Research Program

PHILIP S. CHODROW

I am a data scientist and applied mathematician with interests in machine learning, network data, and modeling of social systems. My research program aims to advance scalable, statistically grounded techniques for the study of complex and networked data. My portfolio includes both single-author publications and collaborations across a range of institutions and disciplines. My work appears in applied mathematics journals [1, 2], network-science journals [3, 4], and high-impact interdisciplinary journals [5, 6, 7]. My primary research directions are:

(D1) **Models and Algorithms for Network Data Science.** Network modeling offers a powerful formalism for studying interconnected systems in many social, biological, and technical domains. I work at the intersection of theory and computation to develop models and algorithms for network data science. I am especially interested in problems involving hypergraphs, which model polyadic relationships in networked systems. My work on hypergraphs addresses spectral theory (ongoing work), random models [3, 4], and clustering algorithms [7]. A key thrust in my work on network algorithms is the use of principled approximations to enable scalable data analysis in settings where exact computation may be infeasible [1, 7].

(D2) **Inference and Dynamics in Biosocial Systems.** One of my longstanding interests is the use of modeling to inform our understanding of social inequality, inequity, and division. In one vein, I develop stochastic models for the formation and impact of hierarchies in human and animal societies ([6] and ongoing work). In another vein, I develop and analyze nonlinear models of opinion polarization in social networks. Many well-known opinion models possess certain mathematical pathologies. I like to approach such models by embedding them within well-behaved parameterized families and studying the long-term behavior of the target model via suitable limits ([2] and ongoing work). In some cases, the parameterized embedding families are of intrinsic interest, displaying rich behaviors beyond those of the original target model.

(D3) **Data Science for Justice, Equity and Sustainability.** I believe deeply in the potential of computational tools to illuminate inequities and promote a more just and inclusive world. I both develop novel techniques for data analysis and directly support equity-focused data-science research. In an early paper, I developed information-geometric tools for studying spatial division, focusing on ethnoracial residential segregation in the U.S. [5]. This work has recently grown into a collaboration on pollution inequality environmental with a multi-institution team of environmental scientists [8]. I am also currently working with undergraduate collaborators on data-driven projects related to gender representation in mathematical subfields and racial disparities in criminal-sentencing decisions.

Long term, I aim to pursue my work at a primarily-undergraduate institution. Many of the problems that I study are accessible to undergraduates, and one of my goals is to develop a robust line of undergraduate research.

D1. Models and Algorithms for Network Data Science

1.1. **Foundations of Hypergraph Clustering.**

A hypergraph consists of a set of nodes and a set of edges, each of which is a set containing any number of nodes. Hypergraphs generalize dyadic graphs and provide a natural modeling framework for systems in which entities interact polyadically in groups of two or more. Common interaction mechanisms include collaboration, communication, and co-presence.

1.1.1. *Spectral Properties and Algorithms.* Spectral graph theory is one of the pillars of the mathematical treatment of relational systems. Spectral methods analyze graphs by assigning to them one or more matrices—such as adjacency or Laplacian matrices—whose eigenvalues and eigenvectors encode information about graph structure. The extension of spectral methods to hypergraphs poses challenges. Given a fixed hypergraph, one approach is to extract from it all pairwise relationships and thereby obtain a dyadic graph. One can then extract matrices from this graph. This procedure, however, destroys polyadic structure [4]. When a hypergraph is *uniform*—with all edges containing the same number of nodes—one can encode its adjacency structure in a symmetric tensor and tensorial spectral methods may be applied [9]. Much polyadic data, however, is nonuniform, and cannot be represented by a single tensor.

It is, however, possible to obtain a matrix representation of a nonuniform hypergraph that retains polyadic relationships. The *Hashimoto matrix* $\mathbf{B}$ of a hypergraph [10] is indexed by node–hyperedge pairs, with entries

$$b_{(ie),(jf)} = \begin{cases} 1\,, & i \in e,\, j \in f,\, i \in f \setminus \{j\} \\ 0\,, & \text{otherwise}\,. \end{cases}$$

That is, $b_{(ie),(jf)} = 1$ if edge f can be reached from edge e by traversing node $i \neq j$. Since no reduction to a pairwise graph is required, the matrix $\mathbf{B}$ maintains polyadic information in edges of differing sizes. The cost of this representation is size: if $\bar{k}$ is the mean size of a hyperedge and m the total number of hyperedges, then $\mathbf{B} \in \{0,1\}^{m\bar{k} \times m\bar{k}}$. Computation on such a matrix can be prohibitive for data sets of even a few hundred nodes.

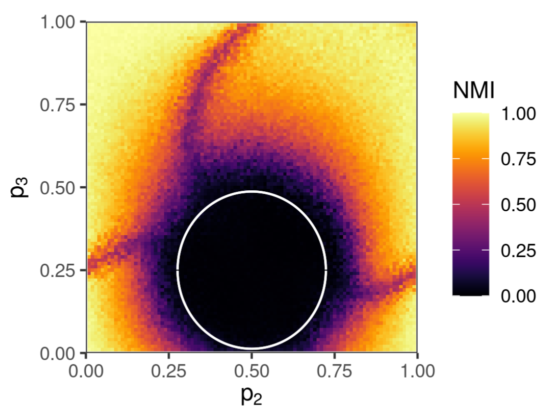
In ongoing work, Jamie Haddock (Harvey Mudd), Nicole Eikmeier (Grinnell), and I are addressing this computational challenge and deploying the Hashimoto matrix for hypergraph clustering tasks. We first prove a generalization of the celebrated Ihara–Bass theorem [11]. This generalization allows us to characterize the spectrum of $\mathbf{B}$ in terms of a smaller matrix that contains only pairwise adjacency information:

Theorem 1 (PSC, JH, NE '21)**.** *Let H be a hypergraph on n nodes with m_k edges of size k for each $k = 2, 3, \ldots, K$, and let $\mathbf{B}$ be its Hashimoto matrix. Then:*

(1) *For each k, if $m_k > n$, then $\lambda = 1 - k$ is an eigenvalue of $\mathbf{B}$ with multiplicity $m_k - n$.*
(2) *If $\sum_k m_k(k-1) > n$, then $\lambda = 1$ is an eigenvalue of $\mathbf{B}$ with multiplicity $\sum_k m_k(k-1) - n$.*
(3) *If the conditions in (1) and (2) are both met, then the remaining eigenvalues of $\mathbf{B}$ are eigenvalues of the $2n(K-1) \times 2n(K-1)$ matrix*

$$\mathbf{B}' = \begin{bmatrix} \mathbf{0}_{n(K-1)} & \mathbf{D} - \mathbf{I}_{n(K-1)} \\ (\mathbf{I}_{K-1} - \mathbf{K}) \otimes \mathbf{I}_n & \mathbf{A} + (2\mathbf{I}_{K-1} - \mathbf{K}) \otimes \mathbf{I}_n \end{bmatrix},$$

where $\mathbf{A}$ is a block pairwise adjacency matrix; $\mathbf{D}$ is a matrix of node degrees; $\mathbf{K}$ is a matrix of edge sizes; $\mathbf{0}_\ell$ and $\mathbf{I}_\ell$ are the zero and identity matrices of size ℓ respectively; and $\otimes$ is the matrix Kronecker product.



We then consider applications to hypergraph spectral clustering. Generalizing results for graphs [12], we show that the Jacobian of the belief-propagation algorithm for inference in clustered hypergraphs can be approximated by a matrix that is composed of weighted blocks of $\mathbf{B}$. Eigenvectors of this Jacobian with eigenvalues larger than 1 in magnitude contain information that can be used to cluster nodes. We propose a simple consensus-clustering algorithm for aggregating this information, and analytically derive approximate conditions under which it is able to recover cluster information with positive normalized mutual information (NMI) relative to planted labels in sparse synthetic hypergraphs. This phenomenon is illustrated above: for combinations of parameters inside the white ellipse, the algorithm fails to detect the true labels, while outside the ellipse success is possible. In planned work, we aim to prove concentration results on the spectrum of $\mathbf{B}$ under random hypergraph models; such results would imply probabilistic guarantees for the success of our algorithm on synthetic data.

1.1.2. *Scalable Clustering.* Even with matrix reductions such as Theorem 1, spectral methods can be limited by the computational complexity of eigenvalue computations. In recent work that was published in *Science Advances*, Nate Veldt (Cornell, Texas A&M), Austin Benson (Cornell), and I introduced a degree-corrected Poisson hypergraph stochastic blockmodel, a random-hypergraph model with multiple, densely-connected subsets of nodes [7]. By analyzing the likelihood of this model, we derived a novel objective function that generalizes the popular modularity functional for graph clustering [13]. We then proposed an efficient optimization heuristic based on the popular Louvain clustering algorithm [14]. In a sequence of experiments, we demonstrated our algorithm on hypergraph data of up to one million nodes. We also showed that our model is able to detect planted partitions in hypergraphs even in regimes in which methods based on pairwise graph representations must necessarily fail due to information-theoretic bounds [15]. By studying several popular hypergraph data sets, we found that our algorithm generates qualitatively different partitions than graph-based methods. We also found that the clusters returned by our proposed algorithm were often more interpretable or better aligned with ground-truth data labels.

1.2. **Configuration-Model Random Graphs and Generalizations.** Null-hypothesis testing is a cornerstone of frequentist statistics. The null hypothesis is a probability distribution over counterfactual data realizations that preserves certain structures and randomizes others. In the setting of graph data analysis, *configuration models* are a popular class of random graph models used as null distributions. The *degree* of a node is the number of edges incident to it, and the *degree sequence* collects the degree of each node in a graph. Configuration models are random graphs that preserve the degree sequence [16]. The analysis and application of configuration models can pose surprising mathematical and statistical challenges.

1.2.1. *Moments of Random Multigraphs.* Let $\mathcal{M}_{\mathbf{d}}$ be the set of multigraphs on n nodes with prescribed degree sequence $\mathbf{d} \in \mathbb{Z}_+^n$, and let $\eta_{\mathbf{d}}$ be the uniform distribution on $\mathcal{M}_{\mathbf{d}}$. Despite the simplicity of its definition, properties of $\eta_{\mathbf{d}}$ can usually be computed only approximately and under sparsity assumptions [17, 18, 19]. The complexity of this distribution has statistical consequences: standard Markov-chain methods for sampling from $\eta_{\mathbf{d}}$ can be impractically slow for data of even moderate size [20, 21]. In many applications, however, it is not necessary to draw complete samples from $\eta_{\mathbf{d}}$; it instead suffices to estimate certain moments. In a recent paper [1] in *SIAM Journal on Mathematics of Data Science*, I showed how to estimate moments of $\eta_{\mathbf{d}}$ while bypassing intensive Markov chain sampling.

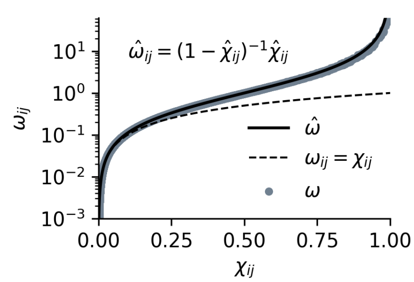
An especially important object in applications is the expectation $\Omega = \mathbb{E}[\mathbf{W}]$ of the adjacency matrix $\mathbf{W}$ of a multigraph that is distributed according to $\eta_{\mathbf{d}}$. By analyzing the dynamics of a Markov-chain sampling algorithm for this model, I obtained estimates for the entries ω_{ij} of Ω. Let $\mathbf{X} = [\mathbb{I}(W_{ij} \geq 1)_{ij}]$, with expected entries $\chi_{ij} = \mathbb{E}[X_{ij}]$. Let $\boldsymbol{\beta} = \sum_j \chi_{ij}$. For fixed $\mathbf{d}$, the component β_i gives the expected number of distinct neighbors of node i under $\eta_{\mathbf{d}}$, and is a function of the degree sequence $\mathbf{d}$. Finally, let $u(\mathbf{d})$ be the Lipschitz constant of $\boldsymbol{\beta}$ restricted to the set $\{\mathbf{d}' : \mathbf{d}' \geq \mathbf{d}\}$ entrywise.

Theorem 2 (PSC '20). *For any pair of nodes $i \neq j$, it holds that*

$$\chi_{ij} = \frac{\beta_i \beta_j}{\|\boldsymbol{\beta}\|_1} + u_*(\mathbf{d})\epsilon_0 \quad \text{and} \quad \omega_{ij} = \frac{\chi_{ij}}{1 - \chi_{ij}} + u_*(\mathbf{d})\epsilon_1\,,$$

where $\epsilon_0 = O\left(\frac{\|\boldsymbol{\beta}\|_\infty}{\|\boldsymbol{\beta}\|_1}\right)$ and $\epsilon_1 = O\left(\sqrt{\frac{\|\boldsymbol{\beta}\|_\infty}{\|\boldsymbol{\beta}\|_1}}\right)$ in the limit $n \to \infty$.

The relationship between ω_{ij} and ξ_{ij} is illustrated at right: grey dots are estimated from a contact network via Markov-chain sampling, while the black line is the estimate provided by Theorem 2. In practice, $\boldsymbol{\beta}$ is not known and must be estimated as a function of $\mathbf{d}$. Theorem 2 implies that $\boldsymbol{\beta}$ must approximately solve the following system of nonlinear equations: for all i,



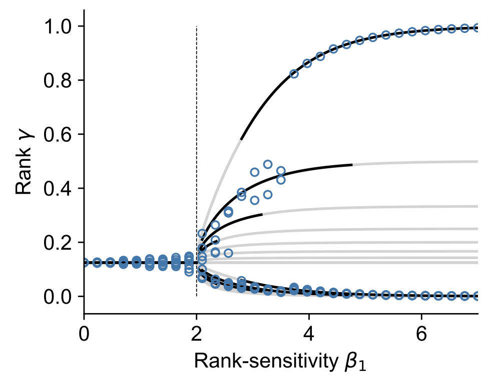
$$h_i(\boldsymbol{\beta}) \triangleq \sum_j \frac{\beta_i \beta_j}{\|\boldsymbol{\beta}\|_1 - \beta_i \beta_j} = d_i\,. \tag{1}$$

I proved a qualified uniqueness result on the solution set of (1), and developed a coordinate-wise Newton method for computing $\boldsymbol{\beta}$. In computational experiments, I showed that the solution of (1) improves on a popular heuristic [13], reducing the relative error in estimating entries of Ω by over an order of magnitude. I conjecture that it is possible to significantly tighten the bounds on the error terms in Theorem 2, and I hope to pursue this problem in future work.

1.2.2. *Null Models for Hypergraphs.* I have recenly proposed extensions of configuration models to the setting of hypergraphs. In work [4] that was published in the *Journal of Complex Networks*, I showed how to define hypergraph configuration models: probability distributions over sets of hypergraphs that fix both the degree of each node and the size of each hyperedge. I formulated and proved Markov-chain sampling algorithms for two such models. I then applied these models to several problems involving null-hypothesis testing in hypergraphs. In follow-up work [3] in *Applied Network Science*, Andrew Mellor (Oxford) and I defined a configuration model on *annotated hypergraphs* in which different nodes can play distinct roles within edges. For example, scientific publications can have *leading* and *senior* authors, each of whom may have contributed in different ways to the collaboration. We formulated and proved a Markov-chain sampling algorithm for our model, and applied it to the problem of important agents and detecting densely connected sets of agents in hypergraph data with annotations.

D2. Inference and Dynamics in Biosocial Systems

2.1. **Dynamics of Prestige-Based Hierarchies.** Many human and animal societies are structured by enduring social hierarchies. One mechanism of hierarchy in human societies is *prestige*. In a prestige-based hierarchy, the social position of an agent depends primarily on how they are perceived by others. In recent work [6] that was published in the *Proceedings of the National Academy of Sciences*, I collaborated with Mari Kawakatsu (Princeton), Nicole Eikmeier (Grinnell), and Dan Larremore (CU Boulder) to study the mathematical conditions under which prestige-based hierarchies form and persist. The state of our model at time t is a matrix $\mathbf{A}(t) \in \mathbb{R}^{n\times n}_{\geq 0}$ of nonnegative real numbers, which represent previous interactions between the n agents. At each discrete time step t, we compute the *score vector* $\mathbf{s}(t) = \sigma(\mathbf{A}(t))$, where $\sigma : \mathbb{R}^{n\times n} \to \mathbb{R}^n$ assigns a real number to each agent based on the current state. Then, each agent i chooses to endorse another agent j with probability $\gamma_{ij}(t) \propto e^{u_{ij}(\mathbf{s}(t))}$. Here, u_{ij} is a utility function $u_{ij}(\mathbf{s}) = \beta_1 s_j + \beta_2 (s_j - s_i)^2$. The first parameter β_1 is interpretable as a prestige preference: high β_1 indicates that agents strongly prefer to endorse agents with high scores. The second parameter β_2 is interpretable as a proximity preference: high β_2 indicates that agents prefer to endorse other agents with scores that are close to their own. These interactions are collected into a matrix $\mathbf{\Delta}(t)$, and the system state is updated as $\mathbf{A}(t+1) = \lambda \mathbf{A}(t) + (1-\lambda)\mathbf{\Delta}(t)$, where $\lambda \in [0,1]$ is a timescale parameter.



When β_1 is large, prestige-based hierarchies emerge. For three distinct score functions σ, we prove the existence and derive the location of a bifurcation in the parameter β_1 in the limit of long system memory. At this bifurcation, the egalitarian regime in which agents have equal scores becomes linearly unstable, and a persistent hierarchy emerges. We also observe the presence of multistable regimes, in which steady states with qualitatively different degrees of hierarchical organization are possible. These phenomena are illustrated at right for the SpringRank score function [22]. Blue points give simulation results, black curves are stable equilibria of a deterministic approximation, and grey curves are unstable equilibria. The critical point at $\beta_1^c = 2$ is highlighted.

A distinctive feature of our model is that it possesses a computationally tractable likelihood, allowing us to fit it to real data in principled fashion. We do this for four data sets from human and animal societies, finding interpretable similarities between them. In all cases, we estimate that $\beta_1 > 0$ (indicating prestige preference) and $\beta_2 < 0$ (indicating proximity preference). We furthermore find that each system lies in a multistable regime near the bifurcation point. In this regime, both weakly or strongly hierarchical long-term outcomes are possible. This finding suggests the intriguing possibility of interventions that might move complex, prestige-based systems from more hierarchical states to less hierarchical ones.

This work was highlighted in a *PNAS* commentary [23] and will appear soon as an invited post on the *SIAM News* blog.

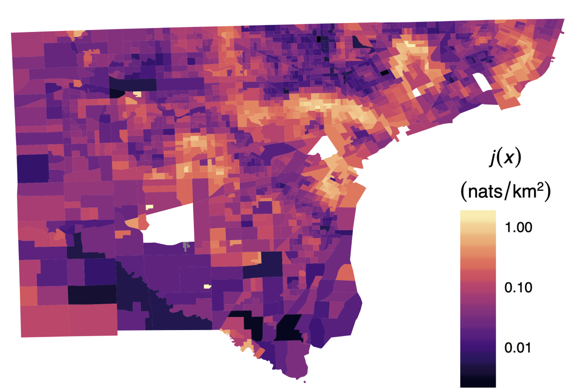
2.2. **Animal Hierarchies from Heterogeneous Behavior.** Inference of dominance hierarchies in animal societies poses methodological challenges. Animal-behavior data is often collected manually, with researchers logging many different types of behavior. These different kinds of behavior carry different information about latent hierarchical structure [24]. In ongoing work, Kelly Finn (Dartmouth), Mason Porter (UCLA), and I propose a novel generative model of heterogeneous dominance interactions between individuals. Our model has parameters that regulate how strongly each type of interaction follows the direction of a latent hierarchy. Both these parameters and the hierarchical ranks themselves can be inferred from data. We plan to develop statistically grounded treatments of several methodological problems in the inference of animal hierarchies. These include the determination of when different interaction types should be aggregated during data collection, power analyses for informing data collection requirements, and systematic accounting for biases that arise from common observational methodologies.

2.3. **New Methods for Opinion-Dynamics Models.** In recent work [2] in *SIAM Journal on Applied Mathematics*, Peter Mucha (UNC Chapel Hill, Dartmouth) and I studied a nonlinear model of opinion fragmentation in social networks. In this model, agents possess a binary opinion and a network of social contacts. When an agent disagrees with a contact, they may either (with probability $\alpha \in [0, 1]$) update their opinion to match that of their contact or (with probability $1 - \alpha$) sever communication and connect with a new contact uniformly at random. Though simple, this model possesses rich behavior, including a critical point α^*. If $\alpha > \alpha^*$, the network remains connected and a consensus is reached with high probability. If $\alpha < \alpha^*$, however, the system fragments into disconnected components within which opinions are homogeneous. By adding random opinion switches to this model, we developed novel approximations for higher-order moments of this system. Considering the limit as the switch-rate goes to 0, we provided state-of-the-art estimates of both α^* and the system behavior past this critical point.

The Hegselmann–Krause (HK) model [25] is a popular model of opinion dynamics with *bounded confidence*. An agent i with opinion x_i who encounters an agent j with opinion x_j will incorporate j's opinion into their own if and only if $|x_i - x_j| < c$ for some constant c. The all-or-nothing nature of this decision rule limits the use of standard linearization techniques to study the model. In ongoing work, Heather Zinn Brooks (Harvey Mudd), Mason Porter (UCLA) and I are studying a parameterized family of smooth models in which both the the HK model and the DeGroot consensus model [26] appear as limiting cases. We focus on a model variant with "zealot" or "media" nodes [27] whose opinions do not change. For several classes of graphs, we analytically characterize the regimes of stability for unpolarized solutions. Outside these regimes, there exist stable stationary states in which agents separate into strongly polarized factions and are able to exchange information only with neighbors within their own faction. An important theme of this work is the role of graph topology in governing the stability of opinion configurations.

D3. Data Science For Justice, Equity and Sustainability

3.1. **Information geometry and spatial segregation.** In a single-author paper [5] in the *Proceedings of the National Academy of Sciences*, I developed a set of information-geometric tools for studying spatial segregation. I proved that a wide range of extant measures of segregation from the sociology literature can be unified under the formalism of Bregman information measures. I also proved a theorem that relates a localized version of the Bregman information to the mean curvature of an information manifold. This manifold is embedded in the space of discrete probability distributions, parameterized by spatial coordinates, and endowed with a Riemannian metric induced by a Bregman divergence. I used this correspondence to measure local scales of spatial segregation, as shown at right for the city of Detroit: larger values of the mean local information $j(x)$ correspond to more locally-segregated regions of the city. I also developed algorithms for hierarchical visualization of segregation in cities. Recently, I joined a team of collaborators in environmental science led by Angelique Demetillo (UVA) and Sally Pusede (UVA) to use this methodology to study racial disparities in air pollution in major U.S. cities [8]. Interest in these techniques appears to be growing in the spatial data-science community, and I look forward to more such collaborations in the future.



3.2. **Equity-Oriented Data Science.** Are some mathematical subfields more gender-diverse than others? In ongoing work, Ben Brill (UCLA, undergraduate mentee), Mason Porter (UCLA), Heather Zinn Brooks (Harvey Mudd), and I are studying this question using data from the Mathematics Genealogy Project. Through data analysis, we also plan to learn how subfields vary in their proportions of gender-minoritized mathematicians. We plan to use stochastic models on networks to study the mechanisms by which subfields succeed or fail in diversifying over time.

In a recent class project, my student Hinal Jajal (UCLA, undergraduate mentee) scraped sentencing records from the Virginia state court system. She obtained a complete set of over 2 million distinct sentences. This Fall, she and I will collaborate to analyze this data set. We are particularly interested in whether these data exhibit systematic racial disparities in sentences and whether these disparities are correlated with particular courts or judges. Long-term, we hope to publish a scholarly article and white paper on racial disparities in state-level criminal sentencing in collaboration with a team at the Institute for the Quantitative Study of Inclusion, Diversity, and Equity (QSIDE).

Mentorship of Undergraduate Research

My research interests are well-suited to collaboration with undergraduates. Many of the problems that I study have a relatively low technical barrier to entry. Students who have taken courses in modeling, probability, or dynamical systems are well-placed to make significant research contributions that are publishable in computer science, mathematics, physics, and interdisciplinary venues.

- (D1) There are rich research opportunities for students who are interested in modeling random networks and developing algorithms to study them. One especially fertile area is the analysis of hypergraphs and other "higher-order" data structures [28]. Important open problems include the further development of spectral theory and random walks on non-uniform hypergraphs, information-theoretic thresholds for statistical algorithms, and generative modeling of polyadic systems. In many cases, progress in these areas may depend more on creative thinking than on a high degree of technical sophistication. Students can contribute either theory or data-driven research, depending on their interests.
- (D2) There are many opportunities for undergraduates to contribute in the mathematical modeling of social and biological systems. Models of both hierarchy formation and opinion dynamics can be pursued from either analytical or computational perspectives, and they can lead to engagement with empirical data. In both cases, these models can be extended in a variety of directions, depending on student interest. Such extensions include change points in parameters, more complex specifications of agent behaviors, and more structured spaces of possible agent states. It is often relatively straightforward to observe the influence of a given modeling decision on macroscopic system behavior, making this a convenient area for students to rapidly test their ideas. Theoretically inclined students may then to seek approximations to the observed behavior, while computationally inclined students may pursue large-scale computational experiments or develop algorithms for learning model parameters from available data.
- (D3) Equity-oriented data science is an exciting and accessible area, with new challenges and data sets emerging regularly. While some of these data sets can be understood using out-of-the-box inference algorithms, others may require the development of custom mathematical and statistical machinery. As described above, I am already working with two undergraduate mentees on such projects. There are numerous opportunities for students to combine interests in computational science and social good, thereby producing interesting, impactful research.

These interests lend themselves well to interdisciplinary collaborations, which I anticipate will closely involve my undergraduate research mentees.

I was given several opportunities as an undergraduate to develop my research interests. I am eager to pay it forward.

References

[1] **PSC**. Moments of uniformly random multigraphs with fixed degree sequences. *SIAM Journal on Mathematics of Data Science*, 2(4):1034–1065, 2020.

[2] **PSC** and Peter J. Mucha. Local symmetry and global structure in adaptive voter models. *SIAM Journal on Applied Mathematics*, 80(1):620–638, 2020.

[3] **PSC** and Andrew Mellor. Annotated hypergraphs: Models and applications. *Applied Network Science*, 5(9), 2020.

[4] **PSC**. Configuration models of random hypergraphs. *Journal of Complex Networks*, 8(3):cnaa018, 2020.

[5] **PSC**. Structure and information in spatial segregation. *Proceedings of the National Academy of Sciences*, 114(44):11591–11596, 2017.

[6] Mari Kawakatsu*, **PSC***, Nicole Eikmeier*, and Daniel B Larremore. Emergence of hierarchy in networked endorsement dynamics. *Proceedings of the National Academy of Sciences*, 118(16):e2015188118, 2021.

[7] **PSC**, Nate Veldt, and Austin R. Benson. Generative hypergraph clustering: From blockmodels to modularity. *Science Advances*, 7:eabh1303, 2021.

[8] Mary A. Demetillo, Colin Harkins, Brian McDonald, **PSC**, Kang Sun, and Sally Pusede. Space-based observational constraints on NO2 air pollution inequality from diesel traffic in major U.S. cities. *Geophysics Review Letters*, 48(17):e2021GL094333, 2021.

[9] Lek-Heng Lim. Singular values and eigenvalues of tensors: a variational approach. In *1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 2005.*, pages 129–132. IEEE, 2005.

[10] Christopher K Storm. The zeta function of a hypergraph. *The Electronic Journal Of Combinatorics*, 13(R84):1, 2006.

[11] Hyman Bass. The Ihara-Selberg zeta function of a tree lattice. *International Journal of Mathematics*, 3(06):717–797, 1992.

[12] Florent Krzakala, Cristopher Moore, Elchanan Mossel, Joe Neeman, Allan Sly, Lenka Zdeborová, and Pan Zhang. Spectral redemption in clustering sparse networks. *Proceedings of the National Academy of Sciences*, 110(52):20935–20940, 2013.

[13] Mark E. J. Newman and Michelle Girvan. Finding and evaluating community structure in networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 69(2):1–16, 2003.

[14] Vincent D. Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal Of Statistical Mechanics - Theory And Experiment*, 10:1–12, 2008.

[15] Aurelien Decelle, Florent Krzakala, Cristopher Moore, and Lenka Zdeborová. Inference and phase transitions in the detection of modules in sparse networks. *Physical Review Letters*, 107(6):065701, 2011.

[16] Bailey K. Fosdick, Daniel B. Larremore, Joel Nishimura, and Johan Ugander. Configuring random graph models with fixed degree sequences. *SIAM Review*, 60(2):315–355, 2018.

[17] Michael Molloy and Bruce Reed. A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms*, 6(2-3):161–180, 1995.

[18] Omer Angel, Remco van der Hofstad, and Cecilia Holmgren. Limit laws for self-loops and multiple edges in the configuration model. *arXiv:1603.07172*, 2016.

[19] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.

[20] Catherine Greenhill. The switch Markov chain for sampling irregular graphs. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1564–1572. SIAM, 2014.

[21] Catherine Greenhill. A polynomial bound on the mixing time of a Markov chain for sampling regular directed graphs. *The Electronic Journal of Combinatorics*, 18(1):234, 2011.

[22] Caterina De Bacco, Daniel B Larremore, and Cristopher Moore. A physical model for efficient ranking in networks. *Science Advances*, 4(7):eaar8260, 2018.

[23] Simon DeDeo and Elizabeth A. Hobson. From equality to hierarchy. *Proceedings of the National Academy of Sciences*, 118(21):e2106186118, 2021.

[24] Kelly R Finn, Matthew J Silk, Mason A Porter, and Noa Pinter-Wollman. The use of multilayer network analysis in animal behaviour. *Animal Behaviour*, 149:7–22, 2019.

[25] Rainer Hegselmann, Ulrich Krause, et al. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3):1–33, 2002.

[26] Morris H DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.

[27] Heather Z. Brooks and Mason A. Porter. A model for the influence of media on the ideology of content in online social networks. *Physical Review Research*, 2(2):023041, 2020.

[28] Federico Battiston, Giulia Cencetti, Iacopo Iacopini, Vito Latora, Maxime Lucas, Alice Patania, Jean-Gabriel Young, and Giovanni Petri. Networks beyond pairwise interactions: Structure and dynamics. *Physics Reports*, 874:1–92, 2020.

*equal first authors.