究背景及现状, 我们研究小组近些年已经开展了一些工作^[21,43–50]. 本文将首先介绍以同步发电机为主导的传统电力系统单机及多机系统暂态分析的基本理论, 然后通过电力电子化电力系统单机无穷大母线系统电压突然跌落和再恢复的具体场景, 结合我们近期完成的一些研究成果, 对其暂态行为进行对比研究, 由此揭示出电力电子化电力系统受扰后的典型特征. 进一步将建立同时包含电力电子设备与同步机的多机系统模型, 突出和比较这两种设备在暂态分析中的异同. 最后, 将归纳和总结电力电子化电力系统暂态分析面临的基本问题与挑战, 以期引起复杂系统和统计物理背景的研究人员对这一重要前沿问题的兴趣.

2 传统电力系统模型及暂态稳定分析

2.1 单机系统暂态分析模型

同步发电机是电力系统中最核心的元件^[7,8], 其转子磁链与定子磁链的交链实现了机械能与电能之间的相互变换. 设转子励磁绕组中心轴为

d 轴,沿转子旋转方向领先 d 轴 90 电角度的坐标轴为 q 轴. 转子在 d 轴上有励磁绕组 f 及等值阻尼绕组 D ,在 q 轴上有等值阻尼绕组 Q . 不考虑零序分量,定子 abc 三相绕组经派克变换等效为 dq 轴下的 d, q 电枢绕组. 若完整考虑同步机的动态过程,则需计及 d, q, f, D, Q 五个绕组的电磁动态以及转子机械动态在内的共七个动态过程. 完整电机方程为七阶模型,这还不包括励磁控制与原动机动态等. 电力系统中常见的暂态故障,如短路故障,往往持续几百毫秒到几秒,这个时间尺度被称为机电时间尺度. 在机电时间尺度内转子的机械过程将主导系统的动态行为,而电磁绕组的过渡过程影响较小. 所以通常在暂态稳定分析中,忽略电磁绕组的动态过程而仅考虑转子的运动. 考虑如图 2 所示的单机无穷大系统,同步发电机采用经典二阶模型,其中 x'_d 为发电机暂态电抗, x_T 为变压器电抗, x_L 为双回线路中单回线的电抗,并做如下假设:

- 1) 近似认为发电机暂态电抗 x'_d 后的电动势 E' 在暂态过程中保持不变;
- 2) 发电机的机械输入功率 P_M 保持不变;
- 3) 网络采用准稳态模型,在这种模型下电感上的电压电流关系为代数关系;
- 4) 忽略变压器电阻及线路电阻,忽略发电机的机械阻尼.

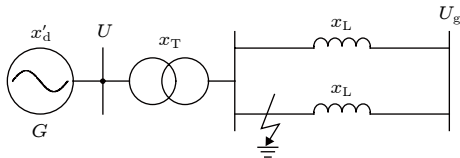


图 2 一台同步发电机经变压器升压后串联双回线路接入无穷大电网的示意图
Fig. 2. Schematic show for a single-machine-infinite-bus (SMIB) system.

发电机转子的运动方程,又称摇摆方程,表示为^[7]

$$\begin{cases} \dot{\delta} = \omega, \\ M\dot{\omega} = P_M - D\omega - P_e, \end{cases} \quad (1)$$

式中 \dot{X} 表示 X 对时间的微分, M 为发电机的惯性时间常数; δ 代表转子角度,又称功角; ω 为功角转速与同步转速的差; D 为同步机的阻尼系数; P_M 与 P_e 分别为发电机的机械输入功率和电磁输出功率,这两者之间的不平衡决定了转子的运动. 这通常也被称为经典模型,可为传统电力系统动态提供基本的物理图像.

图 3 绘制了电磁功率 P_e 与功角 δ 间的功率特性曲线. 不失一般性,这里不考虑阻尼的作用, P_e 与 δ 是最简单的正弦形式. 功率特性曲线与 $P = P_M$ 有两个交点分别为 S 和 U ,对应的功角为 δ_S 和 δ_U ,这两个点为系统的平衡点. 若同步机运行在 S 点,此时输入功率施加一个微小的增量 ΔP_M ,转子加速,功角增加 $\Delta\delta$, S 点处功率特性曲线斜率为正,所以电磁功率增加 ΔP_e ,因此制动性质的电磁转矩将相应地增加,以抑制功角增大,反之也成立. S 点处电磁功率与功角形成了负反馈调节,所以此处为小干扰的稳定平衡点 (stable equilibrium point, SEP). 相应地, U 点为小干扰非稳平衡点 (unstable equilibrium point, UEP). 单机系统的小干扰稳定性可以这样判断,而对于复杂的多机系统,则需依据特征值分析来进一步判断.

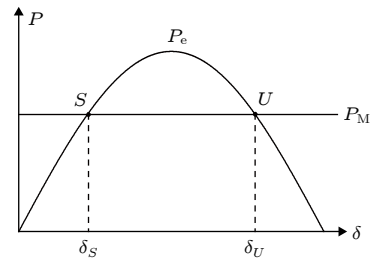


图 3 系统的功率特性曲线^[7]
Fig. 3. Power-angle relationship^[7].

2.2 单机系统等面积定则

小干扰稳定保证电力系统在工作点处线性化区域内的稳定运行,暂态稳定分析则研究电力系统在遭受大扰动下,能否恢复至新的运行状态. 传统电力系统暂态稳定分析主要有两种方法,即时域仿真和直接法. 时域仿真法将电力系统各原件的数学模型联立得到一组微分代数方程组,通过逐步积分求得各状态量随时间变化的轨迹,根据状态量最终是否回到稳定工作点上来判断其稳定性. 时域仿真法适用于各种不同复杂的设备,但是较难进行机理分析. 直接法不必计算系统整个运行轨迹(但通常仍然需对从故障发生到故障恢复过程作数值积分),而通过计算故障恢复时刻的状态满足的条件来直接判断稳定性. 其简化了故障恢复后的轨迹数值积分,可对稳定性作快速判断,并能提供稳定裕度和安全运行警告. 直接法从设备运动的规律出发,更能说明暂态稳定的机理. 本文将重点介绍单机系统和多机系统暂态稳定的几种直接法.

对于暂态稳定分析, 一般认为故障前系统工作于稳定状态, 稳定功角设为 δ_0 . 设在 t_0 时刻, 双回线中的一条线路发生短路故障, 保护装置检测到故障后, 于 t_c 时刻切除故障. 将整个暂态过程分为 (I) 故障前, (II) 故障中和 (III) 故障后. 不同暂态过程的功率特性是不同的, 如 (2) 式所示^[7]. 式中 $x_{\text{I}} - x_{\text{III}}$ 为发电机内电动势到无穷大母线之间的总电抗, 它们之间的关系通常为 $x_{\text{II}} > x_{\text{III}} > x_{\text{I}}$, 所以有 $P_{\text{II}} < P_{\text{III}} < P_{\text{I}}$. 图 4(a) 绘制了故障前、故障中和故障后三个过程中不同电磁功率 P_e 与功角 δ 间的功率特性 (均为简单的正弦形式, 但幅度不同).

$$P_e = \begin{cases} P_{\text{I}} \sin \delta = (E'U_g/x_{\text{I}}) \sin \delta, & t < t_0, \\ P_{\text{II}} \sin \delta = (E'U_g/x_{\text{II}}) \sin \delta, & t_0 \leq t < t_c, \\ P_{\text{III}} \sin \delta = (E'U_g/x_{\text{III}}) \sin \delta, & t_c \leq t, \end{cases} \quad (2)$$

对 (1) 式中第二式两端同乘 $d\delta$, 有

$$M \frac{d^2\delta}{dt^2} d\delta = (P_{\text{M}} - P_e) d\delta, \quad (3)$$

即

$$M\omega d\omega = -(P_e - P_{\text{M}}) d\delta, \quad (4)$$

左右同时积分可得

$$\frac{1}{2}M\omega^2 + \int (P_{\text{M}} - P_e) d\delta = 0. \quad (5)$$

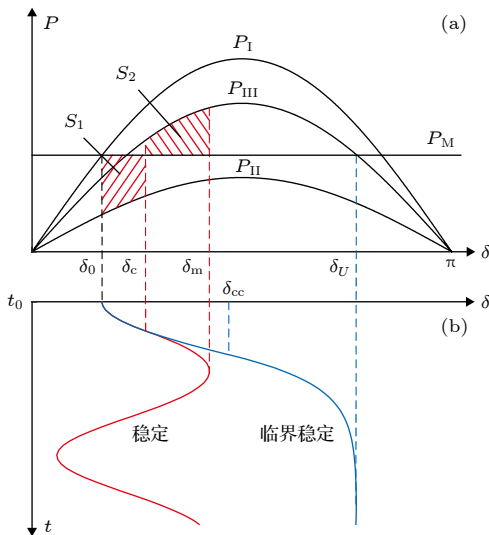


图 4 (a) 简单系统在故障前、故障中、故障后的三条不同功率特性曲线; (b) 暂态稳定 (红色曲线) 和临界稳定 (蓝色曲线) 情况下的功角的时域波形^[9]

Fig. 4. (a) Power-angle relationships for three different states of before-fault (P_{I}), during-fault (P_{II}), and post-fault (P_{III}); (b) time-domain responses of the power angle when the system are stable (red curve) and critical stable (blue curve), respectively^[9].

(5) 式左侧第一式定义为系统动能, 第二式为系统势能, (5) 式实际上表示了动能与势能之间相互转化的关系: 系统增加 (减少) 的动能由减少 (增加) 的势能转化得到, 因此系统的总能量保持恒定.

故障中由于 $P_{\text{M}} > P_e$, 转子加速, 故障切除时系统的动能为

$$V_p = \frac{1}{2}M\omega_c^2 = \int_{\delta_0}^{\delta_c} (P_{\text{M}} - P_{\text{II}} \sin \delta) d\delta = S_1. \quad (6)$$

故障恢复阶段转子将减速, 若系统暂态稳定, 则故障恢复阶段系统增加的势能可以全部消纳故障切除时系统的动能. 令 δ_m 为故障恢复阶段功角的最大值, 则有:

$$S_1 = \int_{\delta_c}^{\delta_m} (P_{\text{M}} - P_{\text{III}} \sin \delta) d\delta = S_2. \quad (7)$$

在 (6) 式和 (7) 式中, S_1 和 S_2 分别被称为加速面积和减速面积, 如图 4(a) 中红色阴影所示. 加速面积与减速面积相等时, 系统暂态稳定. 等面积定则由此得名. 由减速面积的定义可知, 如果当功角运动至 δ_U (故障后的非稳平衡点) 时, 系统增加的势能达到最大值, 若此时仍无法全部消纳故障切除时的动能, 则转子开始加速并最终失去稳定. 由此我们可以通过比较加速面积与减速面积大小来判断暂态稳定性. 运用等面积定则时, 首先通过逐步积分法计算故障中的故障轨迹, 然后比较加速面积是否小于最大减速面积, 若小于, 则暂态稳定, 反之则暂态失稳. 临界情况下的切除角和切除时间为极限切除角 δ_{cc} 和极限切除时间 T_{cc} , 这两个指标代表了系统的稳定裕度, 为继电保护的设计提供了重要参考. 图 4(b) 中绘制了暂态稳定和临界稳定两种情况下的功角时域响应, 分别用红色和蓝色曲线表示.

等面积定则本质上是不考虑阻尼条件下的能量守恒. 我们也可直接从势能函数的角度来解释. 故障恢复阶段系统的势能函数为

$$V_p(\delta) = \int_{\delta_s}^{\delta} (P_e - P_{\text{M}}) d\delta. \quad (8)$$

绘制势能函数图如图 5(a) 所示. 势能函数以 δ_s (故障后的稳定平衡点) 为中心构成一个“盆”状区域, 将转子的运动视为小球在“盆”状区域内受重力影响来回滚动, 只有在“盆”状区域内才是稳定的. 所以系统的临界能量为

$$V_{pc} = \int_{\delta_s}^{\delta_U} (P_e - P_{\text{M}}) d\delta. \quad (9)$$

图 5(b) 绘制了对应的稳定、临界稳定和失稳三种情况下系统在相空间的运行轨迹. 这与等面积定则的结论相一致. 这里假定理想发电机没有阻尼, 所以总能量保持不变, 暂态稳定时将作周期往返运动. 事实上, 同步机中存在机械阻尼和电气阻尼, 若转子第一摇摆稳定, 则之后摆动的最大幅值将不断减小, 系统将作衰减振荡运动. 上述理论分析是针对理想的情况, 若单纯从数学的角度来考虑, 则随着参数取值范围的扩大, 系统 (1) 将拥有更加丰富的动态行为, 如在整个参数空间内, 系统存在两个平衡点、一个极限环以及平衡点和极限环共存的双稳区, 系统也将出现更为广泛的等面积定则形式等 [45,46].

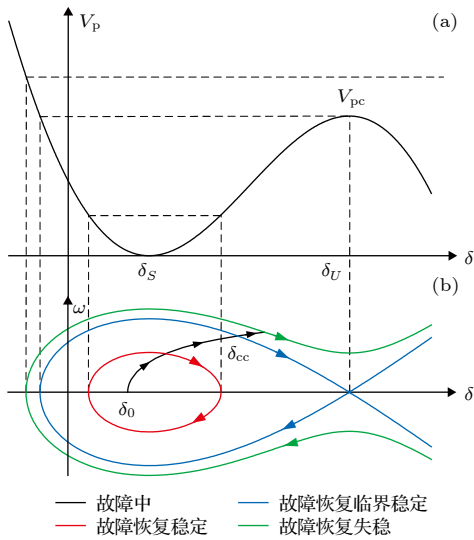


图 5 (a) 势能函数曲线; (b) 不同状态下的轨迹 [10]

Fig. 5. (a) Potential energy function curve; (b) trajectories in state space corresponding different states: stable, critically stable, and unstable [10].

2.3 多机系统暂态稳定分析

多同步机电力系统被描述为一组非线性微分代数方程, 其中微分方程由每台同步机的摇摆方程构成:

$$\begin{cases} \dot{\delta}_i = \omega_i, \\ M_i \dot{\omega}_i = P_{Mi} - D_i \omega_i - P_{ei}, \end{cases} \quad (10)$$

其中 $i = 1, 2, \dots, n$ 表示发电机的编号, 其余符号与 (1) 式中相应符号含义一致. 端电压及电磁功率为:

$$\begin{cases} u_{di} = -U_i \sin(\delta_i - \theta_i) = 0 - i_{qi} x'_{di}, \\ u_{qi} = U_i \cos(\delta_i - \theta_i) = E'_i + i_{di} x'_{di}, \end{cases} \quad (11)$$

$$\begin{cases} P_{ei} = E'_i U_i / x'_{di} \sin(\delta_i - \theta_i), \\ Q_{ei} = -U_i^2 / x'_{di} + E'_i U_i / x'_{di} \cos(\delta_i - \theta_i), \end{cases} \quad (12)$$

其中 u_{di}, u_{qi} 和 U_i, θ_i 分别为端电压在 dq 坐标系和极坐标系的分量.

多台同步机通过网络实现功率交换, 由潮流方程所决定:

$$\begin{cases} P_{ei} = U_i \sum_{j=1}^n G_{ij} U_j \cos(\theta_i - \theta_j) + B_{ij} U_j \sin(\theta_i - \theta_j), \\ Q_{ei} = U_i \sum_{j=1}^n G_{ij} U_j \sin(\theta_i - \theta_j) - B_{ij} U_j \cos(\theta_i - \theta_j), \end{cases} \quad (13)$$

其中 G 和 B 分别表示电导和电纳, $G_{ij} + jB_{ij}$ 为消除同步发电机节点外系统的节点导纳矩阵中第 i 行, j 列元素. 由 (11) 式—(13) 式联立的代数方程与 (10) 式一起组成了多机系统的暂态模型.

直接法仍需对故障中过程运行轨迹进行数值积分, 相比完全数值仿真方法的优点在于, 根据故障切除时刻的状态对稳定性直接判断. 由此其核心为对故障后稳定运行状态的吸引域的估计. 多年来, 电力系统研究人员提出了多种暂态稳定分析直接法, 具有代表性的方法有: 最近不稳定平衡点法 (closest UEP)、主导不稳定平衡点法 (controlling UEP) [10]、位能界面曲线方法 (potential energy boundary surface, PEBS) [12]、基于稳定边界的主导不稳定平衡点法 (boundary of stability region based controlling UEP, BCU) [13]、扩展等面积定则 (extended equal area criterion, EEAC) [14] 等.

通常直接法判断暂态稳定的步骤为: 首先认为系统故障前工作于稳定状态, 通过逐步积分法计算故障中的故障轨迹, 然后根据不同直接法的计算步骤计算相应的临界能量值, 进一步将故障切除时刻的能量函数值与临界能量值相比较, 若小于临界能量值则暂态稳定, 反之则暂态失稳. 其中关键问题在于如何计算相应故障下的临界能量值. 早期方法有最近不稳定平衡点法和主导不稳定平衡点法, 这两种方法基于这样的事实: 稳定平衡点的稳定边界是由系统不稳定平衡点的稳定流形并集所构成. 最近不稳定平衡点法采用计算所有不稳定平衡点中最小的能量函数值为临界能量值. 这种估计方法独立于故障时的轨迹, 结果过于保守而无法广泛应用. 主导不稳定平衡点定义为稳定流形包含故障时轨线在稳定边界上的出口点的平衡点. 主导不稳定

平衡点法采用主导不稳定平衡点的能量函数值为临界能量值. 这种方法针对故障特性类型来估计稳定域, 因而有更高的精度. 在后续发展的位能界面曲线方法中, 位能界面曲线为从稳定平衡点出发沿任意方向联结位能最大值点形成的曲面, 将故障中轨迹与位能界面曲线的交点的能量函数值为临界能量. 但有时可能给出错误的估计, 为克服这一困难, 人们将基于稳定边界的主导不稳定平衡点法结合主导不平衡点法和位能界面曲线法, 利用故障中轨线与位能界面曲线的交点找到原系统的主导不稳定平衡点. 这种方法保证稳定域的精确估计, 并提高了计算速度.

在另一方面, 20 世纪 80 年代中期, 我国薛禹胜院士和比利时 Pavella 教授提出了扩展等面积定则方法, 将系统中的多台机划分为两互补群, 并用每个群所属机相角的加权平均值作为该群等值机的相角, 原多机空间被映射至等值两机空间, 这种概念被称为部分惯量中心 (PCOI) 映射. 不同的划分机群的方法会有不同的两机映像子系统, 其中最严重的映像子系统的临界条件就是原多机系统的稳定极限条件, 所以 PCOI 映射保留了稳定特性和临界条件. 每个映像子系统稳定的充要条件为当次摆动不经过映像加速功率为零值的动态鞍点, 这与传统等面积定则方法类似, 通常被称为扩展等面积定则.

3 电力电子化电力系统模型及暂态稳定分析

3.1 VSC 的系统模型

三相电压源型功率变化器 (VSC) 具有控制交流与直流系统之间功率交换的能力, 是电力电子化系统中最常见的设备, 广泛应用于光伏发电、风力发电、柔性输电、无功补偿中^[4,5]. 本文将以 VSC 为例介绍电力电子设备暂态行为的基本特点. 图 6 为一个典型的单机 VSC 并网系统的电路图和框图. VSC 的控制器工作在 dq 锁相旋转坐标系下, 其中锁相环控制器 (phase-locked loop, PLL) 实现了 d 轴与端电压相位之间的同步, 并为控制器提供相位基准. abc 三相静止坐标系下交流控制量经派克变换为 dq 坐标系下相应的直流信号, 接着通过比例积分控制器 (PI 控制器) 对其进行控制. 电气量中 d 轴分量又称为有功分量, q 轴分量又称为无功分量. 在有功控制支路, 直流电压控制器与有功电流控制器级联, 分别控制有功电流和有功内电动势以实现维持直流电压稳定的功能; 在无功控制支路, 端电压控制器与无功电流控制器级联, 分别控制无功电流和无功内电动势以实现维持端电压稳定的功能. 最后有功和无功内电动势指令值变换回 abc 三相坐标系下, 并经脉冲宽度调制 (PWM) 技

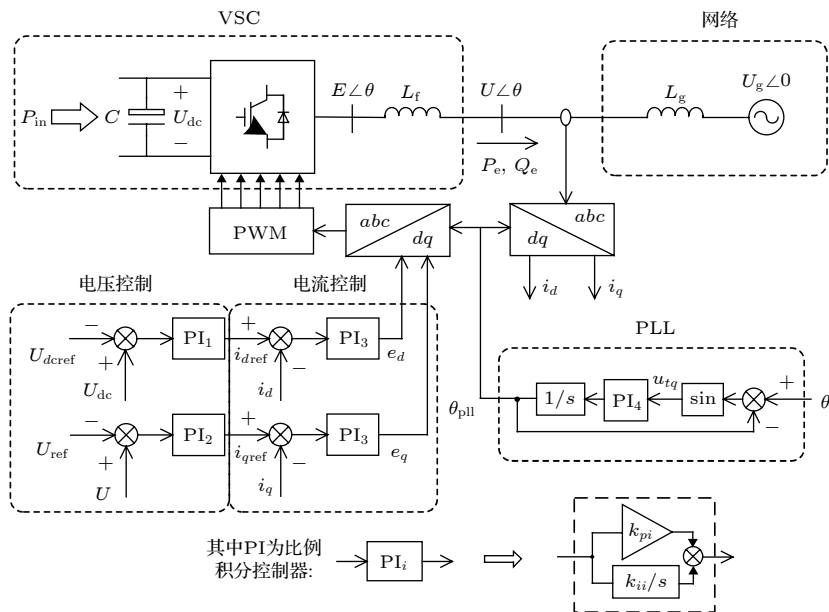


图 6 VSC 单机并网系统控制框图

Fig. 6. Schematic show of a grid-connected VSC system and its controllers.

术控制开关二极管生成实际内电动势电压,从而实现从直流(图6左端)到三相交流(图6右端)的功率转换.在图6所示的矢量控制中,不同带宽的控制环路彼此相互级联,构成了电力电子设备控制系统的基本结构.

VSC的控制环路和所控制的电气量展现出多时间尺度的特性.这里我们强调,多时间尺度是一般非线性系统的普遍特征^[51,52].一般而言,内环电流控制器的响应速度最快,为10 ms左右;外环直流电压控制器、端电压控制器和锁相环的速度稍慢,为100 ms左右.根据不同控制器的响应速度,可将VSC设备划分为电压控制时间尺度和电流控制时间尺度.对于电力系统的短路故障,相应的保护动作时间为几百毫秒到几秒,在这一时间尺度下电压控制环将主导系统动态.本文将重点考虑电压时间尺度的动态,并与全时间尺度模型仿真相比较.模型假设如下:

1) 直流侧的输入功率 P_{in} 保持不变;

2) 忽略电流环动态,有功和无功电流值将始终等于其参考值 ($i_d = i_{dref}, i_q = i_{qref}$);

3) 电感用相量描述;在这一条件下电感的电压电流关系为代数关系, ($X_g = j\omega_0 L_g, X_f = j\omega_0 L_f$);

4) 变流器采用平均模型,忽略变流器损耗和线路电阻损耗.

由此,在电压时间尺度下,模型得到简化.其控制框图如图7所示,符号的意义已在图6中标注.根据该控制框图,以下将推导系统的状态方程.选取状态变量为 $\mathbf{X} = [U, \theta, U_{dc}, \theta_{pll}, x_{pll}]^T$, 其中 x_{pll} 为锁相环控制器相关的中间变量.端电压控制器和直流电压控制器均采用PI控制,其微分方程为:

$$\begin{cases} \dot{i}_d = k_{p1}\dot{U}_{dc} + k_{i1}(U_{dc} - U_{dcref}), \\ \dot{i}_q = k_{p2}\dot{U} + k_{i2}(U - U_{ref}). \end{cases} \quad (14)$$

锁相环控制器由一个积分器和一个PI控制器组成,表现出二阶动态特性:

$$\begin{cases} \dot{\theta}_{pll} = k_{p4}\dot{x}_{pll} + x_{pll}, \\ \dot{x}_{pll} = k_{i4}u_q/U = k_{i4}\sin(\theta - \theta_{pll}), \end{cases} \quad (15)$$

其中 x_{pll} 为中间状态量.

直流电容通过吸收不平衡的功率来维持直流侧电压在正常范围:

$$\frac{d\left(\frac{C}{2}U_{dc}^2\right)}{dt} = P_{in} - P_e = P_{in} - \frac{UU_g}{X_g}\sin\theta, \quad (16)$$

进一步有

$$\dot{U}_{dc} = \frac{P_{in} - P_e}{CU_{dc}}. \quad (17)$$

交流电感采用相量模型,故流经电感的电流相量落后于两端电压相量 90° :

$$\mathbf{U} = \mathbf{U}_g + j\mathbf{X}_g\mathbf{I}, \quad (18)$$

在 dq 坐标系下表示为:

$$\begin{cases} u_d - u_{gd} = -X_g i_q, \\ u_q - u_{gq} = X_g i_d, \end{cases} \quad (19)$$

其中 u_{gd} 和 u_{gq} 分别为无穷大母线电压 U_g 在 dq 坐标轴下的投影,

$$\begin{cases} u_{gd} = U_g \cos\theta_{pll}, \\ u_{gq} = -U_g \sin\theta_{pll}. \end{cases} \quad (20)$$

VSC端电压的幅值与相位为:

$$\begin{cases} U = \sqrt{u_d^2 + u_q^2}, \\ \theta = \arctan(u_q/u_d) + \theta_{pll}. \end{cases} \quad (21)$$

将方程(20)和(21)对时间求导,可得到

$$\begin{cases} \dot{u}_d = -X_g \dot{i}_q - U_g \sin\theta_{pll} \dot{\theta}_{pll}, \\ \dot{u}_q = X_g \dot{i}_d - U_g \cos\theta_{pll} \dot{\theta}_{pll}, \\ \dot{U} = \frac{u_d \dot{u}_d + u_q \dot{u}_q}{U}, \\ \dot{\theta} = \frac{u_d \dot{u}_q - u_q \dot{u}_d}{U^2} + \dot{\theta}_{pll}. \end{cases} \quad (22)$$

联立方程(14), (15), (17), (19), 和(22), 消去中间

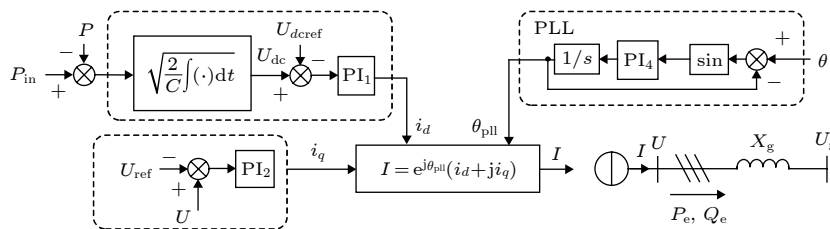


图7 VSC电压时间尺度模型系统控制框图

Fig. 7. Schematic show of the control diagram of the VSC system within the voltage timescale.

变量, 系统的状态方程整理为:

$$\begin{cases} \dot{U} = \left(P_{in} - \frac{UU_g}{X_g} \sin \theta \right) f_2 + (U_{dc} - U_{dc \text{ ref}}) f_3 \\ \quad + (U - U_{\text{ref}}) f_4 + f_5 \sin(\theta - \theta_{\text{pll}}) + x_{\text{pll}} f_6, \\ \dot{\theta} = \left(P_{in} - \frac{UU_g}{X_g \sin \theta} \right) f_7 + (U_{dc} - U_{dc \text{ ref}}) f_8 \\ \quad + (U - U_{\text{ref}}) f_9 + f_{10} \sin(\theta - \theta_{\text{pll}}) + x_{\text{pll}} f_{11}, \\ \dot{U}_{dc} = \frac{1}{CU_{dc}} \left(P_{in} - \frac{UU_g}{X_g} \sin \theta \right), \\ \dot{\theta}_{\text{pll}} = k_{p4} \sin(\theta - \theta_{\text{pll}}) + k_{i4} x_{\text{pll}}, \\ \dot{x}_{\text{pll}} = \sin(\theta - \theta_{\text{pll}}), \end{cases} \quad (23)$$

其中 $f_2 \sim f_{11}$ 均为含 θ , θ_{pll} , U , 和 U_{dc} 的非线性函数, 详见附录 A.

3.2 分岔分析

首先令状态方程右侧等于零来得到平衡点. 观察到方程组 (23) 右侧多项式中的每一项都含有以下因式中的一项: $P_{in} - \frac{UU_g}{X_g} \sin \theta$, $U_{dc} - U_{dc \text{ ref}}$, $U - U_{\text{ref}}$, $\sin(\theta - \theta_{\text{pll}})$, x_{pll} , 因此状态方程等于零的解即为上述因式为零的解. 可得系统存在两个平衡点:

$$\mathbf{X}_{\text{equ}} = \left[1 \quad \arcsin \frac{P_{in} X_g}{UU_g} \quad 1 \quad \arcsin \frac{P_{in} X_g}{UU_g} \quad 0 \right]^T \quad (24)$$

和

$$\mathbf{X}'_{\text{equ}} = \left[1 \quad \pi - \arcsin \frac{P_{in} X_g}{UU_g} \quad 1 \quad \pi - \arcsin \frac{P_{in} X_g}{UU_g} \quad 0 \right]^T. \quad (25)$$

进一步, 我们可对系统做分岔分析; 具体结果如图 8 所示, 用实线表示稳定不动点, 虚线表示非稳不动点. 可发现系统在 U_g 较大时存在两个平衡点 \mathbf{X}_{equ} 和 \mathbf{X}'_{equ} , 其中 \mathbf{X}_{equ} 的 θ 小于 90° , 为系统正常运行时的工作点, \mathbf{X}'_{equ} 中 θ 大于 90° , 为非稳平衡点. 随着 U_g 的减小, 两个平衡点逐渐靠近并碰撞后消失, 此时的分岔可称为广义鞍结点分岔. 此外, 我们发现 \mathbf{X}_{equ} 在此之前就已经通过亚临界霍普夫分岔失稳 (对应的非稳极限环在图中没有显示). 进一步在图 8 的小图中, 围绕分叉参数点附近的本征值分布显示: 在系统正常运行于稳定的 \mathbf{X}_{equ} 参数区域时, 系统的五个特征根 (λ_{1-5}) 全部位于复数

空间的左半平面, 但当 U_g 减少时, 有一对特征根 ($\lambda_{1,2}$) 逐渐靠近并穿过虚轴, 系统发生了霍普夫分岔失稳. 经计算, U_g 的分岔点为 $U_{\text{hopf}} = 0.645$. 系统参数见附录 B.

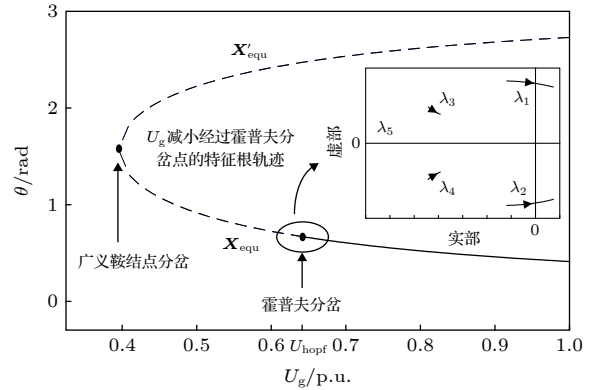


图 8 参数 U_g 变化时的分岔图 (包含亚临界霍普夫分岔和广义鞍结点分岔). 小图中显示 U_g 减小时经过霍普夫分岔点的特征根轨迹

Fig. 8. Bifurcation diagram with the variation of U_g including a sub-critical Hopf bifurcation and a generalized saddle-node bifurcation. The sub-figure shows the eigenvalue traces when U_g decreases and passes through the Hopf bifurcation point.

3.3 暂态稳定分析

3.3.1 暂态现象特征

我们在 Matlab/SIMULINK 中分别按照图 6 和图 7 搭建了 VSC 单机并网系统的详细模型和只考虑电压控制时间尺度的简化模型. 设置典型故障为: $t = 0.1$ s 时无穷大母线电压 U_g 突然从 $U_g = 1.0$ 跌落, 并持续一段时间后电压完全恢复. 持续时间记为 T_{dur} . 不失一般性, 选取参数电压跌落深度 $U_g = 0.65$ 和持续时间 $T_{\text{dur}} = 0.5$ s, VSC 的端电压的幅值和相位的响应如图 9 所示.

如图 9 所示, 电力电子设备在暂态扰动的过程中表现出明显的多时间尺度特征. 显然详细模型的时域波形含有两种不同频率的振荡成分: 即由电压控制环主导的 10 Hz 左右的低频振荡和由电流环主导的 50 Hz 左右的高频振荡, 尤其高频振荡在故障发生 ($t = 0.1$ s) 和故障恢复 ($t = 0.6$ s) 后能够很快衰减. 可见电压尺度模型的时域波形很好地保留了详细模型时域波形中的低频成分, 这也说明该模型在暂态过程分析中的合理性.

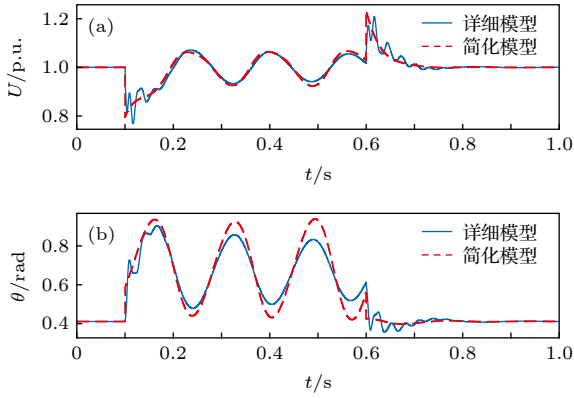


图9 VSC电压时间尺度模型与详细模型在暂态故障后的响应对比

Fig. 9. Response comparison between the voltage timescale VSC system and the detailed VSC system after transient fault.

观察图9中故障发生时和故障切除时的波形, 可发现其端电压幅值和相角展现出不连续跳跃的特征. 进一步从控制框图(图7)可看出, 系统中的积分环节输出是连续的, 而代数环节的输出则完全由当前时刻的输入决定. 在本例中VSC端电压中PI环节的比例支路与网络代数方程形成了反馈控制环路, 构成了完整的代数环. 当该代数环中一个量发生突变(如文中考虑的电压跌落), 其他变量将立即随之发生改变. 故在故障发生和故障切除的两个时刻, 代数环中的状态变量出现突变, 并将为之前的耦合微分方程组(22)提供新的初始值. 由此可见, 微分-代数关系将决定整个系统的暂态行为, 而代数环的存在将使电力电子设备的暂态分析变得更加复杂.

3.3.2 临界故障切除时间

进一步地, 可通过分析不同故障深度和故障持续时间下的系统稳定性, 来获得工程上感兴趣的临界故障切除时间. 显然故障深度越深和故障持续时间越长, 系统越不稳定. 对于跌落较深的故障, 只有及时切除才能保证系统的稳定性. 随着故障深度变浅, 临界故障切除时间将趋向于无穷, 我们通过数值计算发现临界参数 $U_{\text{critical}} \approx 0.654$. 当 $U_g > U_{\text{critical}}$ 时, 即使不切除故障系统也可以保持稳定. 由于小扰动稳定是大扰动稳定的前提条件, 有 $U_{\text{critical}} > U_{\text{hopf}}$.

我们还发现VSC在失稳的过程中表现出与同步机暂态稳定相类似的多摆现象. 由于故障深度越深, θ 发散的速度越快, 摆动的周期越少, 而处于同一摆动周期的临界清除时间几乎相同. 所以对应地

可观察到临界故障切除时间呈现出阶梯状结构. 具体细节可参考我们的近期文章[49].

3.4 VSC与同步机多机系统模型

同2.3节中传统的多机系统一样, 含有VSC的电力系统依然可描述为一组非线性微分代数方程组. 图10(a)所示为一个同时含有VSC和同步机的多机耦合系统. 同步机节点相关方程已在之前讲述, 如(11)式—(13)式. 假设第 j 节点为VSC节点, 由单机模型中(14)式, (15)式和(17)式可进一步整理得微分方程组:

$$\begin{cases} \dot{x}_{1,j} = k_{i1,j} (U_{\text{dc},j} - U_{\text{dc,ref},j}), \\ \dot{x}_{2,j} = k_{i2,j} (U_j - U_{\text{ref},j}), \\ \dot{U}_{\text{dc},j} = (P_{\text{inj}} - P_{e,j}) / (C_j U_{\text{dc},j}), \\ \dot{\theta}_{\text{pll},j} = k_{p4,j} \sin(\theta_j - \theta_{\text{pll},j}) + x_{\text{pll},j}, \\ \dot{x}_{\text{pll},j} = k_{i4,j} \sin(\theta_j - \theta_{\text{pll},j}), \end{cases} \quad (26)$$

及VSC设备的输出电流方程:

$$\begin{cases} i_{di} = k_{p1,j} (U_{\text{dc},j} - U_{\text{dc,ref},j}) + x_{1,j}, \\ i_{qi} = k_{p2,j} (U_j - U_{\text{ref},j}) + x_{2,j}, \end{cases} \quad (27)$$

其中 $x_{1,j}$ 和 $x_{2,j}$ 为电压控制环的中间变量.

因此VSC向网络中注入的有功功率和无功功率为:

$$\begin{aligned} P_{e,j} &= i_{dj} u_{dj} + i_{qj} u_{qj} \\ &= [k_{p1,j} (U_{\text{dc},j} - U_{\text{dc,ref},j}) + x_{1,j}] U_j \cos(\theta_j - \theta_{\text{pll},j}) \\ &\quad + [k_{p2,j} (U_j - U_{\text{ref},j}) + x_{2,j}] U_j \sin(\theta_j - \theta_{\text{pll},j}), \end{aligned} \quad (28)$$

$$\begin{aligned} Q_{e,j} &= i_{dj} u_{qj} - i_{qj} u_{dj} \\ &= [k_{p1,j} (U_{\text{dc},j} - U_{\text{dc,ref},j}) + x_{1,j}] U_j \sin(\theta_j - \theta_{\text{pll},j}) \\ &\quad - [k_{p2,j} (U_j - U_{\text{ref},j}) + x_{2,j}] U_j \cos(\theta_j - \theta_{\text{pll},j}). \end{aligned} \quad (29)$$

结合潮流方程(13), 可构建VSC与同步机多机耦合系统模型. 图10(b)展示了该模型中微分代数方程组之间变量的传递关系. 显然在多机模型中, VSC设备仍然表现出基本的功率交换的形式. 由此可看出VSC与同步机在模型上的统一性, 但同时也可看出, 同步机相关方程之间的变量传递关系相对简单, 非线性项单一, 而对比VSC相关微分方程与代数方程耦合关系强, 非线性项复杂. 据作者们了解, 现在还没有电力电子化电力系统多装备的非线性统一模型, 以及进一步的多电力电子系

统非线性分析方法, 相应的多机暂态分析的研究仍在探索阶段.

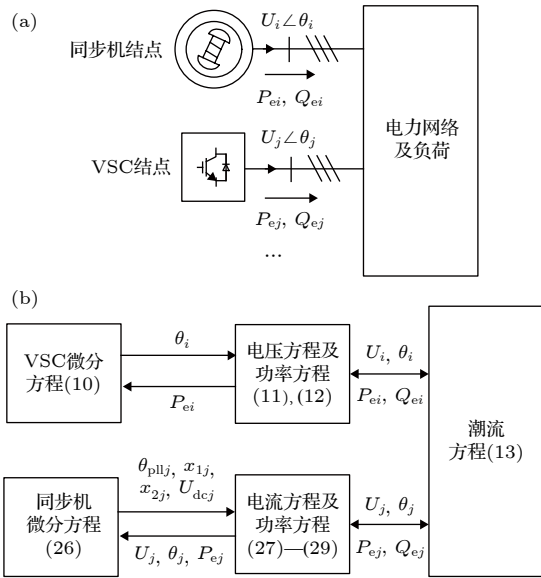


图 10 (a)VSC 与同步机多机耦合系统和 (b) 其微分代数方程组中变量传递关系

Fig. 10. (a) A multi-machine power system with VSC devices and synchronous generators and (b) its variable relations in differential algebraic equations.

4 讨论

电力系统作为目前最复杂的人造系统, 具有非线性、多时标、复杂性的物理本质^[51,52]. 电力电子设备控制方式的多样化及其多机复杂相互作用使得这些性质表现更加明显. 电力电子化电力系统暂态与传统电力系统暂态的区别可初步总结在表 1 中.

传统电力系统中节点电压与注入电磁功率之间关系 (潮流方程) 是最主要的非线性环节. 而在电力电子化电力系统中, PLL 同步方式、保护控制的切换和限幅与坐标变换等非线性环节经过反馈控制环路相互交织在一起, 使得非线性作用关系变得极为复杂. 本文中得到了只考虑电压时间尺度下的五阶微分方程, 从非线性动力学的角度来看已经非常复杂, 其中包括了十余个非线性函数. 其显然不同于非线性物理界熟知的低阶非线性系统, 如洛伦兹振子、卢瑟振子、达芬振子等. 电力电子化电

力系统的非线性环节的处理将非常棘手, 为进一步理论分析带来了巨大困难. 另外, 需强调的, 本文只是针对并网变换器这一小模块在通常矢量控制方式下开展了建模分析, 并未考虑更为复杂的限幅环节、附加暂态辅助控制、风机的机侧变换器、风机的机械转子动态等.

传统电力系统中同步机转子的惯性时间常数远大于电磁元件, 如电感的时间常数, 由此可划分出独立的机电时间尺度与电磁时间尺度. 在通常电力系统动态分析中, 研究人员可以集中针对机电时间尺度下的动力学行为来研究. 由此, 摇摆方程才在传统电力系统动态分析中处于中心地位. 而与之相对应, 电力电子化电力系统级联矢量控制的多样化控制器参数和其控制的物理量共同决定了时间尺度的划分, 如电流控制环与交流电感共同决定了电流时间尺度动态, 电压控制环与直流电容共同决定了电压时间尺度动态, 对于双馈风机还有转速控制器与机械转子共同决定的机电时间尺度等. 这些不同时间尺度受控制参数与系统参数的影响, 相互作用明显. 如何做多时间尺度分离, 或者在什么条件下可以这样做都是值得研究的课题. 当前我们还未有普遍接受的类似于同步发电机的不同阶数的实用动力学模型. 另外值得指出的, 图 10 中的多 VSC 与同步机模型可以看成是传统电力系统暂态模型在电力电子化电力系统中的推广, 具有普遍适用性, 但其还只是适用于电压控制时间尺度下的动态分析, 如果要包含考虑更快的电流时间尺度动态, 不光 VSC 模型部分将变得更为复杂, 原有网络部分准稳态代数模型也将用详细的电磁暂态微分模型来替代, 由此将造成理论分析上的更大的困难. 如何克服电力电子化电力系统动态网络部分的分析问题无疑是一大核心挑战.

同步发电机在受扰后, 表现出由转子运动方程决定的二阶特性, 考虑多机系统, 虽然系统阶数也随着同步机数量的增多而增多, 但是各同步机的特性相似, 相互耦合作用也较为清晰. 而在电力电子化系统中, 装备受到故障扰动后多时标控制器状态量间关系紧密, 系统响应特性将明显区别于二阶系统. 虽然当前有不少研究工作将电力电子装备动态

表 1 电力电子化系统与传统电力系统暂态问题初步比较

Table 1. Comparison of transient problems between Power-electronic-based power systems and traditional power systems.

	复杂性	时间尺度	系统阶数
传统电力系统	非线性项较少且单一	机电与电磁时间尺度较好分离	同步机二阶模型
电力电子化电力系统	非线性项分布广泛且复杂	多时间尺度间强耦合	装备多样化且高阶

特性类比于人们所熟悉的二阶线性动态系统, 如只针对其中的锁相环动态, 或直流电容上的电压动态等, 从而得到简化模型. 但是, 电力电子装备本质上都是由控制器所决定的高维动力系统, 这样简化是值得怀疑的. 而另一方面, 考虑到多机系统时的高维系统特性将更加明显, 若不经任何简化, 又往往会面临维数灾难. 举一个典型的例子, 一个风电场通常由上百台双馈风机或直驱风机所构成, 如何分析其动态相互作用以及可能的振荡无疑是非常有挑战的问题.

综上, 电力电子化电力系统是一个非线性、多时标、复杂性特征明显的高维动力系统, 具有多样化设备跨尺度相互作用的特点. 类似的复杂系统问题也普遍存在于如机械振动分析、气象预报、地震监测、生物神经网络、人类经济行为等众多复杂性学科之中, 不同时间和空间尺度上节点的多样性和其复杂耦合相互作用关系是二十一世纪复杂系统学科的共同挑战. 由此急需尝试借鉴复杂系统等学科的研究成果, 从物理机理出发, 探索电力电子化电力系统稳定性分析和控制方法, 建立起新的动态研究的框架和体系. 本文抛砖引玉, 希望能够得到复杂系统和统计物理研究同行的关注, 共同开展电力网络与复杂系统的交叉研究, 为国家的能源转型贡献一份力量.

附录 A

状态方程 (23) 式中 $f_2 - f_{11}$ 的表达式, 其中 f_1 是 $f_2 - f_{11}$ 中的因式, 未直接出现在状态方程中.

$$f_1 = \frac{1}{1 + X_g k_{p2} \cos(\theta - \theta_{pll})}, \quad f_2 = \frac{f_1 X_g k_{p1} \sin(\theta - \theta_{pll})}{CU_{dc}},$$

$$f_3 = f_1 X_g k_{i1} \sin(\theta - \theta_{pll}), \quad f_4 = -f_1 X_g k_{i2} \cos(\theta - \theta_{pll}),$$

$$f_5 = -f_1 U_g k_{p4} \sin \theta, \quad f_6 = -f_1 U_g k_{i4} \sin \theta,$$

$$f_7 = \frac{X_g k_{p1}}{U C U_{dc}} [\cos(\theta - \theta_{pll}) + f_1 X_g \sin^2(\theta - \theta_{pll})],$$

$$f_8 = \frac{X_g k_{i1}}{U} [\cos(\theta - \theta_{pll}) + f_1 X_g \sin^2(\theta - \theta_{pll})],$$

$$f_9 = f_1 \frac{X_g}{U} [X_g k_{p2} k_{i2} \sin(\theta - \theta_{pll}) \cos(\theta - \theta_{pll}) - k_{i2} \sin(\theta - \theta_{pll})],$$

$$f_{10} = \left[1 - \frac{X_g}{U} f_1 k_{p2} \sin(\theta - \theta_{pll}) - \frac{U_g}{U} \cos \theta \right] k_{p4},$$

$$f_{11} = \left[1 - \frac{X_g}{U} f_1 k_{p2} \sin(\theta - \theta_{pll}) - \frac{U_g}{U} \cos \theta \right] k_{i4}.$$

附录 B

仿真参数如下.

VSC 系统参数: $S_{base} = 2$ MVA, $U_{base} = 690$ V, $f_{base} = 50$ Hz, $\omega_{base} = 2\pi f_{base}$, $U_{dc base} = 1400$ V, $U_{dc ref} = 1$, $C = 0.1$ F, $X_f = 0.1$, $U_{ref} = 1$, $U_g = 1$, and $X_g = 0.5$, where SCR = 2.

控制参数: 1) 直流电压控制 $k_{p1} = 3.5$, $k_{i1} = 140$, 端电压控制 $k_{p2} = 1$, $k_{i2} = 100$; 2) 电流控制 $k_{p3} = 0.3$, $k_{i3} = 160$; 3) PLL $k_{p4} = 50$, $k_{i4} = 2000$.

工作点: $P_{e0} = 0.8$, $Q_{e0} = 0.167$, $U_0 = 1$, and $\theta_0 = 0.412$ rad.

同步机参数: $S_{base} = 200$ MVA, $U_{base} = 13.8$ kV, $f_{base} = 50$ Hz, $\omega_{base} = 2\pi f_{base}$, $x_d = 1$, $x'_d = 0.3$, $x''_d = 0.21$, $x_q = 0.6$, $x''_q = 0.31$, $x_l = 0.15$, $r_s = 0.005$, $H = 3.2$, $D = 0.1$, $T'_d = 1.64$, $T''_d = 0.34$, $T'_{q0} = 1.4$.

参考文献

- [1] Zhou X X, Chen S Y, Lu Z X, Huang Y H, Ma S C 2018 *Proceedings of the CSEE* **38** 1893 (in Chinese) [周孝信, 陈树勇, 鲁宗相, 黄彦浩, 马士聪, 赵强 2018 *电机工程学报* **38** 1893]
- [2] Yuan X M, Cheng S J, Hu J B 2016 *Proceedings of the CSEE* **36** 5145 (in Chinese) [袁小明, 程时杰, 胡家兵 2016 *电机工程学报* **36** 5145]
- [3] Hu J B, Yuan X M, Cheng S J 2019 *Proceedings of the CSEE* **39** 5457 (in Chinese) [胡家兵, 袁小明, 程时杰 2019 *电机工程学报* **39** 5457]
- [4] Biaabjerg F, Ma K 2013 *IEEE J. Emerging Sel. Top. Power Electron.* **1** 139
- [5] Biaabjerg F, Teodorescu R, Liserre M, Timbus A V 2006 *IEEE Trans. Ind. Electron.* **53** 1398
- [6] National Grid ESO LFDD 09/08/2019 Incident Report <https://www.nationalgrideso.com/document/152346/download> [2019-11-10]
- [7] Kundur P, Balu N J, Lauby M G 1994 *Power System Stability and Control* (Vol. 1) (New York: McGraw-hill) p827
- [8] Anderson P M, Fouad A A 2002 *Power System Control and Stability* (Vol. 2) (Boloken: John Wiley & Sons) p13
- [9] Fu S T, Ni Y Y, Xue Y S 1999 *Direct Method Stability Analysis* (Vol. 1) (Beijing: China Electric Power Press) p25 (in Chinese) [傅书遒, 倪以信, 薛禹胜 1999 直接法稳定分析(北京: 中国电力出版社) 第25页]
- [10] Liu S, Chen C 2014 *Energy Function Analysis of Power Aystem Transient Atability: Network Structure Maintenance Model* (Vol. 1) (Beijing: Science Press) p75 (in Chinese) [刘笙, 陈陈 2014 电力系统暂态稳定的能量函数分析: 网络结构保持模型(北京: 科学出版社) 第75页]
- [11] Chiang H D 2010 *Direct Methods for Stability Analysis of Electric Power Systems* (Vol. 1) (Beijing: Science Press) p60
- [12] Kakimoto N, Ohsawa Y, Hayashi M 1978 *Trans. IEE of Japan* **98** 62
- [13] Llamas A, Lopez J, Mili L, Phadke A, Thorp J 1995 *IEEE Trans. Power Syst* **10** 210
- [14] Xue Y, Cutsem T V, Ribbens-Pavella V 1988 *IEEE Trans.*

Power Syst. **3** 400

- [15] Kalcon G O, Adam G P, Anaya-Lara O, Lo S, Uhlen K 2012 *IEEE Trans. Power Syst.* **27** 1818
- [16] Harnefors L, Bongiorno M, Lundberg S 2007 *IEEE Trans. Ind. Electron.* **54** 3323
- [17] Sun J 2011 *IEEE Trans. Power Electron.* **26** 3075
- [18] Rygg A, Molinas M, Zhang C, Cai X 2016 *IEEE J. Emerging Sel. Top. Power Electron.* **4** 1383
- [19] Huang Y, Yuan X, Hu J, Zhou P 2015 *IEEE J. Emerging Sel. Top. Power Electron.* **3** 1193
- [20] Yuan H, Yuan X, Hu J 2017 *IEEE Trans. Power Syst.* **32** 3981
- [21] Yang Z, Mei C, Chen S, Zhan M 2020 *IEEE J. Emerging Sel. Top. Power Electron.* (in press)
- [22] Kabalan M, Singh P, Niebur D 2016 *IEEE Trans. Smart Grid* **8** 2287
- [23] Wu H, Wang X 2018 *IEEE Trans. Ind. Electron.* **66** 6473
- [24] Huang L, Xin H, Wang Z, Zhang L, Wu K, Hu J 2017 *IEEE Trans. Smart Grid* **10** 578
- [25] Huang M, Wong S C, Chi K T, Ruan X 2012 *IEEE Trans. Circuits Syst. I, Reg. Papers* **60** 1062
- [26] Huang M, Peng Y, Chi K T, Liu Y, Sun J, Zha X 2017 *IEEE Trans. Power Electron.* **32** 8868
- [27] Zhang C, Cai X, Li Z, Rygg A, Molinas M 2017 *IET Power Electron.* **10** 894
- [28] Sun J, Wang G, Du X, Wang H 2018 *IEEE Trans. Power Electron.* **34** 3025
- [29] Ying J, Yuan X X, Hu J B 2017 *IEEE Trans. Energy Convers.* **32** 1502
- [30] Vittal E, O'Malley M, Keane A 2010 *IEEE Trans. Power Syst.* **25** 433
- [31] Gautam D, Vittal V, Harbour T 2009 *IEEE Trans. Power Syst.* **24** 1426
- [32] Hu Q, Fu L, Ma F, Ji F 2019 *IEEE Trans. Power Syst.* **34** 3220
- [33] Li M J, Yu Z, Xu T, He J B, Wang C, Xie X X, Liu C 2017 *Proceedings of the CSEE* **41** 1035 (in Chinese) [李明节, 于钊, 许涛, 贺静波, 王超, 谢小荣, 刘纯 2017 电网技术 **41** 1035]
- [34] Menck P J, Heitzig J, Kurths J, Schellnhuber H J 2014 *Nat. Commun.* **5** 3969
- [35] Schultz P, Heitzig J, Kurths J 2014 *New J. Phys.* **16** 125001
- [36] Rohden M, Sorge A, Timme M, Witthaut D 2012 *Phys. Rev. Lett.* **109** 064101
- [37] Schafer B, Witthaut D, Timme M, Latora V 2018 *Nat. Commun.* **9** 1975
- [38] Motter A E, Myers S A, Anghel M, Nishikawa T 2013 *Nat. Phys.* **9** 191
- [39] Yang Y, Nishikawa T, Motter A E 2017 *Science* **358** 3184
- [40] Dorfler F, Chertkov M, Bullo F 2013 *PNAS* **110** 2005
- [41] Grob D, Arghir C, Dorfler F 2018 *Automatica* **90** 248
- [42] Carareto R, Baptista M S, Grebogi C 2013 *Commun. Nonlinear Sci. Numer. Simul.* **18** 1035
- [43] Ma J, Sun Y, Yuan X, Kurths J, Zhan M 2016 *PLoS ONE* **11** e0165943
- [44] Sun Y, Kurths J, and Zhan M 2017 *CHAOS* **27** 083116
- [45] Y Sun, J Ma, J Kurths, M Zhan 2018 *Proc. R. Soc. London, Ser. A* **474** 20170733
- [46] M He, W He, J Hu, X Yuan, M Zhan 2019 *Nonlinear Dynamics* **95** 1965
- [47] Q Yue, J Yu, Zhan M 2019 *IET RPG 2019 International Conference on Renewable Power Generation Shanghai, China*, October 24–25, 2019 p294
- [48] Qiu Q, Ma R, Kurths J, Zhan M 2020 *Chaos* **30** 013110
- [49] Yang Z Q, Ma R, Chen S, Zhan M 2020 *IEEE J. Emerging Sel. Top. Power Electron.* (in press)
- [50] Ma R, Yang Z, Cheng S, Zhan M 2020 Sustained Oscillations and Bifurcations in Three-phase VSC Tied to AC Grid (submitted) *IET Renewable Power Generation*
- [51] Strogatz S H 2018 *Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry, and Engineering* (Vol. 2) (Boulder: Westview Press) p241
- [52] Kuehn C 2015 *Multiple Time Scale Dynamics* (Vol. 1) (Berlin: Springer International Publishing) p53

SPECIAL TOPIC—Statistical physics and complex systems

Problems and challenges of power-electronic-based power system stability: A case study of transient stability comparison*

Yang Zi-Qian Ma Rui Cheng Shi-Jie Zhan Meng[†]*(State Key Laboratory of Advanced Electromagnetic Engineering and Technology, Hubei Electric Power Security and**High Efficiency Key Laboratory, School of Electrical and Electronic Engineering, Huazhong**University of Science and Technology, Wuhan 430074, China)*

(Received 24 December 2019; revised manuscript received 19 March 2020)

Abstract

With the development of power electronic technology and requirement for clean energy, the traditional power systems which are dominated by synchronous generators are gradually changing into the power-electronic-based power systems with diversified power electronic equipment. The power systems are facing a great revolution in their primary equipment, and this has not happened in the past one hundred years. In recent years, with great increasing penetration of power electronic devices into power grids, the large-scale blackouts caused by power electronic devices have been reported, which seriously threatens the safe and stable operation of power systems. Under the above background, in this paper we first introduce several methods of analyzing the traditional power system transient stability from the equal area criterion for the single machine infinite bus system to several Lyapunov function based direct methods for multi-machine systems. Then we introduce some of our recent work on the nonlinear modeling and analysis of a key component of power-electronic-based power systems, voltage source converter (VSC), and propose a multiple machine system model including power electronic equipment and traditional synchronous machines. Finally, we illustrate the transient characteristics of the power electronic devices, and summarize the basic problems and challenges for the transient stability of power-electronic-based power systems. We hope that these basic problems in power-electronic-based power system dynamics including nonlinearity, multi-time-scale, and complexity could arouse the general interest of researchers in the fields of complex systems and statistical mechanics.

Keywords: power-electronic-based power system dynamics, transient stability analysis, grid-connected converter, complex system dynamics

PACS: 89.75.-k, 05.45.-a, 84.70.+p, 84.30.Jc

DOI: [10.7498/aps.69.20191954](https://doi.org/10.7498/aps.69.20191954)

* Project supported by the National Key Research and Development Program of China (Grant No. 2017YFB0902000), the Science and Technology Project of State Grid, China (Grant No. SGXJ0000KXJS1700841), and the International (Regional) Cooperation and Exchange Program of the National Natural Science Foundation of China (Grant No. 11861131011).

[†] Corresponding author. E-mail: zhanmeng@hust.edu.cn

网络直播平台数据挖掘与行为分析综述

郭淑慧 吕欣

Live streaming: Data mining and behavior analysis

Guo Shu-Hui Lu Xin

引用信息 Citation: *Acta Physica Sinica*, 69, 088908 (2020) DOI: 10.7498/aps.69.20191776

在线阅读 View online: <https://doi.org/10.7498/aps.69.20191776>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

基于本征正交分解的气动优化设计外形数据挖掘

Proper orthogonal decomposition-based data mining of aerodynamic shape for design optimization

物理学报. 2017, 66(22): 220203 <https://doi.org/10.7498/aps.66.220203>

基于Hadoop大数据平台和无简并高维离散超混沌系统的加密算法

Encryption algorithm based on Hadoop and non-degenerate high-dimensional discrete hyperchaotic system

物理学报. 2017, 66(23): 230503 <https://doi.org/10.7498/aps.66.230503>

人脑默认模式网络的动力学行为

Dynamics of the default mode network in human brain

物理学报. 2020, 69(8): 080203 <https://doi.org/10.7498/aps.69.20200170>

相对性区域创新指数与经济周期挖掘

Relativistic regional innovation index and novel business cycle

物理学报. 2020, 69(8): 088905 <https://doi.org/10.7498/aps.69.20191970>

光频梳频域干涉测距主要参数分析及一种改进的数据处理方法

Analysis of main parameters of spectral interferometry ranging using optical frequency comb and an improved data processing method

物理学报. 2019, 68(19): 190602 <https://doi.org/10.7498/aps.68.20190836>

专题：统计物理和复杂系统

网络直播平台数据挖掘与行为分析综述*

郭淑慧 吕欣†

(国防科技大学系统工程学院, 长沙 410073)

(2019年11月22日收到; 2020年3月20日收到修改稿)

随着移动通信和互联网技术的不断发展,网络直播逐渐成为了新媒体环境下人们青睐的在线娱乐和信息传播方式.目前广泛应用于课堂教学、真人秀、电竞赛事、品牌营销等方面.数百万主播与数亿计观众的活跃加入和互动,产生了丰富的在线人群行为活动数据,为开展大规模人群行为动力学、平台内容推荐与检测、在线社群演化等研究提供了丰富的实验场景.本文通过梳理国内外网络直播平台数据挖掘与行为分析的相关研究文献,分析了直播平台负载水平、观众行为、主播行为以及社群网络的特征和变化规律,并对直播平台中大规模人群行为表现出的时空规律和重尾效应进行了总结.直播平台中各种社群网络的形成和演化机制、内容推荐与检测等是未来网络直播领域研究的发展趋势.

关键词: 直播平台, 用户行为, 社群网络, 数据挖掘

PACS: 89.70.-a, 89.75.-k, 89.75.Kd

DOI: 10.7498/aps.69.20191776

1 引言

网络直播是一种新型信息交流方式,可以让观众收看到主播所处场景中正在进行的音、视频实况,同时观众可以通过打赏或发表评论的方式与主播进行互动,相对于传统信息传播媒体来说有着互动性强、时空适应性强等优势.网络直播平台则是由公司或组织管理的供主播发布直播以及观众收看直播的网站.随着互联网经济的发展,网络直播日益火爆,斗鱼TV、虎牙TV、抖音等网络直播平台在市场中异军突起,以超低的门槛吸引了大量主播和观众,截至2019年6月,我国直播用户规模已达4.33亿^[1].

网络直播目前的应用领域比较广泛,除了应用于娱乐性的真人秀、电竞赛事之外,还有课堂教学^[2,3]、品牌营销^[4,5]、传统文化与工艺技术传承^[6,7]、政务会议与庭审过程公开^[8,9]等方面.不同领域的网络

直播和观众都会产生大量交互数据,这些数据一方面可以用于挖掘直播平台的负载变化模式和用户参与及交互的内在机制,探究用户行为和偏好,进而对相应情境下的大规模人群行为进行模式分析和规律挖掘^[10-17].另一方面还可以基于直播平台负载的测量结果及从中挖掘的用户行为的特征和偏好,提升网络直播平台内容推荐和内容检测水平^[18-22].

本文从国内外网络直播平台用户行为数据挖掘的研究入手,对直播平台负载水平、观众行为、主播行为以及社群网络的特征和变化规律进行梳理和总结,并讨论网络直播平台研究在当前面临的问题和未来的研究方向.

2 直播平台负载研究

网络直播是通过网络直播平台进行实时信息传输的新媒体形式.随着网络的发展和普及,网络

* 国家自然科学基金(批准号: 82041020, 71771213, 91846301, 71790615)和湖南省科技计划项目(批准号: 2017RS3040, 2018JJ1034)资助的课题.

† 通信作者. E-mail: xin.lu@flowminder.org

直播作为一种学习、娱乐的便捷资源被人们越来越广泛地使用. 直播平台负载的水平差异反映了直播平台用户的分布规律和使用偏好, 对直播平台的负载研究可以从整体上把握直播平台的资源消耗与服务使用情况, 对直播平台优化资源配置、提供经济稳定的负载支持有指导意义.

目前关于平台负载研究主要是通过统计直播平台的运行负载, 挖掘负载水平产生规律性差异的时间因素、空间因素以及其他影响因素, 总结直播平台情境下的大规模人群行为偏好和行为特征.

2.1 时序特征

受时间节律的影响, 人类行为会在诸多方面不同程度地体现出日内效应 (diurnal effect)、周内效应 (weekly effect) 等时序规律, 如金融市场的流动性^[23]、人类的情绪积极程度^[24]、反应灵敏度^[25]、器官工作机能^[26]等都会在一天内不同时段表现出显著差异, 股市收益率和波动还存在明显的周内效应^[27]. 目前对直播平台的负载研究大部分集中于从系统带宽、主播规模、观众数量、打赏额和评论量等方面的时序变化规律中挖掘直播平台负载的日内效应、周内效应和长期规律等时序特征.

Veloso 等^[11]最早根据巴西某网络电视直播平台的网站日志对负载的时序特征进行了研究. 在 2002 年为期 28 天的统计中, 用户的访问模式显示出了明显的日内效应和周内效应, 昼夜模式造成凌晨 4:00—11:00 在线观众数量偏低, 峰值和谷值分别在 3:00 和 9:00 附近取得; 双休日的平均观众数量明显高于工作日. 尽管该直播平台的用户规模超过 69 万、覆盖 65 个国家, 但受当时的网络发展水平限制, 平台使用的带宽峰值仅仅为 80 Mbps.

随着 Twitch, YouTube Live 等专门化网络直播平台的兴起, Kaytoue 等^[28]根据 2011 年末 Twitch 平台的直播间数量和在线观众数量变化对直播平台负载的周内效应进行分析, 发现 Twitch 平台的双休日负载明显高于工作日 (后续学者^[29,30]也得出了—致结论), 原因是 Twitch 平台的主要直播内容是电子竞技, 而大型电子竞技的竞赛通常在双休日举行. Pires 和 Simon^[31]对比了 Twitch 平台和 YouTube Live 平台在 2014 年的系统带宽和直播间数量发现, 两个平台的带宽峰值都超过了 1 Tbps, 但 Twitch 平台的带宽水平更高, 峰值超过了 1.6 Tbps. 两个平台在直播间的数量在

双休日都明显高于工作日, 但 Twitch 平台和 YouTube Live 的日内负载峰值分别在 5:00 和 18:00 附近取得, 而且 Twitch 平台在日内和周内负载变化模式的敏感度均低于 YouTube Live. 原因是 Twitch 平台的开放时间较早, 用户在全球范围内的覆盖范围更广, 减弱了昼夜更替造成的时序差异.

近年来逐渐出现了对国内直播平台负载的研究, Zhu 等^[32]通过收集国内直播平台斗鱼 TV 在 2016 年 12 月为期 14 天的直播数据, 发现观众数量和主播数量在一天中显示出几乎一致的变化规律, 都在 21:00—8:00 减少, 8:00—21:00 增加, 在晚上 9:00—10:00 达到最高水平. Wang 等^[33]通过分析 2016 年 9 月起为期 124 天内的斗鱼 TV 主播开播数量、观众打赏总额和评论量来挖掘斗鱼 TV 平台负载的日内效应和周日效应, 从观众的打赏总额、评论量和主播直播次数、直播时长分别展现观众和主播在一天之中的活跃程度变化趋势. 结果发现观众和主播的活动都表现出很强的昼夜规律, 并且高度同步, 任意两个数据系列之间的皮尔逊相关系数都高于 0.85. 但观众活动的高峰时段出现在 23:00—24:00, 主播最活跃的时段是 20:00—22:00, 说明观众的活跃时间存在一定的时滞现象, 与之前的研究结论^[32]略有差异.

总体上说, 国内外直播平台负载在一天之中都呈现“倒 N 型”^[34], 直播平台负载具有明显的“日内效应”^[25], 负载水平在一天中呈现降低-升高-降低的循环模式 (如图 1 所示).

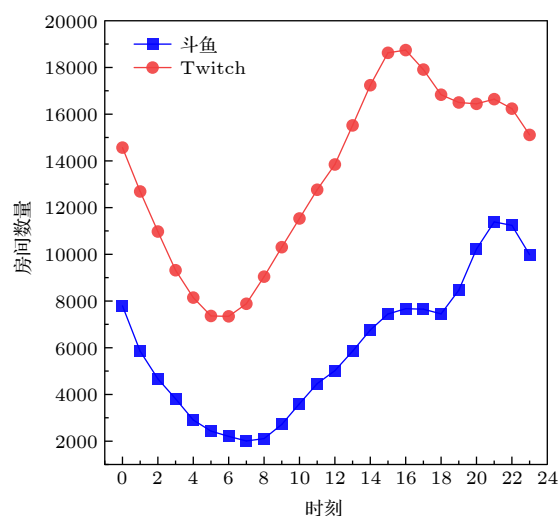


图 1 国内外直播平台负载日内效应^[30,32]

Fig. 1. The diurnal effects of domestic and foreign live streaming workloads^[30,32].

表 1 不同直播平台的负载时序特征

Table 1. The workload changes of different live streaming platforms.

直播平台	年份	日内效应	周内效应
Anonymous ^[11]	2002	3, 19时观众数量最多	双休日观众数量更多
Twitch ^[30]	2015	17, 0时直播间最多	双休日直播间数量显著增多
YouNow ^[35]	2015	直播间和观众最多: 22时	双休日直播间和观众数量更多
斗鱼 ^[32]	2016	直播间和观众最多: 21—22时	—
斗鱼 ^[33]	2018	直播间最多: 20—21时, 观众最多: 23—24时	双休日直播间和观众数量更多

但由于直播平台的直播类型各有侧重, 而且直播平台随着时间在逐渐发展, 即使是同一直播平台在不同的统计期内表现的平台负载时序规律也不完全相同. 多个直播平台、不同时期的负载变化时序特征见表 1.

2.2 空间分布特征

除了分析直播平台负载的时序特征之外, 部分研究通过分析主播、观众和直播平台服务器的位置对直播平台负载产生的影响来挖掘直播平台负载的地理分布特征, 进而对直播负载的资源分配及优化进行指导.

Veloso 等^[11]对早期网络电视直播平台的观众数量和观看次数在所覆盖区域间的数量分布进行了统计分析, 发现观众数量和观看次数在划分的地理区块之间的分布形式都近似 Zipf 分布^[36]:

$$Z(r) = Ar^{-\alpha}, \quad (1)$$

其中参数 r 代表地理区块按照观众数量或观看次数降序排列的排名, $Z(r)$ 则代表排名为 r 的区域中的观众数量或者观看次数, α 的取值分别为 1.29 和 1.49, 展现了早期网络电视直播的观众在地理分布上的不均匀现象. 与之相对的, Li 等^[37]统计分析了 PPTV 直播频道的观众地理分布熵 (viewer geographical entropy) 的累积分布函数, 发现大多数频道的观众地理分布熵超过 0.8, 显示了现代网络直播平台观众观看者在地理位置上均匀分布. 观众地理分布熵形式为

$$e_k = \frac{1}{\log N} \left(- \sum_{i=1}^N p_{ki} \log p_{ki} \right), \quad (2)$$

e_k 是直播间 k 的观众地理分布熵, 其中 N 是直播间全部观众所覆盖区域的数量, p_{ki} 代表直播间 k 在区域 i 的观众数量占全部观众数量的比例.

Kaytoue 等^[28]通过统计分析 Twitch 平台的主播在不同时区的分布, 表明平台中的大多数主播

都来自北美、欧洲和东亚, 与 Twitch 平台的服务器集中布置在北美、欧洲和亚洲的分布规律^[38]相符合, 反映了直播平台负载的地理特征对直播平台服务器设置的指导意义.

Yan 等^[39]对比了用户生成视频、短视频和直播视频三种平台的城市、郊区及整个区域的移动网络用户的观众地理分布熵, 结果显示无论市区、郊区还是整个区域, 用户生成视频的熵都高于其他服务, 表明观看直播的用户比观看用户生成视频的用户在空间分布上更不均匀, 主要原因是直播内容通常耗费的流量更多, 用户更倾向于在固定的场所使用 WiFi 或宽带网络而不是移动流量来观看直播.

2.3 其他影响因素

除了时空对直播平台负载的影响之外, 少量学者对主要直播电子竞技内容的直播平台 Twitch、斗鱼 TV 的负载水平是否受到大型电子竞技竞赛项目直播的影响进行了研究. Kaytoue 等^[28]通过观察 Twitch 平台在 2011 年 12 月 29 日到 2012 年 1 月 9 日每天的观众数量变化情况, 发现在直播一些重要的电子竞技比赛时, 观众数量会出现明显增加, 说明了电子竞技竞赛项目直播对 Twitch 平台的负载有强烈的刺激作用. 类似地, Deng 等^[12]统计了 Twitch 平台电子竞技竞赛项目直播吸引的观众占整个平台观众的比例, 结果显示某些热门的电子竞技竞赛项目直播所产生的观众数量能占直播平台全部观众的 30% 以上, 即使是不太流行的电子竞技赛事的直播也能吸引大量观众, 峰值超过全平台观众 10%. 但 Wang 等^[33]在对斗鱼 TV 观众评论数以及打赏额在 2016 年为期 124 天的统计期中的变化中却并未发现重大赛事对平台负载产生的显著影响, 原因可能是斗鱼 TV 存在大部分娱乐类直播, 受电子竞技竞赛项目直播的影响并不明显.

3 观众行为分析

直播平台内可能出现以下的一种或几种观众行为: 选择直播间进行观看、切换或退出直播间、在直播间中评论或打赏、以及观众观看直播而引发的行为(如被主播引导购买商品). 众多学者对直播平台观众行为中的观看规律进行分析和建模, 研究观众各种行为以及背后的心理, 对于理解用户参与网络直播的原因、提升用户体验、为用户提供更有价值的网络直播服务有重要的决策价值.

已有研究中对观众的观看规律主要从观众的观看次数与时长、频道选择与切换、观众评论与打赏等方面入手, 从观众的观看记录中提取直播平台中观众的各种观看行为, 挖掘其中观众的偏好及心理动机, 进一步开展直播平台观众的行为动力学建模和社群网络演化研究.

3.1 观看次数与时长

众多研究表明, 直播平台内观众的观看时长、观看次数呈现一定的重尾效应, 即直播平台中存在大量观看次数很少、观看时间非常短的用户, 但同时还有极少量用户观看直播的次数很高、观看时长相对非常长.

Veloso 等^[11]于 2002 年的研究结果显示早期直播平台中观众观看次数分布近似 Zipf 分布, 其中参数 r 代表的是按照观看次数降序排列的观众排名, $Z(r)$ 则代表排名为 r 的观众的观看次数, 参数 α 的取值为 0.719; Li 等^[37]于 2016 年发现 PPTV 平台内观众观看次数分布更符合互补 Weibull 分布而不是幂律分布, 累计概率分布函数形式为

$$P(X \geq x) = \exp[-(x/x_0)^c], \quad (3)$$

其中参数 c 为拉伸因子, x_0 为常数参数.

Sripanidkulchai 等^[40]则于 2004 年对早期直播平台观看时长分布进行了探索, 结果显示不同直播间内观众的观看时长分布均近似 Zipf 分布, 其中参数 r 代表的是按照观看时长降序排列的观众排名, $Z(r)$ 则代表排名为 r 的观众的观看时长, 参数 α 的取值在 0.7—2.0 之间. 而 Tang 等^[41]于 2006 年发现 CCTV 多个直播频道的观众观看时长的概率密度函数形式符合对数正态分布:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right], \quad (4)$$

其中 $f(x)$ 是观看时长为 x 分钟的概率密度函数值, μ 和 σ 的取值均在 4—5 和 1—2 之间. Li 等^[37]于 2016 年对 PPTV 中观众观看时长分布的概率密度函数进行了研究, 结果显示观众观看时长的概率密度函数形式为混合指数分布:

$$f(x) = \sum_{i=1}^n \alpha_i \frac{1}{\mu_i} \exp\left(-\frac{x}{\mu_i}\right), \quad (5)$$

其中 $f(x)$ 是观看时长为 x 分钟的概率密度函数值, μ_i 和 α_i 是第 i 个指数分布的均值和权重, $\sum_{i=1}^n \alpha_i = 1$.

特别地, Tang 等^[41]通过分析观众观看直播的时长记录, 发现观众已经观看直播的时长与继续保持观看的时长存在显著的正相关关系, 即观众如果已经花费了比其他观众长的时间观看直播, 则会更倾向于比其他观众花费更长的时间继续观看直播.

3.2 频道选择与切换行为规律

如果在观看直播的过程中发生网络故障、主播关闭直播、不感兴趣等情况, 观众就有可能对当前直播间进行重新连线、切换到其他直播间或者直接退出直播平台. 目前已有相关研究对观众在观看直播中的重连、切换、退出等进行统计建模与系统分析, 展现直播平台内观众流动的动态过程和内在机理.

Li 等^[37]对客观原因造成直播中断时的观众重连行为进行了研究. 由于观看出现中断的原因可能是网络连接失败等客观问题, 也可能单纯是由于观众的兴趣发生变化而主动退出直播间, 所以作者首先定义了由客观问题造成中断的直播段特征, 是观众在某直播间内发出观看请求之后的一小段时间之内对该直播间重复发出了观看请求. 进一步分别统计了观众对直播中断次数的容忍程度和放弃观看的概率分布, 结果显示随着直播中断次数的增加, 观众放弃的概率递增, 但增幅在逐次减小. 在移动网络下观看直播的观众在直播中断时的耐心程度比在 WiFi 或者宽带环境下的更高, 在遇到 2 次连续的直播中断时, 放弃观看的概率小于 50%, 甚至在某些情况下观众能忍受 10 次连续的直播中断.

Nascimento 等^[42]对 Twitch 平台中的观众切换行为进行了建模分析. 作者首先定义了直播间共存在三种状态, 分别是直播中、直播即将结束和直播结束(如图 2 所示). 由于主播下播之后直播间内的观众并不会被强制清空, 所以会出现主播已经下

播但观众数量不为 0 的情况. 直播间三种状态中的“直播即将结束”包含了主播即将关闭直播以及关闭直播后观众数量仍保持一定水平时期,“直播结束”指的是直播间内观众数量非常低甚至为 0 的状态.

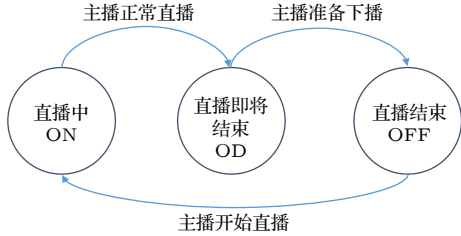


图 2 直播间状态演化图^[42]

Fig. 2. The graph of live streaming channel's state dynamics^[42].

由直播间的三种状态任意组合可以归纳出观众切换行为共包含 9 种类型, 通过统计观众不同类型切换行为的比例发现观众在频道之间的切换行为中, 大约 20% 的 ON-ON 切换 (即从一个正在直播的直播间切换到另一个正在直播的直播间) 和 30% 的 OFF-OFF 切换 (即从一个未开播的直播间切换到另一个未开播的直播间) 持续不到 1 min, 表明观众进入直播间时并不事先知道他们想观看的内容. 77% 的切换行为是 ON-X (X 可能是 ON 或者 OFF), 这说明大部分的直播间切换是由于观众对切换前直播间的内容不满意. 通过对比观众在 ON-ON 切换前后的观看时长的分布, 发现观众在切换后的直播间会观看更长的时间, 同样验证了大部分观众是主动退出了之前不满意的直播间. 由于观众对直播间选择的随意性较强、满意度较低, 部分研究^[18–20] 设计了针对直播平台的内容推荐算法及系统, 为观众选择直播频道进行个性化推荐.

Li 等^[37] 对直播节目中观众的进入和退出过程进行了建模和分析. 通过统计直播过程中观众加入和离开的速率变化, 发现在直播节目开始之前的一小段时间内会出现观众大量进入和立刻退出的现象, 而且观众的加入和离开速率在很大程度上依赖于某些子事件的发生. 考虑到以上因素, 作者使用高斯径向基函数 (RBFs) 之和来拟合直播过程中观众加入和离开行为, 较低的 RMSE 值表明模型与实际过程相符. 形如

$$\lambda(t) = \sum_{i=1}^m a_i \exp \left[\frac{-(t - T_i)^2}{2\sigma_i^2} \right], \quad (6)$$

其中 T_i 是第 i 个子事件发生的时刻, a_i 是子事件 i 的影响权重.

3.3 评论与打赏行为分析

观众向直播间赠送虚拟礼物被称之为“打赏”, 已有研究对国内直播平台内的打赏金额分布规律进行挖掘. Zhao 等^[43] 统计了映客平台内某些主播收到的打赏, 结果显示观众打赏的金额服从长尾分布. 也就是说, 极少比例的观众贡献了大部分打赏, 贡献排名前 20 名的观众的打赏额占有所有观众打赏的 90% 以上. 相似地, Zhu 等^[32] 通过统计斗鱼 TV 中打赏额在主播之间的分布情况, 发现打赏额在主播之间分布的幂律性, 几个最受欢迎的主播拥有非常高数量的打赏, 其他频道分享的很少, 而且仅 2.7% 的打赏观众贡献了 80.2% 的打赏额. 兰荣亨等^[44] 则根据观众的观看、评论和打赏记录对观众群体行为进行特征构建, 成功对不同特征的观众群体进行了聚类区分.

Wang 等^[33] 对打赏行为的时间规律进行分析, 发现不论是第一次打赏出现的时间还是打赏间隔时间的分布形式都高度符合 Weibull 分布, 其累计分布函数形为

$$F(t) = \Pr[X \leq t] = 1 - e^{-(t/\lambda)^k}, \quad (7)$$

其中 λ 和 k 是分布的尺度和形状因子. 拟合结果显示两个分布的形状因子 k 均小于 1, 即说明直播平台中主播已经等待打赏的时间越长, 那么后续打赏到来所需的等待时间越长.

观众评论是观众利用文字和表情符号在直播间中进行交流的一种方式, 目前对观众评论的研究主要是对评论的情感、特征、观众交互进行分析. Poyane^[45] 对 Twitch 平台部分直播 Dota2 的直播间内的观众评论文本数据进行了情感分析, 发现随着直播间观众规模的增加, 观众评论的消极色彩会相应增强. 类似地, Nematzadeh 等^[46] 也发现随着观众数量的增加, 评论区会由正常对话向过载的、不和谐对话转变.

Olejniczak^[47] 对 Twitch 平台观众的评论内容从语句特征上进行了分析. 发现观众更倾向于使用大量的表情符号和重复信息来表达态度, 使用新颖的词汇和独特的表情符号来力求与众不同. 由于评论区只显示最新的几条评论内容, 评论长度会随着观众数量增加而缩短. Li 等^[33] 发现了观众评论与

打赏的周内模式存在很强的相关性 (皮尔森相关系数超过 0.85), 且评论与视频内容同步性很强, 提出了根据评论情感标注直播亮点的算法^[48]. 周钰淇^[22]则提出了根据评论内容对直播内容是否合法进行检测的深度学习算法.

4 主播行为分析

直播平台中的主播行为研究主要集中在主播的直播次数、直播时长以及流行度排名与预测三方面. 通过分析主播群体独特的行为模式, 挖掘直播平台内主播直播规律和活动特征, 对进一步开展直播平台内大规模人群行为分析和研究、优化直播平台系统建设有重要意义.

4.1 直播次数

国内外直播平台的主播直播次数普遍呈现出一定程度的幂律分布特点, 即直播次数较少的主播占了很大部分, 直播次数多的主播占比很少^[14,32,35,42].

对国外直播平台的主播直播次数规律的研究中, Stohr 等^[35]挖掘 Younow 平台的主播直播数据发现, 超过 40% 的主播只直播了一次, 约 10% 的主播在一周内直播了 7 次以上. 这表明, 有一小部分高度活跃主播愿意每天直播多次, 而大多数主播仅进行少量的直播. 类似地, Nascimento 等^[42]发现 Twitch 平台的少部分专业主播团队直播的次数达到每天 19 次, 而大量主播 (40%—50%) 每天直播次数不超过 1 次, 与 Jia 等^[14]的结论基本一致.

对国内直播平台的主播直播次数规律的研究中, Zhu 等^[32]通过统计 14 天内斗鱼 TV 主播开播天数的分布, 发现约 63% 的主播每周直播的天数不超过 1 天, 只有 14% 的主播在统计期中至少直播了一半的时间. 大多数主播并不经常直播, 但整个平台每天约有 4% 的主播会进行直播.

4.2 直播时长

众多研究表明主播在直播时长方面表现出重尾分布规律. Zhu 等^[32]通过收集斗鱼 TV 主播在为期 14 天统计期内的所有开播记录来统计主播直播时长分布, 发现 70% 以上的直播时长都小于 200 min, 但存在极少比例 (小于 1%) 的主播直播时长达到 1000 min. 统计结果显示斗鱼 TV 主播的直播时长中位数是 90 min, 比 Twitch 平台的

45 min^[28] 更长, 原因是相对于 Twitch 平台主要针对游戏内容进行直播, 斗鱼 TV 的直播类型更多样, 许多直播间播放已经制作好的视频或大型活动, 提高了直播的持续时间. 类似地, 对 Younow^[35]和 Twitch 平台^[14,30,42,49]的主播直播时长特征的研究结果均显示主播直播时长分布呈现重尾效应.

研究中通常以直播平台内所有主播直播时长的中位数作为衡量该直播平台主播直播时长的指标, 由于直播内容和针对观众等方面的差异, 各个直播平台的直播时长中位数不尽相同. 即使是相同的直播平台, 不同直播类别的直播时长中位数也会有一定的变化. 部分结论总结如表 2 所列.

表 2 各个网络直播平台的直播时长中位数
Table 2. Median live streaming duration of each live streaming platform.

文献	直播平台	采集年份	中位数
[35]	YouNow	2015	16 min
[42]	Twitch(StarCraftII的较大直播间)	2013—2014	3.7 h
[30]	Twitch(点播的视频)	2015	8 min
[49]	Twitch	2014	150 min
[28]	Twitch	2011—2012	95 min
[32]	斗鱼	2016	90 min

4.3 流行度排名和预测

主播的“流行度”指的是主播吸引观众的能力. 通过某些衡量指标对主播吸引力进行排名, 排名越靠前则说明主播吸引观众的能力越强、主播流行度越高. 这一排名在体现主播在直播平台中的地位 and 水平的同时也反映出了观众的访问模式. 目前的研究中通常以粉丝或观众数量、打赏金额、评论数量作为衡量指标来对直播平台主播的流行度进行排序.

大量对主播流行度排名的研究显示, 主播流行度存在一定的重尾分布规律^[12,14,30,35,40,43,45,49,50]. Pires 等^[31,50]研究发现 Twitch 平台的观众数量分布符合 Zipf 分布, 且参数 α 的值在 1.0—1.5 之间变动, 说明大量观众在很少几个直播间中聚集, 少数主播吸引了绝大部分观众. Stohr 等^[35]对 Younow 平台观众数量分布的研究也得出了类似的结论. 但 Zhang 和 Liu^[49]则发现 Twitch 平台的观众数量分布形式不是标准的幂律分布. 由于著名主播通过直播吸引了极大比例的观众观看, 观众数量分布的尾部出现了明显的下降, 更符合 Gamma 分布或者 Weibull 分布的特点, 与 Wang 等^[33]对斗鱼 TV

内评论数量和打赏金额分布形式一致. Arnett 等^[51]则对主播在社交平台上的公开活动对主播流行度是否产生影响进行了研究, 并没有发现主播的观众和粉丝数量变动与在社交平台上的活动存在统计学上显著的相关性.

对主播流行度预测的研究中, Kaytoue 等^[28]分析线上内容发布后短期和长期的流行度相关性, 提出线性回归模型, 以此来通过前期观众数较准确地预测后期观众数. 还提出了一种流行度的定义(不仅仅只比较在线人数, 还考虑上线时间的早晚等因素), 并以此对主播进行新的排序. 基于主播流行度预测线性回归模型, Netzorg 等^[52]提出了基于主播行为的主播未来流行度预测模型, 发现主播的努力行为(如发布更多直播、定期直播、在其他社交媒体账号上发布直播信息等)在提升主播流行度方面是有效的, 而且职业主播比业余主播更受欢迎. 类似地, Szabo 等^[53]用浏览次数代表视频的流行度, 从前期数据预测视频未来长期的流行度. Zhu 等^[32]对直播间出现的总观众数和礼物总价值进行线性相关分析, 计算得出直播间内观众总数和礼物总价相关系数是 0.6421 ($p < 0.001$). Jia 等^[14]计算得出直播间在线人数和主播直播次数的相关系数也较高, 即直播次数多的主播流行度可能更大.

5 社群网络分析

网络直播吸引了大量的主播与观众参与, 直播平台中大规模人群交互形成了很多独具特色的社群现象. 通过统计分析直播用户的使用特征, 识别和发现直播平台内的社群及社群网络, 进一步分析直播社群网络的节点属性、结构特征以及形成、演化过程, 开展对直播平台大规模人群参与、流动及交互的规律挖掘和动力学研究, 对信息传播、网络营销、舆情监测引导等策略的制定等都有参考和指导意义.

5.1 社群发现与成员识别

直播平台具有的社交属性使得其中存在大规模人群的交互关系, 从而形成了多种类型的用户关系复杂网络. 而直播平台用户网络的节点属性、结构和形成演变机制则体现了直播平台情景下大规模人群活动的交互特征和选择偏好. 部分研究对用

户观看、关注、评论、打赏等关系网络中的社区发现^[54]方法进行了探索, 通过识别用户关系网络中相似用户形成的社区为直播平台用户关系网络发展动态的研究奠定基础.

Churchill 和 Xu^[55]于 2016 年发表了首个对直播平台用户社区的研究并提出了社区发现和成员识别算法. 该社区发现算法是通过可视化主播共享观众关系网络实现主播社区识别. 作者首先收集了游戏直播平台 Twitch 的主播信息, 包括主播直播的游戏类型及粉丝列表. 进一步地, 作者构建了以主播为节点, 共享观众关系为边的主播关系网络, 其中节点大小代表了主播拥有的粉丝数量, 颜色代表主播的直播游戏类型, 主播之间共享的粉丝数量越多, 那么连边越粗、节点之间的距离越短. 通过可视化主播关系网络, 根据节点颜色和距离的分布直观分析主播之间联系的紧密程度, 实现主播关系网络中的社区规模和结构的识别.

作者提出的社区成员识别算法则是根据主播与主播之间的关注关系, 自动识别出社区成员. 作者首先人工挑选出实际属于 Twitch 平台三大主流社区^[56]的四位主播作为种子节点, 种子节点的关注者中粉丝量在 28000 以上的则被程序自动判定与该主播所属同一社区, 从而实现了主播所在社区成员的识别和发现. 识别结果与 Gephi 中的模块化识别结果基本一致, 说明了社区发现算法的有效性. 类似地, Lykousas 等^[57]通过设定违规用户作为种子节点, 从 Live.me 平台和 Loops Live 平台的观看关系网络中自动判定用户是否违规, 实现了违规用户所在社区的成员发现.

5.2 社群网络节点重要性

前文总结了直播平台中观众数量、评论量和打赏金额在直播间的分布普遍呈现出重尾效应, 说明以主播为节点, 以观众观看、评论或打赏为节点重要性衡量指标的网络中, 存在少量中心节点, 它们在整个主播社群网络中的地位和重要性非常高. 但由于网络直播发展的时间较短, 目前尚未出现对直播平台主播社群网络的节点性质进行专门分析的文献.

在观众与主播共同形成的社群网络中, 由于观众对主播有天然的选择权利, 社群中的观众成员对主播的喜爱和认可促使主播成为了社群的意见领袖和核心, 研究普遍认为主播促进了整个社群建立

和发展,对社群的发展方向起决定作用. William 等^[58]对 Twitch 平台的主播和观众进行访谈,发现直播间的社群氛围折射了主播的品质和态度,即主播成为了整个直播间的意见领袖.而且直播间内的核心成员发挥吸引其他参与者、促进互动以及缓和聊天的重要作用以建立社群,也即社群中的意见领袖促进了整个社群建立和发展.庄庆玲和周丽^[59]以斗鱼游戏主播为例研究了弹幕式互动直播平台主播和观众之间形成的社群,发现在直播间的互动中主播会成为意见领袖,与追随者也就是观众基本属于同一阶层,拥有共同的兴趣但意见领袖对该领域有更全面和深入的了解,依靠平台进行有偿信息交流.

5.3 社群网络结构特征

研究复杂系统内的社群网络结构有助于理解或预测系统的表现^[60],少量研究对直播社群网络的结构特征进行了描述,通过社群网络的结构性质反应直播社群的交互特点,以解释直播平台中的社群表现.

Churchill 和 Xu^[55]对 Twitch 平台主播社群的网络结构性质进行了研究.发现主播直播的游戏相同或相似,那么他们拥有的粉丝相似度也很高,体现了观众对直播类型的偏好性.而且通过识别 Twitch 平台中三大主流社区的成员,发现社区规模大小与直播难度相关,成为该类主播的难度越大则该类主播数量越少. Lykousas 等^[57]通过对直播平台违规用户所形成的观看网络结构进行统计分析,对网络的平均度、密度和相互性进行分析发现,违规用户网络的互惠性很差 ($\text{reciprocity} < 0.15$),即违规用户社群网络中节点与节点之间在互惠互利方面的表现并不好.

6 讨论与展望

综上所述,网络视频直播用户行为挖掘的国内外相关研究目前已取得一定的进展.从大量相关文献发现,直播平台负载、观众行为、主播行为和社群网络是本领域的研究重点.其中,直播平台负载水平的变化模式体现出了明显的日内效应和周内效应,大规模人类行为在直播平台中体现出明显的重尾特征,如观众观看次数与时长、打赏额和评论量、主播直播次数和时长、吸引观众的能力等分布

均从不同程度上符合重尾分布.说明直播平台内人群分布的异质性很强,可以据此对直播平台的经营模式如用户(观众、主播)激励模式、虚拟礼物打赏机制等进一步优化.网络直播平台中大规模人群交互形成了多种用户社群网络,识别用户社群和分析社群网络的结构和属性对优化直播平台的发展和應用有重要意义.

由于网络直播以及相关研究发展的时间尚且较短,对网络直播的研究广度和深度有待进一步探索.从研究主题发展轨迹和当前研究重点可以预测,挖掘直播平台中各种社群网络的形成和演化机制、设计针对直播平台的内容推荐和检测算法等是网络直播领域研究的未来发展趋势.

参考文献

- [1] CNNIC http://www.cac.gov.cn/2019-08/30/c_1124938750.htm [2019-8-30]
- [2] Huang S, Hu H 2002 *Proceedings International Symposium on Multimedia Software Engineering* Taiwan, December 11–13, 2000 p411
- [3] Abdous M, Yoshimura M 2010 *Comput. Educ.* **55** 733
- [4] Cao M X 2019 *Mod. Mark.* (10) 82 (in Chinese) [曹孟熙 2019 现代营销(经营版) (10) 82]
- [5] Lu Z, Xia H, Heo S, Wigdor D 2018 *Proceedings of the 2018 CHI conference on human factors in computing systems* Montréal, Canada, April 21–26, 2018 p1
- [6] Yan M 2019 *Southeast. Commun.* (5) 66 (in Chinese) [阎敏 2019 东南传播 (5) 66]
- [7] Dong D L 2019 *Course Educ. Res.* (41) 24 (in Chinese) [董弟林 2019 课程教育研究 (41) 24]
- [8] Bao X 2009 *Gov. Leg. Inst.* (22) 51 (in Chinese) [保旭 2009 政府法制 (22) 51]
- [9] Guo B, Zhang H Y 2008 *Inherit* (22) 18 (in Chinese) [郭波, 张会永 2008 传承 (22) 18]
- [10] Borges A, Gomes P, Nacif J, Mantini R, Almeida J M, Campos S 2012 *Comput. Commun.* **35** 1004
- [11] Veloso E, Almeida V, Meira W, Bestavros A, Jin S 2002 *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurement* Marseille France, November 6–8, 2002 p117
- [12] Deng J, Cuadrado F, Tyson G, Uhlig S 2015 *2015 International Workshop on Network and Systems Support for Games (NetGames)* Zagreb, Croatia, December 3–4, 2015 p1
- [13] Haimson O L, Tang J C 2017 *Proceedings of the 2017 CHI conference on human factors in computing systems* Colorado, USA, May 6–7, 2017 p48
- [14] Jia A L, Shen S, Epema D H, Iosup A 2016 *Acm T. Multim. Comput.* **12** 47
- [15] Bonald T, Massoulié L, Mathieu F, Perino D, Twigg A 2008 *ACM SIGMETRICS Performance Evaluation Review* MD, USA, June 2–6, 2008 p325
- [16] Chang H, Jamin S, Wang W 2009 *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement* Illinois USA, November 4–6, 2009 p417

- [17] Fallica B, Lu Y, Kuipers F, Kooij R, Miegheem P V 2008 *The Second International Conference on Next Generation Mobile Applications, Services, and Technologies* Cardiff, United Kingdom, September 16–19, 2008 p501
- [18] Yang T W, Shih W Y, Huang J L, Ting W C, Liu P C 2013 *Conference on Technologies and Applications of Artificial Intelligence* Taiwan, December 6–8, 2013 p188
- [19] Liu Y W, Lin C Y, Huang J L 2015 *2015 IEEE International Conference on Consumer Electronics-Taiwan* Taiwan, June 6–8, 2015 p118
- [20] Lin C Y, Chen H S 2019 *Multimedia Tools Appl.* **78** 1999
- [21] Zhang L 2019 *Mark. China* (**Z2**) 56 (in Chinese) [张璐 2019 成功营销 (**Z2**) 56]
- [22] Zhou Y Q 2018 *M.S. Thesis* (Nanjing: Nanjing University of Posts and Telecommunications) (in Chinese) (in Chinese) [周钰淇 2018 硕士学位论文 (南京: 南京邮电大学)]
- [23] Han D, Wang C, Yue H Y 2006 *Beijing Hangkong Hangtian Daxue Xuebao* 5 (in Chinese) [韩冬, 王春峰, 岳慧煜 2006 北京航空航天大学学报 (社会科学版) 5]
- [24] Pink D H 2019 *When: The Scientific Secrets of Perfect Timing* (New York: Penguin Press) pp15–20
- [25] Hines C B 2004 *J. Cathol. Educ.* **7** 390
- [26] Bernard T, Giacomoni M, Gavarry O, Seymat M, Falgairrette G 1997 *Eur. J. Appl. Physiol. Occup. Physiol.* **77** 133
- [27] Chen X B, Zhang Z C 2008 *Chin. J. Manage. Sci.* (**4**) 44 (in Chinese) [陈雄兵, 张宗成 2008 中国管理科学 (**4**) 44]
- [28] Kaytoue M, Silva A, Cerf L, Meira Jr W, Raïssi C 2012 *Proceedings of the 21st International Conference on World Wide Web* Lyon France, April 16–20, 2012 p1181
- [29] Farrington D J, Muesch N M https://web.wpi.edu/Pubs/E-project/Available/E-project-031915-220004/unrestricted/Analysis_of_the_Characteristics_and_Content_of_Twitch.tv_Live-streaming.pdf [2015-3]
- [30] Claypool M, Farrington D, Muesch N 2015 *2015 IEEE Games Entertainment Media Conference (GEM)* Toronto, Canada, October 14–16, 2015 p1
- [31] Pires K, Simon G 2015 *Proceedings of the 6th ACM multimedia systems conference* Oregon, USA, March 18–20, 2015 p225
- [32] Zhu Z H, Yang Z, Dai Y F 2017 *International Conference on Social Computing and Social Media* Vancouver, Canada, July 9–14, 2017 p274
- [33] Wang X, Tian Y, Lan R, Yang W, Zhang X 2018 *IEEE T. Circ. Syst. Vid.* **29** 3454
- [34] Peng S L, Wu N N, Zhao G Q 2014 *AMM* **675–677** 1810
- [35] Stohr D, Li T, Wilk S, Santini S, Effelsberg W 2015 *IEEE 40th Local Computer Networks Conference Workshops (LCN Workshops)* Florida, USA, October 26–29, 2015 p673
- [36] Newman M E J 2005 *Contemp. Phys.* **46** 323
- [37] Li Z Y, Kaafar M A, Salamatian K, Xie G G 2016 *IEEE Trans. Circuits. Syst. Video* **27** 2675
- [38] Deng J, Tyson G, Cuadrado F, Uhlig S 2017 *International Conference on Passive and Active Network Measurement* Sydney, Australia, March 30–31, 2017 p60
- [39] Yan H, Lin T H, Zeng M, Wu J, Huang J X, Li Y, Jin D P 2017 *GLOBECOM 2017 : IEEE Global Communications Conference (GLOBECOM)* Singapore City, Singapore, December 4–8, 2017 p1
- [40] Sripanidkulchai K, Maggs B, Zhang H 2004 *Proceedings of the 4th ACM SIGCOMM Conference on Internet Measurement* Sicily, Italy, October 25–27, 2004 p41
- [41] Tang Y, Sun L, Luo J, Zhong Y 2006 *7th Pacific-Rim Conference on Multimedia* Hangzhou, China, November 2–4, 2006 p175
- [42] Nascimento G, Ribeiro M, Cerf L, Cesário N, Kaytoue M, Raïssi C, Vasconcelos T, Meira W 2014 *9th Latin American Web Congress* Minas Gerais, Brazil, October 22–24, 2014 p1
- [43] Zhao J, Ma M, Gong W, Zhang L, Zhu Y, Liu J 2017 *2017 IEEE/ACM 25th International Symposium on Quality of Service (IWQoS)* Vilanova i la Geltrú, Spain, June 14–16, 2017 p1
- [44] Lan R H, Zhu G, Yang W, Tian Y, Zhu M 2019 *Comput. Syst. Appl.* **28** 69 (in Chinese) [兰荣亨, 朱格, 杨文, 田野, 朱明 2019 计算机系统应用 **28** 69]
- [45] Poyane R 2018 *Proceedings of the 22nd International Academic Mindtrek Conference* Tampere, Finland, October 10–11, 2018 p262
- [46] Nematzadeh A, Ciampaglia G L, Ahn Y Y, Flammini A 2016 *R. Soc. Open Sci.* **6** 191412
- [47] Olejniczak J 2018 *Redefining Community in Intercultural Context* Bucharest, Romania, May 11–15, 2018 p329
- [48] Lan R H, Hu Y H, Zhu G, Tian Y, Zhu M 2019 *Comput. Syst. Appl.* **28** 219 (in Chinese) [兰荣亨, 胡雨晗, 朱格, 田野, 朱明 2019 计算机系统应用 **28** 219]
- [49] Zhang C, Liu J 2015 *Proceedings of the 25th ACM Workshop on Network and Operating Systems Support for Digital Audio and Video* Oregon, USA, March 18–20, 2015 p55
- [50] Pires K, Simon G 2014 *Proceedings of the 2014 Workshop on Design, Quality and Deployment of Adaptive Video Streaming* Sydney, Australia, December 2, 2014 p13
- [51] Arnett L, Netzorg R, Chaintreau A, Wu E 2019 *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* Glasgow, Scotland Uk, May 4–9, 2019 pLBW1211
- [52] Netzorg R, Arnett L, Chaintreau A, Wu E 2018 arXiv: 1812.03379 [cs.SI]
- [53] Szabo G, Huberman B A 2008 *Commun. ACM* **53** 80
- [54] Clauset A, Newman M E J, Moore C 2004 *Phys. Rev. E* **70** 6
- [55] Churchill B C, Xu W 2016 *2016 IEEE International Conferences on Big data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom)(BDCloud-SocialCom-SustainCom)* Georgia, USA, October 8–10, 2016 p223
- [56] Smith T, Obrist M, Wright P 2013 *Proceedings of the 11th European Conference on Interactive TV and Video* Como Italy, June, 2013 p131
- [57] Lykousas N, Gómez V, Patsakis C 2018 *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* Barcelona, Spain, August 28–31, 2018 p375
- [58] Hamilton W A, Garretson O, Kerne A 2014 *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems* Toronto, Canada, April 26–May 1, 2014 p1315
- [59] Zhuang Q L, Zhou L 2018 *J. News Res.* **10** 32 (in Chinese) [庄庆玲, 周丽 2018 新闻研究导刊 **10** 32]
- [60] Newman M E J 2003 *SIAM Rev.* **45** 167

SPECIAL TOPIC—Statistical physics and complex systems

Live streaming: Data mining and behavior analysis^{*}Guo Shu-Hui Lu Xin[†]*(College of Systems Engineering, National University of Defense Technology, Changsha 410073, China)*

(Received 22 November 2019; revised manuscript received 20 March 2020)

Abstract

With the rapid development of mobile communication and Internet technologies, online live streaming has gradually become popular for information communication and entertainment in the new media environment. Live streaming has been widely used in teaching, reality show, E-sports games and events, brand marketing and other aspects. With the active participation of millions of streamers and hundreds of millions of viewers, massive online crowd behavior activity data are generated, which offers rich experimental scenarios for large-scale crowd behavior dynamics research, live streaming channel recommendation and online community evolution. In this paper, we summarize the relevant research literature of live streaming, and review current studies from a comprehensive list of aspects: workload pattern, viewers and streamers behavior, community network discovery and analysis, etc. We summarize the temporal and spatial patterns of live streaming platform workload, heavy tailed effect of large-scale crowd behavior in live streaming platform, etc. We believe that the future work on live streaming can be directed in the examination of formation and evolution mechanism of various community networks formed by large-scale users, as well as the recommendation and detection of live streaming content.

Keywords: live streaming platform, human behavior, community network, data mining**PACS:** 89.70.-a, 89.75.-k, 89.75.Kd**DOI:** [10.7498/aps.69.20191776](https://doi.org/10.7498/aps.69.20191776)

^{*} Project supported by the National Natural Science Foundation of China (Grant Nos. 82041020, 71771213, 91846301, 71790615) and the Science and Technology Program of Hunan Province, China (Grant Nos. 2017RS3040, 2018JJ1034).

[†] Corresponding author. E-mail: xin.lu@flowminder.org