



# **NERCCS 2021:**

## **Fourth Northeast Regional**

## **Conference on Complex Systems**

March 31 - April 2, 2021      Online (via Zoom)

Abstracts

**Abstracts for NERCCS 2021:**  
**Fourth Northeast Regional Conference on**  
**Complex Systems**

Complex Systems Society US Northeast Chapter

Center for Collective Dynamics of Complex Systems  
Binghamton University, State University of New York

March 31–April 2, 2021  
Online (via Zoom)

<https://nerccs2021.github.io/>

# Contents

<b>I Invited Talks</b>	<b>1</b>
Erik Boltt: On Geometry of Information Flow for Causal Inference in Complex Systems	2
Alexandra Paxton: From Arguments to Wars and the Scales In-Between: Dynamical Systems Analyses of Human Conflict	3
<b>II Contributed Talk Session 1</b>	<b>4</b>
Li Pi, Paul Expert and Ceire Costelloe: CPE spreading patterns from contact networks of patients in 3 urban hospitals: implications for infection prevention and control	5
Samin Aref and Zachary Neal: Analyzing coalitions, ideology, and effectiveness in the US Congress by optimally partitioning signed networks of legislators	6
Guillaume St-Onge, Hanlin Sun, Antoine Allard, Laurent Hébert-Dufresne and Ginestra Bianconi: Bursty exposure on higher-order networks leads to nonlinear infection kernels	7
Pegah Hozhabrierdi, Raymond Zhu, Maduakolam Onyewu and Sucheta Soundarajan: Network-Based Analysis of Early Pandemic Mitigation Strategies: Solutions, and Future Directions	8
<b>III Contributed Talk Session 2</b>	<b>9</b>
Mateusz Wilinski and Andrey Lokhov: Scalable Learning of Independent Cascade Dynamics from Partial Observations	10
Dj Passey, Ben Webb, Joseph Wilkes and Joseph Jamieson: Understanding Reservoirs: Ideal Networks for Replicating Chaos	11
Andreia Sofia Teixeira, Francisco C. Santos, Alexandre P Francisco and Fernando P. Santos: Fairness in multiplayer ultimatum games through degree-based role assignment	12
Gabriele Pisciotta and Giulio Rossetti: A Network Science approach to Instance Matching	13
<b>IV Contributed Talk Session 3</b>	<b>14</b>
Alfredo Morales, Isabella Loaiza and Sandy Pentland: Modeling Migrant Mobility	15

<b>Ülgen Kilic, Sarah Muldoon and Dane Taylor: Geometrical and topological data analyses reveal that higher-order structures provide flow channels for neuronal avalanches</b>	<b>16</b>
<b>Andrea Allen, Mariah Boudreau, Nicholas Roberts and Laurent Hebertdufresne: Modeling epidemic interventions with probability generating functions</b>	<b>17</b>
<b>Dylan Morris, Daniel Cooney, Simon Levin, Daniel Rubenstein and Pawel Romanczuk: Social Dilemmas of Sociality in the Presence of Beneficial and Costly Contagion</b>	<b>18</b>
<b>V Contributed Talk Session 4</b>	<b>19</b>
<b>Denis Patterson and Jonathan Touboul: A Mathematical Model of Neuronal Identity with Ectopic Domains</b>	<b>20</b>
<b>Eric Peña and Hiroki Sayama: A Life Worth Mentioning: Complexity of Conway’s Game of Life and Other Notable Life-like Rules</b>	<b>21</b>
<b>Lingqi Meng and Naoki Masuda: Epidemic threshold for metapopulation model networks with a second-order mobility rule</b>	<b>22</b>
<b>Caitrin Hall, Ji Chul Kim, Edward Large and Alexandra Paxton: Differential Coordination Patterns During Human-Metronome Phasing</b>	<b>23</b>
<b>VI Contributed Talk Session 5</b>	<b>24</b>
<b>Daniel Stroembom and Alice Antia: Anticipation induces polarized collective motion in attraction based models</b>	<b>25</b>
<b>Zhao Song and Dane Taylor: Asymmetric Coupling of Networks Optimally Accelerates Collective Dynamics</b>	<b>26</b>
<b>Jan Korbel, Simon Lindner, Rudolf Hanel and Stefan Thurner: Thermodynamics of structure-forming systems</b>	<b>27</b>
<b>Virginia Morini, Laura Pollacci and Giulio Rossetti: Towards a Standard Approach for Echo Chamber Detection: Reddit Case Study</b>	<b>28</b>
<b>VII Contributed Talk Session 6</b>	<b>29</b>
<b>Massimo Stella: Cognitive network science and its links with Education, complex systems and ways of thinking</b>	<b>30</b>
<b>Shuang Zhang, Feifan Liu and Haoxiang Xia: Understanding Scholars Mobility in Knowledge Embedded Space</b>	<b>31</b>

<b>Tuan Pham, Andrew Alexander, Jan Korbel, Rudolf Hanel and Stefan Thurner: Balanced and fragmented phases in societies with homophily and social balance</b>	<b>32</b>
<b>Elie Alhajjar, Ryan Fameli and Shane Warren: Are Terrorist Networks Just Glorified Criminal Cells?</b>	<b>33</b>
<b>VIII Contributed Talk Session 7</b>	<b>34</b>
<b>Mengyuan Sun and Naoki Masuda: COVID-19 impacts on domestic airlines via multi-layer network analysis</b>	<b>35</b>
<b>Anca Radulescu: Complex dynamics in networks, templates and mutated systems</b>	<b>36</b>
<b>Bao Huynh, Haimonti Dutta and Dane Taylor: Impact of Community Structure on Consensus Machine Learning</b>	<b>37</b>
<b>Ülgen Kılıç and Sarah Muldoon: Characterization of communities in dynamic functional networks</b>	<b>38</b>
<b>IX Contributed Talk Session 8</b>	<b>39</b>
<b>Marina Vegué, Vincent Thibeault, Patrick Desrosiers and Antoine Allard: Dimension reduction on heterogeneous networks</b>	<b>40</b>
<b>Yiming Che and Changqing Cheng: Active Learning and Relevance Vector Machine in Efficient Estimate for Basin Stability of Dynamic Networks</b>	<b>41</b>
<b>Chanon Thongprayoon, Lorenzo Livi and Naoki Masuda: Visualizing trajectory of tie-decay temporal networks</b>	<b>42</b>
<b>Vincent Thibeault, Marina Vegué, Antoine Allard and Patrick Desrosiers: Dimension reduction of high-dimensional dynamics on networks with adaptation</b>	<b>43</b>
<b>X Poster Session 1</b>	<b>44</b>
<b>Scott Donaldson: Flocc: From Agent-Based Models to Interactive Simulations on the Web</b>	<b>45</b>
<b>Felipe Xavier Costa and Pedro Pessoa: Entropic dynamics of networks</b>	<b>46</b>
<b>Goktug Islamoglu: Exact Formulation of Ising Model Transitions Between Six Magnetic Phases [commercial video]</b>	<b>47</b>

<b>Jun Kataoka, Jayanth Sivakumar, Eric Peña, Yiming Che and Changqing Cheng: Network Reconstruction (Bayesian Sequential Inference of Sparse Network Connectivity)</b>	<b>48</b>
<b>Nicholas Landry and Karen Stengel: The effect of a viral load function on epidemic dynamics</b>	<b>49</b>
<b>Salvatore Citraro, Massimo Stella and Giulio Rossetti: Towards a multiplex network model of word associations and similarity in the human mind</b>	<b>50</b>
<b>Kazuki Nakajima, Kazuyuki Shudo and Naoki Masuda: Configuration models for hypergraphs preserving local quantities of nodes and hyperedges [commercial video]</b>	<b>51</b>
<b>Stephany Rajeh, Marinette Savonnet, Eric Leclercq and Hocine Cherifi: Correlation of community-aware and classical centrality measures: Examining the role of network topology [commercial video]</b>	<b>52</b>
<b>Jake Williams and Diana Solano-Oropeza: The Mixing Law</b>	<b>53</b>
<b>Penghang Liu, Tomomi Kito, Naoki Masuda and A. Erdem Sarıyüce: Temporal Motifs in Patent Opposition and Collaboration Networks [commercial video]</b>	<b>54</b>
<b>Francesca Larosa, Jamie Rickman and Nadia Ameli: Finding the right dam(n) partners for a just energy transition</b>	<b>55</b>
<b>Daiki Miyagawa, Koki Okamoto and Genki Ichinose: Robustness of football passing networks against cascading failure</b>	<b>56</b>
<b>Luka Blagojević and Aleksandra Alorić: Incremental research - a favoured strategy in the “publish or perish” environment [commercial video]</b>	<b>57</b>
<b>Mahsa Bagherikalhor and G.Reza Jafari: Abnormal behavior of the Heider balance theory with heterogeneous triads [commercial video]</b>	<b>58</b>
<b>Srinivas Pandey and Hiroki Sayama: Analyzing Eccentric Behavior of GAB Social Media Users [commercial video]</b>	<b>59</b>
<b>Noha Abdel-Mottaleb, Kashin Sugishita, Naoki Masuda and Qiong Zhang: Bio-Inspired Water Distribution Network Design</b>	<b>60</b>
<b>Tayeb Jamali, Olha Buchel, Yaneer Bar-Yam and Leila Hedayatifar: A Meta-Population Movement-Based Epidemic Model for COVID-19 Pandemic</b>	<b>61</b>
<b>Oleg Pavlov and Evangelos Katsamakas: COVID-19 and Financial Sustainability of Academic Institutions</b>	<b>62</b>

<b>Sangeeta Rani Ujjwal, Ulrike Feudel and Ramakrishna Ramaswamy: Intermittent frequency chimeras in modular network of FitzHugh-Nagumo oscillators</b>	<b>63</b>
<b>XI Poster Session 2</b>	<b>64</b>
<b>Yanchen Liu: approximate network symmetry</b>	<b>65</b>
<b>John Meluso, Laurent Hébert-Dufresne, James Bagrow and Rob Razzante: Masculinity Contest Cultures and Inclusive Cultures: Insights From a Network Model of Organizational Socialization and Promotion</b>	<b>66</b>
<b>Christopher Diggans, Jeremie Fish and Erik Boltt: Emergent hierarchy through conductance-based node constraints</b>	<b>67</b>
<b>Volkan Saribas, Babek Erdebilli and Ibrahim Yilmaz: MULTICRITERIA SELECTION AND EVALUATION OF THE LIFESAVING UAVs FOR A DETERMINED AREA (AEGEAN SEA COASTLINE) IN TURKEY BY COMPARING TODIM&amp;GOAL PROGRAMMING (GP) METHOD</b>	<b>68</b>
<b>Ruodan Liu, Masaki Ogura, Elohim Fonseca Dos Reis and Naoki Masuda: Modeling epidemic spreading in Markovian temporal networks with different degrees of concurrency</b>	<b>69</b>
<b>John O'Meara, Ashwin Vaidya and Stephen Essien: Complexity and Curriculum: Examining the Role of Connectivity in a Precalculus Course</b>	<b>70</b>
<b>Esteră Boncea, Paul Expert and Ceire Costelloe: Exploring the connectivity of a hospital: A network analysis of patient movement community structures [commercial video]</b>	<b>71</b>
<b>Fabiano Lemes Ribeiro: On the relation between transversal and longitudinal scaling in cities</b>	<b>72</b>
<b>Prosenjit Kundu, Chittaranjan Hens, Baruch Barzel and Pinaki Pal: Perfect synchronization in networks of Sakaguchi-Kuramoto oscillators</b>	<b>73</b>
<b>Touria Karite, Adil Khazari and Delfim F. M. Torres: Optimal control problem of variable order fractional systems with Riemann-Liouville derivative</b>	<b>74</b>
<b>Rupam Acharyya, Penghang Liu, A. Erdem Sariyuce and Naoki Masuda: Detecting Fraudulent Users in Online Marketplaces Using Temporal Motifs</b>	<b>75</b>
<b>Matthew Tomlinson, David Greenwood and Marcin Mucha-Kruczynski: Extreme financial losses and gains: mutually-exciting arrivals or conditional volatility?</b>	<b>76</b>
<b>André Martins: Defining extremism in opinion models [commercial video]</b>	<b>77</b>

<b>Cynthia Siew and Anutra Guru: Longitudinal Subject Fluency Networks of Psychology and Biology Students</b>	<b>78</b>
<b>Francesco Bertolotti, Angela Locoro and Luca Mari: Emergence of Risk Sensitivity in a First-principles Agent-based model</b>	<b>79</b>
<b>Andreas Pape and Peter DiCola: A Modular Common Pool Resource Model with Social Planning</b>	<b>80</b>
<b>Elias Fernández Domingos, Jelena Grujić, Juan Carlos Burguillo, Francisco C. Santos and Tom Lenaerts: Modeling behavioral experiments on uncertainty and cooperation with population-based reinforcement learning - Extended Abstract</b>	<b>81</b>
<b>Tong Wu, Matthew McGuire, Sarah Muldoon and David Poulsen: Evolution of brain networks after traumatic brain injury</b>	<b>82</b>
<b>Xinglong Ju and Feng Liu: Optimizing Wind Farm Efficiency with Self-Adaptivity Mechanism in Evolutionary Algorithm</b>	<b>83</b>
<b>Zachary Boyd and Kayvan Lavassani: Resiliency of Global Medical Equipment Supply Chains</b>	<b>84</b>

# **Part I**

## **Invited Talks**

# **On Geometry of Information Flow for Causal Inference in Complex Systems**

Erik Boltt

Department of Mathematics, Clarkson University

eboltt@clarkson.edu

## Abstract

Causal inference is perhaps one of the most fundamental concepts in science, beginning originally from the works of some of the ancient philosophers, through today, but also weaved strongly in current work from statisticians, machine learning experts, and scientists from many fields. We take the perspective of information flow, which includes the Nobel prize winning work on Granger-causality, and the recently highly popular transfer entropy, these being probabilistic in nature. We will briefly describe some examples of our causation entropy principle designed to differentiate direct versus indirect influences, with applications such as from Earth science, and neuro science. Our main contribution today will be our newly developed analysis contrasting geometric interpretation of information flow versus otherwise usually indicated by positive conditional entropies. The dimensionality of data intrinsic in an underlying manifold effectively projects into the outcome space summarizes information flow. Thus contrasting fractal dimensions conditionally applied to competing explanations of future forecasts shed light on causality.

**From Arguments to Wars and the Scales In-Between:  
Dynamical Systems Analyses of Human Conflict**

Alexandra Paxton

Department of Psychological Sciences, University of Connecticut  
[alexandra.paxton@uconn.edu](mailto:alexandra.paxton@uconn.edu)

**Abstract**

A dynamical systems perspective on human interaction might consider the interaction context—that is, the set of physical and social characteristics in which the interaction occurs—as a sort of constraint on the emergent dynamics. In this talk, I consider the impact of conflict as a constraint on human interaction on a range of spatiotemporal scales.

**Part II**

**Contributed Talk Session 1**

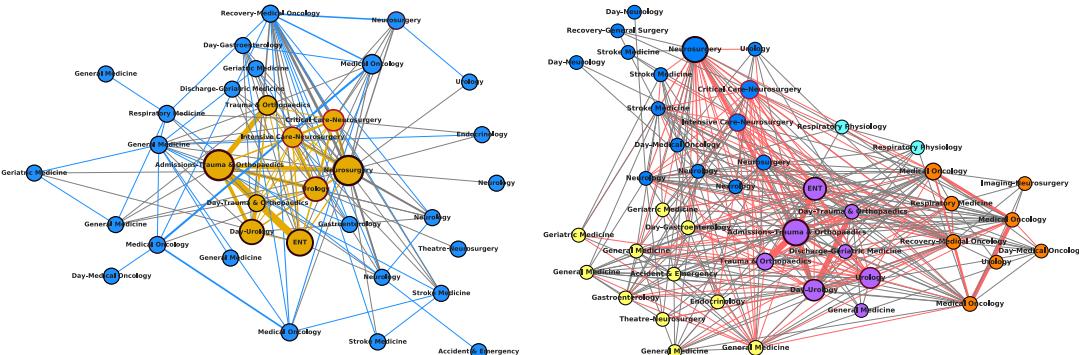
## CPE spreading patterns from contact networks of patients in 3 urban hospitals: implications for infection prevention and control

L Pi<sup>1</sup>, P Expert<sup>2</sup>, C Costelloe<sup>2</sup>

<sup>1</sup>Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, UK

<sup>2</sup>Global Digital Health Unit, Imperial College London, UK.

Healthcare-associated infection (HCAI), defined as infections acquired by patients during care delivery in a healthcare setting, represents one of the most significant challenges for modern medicine. HCAI can significantly impact patients' lives, leading to prolonged hospital stays, long-term disabilities and increased risk of mortality. In a recent modelling study, HCAI was estimated to cause 22,800 deaths in England and cost the health system an extra £2.1 billion in 2016/2017 [1]. Among nosocomial infection organisms, Carbapenemase-producing Enterobacteriaceae (CPE), pose nearly the greatest clinical threat, given the high levels of resistance to carbapenems which are considered as agents of 'last resort' against life-threatening hospital infections. Understanding patterns of CPE infection spreading in hospitals is therefore crucial to inform infection control protocols to mitigate the presence of CPE in hospitals. As CPE spreads in healthcare settings through contacts between patients, healthcare workers, and contaminated environments, we designed a track and trace experiment to reconstruct the contact networks of CPE carriers. We used patient electronic health records of 3 urban hospitals from 2015-2018 to i) track microbiologically confirmed CPE carriers and ii) trace the patients they shared place and time with until their identification using time stamped ward transfer information. We aggregated the contact network yearly to avoid confounding seasonal effects and obtain a weighted statistical representation of the contact networks. The contact networks exhibit a core-periphery structure (Fig 1, left), highlighting the presence of a core set of wards where most carriers-contact interaction occur before being redistributed to peripheral wards. Interestingly, high caseloads are not necessarily associated with core wards, supporting the mixing and disseminating picture. Using modularity, we identified functional communities of wards from the whole patient movement network. The contact networks projected onto the whole patient movement community structure shows a comprehensive coverage of the hospital. Our findings are robust over time periods and across hospitals. They show that CPE infection can reach virtually all parts of hospitals through first-level contacts. The methodology is generalisable to any contact transmitted pathogen, e.g. SARS-CoV-2. They can form the basis for improving infection prevention control strategies.



**Figure 1:** Structures for Hospital C for year 2. Left: Core (yellow) and periphery (blue) of the carriers contact network. Right: Functional modules identified from the whole population movement. Edges used by the contact network in red.

**References:** [1] Guest, J. F *et al.* "Modelling the annual nhs costs and outcomes attributable to healthcare-associated infections in England." BMJ Open10 (2020)

# Analyzing coalitions, ideology, and effectiveness in the US Congress by optimally partitioning signed networks of legislators

Samin Aref<sup>1</sup> and Zachary P. Neal<sup>2</sup>

<sup>1</sup> Max Planck Institute for Demographic Research  
aref@demogr.mpg.de

<sup>2</sup> Department of Psychology, Michigan State University  
zpneal@msu.edu

A signed network is one with positive and negative edges, and is  $k$ -balanced if the vertices can be partitioned into  $k$  subsets ( $k$ -partitioned) such that positive (negative) edges exist within (between) subsets. We analyze signed networks of US Congress from 1979-2018 by solving NP-hard problems to optimally partition the legislators into coalitions which minimize the inconsistent edges for  $k$  and in general. These signed networks are inferred from bill co-sponsorship data using a backbone model such that having significantly high (low) co-sponsorships is recorded as a positive (negative) tie. We then investigate the ideological composition and legislative effectiveness of these coalitions. Our numerical results show that the collaborative and oppositional patterns of US Congress are structured by more than two coalitions, in contrast to the traditional two-party categorization which oversimplifies this complex system. After extensive computational experiments, we focus on the composition of optimal 3-partitions as the most immediate point of departure from the traditional two-group categorization. The optimal 3-partitions are characterized by a large liberal coalition and a large conservative coalition, but also include a smaller, ideologically-fluid coalition of legislators. We observe that in most sessions of the Congress, this third coalition is ideologically aligned with the majority party. This group of outsiders tend to push back against members of their own party and have significantly higher effectiveness in advancing their legislative agenda. This suggests that staying out of the partisan divide could enhance legislative effectiveness.

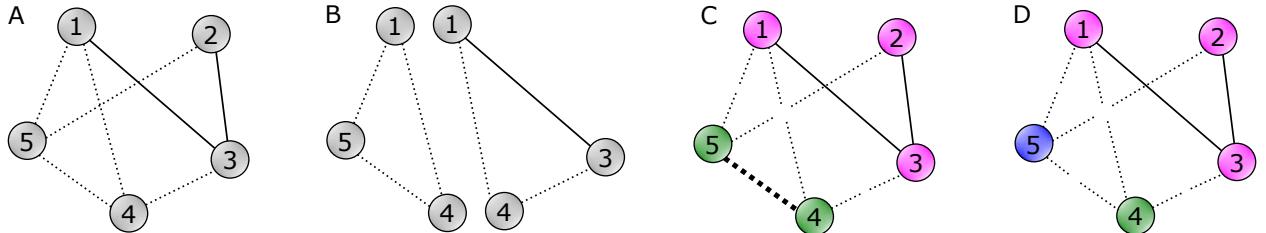


Figure 1: (A) An example signed network with negative (positive) edges shown by dotted (solid) lines. (B) Evaluating classic balance using triads: Triad 1-4-5 is unbalanced and triad 1-3-4 is balanced. Node 2 is disregarded and the macro structure remains unexplored. (C) Evaluating classic balance via 2-partitioning: The 2-partition  $\{\{1, 2, 3\}, \{4, 5\}\}$  minimizes the total number of intra-cluster negative and inter-cluster positive edges to one (edge 4-5). (D) Evaluating generalized balance and clusterability via  $k$ -partitioning: The 3-partition  $\{\{1, 2, 3\}, \{4\}, \{5\}\}$  satisfies the conditions of generalized balance.

## Bursty exposure on higher-order networks leads to nonlinear infection kernels

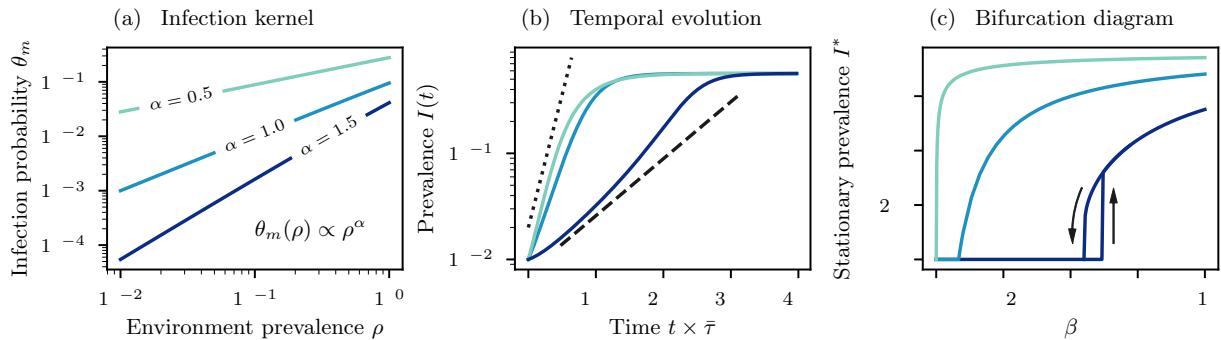
G. St-Onge<sup>1</sup>, H. Sun<sup>2</sup>, A. Allard<sup>1</sup>, L. Hébert-Dufresne<sup>1,3</sup>, and G. Bianconi<sup>2</sup>

1. Université Laval, 2. Queen Mary University of London, 3. University of Vermont

Three properties of human dynamics and disease transmissions are often overlooked in standard epidemic models: the hypergraph structure of contacts, the burstiness of human behavior, and the complex nonlinear relationship between the exposure to infected contacts and the risk of infection.

In this work, we combine these three properties in a hypergraph contagion model, with hyperedges representing environments where individuals can interact, and where a minimal number of interactions with infected individuals are needed to contract the disease. The central result we obtain is that bursty exposure, modeled by a power-law distribution of participation time to environments, can induce a nonlinear relationship between the number of infected participants and the probability to become infected. We then demonstrate how conventional epidemic wisdom can break down with the emergence of discontinuous transitions, super-exponential spread, and regimes of hysteresis.

On a theoretical level, this work formally provides a connection between complex contagions based on nonlinear infection kernels and threshold models. This also allows for a deeper understanding of how higher-order interactions and burstiness affect epidemic spreading. On the epidemiological level, our results challenge a key assumption of most epidemic models and ask: Why assume a linear relationship between the number of infectious contacts and the risk of infection?



**Figure 1:** Bursty exposure induces contagions with nonlinear infection kernels. We assume that an individual must have ( $K = 2$ ) interactions with infected individuals in a hyperedge (environment) to become infected. Each individual participates a time  $\tau$  to a hyperedge, distributed according to  $P(\tau) \propto \tau^{-\alpha-1}$  with  $\tau \in [1, \infty)$ . (a) Effective infection probability per participation to a hyperedge of any size  $m$ . The infection probability has a power-law scaling  $\theta_m(\rho) \propto \rho^\alpha$  if  $K \geq \alpha$ . (b)-(c) We study the consequences of nonlinear infection kernels for a SIS-type contagion dynamics on hypergraphs. The average number of interactions in an environment is proportional to  $\beta$ . (b) Supra-linear kernel  $\alpha > 1$  leads to a super-exponential growth for the global prevalence  $I(t)$ . We adjusted  $\beta$  to obtain similar prevalence levels.  $\bar{\tau}$  is the median participation time associated with  $P(\tau)$ . (c) The bifurcation diagram in the stationary state ( $t \rightarrow \infty$ ) can be continuous or discontinuous with a bistable regime. Sub-linear and linear kernels  $\alpha \leq 1$  lead to a continuous phase transition, and the epidemic threshold  $\beta_c$  vanishes for  $\alpha \rightarrow 0$ . Supra-linear kernels  $\alpha > 1$  can lead to a discontinuous phase transition with a bistable regime.

**Network-Based Analysis of Early Pandemic Mitigation Strategies: Solutions, and Future Directions**

Pegah Hozhabriderdi, Raymond Zhu, Maduakolam Onyewu and Sucheta Soundarajan

See the full paper published in the Northeast Journal of Complex Systems:  
<https://orb.binghamton.edu/nejcs/vol3/iss1/3/>

## **Part III**

# **Contributed Talk Session 2**

# Scalable Learning of Effective Spreading Models on Networks

Mateusz Wilinski<sup>1</sup> and Andrey Lokhov<sup>1</sup>

<sup>1</sup>Los Alamos National Laboratory, Los Alamos, 87545 New Mexico, United States of America

Spreading processes on networks play an increasingly important role in modeling infectious diseases, regulatory networks, marketing and opinion setting. Events like Coronavirus or Cambridge Analytica affair further highlight the need for prediction, optimization, and control of diffusion dynamics. To tackle these tasks, it is essential to learn the effective spreading model and transmission probabilities across the network of interactions. Unfortunately, in most cases the transmission rates are unknown and need to be inferred from the spreading data. Additionally, it is rarely the case that we have full observation of the dynamics. As a result, typical maximum likelihood approach quickly becomes intractable for large network instances.

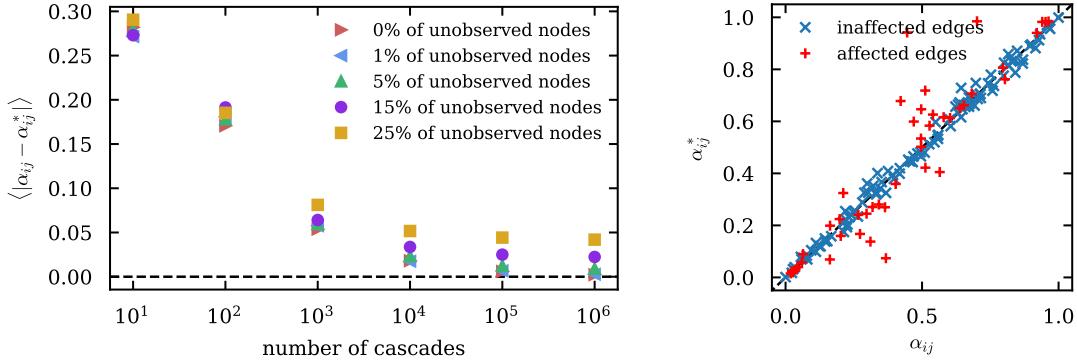


Figure 1: *Left:* Average distance between real and estimated transmission parameters  $\alpha_{ij}$ , for a random 3-regular graph with  $N = 100$  nodes, as a function of the number of available cascades. *Right:* Example of parameters estimation for a random 3-regular graph with  $N = 100$  nodes,  $M = 10^4$  cascades and 15% unobserved nodes. Each point represents a different parameter  $\alpha_{ij}$ . The red crosses represent edges affected by an unobserved node.

We introduce an efficient algorithm, based on a mean-field approximation and the dynamic message-passing approach, which is able to reconstruct parameters of the effective spreading model given only limited information on the activation times of nodes in the network. The proposed method can be easily generalized to a large class of dynamic models on networks and beyond. Exemplary results for Independent Cascade model on a random regular graph, are shown in figure 1. Number of cascades refer to the number of observed spreading realisations. For each simulation we randomly pick unobserved nodes and transmission probabilities. Note that we analyse a general case, where transmission probability can be different for each edge, but the method can easily be adjusted to simpler settings. Our algorithm is linear in both cascade length (number of spreading steps) and size of the network (number of edges), which makes it scalable and efficient when dealing with huge empirical networks. Additionally, we present a scheme to infer the effective model, which is able to approximate marginal activation probabilities even for settings where message passing is usually inaccurate.

# Understanding Reservoirs: Ideal Networks for Replicating Chaos

J Jamieson, D. Passey, D. Smith, B. Webb, and J. Wilkes

## Abstract

Reservoir computers are a machine learning model that, unlike most standard machine learning models, make use of an internal complex network. The effect of network topology on the dynamics of reservoir computers is not well understood. To investigate this, we trained over one-hundred million reservoir computers to replicate the chaotic dynamics of the Lorenz attractor. In our experiments, we varied hyperparameters and internal network structure to determine the combinations that produce the best predictions.

Specifically, we studied how reservoir computer prediction varied with respect to common network topologies across a substantial range of hyperparameters. We also studied the performance of "thinned" versions of the standard topologies, by removing a percent of edges at random.

We found that sparse, e.g. thinned, networks are typically better suited to learn the dynamics of the Lorenz attractor. Our experiments also revealed an interaction between the reservoir computer spectral radius parameter (a rescaling of the reservoir adjacency matrix to produce a desired spectral radius) and the sparsity of the internal network suggesting that spectral radius and sparsity combine in unexpected ways to impact reservoir computer predictive ability.

To help interpret our experimental findings, we analyzed the untrained reservoir dynamical system, studying how reservoir computer fixed points and their stability vary in response to a training signal. This analysis demonstrated that high spectral radius and high connectivity reduce the sensitivity of reservoir computers to input data, making them poorly suited for learning. This analysis offers insight into better design principles for reservoir computers and better understanding of the role of network topology in machine learning.

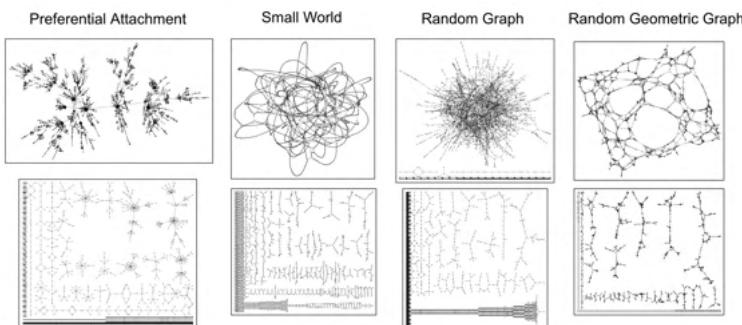


Figure 1: Complex network models before and after removing 80% of the edges.

## Fairness in multiplayer ultimatum games through degree-based role assignment

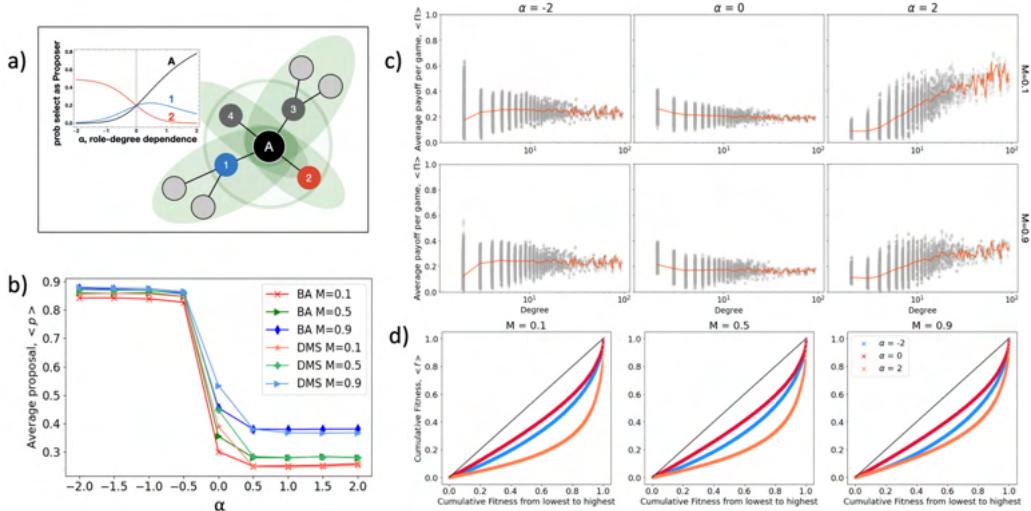
**Andreia Sofia Teixeira<sup>1,4</sup>, Francisco C. Santos<sup>1,2</sup>, Alexandre P. Francisco<sup>1,2</sup>, Fernando P. Santos<sup>3</sup>.** (1) INESC-ID,

Lisboa, Portugal, [steixeira@inesc-id.pt](mailto:steixeira@inesc-id.pt). (2) Instituto Superior Técnico, Universidade de Lisboa, Portugal (3)

Department of Ecology and Evolutionary Biology, Princeton University, Princeton, USA (4) Indiana University

Network Science Institute (IUNI), Indiana University, Bloomington IN, USA

From social contracts to climate agreements, individuals engage in groups that must collectively reach decisions with varying levels of equality and fairness. These dilemmas also pervade Distributed Artificial Intelligence, in domains such as automated negotiation, conflict resolution or resource allocation. As evidenced by the well-known Ultimatum Game – where a Proposer has to divide a resource with a Responder – payoff-maximizing outcomes are frequently at odds with fairness. Eliciting equality in populations of self-regarding agents requires judicious interventions. Here we use knowledge about agents' social networks to assess the impact of a novel fairness mechanism, in the context of Multiplayer Ultimatum Games. We focus on network-based role assignment and show that preferentially attributing the role of Proposer to low-connected nodes increases the fairness levels in a population. The probability that a node  $j$  is selected as Proposer is given by  $p_j = \frac{e^{\alpha k_j}}{\sum_i e^{\alpha k_i}}$ , where  $\alpha$  controls the dependence on the degree (see Fig. 1 a)). We evaluate the effectiveness of low-degree Proposer assignment considering networks with different average connectivity (using the Barabási-Albert algorithm (BA) of growth and preferential attachment and Dorogotsev-Mendes-Samukhin (DMS) duplication model), group sizes, and group voting rules - ( $M$ ) - when accepting proposals (e.g. majority or unanimity) (see Fig. 1b)). We further show that low-degree Proposer assignment is efficient, not only optimizing fairness, but also the average payoff level in the population (see Fig. 1 c)). Finally, we show that stricter voting rules (i.e., imposing an accepting consensus as requirement for collectives to accept a proposal) attenuates the unfairness that results from situations where hubs are the natural candidates to play as Proposers (see Fig. 1 d)). Our results suggest new routes to use role assignment and voting mechanisms to prevent unfair behaviors from spreading on complex networks.



**Figure 1:** Panel a): Example of probability distribution of degree-based assignment of the role of the proposer. Panel b): Average proposal for different networks, different values of  $\alpha$  (low (high)  $\alpha$  means that nodes with low (high) degree are more likely to be the proposers) and different group voting rules ( $M$  stands for the minimum group fraction of acceptance necessary for the proposal to be accepted). Panel c): Distributions of average fitness for different values of  $\alpha$  and group voting rules. Panel d): Lorenz curves for different values of  $\alpha$  and group voting rules.

## A Network Science approach to Instance Matching

Gabriele Pisciotta<sup>1</sup> and Giulio Rossetti<sup>2</sup>

<sup>1</sup> University of Pisa, Italy g.pisciotta1@studenti.unipi.it

<sup>2</sup> KDD Lab, ISTI-CNR, Italy giulio.rossetti@isti.cnr.it

Nowadays, the use of Knowledge Graphs (KG) as background knowledge is widespread in Machine Learning. Companies and users continuously create and share Linked Open Data: as a result it is common that a same real world entity can be described differently by different KGs. Identifying identity relations is a challenging task, usually called as Instance Matching (IM), that enables inter-linking of different KGs with the aim of supporting the expansion of the knowledge available to ML and Intelligent Agents systems. The IM can be seen as a special case of the Link Prediction (LP) task, in which the only edge type to be predicted between individual of different KGs is the `owl:sameAs`, as shown in Figure 1.

Quoting [4], the IM problem requires to find classifier that effectively discriminate between matching instances and non-matching instances.

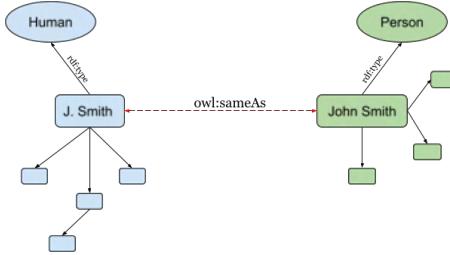


Figure 1: Example of matched instances

	Our Method			CODI		
	F1	P	R	F1	P	R
Value	0.90	0.99	0.85	0.96	0.99	0.93
Structural	0.63	0.72	0.60	0.88	0.95	0.83
Logical	0.92	0.93	0.92	0.97	0.96	0.97
Mixed	0.61	0.75	0.53	0.65	0.86	0.54

Table 1: Results for IIMB 2010 [3]

Considering the network-based nature of the data, we decided to discard any linguistic feature attached to the entities (usually exploited to address the IM task) and to focus on KGs topology alone: for each possible edge between the entities coming from KG pairs, discarding those belonging to disjoint classes, we compute: (i) the Jaccard distance between the two neighbours set, (ii) the Resource Allocation Index [5], (iii) the Adamic Adar coefficient [2], (iv) the Preferential Attachment [2].

We tested our approach on the IIMB 2010 dataset of the Ontology Evaluation Alignment Campaign [1] where the task requires to match entities of an original dataset to the ones of its 80 perturbations embedding various kinds of data transformation (including value, structural, logical, and a mixed ones). We applied AdaBoost to a 80-20 training-test split of the edges for each perturbation sub-task: Table 1 reports the average results for each type of transformation in terms of Precision/Recall and F1. Our results, even with a limited feature set, are in line with SOTA that leverage linguistic features analysis: moreover, our pure topological approach appears robust to value and logical transformation, downgrading its performances only when structural and mixed transformations are applied.

Concluding, our preliminary work underlines how IM can be successfully tackled without involving linguistic features related to the textual description of the entities, while focusing only on Network Science based features. As future works, we plan to extend the experiments to other datasets enriching the feature set and model definition strategies.

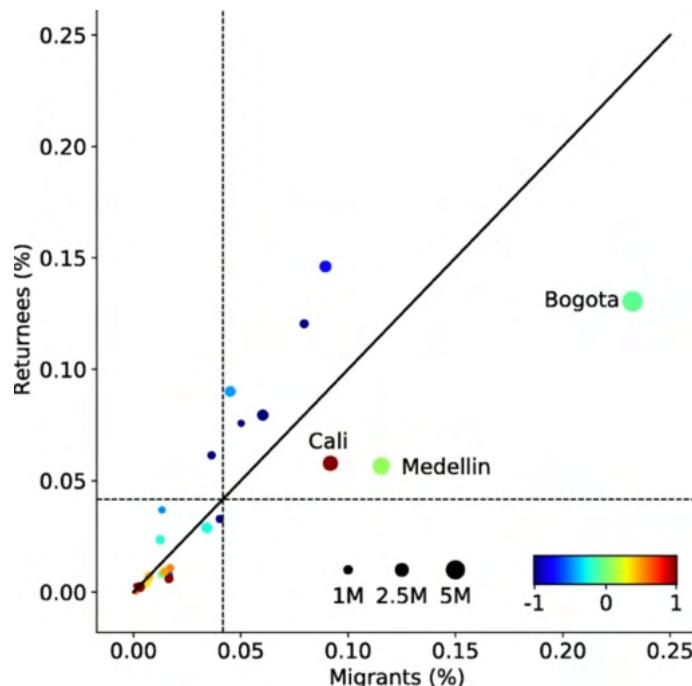
## **Part IV**

# **Contributed Talk Session 3**

# Modeling Migrant Mobility

Alfredo J. Morales, Isabella Loaiza, Alex “Sandy” Pentland  
MIT Media Lab  
alfredom@mit.edu

Massive migrations have become increasingly prevalent over the last decades. A recent example is the Venezuelan migration crisis across South America, which particularly affects neighboring countries like Colombia. It is crucial to map and understand geographic patterns of migration, including spatial mobility and dynamics over time. The aim of this paper is to uncover mobility and economic patterns of migrants that left Venezuela and migrated into Colombia due to the effects of the ongoing social, political and economic crisis. We analyze and compare the behavior of two types of migrants: Venezuelan refugees and Colombian nationals who used to live in Venezuela and return to their home country. We adapt the gravity model for human mobility in order to explain migrants’ dispersion across Colombia and analyze patterns of economic integration. This study is a first attempt at analyzing and comparing two kinds of migrant populations in one destination country, providing unique insight into the processes of mobility and integration after migration [1]. Findings suggest that migrants and returnees have different drivers behind their mobility patterns and destination choices (see Figure 1). While migrants are less constrained by distance and more attracted to larger cities, returnees seem to stay in places where they are from and have social networks. We also analyzed the economic integration of these individuals and found that migrants are less integrated than returnees despite having slightly higher levels of education. We hypothesize that such difference in economic integration is due to the advantages of having pre-existing social networks. The resolution of the available data does not allow a deeper investigation of the effects of social networks on the evolution of spatial and economic integration in migrants. However, given the challenges that hosting societies face in the context of sudden migrations, these findings highlight their importance for future research.



**Figure 1** Percentage of migrants versus returnees per department. Dots represent departments. The x-axis represents the relative number of migrants per department in 2019. The y-axis represents the relative number of returnees per department in 2019. The dot size is proportional to the department population. The dot color is proportional to the department distance to Caracas, Venezuela. The distance has been normalized by subtracting the average distance and dividing by the standard deviation. Blue indicates closer to Caracas. Red and green indicate farther. Scales in Figure. The vertical and horizontal dashed lines show the expected relative number of migrants or returnees if the distributions were uniform. The diagonal line is the identity ( $x=y$ ). The departments hosting Bogotá, Medellín and Cali have been annotated.

[1] Saa, I.L., Novak, M., Morales, A.J. *et al.* Looking for a better future: modeling migrant mobility. *Appl Netw Sci* 5, 70 (2020). <https://doi.org/10.1007/s41109-020-00308-9>

# Geometrical and topological data analyses reveal that higher-order structures provide flow channels for neuronal avalanches

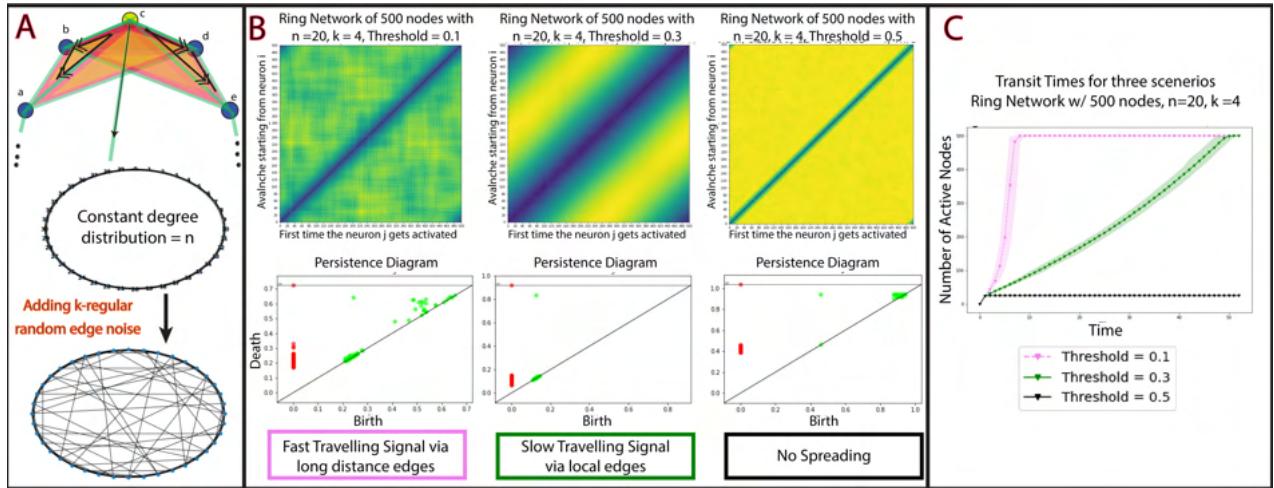
Bengier Ülgen Kılıç<sup>1</sup>, Sarah Feldt Muldoon<sup>1,2,3</sup>, Dane Taylor<sup>1,2</sup>

<sup>1</sup> Department of Mathematics, University at Buffalo, SUNY, Buffalo, NY

<sup>2</sup> CDSE Program, University at Buffalo, SUNY, Buffalo, NY

<sup>3</sup> Neuroscience Program, University at Buffalo, SUNY, Buffalo, NY

Neuronal systems exhibit a diversity of dynamical patterns (e.g., excitatory behavior, oscillations, wave propagation), and recent work has developed a broad understanding of how such dynamics relate to network properties including, for example, modular, geometric, and higher-order structures. Nevertheless, we still lack a mechanistic understanding of how (micro-scale) neuron dynamics reliably and efficiently organize to enable (large-scale) signal propagation and neurocomputation at the systems level. Here, we show that higher-order structures provide channels over which neuronal avalanches flow, which can potentially be used by neuronal systems to strike a balance between local and global signal propagation. To quantitatively study this phenomenon, we leverage mathematical tools from algebraic topology and geometry [1]. Our work complements prior research that has used topological data analysis (TDA) to study brain activity at the macroscopic level [2]. Here, we extend prior work [1] to investigate how neuronal systems balance structural properties (e.g., node degrees, manifold topology, and neighborhood overlap) with dynamical properties (e.g., intrinsic noise, refractory periods, and activation timescale) to yield neuronal avalanches that reliably follow higher-order structures **Figure 1**. Our model is a simple two-state ( $x_i^t \in \{0, 1\}$ ) stochastic threshold model with a non-linear response function in which neuron  $i$  fires at time  $t+1$  with probability  $p(x_i^{t+1} = 1 | x_i^t = 0) = \frac{1}{1 + e^{-C R(i)}}$  where  $R(i) = \left( \sum_{b_i \in B_i^t} \frac{b_i}{n+k} - t_i \right)$ ,  $B_i^t$  is the set of active neighbors of neuron  $i$  at time step  $t$ ,  $C$  is a constant to tune stochasticity of neuronal responses and  $t_i$  is an activation threshold for neuron  $i$ .



**Figure 1:** **A.** We run our experiments on noisy geometric networks that consist of both (i) short-range edges between nodes that are nearby along a ring manifold and (ii) long-range edges that represent a structural perturbation to the ring manifold. The top image highlights that there are higher order structures connecting adjacent neighbors. The middle and top images represent that the noisy geometric network consists of long and short-range edges, and we construct them in a way that is similar to the Watts-Strogatz small-world model, except that we force the network to be  $k$ -regular. **B.** The top panels depict distance matrices  $D$ , where  $D_{ij}$  is a notion of distance that is obtained directly from the dynamics—it is the first activation of the node  $j$  for a neuronal cascade that starts from node  $i$ . The lower panels depict persistence diagrams that summarize the topology of the neuronal cascades.. The three panels capture different behaviors that arise for three parameter choices for the neuron dynamics. From left to right, the neuronal avalanches (i) quickly spread but are not confined to the higher-order structural channels; (ii) slowly spread down the higher-order structural channels; or (iii) do not spread. **C.** To support our characterizations of fast, slow and no spreading, we plot the cascade size versus time for the three experiments shown in B.

## References

- [1] Taylor, D., Klimm, F., Harrington, H. et al. Topological data analysis of contagion maps for examining spreading processes on networks. *Nat. Commun.* 6, 7723 (2015).
- [2] Petri, Giovanni and Expert, Paul and Turkheimer, Federico and Carhart-Harris, Robin and Nutt, David and Hellyer, Peter J and Vaccarino, Francesco Homological scaffolds of brain functional networks. *Journal of The Royal Society Interface* (11) 101, 20140873 (2014).

## Modeling epidemic interventions with probability generating functions

Andrea Allen<sup>1</sup>, Mariah Boudreau<sup>1</sup>, Nicholas Roberts<sup>1</sup>, and Laurent Hébert-Dufresne<sup>1,2</sup>

1. Vermont Complex Systems Center, University of Vermont, Burlington, VT, USA

2. Département de Physique, de Génie Physique et d'Optique, Université Laval, Québec, Canada

Infectious diseases, now more than ever, are a critical area of study within the field of public health. Probability generating functions are a common approach used to analyze percolation models on contact networks,<sup>1</sup> which allow for the inclusion of both the stochastic nature of disease spread and the heterogeneous structure of contact networks. One critical limitation of using percolation models to study epidemics, and the generating function approach more specifically, is that they integrate over time, precluding us from studying temporal disease dynamics. However, there has been recent work<sup>2</sup> which allows for the application of a time varying generating function approach to study disease spread. We extend this theory to study the effect of public health interventions, such as vaccination, on the population-level disease dynamics. We perform simulations of disease spread on synthetic contact networks, and find that our results agree with our framework across a range of degree distributions. This agreement validates the use of our model over computationally expensive simulations and paves the way for extensive computational experiments of time-dependent interventions against emerging epidemics.

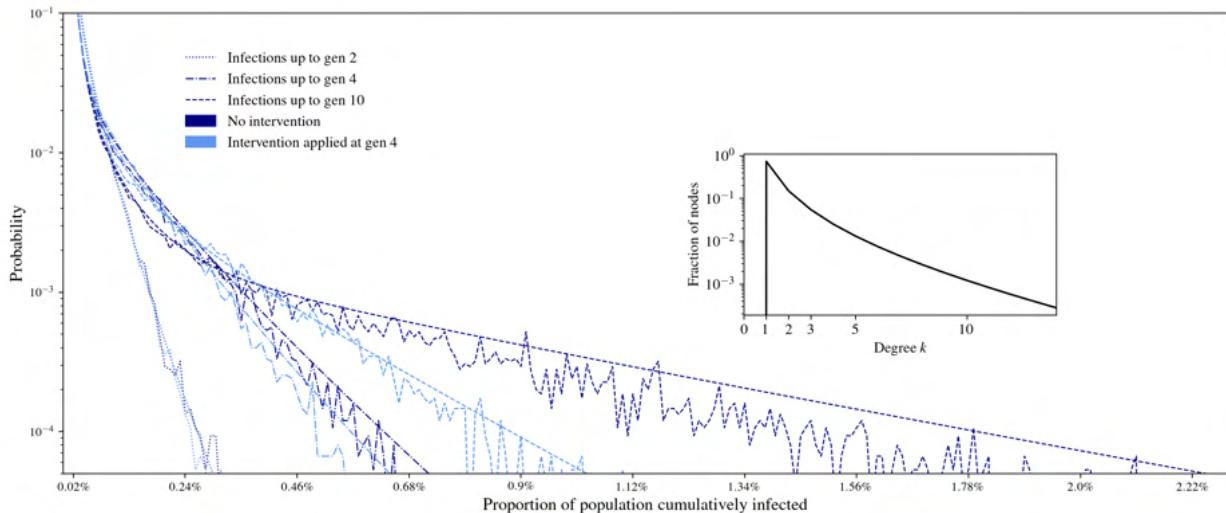


Figure 1: Probability distributions for cumulative outbreak size by epidemic generation on a modified power-law network, validated with simulation results, highlighting the effects of interventions on outbreak size. Smooth curves depict the analytical distributions, noisy curves show empirical results. Disease transmissibility  $T$  was set to  $T = 0.6$ , and in the intervention regime, was reduced to  $T = 0.4$  at generation 4. The non-intervention regime is shown in dark blue, the intervention regime in light blue. The inset plot shows the network degree distribution, given by  $p_k = k^{-2} e^{-k/5}$ .

<sup>1</sup>Mark EJ Newman. “Spread of epidemic disease on networks”. In: *Physical review E* 66.1 (2002), p. 016128.

<sup>2</sup>Pierre-André Noël et al. “Time evolution of epidemic disease on finite and infinite networks”. In: *Physical Review E* 79.2 (2009), p. 026101.

## Social Dilemmas of Sociality in the Presence of Beneficial and Costly Contagion

Dylan Morris<sup>1,\*</sup>, Daniel Cooney<sup>2,\*</sup>, Simon Levin<sup>3</sup>, Daniel Rubenstein<sup>3</sup>, and Paweł Romanczuk<sup>4</sup>

<sup>1</sup> Department of Ecology and Evolutionary Biology, University of California, Los Angeles

<sup>2</sup> Department of Mathematics, University of Pennsylvania, dbcooney@sas.upenn.edu

<sup>3</sup> Department of Ecology and Evolutionary Biology, Princeton University

<sup>4</sup> Institute for Theoretical Biology, Humboldt University of Berlin

Levels of sociality in nature vary widely from solitary species to the formation of complex animal societies. Increased levels of social interaction can allow for the spread of useful innovations and beneficial social information, but can also facilitate the spread of harmful infectious disease. Given the selection pressures that infectious disease impose on a population, it is natural to explore the ways in which contagion processes may shape the evolution of complex social systems [1].

In this talk, we consider a model for the evolution of sociality strategies in the presence of both a beneficial and costly contagion. We study these dynamics at three timescales: using a susceptible-infectious-susceptible (SIS) model to describe contagion spread for given sociality strategies, a replicator equation approach to study the changing fractions of resident and mutant strategies, and an adaptive dynamics approach to study the long-time evolution of the resident level of sociality.

We calculate socially-optimal and evolutionarily-stable sociality strategies through the reproduction numbers  $\mathcal{R}_{\text{opt}}$  and  $\mathcal{R}_{\text{ESS}}$  for the good contagion. For various assumptions on the benefits and costs of infection, we identify the social dilemma displayed in Figure 1:  $\mathcal{R}_{\text{ESS}} < \mathcal{R}_{\text{opt}}$  (less interaction than optimal) when the bad contagion spreads more readily, while  $\mathcal{R}_{\text{ESS}} > \mathcal{R}_{\text{opt}}$  (more interaction than optimal) when the good contagion spreads more readily. In extreme cases, this dilemma can result in complete collapse of sociality or reintroduction of harmful disease.

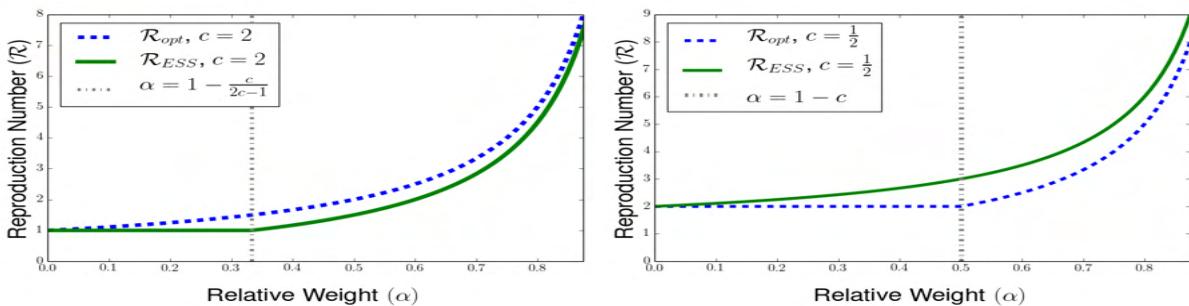


Figure 1:  $\mathcal{R}_{\text{ESS}}$  (green) and  $\mathcal{R}_{\text{opt}}$  (blue) depending on the weight of the good contagion  $\alpha$  to Cobb-Douglas utility when the good contagion spreads less (left) or more (right) than the bad contagion.

## References

- [1] Udiani, O., & Fefferman, N. H. (2020). How disease constrains the evolution of social systems. *Proceedings of the Royal Society B*, 287(1932), 20201284.

**Part V**

**Contributed Talk Session 4**

# A Mathematical Model of Neuronal Identity with Ectopic Domains

Denis Patterson<sup>1</sup> and Jonathan Touboul<sup>2</sup>

<sup>1</sup> High Meadows Environmental Institute, Princeton University,  
denispatterson@princeton.edu

<sup>2</sup> Department of Mathematics, Brandeis University,  
jtouboul@brandeis.edu

Recent experiments studying the development of cortical structures in mice have identified COUP-TF1 as a crucial determinant of both the position and sharpness of the boundary between the neo and entorhinal cortices. When COUP-TFI is under expressed, neocortex (NC) invades into territory occupied by the entorhinal cortex (MEC) in wild-type mice, but the sharp boundary between cortical regions is maintained. However, if COUP-TFI is over-expressed, the boundary fractures and entorhinal cortex invades the neocortical domain, resulting in mice with ectopic regions of misplaced cortex.

We introduce a novel PDE model based on a Keller-Segel-type chemotaxis mechanism to account for both the sharp cortical boundaries of wild-type mice and the ectopic regions observed in mutant mice. Competition between entorhinal and neocortical progenitor cells is mediated by a gradient of COUP-TF1 across the spatial domain and chemotaxis operators model each cell's affinity for cells of their own type. We verify the well-posedness of the system and establish necessary conditions for pattern forming Turing bifurcations; we also numerically study the structure of the Turing space and its dependence on model parameters. Numerical simulations show excellent agreement with experimental observations and we present experimental data verifying the differential adhesion hypothesis underlying the model's phenomenology.

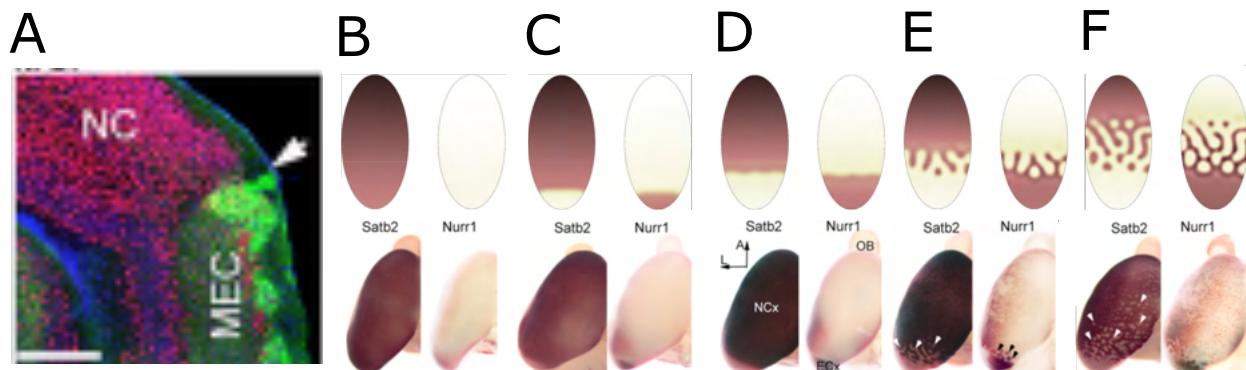


Figure 1: **A:** Standard boundary formation in a wild type mouse brain. **B-F:** Comparison of model solutions (top row) with imaging data from mutant mice brains (bottom row). COUP-TFI expression levels increase from B to F with the wild type brain shown in panel D.

# A Life Worth Mentioning

## Complexity of Conway's Game of Life and Other Notable Life-like Rules

Eric Peña & Hiroki Sayama

Center for Collective Dynamics of Complex Systems & Department of Systems Science  
 Binghamton University, State University of New York  
 eric.pena@binghamton.edu  
 sayama@binghamton.edu

Cellular automata (CA) have been lauded for their ability to generate complex global patterns from simple and local rules. The late English mathematician, John Horton Conway, developed his illustrious Game of Life (Life) CA in 1970 and it has since remained the quintessential CA construction—capable of producing a myriad of complex dynamic patterns and computational universality. Life and several other Life-like rules are usually classified in the same group of aesthetically and dynamically interesting CA characterized by their complex behaviors. However, a rigorous quantitative comparison among similarly classified Life-like rules has not been fully established. We use a modified conditional entropy measure, which considers an information-theoretic distance from randomness, as a measure of complexity ( $\Delta$ ) for two dimensional (2D) CA patterns. We use this measure to show that Life naturally aims to minimize density and maximize information content while being the most parsimonious Life-like rule with this feature. That is, Life is capable of maintaining a consistent amount of long-lasting complexity with the least number of conditions compared to other Life-like rules. We also show that the complexity of higher density Life-like rules, which themselves contain the Life rule (i.e., B3/S23), form a distinct density-complexity relationship whereby B356/S23 is proposed as an optimal complexity candidate (Fig.1). Generally, our results support that Life functions as the basic ingredient for cultivating the balance between structure and randomness thereby maintaining complexity in 2D CA over many evolutions. This work highlights the genius of John Horton Conway and his Game of Life—a testament to the timeless marvel which is certainly a *life worth mentioning*.

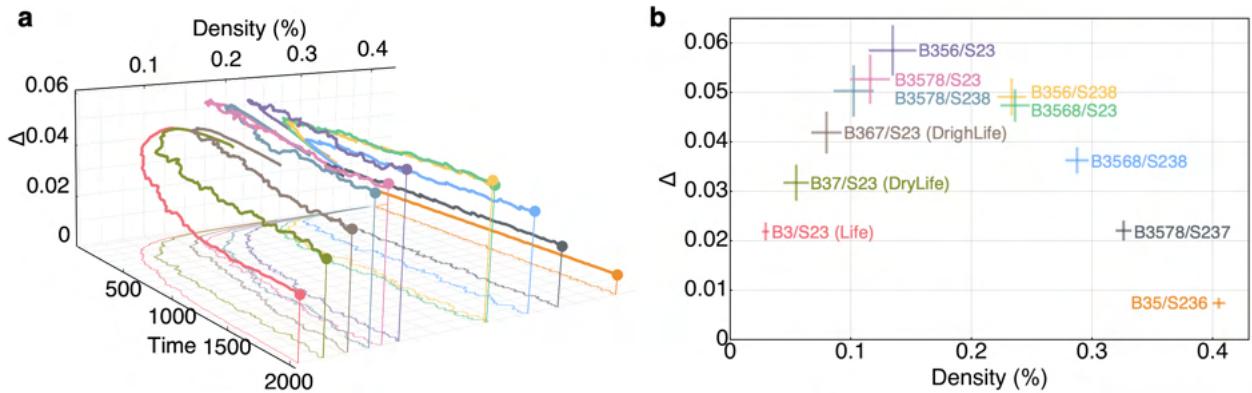


Figure 1: **a**, Temporal evolution of Life and higher density Life-like rules. **b**, Density and complexity concave relationship is revealed among Life-like rules that contain Life whereby each rule is represented with their 95% confidence intervals over 50 simulated runs.

## Epidemic threshold for metapopulation model networks with a second-order mobility rule

Lingqi Meng<sup>1</sup> and Naoki Masuda<sup>1,2</sup>

<sup>1</sup> Department of Mathematics, University at Buffalo

<sup>2</sup> Computational and Data-Enabled Science and Engineering Program, University at Buffalo

Metapopulation models have been proven to be powerful for simulating and predicting the dynamics of disease outbreaks. On a metapopulation network, agents migrate from one subpopulation to another obeying a given mobility rule. How different mobility rules affect epidemic dynamics occurring on metapopulation model networks has been a much debated topic. However, with regard to the effect of mobility on epidemic dynamics, there are relatively few efforts on systematic comparison between simple random walks and more complex mobility rules. Here we study a susceptible-infected-susceptible (SIS) dynamics on metapopulation model networks, where individual agents obey a parametric second-order Markov chain mobility rule, called the node2vec, of which we studied diffusion properties in our previous study [1]. We map the second-order mobility rule of the node2vec to a first-order transition probability matrix among directed edges to derive and analyze the expressions of the epidemic threshold for the SIS model and run numerical simulations for synthetic and empirical networks. We observe that for various networks, the epidemic threshold increases when the agents less frequently backtrack and visit the common neighbor(s) of the currently and the last visited subpopulations than when the agents obey the unbiased random walk. However, the amount of change in the epidemic threshold induced by the node2vec mobility is not as large as the one induced by the change in the diffusion rate.

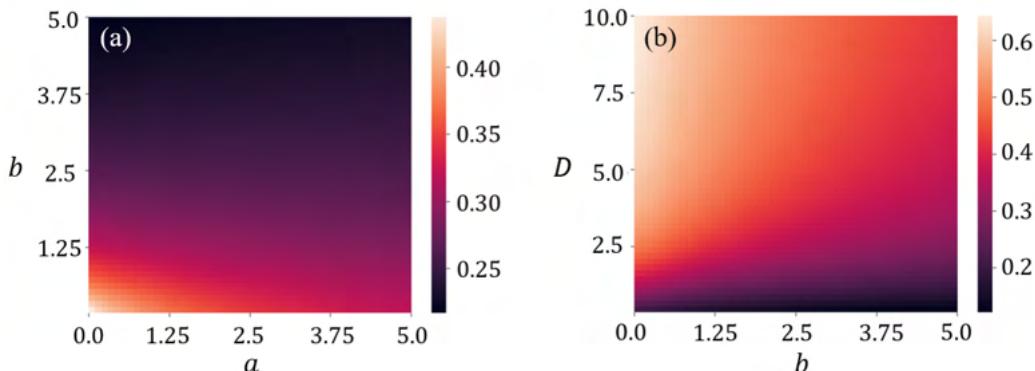


Figure 1: Epidemic threshold for a Barabási–Albert model network with 100 nodes and 291 edges. (a)  $b$  (i.e., the weight of transiting to a common neighbor of the currently visited subpopulation and the subpopulation visited in the last step) versus  $a$  (i.e., the weight of backtracking). (b)  $D$  (i.e., the diffusion rate) versus  $b$ .

## References

- [1] L. Meng, N. Masuda, Analysis of node2vec random walks on networks, *P. Roy. Soc. A* 476 20200447 (2020).

## Differential Coordination Patterns During Human-Metronome Phasing

Caitrín Hall <[caitrin.hall@uconn.edu](mailto:caitrin.hall@uconn.edu)>, Ji Chul Kim, Edward Large, & Alexandra Paxton  
*Department of Psychological Sciences, University of Connecticut*

One way perception-action coordination (sensorimotor synchronization [SMS]) may be studied is the coordination of movement with auditory rhythms. The current study assesses phasing--a musical technique in which two people tap the same rhythm at varying phases by adjusting tempi--to explore how SMS is impacted by situational and personal factors. Building on a study of professional percussionists, this experiment was designed to evaluate phasing in musicians and nonmusicians. To ensure that novice participants could complete the experiment, we simplified the original phasing task in two ways. First, rather than gradually desynchronizing and resynchronizing with a partner, participants phased with an isochronous metronome. Second, we narrowed the tempo range to 80-140 beats per minute (bpm). Participants were familiarized with phasing via video demonstrations, completed two practice sessions, and then engaged in the experimental phasing task with a metronome.

One technique to analyze the coordination characteristics of a system is recurrence quantification analysis (RQA). Coordination research uses RQA to describe potential nonlinearities of dynamical systems that cannot be captured by linear time series methods. We used multidimensional recurrence quantification analysis (MdRQA) to compare the nonlinear dynamics of phasing across tempi, musical experience, and multilingualism. Varying coupling patterns emerged, though not significantly influenced by linguistic or music experience. Only 38% of trials successfully completed the task (“Compliant Trials”); 41% of trials unsuccessfully attempted the task (“Noncompliant Trials”). Participants successful at phasing tapped stably near synchrony and antiphase, while those unskilled at phasing could not always detect synchrony and were less attracted to antiphase, as supported by differences in the predictability of the structure (measured in the determinism MdRQA metric) between Compliant and Noncompliant Trials.

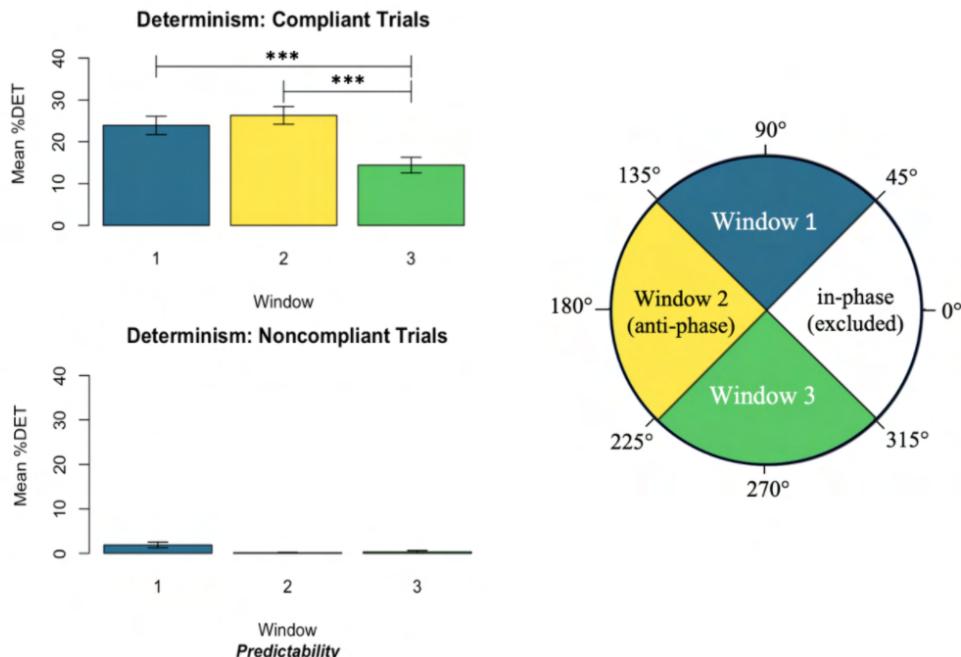


Figure. Bar charts: %DET for Compliant Trials (top) and Noncompliant Trials (bottom) for windows 1 (blue), 2 (yellow), and 3 (green). Pie chart: relative phase range for each window.

## **Part VI**

# **Contributed Talk Session 5**

**Anticipation induces polarized collective motion in attraction based models**

Daniel Stroembom and Alice Antia

See the full paper published in the Northeast Journal of Complex Systems:  
<https://orb.binghamton.edu/nejcs/vol3/iss1/2/>

# Asymmetric Coupling of Networks Optimally Accelerates Collective Dynamics

Zhao Song<sup>1</sup> and Dane Taylor<sup>1</sup>

<sup>1</sup> Department of Mathematics  
 University at Buffalo, State University of New York  
 zhaosong@buffalo.edu  
 danet@buffalo.edu

Networks are often interconnected with one network wielding greater influence over another, yet we lack basic insights into the effects of such asymmetry. Studying the convergence rate  $\text{Re}(\lambda_2)$  for consensus dynamics, we show that coupling asymmetry can monotonically increase/decrease  $\text{Re}(\lambda_2)$  or give rise to different types of optima. Symmetry-optimized systems can be counter-intuitive: when a fast and slow network are optimally coupled, either one can be more dominant. We characterize this and other phenomena and provide theoretical support for the optimization of asymmetrically coupled social, biological, and physical systems.

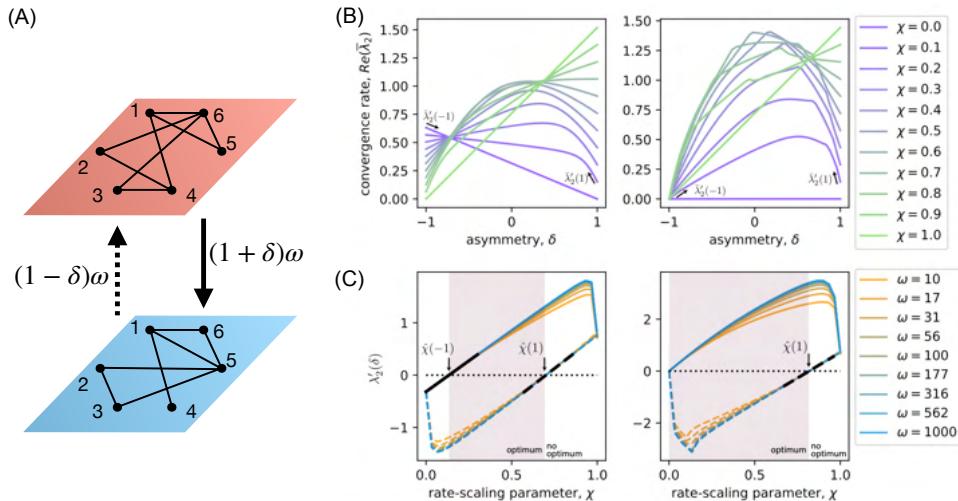


Figure 1: **A robust optimum requires that neither layer dominates.** (B)-(C) The two columns represent two different multiplex networks: a toy network (A) and an empirical social network. They are both 2-layer networks with directed edges within and/or between layers. (B) We plot  $\text{Re}(\lambda_2)$  versus  $\delta$  for several choices of  $\chi$ , which is a rate-scaling parameter that tunes the relative rate of consensus for the two coupled networks. The optima exists only for some values of  $\chi$ . (C) We plot  $\frac{d}{d\delta} \text{Re}(\lambda_2)$  versus  $\chi$ . The colored curves depict different coupled strengths  $\omega$ . The solid and dash black lines depict our theoretical prediction for the boundary on the range of  $\chi$  for which an optimum exists.

## Thermodynamics of structure-forming systems

Jan Korbel<sup>1,2</sup>, Simon David Lindner<sup>1,2</sup>, Rudolf Hanel<sup>1,2</sup> and Stefan Thurner<sup>1,2,3</sup>

<sup>1</sup> Section for the Science of Complex Systems, CeMSIIS, Medical University of Vienna, Vienna, Austria

<sup>2</sup> Complexity Science Hub Vienna, Vienna, Austria

<sup>3</sup> Santa Fe Institute, Santa Fe, NM, USA

Structure-forming systems are ubiquitous in nature, ranging from atoms building molecules to self-assembly of colloidal amphiphilic particles. The understanding of the underlying thermodynamics of such systems remains an important problem. Here we derive the entropy for structure-forming systems that differs from Boltzmann-Gibbs entropy by a term that explicitly captures clustered states. For large systems and low concentrations, the approach is equivalent to the grand-canonical ensemble; for small systems, we find significant deviations. We derive the detailed fluctuation theorem and Crooks' work fluctuation theorem for structure-forming systems. The connection to the theory of particle self-assembly is discussed. We apply the results to several physical systems. We present the phase diagram for patchy particles described by the Kern-Frenkel potential. We show that the Curie-Weiss model with molecule structures exhibits a first-order phase transition.

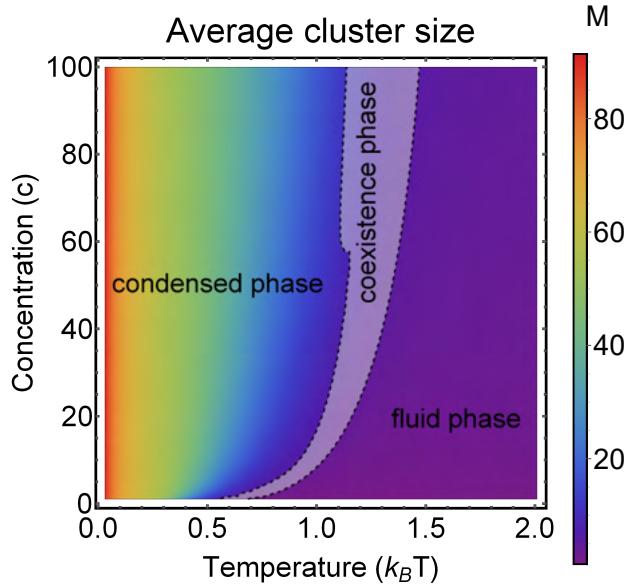


Figure 1: Phase diagram for the self-assembly of patchy particles for  $n = 100$  particles. More details are in Ref. [1]

### References

- [1] J. Korbel, S. D. Lindner, R. Hanel and S. Thurner, Thermodynamics of structure-forming systems, *Nature Communications* 12 (2021) 1127.

## Towards a Standard Approach for Echo Chamber Detection: Reddit Case Study

Virginia Morini<sup>1</sup>, Laura Pollacci<sup>2</sup> and Giulio Rossetti<sup>1</sup>

<sup>1</sup> KDD Lab, ISTI-CNR, Pisa, Italy (virginia.morini, giulio.rossetti)@isti.cnr.it

<sup>2</sup> Department of Computer Science, University of Pisa, Italy laura.pollacci@di.unipi.it

The rise of Social Networking sites (SNS) has drastically changed how people interact and access information. As a side effect, heterogeneity in content and freedom of expression offered by these platforms have led to the diffusion of alarming phenomena, such as Polarization and Echo Chambers (ECs). Up to date, researchers have reached no consensus about a standard methodology for ECs detection in a platform-independent way. In this preliminary work, we propose an approach to identify political ECs on Reddit, easily extendible to other SNS. Based on ECs common definition, such a methodology may be broken down into four key steps.

(i) Firstly, we identify a controversial topic i.e., the debate about Gun Control and Minority Discrimination during the first two and a half years of Donald Trump’s presidency. (ii) Then, since an ECs key feature is the homogeneous group polarization, we build a model to quantify user’s polarization degree from posts shared on the platform. For such a purpose, we define our Ground Truth, selecting strongly polarized subreddits concerning Pro-Trump and Anti-Trump beliefs, and we leverage Word Embeddings and Long Short-Term Memory architecture to compute posts polarization scores. Our results, shown in Table 1, are in line with SOTA, and, moreover, the model is quite able to generalize on less polarized sociopolitical topics. (iii) ECs are also strongly characterized by the dense interactions between their members with respect to the overall Network. To assess this requirement, we define, for each topic, the users’ interaction Network by retrieving all posts’ comments, and we label users with their average posts polarization score. (iv) Finally, we apply Community Detection algorithms (i.e., Louvain, Infomap and Eva) to extract communities, evaluating their cohesion both from a topological and ideological point of view.

From preliminary results, as we can see from Figure 1, with Eva we are able to detect potentially Pro-Trump and Anti-Trump ECs in both sociopolitical topics.

In conclusion, our case study underlines how features shared by most SNS (e.g., the existence of controversial issues, posts, comments) can be easily leveraged to identify ECs in a standard way. Consequently, as future work, we plan to extend this case study to other Social Media Services in order to analyze user behaviour across a wider range of online environments.

Dataset	# posts	Accuracy
Test Set	60,000	0.843
Gun Control	2,411	0.712
Minority Discrimination	4,839	0.732

Table 1: Model Accuracy

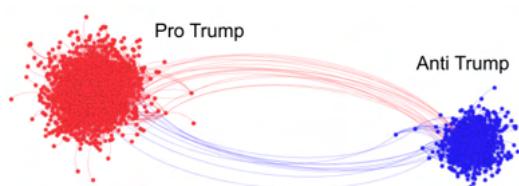


Figure 1: Discrimination - Echo Chambers

## **Part VII**

# **Contributed Talk Session 6**

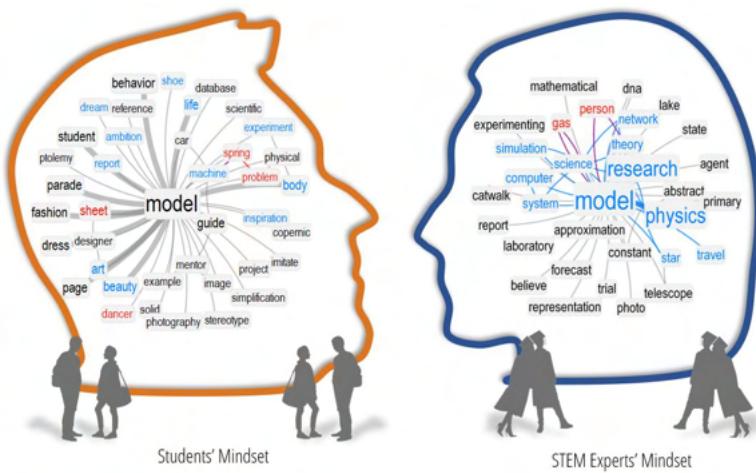
## Cognitive network science and its links with Education, complex systems and ways of thinking

Massimo Stella, Department of Computer Science, University of Exeter, UK - [m.stella@exeter.ac.uk](mailto:m.stella@exeter.ac.uk)

Cognitive network science is a rising field at the interface of complex networks and psycholinguistics aimed at representing knowledge in the human mind through conceptual associations between concepts and ideas. This contribution focuses on a specific type of cognitive networks combining memory patterns and valence norms, quantifying how individuals frame and perceive concepts. *Forma mentis* networks (Stella et al., *PLoS ONE*, 2019) capture the way of thinking of individuals and groups, i.e. their mindset or *forma mentis* in Latin. Relying on decades of psycholinguistic research, forma mentis networks represent how ideas are linked in memory via free associations (e.g. reading “math” makes one think of “anxiety”) and how they are perceived via word valence (e.g. how pleasant “math” is rated by an individual).

Forma mentis networks were successfully used in order to detect key aspects of STEM perceptions in high school students ( $N=159$ ) and STEM professionals ( $N=59$ ) (Stella et al., *PLoS ONE*, 2019). A cognitive dissonance was identified in the mindset of students, who framed positively the idea of “science” but then concentrated negative emotions of anxiety/stress around “physics” and “maths”. This emotional dichotomy was absent in researchers. More recent investigations focused on detecting computational thinking (Stella et al., 2020) and complex systems thinking (Stella, *IJCE*, 2020), as ways of understanding the world through data and connections, respectively. Despite being trained in multiple scientific disciplines, students perceived “computation” and “complexity” as distressing ideas, framed with links to mathematical jargon but crucially lacking links to real-world discovery. Students’ mindsets around “data”, “model” and “simulations” revealed almost no awareness of numerical modelling and its implications for understanding our complex world (see Figure 1). On the contrary, STEM experts perceived numerical models as positive concepts, framed and interpreted as a key way for achieving new knowledge.

The above findings provide concrete evidence that forma mentis networks can provide detailed knowledge about how students shape their mindsets and ways of thinking around Educational concepts. Comparisons with professionals or other target groups can provide crucial data-informed evidence for innovative Educational intervention policies aimed at fostering complex systems and data literacy.



**Figure 1:** Visual comparison of high school students’ and STEM experts’ mindsets around “model”. Associations provided by two or more individuals are thicker. Words perceived by students as positive (negative) are highlighted in cyan (red). Links between positive (positive/negative) words are highlighted in cyan (purple). Larger font-size indicates higher closeness centrality of a concept.

# Understanding Scholars Mobility in Knowledge Embedded Space

Shuang Zhang, Feifan Liu, Haoxiang Xia

Institute of Systems Engineering, Dalian University of Technology, Dalian, China

Despite persistent efforts in understanding human mobility in geographical space, little is known about the spatiotemporal regularity of scholars' mobility in virtual knowledge space owing to the inability to track scholars' overall and fine-grained trajectories in high-dimensional knowledge space. Only a few studies focus on scholars' research interest evolution and the factors that influence topic selection. Here, aiming at revealing the characteristics of scholars' knowledge exploration behavior and its correlation with academic performance, we study the trajectories of 180,345 highly productive scholars in the Computer Science field (CS), considering their publication sequence as walks on the two-dimensional scientific domain map constructed by manifold learning algorithm UMAP based on the semantic proximity of CS papers learned by Doc2vec. We find that the distributions of the exploration scope and the average step lengths of scholars are both left-skewed compared with the results of reshuffled or replaced publication sequence, the visitation frequency of location follows Zipf's law, and the growth of scholars' exploration scope and the number of visited locations both present ultraslow diffusive processes. These results suggest scholars' mobility in knowledge space is characterized by subfield constraint, frequency bias, and proximity bias. Furthermore, we also observe the knowledge boundary effect on the relationship between scholars' exploration scope and academic performance. With the increase of scholars' exploration scope, h-index and the average citations show a trend of firstly increasing and then decreasing, and publication counts and disruption score remain unchanged after increasing, implicating over-exploration will not lead to a sustained increase in academic output and disruption but a decrease in impact. Overall, our study depicts a distinctive portrait of scholars' knowledge exploration behavior and can help design scientific policies.

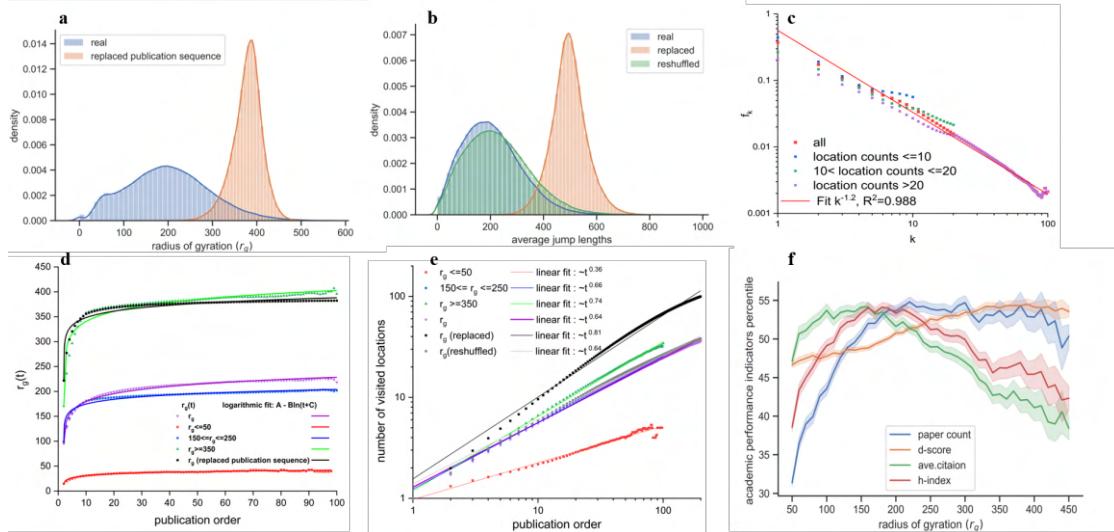


Figure 1. The bounded nature of scholars' trajectories. **a, b.** the distribution of scholars' exploration scope and step lengths, respectively measured by radius of gyration ( $r_g$ ) and jump lengths. **c.** Zipf's plot showing the visitation frequency  $f_k$  of the  $k$ th most visited location. **d,e.**  $r_g$ , number of visited grids on map versus publication sequence. **f.** the correlation between  $r_g$  and academic performance.

## Balanced and fragmented phases in societies with homophily and social balance

T Pham<sup>1,2</sup>, A. C. Alexander<sup>3</sup>, J. Korbel<sup>1,2</sup>, R. Hanel<sup>1,2</sup>, S. Thurner<sup>1,2,4</sup>

<sup>1</sup> CeMSIIS, Medical University of Vienna, Spitalgasse 23, A-1090, Vienna, Austria

<sup>2</sup> Complexity Science Hub Vienna, Josefstadtstrasse 39, A-1090 Vienna, Austria

<sup>3</sup> Department of Mathematics, Princeton University, NJ 08544, Princeton, USA

<sup>4</sup> Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

In the past decades an increasing level of social fragmentation has been observed in societies. We try to understand this phenomenon with an agent-based model where we use spins to represent  $G$ -dimensional vectors of binary opinions of individuals and use a positive (negative) link weight to represent friendship (enmity), respectively. We take into account the joint effects of (1) homophily—the tendency of people with similar opinions to establish positive relations, and (2) social balance—the tendency to establish balanced triadic relations. These two mechanisms are incorporated in a localized Hamiltonian that minimizes social stress through the co-evolution of opinions of individuals and their social networks. Specifically, starting from a random assignment of opinions for each of  $N$  agents, an initial social network is constructed from the similarities between pairs of connected agents. The dynamics is iterated in four steps until the system reaches a steady state: i) we randomly select an agent,  $i$ , and within its social neighbourhood a fraction of  $Q$  triads is randomly chosen among all possible  $N_{\Delta}^{(i)}$  triads of node  $i$ . The social stress  $H$  of  $i$  is calculated; ii) one of  $i$ 's opinions is flipped and the weights of all links which are part of the  $Q$  selected triads are reevaluated. Now we re-calculate the stress of  $i$ ,  $H'$ ; iii) The update is accepted, if either the stress is reduced, i.e.  $H' < H$ , or in case it increases, with probability  $e^{-H'+H}$ , otherwise the opinions and network remain unchanged. We show, for any number of opinions,  $G$ , how the likelihood of social fragmentation increases as individuals care more about social balance. We identify the critical size of the social neighbourhood,  $Q_c$ , above which society must fragment into communities that are internally cohesive and hostile towards other groups.

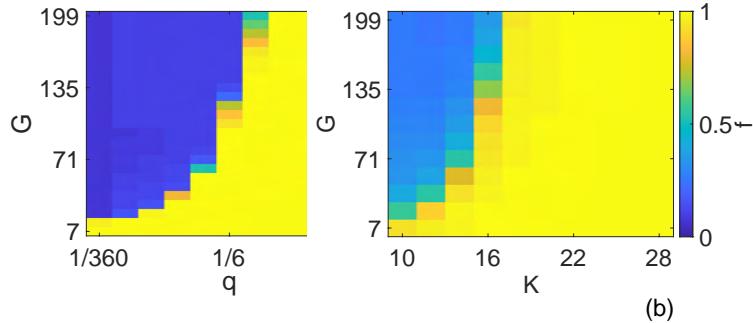


Figure 1: Order parameter,  $f = (n_+ - n_-)/(n_+ + n_-)$ , which measures the level of fragmentation (a) as a function of  $q = Q/N_{\Delta}$  and  $G$  for  $K = 32$  and  $N_{\Delta} = 360$ , and (b) as a function of the degree  $K$  and  $G$  with  $q = 1/3$ . Here  $n_+$ ( $n_-$ ) is the number of balanced (unbalance) triangles in the network. A triangle is called balanced, if the product of its three link states is positive, and unbalanced, if the product is negative.

## References

- [1] T. Pham, A.C. Alexander, J. Korbel, R. Hanel and S. Thurner, arxiv:2012.11221 (2020).

## **Are Terrorist Networks Just Glorified Criminal Cells?**

Elie Alhajjar, Ryan Fameli and Shane Warren

See the full paper published in the Northeast Journal of Complex Systems:

<https://orb.binghamton.edu/nejcs/vol3/iss1/1/>

## **Part VIII**

# **Contributed Talk Session 7**

# COVID-19 impacts on domestic airlines via multilayer network analysis

Mengyuan J. Sun<sup>1</sup> and Naoki Masuda<sup>1,2</sup>

1. Department of Mathematics, University at Buffalo, Buffalo, USA

2. Computational and Data-Enabled Science and Engineering Program, University at Buffalo, Buffalo, USA

The unprecedented pandemic of COVID-19 has cast a major impact on passenger air transportation due to restricted traveling policy and health concerns. Because airline companies are constantly opening and shutting down routes to compete with each other, air transportation networks are temporal networks. To aim to understand such competition, we analyzed an air transportation network data set by month as a multi-layer temporal network in which each layer corresponds to an airline. To examine the possible impacts of COVID-19, we included passenger flight data from January 2015 up to October 2020. The data are from the Bureau of Transportation Statistics. We analyzed the four most major domestic airlines, i.e., American Airlines, United Airlines, Delta Air Lines, and Southwest Airlines. (We analyzed pre-COVID-19 data using a different network analysis method [1].) The airports present in the flight network of any of the four airlines are nodes. A pair of airports is adjacent by an edge if and only if there is a direct commercial flight between them in the month to be analyzed. For simplicity, we assume the network is undirected and unweighted. Not to our surprise, the number of nodes and that of edges considerably decreased during the COVID-19 pandemic (see Fig. 1(a)). To explore the interactions among the four layers, we measured the normalized total overlap, which is defined as the fraction of edges that simultaneously exist in a pair of layers of interest. There was a substantial drop in the normalized total overlap after the start of the COVID-19 pandemic for most pairs of airlines (see Fig. 1(b)), implying that there were fewer routes operated by two airlines at the same time. This might be because the airlines chose to shut down competitive routes to decrease operational costs under the pandemic. Next, we carried out further analysis of the total overlap with the aim of excluding the effect of the decreased number of nodes and edges, which would naturally decrease the total overlap. For each month, we generated 1000 randomized multi-layer networks using the layer-wise configuration model, which preserves the degree of each node within each layer and measured the total overlap for each generated randomized network. Then, we compared the edge overlap for the original multi-layer network with that for the randomized networks, separately for each month, by calculating the Z score. The Z score significantly decreased around the COVID-19 period (Fig. 1(c)). To quantify this observation and account for the problem of multiple comparisons, we calculated the 5% significance level with the Bonferroni correction, which is marked by the horizontal dashed lines in Fig. 1(c). For one airline pair, the Z score is significantly negative during the COVID-19 pandemic but not before. For the other five airline pairs, the Z score is either insignificant or still significantly positive, but its magnitude considerably dropped (i.e., smaller total overlap values) during the pandemic in all cases. Therefore, we conclude that the drop in the edge overlap around June 2020 shown in Fig. 1(b) is not a byproduct of the decrease in the number of the nodes and edges during the same period. In sum, we suggest that the domestic airline network of the US had fewer flight overlaps and less direct competition among the airlines during the COVID-19 pandemic. We are currently measuring other indices that are related to flight route overlaps to try to consolidate this finding.

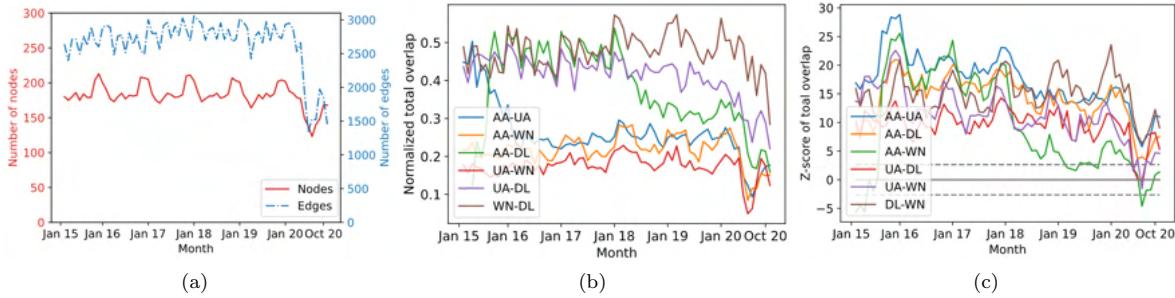


Figure 1: (a) Number of nodes and edges of the aggregated network. (b) Normalized total overlap between different pairs of airlines. (c) Z score of the normalized total overlap. In (b) and (c), each line corresponds to a pair of airlines. In (c), the solid line represents 0, and the dashed lines represent the 5% significance level after the Bonferroni correction.

## References

- [1] K. Sugishita and N. Masuda, *Scientific Reports*, in press (2021); also available at arXiv:2005.14392.

## Complex dynamics in networks, templates and mutated systems

Anca Rădulescu, Mathematics Department, SUNY New Paltz, [radulesa@newpaltz.edu](mailto:radulesa@newpaltz.edu)

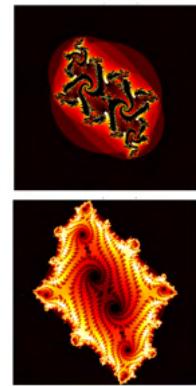
We explore three directions extending the traditional theory of complex quadratic iterations: (1) **complex quadratic networks**; (2) **template iterations** and (3) **mutated iterations**. In all three cases, the system's long-term dynamics is highly non-trivial, and can be represented by asymptotic sets (Mandelbrot and Julia sets) with specific topological signatures, and with properties far beyond those described in the case of single map iterations.

**Network iterations, and applications to brain dynamics [1].** An outstanding question in natural complex networks is how their function depends on the network structure. We use quadratic iterations to develop a new framework for understanding how dynamic behavior emerges in large networks of nodes. Each node is viewed symbolically as an integrator of internal and external inputs. The node-wise dynamics is driven by a traditional complex quadratic map, with coupling specified by the adjacency matrix  $A_{ij}$  and weights  $g_{ij}$ . Then the system takes the form of an iteration in  $\mathbb{C}^n$ , which acts in each component  $z_k$  as  $z_k \mapsto (\sum_{j=1}^n A_{kj} g_{kj} z_j)^2 + c_k$ . This theoretical framework pushes the boundaries of an unexplored field in mathematics, combining theory and computation in complex dynamics and graph theory. We use the asymptotic behavior of multi-dimensional orbits (via the topological and fractal structure of Julia and Mandelbrot sets) to quantify dynamic behavior under perturbation of the network architecture. In particular, topological landmarks of Mandelbrot sets can provide valuable classification criteria for dynamic networks in neuroscience.

**Template iterations, and applications to epigenetics [2].** While single map iterations have been often used to represent natural phenomena, a more realistic mathematical framework is that of time-dependent (non-autonomous) iterations, in which the iterated map may evolve in time. We study the dynamic patterns generated by a pair of complex quadratic functions,  $f_{c_0}$  and  $f_{c_1}$ , applied according to a general binary symbolic sequence  $s$  (which we call *template*), in which the “one” positions correspond to iterating the “correct” function  $f_{c_1}$  and the “zero” positions correspond to iterating the “erroneous” function  $f_{c_0}$ . Template systems generate new aspects of existing problems in non-autonomous iterations. Our interest resides in understanding the dependence of the dynamic behavior on the size and timing of errors (as captured by the complex pair  $(c_0, c_1)$  and by the template  $s$ , respectively). We extended the concepts of Julia and Mandelbrot sets to the case of template iterations, and we study the new topological complexities introduced by this generalization. For example, the Mandelbrot set can be defined in this context as a parameter subset of the product space  $\mathbb{C}^2 \times \{0, 1\}^\infty$ . The structure of Mandelbrot slices in either subspace, can be efficiently used to understand the effects of timing and size of errors in replication systems (such as DNA copying mechanisms).

**Mutated iterations, and application to cell differentiation and tumor growth [3].** Copying errors which occur during mechanisms like cell differentiation, or tumor growth, are highly local, affecting specific genetic loci in the genome structure. If we use the complex plane  $\mathbb{C}$  to symbolically represent a cell’s DNA, each point in the plane corresponds to a single gene. The folding of  $\mathbb{C}$  through the quadratic map  $f_{c_1}(z) = z^2 + c_1$  represents how the emphasis of each gene is recomputed in the replication process, along the specialization path of the cell. This results each time into a new copy of  $\mathbb{C}$ , used as template for subsequent replications. We insert local mutations into the iteration, acting as an erroneous function  $f_{c_0}$  within a small mutation disk around a mutation *center*, and continuously interpolating to  $f_{c_1}$  within a “transition” annulus around the mutation disk. One can then analyze the asymptotic trajectory under iterations of the resulting mutated system  $f$ . Within this symbolic framework, the prisoner set  $\mathcal{P}(f)$  of the system can be seen and studied as the set of sustainable genes, active within the chromosome when the cell had assumed its stable profile, or differentiation.

Examples of network Julia sets for 8 nodes.



## References

- [1] Rădulescu A, Evans S, 2019. Journal of Complex Networks. 7(3); 315-345.
- [2] Rădulescu A, Butera K, Williams B, 2021. Journal of Nonlinear Science. 31(22).
- [3] Rădulescu A, Longbotham A, 2020. ArXiv preprint::2011.14002

# Impact of Community Structure on Consensus Machine Learning

Bao Huynh<sup>1</sup>

Haimonti Dutta<sup>2</sup>

Dane Taylor<sup>1</sup>

<sup>1</sup>Department of Mathematics, University at Buffalo, State University of New York

<sup>2</sup>Management Science and Systems Department, University at Buffalo, State University of New York

February 12, 2021

Consensus dynamics provides a foundation decentralized machine learning algorithms for data that are distributed across a cloud compute cluster or across the internet of things. In these and other settings, one seeks to minimize the time  $\tau_\epsilon$  required to obtain consensus within some  $\epsilon > 0$  margin of error.  $\tau_\epsilon$  typically depends on the topology of the underlying communication network, and for many algorithms  $\tau_\epsilon$  depends on the second-smallest eigenvalue  $\lambda_2 \in [0, 1]$  of the network's normalized Laplacian matrix  $\hat{\mathbf{L}}$ :  $\tau_\epsilon \sim \mathcal{O}(\lambda_2^{-1})$ . We will discuss our recent work [1] in which we analyze the effect on  $\tau_\epsilon$  of network community structure, which can arise when nodes/sensors are spatially clustered into groups, for example. We study consensus learning over networks drawn from stochastic block models (SBMs), and in particular, heterogeneous models that yield random networks containing communities with different sizes and densities. Using recently developed random matrix theory techniques for SBMs [2, 3], we analyze the effects of community structure on  $\lambda_2$  and consensus. Our analysis reveals that there exist two regimes that are separated by a critical amount of community structure as measured by a community prevalence measure (see figure caption). Specifically, above the threshold ( $\Delta > \Delta_1^*$ ), the presence of communities is a significant bottleneck on consensus and consensus-based machine learning algorithms that causes  $\tau_\epsilon$  to grow exponentially. Below the threshold ( $\Delta < \Delta_1^*$ ), communities have little affect on consensus. This observation establishes an important new connection between decentralized machine learning and the network-science literature on detectability theory (see e.g., [3] and the references therein). We support our findings with empirical experiments for decentralized support vector machines (SVM) [4].

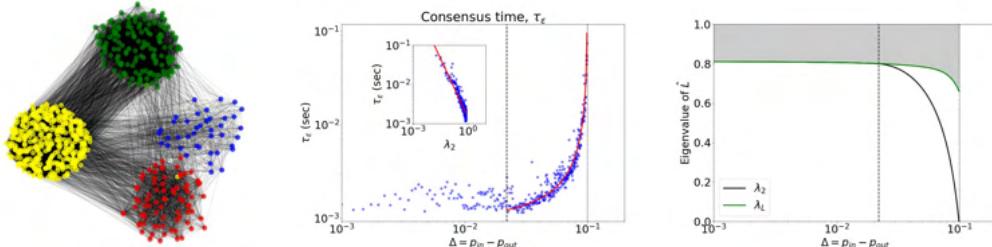


Figure 1: **(left)** Network sampled from a stochastic block model (SBM) with  $K = 4$  communities of sizes  $[n_1, n_2, n_3, n_4] = [40, 80, 150, 400]$ . The edge probabilities within and between communities are given by  $[p_{in}, p_{out}] = [0.5, 0.02]$ . **(center)** Empirical observations of convergence time  $\tau_\epsilon$  (blue dots) versus *community prevalence*  $\Delta = p_{in} - p_{out}$ . The red curve is a line of best fit:  $\tau_\epsilon \approx 0.000087/(0.1 - \Delta)$ . The inset shows  $\tau_\epsilon$  versus  $\lambda_2$ , and the red line indicates the fit  $\tau_\epsilon \approx 0.0015/\lambda_2$ . **(right)** Predicted values for the second-smallest eigenvalue  $\lambda_2$  of  $\hat{\mathbf{L}}$  and the third-smallest eigenvalue  $\lambda_3$ . The vertical lines indicate two critical values: a *spectral bifurcation* occurs at  $\Delta_1^*$  (dashed lines) in that the gap between  $\lambda_2$  and  $\lambda_3$  disappears;  $\tau_\epsilon$  diverges at  $\Delta_2^* = p_{in}$  (dotted lines) since the communities become disconnected and  $\lambda_2 \rightarrow 0$ .

## References

- [1] Bao Huynh, Haimonti Dutta, and Dane Taylor. Impact of community structure on consensus machine learning. *arXiv preprint arXiv:2011.01334*, 2021.
- [2] Tiago P. Peixoto. Eigenvalue spectra of modular networks. *Phys. Rev. Lett.*, 111:098701, Aug 2013.
- [3] Dane Taylor, Saray Shai, Natalie Stanley, and Peter J. Mucha. Enhanced detectability of community structure in multilayer networks through layer aggregation. *Phys. Rev. Lett.*, 116:228301, Jun 2016.
- [4] Haimonti Dutta and Nitin Nataraj. GADGET SVM: A gossip-based sub-gradient solver for linear SVMs. *arXiv preprint arXiv:1812.02261*, 2018.

# Characterization of communities in dynamic functional networks

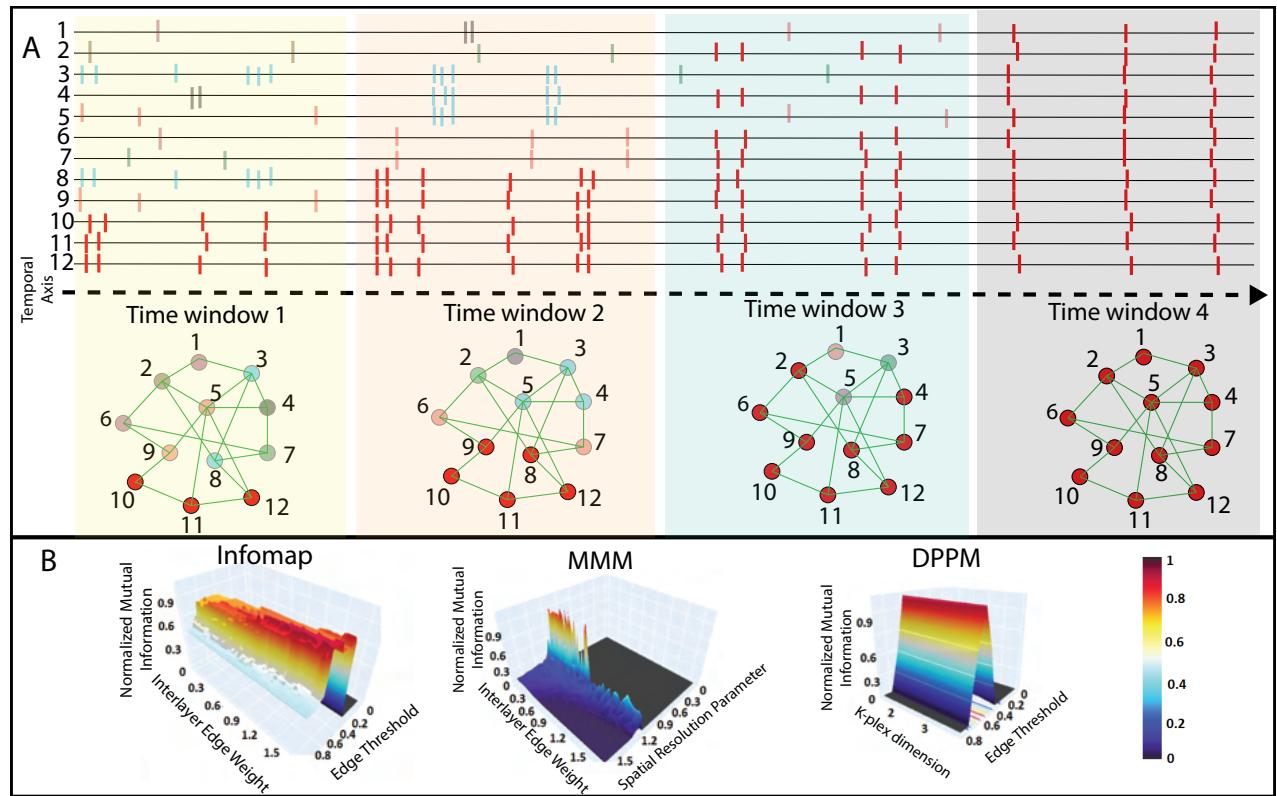
Bengier Ülgen Kılıç<sup>1</sup>, Sarah Feldt Muldoon<sup>1,2,3</sup>

<sup>1</sup> Department of Mathematics, University at Buffalo, SUNY, Buffalo, NY

<sup>2</sup> CDSE Program, University at Buffalo, SUNY, Buffalo, NY

<sup>3</sup> Neuroscience Program, University at Buffalo, SUNY, Buffalo, NY

Grouping similar functioning parts of a system together is a common but difficult task, and doing so for an ongoing dynamic system in which the interactions of these parts vary over time, is even more challenging. Dynamic community detection (DCD) tools for temporal data offer various solutions to this problem. However, how to chose an appropriate DCD algorithm is not always clear, as each algorithm requires making multiple choices about how to (i) model the system, (ii) define a community, and (iii) chose optimal algorithm parameters. Once these assumptions are made, most methods perform a non-unique stochastic partition-space-scan to find the minima of a quality function that mathematically describes the community over a domain. Here, we explored the parameter spaces of a handful of novel-class algorithms for performing community detection in temporal networks—each of which uses different modeling techniques and community definitions. We test algorithms on simulated neuronal spike train data to show that the performance of some algorithms can be improved and parameter spaces can be simplified under specific modeling assumptions extracted from the data. Moreover, algorithms are shown to respond distinct type of community events in the system differently—some capture but others ignore the particular activity. We conclude that the choice of a DCD algorithm must take into consideration the specific properties of the data in order to achieve optimal performance when performing DCD in real-world systems.



**Figure 1:** **A.** Spike trains of a group of neurons leading to fully synchronous activity and the corresponding network evolution where each snapshot is obtained by calculating pairwise synchronization within the respective time window. Each tick indicates a spike and colors represent varying spike rates, whereas node colors correspond to the community assignment (same color within the same time window represents synchronous activity). When dynamic community detection is applied to such a system, different algorithms yield different results because they each define a ‘community’ differently. Nevertheless, these different descriptions tend to spit out similar network partitions in many cases. **B.** Parameter landscapes of different community detection algorithms as a function of normalized mutual information. Note that each algorithm utilizes distinct inputs.

# **Part IX**

# **Contributed Talk Session 8**

## Dimension reduction on heterogeneous networks

Marina Végué<sup>1,2</sup>, Vincent Thibeault<sup>1,2</sup>, Patrick Desrosiers<sup>1,2,3</sup> and Antoine Allard<sup>1,2</sup>

<sup>1</sup> Département de physique, de génie physique et d'optique, Université Laval, Québec, Canada

<sup>2</sup> Centre interdisciplinaire de modélisation mathématique de l'Université Laval, Québec, Canada

<sup>3</sup> CERVO Brain Research Center, Québec, Canada

The complexity inherent to large, non-linear dynamical systems of interacting units (ecological communities, neuronal assemblies, etc.) makes them very difficult to study. One of the big challenges of network science is finding ways to approximate these systems by systems of reduced dimension, which can be more tractable both analytically and computationally [1, 2, 3, 4]. We have addressed this problem for dynamical systems whose units interact through a weighted, directed connectivity matrix. Following the lines of previous work [3, 4], we propose a two-step method for such a dimension reduction that takes into account the properties of the interaction matrix. First, units are partitioned into groups of similar connectivity properties. Each group is associated to one linear observable, that is a weighted average of the node activities within the group. The number of groups thus defines the dimension of the reduced system of observables. Second, we derive a set of conditions that have to be fulfilled for these observables to properly represent the original system's behavior, together with a method for approximately solving them. The result is a reduced interaction matrix and an approximate system of ODEs for the temporal evolution of the observables which is analogous in form to the original system. We show that the reduced system can be used to predict some properties of the complete dynamics (such as bifurcation boundaries in the parameter space) for different types of connectivity structures, including highly heterogeneous networks.

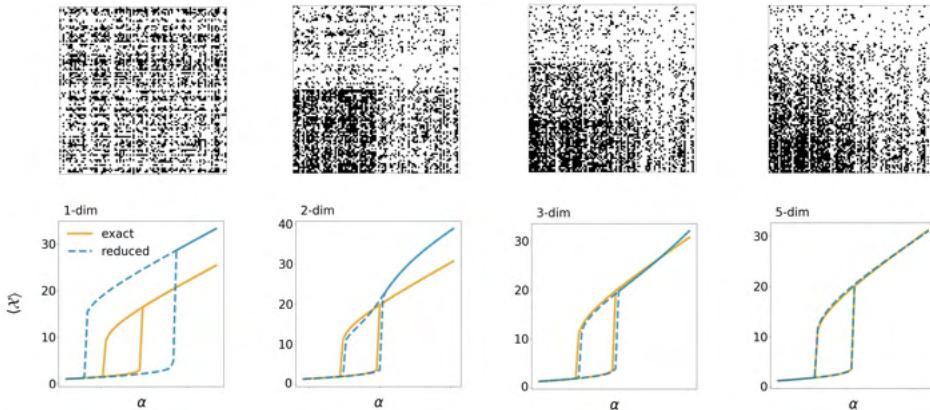


Figure 1: Dimension reduction of a heterogeneous network on 100 nodes with Wilson-Cowan dynamics. From left to right, different node groupings are shown (1, 2, 3, and 5 groups). **Top:** Connectivity matrix, reordered according to the different node groupings. **Bottom:** Bifurcation diagram of the observable dynamics for the different groupings (orange: exact dynamics, blue: approximate, reduced dynamics). The x-axis shows the bifurcation parameter, which controls the overall strength of the connections in the network. The y-axis shows the observables' average at equilibrium.

## References

- [1] J. Gao, B. Barzel, and A.-L. Barabási. Universal resilience patterns in complex networks. *Nature*, 530(7590):307–312, 2016.
- [2] J. Jiang, Z.-G. Huang, T. P. Seager, W. Lin, C. Grebogi, A. Hastings, and Y.-C. Lai. Predicting tipping points in mutualistic networks through dimension reduction. *PNAS*, 115(4):E639–E647, 2018.
- [3] E. Laurence, N. Doyon, L. J. Dubé, and P. Desrosiers. Spectral dimension reduction of complex dynamical networks. *Phys. Rev. X*, 9(1):011042, 2019.
- [4] V. Thibeault, G. St-Onge, L. J. Dubé, and P. Desrosiers. 40 Threefold way to the dimension reduction of dynamics on networks: An application to synchronization. *Phys. Rev. Research*, 2:043215, 2020.

# Active Learning and Relevance Vector Machine in Efficient Estimate for Basin Stability of Dynamic Networks

Yiming Che and Changqing Cheng

Department of Systems Science and Industrial Engineering  
Binghamton University, State University of New York  
ccheng@binghamton.edu

## Abstract

The interconnectivity between constituent nodes gives rise to the cascading failure in most dynamic networks, such as traffic jam in transportation network and large-scale blackout in power grid system. Basin stability (BS) has recently garnered tremendous traction to quantify the reliability of such dynamical systems. In the power grid network, it quantifies the capability of the grid to regain the synchronous state after being perturbated. It is noted that detection of the most vulnerable node or generator with the lowest BS or the N-1 reliability is critical towards the optimal decision making on maintenance. However, the conventional estimation of BS relies on Monte Carlo (MC) method to separate the stable and unstable dynamics originated from the perturbation, which incurs immense computational cost particularly for large-scale networks. As BS estimate is in essence a classification problem, we investigate relevance vector machine (RVM) and active learning to locate the boundary of stable dynamics or the basin of attraction in an efficient manner. This novel approach eschews the large number of sampling points in MC method and reduces over 95% of the simulation cost in the assessment of N-1 reliability of power grid network. The result of experiment on the IEEE 118-Bus system in Fig. 1 demonstrates the effectiveness and efficiency of the proposed method.

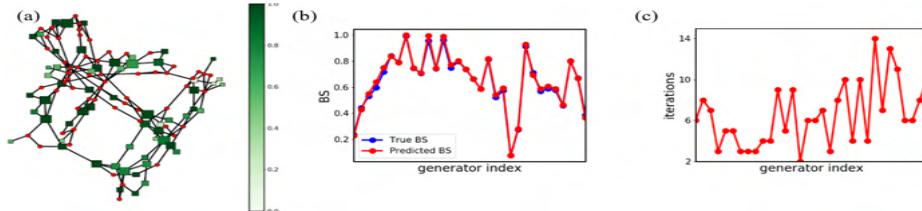


Figure 1: (a) Topology of IEEE 118-Bus system. The red circular nodes are loads and the square nodes are generators. The color of the generators implies the magnitude of the BS and the size of generator indicates the degree, i.e., the number of connected edges; (b) comparison of network BS for different generators in the IEEE 118-Bus system using MC method and the proposed approach; (c) the number of iterations in active learning on each node.

# Visualizing trajectory of tie-decay temporal networks

Chanon Thongprayoon<sup>1</sup>, Lorenzo Livi<sup>2</sup>, Naoki Masuda<sup>1,3</sup>

<sup>1</sup>Department of Mathematics, University at Buffalo, SUNY, New York, USA

<sup>2</sup>Departments of Computer Science and Mathematics, University of Manitoba, Manitoba, Canada

<sup>3</sup>Computational and Data-Enabled Science and Engineering Program, University at Buffalo, SUNY, New York, USA

To aid visualization and data analysis of a set of networks, we often map the networks to a low dimensional vector space (e.g., Euclidean space). This procedure is called network embedding. Generally, when embedding networks, the (dis)similarity in the embedding space is preserved as much as possible. In temporal networks, the network evolves over time, and it is usually the case that we only obtain discrete events occurring at discrete times. Using network embedding, we consider the problem of constructing a continuous-time trajectory from discrete time-stamped events occurring on edges of a network.

We consider embedding of temporal networks, where at least one contact event occurs at time  $t_i$  (with  $i = 0, 1, \dots$ ) in the network. We represent the network evolution over time with a time-dependent adjacency matrix  $B(t) = [b_{ij}(t)]$ . Embedding yields a continuous-time trajectory that represents dynamics of the given temporal network. We use a tie-decay network model in which the weight of each edge exponentially decays at a rate  $\alpha (> 0)$  in the absence of contact events, i.e.,  $b_{ij}(t) = e^{-\alpha(t-t_k)} b_{ij}(t_k)$  for any  $t \in [t_k, t_{k+1})$ , and  $b_{ij}(t_k)$  instantly increases by 1 if there is a contact event between the  $i$ th and  $j$ th nodes at time  $t_k$  [1]. To embed the temporal network into a Euclidean space and determine its trajectory, firstly, we propose to construct a matrix encoding the distance between pairs of networks at different discrete times,  $t_0, t_1, \dots$ , but crucially not for  $t \notin \{t_0, t_1, \dots\}$ . Next, we apply the landmark multidimensional scaling (LMDS) algorithm to the distance matrix [2]. The LMDS, an extension of multidimensional scaling (MDS), uses a selected subset of the data points as reference points (i.e., landmarks) to compute the embedding coordinates of any other data points. In this fashion, we produce trajectories of temporal networks in continuous time. The trajectories for each of two empirical temporal network data and one synthetic, periodic input of contact events are shown in Fig. 1. We chose a Laplacian network distance measure. (We also used two other distance measures; results not shown.) For example, Fig. 1(b) illustrates that the network varies relatively slowly between  $t = t_{350}$  and  $t = t_{400}$  compared to between  $t = t_{50}$  and  $t = t_{100}$ . As another example, Fig. 1(c) captures periodic nature of the input and suggests that the trajectory approaches a limit cycle, corresponding to the periodic nature of  $B(t)$ .

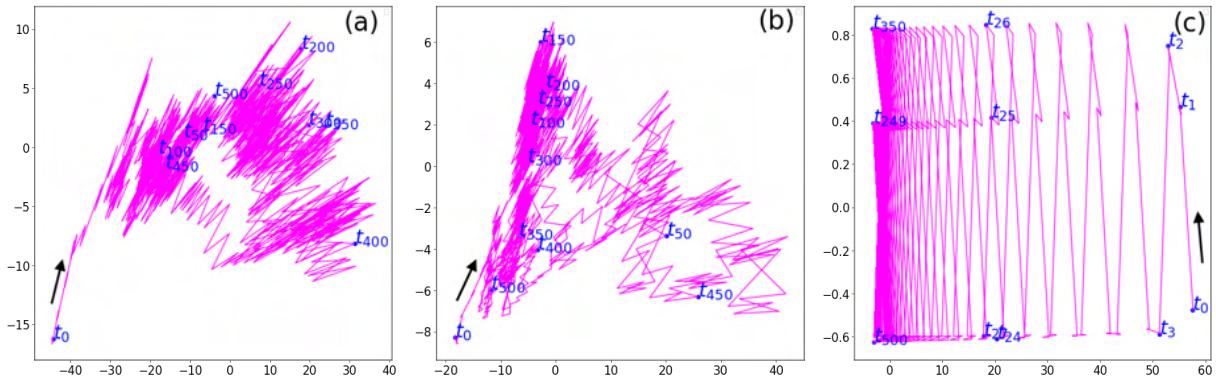


Figure 1: Trajectories of temporal networks. (a) Proximity between students in a primary school. (b) Proximity between students in a high school. In (a) and (b), the data are from the SocioPatterns Project, and we have set  $\alpha = 0.01$  and only used the first 501 time points at which the events are recorded. (c) A periodic, synthetic input of adjacency matrices with 9 nodes that arrive with regular intervals with period four. We set  $\alpha = 0.04$ .

## References

- [1] W. Ahmad, M. A. Porter, and M. Beguerisse-Díaz. arXiv:1805.00193, 2018.
- [2] V. De Silva, and J. B. Tenenbaum. Technical Report, Stanford University, 2004. [Online]. Available: [http://graphics.stanford.edu/courses/cs468-05-winter/Papers/Silva\\_landmarks5.pdf](http://graphics.stanford.edu/courses/cs468-05-winter/Papers/Silva_landmarks5.pdf)

# Dimension reduction of high-dimensional dynamics on networks with adaptation

Vincent Thibeault<sup>1,2</sup>, Marina Vegué<sup>1,2</sup>, Antoine Allard,<sup>1,2</sup> and Patrick Desrosiers<sup>1,2,3</sup>

1. Département de physique, de génie physique et d'optique, Université Laval, Québec (QC), G1V 0A6, Canada

2. Centre interdisciplinaire de modélisation mathématique de l'Université Laval, Québec (QC), G1V 0A6, Canada

3. CERVO Brain Research Center, Québec (QC), G1J 2G3, Canada

Adaptation is a defining property of complex systems. It is characterized by changes in the structure and the dynamics of the system according to its environment and its own behavior [1]. For instance, neurons in the brain are connected by synapses that increase or decrease in strength as a response to neuronal activity. Synaptic plasticity is known to be an essential mechanism of learning and memory, but its precise role and its global influence on brain activity remain unclear [2]. From a theoretical standpoint, the impact of adaptation on the neuronal dynamics' equilibrium points is hard to predict. This is partly due to the high dimensionality of the dynamical system describing adaptative neuronal networks. Indeed, in addition to the ordinary differential equations (ODEs) describing the activity of all neurons (nodes), there is another large set of ODEs governing the evolution of the strength (weight) of all synapses (edges).

Based on our previous works [3-4], we introduce a new dimension reduction framework that systematically yields a low-dimensional (reduced) adaptative dynamics from a high-dimensional (complete) adaptative dynamics. The reduced dynamics accurately describes the complete Wilson-Cowan dynamics with three different adaptation rules: Hebb's rule, Oja's rule, and even the biologically plausible Bienenstock-Cooper-Monroe's (BCM) rule [5]. The Wilson-Cowan dynamics with the BCM rule exhibits rich bifurcation phenomena that are well predicted by the reduced adaptative dynamics [FIG. 1]. For instance, our dimension reduction framework captures the emergence of surprising nonlinear oscillations in the firing rate and the synaptic strength which appear through a supercritical Hopf bifurcation.

Our framework is flexible: we can tune the number of observables as we want to describe adaptative networks with more modules and heterogeneity. It also unlocks the possibility to perform dynamical analysis on large real networks which paves the way towards a better understanding of emergent phenomena in the plastic brain.

[1] M. Mitchell, *Complexity : A Guided Tour*, (Oxford University Press, New York, 2009).

[2] Y. Humeau and D. Choquet, "The next generation of approaches to investigate the link between synaptic plasticity and learning", *Nat. Neurosci.* **22**, 1536 (2019).

[3] V. Thibeault, G. St-Onge, L.J. Dubé, and P. Desrosiers , "Threefold way to the dimension reduction of dynamics on networks: An application to synchronization", *Phys. Rev. Research* **2**, 043215 (2020).

[4] E. Laurence, N. Doyon, L.J. Dubé, and P. Desrosiers, "Spectral Dimension Reduction of Complex Dynamical Networks". *Phys. Rev. X* **9**, 011042 (2019).

[5] L. Cooper and M. Bear "The BCM theory of synapse modification at 30: interaction of theory with experiment", *Nat. Rev. Neurosci.* **13**, 798 (2012).

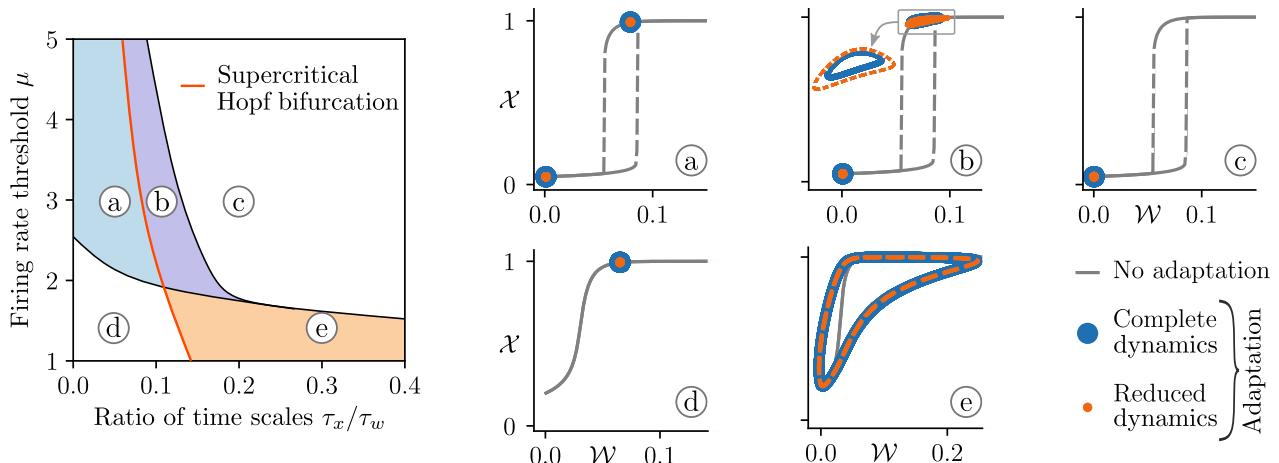


FIG. 1. Bifurcation diagram of the Wilson-Cowan dynamics with the BCM rule (10 200 ODEs) obtained from its reduced dynamics (3 ODEs). The firing rate threshold  $\mu$  is the midpoint of the sigmoidal activation function of the Wilson-Cowan dynamics and  $\tau_x$  is the time scale of the activity while  $\tau_w$  is the one for the weights. Figures a-e illustrate the equilibrium points of the mean global activity observable  $\chi$  vs. the mean global weight observable  $\mathcal{W}$  when there is no adaptation and when there is adaptation for the complete and reduced dynamics. These figures show that the reduced dynamics predicts accurately the equilibrium points in every region of the bifurcation diagram. We choose a weighted Erdős-Rényi network with  $N = 100$  and  $p = 0.5$  as an initial condition for the synaptic dynamics. 43

# **Part X**

## **Poster Session 1**

**Flocc: From Agent-Based Models to Interactive Simulations on the Web**

Scott Donaldson

See the full paper published in the Northeast Journal of Complex Systems:

<https://orb.binghamton.edu/nejcs/vol3/iss1/6/>

**Entropic dynamics of networks**

Felipe Xavier Costa and Pedro Pessoa

See the full paper published in the Northeast Journal of Complex Systems:  
<https://orb.binghamton.edu/nejcs/vol3/iss1/5/>

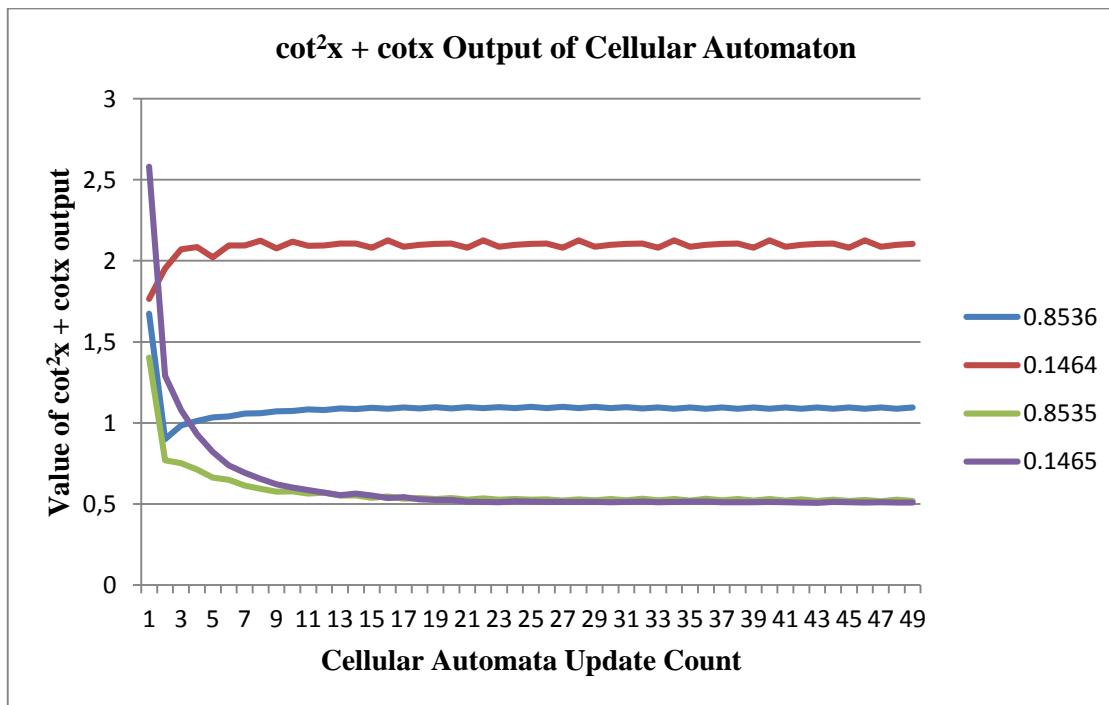
# Exact Formulation of Ising Model Transitions Between Six Magnetic Phases

Goktug Islamoglu

B.Sc. Civil Engineering, Istanbul Technical University

M.Sc. Occupational Health and Safety, Karadeniz Technical University

In a conventional cellular automaton, a cell's state is updated in relation to its neighbors. In this work the author explores has reversed the cellular automaton's cell update rule, allowing a cell dictate the states of its neighbors<sup>1</sup>. The rule forces the corners of the Moore neighborhood to update into von Neumann neighborhood. Author hypothesizes that this is a decimation. By employing configurations that tune the automaton's neighborhood based on the cell's state (by default 6-neighborhood tuning is preferred), six magnetic responses are plotted respectively in the probability range: demagnetization (degaussing), ferrimagnetic, ferromagnetic, paramagnetic, antiferromagnetic and again paramagnetic. The maximum count of cells with state 1 corresponds to an evasion curve (inverse of a pursuit curve) that evades the inverse Ising critical temperature. If tuning is brought to less than 6 neighbors, instead the maximum cell count corresponds to Ising criticality. This is achieved by coupling the system with a  $p^*(1-p) = 1/8$  equation (neighborhood average). The first order transition points are the roots of this coupling equation  $\frac{1}{2} \pm \frac{1}{2\sqrt{2}}$ , which are exactly representable by trigonometric values of  $\cos^2\left(\frac{\pi}{8}\right)$  and  $\sin^2\left(\frac{\pi}{8}\right)$ . This is the basis of the trigonometric transformation, when plotted, a  $\cot^2x + \cot x$  graph emerges in absence of any trigonometric functions present in the automaton's code. Arms of this graph are the roots given above, and the immediate values outside of first order transition boundary. By solving this cotangent equation, the author has come upon a renormalization recursion, thus reaffirming his hypothesis of renormalization simulating cellular automaton. The recursion reveals exact equations that correspond to phase transitions unique to each magnetic phase.



**Keywords:** Cellular Automata, Ising Model, Magnetic Phase Transitions

1. <https://github.com/goktu/ADama/blob/master/cellularautomata.py> 47

# Network Reconstruction

## Bayesian Sequential Inference of Sparse Network Connectivity

Jun Kataoka, Jayanth Sivakumar, Eric Peña, Yiming Che, Changqing Cheng

Department of Systems Science and Industrial Engineering  
Binghamton University, State University of New York

While significant effort has been devoted to the design, control, and optimization of complex networks, most existing work assumes the network structure is known or is readily available. However, the network topology could be radically altered (e.g., due to an adversarial attack or power outage) and remain unknown for subsequent analysis. In this work, we propose a novel Bayesian sequential learning algorithm to adaptively reconstruct network connectivity. A sparse Spike-and-Slab prior distribution is placed on all edges and the connectivity learned from the reconstructed nodes will be incorporated to select the next node and will then update the prior knowledge. Central to our approach is that most realistic networks are sparse, in that the degree of each node is much smaller compared to the total number of nodes in the network. A sophisticated method of sequentially selecting nodes is implemented using the between-node expected improvement. The performance is measured and compared against randomly selecting nodes. Our algorithm is an improvement over traditional reconstruction efforts when faced with limited data and is robust to varying levels of noise. That is, our algorithm accurately reconstructs the network with fewer nodes than the whole network structure. This algorithm has been applied to real network data: IEEE 118-Bus System provided by the Illinois Center for a Smarter Electric Grid (ICSEG) whereby the time-series for each node has been simulated. The performance of our algorithm compared to random node selection is shown in Fig. 1. This algorithm has also been applied to the Barabasi-Albert network for  $m = 1$  and  $m = 2$ . This study has the potential to significantly scale up the reconstruction algorithms to accommodate massive networks and radically transform the operation of various realistic networked systems including power grid, transportation, and communication networks particularly in a host of military operation scenarios.

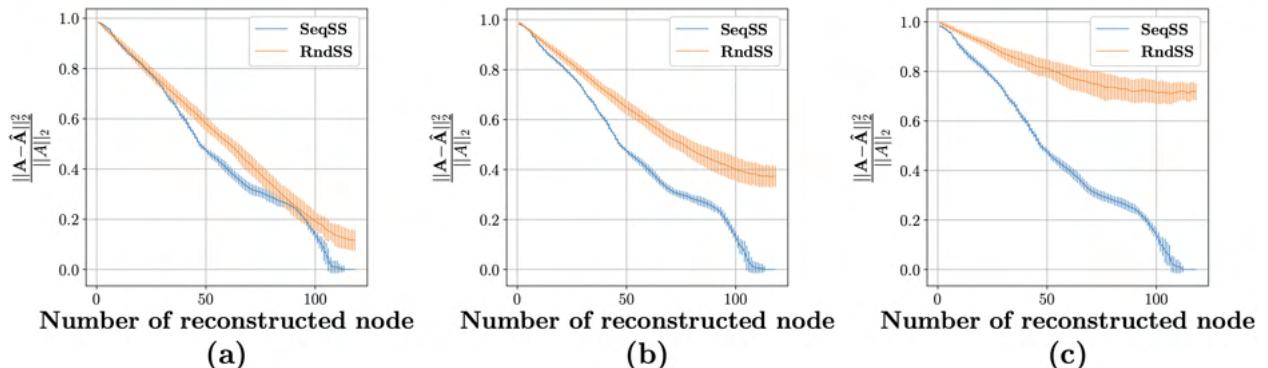


Figure 1: Reconstruction error for the proposed algorithm and random selection for noise levels (a)  $\sigma = 0.03$ , (b)  $\sigma = 0.04$ , and (c)  $\sigma = 0.05$

# The effect of time-dependent infectiousness on epidemic dynamics

Nicholas W. Landry<sup>1</sup> and Karen L. Stengel<sup>2</sup>

<sup>1</sup> Department of Applied Mathematics, University of Colorado Boulder  
nicholas.landry@colorado.edu

<sup>2</sup> Department of Computer Science, University of Colorado Boulder  
karen.stengel@colorado.edu

## Abstract

In contrast to the common assumption in epidemic models that the rate of infection between individuals is constant, in reality, an individual's viral load determines their infectiousness. We compare the dynamics of an epidemic model with a deterministic, time-dependent infection rate to that of a standard SIR model with a constant, Markovian rate of infection and healing. We predict the epidemic threshold in terms of contact structure for fully-mixed populations and category-mixed populations. We compare the trajectories of epidemic spread on both fully-mixed and category-mixed populations using both an SIR model and a time-dependent viral load mean-field model. We find that the reproductive number only depends on the total infectious *exposure* and the largest eigenvalue of the mixing matrix and that the two are independent of each other. We find that when we compare the viral load model to the SIR model, the epidemic peak is delayed but more pronounced. In addition, the specific infection rate function has a strong effect on the time dynamics of the epidemic.

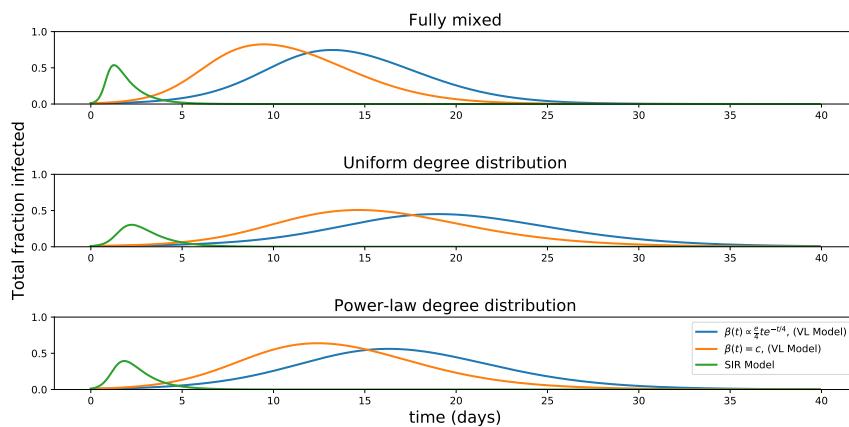


Figure 1: Plot of the infection curves for different population structures and epidemic models, where the total infection and healing rates are identical for each model.

## Towards a multiplex network model of word associations and similarity in the human mind

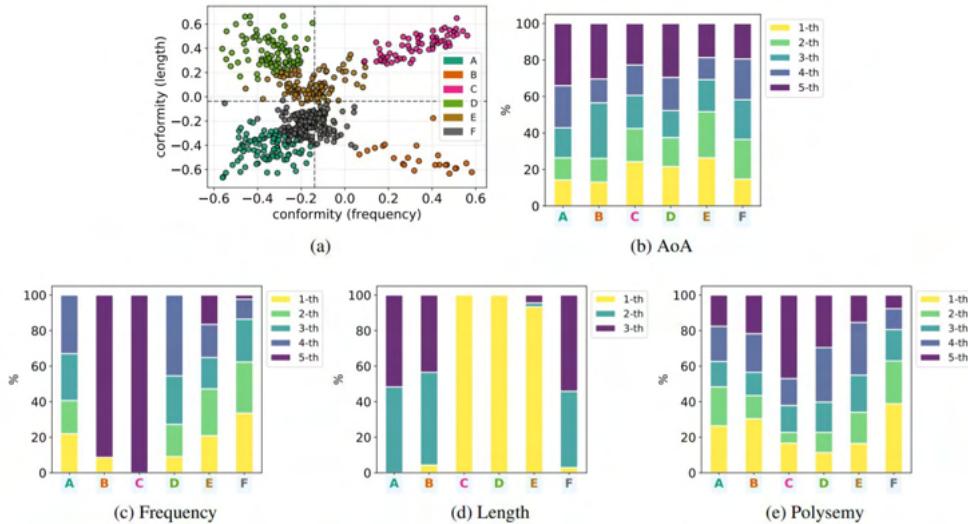
Salvatore Citraro, Department of Computer Science, and ISTI-CNR, Italy - [salvatore.citraro@phd.unipi.it](mailto:salvatore.citraro@phd.unipi.it)

Massimo Stella, Department of Computer Science, University of Exeter, UK - [m.stella@exeter.ac.uk](mailto:m.stella@exeter.ac.uk)

Giulio Rossetti, Istituto di Scienza e Tecnologie dell'Informazione (ISTI), CNR, Italy - [giulio.rossetti@isti.cnr.it](mailto:giulio.rossetti@isti.cnr.it)

This work suggests a unified multiplex network framework enriching network layers of multiple interaction types (structural layers) together with layers of similarity patterns between nodes (similarity layers). This approach gives emphasis to structural links but it also considers potential features of individual nodes that are often neglected in network models (Citraro and Rossetti, APN, 2020). We show the benefit of considering similarity in synergy with structure when modelling the human mental lexicon, a cognitive complex system where conceptual units/words possess features as reflected in language (e.g. frequency, polysemy or length) and can be associated with each other according to mechanisms like recall from memory, overlap in meaning or phonological similarities (see also Stella et al., Sci.Rep., 2017).

Building semantic and phonological structural layers in addition to similarity layers spanning vector spaces across frequency, polysemy and length, we reconstruct a multidimensional representation of the mental lexicon including 529 English words and available to a population of over 1000 English toddlers between 18 and 36 months. Considering the network metric of conformity (Rossetti et al., IEEE Int.Sys., 2021), we investigate correlations between structural and similarity layers. Conformity is able to unveil the heterogeneous mixing patterns hidden in complex attributed graphs, and considers the evidence that nodes with similar characteristics are not divided by long chains. Modelling each node as a vector of its conformity values w/r/t the single features, we highlight the presence of multiple word clusters that could not be observed on structural layers only. As reported in Figure 1 (a), conformity identifies 6 clusters of nodes differing in their tendencies to connect (structurally) with other nodes sharing (similar) features. In particular, a core of words of short length, high frequency and several meanings is highlighted by this methodology (Figure 1 (c-e)). This core emerges early on during word development? Figure 1 (b) can suggest a prevalence of early acquired words, but more statistical tests and analyses are needed to assess some possible roles of such clusters in the development of children's mental lexicon.



**Figure 1:** (a) Conformity w/r/t frequency vs. conformity w/r/t length: point colors denote node memberships to a cluster -- K-Means as method, k=6 according to the elbow method; (c-e) percentages of feature values within the clusters, e.g., the cluster labeled as C contains only (c) very frequent and (d) short words, and (e) nearly half of them convey a high number of meanings; (b) distribution of word age of acquisition is also presented for each cluster.

# Configuration models for hypergraphs preserving local quantities of nodes and hyperedges

Kazuki Nakajima<sup>1</sup>, Kazuyuki Shudo<sup>1</sup>, and Naoki Masuda<sup>2,3</sup>

<sup>1</sup> Department of Mathematical and Computing Science, Tokyo Institute of Technology

<sup>2</sup> Department of Mathematics, University at Buffalo

<sup>3</sup> Computational and Data-Enabled Science and Engineering Program, University at Buffalo

Empirical hypergraph data, where unit interactions include those among three or more nodes in general, are increasingly available [1]. By extending a class of configuration models called the  $dK$ -series that are originally proposed for dyadic graphs [2], we propose a family of configuration models for hypergraphs that preserve local properties of the given hypergraph to different extents. The proposed model specifies the joint degree distributions of nodes in the subgraphs of size  $d_v$  or less and the joint degree distributions of hyperedges in the subgraphs of size  $d_e$  or less for the bipartite graph that corresponds to the given hypergraph. We consider  $d_v \in \{0, 1, 2, 2.5\}$  and  $d_e \in \{0, 1\}$ .

Figure 1 shows four properties of nodes, i.e., (a) node's degree distribution (DD), (b) node's degree correlation (DC), (c) so-called degree-dependent node redundancy coefficient (DRC), and (d) the distribution of the shortest path length (SPL), for an empirical hypergraph data and the corresponding configuration models with  $d_e = 0$ . The empirical data is a drug network from the national drug code directory (NDC-classes) data set with 1,149 nodes and 1,047 hyperedges. The model with  $d_v = 0$  only intends to preserve the number of edges in the original bipartite graph. Therefore, it does not accurately approximate the four node's quantities. The model with  $d_v = 1$  preserves the node's degree distribution of the given hypergraph (Fig. 1(a)), as it intends, but not the other three quantities (Fig. 1(b)–(d)). The model with  $d_v = 2$  preserves the node's degree distribution and roughly preserves the degree correlation (Fig. 1(a)–(b)). The model with  $d_v = 2.5$  further captures the abundance of triadic relationships (Fig. 1(c)). Furthermore, as  $d_v$  increases from 0 to 2.5, the model better approximates the distribution of the shortest path length although the model is not designed to preserve it (Fig. 1(d)). The present family of configuration models is expected to serve as reference models when one investigates the structure and dynamics of empirical and synthetic hypergraphs.

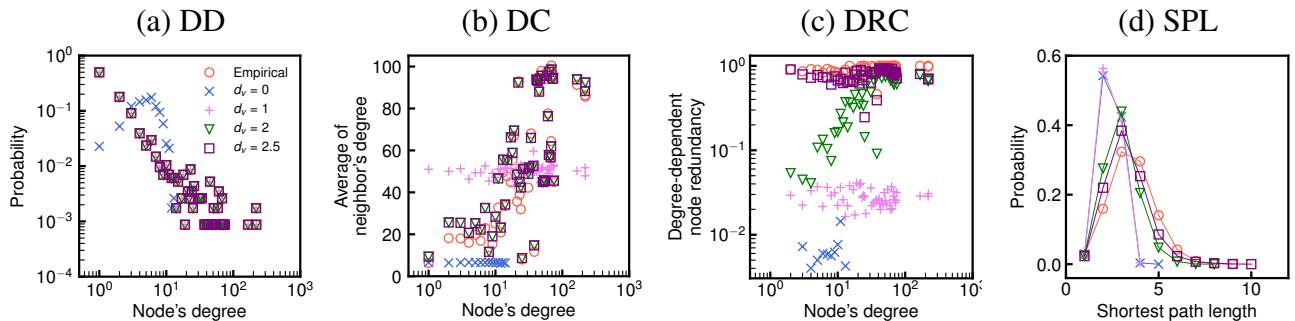


Figure 1: Distributions of node properties for the configuration models for hypergraphs. We set  $d_e = 0$ .

## References

- [1] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J. -G. Young, and G. Petri. Networks beyond pairwise interactions: structure and dynamics. *Physics Reports*, 874:1–92, 2020.
- [2] C. Orsini, M. M. Dankulov, P. Colomer-de Simón, A. Jamakovic, P. Mahadevan, A. Vahdat, K.E. Bassler, Z. Toroczkai, M. Boguñá, G. Caldarelli, S. Fortunato, and D. Krioukov. Quantifying randomness in real networks. *Nature Communications*, 6:8627, 2015.

## Correlation of community-aware and classical centrality measures: Examining the role of network topology

Stephany Rajeh<sup>1</sup>, Marinette Savonnet<sup>1</sup>, Eric Leclercq<sup>1</sup>, Hocine Cherifi<sup>1</sup>

<sup>1</sup> Computer Science Department - University of Burgundy, Dijon, France

This work investigates the relationship between community-aware centrality measures, classical centrality measures, and network topology. Artificial networks with controlled mesoscopic properties and real-world networks originating from various domains are used in the experiments. Kendall's Tau correlation quantifies the interactions between ten classical centrality measures and their corresponding community-aware version based on intra-community links (local measures) and inter-community links (global measures)[1, 2]. Results show that the mixing parameter is the most influential mesoscopic property in artificial networks. Global community-aware centrality measures exhibit low correlation with classical centrality measures in networks with a strong community structure, while local community-aware centralities show high correlation. One observes the inverse in networks with a weak community structure. Furthermore, correlation values are relatively insensitive to the degree distribution's exponent variations and the community size distribution. Simple linear regression between the real-world networks' topological features and the mean correlation between classical and community-aware centralities is computed. Results confirm the influence of the community structure strength. Furthermore, it appears that density, transitivity, and efficiency are the most significant macroscopic features affecting the mean correlation between classical and local community-aware centrality measures.

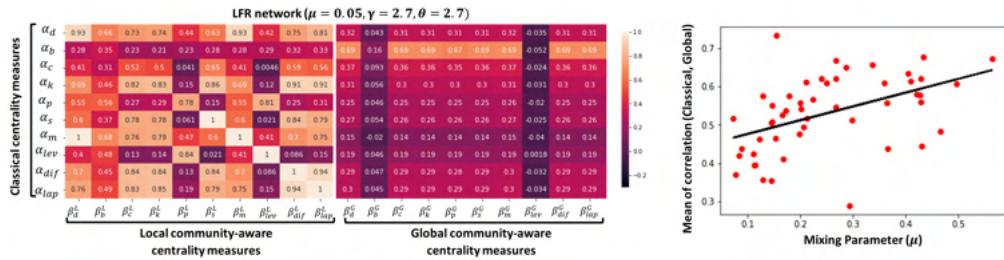


Figure 1: Left: Kendall's Tau correlation between classical ( $x$ -axis) and community-aware ( $y$ -axis) centrality measures. Right: Mean correlation versus mixing parameter of 50 real-world networks.

## References

- [1] N. Gupta, A. Singh and H. Cherifi. Community-based immunization strategies for epidemic control. IEEE COMSNETS (2015), doi: 10.1109/COMSNETS.2015.7098709
- [2] Ghalmame, Z, et al. Centrality in Complex Networks with Overlapping Community Structure. Sci Rep 9, 10133 (2019), doi: <https://doi.org/10.1038/s41598-019-46507-y>

## The Mixing Law

Jake Ryland Williams<sup>1</sup> and Diana Solano-Oropeza<sup>2</sup>

Departments of <sup>1</sup>Information Science and <sup>2</sup>Physics, Drexel University

{jw3477, ds3367}@drexel.edu

For text, Zipf's Law (ZL) is the existence of a harmonic relationship in the usage of a document's vocabulary. So  $f(r_w) \propto r_w^{-1}$  means each word,  $w$ , is ordered by a positive integer rank  $r_w$  that is equal to its waiting time between appearances (in a shuffling of the document). However, documents exhibit variation in the scaling exponent, and so ZL is often weakened as  $f(r_w) \propto r_w^{-\theta}$  for some  $\theta \geq 0$ . Moreover, while smaller samples of text generally uphold ZL with  $\theta \approx 1$ , larger samples are known to exhibit words at higher ranks/lower frequencies that deviate far from ZL. In these circumstances, a *scaling break* is said to occur at some critical rank  $b$ , after which a second, more-severe scaling emerges, i.e.,  $f(r_w) \propto r_w^{-\theta}$  for  $r_w < b$ , but then  $f(r_w) \propto r_w^{-(\theta+\mu)}$  when  $r_w \geq b$ , for some *different* value  $\mu > 0$ . Text-mixing was proposed to explain this non-ZL variation at large ranks and exhibited explanatory power in sampling experiments. But that work's formalism excluded a functional relationship, which we propose as the *Mixing Law* (ML).

Define text mixing mathematically for a set of  $k$  documents  $\mathcal{D} = \{d_i\}_{i=1}^k$  by assuming each upholds ZL as  $f(r_w | d_i) \propto r_{w,i}^{-\theta_i}$  over a vocabulary of  $N_i = |d_i|$  words via a scaling exponent  $\theta_i$ . The goal is to simplify the *mixture frequencies*:  $\bar{f}(r_w | \mathcal{D}) = \sum_{i=1}^k f(r_w | d_i)$ . This is challenged by variation in each word's ranks across the documents, i.e., the distribution of  $\{r_{w,i}\}_{i=1}^k$  (local ranks). Now, ZL constrains the variation in exponents  $\{\theta_i\}_{i=1}^k$  so that there exists  $\theta$  with  $\theta_i \approx \theta \approx 1$  across  $\mathcal{D}$ . If for each word  $w$  of *mixture-rank* (global rank)  $r_w$  one models  $P(r_w | N_i)$  as the chance a document  $d_i \in \mathcal{D}$  of size  $N_i = |d_i|$  contains  $w$ , then:  $\bar{f}(r_w | \mathcal{D}) \propto r_w^{-\theta} \cdot P(r_w | \langle N \rangle)$ , models the mixture frequencies (using  $\langle N \rangle$  to denote centrality for  $\mathcal{D}$ 's vocabulary sizes).

Provided the mixture's vocabulary has size  $N$ , let  $H_N = \sum_{r=1}^N r^{-1}$  be the  $N^{\text{th}}$  harmonic number, so that  $\langle r \rangle = \frac{N}{H_N}$  averages mixture ranks. Then  $\theta \approx 1 - H_N^{-1}$  approximates the replication rate for Simon's model from the strong form of ZL over  $N$  harmonics. We propose approximating the ML's form and role as:

$$\bar{f}(r_w | \mathcal{D}) \approx r_w^{-\theta} \cdot P(r_w | \langle N \rangle) \propto \underbrace{r_w^{-(1-H_N^{-1})}}_{\text{Zipf's Law}} \cdot \underbrace{\left[ 1 - \left( 1 + \frac{\langle r \rangle}{r_w} \right)^{-\frac{\langle N \rangle}{\langle r \rangle}} \right]}_{\text{Mixing Law}}$$

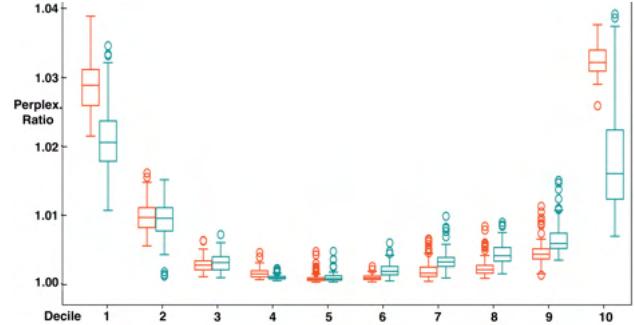


Figure 1: Perplexity ratios for  $\langle N \rangle$  &  $N$  measured by corpus (numer) and regression (denom) on samples of 35 documents by decile. Small/large deciles (1–3/7–10) of a synthetic corpus (orange) are fragmentations/agglomerations of real documents (blue), taken from deciles 4–7.

## Temporal Motifs in Patent Opposition and Collaboration Networks

Penghang Liu<sup>1</sup>, Tomomi Kito<sup>2</sup>, Naoki Masuda<sup>1</sup>, Ahmet Erdem Sarıyüce<sup>1</sup>

<sup>1</sup> University at Buffalo, <sup>2</sup> Waseda University

Patents are intellectual properties that often reflect innovative activities of companies and organizations. Many studies have investigated the citations among the patents, but only a few addressed the relations between the patent owners. We use patent data from Orbis IP database [1] to analyze the adversarial and collaborative relations among the companies, where the former is formed by the patent opposition, a legal activity in which a company challenges the validity of a patent, and the latter is implied by the co-ownership of patents by multiple companies. Characterizing the patent oppositions, collaborations, and the interplay between them is important for understanding how innovation happens. Temporality is an important aspect in this context as the order and frequency of oppositions and collaborations are key to identify and characterize complex relations among companies.

In this study, we construct a two-layer temporal network to model the patent oppositions and collaborations among the companies from 1980 to 2018. We utilize temporal motifs [2] to analyze the oppositions and collaborations from structural and temporal perspectives. We have three main findings. First, the opposition layer contains significantly more wedge motifs than the triangle motifs (see Table 1), which is coherent with the structural balance theory [3] since oppositions are negative relations. Second, if a company receives a burst of oppositions, it is likely to be a large company which is attacked by small companies, and if a company files a burst of oppositions, it is usually a large company targeting other large companies. Third, the companies with collaborations in the past are likely to oppose the same company or be opposed by the same company. In summary, our analysis discovers underlying business relations in the patent data, which we believe will advance the studies in business strategy and management.

449.81	147.60	166.12	26.43	16.75

Table 1: Temporal motifs in patent oppositions with their Z scores. Nodes represent companies and the edges are the patent oppositions. The numbers on the edges denote the temporal order of the oppositions. A higher Z score indicates a more significantly overrepresented pattern.

## References

- [1] T. Kito, N. Moriya, and J. Yamanoi *Japanese Economic Review*, vol. 72, p. 145–166, 2021.
- [2] P. Liu, V. Guarrasi, and A. E. Sarıyüce *Preprint arXiv:2005.11817*, 2020.
- [3] D. Cartwright and F. Harary *Psychological review*, vol. 63, no. 5, pp. 277–293, 1956.

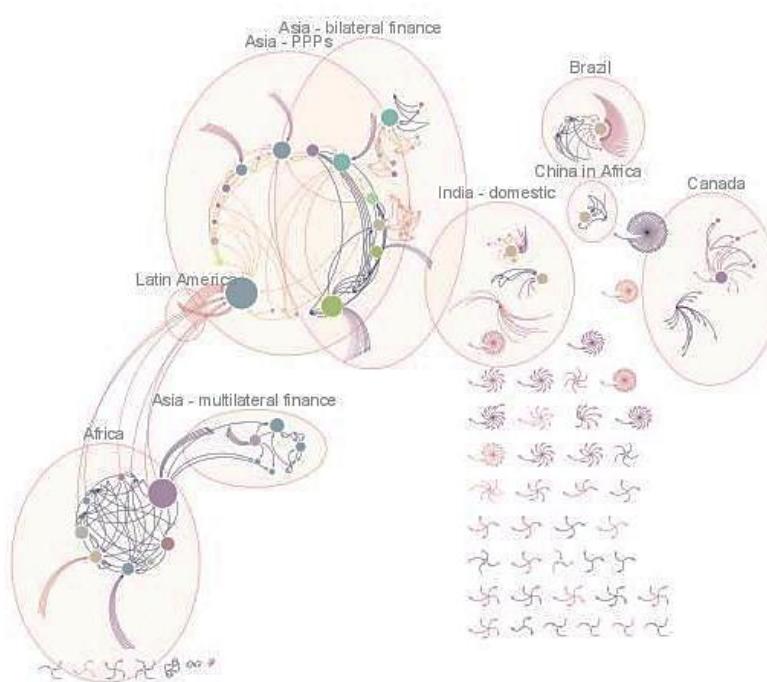
## Finding the right dam(n) partners for a just energy transition

Francesca Larosa, Jamie Rickman and Nadia Ameli

Supporting the transition towards a carbon neutral global economy requires the conversion of our energy systems to fully renewable sources. Over the past decades, solar and wind technologies have witnessed growth in investment and capacity and their role is expected to increase in the future. However, hydropower remains the primary renewable energy technology by capacity and generation accounting for approximately 60% of international investment flows in renewable energy. In particular, hydropower is at core of several developing countries' energy systems, where it provides the base for energy security.

Currently we have limited understanding of the investment architecture of hydropower, as the focus on increasing the financial flows for the transition has diverted attention away from who is

operating in this space and how. Our study responds to this crucial policy question by asking how a just transition can be accelerated around the world through improving access to finance for hydropower. We explore which investment models and co-investment patterns have driven hydropower deployment over the past century (1903-2020). First, we detect how different actors co-mobilised the capital needed for hydropower projects at global level building a bipartite network of project financing investments. Second, we zoom into the most critical investors who hold the project finance landscape together,



allowing capital to flow from developed to developing nations and we reveal the community structure that shapes the landscape for hydropower finance. Finally, we investigate investors' facilitator role in the investment chain and assess their ability to activate public-private partnerships (PPPs) over time.

Our results show that public and international institutions are the critical actors driving hydro finance through multilateral and bilateral financing mechanisms and global PPPs. These investors interact both within and across distant communities allowing finance to reach the most vulnerable areas. They integrate climate objectives into their development goals by mandate, enhancing economic prosperity and sustainable growth, while also boosting clean and affordable energy sources. This virtuous investment cycle is enabled by the diversity of partnerships activated by international institutions, who pave the road for banks and private actors across multiple countries. As one of the biggest challenges in sustainable energy transitions is likely to be in developing countries, our effort serves as a policy base to ensure equitable access and effective capital allocation for hydropower, the principal energy source in these regions. A strong investment architecture and channels will accelerate sustainable power deployment while leaving room for new technologies in the energy transition.

## Robustness of football passing networks against cascading failure

Daiki Miyagawa<sup>1\*</sup>, Koki Okamoto<sup>1</sup>, and Genki Ichinose<sup>1</sup>

<sup>1</sup> Department of Mathematical and Systems Engineering, Shizuoka University,  
3-5-1 Johoku, Naka-ku, Hamamatsu, 432-8561, Japan

\* miyagawa.daiki@shizuoka.ac.jp

Existing real-world systems are often represented by networks. On a network, one unit of the relevant quantity is transmitted between every pair of nodes at each time step. Examples are the flow of packets over the Internet or the flow of electricity in power grids. When those quantities are transmitted between a pair of nodes through a link, loads occur depending on the amount of quantities. If loads exceed nodes' capacities, the nodes collapse. Then, the loads of those failure nodes are redistributed to other nodes. This collapse of nodes and the redistribution of loads will be repeated until the capacities of every node become larger than the loads. These succeeding collapses of nodes are called a *cascading failure*. In this study, we consider cascading failures in football passing networks. Passing in football can be regarded as a network because the ball is passed between players (nodes). As the amount of passes increases, fatigue of players increases and their performances suffer. When this situation happens, other players try to make up for the loads. However, when the performances of some players decrease one after another because of fatigue, passing the ball temporarily does not work. We regard this situation as a cascading failure in a football passing network. Constructing a robust passing network may be related to the victory in a game. To this end, we explore the robustness of football passing networks against cascading failures.

Using the match data of J1 League (the top division league in Japan) in 2019, we constructed one passing network per team for a match. We simulated cascading failures on these networks. In the simulations, a single node with the highest loads is selected and removed (load attack). Here we show the results of the simulations conducted with Yokohama F.Marinos as typical examples. In the year we analyzed, the team aimed at high ball possession. The blue line in Fig. 1 indicates the case that the team made many passes among players in the game (vs Matsumoto Yamaga F.C.). This shows that the network keeps high robustness against cascading failures in the intermediate capacity rate, meaning that all players relatively equally conducted passes in the game. On the other hand, in the other case (vs Sagan Tosu), the passing networks are fragile against load attacks (green line in Fig. 1). In this game, passes were perhaps concentrated to some specific players. In such a case, if the opponent team marks those players, it may be difficult to pass the ball among team members. Our results indicate that even though Yokohama aims at high possession, the quality of their passing networks against cascading failures changes according to opponents' defensive tactics.

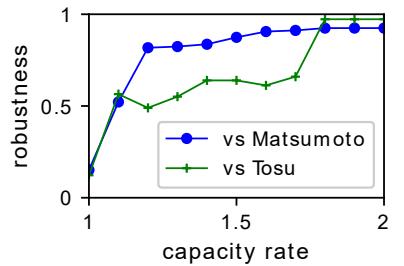


Figure 1: Network robustness of Yokohama F.Marinos against cascading failures.

## Incremental research - a favoured strategy in the “publish or perish” environment

Luka Blagojević<sup>1</sup> and Aleksandra Alorić<sup>2</sup>

<sup>1</sup> Department of Network and Data Science, Central European University, Vienna, Austria

<sup>2</sup> Center for the Study of Complex Systems, Institute of Physics Belgrade, Belgrade, Serbia

To venture into a new research direction and risk slow progress and productivity gaps, or incrementally refine your previous research agenda and have a steady research output is just an academics’ version of a long “to explore or exploit” dilemma researched in various disciplines from computer science to management. Motivated by empirical results [1] that find physicists across different subfields explore with caution and note stronger exploitation among younger researchers, this paper sets to investigate whether seniors’ exploration is only a consequence of survivor bias, or riskiness of exploration and productivity pressures on researchers’ drive them to exploitation strategies. With a stylised agent-based model we investigate agents with a fixed propensity for exploration and show that even with a mild risk (a paper on a new topic only takes twice as long as an incremental research paper) fraction of exploiters increases, and the effect is even stronger if there is productivity related rate of leaving academia (Fig. 1). Dominant exploitation strategies as the consequence have smaller knowledge diversity which is why we ought to find systemic incentives that favour exploration. This is somewhat easier now with the abundance of empirical data, but we shot survivor bi ie research with carefully tuned models.

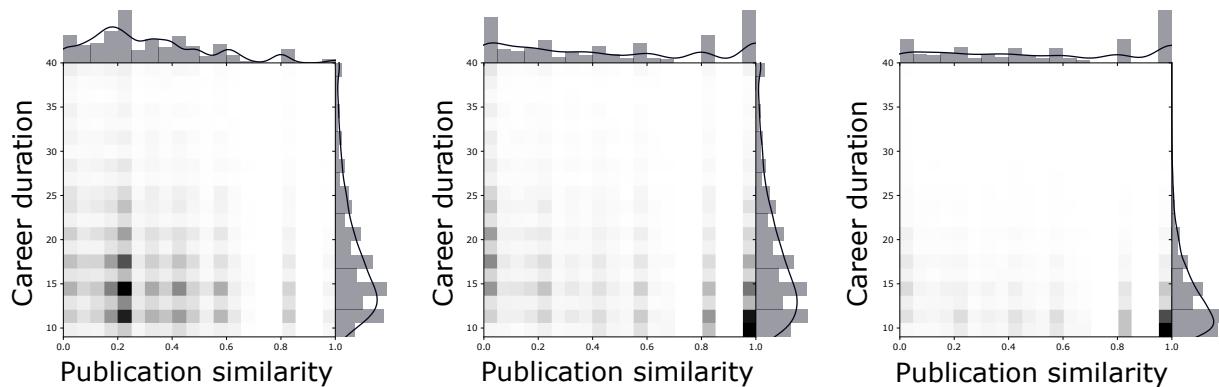


Figure 1: Distribution of agents career length and cosine similarity between first and last research topic in an ABM where exploitation leads to risk free publication, while exploitation leads to publication with 50% chance. Population of agents has exploitation rate  $q = 0.5$  (left) or uniformly drawn  $q \in [0, 1]$  (middle and right). Agents leave academia with fixed rate (left and middle), or with rate dependent on their productivity (right).

### References

- [1] Aleta, A., Meloni, S., Perra, N. and Moreno, Y. Explore with caution: mapping the evolution of scientific interest in physics. *EPJ Data Sci.* 8, 27 (2019).

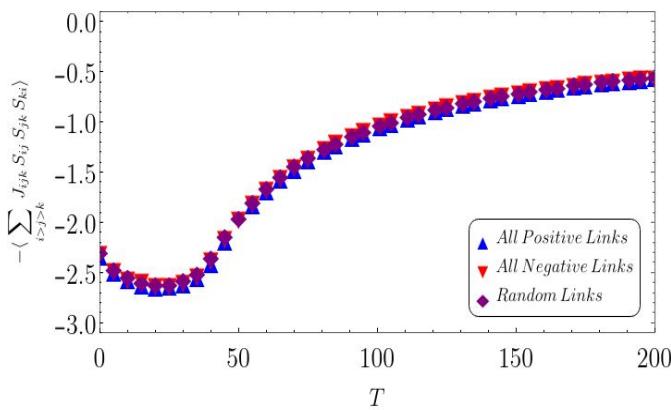
## Abnormal behavior of the Heider balance theory with heterogeneous triads

M. Bagherikalhor<sup>1</sup>, G. R. Jafari<sup>1,2</sup>

<sup>1</sup>Department of Physics, Shahid Beheshti University, G.C., Evin, Tehran 19839, Iran

<sup>2</sup>Institute for Cognitive and Brain Sciences, Shahid Beheshti University, G.C., Evin, Tehran, 19839, Iran

Balance theory was first introduced by Heider as a concept in social psychology[1]. It has a vast application and different comprehensive literature in different branches of science including mathematical sociology[2], ecology[3], and studies of international networks[4]. This theory deals with triplet interactions in which triads are equivalent, but in reality, we are a member of various groups of relationships which are different in priority. Now the question is, does the implementation of heterogeneity in triads confront us with astounding findings? To this purpose, we have studied disordered triadic interactions in the framework of Heider balance theory[5]. In Heider balance theory, the system experiences a first-order phase transition. Above a critical temperature, triads are randomly balanced and unbalanced and below it, all triads are balanced, and the energy of the system reaches its minimum[6]. Here, we are dealing with heterogeneity as a disorder where triads are assigned with different random weights that are quenched and coming from Gaussian probability distribution with mean,  $\mu$  and the variance,  $\sigma$ . By setting the mean value of distribution equals one ( $\mu = 1$ ), we study the network evolution in two regimes: i) weak disorder ( $\sigma = 0.1$ ), ii) strong disorder ( $\sigma = 10$ ). For  $\sigma = 0.1$ , the result is similar to the Heider balance. The minimum energy of the network occurs in the mean value of the probability distribution which is in agreement with our expectation since the network mostly consists of triads with weights around  $\mu = 1$ . The network reaches the global minimum in  $T < T_c$ ,  $T_c \approx 28$  and all triads are balanced. For  $\sigma = 10$ , our results indicate the energy versus temperature increases below a critical temperature due to the diversity of weights and the survival of unbalanced triads (frustrated). Although we expect to see a reduction in the number of unbalanced triads and the minimization of the energy, energy increases below a critical temperature (figure). This result can be more attractive compared to the Kondo effect [7]. The Kondo effect is a phenomenon in superconductivity and states that adding magnetic impurities in a metal causes increasing resistivity while we expect zero resistivity below a critical temperature. Here based on our modeling of a social network with heterogeneous triads, we see there is a schematic similarity in the behavior of these two phenomena.



- [1] F. Heider, *The Psychology of Interpersonal Relations* Wiley, New York, 1958.
- [2] Robert K. Leik, B. F. Meeker, *Mathematical sociology*, Prentice-Hall Englewood Cliffs, N.J, 1975
- [3] H. Saiz, J. Gómez-Gardeñes, P. Nuche, A. Girón, Y. Pueyo, and C. L. Alados, *Ecography* 40, 733, (2017).
- [4] J. Hart, *J. Peace Res.*, 11, 229, (1974).
- [5] M. Bagherikalhor, A. Kargaran, A.H. Shirazi, G.R. Jafari, *Phys. Rev.* 103, 032305 (2021)
- [6] F. Rabbani, Amir H. Shirazi, and G. R. Jafari, *Phys. Rev. E*, 99, 062302, (2019).
- [7] J. Kondo, *Progress of Theoretical Physics* 32, 37, (1964).

# Analyzing Eccentric Behavior of GAB Social Media Users

Sriniwas Pandey<sup>1,+</sup> and Hiroki Sayama<sup>1,2</sup>

<sup>1</sup>Binghamton University, State University of New York

<sup>2</sup>Waseda Innovation Lab, Waseda University, Tokyo, Japan

<sup>+</sup>spandey4@binghamton.edu

This work focuses on extremism in our day-to-day conversations and behavior, termed as ‘eccentricity’. In social media discussions, ‘normal’ is the center of conversations and distance from the center is degree of eccentricity. To analyze the dynamics of eccentricity, we collected posts and connection information for a sample of GAB users. Our sample has approximately three thousand users and a total of one hundred forty seven thousand gabs (posts) from August 2016 till January 2021. Using Doc2Vec and principal component analysis technique, each gab is converted into a numerical vector. For each user, a knowledge base is maintained that contains gab-vectors posted by the user and his neighbors in the past five days. The eccentricity indicates the degree of non-centrality of a post, relative to general discussion going on. We calculate f-score (the weighted (time decay term) average of absolute change in the eccentricity over time) and g-score (change in eccentricity with direction of change). Users are classified into four categories: i) High f-score and g-score: ‘Marginalized’ ii) High f-score and low g-score: ‘Fickle minded’ iii) Low f-score and g-score: ‘Stable’ iv) High f-score and a very low g-score: ‘Mainstreamed’. We also classify users based on degree of deviation from their own previous posts (self-eccentricity).

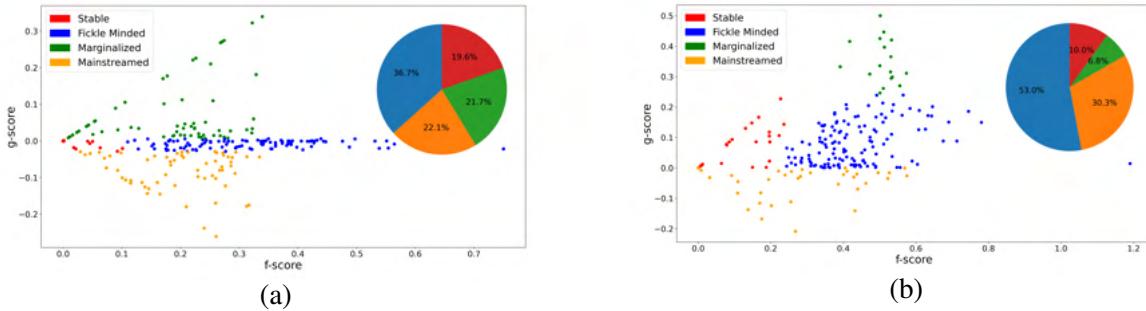


Figure 1: Eccentric behavior of GAB users a) Users represented in ‘f-score’-‘g-score’ space. Pie chart shows proportion of each user class. b) Users in f-score and g-score space of self-eccentricity.

Figure 1a shows the distribution of user posts in f-score and g-score space for eccentricity. 21.7% users get marginalized in their neighborhood, 36.7% users’ behavior keeps fluctuating and 19.6% users are stable. However, if we consider self eccentricity of a user, figure 1b, around 53% users are fickle minded while only 6.8% users are marginalized. To analyze this further, we examined relation between user eccentricity and self eccentricity classes. An interesting observation is that a large proportion of fickle minded users in terms of self eccentricity are either becoming marginalized or mainstream in their neighborhood. This also provides an interesting insight about the eccentricity of overall environment. Many users, who become homogeneous with their own previous ideas, get categorized as stable users in their neighborhood, this implies neighborhood’s eccentricity is decreasing at a faster rate.

## Bio-Inspired Water Distribution Network Design

Noha Abdel-Mottaleb<sup>1</sup>, Kashin Sugishita<sup>2</sup>, Naoki Masuda<sup>2,3</sup>, Qiong Zhang<sup>1</sup>

1. Department of Civil and Environmental Engineering, University of South Florida, Tampa, FL, USA

2. Department of Mathematics, State University of New York at Buffalo, USA

3. Computational and Data-Enabled Science and Engineering Program, State University of New York at Buffalo, USA

Water distribution networks are lifeline infrastructures of urban centers. Increased urbanism has fueled the growth and morphogenesis of cities around the world. Yet it is not clear how to design or reconfigure a water distribution network to accommodate the urban growth, while optimizing for performance (i.e., how adequately a network provides water to users). Given that biological flow networks have been evolving for much longer than water distribution networks, this study evaluates if the application of biological design principles to water distribution networks (i.e., bio-inspired) enhances their performance.

A growth model for *Physarum* mold is applied to a benchmark water distribution network that is based on a real system (Modena, Italy), where the water sources (i.e., root nodes) are predefined. The *Physarum* growth model generates a new network, where the pipe diameters are designed based on the principles of *Physarum* growth. The distribution of the diameters from the generated network is compared with the pipe diameters of the original network. Then, the WaterGEMS criticality tool is used to quantify the system demand shortfall under each individual pipe isolation. For each network, the system demand shortfall is summed over all of the pipe isolations, and the values are compared between the generated network and the original network.

The results showed that the bio-inspired network had a more uniform distribution of pipe diameters than the original network (see Figure below). This is not surprising because previous research has shown that water distribution networks with more uniform diameter distributions have higher redundancy of flow paths which enhances performance of the network under failure scenarios. Additionally, the aggregate system demand shortfall was more than 50 percent lower for the bio-inspired network. Future work will apply this methodology to more water distribution networks and additional performance indicators will be used to more definitively determine the usefulness of applying the biological model to design water distribution networks.

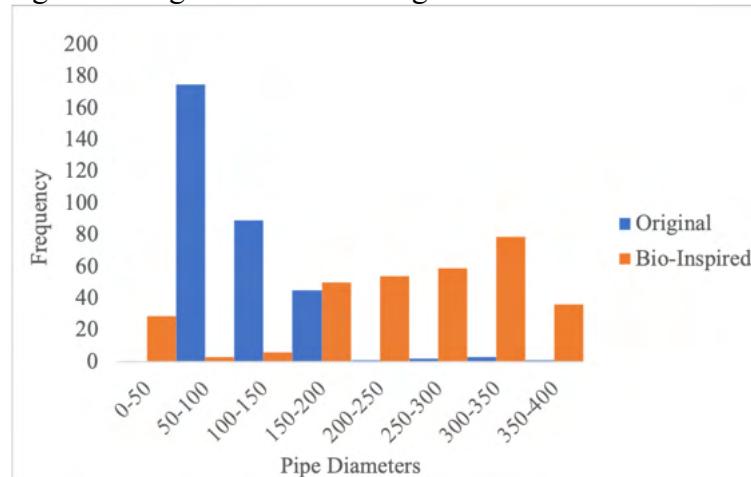


Figure: Histogram of the pipe diameters for the original and bio-inspired networks, where the aggregate system demand shortfall is 466 for the original network and 155 for the bio-inspired network

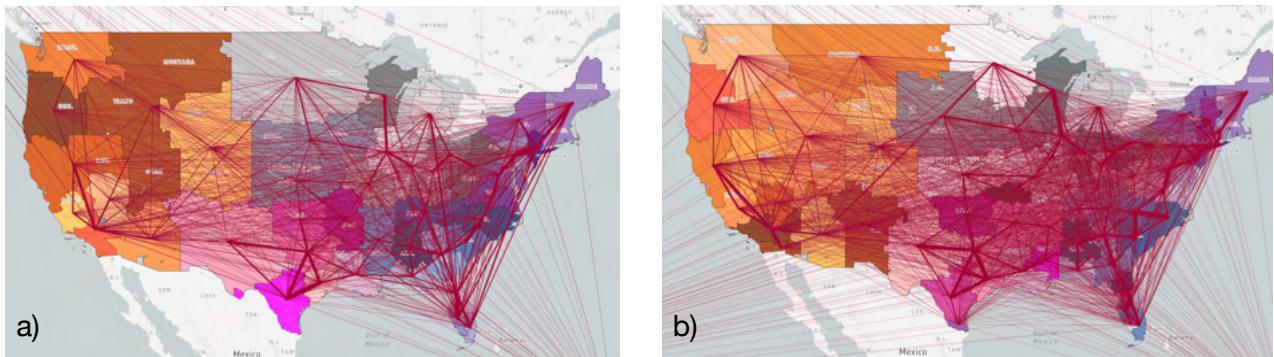
# A Meta-Population Movement-Based Epidemic Model for COVID-19 Pandemic

Tayeb Jamali<sup>1</sup>, Olha Buchel<sup>2</sup>, Yaneer Bar-Yam<sup>2</sup>, Leila Hedayatifar<sup>2</sup>

<sup>1</sup>Department of Physics, Beheshti University, Tehran, Iran

<sup>2</sup>New England Complex Systems Institute, 277 Broadway St, Cambridge, MA, USA

We developed a meta-population movement-based epidemic model to investigate the role of lockdown strategies on slowing down the spread of the COVID-19 outbreak. Previously, we demonstrated analytically that mobility preferences of the US population lead to the formation of self-organized geographical fragmentation patterns, that reveal how areas are connected to each other. We showed that lockdown policies in the US have substantially reduced the long-distance movements, changed mobility patterns and reduced in size the mobility patches compared with the ones before the COVID-19 pandemic (see Figure 1 below). Here, we introduce a generalized Susceptible-Exposed-Infected-Recovered model on geographical networks of mobility patches before/during/after lockdown strategies in the US. Our model has an extra term to account for the mobility flows among the partitions. We study the effect of mobility patches' sizes and connection strengths among them on the disease dynamics. By comparing the results of the model over time, we quantify the effectiveness of preventive mobility policies on slowing down the geographical spread of the disease.

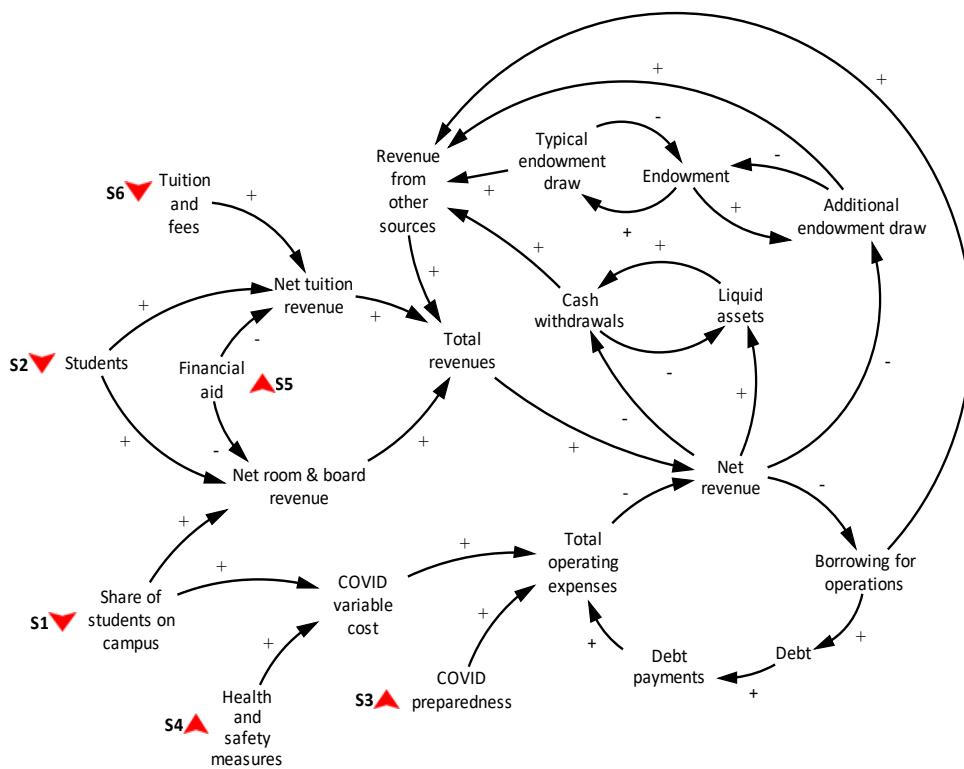


**Figure 1.** Self-organized geographical fragmentation patterns and mobility flows in the US before and during the pandemics: a.) mobility patterns and flows during February 23-29, 2020; b.) mobility patterns and flows during April 5-11, 2020, during the national lockdown. The colors show different communities; similar color hues identify similar groups of closely interconnected communities; line - connections between communities, and lines widths - strengths of the connections.

## COVID-19 and Financial Sustainability of Academic Institutions

Oleg Pavlov, Worcester Polytechnic Institute, opavlov@wpi.edu  
 Evangelos Katsamakas, Fordham University, katsamakas@fordham.edu

The COVID-19 pandemic has had a significant impact on higher education. Steering academic institutions through the pandemic is a complex and multifaceted task that can be supported with model-based scenario analysis. This article studies the short-term and long-term effects of the pandemic on the financial health of a college using *scenario analyses* and *stress testing* with a *system dynamics model* of a representative tuition-dependent college. The figure below shows six effects of COVID-19 (we refer to them as stressors S1 through S6) that are studied in this article and their impact on a college. An upward pointing red arrow indicates that this particular variable increases due to COVID-19 and a downward pointing red arrow means that the value of the variable is likely to decrease due to COVID-19. A positive link between two variables signifies a positive causality between them. Negative causalities are shown with negative links. We find that different combinations of the pandemic mitigation protocols may have varying effects on the financial sustainability of an academic institution. By simulating six individual components of the COVID-19 shock, we learn that due to the causal complexity, nonlinear responses and delays in the system, the negative shocks can propagate widely through the college, sometimes with considerable delays and disproportionate effects. Simulations show that some pandemic mitigation protocol choices may destabilize even financially healthy institutions. The article adds to the literature on the economics of higher education, management of the *pandemic-related risk* on campus, *college stress-testing* and *model-informed decision making* in higher education.



## Intermittent frequency chimeras in modular network of FitzHugh-Nagumo oscillators

Sangeeta Rani Ujjwal<sup>1</sup>, Ulrike Feudel<sup>2</sup> and Ramakrishna Ramaswamy<sup>3</sup>

<sup>1</sup> Department of Physics, Institute of Science  
 Banaras Hindu University, Varanasi, Uttar Pradesh 221005, India  
 sujjwal@bhu.ac.in

<sup>2</sup> Institute for Chemistry and Biology of the Marine Environment  
 University of Oldenburg, 26111 Oldenburg, Germany  
 ulrike.feudel@uni-oldenburg.de

<sup>3</sup>Department of Chemistry, Indian Institute of Technology- Delhi  
 Hauz Khas, New Delhi 110016, India  
 ramakrishna.ramaswamy@gmail.com

We study the dynamics of FitzHugh-Nagumo oscillators in a modular network where the interactions between the modules are nonlocal and include a phase lag. Depending upon the value of phase lag, in addition to global synchrony and modular synchronization we observe a novel type of chimera state which we call as ‘intermittent frequency chimera’ where coherent domains consist of modules separately synchronized to different frequencies, coexist with modules within which the dynamics is desynchronized, and some of these modules show intermittency wherein they oscillate between synchronized and desynchronized state with time. We also observe multistability between these different dynamical states in a sizable region of parameter space.

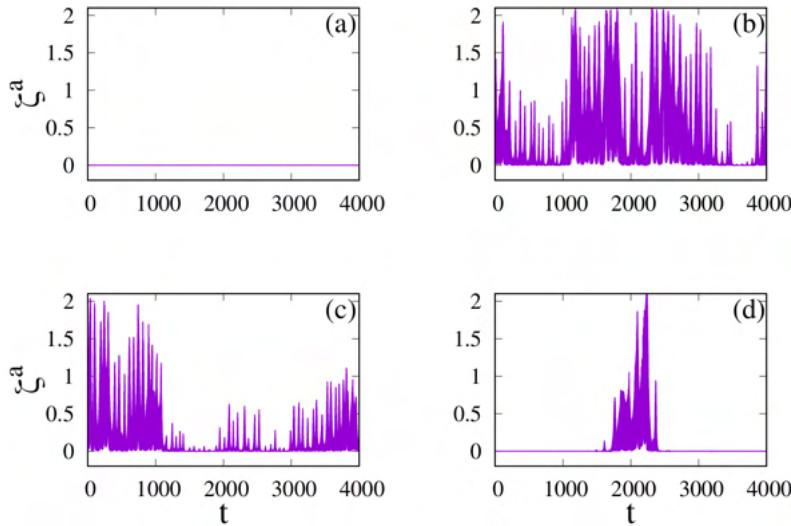


Figure 1: Average synchronization error ( $\zeta^a$ ) for four different modules is plotted with time showing the intermittent frequency chimera where oscillators in module-1 are synchronized (a), in module-2 and 3 (b)-(c) are desynchronized while in module-4 (d) go in and out of synchrony with time.

## **Part XI**

# **Poster Session 2**

## Network Symmetry Extraction

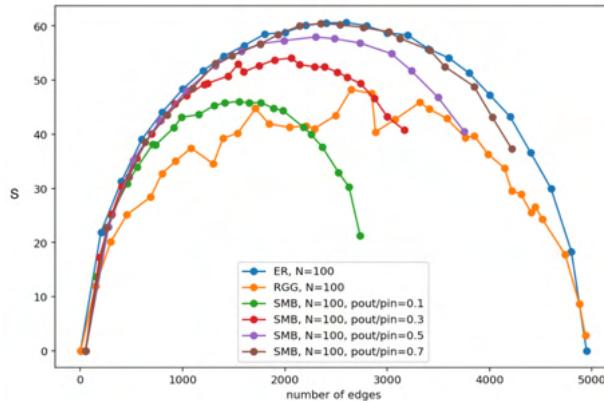
**Y. Liu<sup>1</sup>.** (1) Network Science Institute, Northeastern University, Boston MA, USA, [lycrdfzpk@gmail.com](mailto:lycrdfzpk@gmail.com).

Most literature on measuring the symmetry of networks uses graph automorphism algorithms to find the automorphism group, and defines either the number of elements in the automorphism group or the number of nodes involved in the automorphism group as a measure of the level of symmetry of networks. However, this type of measure has two problems: first, it is a measure for exact symmetry, meaning that it fails to capture a large group of approximate symmetries of networks (for example, a lattice with one missing edge still has a high level of symmetry, intuitively speaking, but according to the present symmetry measure it may have zero symmetry); second, in real-world networks, most of the automorphism transformations found are composed of local permutations, for example two degree-one nodes attached to the same node are symmetric to each other. Therefore if a network has many automorphism transformations it could mean that the network has a lot of local symmetries, instead of global symmetries. For example, a lattice has translational and reflectional symmetries, both of which are global. To overcome these problems, we introduce a new measure of the level of symmetry of a given network with adjacency matrix A:

$$S = \min_Q (\|A - QAQ^T\|),$$

where Q is a permutation operator with properties  $\det(Q) = 1$  and  $QQ^T = I$ , and  $\|\cdot\|$  is order-1 norm.  $QAQ^T$  is the resulting adjacency matrix after a permutation of the node sequence. In order to restrict the symmetries we find to global symmetries, we limit Q such that  $\text{trace}(Q) = 0$ , meaning that all nodes are involved in the permutations. The difference between A and the permuted adjacency matrix  $\widehat{Q}A\widehat{Q}^T$  with permutation  $\widehat{Q} = \arg\min(\|A - QAQ^T\|)$  is a measure of how close  $\widehat{Q}$  is to exact symmetry.

Figure 1 shows S for networks with  $N = 100$  nodes, generated from three different network models: Erdős Rényi (ER), Random Geometric Graphs (RGG), and Stochastic Block Models (SBM), as a function of the number of edges M. For ER networks (blue) and RGGs (orange), when they are empty ( $M = 0$ ) or complete ( $M = 5,000$ ), they have complete symmetry, which is reflected by  $S = 0$ . When half of the potential edge slots are filled ( $M = 2,500$ ), the networks have the least level of symmetry, due to the high uncertainty of edge placements.



**Figure 1:** S for networks with  $N = 100$  nodes, generated from three different network models.

# Masculinity Contest Cultures and Inclusive Cultures: Insights From a Network Model of Organizational Socialization and Promotion

John Meluso<sup>1</sup>, Laurent Hébert-Dufresne<sup>1,2</sup>, James Bagrow<sup>1,3</sup>, Rob Razzante<sup>4</sup>

<sup>1</sup>Vermont Complex Systems Center, University of Vermont

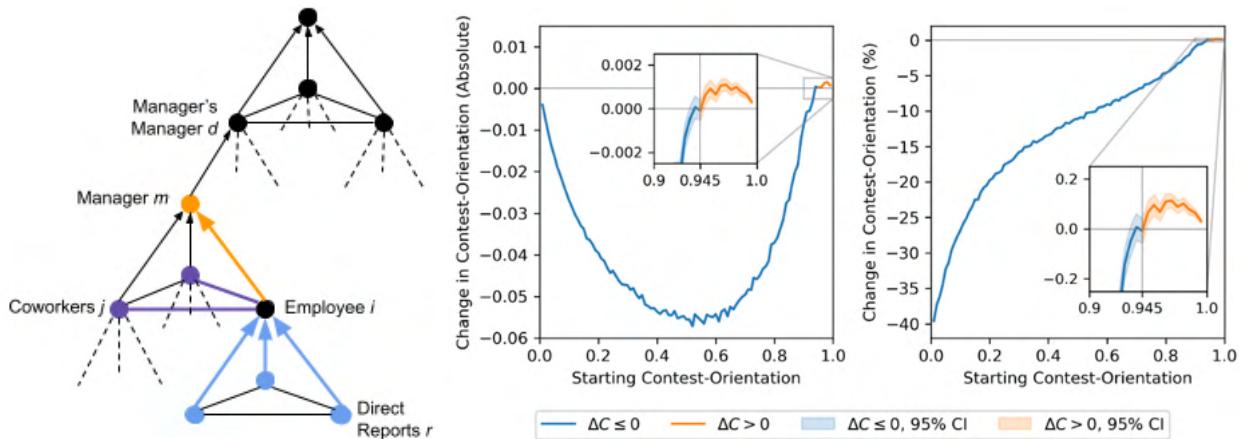
<sup>2</sup>Department of Computer Science, University of Vermont

<sup>3</sup>Department of Mathematics & Statistics, University of Vermont

<sup>4</sup>Department of Communication Studies, The College of Wooster

Recent organizational scholarship finds that “work” becomes a masculinity contest when organizations valorize acts of dominance, acts which likely perpetuate barriers to the advancement of marginalized groups in organizations (as with the “glass ceiling”). But how such Masculinity Contest Cultures (MCCs) develop and sustain themselves remain open questions. This study juxtaposes MCCs with Inclusive Cultures (ICs) to examine how each culture spreads in organizations’ social networks. Drawing upon systems theory, we simulated the processes of socialization and promotion in organization networks via an agent-based model. Varying the hiring pools for different organizations from inclusive to contest-oriented revealed that inclusiveness emerged as self-reinforcing in all but organizations with the most contest-oriented hiring pools. In contrast, hiring pools socialized into hegemonic masculinity made organizations more likely to resist ICs and showed potential to evolve into MCCs in productivity-oriented, hierarchical organizations. Furthermore, organizations tended to see greater cultural change in higher ranks of the organizational hierarchy than in lower ranks, regardless of the more-prevalent culture. Such cultural stratification demonstrates the challenge in transforming the culture of every organizational level toward inclusiveness without further study of socialization processes.

**Keywords:** Agent-Based Modeling, Social Networks, Organizational Culture, Hierarchy, Inclusiveness, Socialization, Masculinity Contest Culture, Hegemonic Masculinity



*Figure: (Left) Organization structure and relationships between employees. (Center/Right) Average change in contest-orientation (vertical axis) over 100 runs for different starting points for contest-orientation (horizontal axis) in absolute and percentage terms, respectively. Insets show the transition to contest culture reinforcement after an additional 500 runs for each starting contest-orientation  $\geq 0.85$ .*

**Emergent hierarchy through conductance-based node constraints**

Christopher Diggans, Jeremie Fish and Erik Bollt

See the full paper published in the Northeast Journal of Complex Systems:  
<https://orb.binghamton.edu/nejcs/vol3/iss1/4/>

# MULTICRITERIA SELECTION AND EVALUATION OF THE LIFESAVING UAVS FOR A DETERMINED AREA (AEGEAN SEA COASTLINE) IN TURKEY BY COMPARING TODIM&GOAL PROGRAMMING (GP) METHOD

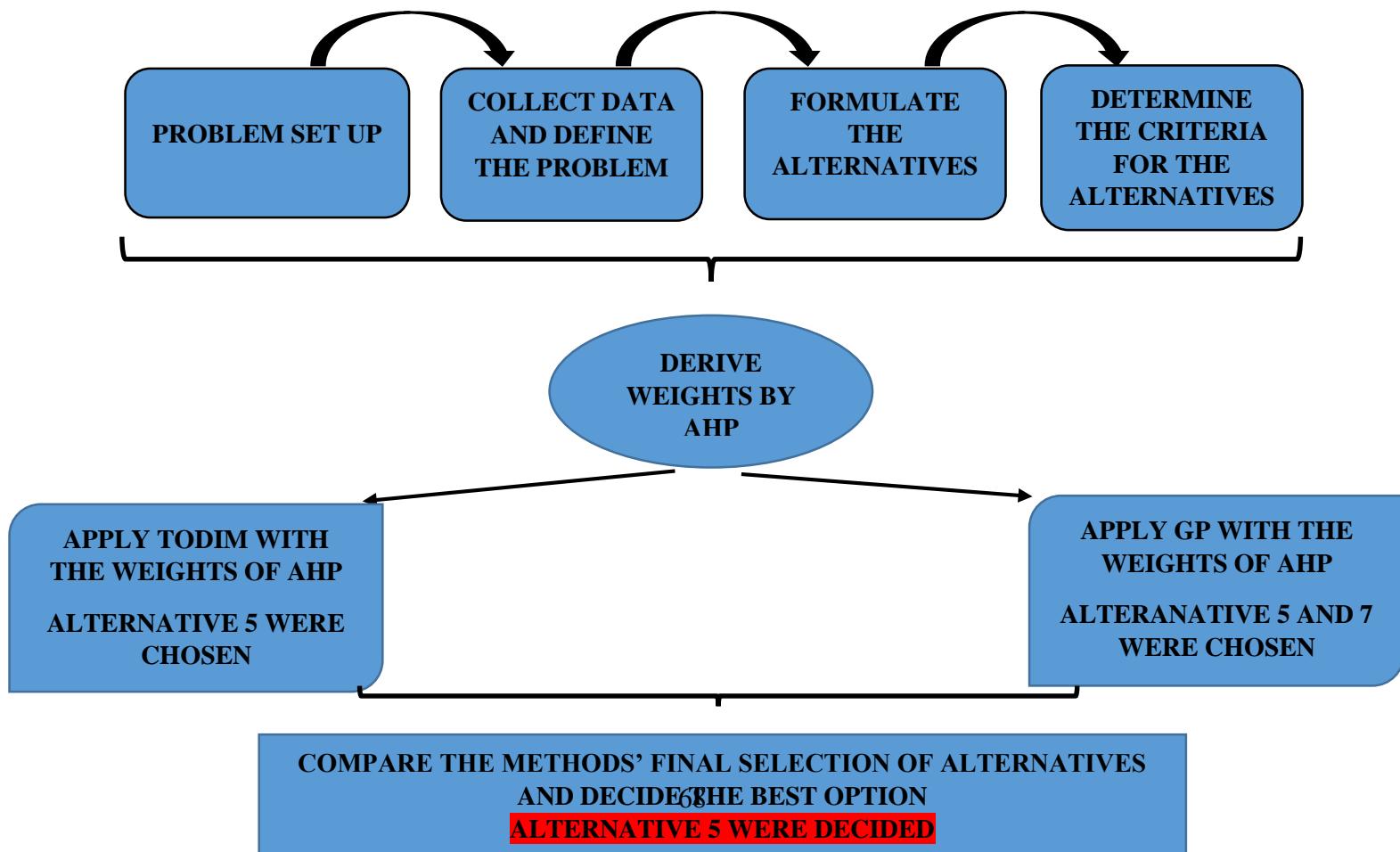
Babek ERDEBİLLİ, İbrahim YILMAZ, Volkan SARIBAŞ

## Abstract.

Unmanned Aerial Vehicles (UAVs) are taking part in our life day by day. UAVs are preferred almost everywhere from the military, defense to surveillance, agriculture. Thus, the usage area almost is expanding every day; however, there is a crucial topic to use UAVs in lifesaving for the victims drowning. There are kinds of UAVs for lifesaving. It is aimed in this study to select the best UAV. Seven UAV alternatives related to lifesaving for the victims drowning and ten criteria proposed by the experts to evaluate the alternatives and select the best alternative. For this purpose, AHP-TODIM (Analytic Hierarchy Process - Tomada de Decisão Iterativa Multicritério) and AHP-GP (Analytic Hierarchy Process – Goal Programming) models are applied to solve this problem with the same values. The calculations indicate that Alternative UAV-5 was chosen as the best of them by TODIM and Alternatives UAV-5 and UAV-7 were selected by Mixed Integer Goal Programming. Finally, the results are compared and the solution sensitivity analysis is addressed.

Keywords: Unmanned Aerial Vehicles (UAVs), Rescue operations, drowning, floatation devices, Multiple criteria analysis, method, Goal Programming, TODIM, weight consistency, decision making.

## Process Flow



# Modeling epidemic spreading in Markovian temporal networks with different degrees of concurrency

RUODAN LIU<sup>1</sup>, MASAKI OGURA<sup>2</sup>, ELOHIM FONSECA DOS REIS<sup>1</sup>, AND NAOKI MASUDA<sup>1,3</sup>

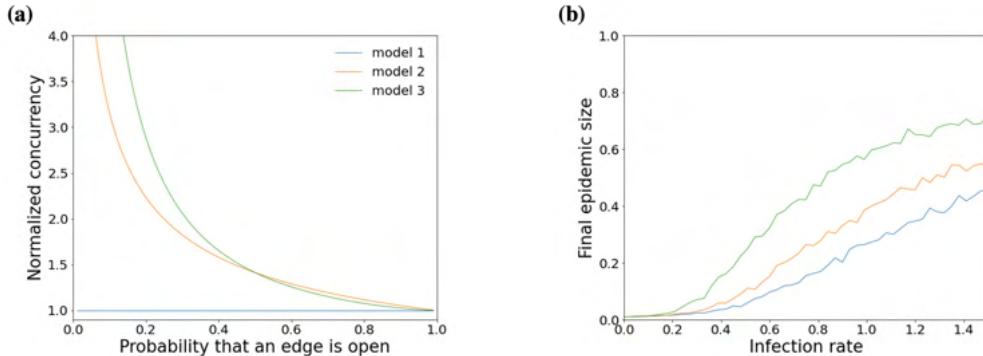
<sup>1</sup>*Department of Mathematics, University at Buffalo, SUNY*

<sup>2</sup>*Graduate School of Information Science and Technology, Osaka University*

<sup>3</sup>*Computational and Data-Enabled Sciences and Engineering Program,  
University at Buffalo, SUNY*

The concurrency of edges, quantified by the number of edges that share a node at a given time, is considered to influence epidemic processes, which has been a topic particularly debated for HIV/AIDS infection in sub-Saharan Africa. More broadly, concurrency is likely to be a determinant of epidemic processes in temporal networks [1]. We study effects of concurrency on epidemic spreading in the stochastic susceptible-infected-recovered (SIR) dynamics on three temporal network models, with which we can tune the amount of concurrency while keeping the probability that each edge is present and the structure of the aggregate (i.e., static) network the same across comparisons. In model 1, the appearance and disappearance of edges are governed by independent continuous-time Markov processes [2], which lacks surplus concurrency. In our model 2, each node independently obeys a two-state (i.e., active/inactive states) continuous-time Markov process, and the edge between two nodes appears if and only if both nodes are active. We also propose model 3, in which the edge between two nodes appears if and only if either node is active.

Using a concurrency measure, we analytically calculated the degree of concurrency of edges sharing a node for the three models under the condition that each edge appears with the same stationary probability in the different models. We have found that models 2 and 3 are more concurrent than model 1. Furthermore, model 2 is more concurrent than model 3 or vice versa depending on whether the stationary probability that an edge exists is larger than 0.5 (Figure 1(a)). We then numerically simulated the stochastic SIR dynamics on the three temporal network models on top of an underlying static network generated by the Barabási-Albert model with  $N = 100$  nodes and average degree  $\langle k \rangle = 7.68$  (and other networks; results not shown). We started each simulation from an arbitrarily chosen single infected node, while the other  $N - 1$  nodes were initially susceptible. We set the parameters of the three models such that the degree of concurrency for models 1, 2, and 3 are  $\frac{9}{25} = 0.36$ ,  $\frac{3}{5} = 0.6$ , and  $\frac{29}{45} \approx 0.64$ , respectively. Figure 1(b) suggests that the final epidemic size is large when the degree of concurrency is large across the infection rate values. Our result affirmatively confirms the hypothesis that concurrency enhances epidemic spreading on networks, and in a model of temporal networks for which one can strictly control the probability of the edge being active and carry out some theoretical analysis.



**Figure 1:** (a) Normalized concurrency index as a function of the stationary probability that an edge is active. (b) The relationship between the infection rate and the final epidemic size. We used the Barabási-Albert network generated with  $N = 100$  and  $\langle k \rangle = 7.68$ . The final size shown is the average over 1000 simulations.

## References

- [1] N. Masuda, J. C. Miller, P. Holme (2020), arXiv: 2012.13317.
- [2] X. Zhang, C. Moore, and M. E. J. Newman (2017), Eur. Phys. J. B, 90: 200.

# Complexity and Curriculum:

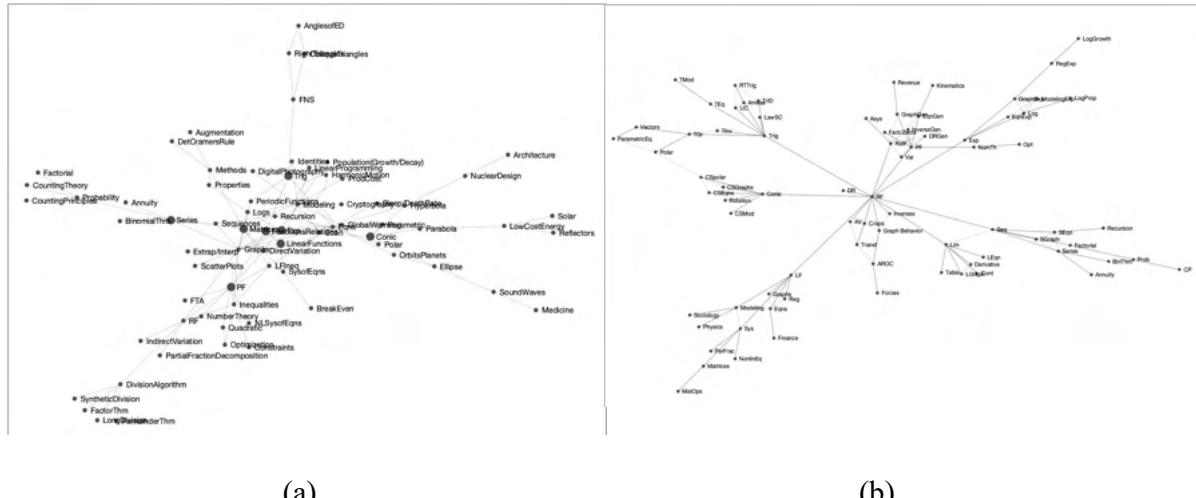
## Examining the Role of Connectivity in a Precalculus Course

John O'Meara, Ashwin Vaidya, Stephen Essien

Department of Mathematical Sciences, Montclair State University, Montclair, NJ 07042

This project aims to utilize the mathematical theory of networks and statistical analysis to understand the significance of connectivity in mathematics education. *We collectively hypothesize that education is a complex system, at various scales, composed of different ideas and concepts which are connected. Therefore, the task of a good teacher lies in elucidating these connections and helping students make their own connections. Such a pedagogy allows students and educators alike to personalize learning and strive to be ‘creative’, which essentially amounts to ‘meaning making’ or assigning new meaning from the foundation of old ideas.*

In this project, we test our hypothesis by examining the network structure of a precalculus course. We identify the curricular network of different, well known precalculus textbooks by identifying various topics covered in this course and their connectivity, i.e. relationships between different topics. The topics covered in each respective text are nodes, and explicit relationships between topics are edges. Connections are made using a uniform procedural rubric which essentially allows us to link concepts if they are related sufficiently strongly. Figure 1 below exemplifies such a graph based on a standard textbook on precalculus.



Through examination of network indices such as connectivity index, average path length and degree distribution of these networks, we may gather a measure of qualitative efficacy through a highly quantitative investigation to objectively analyze the effectiveness of a given textbook and the pedagogy in a course that emphasizes a similar approach. We also investigate statistical correlations between the computed network indices and ratings of these texts which give us some indication of their effectiveness. Effective textbooks, as recognized by publicly available user ratings, are indicated not only by expected magnitude of relevant metrics, but also by the consideration of the synthesis of global and local connectivity, within and between topics. Fitting degree distributions to a probability model allows insight into concentration of the presentation of curriculum, and the family of distributions to which precalculus curricula belong.

# Exploring the connectivity of a hospital: A network analysis of patient movement community structures

EE. Boncea<sup>1</sup>, P. Expert<sup>1</sup>, CE. Costelloe<sup>1</sup>

<sup>1</sup> Global Digital Health Unit, Imperial College London, UK.

A hospital patient's movement trajectory, which encompasses intrahospital transfers through various locations, involves many interactions between heterogenous, autonomous yet interdependent healthcare staff. Diffusion of control is characteristic of a complex system, meaning that the coordination of patient transfers can be a 'problem of many hands', with no single actor held responsible [1]. Without an overarching view of the connected whole, system-level improvements are difficult to achieve. We present a network based exploration of patient movement structures within a London hospital. Ward-transfer data from electronic health records pertaining to patients from 2015-2018 were used to create a weighted patient movement network. The communities found using Newman-Girvan modularity were clearly interpretable, comprising of maternity, paediatric, surgical and medical specialities (Fig 1 left). As patient specialities are a determinant of movement [2], the primary speciality of patients on wards informed an alternative null model adapted from [3], based on similarity of ward specialities, aiming to identify speciality independent communities. The speciality null model yielded more granular communities correlating with patient illness severity (Fig 1 right): e.g., low severity general medical wards (yellow), medium severity wards equipped to treat deteriorating patients (orange) and high severity wards, including the adult intensive care unit and a major trauma ward (dark blue). These results demonstrate that meaningful ward communities can be found within a hospital and give insights into the system's complex topology, highlighting patient severity and speciality as mechanisms shaping movement patterns. Examining their interconnectedness and understanding their emergence is the first step in building a data-driven, shared perception of ward functionality and linking these to micro-level inefficiencies, as well as new interventions to optimise patient movement.

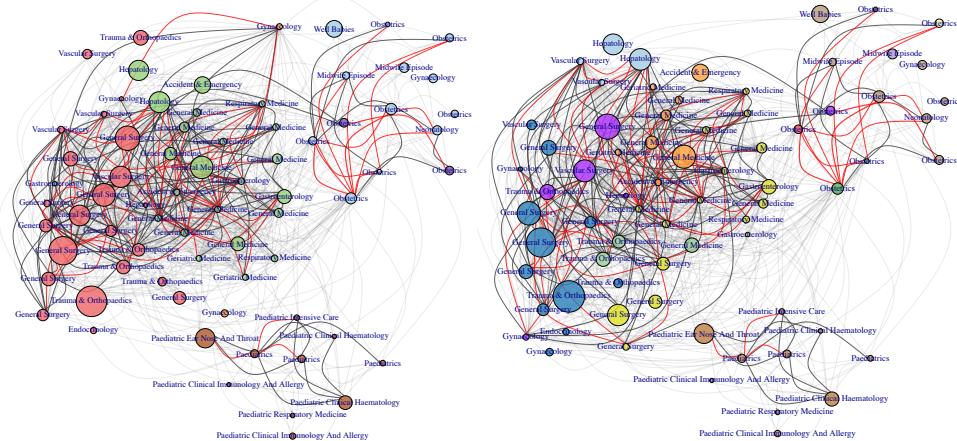


Figure 1: Communities found in the patient movement network with the Newman-Girvan modularity (left) and the speciality modularity (right). Node labels pertain to specialities.

## References:

- [1] M.Dixon-Woods, PJ Pronovost., "Patient safety and the problem of many hands". *BMJ Qual. Saf.* 2016;25(7): 485–488.
- [2] H. Xu, W. Wu, S. Nemati *et al.*, "Patient Flow Prediction via Discriminative Learning of Mutually-Correcting Processes. IEEE TKDE. 2017; 29(1):157-17
- [3] P. Expert, *et al.*, PNAS. 2011;108(19):7663-8

# On the relation between transversal and longitudinal scaling in cities

Fabiano L. Ribeiro<sup>1</sup>

<sup>1</sup> Department of Physics (DFI)  
Universidade Federal de Lavra (UFLA), Lavras, MG, Brazil  
fribeiro@ufla.br

## Abstract

Empirical evidence has been shown that some urban variables scale non-linearly with the city population size. More specifically, some socio-economic variables, such as the number of patents, wages and GDP, show a super-linear behaviour with the city's population. On the other hand, infrastructure variables, such as the number of gas stations and streets' length, scale sub-linearly with the city population, generating a scale economy. However, do these scaling properties observed in a system of cities (transversal scaling) also work for individual cities in their growth process (longitudinal scaling)? The answer to this question has important policy implications, but the lack of relevant data has hindered rigorous empirical tests. The work that will be presented was developed looking at the evolution of two urban variables, GDP (see Fig. (1)) and water network length, for over 5500 cities in Brazil. It will be shown that longitudinal scaling exponents are city-specific; however, they are distributed around an average value that approaches the transversal scaling exponent provided that the data is decomposed to eliminate external factors, and only for cities with a sufficiently high growth rate. This result adds complexity to the idea that longitudinal dynamics is a micro-scaling version of the transversal dynamics of the entire urban system. For more details about this work, see the references [1, 2, 3].

## keywords:

Cities; urban scaling; transversal scaling; longitudinal scaling.

## Acknowledgments

I would like to acknowledge the financial support from *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (process number: 405921/2016-0), and *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (process number: 88881.119533/2016-01).

## References

- [1] Ribeiro, F. L., Meirelles, J., Netto, V. M., Neto, C. R. & Baronchelli, A. On the relation between Transversal and Longitudinal Scaling in Cities. PLoS One 12(1), e0233003 (2020). doi:10.1371/journal.pone.0233003

# Perfect synchronization in networks of Sakaguchi-Kuramoto oscillators

Prosenjit Kundu<sup>1</sup>, Chittaranjan Hens<sup>2</sup>, Baruch Barzel<sup>3</sup>, and Pinaki Pal<sup>4</sup>

<sup>1</sup> Department of Mathematics, State University of New York at Buffalo, Buffalo, New York, USA

<sup>2</sup> Physics and Applied Mathematics Unit, Indian Statistical Institute, Kolkata, India

<sup>3</sup> Department of Mathematics, Bar-Ilan University, Ramat-Gan, Israel

<sup>4</sup> Department of Mathematics, National Institute of Technology, Durgapur, India

From the power grid to neuronal networks, synchronization phenomena have found diverse and pertinent applications capturing the emergence of collective behavior in complex systems. However, the common scenario where the coupling between the system's components induces phase-lags avoid global synchronization even under extremely strong coupling or with a homogeneous frequency.

Here we show that such systems

$$\frac{d\theta_i}{dt} = \omega_i + K \sum_j A_{ij} \sin(\theta_j - \theta_i - \alpha_{ij})$$

can still exhibit perfect synchronization with the appropriate choice of the nodes' frequencies  $\omega$

$$\omega_i = - \sum_j A_{ij} \sin(-\alpha_{ij}) + \frac{1}{N} \sum_{i,j} A_{ij} \sin(-\alpha_{ij})$$

Where,  $A$  is the connectivity matrix.  $\theta, \alpha$  &  $K$  are phase, lag and coupling strength respectively.

We find that high levels of synchronization are sustained in the vicinity of the optimal set, allowing for some level of deviation in the frequencies without significant degradation of synchronization. Demonstrating our results on first and second-order phase-frustrated Kuramoto dynamics, we implement them on both synthetic and real power grid networks, showing how to achieve synchronization in a phase-frustrated environment.

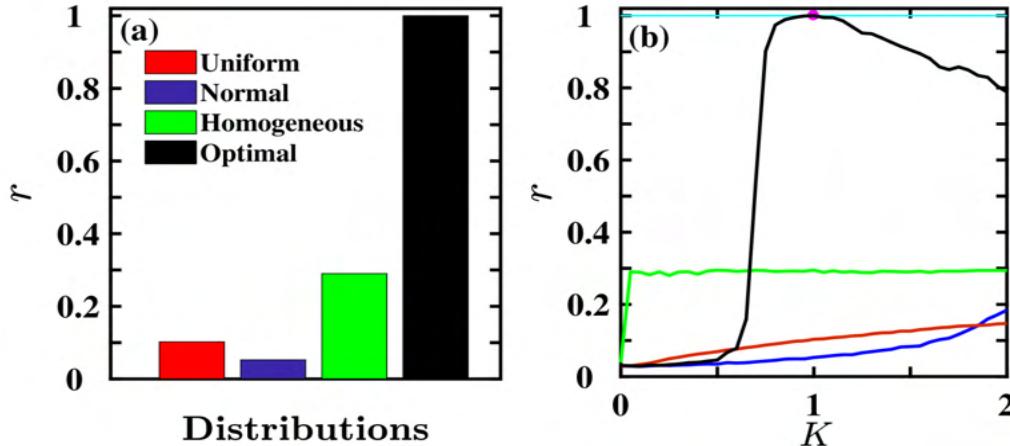


Figure 1: (a) Synchronization order parameter  $r$  of a phase-frustrated scale-free oscillator network of SK oscillators with different selections of  $\omega$  by numerically solving the SK model. (b) Increasing the coupling strengths  $K$  we find that the system persistently avoids synchronization even for large  $K$ , under the general frequency distributions. Synchronization is only obtained using our optimal  $\omega$ .

[1] Pikovsky A., Rosenblum M. and Kurths J., *Synchronization: A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, England) 2003.

[2] Kundu P., Hens C., Barzel B. and Pal P., *Europhys. Lett.* **120** (2017) 40001.

# **Optimal control problem of variable order fractional systems with Riemann–Liouville derivative**

Touria Karite<sup>1</sup>, Adil Khazari<sup>2</sup> and Delfim F. M. Torres<sup>3</sup>

<sup>1</sup> LISA Laboratory, Electrical Engineering and Informatics Department,  
National School of Applied Sciences, Sidi Mohamed Ben Abdellah University, Fez, Morocco  
touria.karite@usmba.ac.ma

<sup>2</sup> LAMA Laboratory, National School of Commerce and Management,  
Sidi Mohamed Ben Abdellah University, Fez, Morocco  
adil.khazari@usmba.ac.ma

<sup>3</sup> CIDMA center, Department of Mathematics, University of Aveiro,  
3810-193 Aveiro, Portugal  
delfim@ua.pt

## **Abstract**

System analysis consists of studying a set of concepts allowing a better knowledge of its properties. Among these notions, there are controllability, observability, stability, stabilization, detectability, spreadability, and so forth. The ability to formulate evolutionary governing equations has led to the successful application of variable order fractional operators to the modeling of complex real-world problems ranging from mechanics, to transport processes, to control theory, to biology. Variable-order fractional computation is a relatively less known branch of computation that offers remarkable opportunities to simulate interdisciplinary processes. They are a powerful mathematical tool that are acknowledged as an alternative and precise approach in effectively describing real-world phenomena. In this paper we are interested in studying the problem of optimal control of fractional system with variable order operator. The fractional time derivative is considered in Riemann–Liouville sense. We first study the problem by using the Lax-Milgram Theorem, the existence and the uniqueness of the solution of the variable order fractional differential system in a Hilbert space. Then we show that the considered optimal control problem has a unique solution.

## **References**

- [1] C.F. Lorenzo and T.T. Hartley, Variable order and distributed order fractional operators, *Nonlinear Dynam.* Vol. 29, No. 1 (2002) 57-98.
- [2] H.G. Sun, W. Chen, H. Wei and Y. Chen, A comparative study of constant-order and variable-order fractional models in characterizing memory property of systems. *Eur. Phys. J. Spec. Top.* Vol. 193, No. 1 (2011) 185-192.
- [3] R. Almeida, D. Tavares and D.F.M. Torres, *The Variable-Order Fractional Calculus of Variations*, Springer International Publishing, 2019

# Detecting Fraudulent Users in Online Marketplaces Using Temporal Motifs

Rupam Acharyya<sup>1</sup>, Penghang Liu<sup>2</sup>, Ahmet Erdem Sariyüce<sup>2</sup>, Naoki Masuda<sup>1,3</sup>

<sup>1</sup> Department of Mathematics, University at Buffalo, Buffalo, NY, USA

<sup>2</sup> Department of Computer Science and Engineering, University at Buffalo, Buffalo, NY, USA

<sup>3</sup> Computational and Data-Enabled Science and Engineering Program, University at Buffalo, Buffalo, NY, USA

Over the years online marketplaces have become a fast moving and popular part of the e-commerce industry. Unfortunately, such rapid growth often results in increases in fraudulent transactions. Detecting fraudulent transactions has since been a challenge in the field of AI when designing applications for the e-commerce industry. Researchers have borrowed techniques and algorithms from, for example, natural language processing or computer vision in determining fraudulent behavior. One prime drawback of these techniques is that fraudsters can easily outwit them with the help of encrypted communication. Another complementary approach towards detecting fraudulent users is to explore the local (i.e., egocentric) transaction network of individual users. (By convention, each transaction is a directed edge from seller to buyer.) In Ref. [1], the authors analyzed egocentric networks of users in an online marketplace, Mercari Inc. They showed that indices derived from static local network fed to standard machine learning algorithms were capable of distinguishing fraudulent users from normal users with a high accuracy. In the current work, we extend their work by analyzing the temporal transaction network of Mercari.

We start with a user which is manually labeled as normal or fraudulent (known as seed user) and create the egocentric network for them by adding its neighbors and the connections among the neighbors. Crucially, each transaction between users is time-stamped. Then, we examine the abundance of consecutive transactions in this network that occur within a small time span (i.e., 6 days for this work), which are known as temporal network motifs. We consider motifs with two nodes and two edges that contain a seed user. Note that these seed users may also be involved in other selling and buying transactions. First, we show that the abundance of static counterpart of such motifs (in particular, the one to the left in Fig. 1A) is considerably different between the normal and fraudulent users. Then, we observe that the number of each type of temporal motifs in the local network of fraudulent users is considerably higher than that of normal users except for the most frequent type of temporal motif (see Fig. 1B). Finally, we use these counts of each type of static and temporal motifs in the local network as features to a random forest classifier. We find that the classification performance improves by adding the temporal motif features (see Fig. 1C). We are currently further exploring temporal motifs of larger size to obtain more insight and better classification accuracy.

## References

- [1] S. Kodate, R. Chiba, S. Kimura, and N. Masuda. *Applied Network Science*, 5:90, 2020.

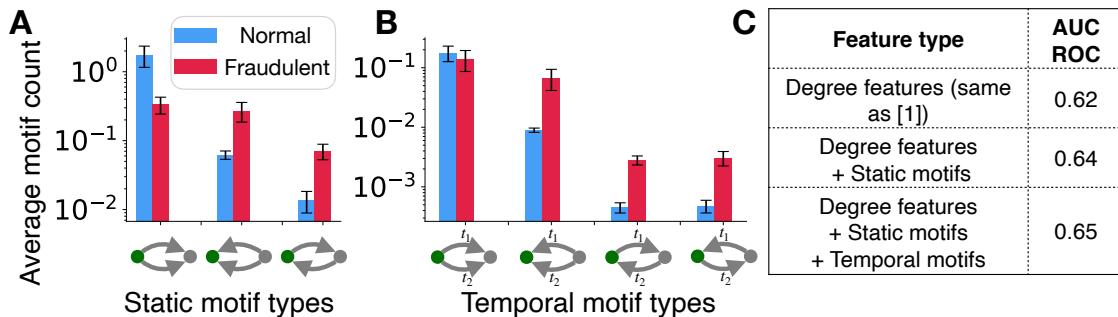


Figure 1: Average count of static (A) and temporal (B) network motifs with two nodes and two edges. The green circle denotes the seed user, and  $t_1$  and  $t_2$  are the timestamps of the first and second events in the temporal motif, respectively. The error bar indicates the standard error. (C) Area under the curve receiver operating characteristics (AUC-ROC) curve as a measure of classification performance. It improves when temporal and static motif counts are added to the degree features from [1].

## Extreme financial losses and gains: mutually-exciting arrivals or conditional volatility?

Matthew F. Tomlinson<sup>1\*</sup>, David Greenwood<sup>2</sup>, and Marcin Mucha-Kruczyński<sup>1</sup>

<sup>1</sup> Department of Physics, University of Bath, Bath BA2 7AY, United Kingdom

<sup>2</sup> CheckRisk LLP, 4 Miles's Buildings, George Street, Bath BA1 2QS, United Kingdom

\* [mft28@bath.ac.uk](mailto:mft28@bath.ac.uk)

Self-exciting bursts of extreme activity are observed in many disparate complex systems, including seismic shocks, neural networks, and financial returns. Statistical models that describe this clustering offer potential insight into the mechanisms behind such extreme events. We construct a two-tailed peaks-over-threshold Hawkes model that captures asymmetric self- and cross-excitation in and between left- and right-tail extreme values within a time series and apply it to extreme gains and losses within the daily log-returns of developed market equity indices [1]. We find that the arrivals of extreme losses and gains are described by a common conditional intensity to which losses contribute twice as much as gains. However, the contribution of the former decays almost five times more quickly than that of the latter. We attribute these asymmetries to the different reactions of market traders to extreme upward and downward price movements: an example of negativity bias, wherein trauma is more salient than euphoria. When compared against standard conditional volatility description of returns, our model of mutually-exciting extremes is found to more accurately forecast out-of-sample extreme values. This suggests a fundamental distinction in the data-generating process for extreme versus bulk returns – a signature that could potentially constrain different classes of systems or characterise different regimes within the same system.

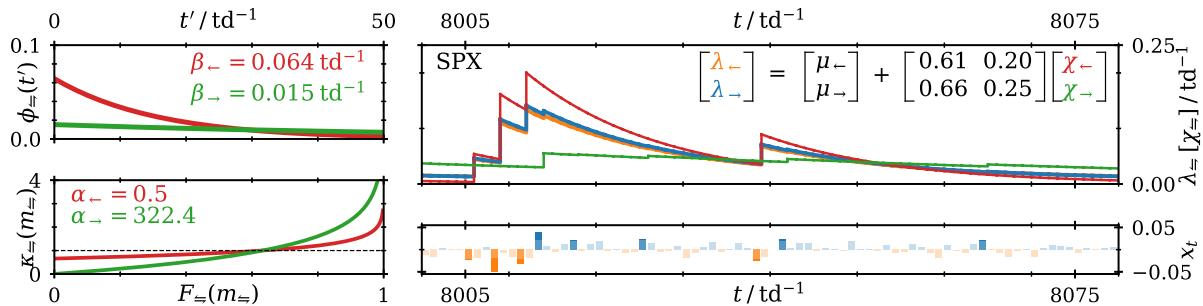


Figure 1: Bivariate 2T-POT Hawkes model fitted to historic S&P 500 daily log-returns [1]. Left: endogenous excitement time kernels  $\phi$  and excess magnitude impact functions  $\kappa$ . Right: response (top, equation shows fitted values of the branching matrix  $\Gamma$ ) of the conditional intensities  $\lambda$  and endogenous excitements  $\chi$  to a cluster of extreme S&P 500 log-returns (bottom, excesses in bold).

## References

- [1] M. F. Tomlinson et al., [arXiv:2011.12291 \[q-n.ST\]](https://arxiv.org/abs/2011.12291) (2020).

## Defining extremism in opinion models

André C. R. Martins<sup>1</sup>

<sup>1</sup> Universidade de São Paulo, São Paulo, Brazil  
amartins@usp.br

Extremism, how it appears and spreads, is one of the main questions studied by opinion dynamics models. And yet, there is no consensus on how to define it mathematically. Even more worrying, the concept of what being an extremist means can be very different when we look at distinct classes of models. When opinions are represented by a continuous number, extremists are often associated with views that lie on the edges of the range of possible values. On the other hand, discrete models, where there is no range, identify extremism as an inability to change opinions. Those definitions are inspired by the mathematical details and limitations of each class. To better understand how those concepts relate to each other, or how they do not, we need more general models where the more traditional and simpler versions can be seen as special cases.

An extension of the Continuous Opinions and Discrete Actions (CODA) model has been suggested as a generalization of both discrete and continuous opinion models. Those extensions explore the Bayesian-inspired ideas of the initial model to create new models from assumptions about the agents' view of the world. In this talk, I will show how to generalize the CODA model so that the general case where there is any number of discrete choices can be treated. The model parameters allow us to describe situations where those choices are all equivalent and the case where they can be better described as positions over a political decision space. Since there is a strength of opinion associated with those choices, this model allows us to distinguish what it means to be at the edge of the spectrum versus an inability to learn due to strong positions. As extremism is a problem related to tragic actions, I will discuss why it might be more related, from a mathematical point of view, to an inability to change one's mind, represented in the model by the strength of the agent opinion, then by the position over a range of possible opinions.

### Acknowledgments

This work was supported by the Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) under grant 2019/26987-2.

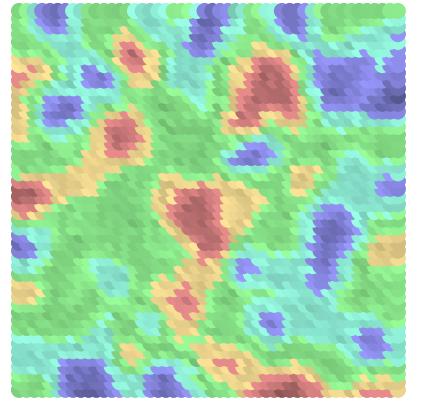


Figure 1: Opinion landscapes showing final preferences for  $M = 5$  choices model after  $T = 50$  average interactions per agent. Each color corresponds to a specific choice, and opinion strength is shown by a lighter (weaker opinion) to darker (stronger) hue.

## Longitudinal Subject Fluency Networks of Psychology and Biology Students

Cynthia S. Q. Siew<sup>1,\*</sup> & Anutra Guru<sup>1</sup>

<sup>1</sup>Department of Psychology, National University of Singapore

\*corresponding author; email: [cynthia@nus.edu.sg](mailto:cynthia@nus.edu.sg)

Within the cognitive sciences, the category fluency task (i.e., list as many items from the semantic category of animals) is commonly used to investigate the representation of concepts within semantic memory [1]. Fluency responses can be represented as networks of interconnected concepts [2]. We attempted to use this task to investigate how knowledge representation of an academic domain (e.g., *biology* and *psychology*) might change over the course of a semester. 200 students enrolled in introductory psychology and 200 students enrolled in introductory biology participated. Students were asked to complete a subject fluency task where they listed as many concepts as they could (in 2 minutes) for their respective course (i.e., psychology students provided fluency responses for the domain of *psychology*). The same group of students completed the task at the beginning (pre-network) and at the end of the semester (post-network). Subject fluency networks were estimated using various network estimation methods (e.g., [3]) and bootstrapping analyses were conducted to statistically compare networks on global network measures (ASPL, CC) and community structure (using the Louvain method [4]). Although post-networks tended to have lower ASPL and higher CC than pre-networks, possibly reflecting increased network connectivity that we had hypothesized to reflect increasing domain expertise, bootstrapping analyses indicated that these differences were not stable across different network estimation methods. This work illustrates both the potential and challenges associated with investigating knowledge representations using the category fluency task.

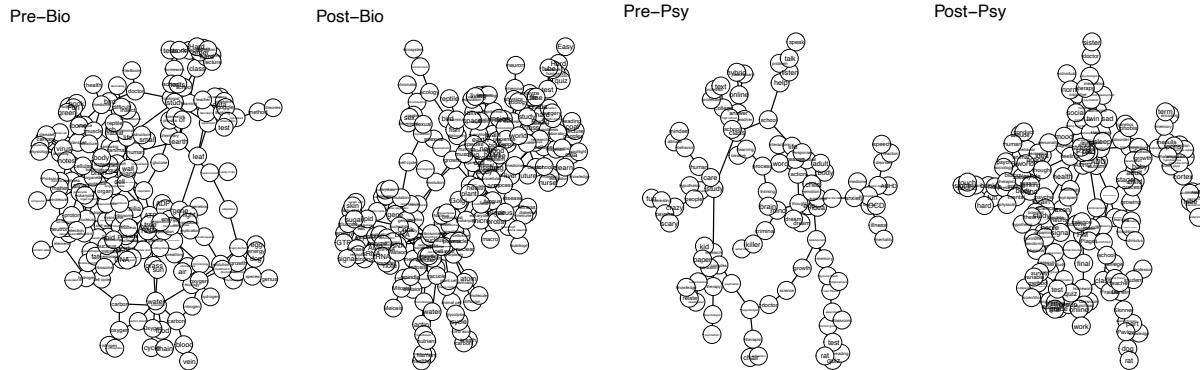


Figure 1. Biology (left) and psychology (right) fluency networks at the beginning and the end of the semester as generated by the community network estimation method.

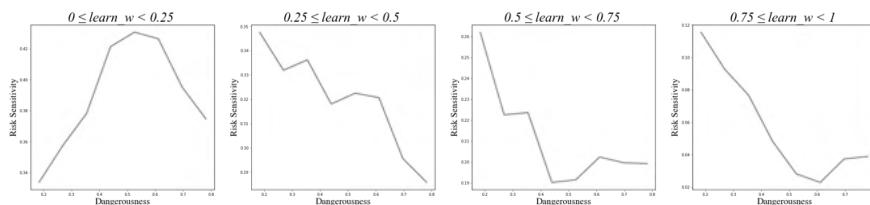
- [1] Bousfield, W. A., & Sedgewick, C. H. W. (1944). An analysis of sequences of restricted associative responses. *The Journal of General Psychology*, 30, 149–165.
- [2] Kenett, Y. N., Anaki, D., & Faust, M. (2014). Investigating the structure of semantic networks in low and high creative persons. *Frontiers in Human Neuroscience*, 8, 407.
- [3] Goñi, J., Arriondo, G., Sepulcre, J., Martincorena, I., de Mendizábal, N. V., Corominas-Murtra, B., ... Villoslada, P. (2011). The semantic organization of the animal category: Evidence from semantic verbal fluency and network theory. *Cognitive Processing*, 12, 183–196.
- [4] Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, P10008.

## Emergence of Risk Sensitivity in a First-principles Agent-based model

F. Bertolotti<sup>l</sup>, A. Locoro<sup>l</sup>, L. Mari<sup>l</sup>. (I)Università Cattaneo - LIUC, Castellanza (VA), Italy, {fbertolotti, alocoro, lmari}@liuc.it

Risk aversion and risk-seeking are acknowledged features of human and non-human behaviour. Biology identifies risk-sensible behaviours in various species like non-human animals, plants and bacteria. Risk preferences also appear at a super-human level, such as in organizations. The risk sensitivity of different entities seems to be shaped by the same underlying principles. Thus, there could be a regularity concerning wherewith living entities originate risk-sensible behaviour. Some existent works examine this occurrence through computer simulations. They show that risk preferences can arise under evolutive forces when individuals interact in unsafe conditions. Other works propose a perspective that includes the role of cultural processes.

The goal of this work is to identify risk sensitivity emergence from both cultural and genetic adaptation. Moreover, it aims at finding relationships between this emergence, the adaptation styles and the environmental variables. It presents an agent-based model in which a single variety of agents exists on a two-dimensional toroidal surface divided into squared cells. Each cell can either contain an energy source or be empty. Energy sources have an initial level of energy, progressively drained by agents. Once an energy source is consumed, it disappears and another one appears in a different cell. Agents use a given amount of energy at each time step. When an agent terminates energy, it dies. It implies that agents necessitate recharging to survive. There are two alternatives to do it. An agent can head towards an energy source and supply from it. Otherwise, it can move towards a different agent and attack it with a given probability of success. Therefore, both attacking and defending agents can win. We analysed scenarios with various chances of victory. A winning agent subtracts a share of energy from the disputant. So, the model forces agents to decide between a risky and safe opportunity. For each possibility, agents compute a payoff, which is the distance from the option times the desirability of that specific event. Risk sensitivity is the difference between risk proneness and risk aversion of agents. Each agent has two risk sensitivities. One is genetic, and it is inherited from parents. The other comes from individual learning and the transmission (partial or complete) of the experience to offsprings. Distinct simulations consider populations that give different importance to knowledge and genotype (i.e. the adaptation style).



**Figure 1:** Risk sensitivity of population plotted on the dangerousness of the system with respect to different style of adaptation, from evolution (left) to learning (right)

The simulation of the model confirms both risk aversion and risk-seeking behaviours emerge from the interaction of agents. A relationship between environmental variables and the emergence of risk sensitivity is found. This link seems to be affected by the adaptation style. Besides, it seems to exist a non-monotonic non-linear relationship between environmental dangerousness and the emergence of risk sensitivity (Figure 1). We propose this is due to the combination of two elements: the existence of incentives to develop risk preferences, which varies with the dangerousness of the surroundings; the direct relationship between the harshness of the environment and how beneficial is the development of risk preferences. The simplicity of this model indicates that these findings could apply to different application fields. Future studies include the generalization of the results and the identification of connections between the pace of life and the emergence of risk preferences.

## **A Modular Common Pool Resource Model with Social Planning**

Andreas Pape and Peter DiCola

Ostrom's famous Principles of Polycentric Governance (Ostrom, 1990) are characteristics of institutions that manage common pool resources (CPRs) that she found, from case studies, tend to describe those that successfully administer those CPRs. This prompts the observation: If these principles tend to improve management of CPRs by a community, and successful management would increase the long-term viability of the community, then in an appropriately specified evolutionary, agent-based model, we should see the emergence of these principles. This project attempts to validate this observation.

In this aspect of this larger project, we share a general, modular Common Pool Resource model with GA-driven agents who can seek to maximize their own or social benefit. We also provide a "Policy" framework and introduce a GA-driven social planner who selects among policies. We show that in some CPRs, the social planner is able to find a policy to persuade "selfish" agents to behave like "altruistic" ones; one aspect of the emergent policies is that they feature graduated sanctions, which is one of Ostrom's principles.

# Modeling behavioral experiments on uncertainty and cooperation with population-based reinforcement learning – Extended Abstract

Elias Fernández Domingos<sup>1,2,\*</sup>, Jelena Grujić<sup>1</sup>, Juan C. Burguillo<sup>3</sup>, Francisco C. Santos<sup>4</sup>, Tom Lenaerts<sup>1,2</sup>

<sup>1</sup> AI-Lab, Computer Science Department, Vrije Universiteit Brussel, Brussels, 1050, Belgium

<sup>2</sup> MLG, Département d'Informatique, Université Libre de Bruxelles, Brussels, 1050, Belgium

<sup>3</sup> atlanTTic Research Center, Universidade de Vigo, 36310 Vigo, Spain

<sup>4</sup> INESC-ID & IST, Universidade de Lisboa, 2744-016 Porto Salvo, Portugal

\* elias.fernandez.domingos@vub.be

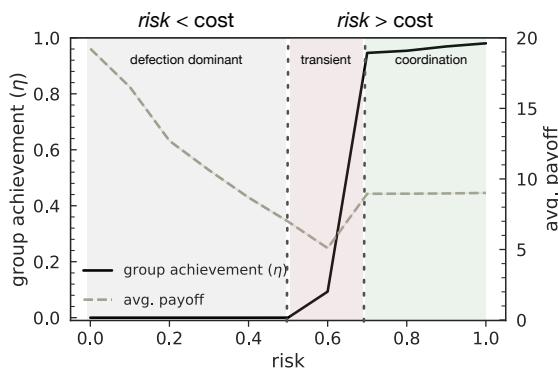


Figure 1. Effect of risk on group achievement and average normalized payoff.

From climate action to public health measures, human collective endeavors are often shaped by different uncertainties. In a manuscript recently published in the Journal of Simulation Practice and Theory [1] we introduce a novel population-based learning model wherein a group of individuals facing a collective risk dilemma [2] acquire their strategies over time through reinforcement learning, while handling different sources of uncertainty. In such an N-person collective risk dilemma players make step-wise contributions to avoid a catastrophe that would result in a loss of wealth for all players. Success is attained if they collectively reach a certain contribution level over time. The dilemma lies in the trade-off between the proportion of personal contributions that players wish to give to collectively reach the goal and the remainder of the wealth they can keep at the end of the game. This tradeoff is affected by the risk that a cataclysm will occur when the collective target is not reached, and all participants lose their remaining wealth. Figure 1 shows that groups only start to successfully achieve the target when the risk is higher than 0.5. We also show that the strategies learned with the model correspond to those observed experimentally, even when there is uncertainty about either the risk of failing when the goal is not reached, the target that should be attained or the time available to reach it. Furthermore, we confirm that being unsure about the time-window favors more extreme reactions and polarization, diminishing the number of agents that contribute fairly. The population-based on-line learning framework we propose is general enough to be applicable in a wide range of collective action problems and arbitrarily large sets of available policies.

## References

- [1] Fernández Domingos, E., Grujić, J., Burguillo, J. C., Santos, F. C., & Lenaerts, T. (2021). Modeling behavioral experiments on uncertainty and cooperation with population-based reinforcement learning. *Simulation Modelling Practice and Theory*, 109, 102299.
- [2] Milinski, M., Sommerfeld, R. D., Krambeck, H. J., Reed, F. A., & Marotzke, J. (2008). The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences*, 105(7), 2291-2294.

# Evolution of brain networks after traumatic brain injury

Tong Wu<sup>1</sup>, Matthew McGuire<sup>2,3</sup>, David Poulsen<sup>2,3</sup>, and Sarah F. Muldoon<sup>1,2,4</sup>

<sup>1</sup> CDSE Program

<sup>2</sup> Neuroscience Program

<sup>3</sup> Department of Neurosurgery

<sup>4</sup> Department of Mathematics

University at Buffalo, State University of New York

The brain can be described as a complex network that evolves over time in both health and disease states. However, how to track meaningful or predictable changes in brain networks remains an unsolved problem. In particular, brain dynamics are significantly altered in many brain diseases (e.g. in epilepsy or Parkinson's disease) which suggests that for early prevention and treatment, it is crucial to study changes in underlying brain network structure that can give rise to these observed changes in brain network dynamics. Changes in brain network structure associated with epileptogenesis are particularly pertinent in regards to the development of post-traumatic epilepsy (PTE) resulting from traumatic brain injury (TBI). Diffusion tensor imaging (DTI) can be used to measure such changes in white matter integrity measured through the Fractional Anisotropy (FA) index and/or tractography analysis. In our study, we acquired *in vivo* DTI data from both TBI and uninjured control (sham) groups of rats in hopes of linking specific changes in brain network structure to the process of epileptogenesis and the development of PTE. We find both increases and decreases in FA values in certain regions, indicated network induced changes in brain structure resulting from injury. Specifically, we see a reduction of the FA value in the right corpus callosum and primary auditory cortex of TBI animals, likely representing the initial injury. However, also observe an increase in the FA value for the left side parietal cortex, indicating that the injury results in more widespread changes in brain network structure.

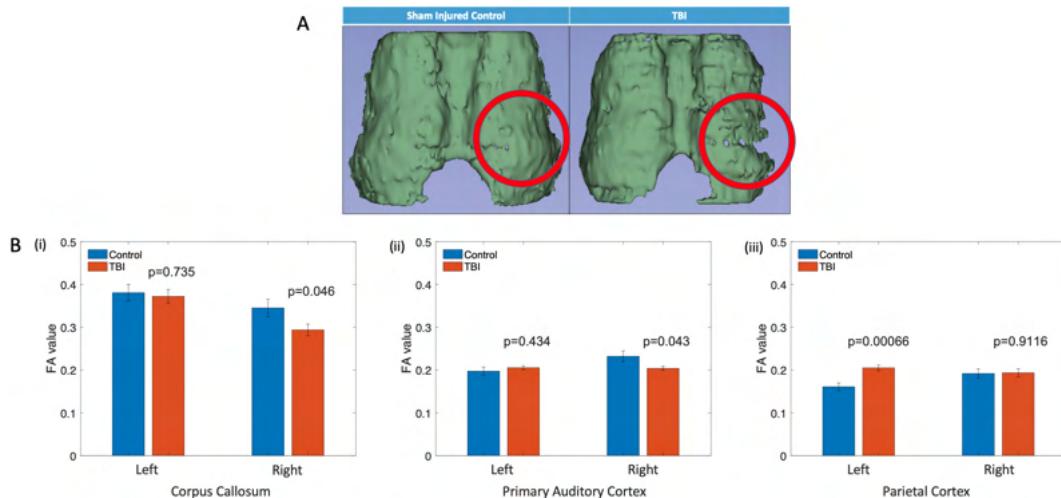


Figure 1: A) 3D reconstructed models of the corpus callosum from FA values. Red circle represents site of injury. B) (i) Mean FA values for the Corpus Callosum, (ii) Primary Auditory Cortex, and (iii) Parietal Cortex in Control (blue) and TBI (red) groups for left and right hemispheres.

# Optimizing Wind Farm Efficiency with Self-Adaptivity Mechanism in Evolutionary Algorithm

Xinglong Ju<sup>1</sup> and Feng Liu<sup>2</sup>

<sup>1</sup> School of Civil and Environmental Engineering, Cornell University

<sup>2</sup> School of Systems and Enterprises, Stevens Institute of Technology

## 1 Introduction

In the past decades, a significant amount of wind farms across the world have been constructed to reduce greenhouse gas emission across the world. Numerous algorithms have been proposed to optimize layout of wind farms and reducing the impact of *wake effect*. Wake effect refers to the input wind speed reduction after wind passing through the turbines in the upwind direction. Genetic algorithm has been widely used to find the optimal solution of wind farm layout.

## 2 Methods

In this study, we devised an insightful bi-criteria identification and relocation (BCIR) mechanism in genetic algorithm to alleviate the solutions trapping into suboptimal conditions. We introduce the Influence Matrix (Fig.1(a)) to guide the bi-criteria wind turbine selection without adding additional computational cost.

## 3 Results

The cost of energy comparison between existing algorithms [1] CGA, AGA, IGA, SIGA, IAGA, and our proposed method (ISIGA) are illustrated in Fig.1(b), and ISIGA achieved the lowest energy cost than other comparison algorithms.

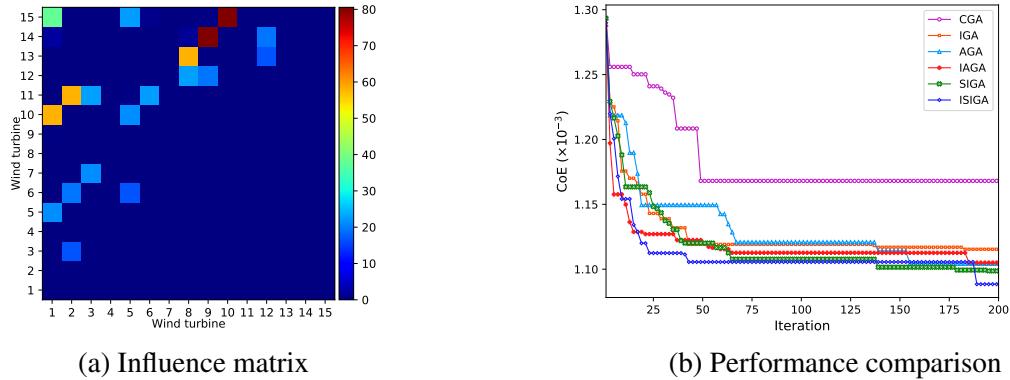


Figure 1: Illustration of Influence Matrix (IM) (a) and cost of energy convergence curves (b).

## 4 Conclusions

Our algorithm demonstrated superior performance than the benchmark algorithms, which demonstrates the effectiveness of self-adaptivity (BCIR mechanism) in genetic algorithm.

## References

- [1] Liu, F., Ju, X., etc. (2020). Wind farm macro-siting optimization with insightful bi-criteria identification and relocation mechanism in genetic algorithm. ECM, 217, 112964.

## Resiliency of Global Medical Equipment Supply Chains

Zachary M. Boyd (Mathematics, UNC—Chapel Hill)  
and Kayvan Lavassani (Business, NC Central Univ.)

Several contemporary events, including cyber-attacks, terrorism, and COVID-19, have highlighted the fragility of global supply chains (Sherkarian et al. 2020). One significant challenge in studying the global supply chains is that companies are highly protective of their trade information, including the identity of suppliers and customers (Okeagu et al. 2021). Using U.S. Security and Exchange Commission filings of all publicly-traded firms, we determined the global network of notable suppliers and customers involved in medical equipment firm (MEF) supply across ten tiers. That is, starting at MEFs (tier 0), we found their suppliers (tier 1), then these suppliers' suppliers (tier 2), and so on, capturing the complex set of relationships required to go from raw materials to usable medical equipment. This is the most extensive medical supply chain network in existence and is the largest possible data set of this kind.

We investigated network resiliency under random disruptions, as well as centrality-based and firm-size-based failures. This included individual firm disruptions and disruptions at the level of industries, countries, and country-industries. In order to understand long-range supply dependencies, we focused on *reachability* of MEFs' *end suppliers*. A firm's end suppliers are the highest-tier (i.e. most distant in the network) suppliers on which they rely. These end suppliers may be producing raw materials that must go through several intermediate parties before being used in medical devices, and disruption in those intermediate links interferes with MEF operations. We say that an end supplier B is *reachable* by a particular firm A if there is a chain of supply relationships from A to B.

We find that the MEF supply chain is vulnerable to both random and targeted disruptions. For example, random disruption of 10% of firms causes MEFs to be able to reach only 60% of their end suppliers, with 20% of MEFs unable to reach *any* end suppliers. For targeted attacks or random failures correlated with firm size, the situation is much worse: with 10% firm disruption, MEFs can only reach 20% of their end suppliers on average. We also consider more specific scenarios. For example, if US-China trade relationships broke down, China's MEFs could reach on average 57% of their end suppliers, while the US MEFs are essentially unaffected.

### References

- Shekarian M., Nooraie S.V.R., Parast M.M. (2020). An examination of the impact of flexibility and agility on mitigating supply. *Intr. J. Production Econ.*, 220, pp 1-16.