Proceedings of Symposia in
Applied Mathematics

Volume 80

Mathematical and Computational Methods for Complex Social Systems

AMS Short Course
Mathematical and Computational Methods
for Complex Social Systems
January 3–5, 2021
Virtual

Heather Z. Brooks
Michelle Feng
Mason A. Porter
Alexandria Volkening
Editors

American Mathematical Society

Mathematical and Computational Methods for Complex Social Systems

Proceedings of Symposia in

APPLIED MATHEMATICS

Volume 80

Mathematical and Computational Methods for Complex Social Systems

AMS Short Course
Mathematical and Computational Methods
for Complex Social Systems
January 3–5, 2021
Virtual

Heather Z. Brooks
Michelle Feng
Mason A. Porter
Alexandria Volkening
Editors

AMS AMERICAN MATHEMATICAL SOCIETY
Providence, Rhode Island

LECTURE NOTES PREPARED FOR THE AMERICAN MATHEMATICAL SOCIETY SHORT COURSE ON MATHEMATICAL AND COMPUTATIONAL METHODS FOR COMPLEX SOCIAL SYSTEMS HELD VIRTUALLY JANUARY 3–5, 2021

The AMS Short Course Series is sponsored by the Society's Program Committee for National Meetings. The series is under the direction of the Short Course Subcommittee of the Program Committee for National Meetings.

2020 *Mathematics Subject Classification.* Primary 91F99, 91D30, 11F33, 11F37, 97M10, 97M70, 97B70, 97P80, 55N31.

Library of Congress Cataloging-in-Publication Data

Cataloging-in-Publication Data has been applied for by the AMS.
See http://www.loc.gov/publish/cip/.

Proceedings of Symposia in Applied Mathematics ISSN: 0160-7634 (print); 2324-7088 (online)
DOI: https://doi.org/10.1090/psapm/80

---

10 9 8 7 6 5 4 3 2 1 30 29 28 27 26 25

Contents

Preface

The spread of memes and misinformation on social media, political redistricting, gentrification in urban communities, pedestrian movement in crowds, and the dynamics of voters are among the many social phenomena that researchers investigate in the field of complex systems. In the study of complex social systems, there is often also societal relevance to improving our understanding of how individuals interact with each other and their environment, giving rise to collective group dynamics. The mathematical and computational study of complex social systems relies on and motivates the development of methods in many topics, including mathematical modeling, data analysis, network science, and topology and geometry. In this volume of *Proceedings of Symposia in Applied Mathematics*, which is associated with our 2021 AMS Short Course, we present a variety of articles about complex social systems. Our collection includes both (1) survey and tutorial articles that introduce complex social systems and methods to study them and (2) manuscripts with original research that highlight a variety of mathematical areas and applications.

1. Introduction

How do cells organize during organism development to form tissues and patterns? What can a fish school's response to a predator tell us about how individual fish interact? What boarding practices speed up the process of passengers entering and settling into their seats in an airplane? How do echo chambers form on social media? What are the social dynamics that give rise to gentrification? How does one reduce the spread of diseases in human populations? These are all examples of questions about so-called "complex systems" [**8**, **12**, **33**]. A complex system has multiple components—typically a large number of them—that interact with each other to produce "emergent" macroscale phenomena [**8**, **12**]. The study of complex systems seeks insight into how interactions between individual entities (i.e., microscale interactions) lead to the emergence of collective dynamics (i.e., macroscale phenomena) in populations and subpopulations of those entities.

As the questions above highlight, research in complex systems involves many disciplines, and it often benefits from cross-disciplinary collaborations. Studying complex systems can lead to valuable theoretical developments (e.g., in mathematics, physics, and related disciplines), as the puzzle of collective behavior can

inspire and challenge researchers to think about new mathematical problems [**49**]. Research on complex social systems can also have a real-world impact on society and can sometimes influence policy decisions. To give one particularly noteworthy example, mathematicians have been involved in detecting gerrymandering of congressional districts, and they have occasionally even served as expert witnesses in court cases [**4**, **17**, **18**, **25**].

The questions that arise in the study of complex systems extend beyond any one mathematical or scientific domain, and researchers who are interested in mathematics and computation in complex systems often combine approaches from several areas, including data-driven modeling, network science, dynamical systems, stochastic processes, machine learning, probability, statistics, topology, and geometry. For example, a study of the question "How do echo chambers form on social media?" may involve (1) gathering and cleaning data (e.g., friendship networks from a social-media platform); (2) applying tools from network theory or topological data analysis to analyze and describe the data; (3) building and simulating a model of opinion dynamics on the networks; (4) deriving analytically tractable mean-field models; and (5) interpreting the findings of calculations and other analyses in the context of the original social-science question. Part (3) may also necessitate the development of new methods or novel extensions of existing methods. Importantly, when studying a complex system, one seeks to understand emergent phenomena (e.g., the emergence of collective group behavior in the form of echo chambers) that arise from the interactions between many entities (e.g., individual social-media accounts) in the system. Regardless of whether one builds a network model, develops an agent-based model, analyzes a system of partial differential equations, creates new topological techniques, or uses other mathematical approaches to analyze echo-chamber formation, the process is centered around a complex system. Accordingly, the "mathematics of complex social systems," which typically has close ties with applications, encompasses the use and development of mathematical methods to provide insight into collective phenomena.

The study of complex systems is typically interdisciplinary in nature and benefits from domain expertise in biology, social science, physics, engineering, mathematics, computer science, public policy, and other areas. Research in complex systems appears in a diverse spectrum of conferences and scholarly journals. The organizations with conferences that feature complex-systems research include the Society for Industrial and Applied Mathematics (SIAM), the American Physical Society (APS), the Network Science Society (NSS), and of course the American Mathematical Society (AMS). Some of these conferences seemingly focus on a specific area of mathematics (e.g., partial differential equations). However, on closer inspection, even such focused conferences often feature plenary talks and various sessions that discuss complex social systems from their subdiscipline's perspective. For example, the last three SIAM "Snowbird" Conferences on Applications of Dynamical Systems (in 2019, 2021, and 2023) listed "social dynamics" (or "dynamics of social systems") and "data and dynamics" among their focus themes. In the last few years, there has also been increasing emphasis on the mathematics of social justice and advocacy. Such studies often raise questions that connect directly with the study of complex social systems. As examples, we highlight recent semester-long programs on *Data Science and Social Justice: Networks, Policy, and Education*

at the Institute for Computational and Experimental Research in Mathematics (ICERM)[1] and on *Algorithms, Fairness, and Equity* at the Simons Laufer Mathematical Sciences Institute (SLMath).[2] Both ICERM and SLMath (formerly MSRI) are National Science Foundation (NSF) mathematical-sciences research institutes. The activities of the recently established Institute for the Quantitative Study of Inclusion, Diversity, and Equity (QSIDE) [**39**] also involve research on complex social systems.

In this volume of *Proceedings of Symposia in Applied Mathematics*, which is an outgrowth of our associated 2021 AMS Short Course,[3] we focus our attention on the intersection of complex social systems and mathematics. We provide survey and tutorial articles to introduce readers to key ideas in the mathematics of complex social systems, and we also highlight original research in this area.

2. Summary of the articles in this volume

A key aim of our 2021 AMS Short Course, which is associated with the present volume, was to provide a starting point for participants from all areas of mathematics to become involved in interdisciplinary research on complex social systems. With this motivation, this volume combines tutorials [**10**, **21**, **38**, **49**] on a few themes in the mathematical and computational study of complex social systems with research articles [**16**, **27**] that dive deeply into specific applications and mathematical methods. The research articles illustrate how some of the ideas and methods in the tutorial articles work in practice. In this section, we briefly overview each of our six chapters and discuss their relationship to each other. We also include a few references for additional discussions of various topics. The chapters in the present volume include many more salient references, and we encourage readers to peruse each chapter's references.

This volume starts with a tutorial [**49**] and a research article [**27**] that focus on modeling. First, Volkening [**49**] introduces data-driven modeling of complex systems. She discusses general modeling principles and methods, and then she illustrates how to develop a series of models of two specific complex social systems: voter dynamics in elections [**1**, **46**] and pedestrian movement in crowds [**5**, **13**, **41**]. Opinion dynamics, voter dynamics, and crowd behaviors are macroscale phenomena that emerge from interactions between many individuals. Volkening's tutorial [**49**] overviews the mathematical-modeling process [**3**, **6**, **24**, **26**] and stresses that making choices is a major part of building models.[4] Different types of models (e.g., population-scale continuum models, agent-based models, and cellular automata) have different benefits and drawbacks. They provide complementary perspectives for studying complex social systems. It is thus common for researchers to build multiple models, as we see in this volume's second chapter [**27**]. In that chapter, Halev et al. develop and analyze models of financial wealth concentration in human societies. Their findings provide insight into some of the dynamics that may

[1] See `https://icerm.brown.edu/programs/ep-22-dssj/`.

[2] See `https://www.slmath.org/programs/353`.

[3] See [**2**, **9**, **14**, **20**, **31**, **37**, **40**, **48**] for lecture recordings and other resources from our AMS Short Course. These items include lectures on data-driven modeling and data ethics, lectures and tutorials on network science and topological data analysis, and panels on cross-disciplinary collaboration.

[4] See Alexandria Volkening's associated lecture in the AMS Short Course [**48**] and her related biological-modeling lecture [**47**].

underlie gentrification (a complex process that arises from individual decisions to invest in or move to specific areas due to economic pressure, historical factors, and many other things). Halev et al. begin by building an agent-based model that describes the dynamics of wealth and amenities in two dimensions. The macroscopic dynamics of wealth and amenities are collective phenomena that emerge from many individual decisions about investment, moving, and other things. Halev et al. also derive an analytically tractable continuum model (i.e., a macroscale model) in the form of coupled partial differential equations. Depending on the model parameters, Halev et al. [**27**] observe the formation of spatially concentrated regions with many amenities and much wealth.

Building models often involves working with data, but finding and analyzing data comes with significant challenges and concerns. The third and fourth chapters of this volume explore the questions, challenges, and potential of data in complex social systems. In the third chapter, Porter [**38**] surveys some of these challenges and suggests several best practices for data ethics from both education and research perspectives.[5] He emphasizes that not all data should be used and that it is important to take particular care with human data. As in model building, the collection, presentation, and analysis of data involves many choices. Consequently, whenever it is possible, researchers should make model code, algorithms, and data publicly available to provide transparency about their choices and any potential biases. Questions related to transparency and algorithmic bias [**29**] are also the focus of the fourth chapter of this volume. In that chapter, Dhar et al. [**16**] study fairness and uncertainty in the Pennsylvania Additive Classification Tool (PACT). The PACT is an algorithm that is used by the Department of Corrections in Pennsylvania to determine custody levels for incarcerated individuals [**15**, **28**]. Systemic unfairness has emerged through many complex factors and interacting entities, and the algorithms that are involved in this process are also the collective result of many efforts. Unfortunately, the full PACT algorithm—despite its key role in influencing the experience of incarcerated people—is not publicly available [**16**]. Dhar et al. work with a large data set from Pennsylvania's Department of Corrections to investigate how perturbations to the input data lead to variability in the outputs of random-forest models. Their investigation thereby sheds light on the level of certainty (or uncertainty) in conclusions that are generated using the PACT.

Whether working with social data or the output of models, researchers face questions about how to quantitatively describe their data. The last two chapters [**10**, **21**] in this volume give tutorials on related, active research areas in complex social systems.[6] In the fifth chapter [**10**], Brooks introduces networks [**34**, **36**, **42**], which are a pillar of research in complex systems. Social data are often networked systems (i.e., collections of nodes and edges that encode sets of individuals and interactions between them). Examples include social-media platforms like Facebook and 𝕏 [**43**]. Brooks overviews how to represent networks mathematically

[5]See Mason Porter's associated lecture in the AMS Short Course [**37**].

[6]These tutorial chapters are associated with lectures by Heather Z. Brooks [**9**] (on networks) and Michelle Feng [**20**] (on topological data analysis) in our AMS Short Course. These lectures have accompanying software tutorials by Daryl DeFord [**14**] and Elizabeth Munch [**31**], respectively. DeFord's tutorial gives an introduction to NetworkX, which is a widely used Python package to analyze networks. Munch's tutorial provides guidance on doing topological data analysis in Python [**32**].

using graphs and matrices, introduces many basic properties of networks (e.g., degree distributions, shortest paths, and clustering coefficients), and discusses models to generate synthetic networks that incorporate various features of real-world social systems. In the final chapter [**21**], Feng introduces topological data analysis (TDA) [**11**, **19**, **30**, **35**], which can provide insight into the structure of networks and other social data [**44**]. Feng discusses how to use persistent homology (PH) [**35**] to describe the "shape" of data and describes how to apply TDA to social data to better understand its multiple-scale structure. It has also become common to combine TDA with machine learning (e.g., to identify parameter regimes that lead to different model behavior) [**7**, **45**]. TDA has yielded insights into several social systems (see, e.g., [**22**, **23**]), and we expect that it will also lead to informative insights that quantitatively link models and social data.

3. Conclusion

One of the main goals of our 2021 AMS Short Course was to invite people from across (both applied and theoretical) mathematics to become involved in research on complex social systems. We sought to provide a starting point and a warm welcome to research in this area. With this goal as motivation, the present volume combines four tutorial chapters on major themes in complex social systems (data-driven modeling [**49**], data ethics [**38**], network science [**10**], and topological data analysis [**21**]) and two research articles on specific complex social systems (gentrification [**27**] and legal-system dynamics [**16**]) that highlight different choices in methods, goals, and interfacing with data.

In concert, the survey and research articles in the present volume illustrate a small portion of the large, vibrant community of researchers who study complex social systems. Addressing the diverse questions that arise from careful consideration of complex social systems has both mathematical and societal value. It also necessitates combining a variety of perspectives, scientific disciplines, and mathematical topics. We hope that readers view this volume of articles as an invitation to build new collaborations, step outside their comfort zones, interact with people outside their home disciplines and specialties, and become involved in research on complex social systems.

Heather Z. Brooks
Michelle Feng
Mason A. Porter
Alexandria Volkening

References

[1] A. I. Abramowitz, *Forecasting the 2008 presidential election with the Time-for-Change model*, PS: Political Science & Politics **41** (2008), no. 4, 691–695.

[2] R. Abebe, S. González-Bailón, and J. H. Tien, *Collaborating across disciplines*, 2021, accessed February 1, 2024, `https://zerodivzero.com/short_course/aaac8c66007a4d23a7aa14857a3b778c/title/664a3995a4124f56816824fa270c027d`.

[3] K. M. Bliss, K. R. Fowler, and B. J. Galluzzo, *Math modeling: Getting started and getting solutions*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2014, `https://m3challenge.siam.org/resources/modeling-handbook`.

[4] S. Bangia, C. V. Graves, G. Herschlag, H. S. Kang, J. Luo, J. C. Mattingly, and R. Ravier, *Redistricting: Drawing the line*, `arXiv:1704.03360`, 2017.

[5] N. Bellomo, L. Gibelli, A. Quaini, and A. Reali, *Towards a mathematical theory of behavioral human crowds*, Mathematical Models and Methods in Applied Sciences **32** (2022), no. 2, 321–358. MR4396158

[6] K. M. Bliss, K. F. Kavanagh, B. J. Galluzzo, and R. Levy, *Math modeling: Computing and communicating*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2018, `https://m3challenge.siam.org/resources/modeling-handbook`.

[7] D. Bhaskar, A. Manhart, J. Milzman, J. T. Nardini, K. M. Storey, C. M. Topaz, and L. Ziegelmeier, *Analyzing collective motion with machine learning and topology*, Chaos: An Interdisciplinary Journal of Nonlinear Science **29** (2019), no. 12, 123125. MR4043359

[8] D. Brockmann, *Complexity exporables*, accessed June 17, 2022, `https://www.complexity-explorables.org`.

[9] H. Z. Brooks, *Networks in social systems*, 2021, accessed September 9, 2022, `https://zerodivzero.com/short_course/aaac8c66007a4d23a7aa14857a3b778c/title/5dd029b5e02146d1926c17d5184d8b63`.

[10] H. Z. Brooks, *A tutorial on networks of social systems: A mathematical modeling perspective*, Mathematical and computational methods for complex social systems, Proceedings of Symposia in Applied Mathematics, vol. 80, American Mathematical Society, Providence, RI, USA, 2025, pp. 115–139.

[11] G. Carlsson, *Topological methods for data modelling*, Nature Reviews Physics **2** (2020), 697–708.

[12] M. De Domenico, D. Brockmann, C. Camargo, C. Gershenson, D. Goldsmith, S. Jeschonnek, L. Kay, S. Nichele, J. R. Nicolás, T. Schmickl, M. Stella, J. Brandoff, A. J. Martínez Salinas, and H. Sayama, *Complexity explained*, 2019, accessed October 15, 2022, `https://complexityexplained.github.io`.

[13] D. C. Duives, W. Daamen, and S. P. Hoogendoorn, *State-of-the-art crowd motion simulation models*, Transportation Research Part C: Emerging Technologies **37** (2013), 193–209.

[14] D. R. DeFord, *Python tutorial on networks*, 2021, accessed September 9, 2022, `https://zerodivzero.com/short_course/aaac8c66007a4d23a7aa14857a3b778c/title/628602c8994746e491872a9380676b62`.

[15] Department of Corrections, *Department of Corrections Procedures Manual: Reception and Classification*, Policy number 11.2.1, Harrisburg, PA, USA, June 19, 2023.

[16] S. Dhar, V. Massaro, D. Mir, and N. C. Ryan, *Uncertainty in criminal justice algorithms: Simulation studies of the Pennsylvania Additive Classification Tool*, Mathematical and computational methods for complex social systems, Proceedings of Symposia in Applied Mathematics, vol. 80, American Mathematical Society, Providence, RI, USA, 2025, pp. 89–113.

[17] M. Duchin, *Outlier analysis for Pennsylvania congressional redistricting*, 2018, accessed February 2, 2024, `https://mggg.org/uploads/md-report.pdf`.

[18] M. Duchin and O. Walch (eds.), *Political geometry: Rethinking redistricting in the US with math, law, and everything in between*, Birkhäuser, Cham, Switzerland, 2022. MR4428558

[19] H. Edelsbrunner and J. Harer, *Persistent homology—A survey*, Surveys on discrete and computational geometry (J. E. Goodman, J. Pach, and R. Pollack, eds.), Contemporary Mathematics, vol. 453, American Mathematical Society, Providence, RI, USA, 2008, pp. 257–282. MR2405684

[20] M. Feng, *Topological techniques*, 2021, accessed March 14, 2023, `https://zerodivzero.com/short_course/aaac8c66007a4d23a7aa14857a3b778c/title/49b44349232746f7b64923fd4a6a2380`.

[21] M. Feng, *Interpreting topology in the context of social science*, Mathematical and computational methods for complex social systems, Proceedings of Symposia in Applied Mathematics, vol. 80, American Mathematical Society, Providence, RI, USA, 2025, pp. 141–163.

[22] M. Feng, A. Hickok, and M. A. Porter, *Topological data analysis of spatial systems*, Higher-order systems: Understanding complex systems (F. Battiston and G. Petri, eds.), Springer, Cham, Switzerland, 2022, pp. 389–399.

[23] M. Feng and M. A. Porter, *Persistent homology of geospatial data: A case study with voting*, SIAM Review **63** (2021), no. 1, 67–99. MR4209654

[24] *GAIMME: Guidelines for assessment & instruction in mathematical modeling education,* Second edition, S. Garfunkel and M. Montgomery (eds.), Consortium for Mathematics and its Applications (COMAP) and Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, USA, 2019, https://m3challenge.siam.org/resources/teaching-modeling.

[25] S. Hershberger, *Courts, commissions, and consultations: How mathematicians are working to end gerrymandering*, Notices of the American Mathematical Society **69** (2022), no. 4, 616–623. MR4398072

[26] J. Humpherys, R. Levy, and T. Witelski, *Directions for graduate and undergraduate modeling courses*, minitutorial, 2016 SIAM Annual Meeting, accessed October 13, 2022, https://www.pathlms.com/siam/courses/3028/sections/4132.

[27] A. Halev, K. Patel, N. Rodríguez, M. Tewari, and L. Wong, *A model for wealth concentration: From a discrete system to a PDE*, Mathematical and computational methods for complex social systems, Proceedings of Symposia in Applied Mathematics, vol. 80, American Mathematical Society, Providence, RI, USA, 2025, pp. 41–64.

[28] V. A. Massaro, S. Dhar, D. Mir, and N. C. Ryan, *Carceral algorithms and the history of control: An analysis of the Pennsylvania Additive Classification Tool*, Big Data & Society **9** (2022), no. 1, https://doi.org/10.1177/20539517221094002.

[29] S. Mitchell, E. Potash, S. Barocas, A. D'Amour, and K. Lum, *Algorithmic fairness: Choices, assumptions, and definitions*, Annual Reviews of Statistics and its Applications **8** (2021), 141–163. MR4243544

[30] E. Munch, *A user's guide to topological data analysis*, Journal of Learning Analytics **4** (2017), no. 2, 47–61.

[31] E. Munch, *Python tutorial on topological data analysis*, 2021, accessed March 14, 2023, https://zerodivzero.com/short_course/aaac8c66007a4d23a7aa14857a3b778c/title/cffb85f269864df08b61382c00c77c2d.

[32] E. Munch, *TDA-Python-Workshop-JMM21*, 2021, https://github.com/lizliz/TDA-Python-Workshop-JMM21.

[33] M. E. J. Newman, *Complex systems: A survey*, American Journal of Physics **79** (2011), no. 8, 800–810.

[34] M. Newman, *Networks*, Oxford University Press, Oxford, UK, 2018. Second edition of [MR2676073]. MR3838417

[35] N. Otter, M. A. Porter, U. Tillmann, P. Grindrod, and H. A. Harrington, *A roadmap for the computation of persistent homology*, European Physical Journal – Data Science **6** (2017), 17.

[36] M. A. Porter and J. P. Gleeson, *Dynamical systems on networks: A tutorial*, Frontiers in Applied Dynamical Systems: Reviews and Tutorials, vol. 4, Springer, Cham, Switzerland, 2016. MR3468887

[37] M. A. Porter, *Data ethics*, 2021, accessed February 1, 2024, https://zerodivzero.com/short_course/aaac8c66007a4d23a7aa14857a3b778c/title/181b207c7d2941278be4641ea5fe0e21.

[38] M. A. Porter, *A non-expert's introduction to data ethics for mathematicians*, Mathematical and computational methods for complex social systems, Proceedings of Symposia in Applied Mathematics, vol. 80, American Mathematical Society, Providence, RI, USA, 2025, pp. 65–88.

[39] *Institute for the Quantitative Study of Inclusion, Diversity, and Equity, Inc. (QSIDE)*, accessed March 15, 2025, https://qsideinstitute.org.

[40] N. Rodríguez, S. Scott, C. Topaz, and J. N. Victor, *Collaborating across disciplines*, 2021, accessed February 1, 2024, https://zerodivzero.com/short_course/aaac8c66007a4d23a7aa14857a3b778c/title/58292d64445d4290a17b716436155316.

[41] A. Schadschneider, M. Chraibi, A. Seyfried, A. Tordeux, and J. Zhang, *Pedestrian dynamics: From empirical results to modeling*, Crowd dynamics, volume 1: Theory, models, and safety problems (L. Gibelli and N. Bellomo, eds.), Birkhäuser, Cham, Switzerland, 2018, pp. 63–102.

[42] S. H. Strogatz, *Exploring complex networks*, Nature **410** (2001), no. 6825, 268–276.

[43] J. H. Tien, M. C. Eisenberg, S. T. Cherng, and M. A. Porter, *Online reactions to the 2017 'Unite the Right' rally in Charlottesville: Measuring polarization in Twitter networks using media followership*, Applied Network Science **5** (2020), 10.

[44] D. Taylor, F. Klimm, H. A. Harrington, M. Kramár, K. Mischaikow, M. A. Porter, and P. J. Mucha, *Topological data analysis of contagion maps for examining spreading processes on networks*, Nature Communications **6** (2015), no. 1, 7723.

[45] M. Ulmer, L. Ziegelmeier, and C. M. Topaz, *A topological approach to selecting models of biological experiments*, PLoS ONE **14** (2019), no. 3, e0213679.

[46] A. Volkening, D. F. Linder, M. A. Porter, and G. A. Rempala, *Forecasting elections using compartmental models of infection*, SIAM Review **62** (2020), no. 4, 837–865. MR4167616

[47] A. Volkening, *Intro to building models*, 2020, accessed June 30, 2022, `https://northwestern.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=7d04a874-a292-4ff2-bd66-ac2500daeea1`.

[48] A. Volkening, *Data-driven modeling*, 2021, accessed June 30, 2022, `https://zerodivzero.com/short_course/aaac8c66007a4d23a7aa14857a3b778c/title/d56faebff3a24f77a76085c1427038d8`.

[49] A. Volkening, *A primer on data-driven modeling of complex social systems*, Mathematical and computational methods for complex social systems, Proceedings of Symposia in Applied Mathematics, vol. 80, American Mathematical Society, Providence, RI, USA, 2025, pp. 1–39.

Proceedings of Symposia in Applied Mathematics
Volume **80**, 2025
https://doi.org/10.1090/psapm/080/00726

A primer on data-driven modeling of complex social systems

Alexandria Volkening

Abstract. Traffic jams on roadways, echo chambers on social media, crowds of moving pedestrians, and opinion dynamics during elections are all complex social systems. These applications may seem disparate, but some of the questions that they motivate are similar from a mathematical perspective. Across these examples, researchers seek to uncover how individual agents—whether drivers, social-media accounts, pedestrians, or voters—are interacting. By better understanding these interactions, mathematical modelers can make predictions about the group-level features that will emerge when agents alter their behavior. In this tutorial, which is based on the lecture that I gave at the 2021 American Mathematical Society Short Course, I introduce some of the terms, methods, and choices that arise when building such data-driven models. I discuss the differences between models that are statistical or mathematical, static or dynamic, spatial or non-spatial, discrete or continuous, and phenomenological or mechanistic. For concreteness, I also describe models of two complex systems, election dynamics and pedestrian-crowd movement, in more detail. With a conceptual approach, I broadly highlight some of the challenges that arise when building and calibrating models, choosing complexity, and working with quantitative and qualitative data.

A complex system might be defined as a system
for which no single model is appropriate.

and

As Picasso said of art, a good model
"is a lie that helps us see the truth."

(Lee A. Segel and Leah Edelstein-Keshet [**SEK13**])

1. Introduction

Traffic jams on roads [**SFK+08**, **SCDM+18**, **BD11**, **JHZ+14**, **BHN+95**], pedestrian crowds [**BCD18**, **HM95**], swarming locusts [**AA15**, **BCME+20**], animal aggregations [**DAB+20**, **CKJ+02**, **PEK99**, **LLEK10**, **BCC+08**, **KTI+11**], collections of cells [**BEK20**, **Vol20b**, **GBKM20**, **GG93**], and echo chambers [**SCP+21**, **EF18**, **CDFMG+21**, **CFPSS19**] are examples of complex systems. In

2020 *Mathematics Subject Classification.* Primary 97M10, 97M70; Secondary 91C99, 34F99.
Key words and phrases. Complex social systems, complex systems, mathematical modeling, data-driven modeling, election forecasting, pedestrian movement.

each of these cases, rich, group-level dynamics emerge from the interactions of smaller components—e.g., drivers, people, locusts, animals, or cells—with one another and with their environment [**DBC+19**, **Bro22**]. The interdisciplinary field of *complex systems* [**New11**] centers on the questions that arise from these emergent dynamics. Complementing experimental approaches to complex systems, mathematicians develop methods in dynamical systems, topology, network science, numerical analysis, probability, partial differential equations, and many other areas. Here I focus on data-driven mathematical modeling, mainly for complex social systems. My goal for this tutorial chapter is to help provide a starting point for folks who are new to this area, and I reflect on some broad questions and choices that emerge when combining models and data.

Figure 1 highlights several complex social systems, ranging from traffic flow [**SFK+08**, **BD11**, **NS92**] to Brexit voter dynamics [**SHP16**]; I also recommend the supplementary material of [**SSS17**, **SFK+08**] and the websites [**Loc**, **PMS17**] for related animations. Across these applications, one interesting feature is the common challenges that they raise from a modeling perspective. For example, in each of the images in Figure 1, a researcher may want to characterize alignment. This can be physical alignment, with pedestrians, locusts, or drivers adjusting how they move in response to other individuals or obstacles in their environment. A different type of alignment is present in Figure 1(d)–(e): people are forming opinions and may be influenced to align with (or against) the beliefs of others. Another thread in complex systems is heterogeneity [**MP07**, **BBD+21**]: each person, animal, or social-media account in Figure 1 is unique. Guided by the data available, each modeler must choose how much detail to include. Should we model voters as having a binary opinion (e.g., "for Brexit" or "against Brexit") or allow opinions to live on a spectrum? Changes in behavior are also present: for example, in evacuation conditions, an emotional contagion can propagate through a crowd, changing how pedestrians act [**BDM+09**, **BHK+11**, **TFB+11**, **BRSW15**].

Higher-order interactions are widespread in complex systems: peer influence and social reinforcement from multiple friends may cause someone to change their opinion or adopt a new technology, when an isolated or pairwise interaction might not [**OT12**, **BR06**, **IPBL19**, **GBC18**, **Sch73**]. In a related vein, the presence of short- and long-range interactions in complex systems leads to rich dynamics. In Figure 1(c), drivers are interacting locally, basing their acceleration on the cars near them. The addition of autonomous vehicles allows for long-range dynamics. Stern *et al.* [**SCDM+18**] have shown that judiciously modulating the speed of one autonomous vehicle can result in the disruption of self-emergent traffic jams and improved fuel usage in some experiments. (Sometimes called "phantom traffic jams", these are jams that appear to emerge from drivers, rather than through external forces [**JHZ+14**, **SCDM+18**].)

Modeling complex social systems stems from and leads to questions that are of societal and mathematical interest. Because this research area is interdisciplinary, I suggest that it is particularly important to identify the driving questions and think about where one is aiming to make a contribution proactively. From an applied perspective, in the case of traffic flow, we might want to shed light on what driver behaviors cause jams or suggest how to use external controls—e.g., time-dependent gating at ramps—to improve traffic. Models can also provide insight into how echo

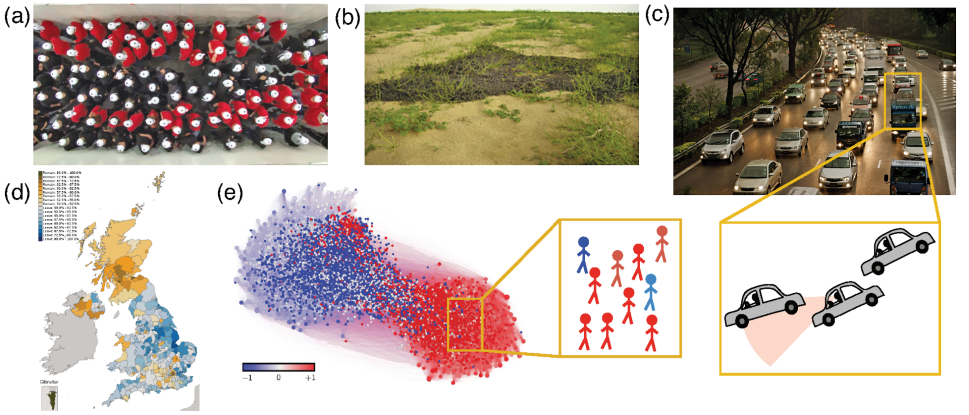


FIGURE 1. Examples of complex social systems: (a) pedestrians forming lanes in a corridor, (b) locusts organizing into bands, (c) drivers and external forces producing vehicular traffic jams, (d) opinion and voting dynamics, and (e) echo-chamber formation. In (a), we see lanes emerge from the interactions of pedestrians moving to the right (in black) and left (in red) in a corridor [**SSS17**, **ZKSS12**]; see Section 5.2. In (b), a large group of locusts organizes into a band as they move over the ground, destroying crops [**Cre16**]. In (c), drivers react to one another and external signals, producing emergent behavior. For example, in an experiment on a circular road [**SFK$^+$08**], Sugiyama *et al.* instructed originally equidistant drivers to drive normally. Despite the lack of external signals, a self-emergent traffic jam formed; see the supplementary material of [**SFK$^+$08**] for an animation. In (d) and (e), election outcomes [**SHP16**, **Mir16**] and echo chambers [**CFPSS19**] may emerge from conversations, news coverage, interactions on social media, or other factors, making opinion dynamics an example of a complex social system. Image (a) adapted (cropped) from [**SSS17**] and licensed under CC-BY 4.0 (`https://creativecommons.org/licenses/by/4.0/`); image (b) reproduced from [**Cre16**] with permission from Elsevier, Copyright (2016) Elsevier Inc.; image (c) reproduced from [**epS11**] and by epSos.de, licensed under CC-BY 2.0 (`https://commons.wikimedia.org/w/index.php?curid=27942335`); image (d) reproduced from [**Mir16**] and by Mirrorme22, Brythones, Nilfanion (English and Scottish council areas), TUBS (Welsh council areas), and Sting (Gibraltar), CC-BY-SA 3.0 (`https://commons.wikimedia.org/w/index.php?curid=47077445`); image (e) reproduced from [**CFPSS19**] and licensed under CC-BY 4.0 (`http://creativecommons.org/licenses/by/4.0/`); I added the boxes and cartoons with detail in (c) and (e).

chambers form or suggest interventions to help dissipate divisions. These goals fall into the framework of seeking to understand normal and altered agent interactions, and to predict resulting group-level features. The commonality between seemingly disparate complex systems also invites cross-fertilization between fields as researchers work toward addressing these goals in their application areas.

From a mathematical perspective, modeling complex systems can be a starting point to drive the development of new methods and inspire researchers to combine subfields in new ways. For example, there is a rich history of deriving analytically tractable continuum models from computational particle-based systems; in this case, researchers move from considering the positions of particles in space (e.g., locusts or pedestrians) to describing their density distribution in the limit as the number of particles goes to infinity. I highlight [**BV05**, **BT11**, **CCH14**, **TBL06**, **MEK99**] as a few examples. In particular, drawing on methods from probability, Oelschläger [**Oel89**] rigorously derived a system of reaction-diffusion equations to describe general agents interacting stochastically through movement, birth, death, and changes in subpopulation. As another example of complex systems driving novel methods and combinations of fields, equation-free modeling approaches [**KGH04**, **GHK$^+$03**, **KS09**, **CRS$^+$22**] offer a different perspective on the challenges associated with detailed agent-based models. This computational framework [**KGH04**, **GHK$^+$03**, **KS09**] provides macroscopic understanding without finding explicit differential equations that govern the evolution of macroscopic features of a complex system; instead, information is extracted from short simulations of a microscopic model.

Motivated by the breadth of complex social systems, my tutorial lecture "Data-driven modeling" kicked off the 2021 American Mathematical Society (AMS) Short Course on Mathematical and Computational Methods for Complex Social Systems, and this chapter is an offshoot of that talk. Many of the figures in this tutorial are related to slides in my presentation; these slides and my talk recording are available at [**Vol21**]. Following the structure of my Short Course presentation, this chapter has three main parts and takes a conceptual approach throughout. First, in Section 2, I highlight some resources, including those that I drew on when preparing my lecture. Second, in Sections 3–4, I overview mathematical modeling and discuss some of the approaches, challenges, and choices that can arise when working with data. Third, in Section 5, I discuss two case studies—election forecasting and pedestrian movement—in more depth.

Mathematical modeling is a big field, and data-driven modeling can be defined in different ways. The array of approaches that modelers can choose from is a strength, since different perspectives contribute in complementary ways to our understanding of complex systems. As a central theme, I want to acknowledge these choices and use the quotations from Segel and Edelstein-Keshet [**SEK13**] at the start of this chapter as a guide. The abundance of modeling approaches to complex systems, coupled with their multidisciplinary nature, also means that communication is more challenging; researchers may not mean the same thing when they say the same term. With this in mind, I discuss some of the things that I—from my perspective as an applied mathematician and math biologist—consider when I think about modeling complex systems. There are many, many perspectives on modeling, and this tutorial represents one, informed by the references herein.

2. Some resources on modeling

I point out some resources below, including the materials that I drew on for my Short Course lecture [**Vol21**].

2.1. Free online resources. The websites [**Bro22**, **DBC+19**] provide dynamic examples of research in complex systems and are an excellent place to gain intuition and explore this field. The Society for Industrial and Applied Mathematics (SIAM) hosts two modeling handbooks [**BKGL18**, **BFG14**]; and SIAM and the Consortium for Mathematics and its Applications provide guidelines on teaching mathematical modeling [**GAI19**]. Humpherys, Levy, and Witelski organized a very useful minitutorial discussing graduate and undergraduate education in modeling at the 2016 SIAM Annual Meeting; both their slides and a recording of their presentation are available online [**HLW16**]. Kutz and Brunton have posted a rich collection of videos [**Bru**, **Kut**] on YouTube, discussing topics including data-driven model discovery. For a demonstration of how to go from a biological paper to making simplifications to building different models, my tutorial lecture [**Vol20a**] for a broad audience may be of interest. Also geared toward a biological audience, the course "What do Your Data Say?" [**MJ**] includes a large collection of video lectures with a statistical, data-driven perspective. To see examples of research talks related to modeling complex systems, I highlight some of the BIRS workshop videos [**BIR**] (this collection from the University of British Columbia library contains a wider selection of topics than just modeling), as well as videos in the virtual SIAM Data Science minisymposium "Topological Techniques and Data-Driven Modeling in Complex Systems" organized by Brooks and Porter [**BP20b**].

2.2. Books. I found the books [**KBBP16**, **Kut13**, **BK19**] to be especially helpful as I developed my Short Course lecture, and the book [**SEK13**] by Segel and Edelstein-Keshet provides the quotations that open this chapter. Additional books related to complex systems and modeling include [**Mit09**, **THK18**, **Boc10**, **MP07**].

2.3. Publicly available data. Accessing data can be a challenge in complex-systems research. As a starting point, I highlight some publicly available data for a few specific applications here. For studies on elections and political opinions in the United States, I recommend the breadth of polling data aggregated by FiveThirtyEight [**BBG+22**]. HuffPost Pollster also curates a broad collection of public polls, with a search bar for finding data [**Huf22a**, **Huf22b**]. At a finer scale, the 2016 presidential election results in California are available at the precinct level from the *Los Angeles Times* [**SFK16**]. Ciocanel, Topaz, and other researchers through the Institute for the Quantitative Study of Inclusion, Diversity, and Equity (QSIDE) [**QSI**] developed a large-scale database (called JUSTFAIR, for Judicial System Transparency through Federal Archive Inferred Records) holding over 500,000 federal district court records [**CTS+20**]. Data from the social-media platform Twitter, as well as tutorials, are available from sources including [**AAA+21**, **Sto**, **KS20**, **Twi**].

3. Some perspectives on data and models

Because terminology can vary across fields, I survey some terms for describing models (Section 3.1) and data (Section 3.2), and then define data-driven modeling for the purposes of this chapter (Section 3.3). If you are coming to this tutorial with an applied question that you want to address, I encourage you to keep your complex

system in mind as you read—what are the parameters in your system, what data could you use to constrain your model, and at what scale do you want to make predictions or describe the system? If you are a mathematician new to modeling, what mathematical challenges does thinking from the perspective of complex systems raise? If you are from a different disciplinary background than mine, how does what we mean by "data-driven modeling" differ from and complement each other? And, if you happen to be a modeler who—like me—was introduced to modeling through research, it might be interesting to reflect on how we teach modeling.

3.1. Types of models. The term "model" means different things in different fields. In the life sciences, "model" may refer to a model organism (e.g., zebrafish, fruit flies, or worms) [**HLW16**] or a schematic hypothesizing the relationship between things. In mathematics, we may think of models that take the form of differential equations or stochastic rules, for example. Mathematical models are described using many terms, and I include a few in Figure 2. Figure 2 also highlights some of the initial choices that modelers face, often constrained by their data. Importantly, the distinctions in Figure 2 are not sharp: models often fall on a spectrum and this can depend heavily on the perspective that one takes.

Models can be described as deterministic or stochastic; stochastic models include variability. Depending on their goals and data, researchers must choose whether to build models that are static (time-independent) or dynamic. Similarly, scientists are faced with the choice of building models that are spatial or non-spatial. Do we need to understand where individual cars are located on a road, or is it sufficient to know how the number of cars evolves? Multiscale approaches (e.g., [**BBC$^+$20**]) are also possible, and I provide an example for the case of pedestrian movement in Section 5.2. We can think of models as being discrete or continuous in time or in space (e.g., so-called "on-lattice" or "off-lattice" microscopic models; see Section 5.2), but models can also be discrete in terms of types of agents; for example, do we assume voters live on an ideological spectrum or assign them a binary opinion? Whereas the choice of making a model discrete or continuous in space and time is often a choice of mathematical and computational implementation, the choice of modeling agents as having discrete or continuous features can be particularly meaningful from the perspective of the application. Understanding how choices of implementation impact model predictions is an important area of research (e.g., [**KBF17**]), as is uncovering how different modeling approaches—such as microscopic and macroscopic (see Section 5.2)—are related (e.g., [**BT11**, **CP21**, **BV05**]).

Some researchers distinguish between mathematical and statistical models, and others see statistical models as a type of mathematical model. A related categorization is phenomenological or mechanistic. These are difficult distinctions, and, in my opinion, scientists use the terms "phenomenological" and "mechanistic" in different ways. Mechanistic models of complex systems get at the mechanism underlying agent behavior. For example, the drivers in Figure 1(c) want to avoid running into one another, and we could model this by specifying repulsive forces between cars. This model can be seen as phenomenological since it describes the effect (e.g., drivers avoid one another) without getting at the mechanism of how the repulsion occurs. If we modeled the physics of the vehicles, the vision cone of individual drivers, and each driver's internal decision process, this would be more

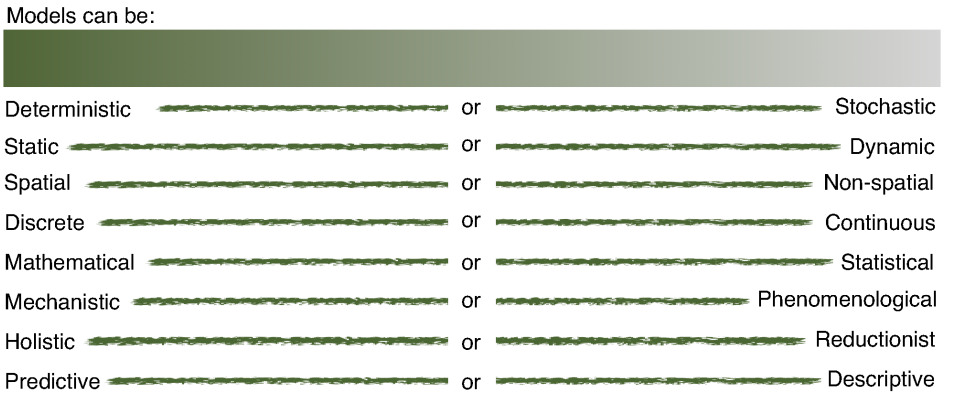


FIGURE 2. Example modeling perspectives [**Vol21**]. Many models fall somewhere in the middle of each of these scales. For example, a model may have both stochastic components (e.g., stochastic rules for when new pedestrians enter the corridor in Figure 1(a)) and deterministic components (e.g., differential equations for pedestrian movement). Models can be discrete or continuous in many ways: they can be discrete in terms of types of opinions (e.g., Republican or Democratic voting opinions in the United States), physical space, or time, for example. The distinction between phenomenological and mechanistic models is difficult, and folks have different opinions on what this means, as I discuss in Section 3.1. Somewhat related, models can be holistic or reductionist—for example, in systems biology, modelers draw together the intricate details of complex biological systems across scales, whereas reductionist models minimize complexity by focusing on individual pieces of the larger puzzle [**SEK13**]. Researchers also build models with different purposes in mind, including describing a complex system, summarizing data, or predicting future behavior under perturbation [**Shm10**].

mechanistic. However, what are the variables in a model of how people make decisions? This is in some sense a phenomenological model as well, raising further questions that involve neuroscience.

These considerations are related to choices of model complexity at their core. When, how, and where do we simplify to formulate our problem and model? In order to understand a complex system, do we choose to take a highly holistic approach—e.g., in the toy traffic example above, coupling models across scales for vehicle position in space, driver decisions, neural dynamics, and more—or adopt a reductionist approach, instead minimizing complexity and focusing on each piece of this puzzle individually? In the case of complex biological systems, a holistic approach is associated with systems biology [**SEK13**]. Given that complex systems have many layers of complexity, I suggest that the meanings of "mechanistic" and "phenomenological" depend on the question that we want to answer and the perspective from which we are studying an application. In my opinion, many models

are mechanistic at one scale, and phenomenological as soon as we step deeper into the complex system.

Lastly, models can serve many purposes, and identifying the "why?" for building a model impacts the other modeling perspectives that we choose to take. Our goal may be to describe the relationship between variables succinctly [**Shm10**] or extract governing features that help us understand a complex system broadly. We may want to build a predictive model to provide forecasts of future events (e.g., an upcoming election). A model can also be predictive about current phenomena that are poorly understood: for example, in a complex biological system, we might build a model that predicts how cell behaviors are altered by a genetic mutation. The genetic mutation exists, and the future event is when biologists discover the altered cell behaviors experimentally. Looking across applications, we may want to identify whether or not there are general principles that are upheld by disparate complex systems [**MB11**].

Starting form simple models, we may delve more and more deeply into the intricacies of our complex system through iteratively more complex models; alternatively, we may want to extract simpler, more universal models from complex ones. Models can also be built for the purpose of estimating the value of a critical, application-meaningful parameter from data, and we may use models to inform policy decisions in real time. Is our goal to gain qualitative intuition or generate concrete quantitative predictions? Do we need to explicitly write down rules or equations governing our system, or is it sufficient to know the outcome of a change in parameters? Is our goal to generate mathematically-rich models that we can use as the basis for the development of novel methods that will be broadly applicable? There are many (often complementary) reasons for modeling, and the purpose of a model can evolve as we delve deeper into the modeling process.

3.2. Types of data. The methods that modelers use to build predictive models that balance model and data complexity look different depending on the form of their data. However, the core concepts are the same when building and validating data-driven models if we look more closely, and, for this reason, I overview some types of data here. For example, data may be quantitative (e.g., the speed of the ith car in Figure 1(c)) or qualitative (e.g., the presence of lanes emerging from pedestrian behavior in Figure 1(a)). See Section 5.1 for an example of modeling with quantitative data, and Section 5.2 for a discussion of the challenges that qualitative data introduce to the modeling process. Textual data also emerge from many complex social systems (e.g., [**AAA$^+$21**, **MCM$^+$22**]).

Sometimes we find ourselves with so much data that we cannot open the files, and other times there is nearly no data. In the first case, the "black-box" modeling approaches that I discuss in Section 3.3 may be useful; for example, if we are working with a huge set of social-media posts, we could complete some data analysis to identify meaningful categories of accounts. It is also common to have rich measurements of some variables in our system, but lack measurements of other variables that are difficult to observe (e.g., different species in a population) [**MBSR19**]. Methods based on Takens' theorem of time-delay embedding [**Tak81**] can help address this challenge in some settings; this approach involves using time-lagged versions of the observed quantities to help fill in for the missing information [**MBSR19**, **DS11**]. If we are working with large sets of qualitative data (e.g., many images), this may motivate the development of new computational and mathematical approaches for

extracting quantitative information from our data. On the other hand, if we have nearly no data, it can be challenging to know where to start. In this case, it is a matter of making many simplifications (and being actively aware of the choices that we make in this process), so that the number of assumptions that we build into our model is balanced with the small amount of data available.

On a related note to amount, data for some complex systems describe rare events. For example, a model may be fit to measurements of average traffic flow, but how do we account for events that are relative outliers, like car accidents? In the case of election forecasting, we might judge a model as wrong if, despite giving Candidate A a 75% chance of winning and Candidate B a 25% chance of winning, Candidate B wins. The reality is that we do not have enough information to determine whether the model is good or bad. Forecasts are more meaningfully judged in aggregate across many elections, but limited polling data are available. Like models, data can also be time-independent (e.g., a social-media followership network at one snapshot in time) or dynamic (e.g., the timeline of posts from a given account) and spatial or non-spatial. The initial form of data is often messy, and in some cases a large portion of the time that researchers spend modeling complex systems is focused on cleaning [**BKGL18**], gathering, and tracking down the oddities in their data.

All forms of data can have bias and require human choices, particularly in the case of complex social systems. I point the reader to the chapter [**Por25**] by Porter and references therein in this volume for a discussion of data ethics. Importantly, just because data exist does not mean they should be used, and as the author mentions in [**Por25**], determining when to use or not use data is a critical step in research on complex social systems. Modelers need to be actively aware of the choices that they make when handling data, and of the presence of any choices made prior to the time that they gained access to the data. For example, if we are interested in understanding the online conversation about a recent event, we might start by downloading a large set of social-media posts using hashtags associated with that event. There are multiple choices wrapped up in this process, and I name a few here [**CY16**, **MPLC14**, **Tuf14**, **TECP20**]. First, we chose one of many social-media platforms, so our analysis will be specific to the groups that use that platform (e.g., Twitter) [**Tuf14**]. Second, we had to select what hashtags to search for and how we would identify posts "associated with" our recent event [**MPLC14**, **TECP20**]. Third, while the Twitter API provides a rich sample of posts, it is not fully clear how this selection is made [**MPLC14**]. All of these choices will affect the results of our model.

3.3. Perspectives on modeling with data. In their 2016 SIAM Annual Meeting minitutorial [**HLW16**], Humpherys, Levy, and Witelski discussed a useful classification of models based on "shades of model uncertainty". As I highlight in Figure 3, black-box, gray-box, and white-box models have different levels of dependency on data [**HLW16**], and their parameters mean different things. According to the classification system in [**HLW16**], black-box models are based heavily on data and can be thought of as maps between inputs and outputs; these models include regression, classification, and machine learning. For example, Tien *et al.* [**TECP20**] applied principal component analysis to Twitter data (the input) to distinguish groups of accounts (the output) based on their media followership.

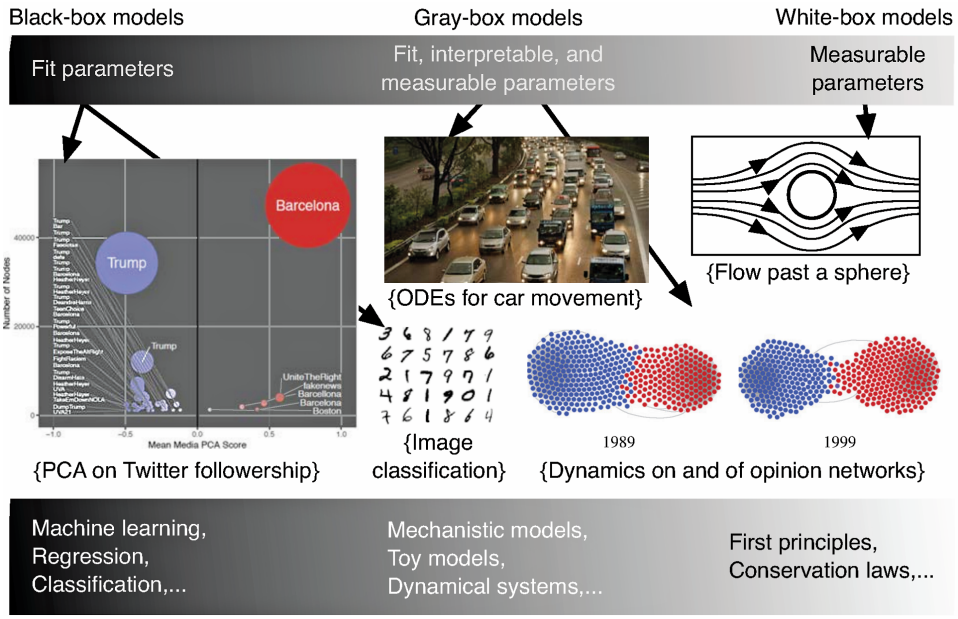


FIGURE 3. Shades of modeling with data [**HLW16**]. Black-box, gray-box, and white-box models depend on data to varying degrees and have different relationships with parameters. Black-box modeling approaches rely on data and often have internal parameters, while white-box models are largely dictated by first principles and have measurable parameters (e.g., conductivity of a material). Gray-box modeling involves visible, interpretable parameters that are fit, specified, or measured using data. All of these approaches require domain expertise. As some examples, I highlight principal component analysis (PCA, a statistical method for reducing the dimensionality of data) applied to media followership on Twitter [**TECP20**], and recognizing handwritten numbers [**LBBH98**] (black-box modeling); deterministic models of traffic flow and game-theoretic models of opinion dynamics on networks [**EF18**] (gray-box modeling); and fluid flow past a sphere (white-box modeling). Image based on a slide in the presentation [**HLW16**] by Humpherys, Levy, and Witelski. PCA–Twitter image and opinion-network images reproduced from [**TECP20**] and [**EF18**], respectively, and licensed under CC-BY 4.0 (`https://creativecommons.org/licenses/by/4.0/`); image-classification image reproduced from [**LBBH98**] with permission, Copyright (1998) IEEE; traffic (car movement) image reproduced from [**epS11**] and by epSos.de, CC-BY 2.0 (`https://commons.wikimedia.org/w/index.php?curid=27942335`).

The parameters in black-box models may be internal or hidden, and it is the model output—rather than the model structure—that is often of most interest. On the other hand, white-box models are based on first principles; these include equations from physics, such as those describing fluid dynamics or optics [**HLW16**]. The parameters in white-box models are measurable, and examples are viscosity and conductivity. Gray-box models depend on a combination of data, first principles, and domain expertise. For example, an ordinary differential equation (ODE) model for driver movement could include equations for velocity and acceleration that are based on phenomenological descriptions of repulsion and attraction between cars (i.e., domain expertise) and measurements of speeds (i.e., data).

The distinctions in Figure 3, like the distinctions in Figure 2, are not perfect. For example, equation-learning and model-selection approaches (e.g., [**MKBP17**, **BPK16**, **NBSF21**, **KBT+22**, **MBPK16**, **LNB+20**, **KAE+23**, **GGRMK98**, **RPK19**]) might be thought of as "dark gray". These approaches rely on sparse regression or machine learning to identify governing equations directly from data. It is also important to keep in mind that there are choices present and domain expertise needed across the spectrum in Figure 3. This is especially true when working with data from complex social systems, since even the data that are selected for training black-box models rely on a modeler's choice to use those data [**Por25**]. For the purposes of this tutorial, I thus think of data-driven modeling as being mathematical modeling that is driven by data, motivated by a given question, and combined with domain expertise. This encompasses developing predictive, mechanistic models based on data; equation learning and model selection;[1] machine learning, regression, or classification to understand data; and using models to raise questions and drive further data collection. Both black-box and gray-box models fit this description, but I predominantly focus on gray-box models in this survey, though again I stress that the distinctions are not sharp.

4. Challenges, choices, and creativity in data-driven modeling

Data-driven modeling involves creativity and choices, informed by the modeler's driving question, data, and domain expertise. In Section 4.1, I provide an example modeling process and highlight some of the places where modelers make choices. In Section 4.2, I then discuss challenges related to data and model calibration. See Sections 5.1 and 5.2 for illustrations of these topics for two specific applications. I take a conceptual approach throughout.

4.1. Building data-driven mathematical models. As an example data-driven, gray-box modeling process, we might follow the steps below [**GAI19**, **BFG14**, **BKGL18**]:

(1) formulate our broad motivation and specific goals
   - get to know the application area or talk to domain experts
   - search for data (qualitative or quantitative) and prior work
   - identify hypotheses to be tested or proposed and questions to be "answered" or raised

[1] Equation learning and model selection—sometimes referred to as "data-driven modeling"—are outside the scope of this survey. See, for example, [**MKBP17**, **BPK16**, **NBSF21**, **KBT+22**, **MBPK16**, **LNB+20**, **KAE+23**] for more discussion of these topics.

(2) come up with a plan for building and evaluating our model
    - determine baseline assumptions and simplify where possible
    - identify our variables, parameter names, timescales, and units
    - specify the values of measurable parameters and determine what parameters need to be fit
    - handle formatting, cleaning, and quantifying our data as needed
    - break our data into sets for fitting parameters, testing, and predicting
(3) simulate, analyze, and use our model
    - identify remaining parameter values using data for fitting
    - validate our model on test data
    - perform a sensitivity analysis or bifurcation analysis if possible
    - use our model to gain intuition, raise questions, and make predictions
    - communicate results to an interdisciplinary audience
(4) iterate to improve

These steps are not necessarily linear and data-driven modeling is iterative, as I note in step (4) [**GAI19**, **BFG14**, **BKGL18**]. The starting point may be data, domain expertise, or questions, and step (1) involves research to begin filling in gaps in our knowledge of these three areas and to formalize our goals. I often review literature in step (1) with step (2) in mind, tagging papers with quantitative data that I can use later for parameter fitting and noting studies that show alternative experimental conditions that could be used for model testing. Steps (2) and (3) then treat complementary parts of model building.

In step (2), we select our overall approach and the variables, parameters, group dynamics, and agent behaviors in which we are most interested. This means making choices related to the concepts in Figures 2 and 3: for example, if we are studying traffic flow on a stretch of roadway, will we track the number $N(t)$ of cars on the road in time, or the position $\mathbf{x}_i(t)$ and velocity $\mathbf{v}_i(t)$ of each vehicle i in time? If we are accounting for driver differences, will we assume that each driver's phenomenological "level of cautiousness" is time-dependent or static? It is important to make these choices in a way that accounts for the complexity of the problem and our data, so step (2) involves making a plan for how we will use data to develop (or train, or fit) our model and later test (or validate) our model, as I discuss in Section 4.2. At the end of step (2), our model is written down (e.g., as a system of differential equations on paper or as a set of stochastic rules in code).

In step (3), we turn to filling in parameter names with parameter values, setting initial conditions, and determining our boundary conditions, as needed. Step (3) involves validating our model to test its predictive value and performing various analyses to check how sensitive our model is to uncertainty in parameter values, initial conditions, boundary conditions [**Woo22**], or data. Depending on the form of our model, we may be able to perform a bifurcation analysis to understand how changes in parameters influence our results. We may also brainstorm alternative ways of judging model output and comparing this with data, since how we choose to describe model output can impact how we interpret our results. At the end of step (3), we use our model to gain understanding and, if possible, suggest new experiments, resulting in model-driven data collection as we regroup and iterate to improve in step (4).

More broadly, step (1) is where we realize that a model can help us accomplish our goals, step (2) is the place where we build the model structure, and step (3) is where we test and prod this structure. Data enter the picture in step (1) as motivation. In steps (2) and (3), we work closely with data to build, test, and use our model in a way that balances model and data complexity to accomplish our goals. In the remainder of this tutorial, I focus primarily on the later parts of step (2) and broadly discuss the early parts of step (3). To learn more about some of the analyses and computational approaches possible in step (3), I suggest the books [**Smi14**, **Str15**, **Kut13**, **SEK13**].

4.2. Balancing model and data complexity. While data-driven models take many forms and scientists use a range of methods to understand them, the overarching theme of balancing model and data complexity is present throughout. Depending on our goals and data, what modeling approach do we choose? How do we build a data-driven, *data-appropriate* model? If we have access to a wealth of domain expertise and a rich set of data, it may make sense to build a complex model, since, in this case, the majority of the model will be purely descriptive, framing known agent interactions in a mathematical way. The new hypothesis that we are testing, along with its few parameters, enters the picture as our assumption. On the other hand, if we are leading the way to model a poorly understood complex system, our model needs to be very simple, again so that the assumptions and hypotheses that we introduce match the amount of data available.

In either case, it is helpful to break our data up into sets for model training (or development) and testing. Training/development data are the data that we use to build our model, specifying parameter values as well as the form of model rules and terms as appropriate. After this, we take a step back and test whether or not our model behaves well on the data that we withheld—our testing data. If the model does well on the testing set, we can use our model to predict future dynamics or shed light on poorly understood behaviors. If the model does not do well on the testing set, we need to return to model development. As a guiding principle, the more parameters and assumptions that we build into a model, the more that it needs to be able to reproduce in order to have predictive value.

I find it useful to keep in mind the old adage about being able to fit an elephant with enough parameters (and make it wag its tail, given one more parameter knob to turn) [**SEK13**]. (See [**BS03**] for some comments on the attribution of this idea and the corresponding quotation.) Depending on who you ask, what constitutes "many parameters" in a model differs significantly. In the case of biological applications, which often necessitate models with nonlinear relationships, Segel and Keshet [**SEK13**] suggest fitting to several different experiments as a means of constraining parameters and helping confront the fitting-to-an-elephant concern. Colloquially, we might say it is easy to fit to an elephant with enough parameters, but it is much harder to take that same model and use it to produce a giraffe, a tiger, and a fish by simply tuning the existing parameters in a way that respects what is known about the relationships between those organisms. At the same time, it is important to remember that no model is a proof on its own, even if it shows wide agreement with different types of data.

Figure 4(c)–(d) highlights two concepts that are related to balancing model and data complexity: underfitting and overfitting. For illustrative purposes, I consider the example of population growth of some organism in time, given some (imperfect)

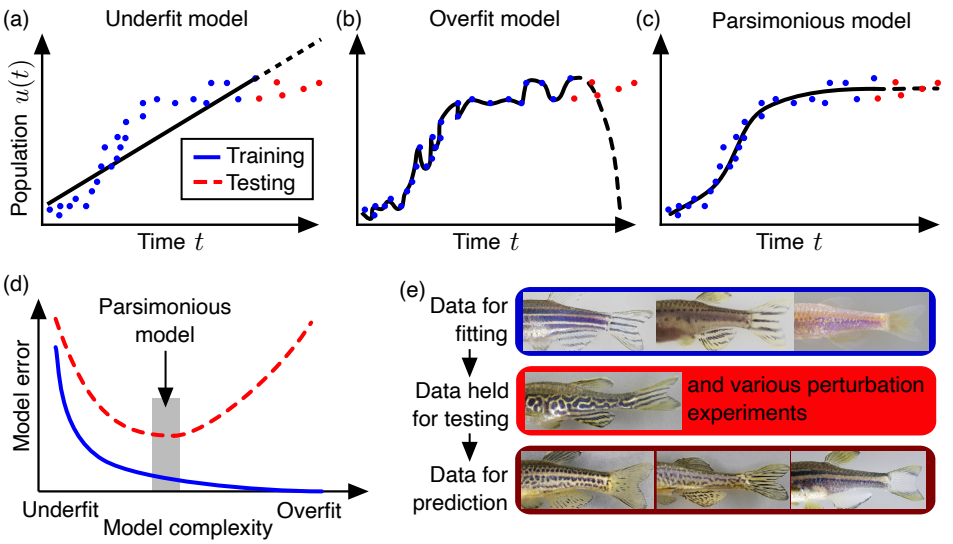


FIGURE 4. Balancing model and data complexity. (a) Underfit models miss meaningful features in data, (b) overfit models include too many assumptions, and (c) parsimonious models balance model and data complexity. In (a)–(c), blue points denote data that we use to develop our model and fit parameters, and red points denote our testing set. (d) Underfit models agree poorly with both our training and testing data, while overfit models represent our training data well and our testing data poorly [**BK19**, **MKBP17**]. Parsimonious models perform well on both sets of data. (e) Creating a plan for model training/development and testing is key to data-driven modeling. This involves breaking data into sets for training and testing, a process that depends on our complex system. For example, if our goal is to understand how cells interact to form patterns in fish skin, this could mean breaking our (qualitative) data into images of fish that are well understood (and involve setting parameters in a model to zero in a clear way), more images of fish that are well understood (and involve changing the values of nonzero parameters in a clear way), and images of poorly understood fish (which involve changing parameters in unknown ways) [**VS18**, **Vol17**]. The first set is used for model development, the second for testing, and the third as a place where we can make predictions [**VS18**, **Vol17**]. Image (d) is based partly on [**BK19**, **MKBP17**]; fish images in (e) adapted from [**SNV15**] with permission from Elsevier, Copyright (2015) Elsevier Ltd.

measurements of the number of organisms at discrete time points. At one extreme, I could assume a linear relationship between population size and time, fitting a line to the data. This involves few parameters, and the difference between the model and training data is high. At the other extreme, I could draw a curve that goes through every single data point [**BKGL18**]—this would mean introducing

many parameters. In terms of these models' ability to approximate population size at some new time in our testing set, neither will do well [**MKBP17**]. A better model lies somewhere in between these two extremes. What we are after is a "parsimonious" model [**MKBP17**, **BK19**]: a model that is supported by our data and no more complex than it needs to be.

Building predictive, data-appropriate models that avoid overfitting and have strong predictive value looks different based on the problem and relies on domain expertise. (See e.g., [**Smi14**, **BK19**] for a more detailed discussion of methods—I focus on broad concepts here.) If our goal is to understand social-media engagement in time, for example, we might build a gray-box model driven by some data $\{w_i\}_{i=1,\ldots,T}$, where w_i is the number of accounts on a social-media platform on day $t = i$. As one approach, we could split the data into a training set $\{w_i\}_{i=1,\ldots,\tilde{T}}$ and a testing set $\{w_i\}_{i=\tilde{T}+1,\ldots,T}$ with $\tilde{T} < T$. We could develop our model and specify its parameters using the training set and then run our model until $t = T$ to evaluate how well it does on the testing set. If our model does well in testing, we could use it to predict future social-media engagement.

When working with qualitative data, the process of balancing model and data complexity looks different, but it is the same at its core. In Figure 4(e), I highlight the complex biological system that most of my work is on: pattern formation in zebrafish skin [**VS15**, **VS18**]. Wild-type and mutant zebrafish feature different patterns, which form through the interactions of pigment cells [**SNV15**, **PT03**]. Although there are some quantitative data (e.g., cell speeds), most take the form of images of fish. To build the model [**VS18**], we broke these qualitative data into three sets. The first set of images contains patterns that correspond to setting specific parameters to zero in a mathematical model (e.g., setting the birth rate of black cells to zero). The second set holds some fish patterns that are relatively well understood; in this case, we know simulating them means changing parameters in a clear way (e.g., slowing domain growth). The final set contains mutant patterns that are poorly understood, patterns that form due to cell interactions that are altered in unknown ways. The first set serves as a natural model development/training set, and once we identified a model that could reproduce these fish patterns, the next step was to step back and break it down, checking if there were any ways that we could simplify the model and still maintain consistency [**VS18**]. "Minimal" model in hand, we used the second set of images for testing, asking whether or not the model could reproduce data that we did not build into it. And, finally, the tested model now serves as a predictive tool to understand the fish in the third set: at this stage, we change parameters in the model with the goal of identifying altered cell interactions that may lead to mutant patterns [**Vol17**].

In order to further improve predictive value and avoid overfitting, there are a wealth of other approaches modelers can take. We can test how uncertainty in our parameters, boundary conditions [**Woo22**], or initial conditions affect our results, and we can explore whether other models lead to the same conclusions to address questions about structural uncertainty [**Llo09**]. We can set parameters in our models to zero or remove rules, checking to see if our models can be made simpler without losing agreement with the training set. We can also ask questions about whether the methods that we use to judge our model influence our results: what alternative methods for measuring agreement between model output and data can we test? The goal is to critically investigate our modeling assumptions as we build a

parsimonious—or minimally complex for the purposes of our goal and application—model based in our data. The specific approaches that we choose to use to prod our model will depend on the form of our model, the number of parameters in our system, our data, and our goals.

It is important to account for the presence of noise in our data and to think carefully about what this means in terms of uncertainty in our predictions and parameters. This depends in part on the purpose of our model (see Section 3.1). In some modeling studies, the goal is to estimate the values of the parameters (as well as our confidence in those estimates) from data. In other settings, the values of the parameters have secondary importance, and it is the model structure and predictions that are most meaningful. For example, many complex biological systems exhibit a relationship with parameters called "sloppiness". A system is sloppy when there is a large regime in parameter space where the model is fairly insensitive to change, except for a few "stiff" parameter combinations with strong influence [**GWC+07**, **DCS+08**, **MCTS13**, **BS03**]. I highlight the work of Gutenkunst *et al.* [**GWC+07**] for a nice discussion of how sloppiness can affect measurements of uncertainty in predictions and parameter values; importantly, this is dependent on how one fits the model to data (e.g., fitting all of the parameters together, or focusing on a few parameters in which we are least confident). Moreover, the presence of sloppiness in some complex systems means that just because we may have high uncertainty in individual parameters, this does not necessarily mean that our model predictions are uncertain [**GWC+07**, **DCS+08**]. Whether or not this is an issue or a benefit comes down to the goals of our specific modeling study.

5. Illustrative case studies

In the remainder of this tutorial, I turn to two case studies of complex social systems: opinion dynamics during elections (Section 5.1) and pedestrian movement in crowds (Section 5.2). These examples illustrate some of the types of models and data from Section 3 in the broader framework of the challenges and choices introduced in Section 4. I highlight the benefits and drawbacks of different modeling choices, with the quotations from Segel and Edelstein-Keshet [**SEK13**] at the beginning of this chapter as a guide.

5.1. Forecasting elections. Political opinion dynamics are a complex social system, and here I focus on the goal of forecasting elections in the United States. Election forecasting is highly interdisciplinary, drawing on probability, geometry, dynamical systems, topology, and statistics, as well as political science, history, economics, computer science, and sociology more broadly. It naturally involves communication and public science, and different forms of data (Section 5.1.1). Framed by this interdisciplinarity, I illustrate a statistical, static modeling approach to elections in Section 5.1.2 and a dynamic, mathematical model in Section 5.1.2.

Many other models and methods for incorporating data into forecasts exist beyond the scope of this survey (e.g., data-assimilation techniques [**LSZ15**]). Election forecasting raises questions at many different scales; for example, using a

compartmental model, Restrepo *et al.* [**RRH09**] investigated how polling data affect whether potential voters decide to vote, and Biondo *et al.* [**BPR18**] developed an agent-based model to better understand how surveys influence opinions. Election forecasting is related to the broader field of opinion dynamics [**CFL09**, **PG16**], which includes the formation and dynamics of echo chambers (e.g., [**SCP+21**, **EF18**, **CDFMG+21**]) and polarization (e.g., [**SMA20**, **YAKM20**]). There are many approaches to opinion formation, such as voter models [**FGSR+14**, **BdA17**, **HL75**], which describe agents randomly adopting the opinions of other agents based on update rules, and threshold models [**LYY18**], which account for "peer pressure" in interactions (e.g., each individual having a threshold for how many of their neighbors must adopt an opinion before they do).

Because elections receive attention so widely and forecasts have the potential to impact turnout, the example of election forecasting highlights a place in complex-systems research where carefully presenting the results of data-driven models is especially important. Communicating probabilistic forecasts in a tangible, interpretable way itself leads to questions, and I suggest [**GHWM20**, **FPS+21**] for further discussion about visualizing and communicating uncertainty. Election forecasting also presents interesting challenges when it comes to evaluating model success and forecast accuracy [**GHWM20**], as I mentioned in Section 3.2.

5.1.1. *Election data.* In terms of step (1) in Section 4.1, as a starting point, the data used to build election-forecasting models include historical results, approval ratings, economic indicators, information about incumbency, and polls [**HR14**, **Sil12**, **Abr08**]. Analysts often separate these data into two types: polls and "fundamental data" (or "fundamentals"). Fundamental data are the data from which voters may form their opinions and determine how they will vote [**GK93**]; for example, economic data fall into the fundamentals category. Regardless of the type, all data come with challenges: data may not go back in time as far as a we would hope or may not be as fine-scale as we would like (e.g., data at the national or state level, rather than the district level).

For forecasts that depend on historical data, one assumption is that the past and the future will behave similarly. Modeling with fundamental data allows forecasters to produce early predictions, prior to when accurate polls may be available [**HR14**, **Lin13**]. However, opinions are dynamic—both across years and within the same election year—and past elections may not be representative of how voters will behave in the future. On the other hand, it is not always clear whether shifts in polls in a given year represent real shifts in opinion or just differences in pollster methods [**GK93**, **WE02**, **Jac05**]. Moreover, polling data are often bias [**Jac05**] and adjusted in proprietary ways; for example, pollsters make decisions such as how to define "likely voters". Polling data can be spotty, with some states being polled more frequently than others [**Lin13**]. Adding another layer of complexity, pollster herding is a phenomenon in which polling organizations adjust their results when their data do not align with other polls [**Sil14**, **CR13**, **GHWM20**].

5.1.2. *Example statistical approach.* In Figure 5(a), I reproduce a plot from [**Abr08**] of net presidential approval ratings[2] in June versus the percentage of the vote that went for the incumbent president's party in November of the same year. This motivates a statistical modeling approach to forecasting U.S. elections that

[2]This is defined as approval minus disapproval (see [**Abr08**] for details), so it can be negative.

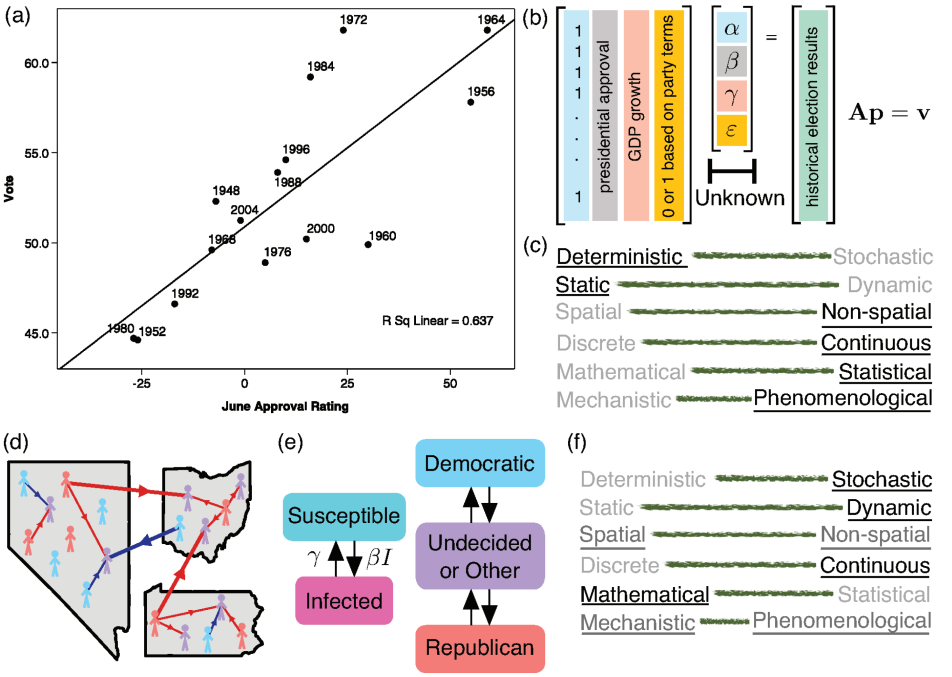


FIGURE 5. Example models for election forecasting: (a)–(c) statistical, (d) network, and (e)–(f) compartmental models. (a) To illustrate how fundamental data can be used to forecast elections, I show that the president's June approval rating and the percentage of the national vote for their party tend to be related [**Abr08**]. (b) Abramowitz's [**Abr08**] model is driven by data like those in (a). (c) This statistical model [**Abr08**] is deterministic and continuous: once the parameter values are set, the result is one prediction of the national vote for the incumbent party. (d) Alternatively, in a network model, one could investigate the interactions between undecided (purple), Republican (red), and Democratic (blue) voters, as I illustrate in this cartoon. Networks [**PG16**, **Str01**, **New18**] are the focus of a chapter [**Bro25**] in this volume; I also suggest the Short-Course lectures [**Bro21**, **DeF21**] to learn more. (e) As a third example, compartmental models [**Het00**, **DH00**, **BCC12**] group individuals and describe how folks change compartments. (f) The compartmental model [**VLPR20**] of election dynamics is stochastic and mathematical. It is spatial in that it produces state-level forecasts, and non-spatial in that it does not track the locations of individual voters. Image (a) reproduced from [**Abr08**] with permission, published by Cambridge University Press, and Copyright (2008) The American Political Science Association; images (d)–(e) adapted from [**VLPR20**].

is driven by fundamental, historical data. As an example of such an approach, I highlight some of the ideas in Abramowitz's "time-for-change" model [**Abr08**, **Abr88**]:

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ . \\ . \\ v_m \end{bmatrix} = \begin{bmatrix} 1 & a_1 & g_1 & c_1 \\ 1 & a_2 & g_2 & c_2 \\ 1 & a_3 & g_3 & c_3 \\ . & . & . & . \\ . & . & . & . \\ 1 & a_m & g_m & c_m \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \varepsilon \end{bmatrix}, \tag{5.1}$$

where v_i is the percentage of the national vote that went for the presidential candidate from the incumbent party in the ith election in the data set; m is the number of years for which data are available; a_i is a measurement of presidential approval before the ith election; g_i includes information about economic growth in the year leading up to the ith election; and c_i is a variable related to incumbency. Once the parameters α, β, γ, and ε are determined from historical data (e.g., using regression), the time-for-change model [**Abr08**, **Abr88**] can predict an election $m+1$ by computing $v_{m+1} = \alpha + \beta a_{m+1} + \gamma g_{m+1} + \varepsilon c_{m+1}$.

Equation (5.1) has the general form $\mathbf{v} = \mathbf{Ap}$, where $\mathbf{p}$ corresponds to parameters, $\mathbf{A}$ contains fundamental data, and $\mathbf{v}$ holds m past election outcomes. If we were to introduce more types of historical data, the number of parameters n would grow. With more parameters, we would expect to get a better fit between the model predictions and past election results. As Figure 4 highlights, however, this does not necessarily correspond to better predictions of future elections, since allowing n to become too large can lead to overfitting. This raises questions about model complexity. How many kinds of fundamental data should a modeler include? How many terms in the model is the "right" number of terms?

To address these questions, we need to define what a good model means and choose how to measure error. For example, consider the function [**BK19**]:

$$E(\mathbf{p}) = \underbrace{\|\mathbf{Ap} - \mathbf{v}\|_2}_{\text{least-squares term}} + \underbrace{\lambda_1 \|\mathbf{p}\|_1}_{\text{LASSO term}} + \underbrace{\lambda_2 \|\mathbf{p}\|_2}_{\text{ridge-regression term}}. \tag{5.2}$$

We can minimize $E(\hat{\mathbf{p}})$ to find the parameter values most consistent with our data:

$$\mathbf{p} = \underset{\hat{\mathbf{p}}}{\operatorname{argmin}}\ E(\hat{\mathbf{p}}).$$

When $\lambda_1 = \lambda_2 = 0$ in equation (5.2), $E(\mathbf{p})$ is the least-squares difference between the model's predictions and the election outcomes under the parameters $\mathbf{p}$. This method for measuring goodness-of-fit is sensitive to variability [**BK19**]. If $\lambda_1 > 0$ and $\lambda_2 = 0$, we instead implement LASSO regression [**Tib96**], which selects sparse models and helps prevent overfitting by forcing some parameters to zero [**BK19**]. When $\lambda_1 > 0$ and $\lambda_2 > 0$, equation (5.2) corresponds to elastic-net regularization.

Importantly, λ_1 and λ_2 provide a means of calibrating model complexity. We can choose to minimize equation (5.2) for different values of the hyper-parameters λ_1 and λ_2, resulting in different models (in the form of the parameter values $\mathbf{p}$) for each choice. Information criteria, such as Akaike information criteria (AIC) and Bayes information criteria (BIC), can come in handy to select the best model from among these alternatives [**MKBP17**, **Aka98**, **Aka74**, **Sch78**]. (In short, AIC and BIC methods score models by combining the log likelihood of the model with a term that

penalizes models with more parameters; this provides one way of comparing models with different numbers of parameters alongside one another [**MKBP17**, **BK19**, **dVHL$^+$06**].) *The Economist*'s 2020 forecasts [**eGH20**], for example, depend in part on a statistical model of the form $\mathbf{Ap} = \mathbf{v}$ with a matrix $\mathbf{A}$ that contains many types of fundamental data. To help prevent overfitting, *The Economist* [**eGH20**] team combines leave-1-out cross validation [**BK19**] and elastic-net regularization with a range of λ_1 and λ_2 values.

Broadly, leave-k-out cross validation is a means of breaking data into training and validation sets. To implement this method, one removes k samples of the training data; the removed data then become the validation set, and the remaining data are used for training [**BK19**]. For example, if $k = 1$ in the presidential election setting and the available data are for the years 2004, 2008, 2012, and 2016, one first removes one year of data (e.g., the 2012 data). The next step is determining the parameter values $\mathbf{p}_{2012}$ that result from fitting based on the data for the remaining years (2004, 2008, and 2016, in this example). Repeating this for the other years leads to four sets of parameter values. One option is to define the final parameter values $\mathbf{p}$ as the mean of these four sets of parameters. Other approaches to testing and validation include k-fold cross validation [**BK19**].

In Figure 5(b)–(c), the statistical, phenomenological approach of this section has benefits and drawbacks, like all models do. Because it is driven by fundamental data, the time-for-change model [**Abr08**, **Abr88**] is not dependent on noisy polling data; instead, it is able to generate forecasts as early as approval, economic, and incumbency data are available. Moreover, this model is simple and has few parameters. On the other hand, the model [**Abr08**, **Abr88**] in Figure 5(b) is static, and it does not add mechanistic understanding of what causes opinions to change in time during an election year.

5.1.3. *Example dynamical-systems approach.* As a more mechanistic approach, one example is the mathematical model [**VLPR20**] that my collaborators and I developed for forecasting U.S. elections. This model, driven by polling data [**Huf22a**, **Huf22b**, **Rea22**, **BBG$^+$22**], has a compartmental Susceptible–Infected–Susceptible (SIS) model at its core. Compartmental modeling is a widely used method for describing disease dynamics (e.g., [**KM27**, **KM32**, **KM33**, **Het00**, **DH00**, **BCC12**]), and it has also been applied to social contagions (e.g., [**BCAKCC06**, **BGBD$^+$18**]). The central concept is that the population of interest can be grouped into compartments.[3] In the SIS setting (Figure 5(e)), there are two compartments: susceptible and infected. Susceptible individuals become infected through interactions with infected folks (i.e., transmission), and infected individuals recover, becoming susceptible. If we track the fraction of the population that is susceptible or infected in time, the result is a gray-box model in the form of differential equations.

[3]This general structure is very flexible: for example, in Figure 6(d), I highlight one way that compartmental modeling could be used to describe pedestrian dynamics. Here the compartments are leading pedestrians moving to the right, following pedestrians moving to the right, leading pedestrians moving to the left, and following pedestrians moving to the left. If we are mainly interested in understanding how many leaders and followers are present, this approach could suffice. We might consider the transition of left-moving leaders to left-moving followers as dependent on interactions with other leaders.

In the approach [**VLPR20**], we adapt the traditional SIS compartmental model by introducing two "contagions" (Democratic and Republican voting inclinations) and replacing susceptible individual with undecided or other voters. For each state or region i, we track the fraction of undecided $S^i(t)$, Democratic $I^i_{\mathrm{D}}(t)$, and Republican I^i_{R} voters in time according to the stochastic ordinary differential equations:

$$dI^i_{\mathrm{D}}(t) = \underbrace{-\gamma^i_{\mathrm{D}} I^i_{\mathrm{D}}}_{\text{Dem. recovery}}\, dt + \underbrace{\sum_{j=1}^{M} \beta^{ij}_{\mathrm{D}} \frac{N^j}{N} S^i I^j_{\mathrm{D}}\, dt}_{\text{Dem. transmission}} + \underbrace{\sigma dW^i_{\mathrm{D}}(t)}_{\text{uncertainty}}, \tag{5.3}$$

$$dI^i_{\mathrm{R}}(t) = \underbrace{-\gamma^i_{R} I^i_{\mathrm{R}}}_{\text{Rep. recovery}}\, dt + \underbrace{\sum_{j=1}^{M} \beta^{ij}_{\mathrm{R}} \frac{N^j}{N} S^i I^j_{\mathrm{R}}\, dt}_{\text{Rep. transmission}} + \underbrace{\sigma dW^i_{\mathrm{R}}(t)}_{\text{uncertainty}}, \tag{5.4}$$

where we use that $S^i(t) = 1 - I^i_{\mathrm{D}}(t) - I^i_{\mathrm{R}}(t)$ to reduce the number of equations. Here I^i_{D}, I^i_{R}, and S^i are stochastic processes; W^i_{D} and W^i_{R} are Wiener processes; M is the number of states or regions; N^j is the number of voting-age individuals in state j; and N is the total number of voting-age individuals across our M regions. This model involves the simplifying assumption that we can bin voters as Democratic, Republican, or undecided. Bounded-confidence and related models (e.g., [**WPCG$^+$14**, **DNAW00**, **HK02**, **BP20a**]) account for opinions existing on a continuous spectrum.

The parameters in equations (5.3)–(5.4) call for special attention. There are $2 \times M$ parameters $\{\gamma^i_{\mathrm{D}}, \gamma^i_{\mathrm{R}}\}_{i=1,\ldots,M}$ that describe the rates at which committed voters become undecided. There are also $2 \times M^2$ parameters $\{\beta^{ij}_{\mathrm{D}}, \beta^{ij}_{\mathrm{R}}\}_{i,j=1,\ldots,M}$ for the rates at which Democratic (Republican) voters in state j "infect" undecided voters. To find the values of these parameters, we [**VLPR20**] relied on polling data. For the ODEs associated with equations (5.3)–(5.4) (with $\sigma = 0$), we minimized the least-squares difference between our model output under parameters $\mathbf{p}$,

$$\mathbf{X}(t_k; \mathbf{p}) = [I^1_{\mathrm{R}}(t_k; \mathbf{p}), \ldots, I^M_{\mathrm{R}}(t_k; \mathbf{p}), I^1_{\mathrm{D}}(t_k; \mathbf{p}), \ldots, I^M_{\mathrm{D}}(t_k; \mathbf{p}), S^1(t_k; \mathbf{p}), \ldots, S^M(t_k; \mathbf{p})],$$

and the averaged state- or region-level polling data,

$$\mathbf{x}(t_k) = [R^1(t_k), \ldots, R^M(t_k), D^1(t_k), \ldots, I^M_{\mathrm{D}}(t_k), S^1(t_k), \ldots, S^M(t_k)],$$

where $k = 1, \ldots, T$ with T months of polling data considered. The parameter values are different for each election year and race, depending on the associated polls.

The goal of forecasting elections provides a natural means of building and testing a model. By using only the polling data (but not the election results) for past races, we can test equations (5.3)–(5.4) by retroactively forecasting past elections [**VLPR20**]. For the statistical model in Figure 5(b), one of the challenges is selecting what types of fundamental data to include in the model, and this comes down to determining what parameters are zero or nonzero. In contrast, for the mathematical model here, it is more the format of the differential equations and the assumptions of an SIS-style model, rather than the values of the parameters, that we want to evaluate. Because the parameters in equations (5.3)–(5.4) depend only on the polls for a given election year, this model can be tested by applying it to forecast previous elections, one at a time. This step in some sense combines model training and validation together. In terms of predictions, there is also a

natural—and high-stakes—opportunity: the model can be used to forecast upcoming elections. Because polling data are inherently noisy, it is also important to estimate the uncertainty in our parameters and predictions, and better understand how sensitive our model is to changes in the parameter values.

One of the benefits of the continuous, stochastic mathematical model in equations (5.3)–(5.4) is that it includes some mechanistic hypotheses about opinion dynamics. The model [**VLPR20**] is also dynamic in time; see Figure 5(f). Once polls becomes available, equations (5.3)–(5.4) can forecast a new U.S. election with parameters that are specific to that election. However, opinion dynamics are not the same as biological disease transmission. Instead, we might think of the transmission terms in equations (5.3)–(5.4) as capturing interactions between committed voters in state j and undecided voters in state i in a phenomenological way. These interactions could be direct (e.g., via conversations between a committed voter in one state and an undecided voter in another state) or indirect (e.g., through news coverage). As another drawback, the model [**VLPR20**] has many more parameters than the statistical approach [**Abr88**, **Abr08**].

5.2. Modeling pedestrian movement. Crowds of people exhibit rich collective behavior, including lane formation and oscillating flows [**HM95**, **HBJW05**, **HJ09**, **SSS17**]. For example, as I show in Figure 6(a), pedestrians may form lanes when two groups walk in opposing directions in a narrow corridor. Like the application of election forecasting in Section 5.1, studying the dynamics of crowds touches on many fields, including engineering, sociology, psychology, physics, computer science, and mathematics [**BCG$^+$16**, **SSS17**, **BR19**, **HJ09**]. This interdisciplinarity stems from the goals that can motivate models of pedestrian movement. Researchers may be interested in designing functional buildings, testing how guidelines influence disease transmission in a crowd, developing methods to improve evacuation in emergency settings, or something else. Here I focus on the goal of understanding under what conditions lanes emerge from pedestrian interactions, and I assume accounting for the spatial organization of individuals in time is important.

For this tutorial, I use crowd movement as a venue for discussing approaches to modeling agent behavior in space, and highlighting some challenges associated with qualitative data (Section 5.2.1). Pedestrian movement provides an opportunity to illustrate a range of gray-box, spatial models, including continuum models, cellular-automaton perspectives, and agent-based approaches (Section 5.2.2). There are other data-driven approaches to crowd dynamics, and I highlight [**BR19**] for a review of statistical models. From the perspective of building simplified models (in particular, models that do not include concepts from social psychology [**SSS17**]), similar challenges and approaches arise in diverse examples of pattern formation and self-organization, including migrating cells (e.g, [**BEK20**, **Vol20b**, **GBKM20**, **HRM17**, **GG93**]), animal aggregations (e.g., [**CKJ$^+$02**, **PEK99**, **LLEK10**]), swarming locusts (e.g., [**AA15**, **BCME$^+$20**, **BT11**]), and more general agents interacting in space (e.g., [**CDM$^+$07**, **VCBJ$^+$95**, **LRC01**, **DCBC06**, **MEK99**, **TBL06**, **CMW16**]).

5.2.1. *Pedestrian data.* Data on pedestrian movement come in quantitative and qualitative forms, including measurements of velocity [**ZKSS12**], questionnaires about pedestrian experience [**SSS17**], and images of crowds [**BHK$^+$11**]. This information may stem from observations in the field or in controlled lab

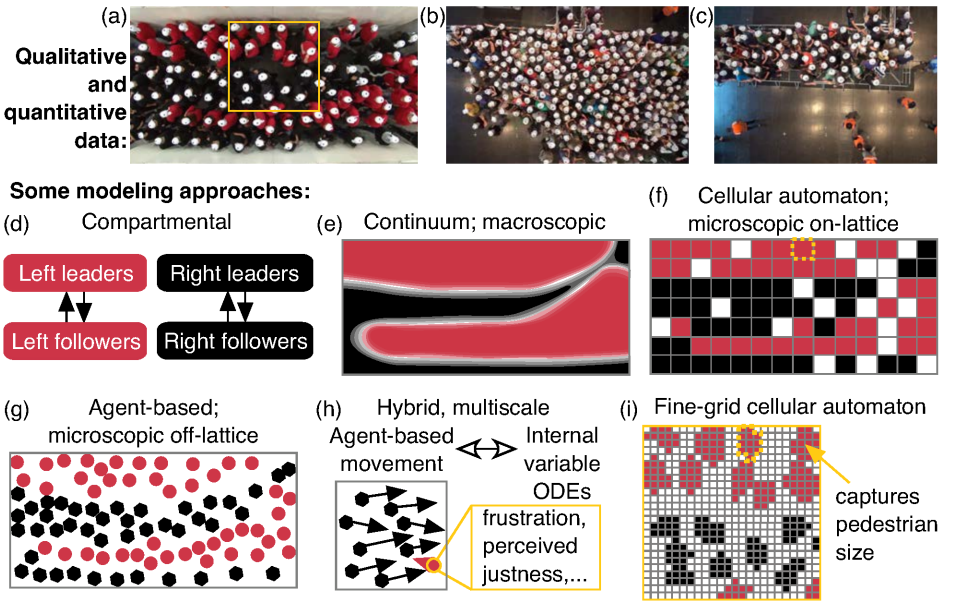


FIGURE 6. Example data and models for pedestrian movement: (a)–(c) experiments, as well as (d) compartmental, (e) continuum, (f) cellular automaton, (g) agent-based, (h) hybrid, and (i) fine-grid cellular automaton models. Experiments considering (a) bidirectional movement in corridors [**ZKSS12**], (b) movement through through an entrance [**SSS17**], and (c) movement through an entrance with spatial constraints [**SSS17**] produce quantitative and qualitative data. There are many approaches that we could take to describe pedestrians forming lanes in (a). In (a) and (d)–(i), red and black denote pedestrians moving to the left and right, respectively. (d) For example, non-spatial compartmental models could track the fraction of people who are following others or leading; see Section 5.1.3. (e) Macroscopic models generally take the form of PDEs for pedestrian density. (f) Microscopic on-lattice models consider the positions of individuals in discrete space and involve stochastic, computational rules. (g) Microscopic off-lattice models track the positions of individuals in continuous space through differential equations. (h) Hybrid, multiscale approaches come in many forms; for example, we could couple an agent-based model of pedestrian movement with a compartmental model for each pedestrian's emotions, which influence their movement. (i) Fine-grid cellular automaton models use multiple grid squares to represent each pedestrian, providing a more detailed perspective on pedestrian size. Images (a)–(c) adapted (cropped) from [**SSS17**] and licensed under CC-BY 4.0 (`https://creativecommons.org/licenses/by/4.0/`).

settings. For example, Zhang *et al.* [**ZKSS12**] performed a series of experiments in which study participants were instructed to move through corridors of different widths. As I show in Figure 6(a), participants in red were asked to move to the left through the corridor, and pedestrians wearing black were asked to move to the right. Lanes—visible as red and black stripes in Figure 6(a)—emerged from the interactions of the pedestrians in some settings [**ZKSS12**]. In addition to this qualitative data, the experiments [**ZKSS12**] produced trajectories of each participant's position, along with measurements of velocity and density.

As another example, Sieben *et al.* [**SSS17**] performed a series of experiments to better understand how pedestrians respond to different barriers as they seek to pass through an entrance. The setups [**SSS17**] in Figure 6(b)–(c) are meant to represent what might happen when people are entering a concert venue. After extracting the positions of the white caps worn by pedestrians, Sieben *et al.* [**SSS17**] collected trajectories of individuals. The authors [**SSS17**] also asked the study participants questions about their experience of walking through the entrance before and after watching a video of the experiment. This survey [**SSS17**] produced data on how comfortable the heterogeneous participants reported feeling and how just they felt the entrance process was, among other things.

When we view Figure 6(a), the presence of lanes of left- and right-moving pedestrians (in red and black shirts, respectively) is striking; while it is not as visible in Figure 6(c), the trajectories of pedestrian movement that Sieben *et al.* [**SSS17**] extracted from these experiments also show lanes in some cases. This highlights one of the challenges associated with spatial complex systems: many of the features in Figure 6(a)–(c) are qualitative. We may see lanes or streams of pedestrians, but how do we define these lanes objectively and quantitatively in large sets of images? At different timepoints in the experiment (see the videos in the supplementary material of [**SSS17**]), the lanes of red-shirt, left-moving and black-shirt, right-moving pedestrians are not as clear and do not extend across the full length of the corridor. How do we define lane width or the time when these bands start or end along the length of the corridor? The qualitative nature of data in spatial complex systems presents new challenges when fitting and testing models.

5.2.2. *Example spatial modeling approaches.* Figure 6 shows some approaches to spatial modeling of complex systems, including crowd movement, at different levels of detail. Here I focus on introducing some broad gray-box, mathematical modeling approaches that we could take to study lane formation, rather than discussing specific references. (See the reviews [**BCG$^+$16**, **SCS$^+$18**, **DDH13**, **BGQR22**] and references therein for more information about crowd dynamics.) These approaches—namely macroscopic, microscopic on-lattice, microscopic off-lattice, and hybrid (e.g., [**KHB13**]) models—are used to study a wealth of spatial dynamics. There are also many perspectives that I do not discuss, including mesoscopic (e.g., [**FTW18**, **BBK13**]) and game-theoretic (e.g., [**Dog10**, **LW11**, **BCD18**]) approaches.

Macroscopic, continuum models of pedestrian movement often take the form of partial differential equations (PDEs). As I show in Figure 6(e), this approach stems from a zoomed-out perspective: instead of tracking the locations of individuals in Figure 6(a), continuum models describe the evolution of density in time. If we make the assumption that there are two populations in our corridor example, a continuum

model would track the density $r(\mathbf{x}, t)$ of "red-shirt-wearing" and $b(\mathbf{x}, t)$ "black-shirt-wearing" pedestrians in space $\mathbf{x}$ and time t. One benefit of macroscopic models is that they are often analytically tractable, and they provide a broad perspective on overarching features that may be at work in a complex system. These models often have few parameters, and researchers can perform bifurcation analysis to understand how these parameters influence group dynamics. The drawback is that PDE approaches may simplify the complex dynamics of heterogeneous pedestrians significantly, and it can be challenging to relate the few parameters in these models to specific agent behaviors.

In contrast to macroscopic models, microscopic approaches focus on the positions or features of individuals, and two prominent frameworks are on-lattice and off-lattice models. These models provide more detailed perspectives at the scale of individual agents, which comes at the cost of more parameters. Spatial modeling is a place where vocabulary differs some between fields, particularly in the case of microscopic models. Depending on one's perspective, the microscopic models in Figure 6(f)–(g) may be described as individual- or agent-based models (IBMs or ABMs), since these models track changes in the positions of agents. The term "agent-based" also refers to more detailed models such as [**BHK+11**, **BDM+09**, **TFB+11**]. Miller and Page [**MP07**] describe agent-based models as "bottom-up" approaches, because the starting point is interactions of individuals. In interdisciplinary—or even within-discipline—conversations, I suggest asking questions to clarify what folks mean by ABMs and IBMs in their setting.

Microscopic on-lattice (cellular automaton) models consider space as a lattice, and pedestrians can either occupy or not occupy positions on a grid (e.g., [**BKSZ01**, **VCM+07**, **BA01**]); see Figure 6(f). Movement, as well as arrival and exit, takes the form of stochastic, computational rules. Notationally, we could denote whether the grid square in row i and column j at time t_k is red (i.e., containing a pedestrian moving to the left in Figure 6(a)), black (i.e., holding a right-moving pedestrian), or white (empty) by:

$$x_{i,j}(t_k) = \begin{cases} -1 & \text{if grid square is red,} \\ 0 & \text{if grid square is empty,} \\ 1 & \text{if grid square is black.} \end{cases}$$

For example, to model right-traveling pedestrians stepping to the side to avoid collisions with left-moving study participants, we might select a grid square (i, j) uniformly at random from Figure 6(a) and implement the rule:

$$\begin{aligned} &\text{if } \underbrace{x_{i,j}(t_k) = -1 \text{ and } x_{i,j+1}(t_k) = 1}_{\text{conditions for a head-on collision}} \text{ and } \underbrace{x_{i+1,j}(t_k) = 0}_{\text{space available}}, \\ &\text{then } \underbrace{x_{i,j}(t_{k+1}) = 0 \text{ and } x_{i+1,j}(t_{k+1}) = 1 \text{ with probability } p}_{\text{pedestrian may step to the side}}. \end{aligned} \tag{5.5}$$

In one time step, we could iterate through a random perturbation of all of the grid squares, implementing this and other model rules. There are many choices and parameters in rule (5.5), including the choice of probability p and the choice of neighborhoods considered (e.g., why should the pedestrian at space (i, j) only look one grid step ahead to space $(i, j + 1)$? Maybe $(i, j + 2)$ is more appropriate?).

Microscopic off-lattice models (e.g., [**HM95**, **HBJW05**]), in comparison, assume that individuals move continuously in space; see Figure 6(g). In this case, movement is modeled through coupled ordinary or stochastic differential equations, for example, of the form:

$$\frac{d\mathbf{V}_i}{dt} = \underbrace{\mathbf{g}(\mathbf{X}_i, \mathbf{V}_i)}_{\text{pedestrian } i\text{'s inherent goals}} + \underbrace{\sum_{j=1}^{N} \mathbf{f}(\mathbf{X}_i, \mathbf{X}_j, \mathbf{V}_i, \mathbf{V}_j)}_{\text{interactions between pedestrians}}, \tag{5.6}$$

$$\frac{d\mathbf{X}_i}{dt} = \mathbf{V}_i, \tag{5.7}$$

where $\mathbf{X}_i(t)$ is the position of the ith pedestrian (e.g., a point mass marking the (x, y) coordinates of the pedestrian's center of mass) and $\mathbf{V}_i(t)$ is that pedestrian's velocity. So called "social-force" models are a prominent off-lattice microscopic approach to pedestrian dynamics [**HBJW05**, **HM95**]. In both on-lattice and off-lattice models, arrival and exit of pedestrians from either side of the corridor in Figure 6(a) could take the form of stochastic rules. Computationally, we might assume that a new pedestrian enters the corridor at a randomly selected (x, y) position near the left or right edge of Figure 6(a) with probability $\alpha\Delta t$, where Δt is the time step of our simulations.

While microscopic models offer detailed perspectives on the behavior of individuals and can make experimentally testable predictions, they have many more parameters than macroscopic models do. In order to avoid overfitting and improve predictive value, it is thus important to break our data into separate sets for model development and testing. For example, we could fit the parameters in the functions in equations (5.6)–(5.7) based on measurements of pedestrian–pedestrian distances and pedestrian velocities. We could specify the rates at which pedestrians enter the corridor based on empirical data, and we could use lane width to determine any unmeasurable parameters or guide the form of model rules. To test our model, we could set aside certain experiments (e.g., experiments with wider corridors) to simulate with our final model. We could, for example, use our validated model to predict how the dynamics will change when a pushier agent is introduced or when the structure of the barriers and walls in Figure 6(a)–(c) is changed.

Adding further difficulty, microscopic models are often stochastic and not analytically tractable, and they face some of the same challenges as qualitative data: how do we define and quantitatively describe the lanes of moving pedestrians in Figure 6(f)–(g) in an automated, objective way? To help address this challenge, it can be helpful to use pair correlation functions [**DBG18**, **JC19**, **TSB+14**], which allow researchers to better understand how likely it is to see pairs of points separated by different distances in spatial data. As an alternative approach, topological techniques, especially persistent homology [**OPT+17**, **EH08**, **Car20**], have recently been combined with modeling to study complex systems, including aggregation [**UZT19**, **TZH15**]. Broadly, topological data analysis (TDA) is a means of extracting the "shape" of large data sets, and persistent homology is a widely used topological approach to identifying connected components, holes, trapped volumes, and higher-dimensional features in data across scales. Additional recent examples of TDA techniques applied to biological and social complex systems include [**BMM+19**, **BCT20**, **MVS20**, **CJDM21**, **AQO+20**, **NSF+21**] and [**FHP22**, **HJJ+22**], respectively. I suggest the chapter [**Fen25**] by Feng in this

volume, the Short-Course lectures [**Fen21**, **Mun21a**], and the associated GitHub repository [**Mun21b**] by Munch to learn more about and get started with computing persistent homology.

Depending on our goals and what our data suggest, building a hybrid, multiscale model that accounts for dynamics within pedestrians may be appropriate; see Figure 6(h). For example, in the off-lattice microscopic setting, we could introduce a variable $P_i(t)$ that tracks how frustrated each individual is based on their perceived justness of the crowd dynamics around them. We could define $P_i(t)$ by comparing the distance that pedestrian i has moved toward their goal in some time interval to the estimated distance that the individuals in a local neighborhood around i are moving. There are many other ways that we could define $P_i(t)$, and we could include feedback between $P_i(t)$ and how pushy pedestrians choose to be, influencing our ODEs for movement in an associated agent-based model.

As a last example, similar to cellular Potts models in biology [**GG93**, **HRM17**], fine-grid cellular automata represent each individual with a collection of grid squares (e.g., [**SHT10**]); see Figure 6(i). These detailed approaches are appropriate when folks are interested in the spatial extent of agents. Representing each pedestrian with $N > 1$ grid squares, instead of just one as in Figure 6(f), increases the number of parameters and the time that it will take to simulate the model. This means fine-grid cellular automata may make more sense when the goal is to describe the behavior of a few pedestrians in a detailed way; as we consider a larger crowd, agent-based or cellular automaton models become more appropriate; and, as we zoom out further into very large, densely packed crowds, macroscopic models are especially helpful.

In crowd dynamics, as for other complex social systems, there are many useful modeling approaches that we could take, and it is a matter of choosing one that is parsimonious and appropriate for our goals. And then—after that first modeling study—we iterate to improve, going back to the drawing board to build our next model with new goals and refined questions in mind. Since different types of models offer complementary perspectives and are amenable to different methods, it is also valuable to consider the relationships between models during this process. For example, as I mentioned in Section 1, many studies (e.g., [**BV05**, **BT11**, **CCH14**, **TBL06**]) link microscopic and macroscopic models of agent behavior in space to combine the benefits of analytically tractable approaches with the detail of agent-based models. On the other hand, comparatively less research has focused on elucidating how alternative microscopic frameworks—such as on-lattice and off-lattice approaches—are related, and I highlight [**NSnVBH$^+$20**, **PS12**, **OFPF$^+$17**] as examples in the case of complex biological systems.

6. Conclusions

I conclude with the best piece of advice that I have been given as a modeler: don't be afraid to be wrong. In particular, developing a model that correctly describes all of the unknown, intricate details of a complex social system would come down to sheer luck, since the space of possible models is huge. This can be discouraging. Instead, I have found it freeing to recognize that all of my models have been and will continue to be "wrong" in some sense. What matters is getting it wrong in a meaningful way. As Box's saying goes after all, "all models are wrong, but some are useful" [**BD87**, **Box76**, **Box79**]. By building a parsimonious model,

balancing our assumptions with the amount of data available, and designing a clear method for testing the model, we can make a meaningful contribution and generate new insights despite being inevitably "wrong" (or "right" in a simplified way). If the first model of a complex system does not cross disciplinary boundaries, it can lay the groundwork for a bridge that brings disciplines together in the future.

Whether our starting point is a rich data set or a nearly blank space, modeling complex systems is an iterative, creative, and interdisciplinary process. It involves being aware of the choices that we are making to simplify the problem, choosing model complexity based on our data, carefully considering the bias in the data and model, and identifying a plan for model building and validation. Through data collection, model development, prediction, communication, and generating new questions, we can push the field forward, help address societal challenges, develop mathematical approaches, and bring disciplines together in new ways.

Acknowledgments

In putting together my lecture for the 2021 AMS Short Course, I acknowledge Heather Zinn Brooks, Jonathan Desponds, Simon Freedman, Brian Hsu, Kara Maki, Niall Mangan, and Bridget Torsey for helpful examples in their earlier talks or pointers to references. Special thanks to Jeffrey Humpherys, Rachel Levy, and Thomas Witelski for their minitutorial [**HLW16**] on modeling courses at the 2016 SIAM Annual Meeting; I drew from their minitutorial for several of the concepts in Section 3. I am also grateful to the anonymous referee for thought-provoking feedback and reference ideas. Thanks to my Short Course co-organizers Heather, Mason, and Michelle for their encouragement and for being a terrific team, and to Mason for introducing me to the term "data-driven modeling of complex systems" in the first place.

References

[AA15] G. Ariel and A. Ayali, *Locust collective motion and its modeling*, PLOS Comput. Biol. **11** (2015), no. 12, e1004522.

[AAA+21] T. Alshaabi, J. L. Adams, M. V. Arnold, J. R. Minot, D. R. Dewhurst, A. J. Reagan, C. M. Danforth, and P. S. Dodds, *Storywrangler: A massive exploratorium for sociolinguistic, cultural, socioeconomic, and political timelines using Twitter*, Sci. Adv. **7** (2021), no. 29, eabe6534.

[Abr88] A. I. Abramowitz, *An improved model for predicting presidential election outcomes*, PS Political Sci. Politics **21** (1988), no. 4, 843–847.

[Abr08] A. I. Abramowitz, *Forecasting the 2008 presidential election with the time-for-change model*, PS Political Sci. Politics **41** (2008), no. 4, 691–695.

[Aka74] H. Akaike, *A new look at the statistical model identification*, IEEE Trans. Automat. Contr. **19** (1974), no. 6, 716–723.

[Aka98] H. Akaike, *Information theory and an extension of the maximum likelihood principle*, Selected Papers of Hirotugu Akaike (E. Parzen, K. Tanabe, and G. Kitagawa, eds.), Springer, New York, 1998, pp. 199–213.

[AQO+20] E. J. Amézquita, M. Y. Quigley, T. Ophelders, E. Munch, and D. H. Chitwood, *The shape of things to come: Topological data analysis and biology, from molecules to organisms*, Dev. Dyn. **249** (2020), no. 7, 816–833.

[BA01] V. J. Blue and J. L. Adler, *Cellular automata microsimulation for modeling bi-directional pedestrian walkways*, Transp. Res. B: Methodol. **35** (2001), no. 3, 293–312.

[BBC+20] N. Bellomo, R. Bingham, M. A. J. Chaplain, G. Dosi, G. Forni, D. A. Knopoff, J. Lowengrub, R. Twarock, and M. E. Virgillito, *A multiscale model of virus pandemic: heterogeneous interactive entities in a globally connected world*, Math. Models Methods Appl. Sci. **30** (2020), no. 8, 1591–1651, DOI 10.1142/S0218202520500323. MR4144366

[BBD+21] N. Bellomo, D. Burini, G. Dosi, L. Gibelli, D. Knopoff, N. Outada, P. Terna, and M. E. Virgillito, *What is life? A perspective of the mathematical kinetic theory of active particles*, Math. Models Methods Appl. Sci. **31** (2021), no. 9, 1821–1866, DOI 10.1142/S0218202521500408. MR4317555

[BBG+22] R. Best, A. Bycoffe, C. Groskopf, R. King, E. Koeze, D. Mehta, J. Mithani, M. Radcliffe, A. Wiederkehr, J. Wolfe, A. Jones-Rooy, N. Rakich, D. Shan, S. Frostenson, J. Mason, A. Mangan, and C. Yee, *FiveThirtyEight: Latest polls*, 2022, last accessed June 30, 2022. `https://projects.fivethirtyeight.com/polls/`

[BBK13] N. Bellomo, A. Bellouquid, and D. Knopoff, *From the microscale to collective crowd dynamics*, Multiscale Model. Simul. **11** (2013), no. 3, 943–963, DOI 10.1137/130904569. MR3105783

[BCAKCC06] L. M. A. Bettencourt, A. Cintrón-Arias, D. I. Kaiser, and C. Castillo-Chavéz, *The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models*, Physica A **364** (2006), 513–536.

[BCC+08] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic, *Empirical investigation of starling flocks: A benchmark study in collective animal behaviour*, Anim. Behav. **76** (2008), no. 1, 201–215.

[BCC12] F. Brauer and C. Castillo-Chavez, *Mathematical models in population biology and epidemiology*, 2nd ed., Texts in Applied Mathematics, vol. 40, Springer, New York, 2012, DOI 10.1007/978-1-4614-1686-9. MR3024808

[BCD18] R. Bailo, J. A. Carrillo, and P. Degond, *Pedestrian models based on rational behaviour*, Crowd Dynamics, Vol. 1 (L. Gibelli and N. Bellomo, eds.), Modeling and Simulation in Science, Engineering and Technology, Birkhäuser/Springer, Cham, 2018, pp. 259–292. MR3965296

[BCG+16] N. Bellomo, D. Clarke, L. Gibelli, P. Townsend, and B. J. Vreugdenhil, *Human behaviours in evacuation crowd dynamics: From modelling to "big data" toward crisis management*, Phys. Life Rev. **18** (2016), 1–21.

[BCME+20] A. J. Bernoff, M. Culshaw-Maurer, R. A. Everett, M. E. Hohn, W. C. Strickland, and J. Weinburd, *Agent-based and continuous models of hopper bands for the Australian plague locust: How resource consumption mediates pulse formation and geometry*, PLOS Comput. Biol. **16** (2020), no. 5, e1007820.

[BCT20] L. L. Bonilla, A. Carpio, and C. Trenado, *Tracking collective cell motion by topological data analysis*, PLOS Comput. Biol. **16** (2020), no. 12, e1008407.

[BD87] G. E. P. Box and N. R. Draper, *Empirical model-building and response surfaces*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons, Inc., New York, 1987. MR861118

[BD11] N. Bellomo and C. Dogbe, *On the modeling of traffic and crowds: a survey of models, speculations, and perspectives*, SIAM Rev. **53** (2011), no. 3, 409–463, DOI 10.1137/090746677. MR2834083

[BdA17] D. Braha and M. A. M. de Aguiar, *Voting contagion: Modeling and analysis of a century of U.S. presidential elections*, PLOS ONE **12** (2017), no. 5, e0177970.

[BDM+09] T. Bosse, R. Duell, Z. A. Memon, J. Treur, C. N. van der Wal, J. Otamendi, A. Bargiela, J. L. Montes, and L. M. D. Peera, *A multi-agent model for mutual absorption of emotions*, Proceedings of the 23rd European Conference on Modelling and Simulation (ECMS'09) (J. Otamendi, A. Bargiela, J. L. Montes, and L. M. Donkey Pedrera, eds.), European Council for Modeling and Simulation, 2009, pp. 212–218.

[BEK20] A. Buttenschön and L. Edelstein-Keshet, *Bridging from single to collective cell migration: A review of models and links to experiments*, PLOS Comput. Biol. **16** (2020), no. 12, e1008411.

[BFG14] K. M. Bliss, K. R. Fowler, and B. J. Galluzzo, *Math modeling: Getting started and getting solutions*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2014. https://m3challenge.siam.org/resources/modeling-handbook

[BGBD+18] L. Bonnasse-Gahot, H. Berestycki, M.-A. Depuiset, M. B. Gordon, S. Roché, N. Rodriguez, and J.-P. Nadal, *Epidemiological modelling of the 2005 French riots: A spreading wave and the role of contagion*, Sci. Rep. **8** (2018), no. 107.

[BGQR22] N. Bellomo, L. Gibelli, A. Quaini, and A. Reali, *Towards a mathematical theory of behavioral human crowds*, Math. Models Methods Appl. Sci. **32** (2022), no. 2, 321–358, DOI 10.1142/S0218202522500087. MR4396158

[BHK+11] T. Bosse, M. Hoogendoorn, M. C. A. Klein, J. Treur, and C. N. van der Wal, *Agent-based analysis of patterns in crowd behaviour involving contagion of mental states*, Modern Approaches in Applied Intelligence. IEA/AIE 2011. Lecture Notes in Computer Science (K. G. Mehrotra, C. K. Mohan, J. C. Oh, P. K. Varshney, and M. Ali, eds.), vol. 6704, Springer, Berlin Heidelberg, 2011, pp. 566–577.

[BHN+95] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama, *Dynamical model of traffic congestion and numerical simulation*, Phys. Rev. E **51** (1995), no. 2, 1035–1042.

[BIR] *The University of British Columbia: BIRS workshop lecture videos*, last accessed June 30, 2022. https://open.library.ubc.ca/cIRcle/collections/48630

[BK19] S. L. Brunton and J. N. Kutz, *Data-driven science and engineering: Machine learning, dynamical systems, and control*, Cambridge University Press, Cambridge, 2019, DOI 10.1017/9781108380690. MR3930582

[BKGL18] K. M. Bliss, B. J. Galluzzo, K. R. Kavanagh, and R. Levy, *Math modeling: Computing and communicating*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2018. https://m3challenge.siam.org/resources/modeling-handbook

[BKSZ01] C. Burstedde, K. Klauck, A. Schadschneider, and J. Zittartz, *Simulation of pedestrian dynamics using a two-dimensional cellular automaton*, Physica A **295** (2001), no. 3–4, 507–525.

[BMM+19] D. Bhaskar, A. Manhart, J. Milzman, J. T. Nardini, K. M. Storey, C. M. Topaz, and L. Ziegelmeier, *Analyzing collective motion with machine learning and topology*, Chaos **29** (2019), no. 12, 123125, DOI 10.1063/1.5125493. MR4043359

[Boc10] N. Boccara, *Modeling complex systems*, 2nd ed., Graduate Texts in Physics, Springer, New York, 2010, DOI 10.1007/978-1-4419-6562-2. MR2676220

[Box76] G. E. P. Box, *Science and statistics*, J. Amer. Statist. Assoc. **71** (1976), no. 356, 791–799. MR431440

[Box79] G. E. P. Box, *Robustness in the strategy of scientific model building*, Robustness in Statistics (R. L. Launer and G. N. Wilkinson, eds.), Academic Press, 1979, pp. 201–236.

[BP20a] H. Z. Brooks and M. A. Porter, *A model for the influence of media on the ideology of content in online social networks*, Phys. Rev. Research **2** (2020), no. 2, 023041.

[BP20b] H. Z. Brooks and M. A. Porter, *Topological data analysis and data-driven modeling in complex systems: A minisymposium in the 2020 SIAM Conference on Mathematics of Data Science*, 2020 last accessed June 30, 2022. https://www.youtube.com/playlist?list=PLnzqyg_akFM1U4KVLOE5IlhyVjAsvJ2Od

[BPK16] S. L. Brunton, J. L. Proctor, and J. N. Kutz, *Discovering governing equations from data by sparse identification of nonlinear dynamical systems*, Proc. Natl. Acad. Sci. USA **113** (2016), no. 15, 3932–3937, DOI 10.1073/pnas.1517384113. MR3494081

[BPR18] A. E. Biondo, A. Pluchino, and A. Rapisarda, *Modeling surveys effects in political competitions*, Physica A **503** (2018), 714–726, DOI 10.1016/j.physa.2018.02.211. MR3886800

[BR06] O. Bandiera and I. Rasul, *Social networks and technology adoption in northern Mozambique*, Econ. J. **116** (2006), no. 514, 869–902.

[BR19] N. W. F. Bode and E. Ronchi, *Statistical model fitting and model selection in pedestrian dynamics research*, Collective Dynamics **4** (2019), 1–32.

[Bro21] H. Z. Brooks, *Networks in social systems*, 2021, last accessed September 09, 2022. https://zerodivzero.com/short_course/aaac8c66007a4d23a7aa14857a3b778c/title/5dd029b5e02146d1926c17d5184d8b63

[Bro22] D. Brockmann, *Complexity exporables*, 2022, last accessed June 17, 2022. https://www.complexity-explorables.org

[Bro25] H. Z. Brooks, *A tutorial on networks of social systems: A mathematical modeling perspective*, Mathematical and computational methods for complex social systems, Proceedings of Symposia in Applied Mathematics, vol. 80, American Mathematical Society, Providence, RI, USA, 2025, pp. 115–139.

[BRSW15] A. L. Bertozzi, J. Rosado, M. B. Short, and L. Wang, *Contagion shocks in one dimension*, J. Stat. Phys. **158** (2015), no. 3, 647–664, DOI 10.1007/s10955-014-1019-6. MR3304346

[Bru] S. Brunton, *Data-driven dynamical systems with machine learning*, last accessed June 30, 2022. https://www.youtube.com/playlist?list=PLMrJAkhIeNNR6DzT17-MM1GHLkuYVjhyt

[BS03] K. S. Brown and J. P. Sethna, *Statistical mechanical approaches to models with many poorly known parameters*, Phys. Rev. E **68** (2003), 021904.

[BT11] A. J. Bernoff and C. M. Topaz, *A primer of swarm equilibria*, SIAM J. Appl. Dyn. Syst. **10** (2011), no. 1, 212–250, DOI 10.1137/100804504. MR2788924

[BV05] M. Bodnar and J. J. L. Velazquez, *Derivation of macroscopic equations for individual cell-based models: a formal approach*, Math. Methods Appl. Sci. **28** (2005), no. 15, 1757–1779, DOI 10.1002/mma.638. MR2166611

[Car20] G. Carlsson, *Topological methods for data modelling*, Nat. Rev. Phys. **2** (2020), 697–708.

[CCH14] J. A. Carrillo, Y.-P. Choi, and M. Hauray, *The derivation of swarming models: mean-field limit and Wasserstein distances*, Collective dynamics from bacteria to crowds (A. Muntean and F. Toschi, eds.), CISM Courses and Lect., vol. 553, Springer, Vienna, 2014, pp. 1–46, DOI 10.1007/978-3-7091-1785-9_1. MR3331178

[CDFMG+21] M. Cinelli, G. De Francisci Morales, A. Galeazzi, W. Quattrociocchi, and M. Starnini, *The echo chamber effect on social media*, Proc. Natl. Acad. Sci. USA **118** (2021), no. 9, e2023301118.

[CDM+07] Y. Chuang, M. R. D'Orsogna, D. Marthaler, A. L. Bertozzi, and L. S. Chayes, *State transitions and the continuum limit for a 2D interacting, self-propelled particle system*, Physica D **232** (2007), no. 1, 33–47, DOI 10.1016/j.physd.2007.05.007. MR2369988

[CFL09] C. Castellano, S. Fortunato, and V. Loreto, *Statistical physics of social dynamics*, Rev. Mod. Phys. **81** (2009), no. 2, 591–646.

[CFPSS19] W. Cota, S. C. Ferreira, R. Pastor-Satorras, and M. Starnini, *Quantifying echo chamber effects in information spreading over political communication networks*, EPJ Data Science **8** (2019), no. 35, 1–13.

[CJDM21] M.-V. Ciocanel, R. Juenemann, A. T. Dawes, and S. A. McKinley, *Topological data analysis approaches to uncovering the timing of ring structure onset in filamentous networks*, Bull. Math. Biol. **83** (2021), no. 3, Paper No. 21, DOI 10.1007/s11538-020-00847-3. MR4200885

[CKJ+02] I. D. Couzin, J. Krause, R. James, G. D. Ruxton, and N. R. Franks, *Collective memory and spatial sorting in animal groups*, J. Theoret. Biol. **218** (2002), no. 1, 1–11, DOI 10.1006/jtbi.2002.3065. MR2027139

[CMW16] J. A. Carrillo, S. Martin, and M.-T. Wolfram, *An improved version of the Hughes model for pedestrian flow*, Math. Models Methods Appl. Sci. **26** (2016), no. 4, 671–697, DOI 10.1142/S0218202516500147. MR3460619

[CP21] J. A. Carrillo and Y.-P. Choi, *Mean-field limits: from particle descriptions to macroscopic equations*, Arch. Ration. Mech. Anal. **241** (2021), no. 3, 1529–1573, DOI 10.1007/s00205-021-01676-x. MR4284530

[CR13] J. D. Clinton and S. Rogers, *Robo-polls: Taking cues from traditional sources?*, PS Political Sci. Politics **46** (2013), no. 2, 333–337.

[Cre16] K. Cressman, *Chapter 4.2 - desert locust*, Biological and Environmental Hazards, Risks, and Disasters (J. F. Shroder and R. Sivanpillai, eds.), Academic Press, Boston, 2016, pp. 87–105.

[CRS+22] T. Chin, J. Ruth, C. Sanford, R. Santorella, P. Carter, and B. Sandstede, *Enabling equation-free modeling via diffusion maps*, J. Dynam. Differential Equations **36** (2024), no. suppl. 1, 415–434, DOI 10.1007/s10884-021-10127-w. MR4710818

[CTS+20] M.-V. Ciocanel, C. M. Topaz, R. Santorella, S. Sen, C. M. Smith, and A. Hufstetler, *JUSTFAIR: Judicial System Transparency through Federal Archive Inferred Records*, PLOS ONE **15** (2020), no. 10, e0241381.

[CY16] P. Cihon and T. Yasseri, *A biased review of biases in Twitter studies on political collective action*, Front. Phys. **4** (2016).

[DAB+20] S. Dodson, B. Abrahms, S. J. Bograd, J. Fiechter, and E. L. Hazen, *Disentangling the biotic and abiotic drivers of emergent migratory behavior using individual-based models*, Ecol. Model **432** (2020), 109225.

[DBC+19] M. De Domenico, D. Brockmann, C. Camargo, C. Gershenson, D. Goldsmith, S. Jeschonnek, L. Kay, S. Nichele, J. R. Nicolás, T. Schmickl, M. Stella, J. Brandoff, A. J. Martínez Salinas, and H. Sayama, *Complexity explained*, 2019, last accessed October 15, 2022. `https://complexityexplained.github.io`

[DBG18] S. Dini, B. J. Binder, and J. E. F. Green, *Understanding interactions between populations: individual based modelling and quantification using pair correlation functions*, J. Theoret. Biol. **439** (2018), 50–64, DOI 10.1016/j.jtbi.2017.11.014. MR3739970

[DCBC06] M. R. D'Orsogna, Y. L. Chuang, A. L. Bertozzi, and L. S. Chayes, *Self-propelled particles with soft-core interactions: Patterns, stability, and collapse*, Phys. Rev. Lett. **96** (2006), 104302.

[DCS+08] B. C. Daniels, Y.-J. Chen, J. P. Sethna, R. N. Gutenkunst, and C. R. Myers, *Sloppiness, robustness, and evolvability in systems biology*, Curr. Opin. Biotechnol. **19** (2008), no. 4, 389–395.

[DDH13] D. C. Duives, W. Daamen, and S. P. Hoogendoorn, *State-of-the-art crowd motion simulation models*, Transp. Res. C: Emerg. Technol. **37** (2013), 193–209.

[DeF21] D. R. DeFord, *Python tutorial on networks*, 2021, last accessed September 09, 2022. `https://zerodivzero.com/short_course/aaac8c66007a4d23a7aa14857a3b778c/title/628602c8994746e491872a9380676b62`

[DH00] O. Diekmann and J. A. P. Heesterbeek, *Mathematical epidemiology of infectious diseases: Model building, analysis and interpretation*, Wiley Series in Mathematical and Computational Biology, John Wiley & Sons, Ltd., Chichester, 2000. MR1882991

[DNAW00] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch, *Mixing beliefs among interacting agents*, Adv. Complex. Syst. **3** (2000), no. 1n04, 87–98.

[Dog10] C. Dogbé, *Modeling crowd dynamics by the mean-field limit approach*, Math. Comput. Modelling **52** (2010), no. 9-10, 1506–1520, DOI 10.1016/j.mcm.2010.06.012. MR2719535

[DS11] E. R. Deyle and G. Sugihara, *Generalized theorems for nonlinear state space reconstruction*, PLOS ONE **6** (2011), no. 3, e18295.

[dVHL+06] G. de Vries, T. Hillen, M. Lewis, J. Müller, and B. Schönfisch, *Chapter 7: Estimating parameters*, A Course in Mathematical Biology, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006, pp. 175–197.

[EF18] T. Evans and F. Fu, *Opinion formation on dynamic networks: Identifying conditions for the emergence of partisan echo chambers*, Royal Soc. Open Sci. **5** (2018), 181122.

[eGH20] *The Economist*, A. Gelman, and M. Heidemanns, *Forecasting the US elections: How The Economist presidential forecast works*, 2020, last accessed September 07, 2022. `https://projects.economist.com/us-2020-forecast/president/how-this-works`

[EH08] H. Edelsbrunner and J. Harer, *Persistent homology—a survey*, Surveys on discrete and computational geometry, Contemp. Math., vol. 453, Amer. Math. Soc., Providence, RI, 2008, pp. 257–282, DOI 10.1090/conm/453/08802. MR2405684

[epS11] epSos.de, *Driving cars in a traffic jam*, 2011, accessed October 12, 2022. `https://commons.wikimedia.org/wiki/File:Driving_Cars_in_a_Traffic_Jam.jpg`

[Fen21] M. Feng, *Topological techniques*, 2021, last accessed March 14, 2023. https://zerodivzero.com/short_course/aaac8c66007a4d23a7aa14857a3b778c/title/49b44349232746f7b64923fd4a6a2380

[Fen25] M. Feng, *Interpreting topology in the context of social science*, Mathematical and computational methods for complex social systems, Proceedings of Symposia in Applied Mathematics, vol. 80, American Mathematical Society, Providence, RI, USA, 2025, pp. 141–163.

[FGSR+14] J. Fernández-Gracia, K. Suchecki, J. J. Ramasco, M. San Miguel, and V. M. Eguiluz, *Is the voter model a model for voters?*, Phys. Rev. Lett. **112** (2014), 158701.

[FHP22] M. Feng, A. Hickok, and M. A. Porter, *Topological data analysis of spatial systems*, Higher-Order Systems, Understanding Complex Systems (F. Battiston and G. Petri, eds.), Springer, Cham, 2022, pp. 389–399.

[FPS+21] S. L. Franconeri, L. M. Padilla, P. Shah, J. M. Zacks, and J. Hullman, *The science of visual data communication: What works*, Psychol. Sci. Public Interest **22** (2021), no. 3, 110–161.

[FTW18] A. Festa, A. Tosin, and M.-T. Wolfram, *Kinetic description of collision avoidance in pedestrian crowds by sidestepping*, Kinet. Relat. Models **11** (2018), no. 3, 491–520, DOI 10.3934/krm.2018022. MR3810836

[GAI19] *GAIMME: Guidelines for Assessment and Instruction in Mathematical Modeling Education, Second Edition*, S. Garfunkel and M. Montgomery (eds.), Consortium for Mathematics and its Applications (COMAP) and Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2019. https://m3challenge.siam.org/resources/teaching-modeling

[GBC18] D. Guilbeault, J. Becker, and D. Centola, *Complex contagions: A decade in review*, Complex Spreading Phenomena in Social Systems, Computational Social Sciences (S. Lehmann and Y. Y. Ahn, eds.), Springer, Cham, 2018, pp. 3–25.

[GBKM20] R. Giniūnaitė, R. E. Baker, P. M. Kulesa, and P. K. Maini, *Modelling collective cell migration: neural crest as a model paradigm*, J. Math. Biol. **80** (2020), no. 1-2, 481–504, DOI 10.1007/s00285-019-01436-2. MR4062827

[GG93] J. A. Glazier and F. Graner, *Simulation of the differential adhesion driven rearrangement of biological cells*, Phys. Rev. E **47** (1993), 2128–2154.

[GGRMK98] R. González-García, R. Rico-Martínez, and I. G. Kevrekidis, *Identification of distributed parameter systems: A neural net based approach*, Comput. Chem. Eng. **22** (1998), S965–S968.

[GHK+03] I. G. Kevrekidis, C. W. Gear, J. M. Hyman, P. G. Kevrekidis, O. Runborg, and C. Theodoropoulos, *Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level analysis*, Commun. Math. Sci. **1** (2003), no. 4, 715–762. MR2041455

[GHWM20] A. Gelman, J. Hullman, C. Wlezien, and G. E. Morris, *Information, incentives, and goals in election forecasts*, Judgm. Decis. Mak. **15** (2020), no. 5, 863–880.

[GK93] A. Gelman and G. King, *Why are American presidential election campaign polls so variable when votes are so predictable?*, Br. J. Political Sci. **23** (1993), no. 4, 409–451.

[GWC+07] R. N. Gutenkunst, J. J. Waterfall, F. P. Casey, K. S. Brown, C. R. Myers, and J. P. Sethna, *Universally sloppy parameter sensitivities in systems biology models*, PLOS Comput. Biol. **3** (2007), no. 10, e189, DOI 10.1371/journal.pcbi.0030189. MR2369325

[HBJW05] D. Helbing, L. Buzna, A. Johansson, and T. Werner, *Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions*, Transp. Sci. **39** (2005), no. 1, 1–24.

[Het00] H. W. Hethcote, *The mathematics of infectious diseases*, SIAM Rev. **42** (2000), no. 4, 599–653, DOI 10.1137/S0036144500371907. MR1814049

[HJ09] D. Helbing and A. Johansson, *Pedestrian, crowd and evacuation dynamics*, Encyclopedia of Complexity and Systems Science (R. A. Meyers, ed.), Springer, New York, 2009, pp. 6476–6495.

[HJJ+22] A. Hickok, B. Jarman, M. Johnson, J. Luo, and M. A. Porter, *Persistent homology for resource coverage: a case study of access to polling sites*, SIAM Rev. **66** (2024), no. 3, 481–500, DOI 10.1137/22M150410X. MR4783075

[HK02] R. Hegselmann and U. Krause, *Opinion dynamics and bounded confidence: Models, analysis and simulation*, J. Artif. Soc. Soc. Simul. **5** (2002), no. 3.

[HL75] R. A. Holley and T. M. Liggett, *Ergodic theorems for weakly interacting infinite systems and the voter model*, Ann. Probability **3** (1975), no. 4, 643–663, DOI 10.1214/aop/1176996306. MR402985

[HLW16] J. Humpherys, R. Levy, and T. Witelski, *Directions for graduate and undergraduate modeling courses*, 2016, last accessed October 13, 2022. `https://www.pathlms.com/siam/courses/3028/sections/4132`

[HM95] D. Helbing and P. Molnár, *Social force model for pedestrian dynamics*, Phys. Rev. E **51** (1995), 4282–4286.

[HR14] P. Hummel and D. Rothschild, *Fundamental models for forecasting elections at the state level*, Elect. Stud. **35** (2014), 123–139.

[HRM17] T. Hirashima, E. G. Rens, and R. M. H. Merks, *Cellular Potts modeling of complex multicellular behaviors in tissue morphogenesis*, Dev. Growth Differ. **59** (2017), no. 5, 329–339.

[Huf22a] *HuffPost Pollster*, 2022, last accessed October 16, 2022. `https://elections.huffingtonpost.com/pollster`

[Huf22b] *HuffPost Pollster API v2*, 2022, last accessed October 16, 2022. `https://elections.huffingtonpost.com/pollster/api/v2`

[IPBL19] I. Iacopini, G. Petri, A. Barrat, and V. Latora, *Simplicial models of social contagion*, Nat. Commun. **10** (2019), 2485.

[Jac05] S. Jackman, *Pooling the polls over an election campaign*, Aust. J. Political Sci. **40** (2005), no. 4, 499–517.

[JC19] S. T. Johnston and E. J. Crampin, *Corrected pair correlation functions for environments with obstacles*, Phys. Rev. E **99** (2019), 032124.

[JHZ+14] R. Jiang, M.-B. Hu, H. M. Zhang, Z.-Y. Gao, B. Jia, Q.-S. Wu, B. Wang, and M. Yang, *Traffic experiment reveals the nature of car-following*, PLOS ONE **9** (2014), no. 4, e94351.

[KAE+23] F. P. Kemeth, S. Alonso, B. Echebarria, T. Moldenhawer, C. Beta, and I. G. Kevrekidis, *Black and gray box learning of amplitude equations: Application to phase field systems*, Phys. Rev. E **107** (2023), 025305.

[KBBP16] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic mode decomposition: Data-driven modeling of complex systems*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2016, DOI 10.1137/1.9781611974508. MR3602007

[KBF17] J. Kursawe, R. E. Baker, and A. G. Fletcher, *Impact of implementation choices on quantitative predictions of cell-based computational models*, J. Comput. Phys. **345** (2017), 752–767, DOI 10.1016/j.jcp.2017.05.048. MR3667637

[KBT+22] F. P. Kemeth, T. Bertalan, T. Thiem, F. Dietrich, S. J. Moon, C. R. Laing, and I. G. Kevrekidis, *Learning emergent partial differential equations in a learned emergent space*, Nat. Commun. **13** (2022), 3318–3318.

[KGH04] I. G. Kevrekidis, C. W. Gear, and G. Hummer, *Equation-free: The computer-aided analysis of complex multiscale systems*, AIChE Journal **50** (2004), no. 7, 1346–1355.

[KHB13] A. Kneidl, D. Hartmann, and A. Borrmann, *A hybrid multi-scale approach for simulation of pedestrian dynamics*, Transp. Res. C: Emerg. Technol. **37** (2013), 223–237.

[KM27] W. O. Kermack and A. G. McKendrick, *A contribution to the mathematical theory of epidemics*, Proc. R. Soc. London **115** (1927), 700–721.

[KM32] W. O. Kermack and A. G. McKendrick, *Contributions to the mathematical theory of epidemics. II.—The problem of endemicity*, Proc. R. Soc. London **138** (1932), 55–83.

[KM33] W. O. Kermack and A. G. McKendrick, *Contributions to the mathematical theory of epidemics. III.—Further studies of the problem of endemicity*, Proc. R. Soc. London **141** (1933), 94–122.

[KS09] I. G. Kevrekidis and G. Samaey, *Equation-free multiscale computation: Algorithms and applications*, Annu. Rev. Phys. Chem. **60** (2009), 321–344.

[KS20] M. Kalt and D. Scott, *Twitter data curation primer: Data curation network GitHub repository*, 2020, last accessed October 16, 2022. `https://github.com/DataCurationNetwork/data-primers/blob/master/Twitter%20Data%20Curation%20Primer/twitter-data-curation-primer.md`

[KTI+11] Y. Katz, K. Tunstrom, C. C. Ioannou, C. Huepe, and I. D. Couzin, *Inferring the structure and dynamics of interactions in schooling fish*, Proc. Natl. Acad. Sci. USA **108** (2011), no. 46, 18720–18725.

[Kut] N. Kutz, *Nathan Kutz Videos*, last accessed June 30, 2022. `https://www.youtube.com/c/NathanKutzAMATH/videos`

[Kut13] J. N. Kutz, *Data-driven modeling & scientific computation: Methods for complex systems & big data*, Oxford University Press, Oxford, England, 2013.

[LBBH98] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, *Gradient-based learning applied to document recognition*, Proc. IEEE **86** (1998), no. 11, 2278–2324.

[Lin13] D. A. Linzer, *Dynamic Bayesian forecasting of presidential elections in the states*, J. Amer. Statist. Assoc. **108** (2013), no. 501, 124–134, DOI 10.1080/01621459.2012.737735. MR3174607

[LLEK10] R. Lukeman, Y.-X. Li, and L. Edelstein-Keshet, *Inferring individual rules from collective behavior*, Proc. Natl. Acad. Sci. USA **107** (2010), no. 28, 12576–12580.

[Llo09] A. L. Lloyd, *Sensitivity of model-based epidemiological parameter estimation to model assumptions*, Mathematical and Statistical Estimation Approaches in Epidemiology (G. Chowell, J. M. Hyman, L. M. A. Bettencourt, and C. Castillo-Chavez, eds.), Springer, Dordrecht, 2009, pp. 123–141.

[LNB+20] J. H. Lagergren, J. T. Nardini, R. E. Baker, M. J. Simpson, and K. B. Flores, *Biologically-informed neural networks guide mechanistic modeling from sparse experimental data*, PLOS Comput. Biol. **16** (2020), no. 12, e1008462.

[Loc] *Plague of Locusts Timelapse — Wild Africa — BBC Earth*, uploaded by BBC Earth on August 21, 2009, last accessed June 17, 2022. `https://www.youtube.com/watch?v=1YNy2R3hg2Q`

[LRC01] H. Levine, W. J. Rappel, and I. Cohen, *Self-organization in systems of self-propelled particles*, Phys. Rev. E **63** (2001), 017101.

[LSZ15] K. Law, A. Stuart, and K. Zygalakis, *Data assimilation: A mathematical introduction*, Texts in Applied Mathematics, vol. 62, Springer, Cham, 2015, DOI 10.1007/978-3-319-20325-6. MR3363508

[LW11] A. Lachapelle and M.-T. Wolfram, *On a mean field game approach modeling congestion and aversion in pedestrian crowds*, Transp. Res. B: Methodol. **45** (2011), no. 10, 1572–1589.

[LYY18] S. Lehmann and Y.-Y. Ahn (eds.), *Complex spreading phenomena in social systems: Influence and contagion in real-world social networks*, Computational Social Sciences, Springer, Cham, 2018, DOI 10.1007/978-3-319-77332-2. MR3966401

[MB11] T. Mora and W. Bialek, *Are biological systems poised at criticality?*, J. Stat. Phys. **144** (2011), no. 2, 268–302, DOI 10.1007/s10955-011-0229-4. MR2823156

[MBPK16] N. M. Mangan, S. L. Brunton, J. L. Proctor, and J. N. Kutz, *Inferring biological networks by sparse identification of nonlinear dynamics*, IEEE Trans. Mol. Biol. Multi-Scale Commun. **2** (2016), no. 1, 52–63.

[MBSR19] S. B. Munch, A. Brias, G. Sugihara, and T. L. Rogers, *Frequently asked questions about nonlinear dynamics and empirical dynamic modelling*, ICES J. Mar. Sci. **77** (2019), no. 4, 1463–1479.

[MCM+22] J. R. Minot, N. Cheney, M. Maier, D. C. Elbers, C. M. Danforth, and P. S. Dodds, *Interpretable bias mitigation for textual data: Reducing genderization in patient notes while maintaining classification performance*, ACM Trans. Comput. Healthcare **3** (2022), no. 4, 1–41.

[MCTS13] B. B. Machta, R. Chachra, M. K. Transtrum, and J. P. Sethna, *Parameter space compression underlies emergent theories and predictive models*, Science **342** (2013), no. 6158, 604–607.

[MEK99] A. Mogilner and L. Edelstein-Keshet, *A non-local model for a swarm*, J. Math. Biol. **38** (1999), no. 6, 534–570, DOI 10.1007/s002850050158. MR1698215

[Mir16] Mirrorme22, Brythones, Nilfanion (English and Scottish council areas), TUBS (Welsh council areas), and Sting (Gibraltar), *United Kingdom EU referendum 2016 area*, 2016, accessed October 12, 2022. https://commons.wikimedia.org/wiki/File:United_Kingdom_EU_referendum_2016_area_results.svg

[Mit09] M. Mitchell, *Complexity: A guided tour*, Oxford University Press, Oxford, 2009. MR2641048

[MJ] M. Madhav and E. Johnson, *What do your data say? A course to help you better understand your data*, last accessed June 30, 2022. https://www.whatdoyourdatasay.com

[MKBP17] N. M. Mangan, J. N. Kutz, S. L. Brunton, and J. L. Proctor, *Model selection for dynamical systems via sparse regression and information criteria*, Proc. R. Soc. A **473** (2017), 20170009.

[MP07] J. H. Miller and S. E. Page, *Complex adaptive systems: An introduction to computational models of social life*, Princeton Studies in Complexity, Princeton University Press, Princeton, NJ, 2007. MR2307118

[MPLC14] F. Morstatter, J. Pfeffer, H. Liu, and K. M. Carley, *Is the sample good enough? Comparing data from Twitter's Streaming API with Twitter's Firehose*, Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media **7** (2021), no. 1, pp. 400–408.

[Mun21a] E. Munch, *Python tutorial on topological data analysis*, 2021, last accessed March 14, 2023. https://zerodivzero.com/short_course/aaac8c66007a4d23a7aa14857a3b778c/title/cffb85f269864df08b61382c00c77c2d

[Mun21b] E. Munch, *Tda-python-workshop-jmm21*, 2021. https://github.com/lizliz/TDA-Python-Workshop-JMM21

[MVS20] M. R. McGuirl, A. Volkening, and B. Sandstede, *Topological data analysis of zebrafish patterns*, Proc. Natl. Acad. Sci. USA **117** (2020), no. 10, 5113–5124, DOI 10.1073/pnas.1917763117. MR4225019

[NBSF21] J. T. Nardini, R. E. Baker, M. J. Simpson, and K. B. Flores, *Learning differential equation models from stochastic agent-based model simulations*, J. R. Soc. Interface **18** (2021), 20200987.

[New11] M. E. J. Newman, *Complex systems: A survey*, Am. J. Phys. **79** (2011), no. 8, 800–810.

[New18] M. E. J. Newman, *Networks*, 2nd ed., Oxford University Press, Oxford, UK, 2018.

[NS92] K. Nagel and M. Schreckenberg, *A cellular automaton model for freeway traffic*, J. Phys. I **2** (1992), no. 12, 2221–2229.

[NSF+21] J. T. Nardini, B. J. Stolz, K. B. Flores, H. A. Harrington, and H. M. Byrne, *Topological data analysis distinguishes parameter regimes in the Anderson-Chaplain model of angiogenesis*, PLOS Comput. Biol. **17** (2021), no. 6, e1009094.

[NSnVBH+20] J. M. Nava-Sedeño, A. Voß-Böhme, H. Hatzikirou, A. Deutsch, and F. Peruani, *Modelling collective cell motion: are on- and off-lattice models equivalent?*, Phil. Trans. R. Soc. B **375** (2020), 20190378.

[Oel89] K. Oelschläger, *On the derivation of reaction-diffusion equations as limit dynamics of systems of moderately interacting stochastic processes*, Probab. Theory Related Fields **82** (1989), no. 4, 565–586, DOI 10.1007/BF00341284. MR1002901

[OFPF+17] J. M. Osborne, A. G. Fletcher, J. M. Pitt-Francis, P. K. Maini, and D. J. Gavaghan, *Comparing individual-based approaches to modelling the self-organization of multicellular tissues*, PLOS Comput. Biol. **13** (2017), no. 2, e1005387.

[OPT+17] N. Otter, M. A. Porter, U. Tillmann, P. Grindrod, and H. A. Harrington, *A roadmap for the computation of persistent homology*, EPJ Data Science **6** (2017), no. 17.

[OT12] E. Oster and R. Thornton, *Determinants of technology adoption: Peer effects in menstrual cup take-up*, J. Eur. Econ. Assoc. **10** (2012), no. 6, 1263–1293.

[PEK99] J. K. Parrish and L. Edelstein-Keshet, *Complexity, pattern, and evolutionary trade-offs in animal aggregation*, Science **284** (1999), no. 5411, 99–101.

[PG16] M. A. Porter and J. P. Gleeson, *Dynamical systems on networks: A tutorial*, Frontiers in Applied Dynamical Systems: Reviews and Tutorials, vol. 4, Springer, Cham, 2016, DOI 10.1007/978-3-319-26641-1. MR3468887

[PMS17] H. Peng, F. Menczer, and K. Sasahara, *EchoDemo: How echo chambers emerge from social media*, 2017, last accessed June 17, 2022. `https://osome.iu.edu/demos/echo/`

[Por25] M. A. Porter, *A non-expert's introduction to data ethics for mathematicians*, Mathematical and computational methods for complex social systems, Proceedings of Symposia in Applied Mathematics, vol. 80, American Mathematical Society, Providence, RI, USA, 2025, pp. 65–88.

[PS12] M. J. Plank and M. J. Simpson, *Models of collective cell behaviour with crowding effects: Comparing lattice-based and lattice-free approaches*, J. R. Soc. Interface **9** (2012), 2983–2996.

[PT03] D. M. Parichy and J. M. Turner, *Temporal and cellular requirements for Fms signaling during zebrafish adult pigment pattern development*, Development **130** (2003), no. 5, 817–833.

[QSI] *Institute for the Quantitative Study of Inclusion, Diversity, and Equity*, last accessed June 30, 2022. `https://qsideinstitute.org`

[Rea22] *RealClearPolitcs: Polls*, 2022, last accessed October 13, 2022. `https://www.realclearpolitics.com/epolls/latest_polls/elections/`

[RPK19] M. Raissi, P. Perdikaris, and G. E. Karniadakis, *Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations*, J. Comput. Phys. **378** (2019), 686–707, DOI 10.1016/j.jcp.2018.10.045. MR3881695

[RRH09] J. M. Restrepo, R. C. Rael, and J. M. Hyman, *Modeling the influence of polls on elections: A population dynamics approach*, Public Choice **140** (2009), no. 3/4, 395–420.

[SCDM+18] R. E. Stern, S. Cui, M. L. Delle Monache, R. Bhadani, M. Bunting, M. Churchill, N. Hamilton, R. Haulcy, H. Pohlmann, F. Wu, B. Piccoli, B. Seibold, J. Sprinkle, and D. B. Work, *Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments*, Transp. Res. Part C Emerg. Technol. **89** (2018), 205–221.

[Sch73] T. C. Schelling, *Hockey helmets, concealed weapons, and daylight saving: A study of binary choices with externalities*, J. Conflict Resolut. **17** (1973), no. 3, 381–428.

[Sch78] G. Schwarz, *Estimating the dimension of a model*, Ann. Statist. **6** (1978), no. 2, 461–464. MR468014

[SCP+21] K. Sasahara, W. Chen, H. Peng, G. L. Ciampaglia, A. Flammini, and F. Menczer, *Social influence and unfollowing accelerate the emergence of echo chambers*, J. Comput. Soc. Sci. **4** (2021), 381–402.

[SCS+18] A. Schadschneider, M. Chraibi, A. Seyfried, A. Tordeux, and J. Zhang, *Pedestrian dynamics: From empirical results to modeling*, Crowd Dynamics, Vol. 1 (L. Gibelli and N. Bellomo, eds.), Modeling and Simulation in Science, Engineering and Technology, Birkhäuser, Cham, 2018, pp. 63–102.

[SEK13] L. A. Segel and L. Edelstein-Keshet, *A primer on mathematical models in biology*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2013, DOI 10.1137/1.9781611972504. MR3076320

[SFK+08] Y. Sugiyama, M. Fukui, M. Kikuchi, K. Hasebe, A. Nakayama, K. Nishinari, S. Tadaki, and S. Yukawa, *Traffic jams without bottlenecks—experimental evidence for the physical mechanism of the formation of a jam*, New J. Phys. **10** (2008), 033001.

[SFK16] J. Schleuss, J. Fox, and P. Krishnakumar, *California 2016 election precinct maps*, accessed March 14, 2019. `https://github.com/datadesk/california-2016-election-precinct-maps`

[Shm10] G. Shmueli, *To explain or to predict?*, Statist. Sci. **25** (2010), no. 3, 289–310, DOI 10.1214/10-STS330. MR2791669

[SHP16] B. J. Stolz, H. A. Harrington, and M. A. Porter, *The topological "shape" of Brexit*, `arXiv:1610.00752`, 2016.

[SHT10] S. Sarmady, F. Haron, and A. Z. Talib, *Simulating crowd movements using fine grid cellular automata*, 2010 12th International Conference on Computer Modelling and Simulation, IEEE, 2010, pp. 428–433.

[Sil12] N. Silver, *The signal and the noise: Why so many predictions fail—but some don't*, Penguin Press, New York City, NY, USA, 2012.

[Sil14] N. Silver, *FiveThirtyEight: Here's proof some pollsters are putting a thumb on the scale*, 2014, last accessed June 17, 2022. `https://fivethirtyeight.com/features/heres-proof-some-pollsters-are-putting-a-thumb-on-the-scale/`

[SMA20] D. Sabin-Miller and D. M. Abrams, *When pull turns to shove: A continuous-time model for opinion dynamics*, Phys. Rev. Research **2** (2020), 043001.

[Smi14] R. C. Smith, *Uncertainty quantification: Theory, implementation, and applications*, Computational Science & Engineering, vol. 12, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2014. MR3155184

[SNV15] A. P. Singh and C. Nüsslein-Volhard, *Zebrafish stripes as a model for vertebrate colour pattern formation*, Curr. Biol. **25** (2015), no. 2, R81–R92.

[SSS17] A. Sieben, J. Schumann, and A. Seyfried, *Collective phenomena in crowds–Where pedestrian dynamics need social psychology*, PLOS ONE **12** (2017), no. 6, e0177328.

[Sto] *storywrangler: From the University of Vermont Computational Story Lab*, last accessed June 30, 2022. `https://storywrangling.org`

[Str01] S. H. Strogatz, *Exploring complex networks*, Nature **410** (2001), 268–276.

[Str15] S. H. Strogatz, *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering*, 2nd ed., CRC Press, 2015.

[Tak81] F. Takens, *Detecting strange attractors in turbulence*, Dynamical Systems and Turbulence, Warwick 1980, Lecture Notes in Math. (D. Rand and L. S. Young, eds.), vol. 898, Springer, Berlin, 1981, pp. 366–381.

[TBL06] C. M. Topaz, A. L. Bertozzi, and M. A. Lewis, *A nonlocal continuum model for biological aggregation*, Bull. Math. Biol. **68** (2006), no. 7, 1601–1623, DOI 10.1007/s11538-006-9088-6. MR2257718

[TECP20] J. H. Tien, M. C. Eisenberg, S. T. Cherng, and M. A. Porter, *Online reactions to the 2017 'Unite the right' rally in Charlottesville: Measuring polarization in Twitter networks using media followership*, Appl. Netw. Sci. **5** (2020), no. 10.

[TFB+11] J. Tsai, N. Fridman, E. Bowring, M. Brown, S. Epstein, G. Kaminka, S. Marsella, A. Ogden, I. Rika, A. Sheel, M. Taylor, X. Wang, A. Zilka, and M. Tambe, *ESCAPES - Evacuation simulation with children, authorities, parents, emotions, and social comparison*, Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems - Innovative Applications Track (AAMAS 2011) (Tumer, Yolum, Sonenberg, and Stone, eds.), 2011, pp. 457–464.

[THK18] S. Thurner, R. Hanel, and P. Klimek, *Introduction to the theory of complex systems*, Oxford University Press, Oxford, 2018. MR3889059

[Tib96] R. Tibshirani, *Regression shrinkage and selection via the lasso*, J. Roy. Statist. Soc. Ser. B **58** (1996), no. 1, 267–288. MR1379242

[TSB+14] K. K. Treloar, M. J. Simpson, B. J. Binder, D. L. S. McElwain, and R. E. Baker, *Assessing the role of spatial correlations during collective cell spreading*, Sci. Rep. **4** (2014), 5713.

[Tuf14] Z. Tufekci, *Big questions for social media big data: Representativeness, validity and other methodological pitfalls*, Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media, 2014, pp. 505–514.

[Twi] *Twitter Developer Platform: Tutorials*, last accessed June 30, 2022. `https://developer.twitter.com/en/docs/tutorials`

[TZH15] C. M. Topaz, L. Ziegelmeier, and T. Halverson, *Topological data analysis of biological aggregation models*, PLOS ONE **10** (2015), no. 5, e0126383.

[UZT19] M. Ulmer, L. Ziegelmeier, and C. M. Topaz, *A topological approach to selecting models of biological experiments*, PLOS ONE **14** (2019), no. 3, e0213679.

[VCBJ+95] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, *Novel type of phase transition in a system of self-driven particles*, Phys. Rev. Lett. **75** (1995), no. 6, 1226–1229, DOI 10.1103/PhysRevLett.75.1226. MR3363421

[VCM+07] A. Varas, M. D. Cornejo, D. Mainemer, B. Toledo, J. Rogan, V. Muñoz, and J. A. Valdivia, *Cellular automaton model for evacuation process with obstacles*, Physica A **382** (2007), no. 2, 631–642.

[VLPR20] A. Volkening, D. F. Linder, M. A. Porter, and G. A. Rempala, *Forecasting elections using compartmental models of infection*, SIAM Rev. **62** (2020), no. 4, 837–865, DOI 10.1137/19M1306658. MR4167616

[Vol17] A. Volkening, *Modeling pattern formation on zebrafish*, PhD diss., Brown University, 2017.

[Vol20a] A. Volkening, *Intro to building models*, 2020, last accessed June 30, 2022. `https://northwestern.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=7d04a874-a292-4ff2-bd66-ac2500daeea1`

[Vol20b] A. Volkening, *Linking genotype, cell behavior, and phenotype: Multidisciplinary perspectives with a basis in zebrafish patterns*, Curr. Opin. Genet. Dev. **63** (2020), 78–85.

[Vol21] A. Volkening, *Data-driven modeling*, 2021, last accessed June 30, 2022. `https://zerodivzero.com/short_course/aaac8c66007a4d23a7aa14857a3b778c/title/d56faebff3a24f77a76085c1427038d8`

[VS15] A. Volkening and B. Sandstede, *Modelling stripe formation in zebrafish: An agent-based approach*, J. R. Soc. Interface **12** (2015), no. 112, 20150812.

[VS18] A. Volkening and B. Sandstede, *Iridophores as a source of robustness in zebrafish stripes and variability in* Danio *patterns*, Nat. Commun. **9** (2018), no. 3231.

[WE02] C. Wlezien and R. S. Erikson, *The timeline of presidential election campaigns*, J. Politics **64** (2002), no. 4, 969–993.

[Woo22] T. E. Woolley, *Boundary conditions cause different generic bifurcation structures in Turing systems*, Bull. Math. Biol. **84** (2022), no. 9, Paper No. 101, DOI 10.1007/s11538-022-01055-x. MR4468545

[WPCG+14] C. H. Weiss, J. Poncela-Casasnovas, J. I. Glaser, A. R. Pah, S. D. Persell, D. W. Baker, R. G. Wunderink, and L. A. Nunes Amaral, *Adoption of a high-impact innovation in a homogeneous population*, Phys. Rev. X **4** (2014), no. 4, 041008.

[YAKM20] V. C. Yang, D. M. Abrams, G. Kernell, and A. E. Motter, *Why are U.S. parties so polarized? A "satisficing" dynamical model*, SIAM Rev. **62** (2020), no. 3, 646–657, DOI 10.1137/19M1254246. MR4131342

[ZKSS12] J. Zhang, W. Klingsch, A. Schadschneider, and A. Seyfried, *Ordering in bidirectional pedestrian flows and its influence on the fundamental diagram*, J. Stat. Mech. **2012** (2012), P02002, DOI 10.1088/1742-5468/2012/02/P02002.

Department of Mathematics, Purdue University, West Lafayette, Indiana 47907
Email address: `avolkening@purdue.edu`

Proceedings of Symposia in Applied Mathematics
Volume **80**, 2025
https://doi.org/10.1090/psapm/080/00727

A model for wealth concentration: From a discrete system to a PDE

A. Halev, K. Patel, N. Rodríguez, M. Tewari, and L. Wong

Abstract. The evolution of urban areas plays a major role in crafting public policy and spurring investment. We propose a discrete model to study the existence and dynamics of spatial wealth concentration, centered on a two-dimensional lattice with wealth and an inherent amenities interacting via a feedback process. Various parameter regimes beget distinct dynamics, including that of spatial wealth concentration—as consistent with empirical observation. To enable a more rigorous analysis we derive a continuum model of partial differential equations (PDEs) from the discrete model and analyze the instability regime of the continuum system; the resultant regime agrees with observations of simulation of the discrete model. We also perform a sensitivity analysis on the continuum model to determine how small changes in parameters affect solutions.

1. Introduction

Gentrification is well-documented as a major factor in modern urban development. This residential phenomenon is associated with major increases in housing prices and upgrades of local amenities, leading to an emigration of low-income residents and an influx of wealthier community members. This wealthier contingent is generally whiter, more educated, and younger compared to the low-income residents they replace [**39**].

While the existence of gentrification is a hot topic in political circles, its underlying causes and effects are widely disputed. An inability to study the problem experimentally and the potential for a wide variety of motivating factors complicate any deep understanding of the issue. Hamnett (1991) suggested three main drivers for gentrification—the existence of middle-class potential gentrifiers, an availability of urban housing, and a tendency among these potential gentrifiers to prefer to live in an urban setting [**20**]. Other proposed drivers of gentrification include falling crime rates in inner city neighborhoods [**12**, **16**], demanding work schedules and lack of free time among the young middle class [**15**], proximity to social amenities, such as coffee shops, beer gardens, bike shares, gyms and restaurants [**10**], and increased racial tolerance among Millennials [**22**].

2020 *Mathematics Subject Classification.* Primary 35B36, 00A71.

The third and fifth authors were partially funded by the National Science Foundation (NSF, No. DMS-1516778). The third and fourth authors were partially funded by the UNC Fire grant.

Further obfuscating an empirical understanding of gentrification is the potential presence of inherently chaotic dynamics [**1**, **24**, **31**]. Despite this, the majority of theoretical analysis of gentrification has focused on binary divisions—blacks and whites, flows of capital and flows of people, macro-forces of capital accumulation—concentrating on subsets of the potential dynamics involved in gentrification [**3**, **39**].

Typically, modeling gentrification involves agent-based models that allow virtual simulation in lieu of experiment; the seminal "Schelling model" of residential segregation utilized an agent-based model inhabiting an eight by eight lattice with two classes of agents to represent an arbitrary binary social division [**32**, **33**]. The Schelling model found that segregation was rampant even in situations where agents were willing to inhabit neighborhoods that consisted of up to two thirds of the other group [**32**, **33**].

Extensions of the Schelling model to examine a variety of issues related to residential segregation and gentrification focused on similar agent-based approaches [**7**, **23**, **30**, **38**]. However, the discrete nature of the agent-based model prevents the implementation of various analytical techniques to better understand such a system.

PDEs are a valuable tool to model the spatiotemporal dynamics of ecological and sociological systems [**4**, **25**], and the derivation of PDE systems from agent-based models has proven to be effective in a range of mathematical applications, particularly in mathematical biology [**2**, **14**, **17**]. In more recent work with gentrification modeling, Hasan et al. (2020) used transport theory to derive a PDE model of the dynamics of wealth and amenities [**21**]. Short et al. (2008) considered agent-based and continuum (PDE) models of criminal behavior and showed that the two systems were in agreement in the limit of large system sizes [**34**].

In contrast to Hasan et al. (2020), we use a similar approach to that of Short et al. (2008) to model the dynamics of wealth and amenities, starting from first principles. We begin with a discrete model based on basic assumptions about what attracts wealth and show that this model exhibits qualitative similarity to the dynamics of gentrification seen in actual cities. We then derive a system of PDEs from this discrete model, and show that the resultant reaction-advection-diffusion system exhibits qualitative similarity to the discrete model for relevant values of our parameters. Thus, analysis of the continuum model is an effective tool to examine the dynamics of the discrete model. We examine the existence of hotspots, small areas or regions with a relatively high level of wealth in comparison to their surroundings, and the parameter regimes under which they occur to be able to decompose the variety of factors involved in the incipient stages of gentrification. We also determine which parameters are most and least sensitive for the system through a mathematical sensitivity analysis.

2. Discrete model

In this section we derive a simple discrete model that describes the coupling of amenities and wealth. Our objective is show that the mathematical framework presented here can be useful to study wealth concentration, and consequently important social issues such as gentrification. In particular, we show that in this mathematical framework, models based on simple principles can lead to the concentration of wealth. Thus, we present a first model, as a proof of concept, to

initiate the process of iterating and refining the model until a suitable one is obtained. With this in mind, we expect that in the process of model validation, the model presented here will likely have to be refined.

Our discrete model focuses on the dynamics of wealth in a given community, for example, measured in dollars as incomes or net worth—-and on an intrinsic, dimensionless amenities, meant to encompass the variety of spatial factors involved in gentrification—from proximity to work to the density of coffee shops and restaurants [**10**, **15**].

We consider a community to be a two-dimensional square lattice with lattice spacing ℓ. For simplicity's sake, we take ℓ to be constant, although it can be varied to incorporate more complex geometries. Within the community, each lattice site, denoted by s, is equipped with a given wealth, denoted $W_s(t)$, and some density of amenities, denoted by $A_s(t)$, at time t. The dynamics will be updated discretely in time with constant time step δt.

If unmaintained, we expect amenities to decay in time with some decay rate ω. On the other hand, we assume that amenity features will increase proportionally to the wealth at a given site due to institutional factors such as increases in property taxes and effectiveness of home owners associations. Newer, wealthier residents may also demand improved or different goods and services, prompting an influx of new retailers who expand, provided residents have the capital to sustain them [**8**, **29**]. Thus, it is natural to assume that residents will invest in their neighborhood amenities at a rate proportional to their wealth. For this purpose, we thus introduce a parameter ϕ, the rate of increase in amenities, per dollar, per unit time; as a result, we model this growth and decay by:

$$A_s(t+\delta t) = A_s(t)(1-\omega\delta t) + \phi W_s(t)\delta t. \tag{2.1}$$

Certain sources of investment—in particular, those originating in the private sector—are highly sensitive to the likelihood of return on investment and thus on the level of current wealth. Other sources, however, are more stable; public investment in infrastructure—such as schools, parks, highways, and rail transit—play a role in the gentrification of neighborhoods [**6**, **39**]. We model this stable, external investment by allowing our amenities to grow at a constant rate $\Gamma > 0$.

As a phenomenon, gentrification does not occur at isolated sites; rather, a wealth of literature shows that neighborhoods are often segregated by socioeconomic class [**9**, **18**, **19**, **28**]. Similarly, individuals do not only invest in amenities at their specific location, but rather in the amenities of the neighborhood. There are a variety of ways one can incorporate this into our model. For example, one could model a direct investment, proportional to the wealth at site s, to their neighboring lattice sites s'. This would lead to a cross-diffusion term [**27**], that is, the wealth density would induce a flux in the amenities. An alternative way to model this effect, which is inspired from the work of Short *et al.* [**34**], is to assume that the level of amenities diffuses at some rate. Given the analytical complications brought about by cross-diffusion terms, in the paper, we choose the latter way to model this effect because it leads to a more analytically tractable model. This leads to an update rule for A_s that allows dispersal to its neighboring sites. Specifically, we introduce a parameter $0 < \eta < 1$ that measures the relative strength of neighborhood

effects and alter our update rule:

$$A_s(t+\delta t) = \left[(1-\eta)A_s(t) + \frac{\eta}{z}\sum_{s'\sim s} A_{s'}(t)\right](1-\omega\delta t) + \phi W_s(t)\delta t + \Gamma\delta t, \tag{2.2}$$

where z is the coordination number—denoting the number of sites adjacent to s—and the sum is taken over all sites s' that neighbor s. Note that any given site s has a fixed amount of wealth to invest, and the parameter η is introduced to guarantee that the investment on site s is the total investment minus what will be invested in neighboring sites of s.

We devise a similar update rule to model the wealth dynamics. We begin with the natural assumption that reinvestment occurs at a rate proportional to the wealth at that site. In lieu of a constant of proportionality, we consider a rate function $f(W_s, A_s)$:

$$W_s(t+\delta t) = W_s(t)(1-\omega\delta t) + f(W_s, A_s)W_s(t)\delta t. \tag{2.3}$$

Similar to population dynamics models, it is possible that the growth/decay rate of wealth is proportional to the wealth and amenities level. Different forms of $f(W_s, A_s)$ may be adopted to better model particular municipalities or underlying factors. As examples, one may consider:

$$f(W_s) = \left(1 - \frac{W_s}{M}\right), \tag{2.4}$$

wherein reinvestment is highest when wealth is lowest and steadily decreases as it approaches some carrying capacity M, modeling logistic growth, or:

$$f(W_s, A_s) = r\left(1 - \frac{W_s}{M}\right)\left(\frac{W_s}{M} - A_s\right), \tag{2.5}$$

where there is growth in middle class areas ($1 < W_s/M < A_s$) and decay otherwise, modeling bistable growth as was proposed in [**5**]. In both of these examples r and M parameters. To aid in the derivation of the continuum model in Section 3, we assume that there is a decay rate of wealth that is proportional to W. Recall, that growth is modeled in the function f. To minimize the number of parameters introduced, this rate is chosen also be ω. This is a mathematical approximation for the belief of some sociologist that neighborhood decline, the deterioration of neighborhoods often caused by lack of investment and maintenance, leads to loss of jobs and thus wealth [**36**].

We implement neighborhood effects in a similar manner as we have done for the amenities. Research has shown that property values, in particular, rise in accordance to their proximity to quality schools and parks [**26**, **35**] and highways [**37**]. Moreover, here is were we incorporate the hypothesis that investment in neighborhoods are skewed towards sites with amenities A_s that are high relative to all neighbors of a given site s' [**10**].

Our update rule for wealth is thus:

$$\begin{aligned} W_s(t+\delta t) = &\left[(1-\eta)W_s(t) + \eta A_s(t)\sum_{s'\sim s}\frac{W_{s'}(t)}{\sum_{s''\sim s'} A_{s''}(t)}\right](1-\omega\delta t) \\ &+ r(W_s, A_s)W_s(t)\delta t. \end{aligned} \tag{2.6}$$

Note that the dispersal of wealth is assumed to be equal to the dispersal of the amenities. Thus, the neighborhood effects for both amenities and wealth are

equal. This can be generalized and would simply result in the introduction of a new dispersal rate for wealth.

From this point, we consider the logistic rate of reinvestment in the form of (2.4), where r is a constant growth rate and M a carrying capacity. This choice of reinvestment rate is motivated by our expectation of two distinct regimes:

(1) Direct, external sources of investment—such as government funding and philanthropy—are targeted towards areas of lower wealth. These sources are highly motivated by return on investment; if low-wealth areas are seen to improve with direct investment, investment will increase up to a certain point.
(2) Once a certain wealth threshold has been achieved, these external sources redirect their focus towards different areas. In fact, those same government entities that previously served as a source of wealth may now serve as a sink through various agents, taxation chiefly among them.

As such, we have the following update rule:

$$W_s(t+\delta t) = \left[(1-\eta)W_s(t) + \eta A_s(t)\sum_{s'\sim s}\frac{W_{s'}(t)}{\sum_{s''\sim s'}A_{s''}(t)}\right](1-\omega\delta t) + r\left(1-\frac{W_s(t)}{M}\right)W_s(t)\delta t. \tag{2.7}$$

Equations (2.2) and (2.7) form the main components of our discrete system with logistic rate of reinvestment. We employ no flux boundary conditions, so that no wealth or amenities is lost through the boundaries; all sources and sinks are contained within the governing equations. The relevant parameters are summarized in Table 1.

$$\begin{cases} A_s(t+\delta t) = \left[(1-\eta)A_s(t) + \dfrac{\eta}{z}\displaystyle\sum_{s'\sim s}A_{s'}(t)\right](1-\omega\delta t) + \phi W_s(t)\delta t + \Gamma\delta t, & (2.2)\\ W_s(t+\delta t) = \left[(1-\eta)W_s(t) + \eta A_s(t)\displaystyle\sum_{s'\sim s}\frac{W_{s'}(t)}{\sum_{s''\sim s'}A_{s''}(t)}\right](1-\omega\delta t) & \\ \qquad\qquad + r\left(1-\dfrac{W_s(t)}{M}\right)W_s(t)\delta t. & (2.7) \end{cases}$$

TABLE 1. Discrete Parameters

Parameter	Interpretation
ω	Wealth and amenities decay rate
δt	Time step size
ϕ	Rate of investment in amenities per dollar
η	Strength of neighborhood effects (between zero and unity)
Γ	External investment in amenities
r	Wealth growth rate
M	Wealth regulating factor
ℓ	Lattice spacing

Note that all parameters are nonnegative.

2.1. Discrete solutions. With this system, we have two spatially homogeneous solutions:

$$\begin{pmatrix} \overline{W}_s \\ \overline{A}_s \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{\Gamma}{\omega} \end{pmatrix} \tag{2.8}$$

and

$$\begin{pmatrix} \overline{W}_s \\ \overline{A}_s \end{pmatrix} = \begin{pmatrix} M\left[1-\rho^{-1}\right] \\ \frac{M\phi}{r}\left[\rho-1\right]+\frac{\Gamma}{\omega} \end{pmatrix}, \tag{2.9}$$

where $\rho = \frac{r}{\omega}$. Note that if $\rho = 1$ the two equilibrium solutions collide, whereas they are otherwise distinct. As such, we can predict a transcritical bifurcation at this point, pending the stability of these spatially homogeneous solutions. To determine whether these spatially homogeneous solutions are stable and analyze the potential bifurcation about the point $\rho = 1$, we run simulations on this system for various parameter regimes as detailed below.

2.2. Discrete simulations. Simulations of our discrete system, (2.2) and (2.7), are run in MATLAB on a square lattice with spacing $\ell = 0.04$ and time step $\delta t = 0.01$ with periodic boundary conditions. By varying our parameters, we observe three distinct sets of dynamics:

(1) *Homogeneous destitution.* In this case, both the amenities and wealth decay throughout, and the entire domain quickly approaches our first spatially homogeneous solution given in (2.8). In particular, we observe homogeneous destitution persists in the regime $\rho \leq 1$, considering only positive values of these parameters.
(2) *Homogeneous wealth.* Here, amenities and wealth quickly converge to the solution given in (2.9). This regime occurs for $\rho \geq 1$. While the homogeneous wealth solution does exist for $\rho < 1$, it takes on negative values in this regime and appears to be unstable to small perturbations; simulations with $\rho < 1$ and with the homogeneous wealth solution as initial condition eventually diverge as numerical errors accumulate.
(3) *Wealth hotspots.* In this regime, spatial homogeneity is not achieved in reasonable time scales. Small pockets of wealth and amenities are surrounded by large areas of destitution. These hotspots form early and quickly become circular. However, achieving temporal stability can take significant time, with hotspots deforming and merging in the process before returning to their circular state. The parameter regime leading to hotspots appears to be $1 < \rho < 1 + \epsilon$, where $\varepsilon > 0$ is small and may depend on other parameters or initial conditions. We will see that these hotspots are easier to see and analyze in the continuum case.

The time evolution for one set of parameters in each respective regime is displayed in Figure 1.

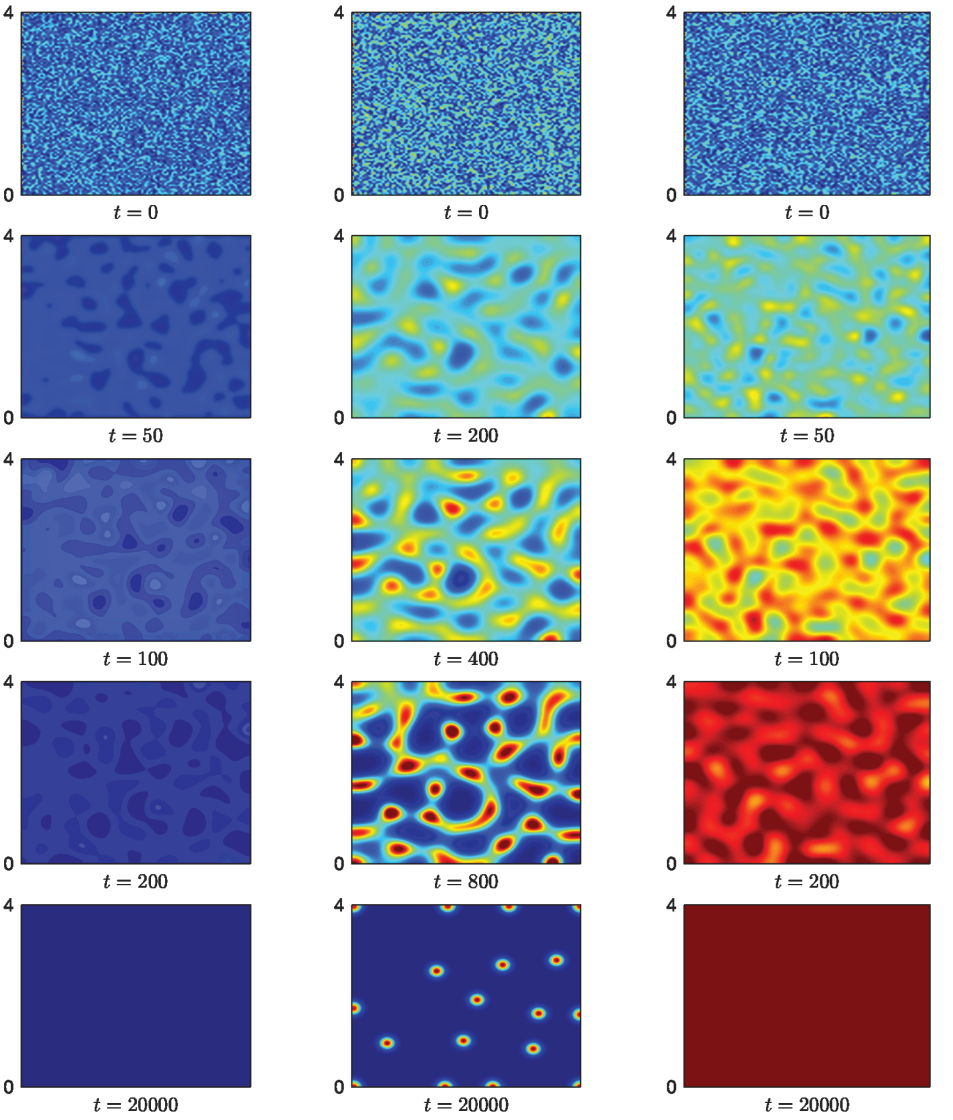


FIGURE 1. Output of the discrete simulation for $\Gamma = 0$ in the three distinct parameter regimes. Left, homogeneous destitution, $\rho = 0.9$. Center, hotspots, $\rho = 1.1$. Right, homogeneous wealth, $\rho = 1.3$.

3. Continuum limit

To examine the dynamics of the system in greater detail, we derive a continuum system from the discrete model. Rewriting equation (2.2) as:

$$A_s(t+\delta t) = \left[A_s(t) + \frac{\eta \ell^2}{z} \Delta A_s(t)\right](1 - \omega \delta t) + \phi W_s(t)\delta t + \Gamma \delta t, \tag{3.1}$$

where $\Delta A_s(t)$ is the discrete spatial Laplacian:

$$\Delta A_s(t) = \left(\sum_{s' \sim s} A_{s'}(t) - z A_s(t) \right) / \ell^2. \tag{3.2}$$

We now subtract $A_s(t)$ from both sides, convert $W_s(t)$ and $A_s(t)$ into wealth and amenities densities, denoted by W and A respectively, by dividing by ℓ^2, and divide through by δt. Taking limits as δt, $\ell^2 \to 0$ and requiring that $D = \ell^2/\delta t$ and $\tilde{\Gamma} = \Gamma/\ell^2$ remain constant, we arrive at our continuous amenities equation:

$$\frac{\partial A}{\partial t} = \frac{\eta D}{z} \Delta A - \omega A + \phi W + \tilde{\Gamma}. \tag{3.3}$$

By performing a series of similar—if slightly more involved—operations, we arrive at our continuous equation for the wealth density:

$$\frac{\partial W}{\partial t} = \frac{\eta D}{z} \Big[\Delta W - 2\nabla \cdot (W \nabla \log A) \Big] + r \left(1 - \frac{W}{\tilde{M}} \right) W - \omega W. \tag{3.4}$$

Note that $\tilde{M} = \frac{M}{\ell^2}$. Equations (3.3) and (3.4) combine to form a system that serves as a continuous counterpart to the discrete model; they are of the general form of a reaction-advection-diffusion system, systems that often beget pattern formation [**11**]. The parameters of the continuum system are summarized in Table 2.

In the continuum system, amenities diffuse spatially while decaying in time; simultaneously, higher levels of wealth and external sources lead to investment in the amenities of a community. Wealth also diffuses spatially and decays in time; wealth reinvestment occurs if the level of wealth is below a certain carrying capacity, that, if exceeded, leads to an additional sink of wealth.

3.1. Dimensionless equations. To better understand the intrinsic properties of the system, we nondimensionalize using the characteristic time scale $\tau \equiv 1/\omega$ and length scale $\ell_c \equiv \sqrt{\eta D/\omega}$, arriving at the following scaled variables:

$$\tilde{A} = \frac{\omega}{\phi} \tilde{M} A, \quad \tilde{W} = \frac{1}{\tilde{M}} W, \quad \tilde{\mathbf{x}} = \frac{\sqrt{z}}{\ell_c} \mathbf{x}, \quad \tilde{t} = \omega t. \tag{3.5}$$

Our dimensionless equations are thus

$$\begin{aligned} \frac{\partial W}{\partial t} &= \Delta W - 2\nabla \cdot (W \nabla \log A) + \rho(1 - W)W - W, \\ \frac{\partial A}{\partial t} &= \Delta A - A + W + \gamma, \end{aligned} \tag{3.6}$$

where $\rho = r/\omega$, as defined earlier, and $\gamma \equiv \tilde{\Gamma}/(\phi M)$.

This nondimensionalization has reduced our parameter space from the original eight dimensional parameters to two nondimensional. Our two dimensionless parameters have clear interpretations in terms of our dimensional variables, reinforcing our choice of nondimensionalization. Specifically, the parameter ρ measures the relative rates of investment and decay; $\rho > 1$ implies excess wealth should be available in the domain while $\rho < 1$ suggests decay may overwhelm investment. Meanwhile the parameter γ is the nondimensional external investment in amenities, which in this case is spatially and temporally homogeneous. Note that it is a rate of investment that is relative to the carrying capacity of the wealth.

TABLE 2. Continuum Parameters

Parameter		Interpretation
Dimensionless	Dimensional	
γ	$\frac{\tilde{\Gamma}}{\phi\tilde{M}}$	External investment into amenities
ρ	$\frac{r}{\omega}$	Relative rate of investment and decay
	D	Diffusion constant

In terms of our nondimensional parameters, the spatially homogeneous solutions are:

$$\begin{pmatrix}\overline{W}\\ \overline{A}\end{pmatrix} = \begin{pmatrix}0\\ \gamma\end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix}\overline{W}\\ \overline{A}\end{pmatrix} = \begin{pmatrix}1-\frac{1}{\rho}\\ 1-\frac{1}{\rho}+\gamma\end{pmatrix}, \tag{3.7}$$

henceforth denoted as *homogeneous destitution* and *homogeneous wealth*, respectively. Once again, depending on the stability of spatially homogeneous solutions about $\rho = 1$, we can predict a transcritical bifurcation at this point.

3.2. Numerical simulation. Numerical simulation is performed via the MATLAB PDE Toolbox using no-flux boundary conditions on a square grid. Simulations of our continuum system exhibit striking similarity to their counterpart in the discrete system. In particular, systems with $\rho < 1$ exhibit decay to the homogeneous destitution solution, while taking $\rho > 1$ will either lead to hotspots or homogeneous wealth for ρ sufficiently close to unity and larger ρ respectively. These three regimes correlate with the dynamics of the discrete system, both in terms of the observed dynamics and the parameters that give rise to them.

Spatial dynamics for various sets of parameters corresponding to those used in the discrete simulation can be seen in Figure 2. The visible similarity indicates that our continuum model is an accurate approximation of our discrete equations for small time steps and spacing. Encouraged by the ability of our continuum system to predict results of our discrete system, we turn to analysis of our continuum equations to attempt to distinguish systems that exhibit hotspots from those that do.

3.3. Linear order analysis. To better understand the behavior of our system surrounding these equilibrium solutions, we consider values of our variables slightly perturbed from their steady states. We start with our homogeneous wealth solution $(\overline{W}, \overline{A}) = (1-\frac{1}{\rho}, 1-\frac{1}{\rho}+\gamma)$ by analyzing perturbations of the form

$$W(\mathbf{x},t) = 1 - \frac{1}{\rho} + \delta_W e^{\sigma t} e^{i\mathbf{k}\cdot\mathbf{x}},$$
$$A(\mathbf{x},t) = 1 - \frac{1}{\rho} + \gamma + \delta_A e^{\sigma t} e^{i\mathbf{k}\cdot\mathbf{x}}.$$

By inserting these perturbations into our dynamical equations (3.6) and discarding nonlinear terms, we arrive at the following eigenvalue equation:

$$\begin{bmatrix} -|\mathbf{k}|^2 - \rho + 1 & \frac{2|\mathbf{k}|^2(1-\rho)}{1-\rho(\gamma+1)} \\ 1 & -|\mathbf{k}|^2 - 1 \end{bmatrix} \begin{bmatrix}\delta_W\\ \delta_A\end{bmatrix} = \sigma \begin{bmatrix}\delta_W\\ \delta_A\end{bmatrix}. \tag{3.8}$$

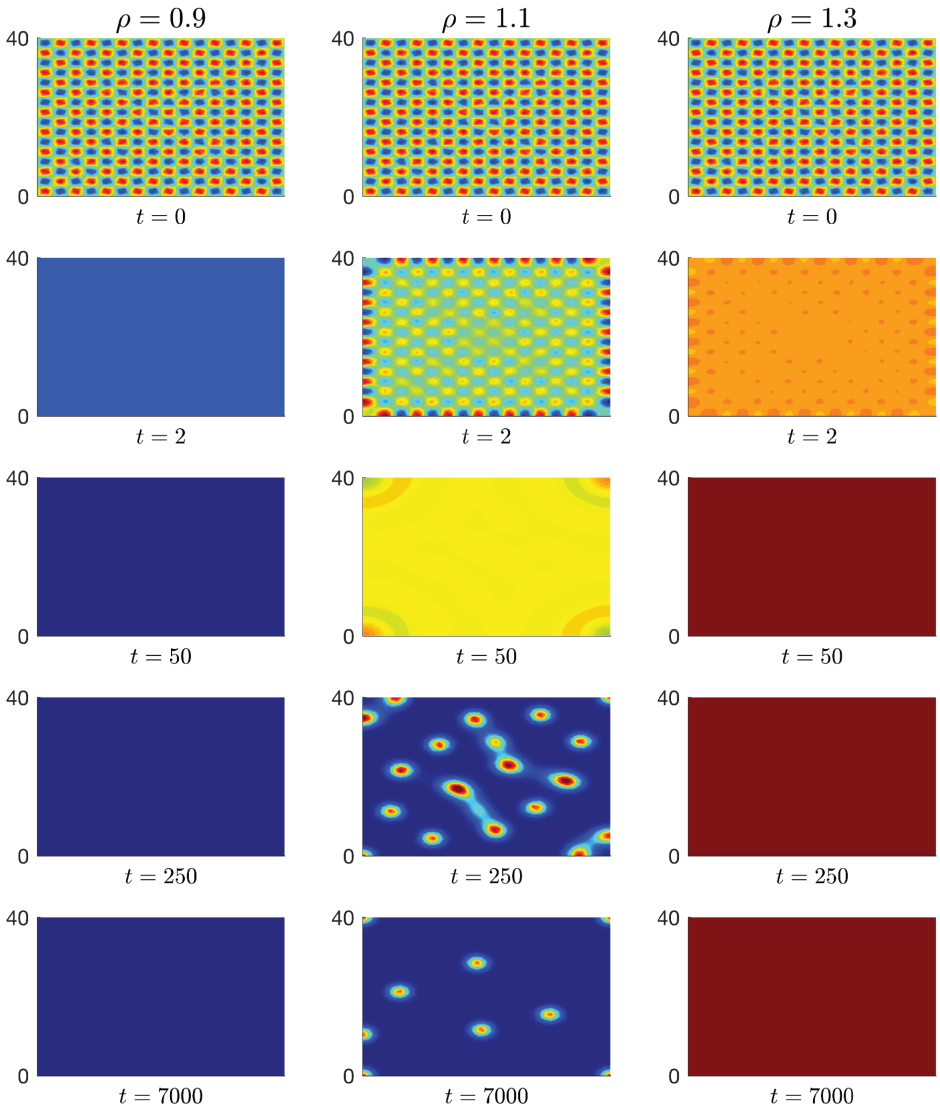


FIGURE 2. Output of the continuum simulation for $\eta = 0.01$, $\gamma = 0$. (a) Homogeneous destitution, $\rho = 0.9$. (b) Hotspots, $\rho = 1.1$. (c) Homogeneous wealth, $\rho = 1.3$.

For our system, linear order instabilities exist if the determinant of this matrix is negative, that is,

$$|\mathbf{k}|^4 + |\mathbf{k}|^2 \left[\rho + 2\frac{1-\rho}{\gamma\rho + \rho - 1}\right] + \rho - 1 < 0. \tag{3.9}$$

Note that the case when $\rho < 1$ is not a physically relevant regime. Moreover, in such case, we can see from (3.9) that the equilibrium solution will be unstable for all $\gamma \geq 0$. This result reinforces the value of our model as the homogeneous wealth solution—which implies uniform, negative wealth if $0 < \rho < 1$—is never attracting

in this case. The case $\rho = 1$ leads to stability. Hence, from here on we focus on the case $\rho > 1$, in which case (3.9) holds if γ lies within the range:

$$\frac{1}{\rho} - 1 < \gamma < -\frac{(\rho-1)\left(\rho^2 - 6\rho + 4\sqrt{\rho-1} + 4\right)}{(\rho-2)^2\rho}. \tag{3.10}$$

Given that we are considering the physically relevant case of $\gamma \geq 0$, then (3.10) can be simplified to:

$$0 \leq \gamma < -\frac{(\rho-1)\left(\rho^2 - 6\rho + 4\sqrt{\rho-1} + 4\right)}{(\rho-2)^2\rho}. \tag{3.11}$$

The inequality (3.11) is only satisfied for the bounded band given by $1 < \rho < 4 - 2\sqrt{2}$. In this regime, the maximally growing mode is given by:

$$|\mathbf{k}^*|^2 = -\frac{\rho(\gamma(\rho-2)+\rho-1)(\rho(\gamma(\rho-2)+\rho-5)+4)}{8(\rho-1)(\gamma\rho+\rho-1)} \tag{3.12}$$

and we have:

$$\begin{aligned} \sigma_{max} &= \sigma\left(|\mathbf{k}^*|^2\right) \\ &= \frac{-2\left(2\gamma^2+7\gamma+5\right)\rho^3 + (\gamma+1)^2\rho^4 + (2\gamma+5)^2\rho^2 - 8(\gamma+3)\rho + 8}{8(\rho-1)(\gamma\rho+\rho-1)}. \end{aligned} \tag{3.13}$$

The maximal eigenvalue for a set of parameters that leads to instability of the homogeneous wealth solution is plotted in Figure 3.

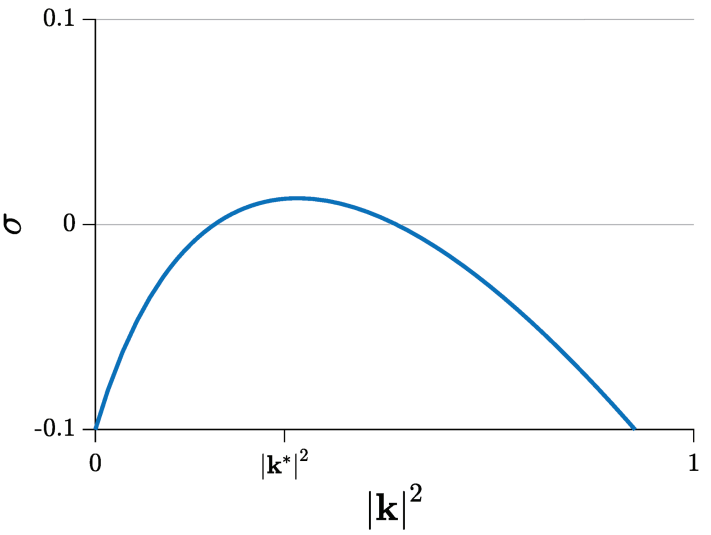


FIGURE 3. Maximal eigenvalue of perturbations of the homogeneous wealth solution is plotted for a set of parameters $\rho = 1.1$, $\gamma = 10^{-2}$. The maximal wavenumber $|\mathbf{k}^*|$ sets the final size of hot spots.

Close to the upper boundary given in (3.10), we expect the maximally growing mode $|\mathbf{k}^*|$ to dictate the size of hotspots; in particular, we expect $2\pi/|\mathbf{k}^*|$ to be the distance between hotspots. To gain a complete picture of the stability of solutions, we perform a similar linear stability analysis on the solution $(\overline{W}, \overline{A}) = (0, \gamma)$. If $\gamma > 0$, we arrive at the following matrix equation:

$$\underbrace{\begin{bmatrix} -|\mathbf{k}|^2 + \rho - 1 & 0 \\ 1 & -|\mathbf{k}|^2 - 1 \end{bmatrix}}_{B} \begin{bmatrix} \delta_W \\ \delta_A \end{bmatrix} = \sigma \begin{bmatrix} \delta_W \\ \delta_A \end{bmatrix}. \tag{3.14}$$

We see that the eigenvalues of B are its diagonal terms and it follows that $(0, \gamma)$ solution exhibits instabilities only if

$$\rho > 1. \tag{3.15}$$

The stability matrix B changes if $\gamma = 0$. Specifically, the logarithmic flux term $(\nabla \cdot (W \nabla \log A))$ in (3.6) contributes to linear order in this case, whereas its expansion is solely higher order if $\gamma > 0$. In this case, we have that:

$$\sigma(k) = |\mathbf{k}|^2 + \rho - 1. \tag{3.16}$$

Here, unstable modes exist for all ρ. In addition, all modes are unstable if $\rho > 1$; in this case, we have an ill-posed problem. Taken together, equations (3.10) and (3.15) allow us to paint a broad picture of hotspots dependent on our reinvestment parameters ρ and γ. For $\rho < 1$, there is insufficient investment in the neighborhood for anyone to maintain a modicum of wealth, and both wealth and amenities decay in time until the neighborhood is destitute in wealth if not amenities.

Two situations arise in the regime $1 < \rho < 4 - 2\sqrt{2}$, dependent on our amenities reinvestment γ. For sufficiently small γ—as defined in (3.10)—some families are able to retain their wealth; however, reinvestment is insufficient to allow wealth to prevail domain-wide and wealth is concentrated in hotspots. On the other hand, larger values of γ allow homogeneous wealth throughout for all $\rho > 1$. In the regime $\rho > 4 - 2\sqrt{2}$, wealth reinvestment is sufficient to maintain wealth; this is independent of amenity reinvestment provided amenity reinvestment is nonnegative. These results are summarized in Table 3, Figure 4, and Figure 5.

TABLE 3. Distinct Parameter Regimes

Region	Parameter Regime		Dest. Solution	Wealth Solution
Homogeneous Destitution	$\rho < 1$		Stable	Unstable
Hotspots	$\mathbf{1 < \rho < 4 - 2\sqrt{2}}$	$\boldsymbol{\gamma < \gamma_*(\rho)}$	**Unstable**	**Unstable**
Homogeneous Wealth		$\gamma > \gamma_*(\rho)$	Unstable	Stable
	$\rho > 4 - 2\sqrt{2}$		Unstable	Stable

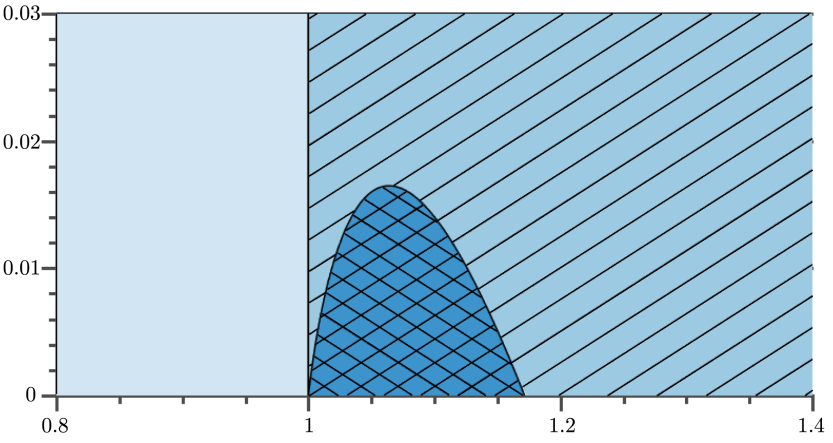


FIGURE 4. Regions of homogeneous destitution, homogeneous wealth and hotspots, shown here in increasing opacity.

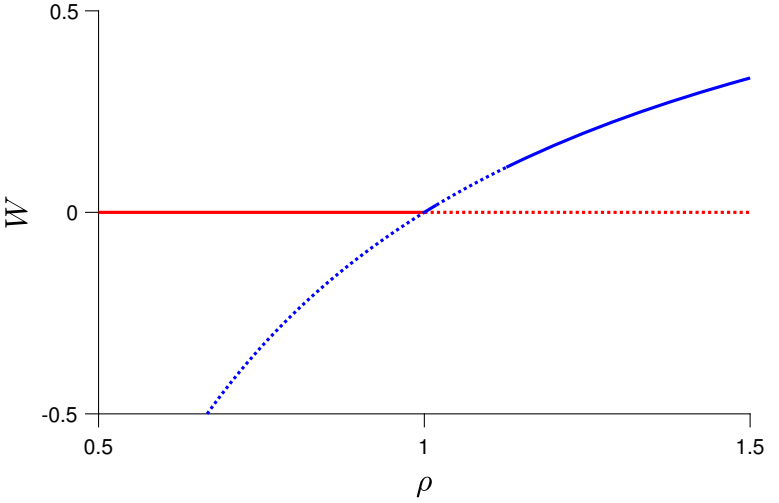


FIGURE 5. Transcritical bifurcation in amenities, shown here for $\gamma = 10^{-2}$. Red and blue lines represent homogenous destitution and homogeneous wealth, respectively; solid and dotted lines denote regions of stable and unstable solutions, respectively. Note that this figure can be seen as a subset of Figure 4 taken along the line $\gamma = 10^{-2}$.

4. Sensitivity analysis

In this section we perform a sensitivity analysis for the steady-state of the model in one dimensional space via a direct method introduced by Dickinson (1976).[**13**]. The steady-state solutions provide the final distribution and will tell us where wealth hotspots form. Here, we describe the results of both a linear and a quadratic sensitivity analysis. The former will tell us which parameters have the greatest effect on the solutions when we perturb their values. The latter tells us more about the nonlinear relationship between parameters. The insight obtained through this analysis is useful in determining the key parameters in the data fitting process.

Note that the emergence of patterns in our solutions will depend on the size of the domain. Therefore, we are also interested in how sensitive our solutions are to changes in the domain. For this reason, we introduce a new parameter to (3.6) to scale the domain. In particular, let $\tilde{t} = \alpha t$ and $\tilde{x} = \sqrt{\alpha}x$. Then after dropping the tilde notation, (3.6) becomes

$$\begin{aligned} \partial_t W &= \partial_{xx} W - 2\partial_x \left(W \partial_x \log A\right) + \alpha(\rho(1-W)W - W), \\ \partial_t A &= \partial_{xx} A + \alpha(-A + W + \gamma) \end{aligned} \tag{4.1}$$

for one dimensional $x \in [-20, 20]$ and $t \in [0, \infty)$. For this section we consider no-flux boundary conditions

$$\begin{aligned} \partial_x W - 2W\partial_x \log A|_{x=\pm 0.5} &= 0, \\ \partial_x A|_{x=+0.5} &= 0, \end{aligned} \tag{4.2}$$

and the following initial conditions:

$$\begin{aligned} W(x,0) &= 0.2 + 0.01\sin^2(16\pi x), \\ A(x,0) &= 0.2 - 0.01\sin^2(16\pi x), \end{aligned} \tag{4.3}$$

which were chosen since realistic solutions have hotspots.

To apply the direct method to the steady-state of (4.1), we set the time derivatives equal to zero and transform the resulting system into a system of four first order ordinary differential equations (ODE). Consider the following change of variables:

$$\begin{cases} Y_1(x) = W(x), \\ Y_2(x) = Y_1'(x) - 2\dfrac{Y_1(x)Y_4(x)}{Y_3(x)}, \\ Y_3(x) = A(x), \\ Y_4(x) = Y_3'(x). \end{cases} \tag{4.4}$$

This transformation leads to the following system of ODEs:

$$\begin{cases} Y_1'(x) = Y_2(x) + 2\frac{Y_1(x)Y_4(x)}{Y_3(x)}, \\ Y_2'(x) = -\alpha\left(\rho\left(1 - Y_1(x)\right)Y_1(x) - Y_1(x)\right), \\ Y_3'(x) = Y_4(x), \\ Y_4'(x) = \alpha\left(Y_3(x) - Y_1(x) - \gamma\right) \end{cases} \tag{4.5}$$

with boundary conditions $Y_2(\pm 20) = 0$ and $Y_4(\pm 20) = 0$. Define

$$\boldsymbol{\theta} = \begin{bmatrix} \rho & \gamma & \alpha \end{bmatrix}^T$$

to be the vector of unknown parameters. A parameter, $\theta_j \in \boldsymbol{\theta}$, is considered to be sensitive if small changes in θ_j lead to large changes in the solution. For any parameter θ_j and any Y_i we can define

$$Z_{Y_i}^{\theta_j}(x) := \frac{\partial Y_i(x)}{\partial \theta_j} \quad \text{for } i = 1,2,3,4 \text{ and } j = 1,2,3, \tag{4.6}$$

which we can think of as a measure of the sensitivity of Y_i with respect to parameter θ_j.

For each θ_j we can derive a system of ODEs for $Z_{Y_i}^{\theta_j}$. The details of this derivation are given in Appendix A. We then solve the system of eight ODEs defined in (4.5) and (A.1) for

$$\left[Y_1(x) \quad Y_2(x) \quad Y_3(x) \quad Y_4(x) \quad Z_{Y_1}^{\theta_j}(x) \quad Z_{Y_2}^{\theta_j}(x) \quad Z_{Y_3}^{\theta_j}(x) \quad Z_{Y_4}^{\theta_j}(x)\right]^T. \tag{4.7}$$

Note that the boundary conditions for (A.1) follow from the boundary conditions for (4.5) since

$$Z_{Y_2}^{\theta_j}(\pm 20) = \frac{\partial}{\partial \theta_j} Y_2(\pm 20) = 0,$$
$$Z_{Y_4}^{\theta_j}(\pm 20) = \frac{\partial}{\partial \theta_j} Y_4(\pm 20) = 0.$$

The system of ODEs described in (4.7) was solved using MATLAB. First, since the boundary value problem solver, *bvp5c*, requires initialization, we obtained an approximate steady-state solution by solving the time-dependent system, (4.1), numerically and scaling t. Namely, we made the substitution $\tau/t_c = t$ for $0 < t_c \ll 1$ so that we could achieve an approximate steady-state solution for $\tau \in [0,1]$. We used *pdepe* to obtain this solution, (W^*, A^*) and then used the solution for W^* and A^* at the final time step as the initial guess for *bvp5c* to solve for the solution to the system of ODEs in (4.7).

4.1. Linear and quadratic sensitivity. After solving the system of ODEs described above, we can compute both the linear and quadratic sensitivity. The linear sensitivity of $Y_i(x)$ with respect to parameter θ_j is defined as the absolute value of

$$\ell_{\theta_j}(x) = \frac{\theta_j}{Y_i(x)} Z_{Y_i}^{\theta_j}(x) \tag{4.8}$$

so that a larger $|\ell_{\theta_j}(x)|$ implies that θ_j is more sensitive. Note that we choose ℓ_{θ_j} to denote linear sensitivity, but this is different from ℓ in Section 2, which represents lattice spacing. We are also interested in the second order sensitivity in addition to the linear sensitivity because it will tell us more about the nonlinear relationship between parameters and their effects on the solutions. The quadratic sensitivity of $Y_i(x)$ with respect to parameters θ_k and θ_j is defined as the absolute value of

$$q_{\theta_k \theta_j}(x) = \frac{\theta_k \theta_j}{Y_i(x)} \frac{\partial^2 Y_i(x)}{\partial \theta_k \partial \theta_j}. \tag{4.9}$$

For the details on the derivations of (4.8) and (4.9), see Appendix B.

Figures 6 and 7 show the results of the first order sensitivity analysis for the steady-state of the wealth solution, $W^*(x) = Y_1(x)$, and the amenities solution, $A^*(x) = Y_3(x)$ respectively. We observe that for ρ near 1.001, γ near 10^{-8}, and

α near 10, both W^* and A^* are most sensitive to changes in ρ and least sensitive to changes in γ. Also, we see that ρ is less sensitive at places where wealth and amenities are lowest; meanwhile, α and γ are least sensitive where wealth and amenities are concentrated. Recall from Table 2, that ρ represents the ratio of investment to decay, so we can conclude that a large change in this ratio will lead to larger changes in the steady-state solutions. On the other hand, the parameter γ represents the amenities growth rate, so changes in the growth rate of amenities have the smallest effect on solutions.

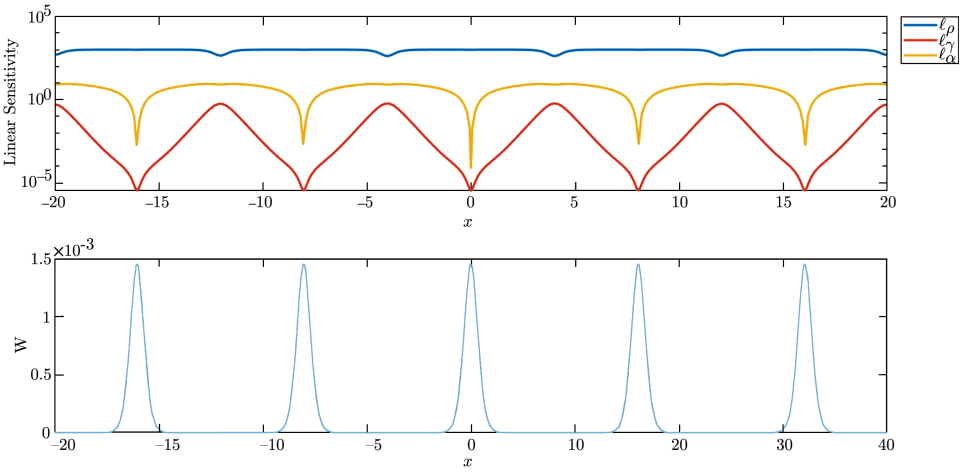


FIGURE 6. The top figure shows the semi-log plots of the linear sensitivity, $|\ell_{\theta_j}|\ \forall \theta_j \in \boldsymbol{\theta}$, for wealth computed for each parameter when $\rho = 1.001$, $\gamma = 10^{-8}$, and $\alpha = 10$.

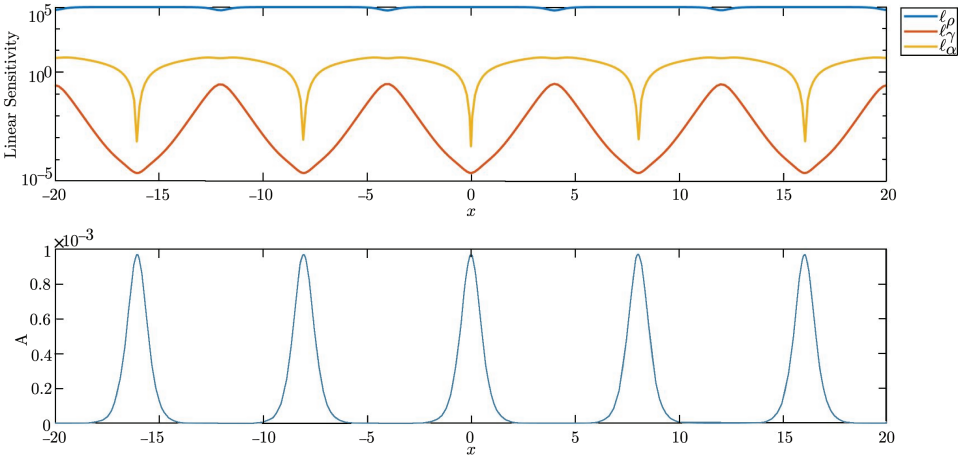


FIGURE 7. The top figure shows the semi-log plots of the linear sensitivity, $|\ell_{\theta_j}|\ \forall \theta_j \in \boldsymbol{\theta}$, for amenities computed for each parameter when $\rho = 1.001$, $\gamma = 10^{-8}$, and $\alpha = 10$.

Figures 8 and 9 show the results after we computed the coefficients for the second order terms. Notice that the coefficients involving ρ tend to be larger and the coefficients involving γ are smaller, which is consistent with our conclusion from the linear sensitivity analysis that solutions are most sensitive to changes in ρ and least sensitive to changes in γ. Also, we see that $q_{\rho\alpha}(x)$ tends to be the largest. In addition, despite ρ being the most sensitive in the linear results, we can see that $q_{\gamma\alpha}(x)$ tends to be larger than $q_{\rho\gamma}(x)$. This suggests that α is more sensitive than implied by the linear sensitivity analysis.

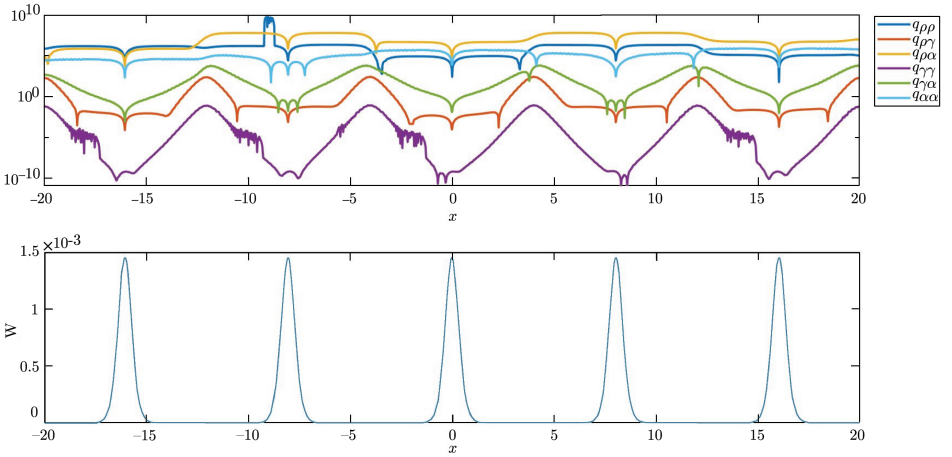


FIGURE 8. The top graph is a plot of the quadratic sensitivities, $|q_{\theta_j,\theta_k}|\ \forall\theta_j,\theta_k \in \boldsymbol{\theta}$, in wealth for each second order term with $(\rho,\gamma,\alpha) = (1.001, 1e-8, 10)$ and $\Delta\theta_j = 1e-8$ for each parameter $\theta_j \in \boldsymbol{\theta}$. The bottom plot is the wealth solution.

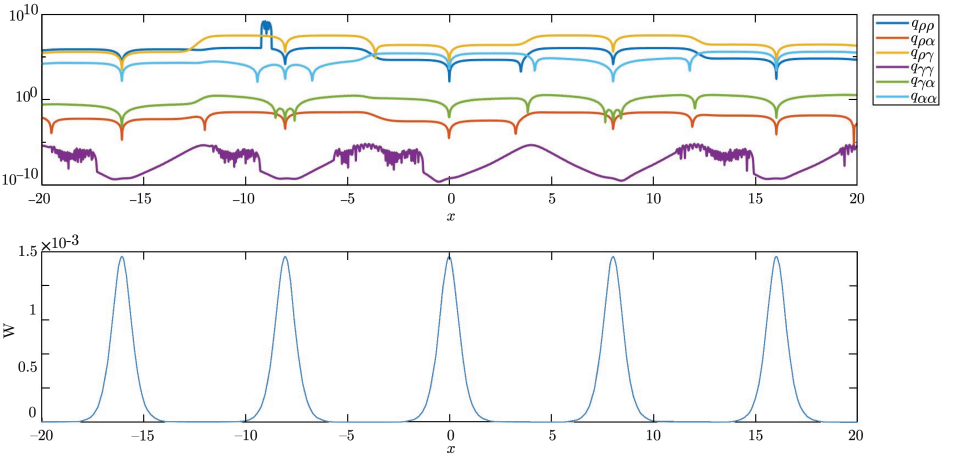


FIGURE 9. The top graph is a plot of the quadratic sensitivities, $|q_{\theta_j,\theta_k}|\ \forall\theta_j,\theta_k \in \boldsymbol{\theta}$, in amenities for each second order term with $(\rho,\gamma,\alpha) = (1.001, 1e-8, 10)$ and $\Delta\theta_j = 1e-8$ for each parameter $\theta_j \in \boldsymbol{\theta}$. The bottom plot is the amenities solution.

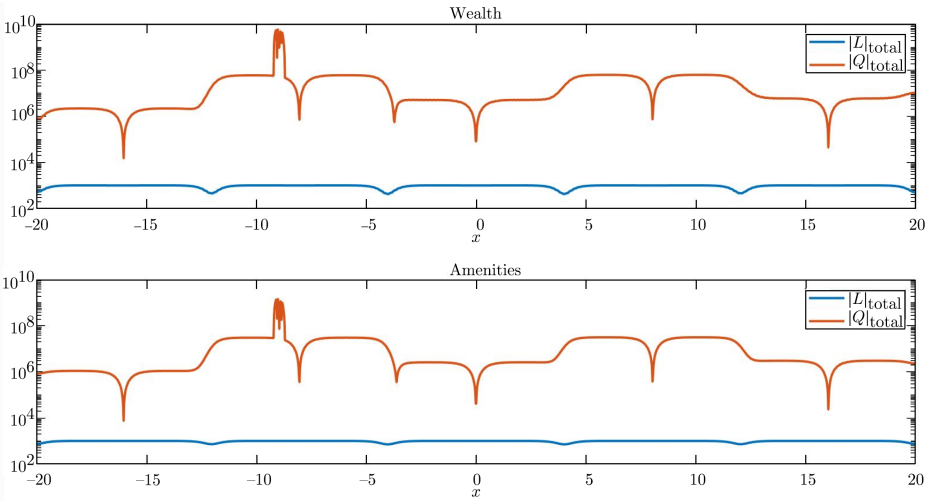


FIGURE 10. Plots of $|L|_{\text{total}}$ defined in (4.10) and $|Q|_{\text{total}}$ defined in (4.11) for the wealth solution (top) and the amenities solution (bottom) with $(\rho, \gamma, \alpha) = (1.001, 1e-8, 10)$ and $\Delta\theta_j = 1e-8$ for each parameter in $\theta_j \in \boldsymbol{\theta}$.

Observe that results for linear and quadratic sensitivity in both wealth and amenities solutions are similar. This similarity is expected since these solutions are cooperative by design. When the wealth solution changes, the amenities solution must also change in a similar way since areas with higher wealth are expected to have more amenities while areas with lower wealth are expected to have fewer amenities.

To understand the relative error, in Figure 10 we consider

$$|L(x)|_{\text{total}} = \sum_{\theta_j \in \boldsymbol{\theta}} |\ell_{\theta_j}(x)| \tag{4.10}$$

and

$$|Q(x)|_{\text{total}} = \sum_{\theta_j, \theta_k \in \boldsymbol{\theta}} |q_{\theta_j, \theta_k}(x)|. \tag{4.11}$$

In particular, we know

$$\left| \frac{\Delta Y_i}{Y_i} \right| \leq |L(x)|_{\text{total}} + |Q(x)|_{\text{total}}.$$

See Appendix C for details on the derivation of (4.10) and (4.11). Since Figure 10 shows that $|Q|_{\text{total}}$ is significantly larger than $|L|_{\text{total}}$ for both the wealth and amenities solutions, we can observe that the error in the linear sensitivity analysis is significant. This implies that sensitivity is nonlinear at every $x \in [-20, 20]$ and it confirms that the extra computations we performed for the quadratic sensitivity analysis were important to include.

5. Discussion

Starting from basic sociological assumptions surrounding the spread of gentrification, we derive a discrete model of this phenomenon. Namely, we consider a system wherein wealth in a community begets an increase in community investment, and vice versa. To this end we define the amenities of a community to be a measure of the results of this investment; these amenities may include various factors such as coffee shops, parks, tapas restaurants, and local festivals and events.

We argue that these amenities both diffuse spatially over time and decay temporally if not maintained; reinvestments to maintain these amenities are proportional to the amount of wealth at a given time. In particular, we focused on a logistic rate of reinvestment; future analyses may consider the effects of different such rates. Similarly, wealth travels up amenities gradients and concentrates in areas of high amenities. From this discrete empirical model, justified on sociological observations, we derive a continuous model; the resultant system of partial differential equations is of the general form of a reaction-diffusion system. For corresponding parameters we observe similar dynamics in the discrete and continuous model, and the homogeneous solutions of both discrete and continuum models are identical, in their respective parameter spaces.

In both discrete and continuous models, the interplay between wealth and amenities creates a feedback loop that, for certain parameter regimes, leads to hotspot formation evocative of those observed in true gentrified neighborhoods. For the logistic rate of reinvestment under consideration, this regime is $1 < \frac{r}{\omega} < 4-2\sqrt{2}$ under the assumption that there is no wealth-independent reinvestment in amenities. By recognizing the term $\frac{r}{\omega}$ as the ratio of investment and decay rates, we are able to qualitatively interpret the relevant regimes of this ratio. For $0 < \frac{r}{\omega} < 1$, decay dominates reinvestment throughout our domain and wealth vanishes throughout for large time scales. For $\frac{r}{\omega} > 4 - 2\sqrt{2}$, reinvestment sufficiently overcomes decay so that wealth approaches the solution $1 - \frac{\omega}{r}$ in a spatially homogeneous fashion. The regime $1 < \frac{r}{\omega} < 4 - 2\sqrt{2}$ begets wealth hotspot formation for γ small; certain pockets of the domain are able to maintain wealth but are surrounded by large areas of destitution.

We have succeeded in designing a model that exhibits qualitative similarities with empirical observations; areas of future work may involve analysis of the efficacy of our model in mirroring actual gentrification. The difficulty in devising effective measures and compiling empirical data sets is an immediate obstacle to accomplishing this; this goal is additionally complicated by the wide variety of underlying factors of gentrification. We have unified these in the "amenities" metric but it is unclear what exactly these factors are and how significant of a role each plays in combining to form amenities.

Despite this, a refined model of gentrification—tuned to empirical data—can be an invaluable tool in both the urban planning of the public sector and the expansion strategies of private entities. The ability to accurately predict the results of changes in investment and policy has broad implications and would allow for more efficient distributions of resources. This model serves as a basis to accomplish this task.

Splitting amenities into public policy driven amenities and private investment driven amenities is another avenue worth exploring; the contrast in effective time scales between the two make this a natural division. The former would act on longer time scales and concentrate mainly on areas of low wealth while the latter

would play out in shorter periods of time and invest mainly in areas where high returns on investment would be expected.

Another natural area of further inquiry would be to analyze different forms of the rate of reinvestment $f(W_s, A_s)$. Modifications of $f(W_s, A_s)$ would lead to different equilibrium solutions and sets of dynamics and could better factor in the myriad of causes of gentrification in actual cities, as well as the unique domains of particular urban areas.

Lastly, our sensitivity analysis determined the most sensitive parameters in the system. Namely, we saw that perturbations in ρ, the relative rate of investment and decay, would lead to the largest changes in solutions. On the other hand, changes in γ, the amenities growth rate, have a much smaller effect on the system. Future work could involve utilizing these results to fit parameter values to data.

In summary, our model of gentrification effectively models the creation of wealth hotspots seen in actual cities. Our ideas can serve as a basis for further inquiry into the factors leading to gentrification, which continue to be elusive. Better understanding of these factors can shape public policy to better serve those effected by this sociological phenomenon.

Appendix A

To calculate the ODEs for $Z_{Y_i}^{\theta_j}(x)$, note that

$$\frac{d}{dx}\left(Z_{Y_i}^{\theta_j}(x)\right) = \frac{\partial}{\partial x}\left(\frac{\partial Y_i(x)}{\partial \theta_j}\right) = \frac{\partial}{\partial \theta_j}\left(\frac{\partial Y_i(x)}{\partial x}\right).$$

Define $f_i(x)$ for $i = 1, 2, 3, 4$ to be the right-hand side of the ode for $Y_i(x)$ defined in (4.5). Then since $Y_i'(x) = f_i(Y_1(\boldsymbol{\theta}), Y_2(\boldsymbol{\theta}), Y_3(\boldsymbol{\theta}), Y_4(\boldsymbol{\theta}), \boldsymbol{\theta})$, we have

$$\begin{aligned}\frac{d}{dx}Z_{Y_i}^{\theta_j}(x) &= \frac{\partial f_i}{\partial \theta_j} + \frac{\partial f_i}{\partial Y_1}\frac{\partial Y_1}{\partial \theta_j} + \frac{\partial f_i}{\partial Y_2}\frac{\partial Y_2}{\partial \theta_j} + \frac{\partial f_i}{\partial Y_3}\frac{\partial Y_3}{\partial \theta_j} + \frac{\partial f_i}{\partial Y_4}\frac{\partial Y_4}{\partial \theta_j} \\ &= \frac{\partial f_i}{\partial \theta_j} + \frac{\partial f_i}{\partial Y_1}Z_{Y_1}^{\theta_j}(x) + \frac{\partial f_i}{\partial Y_2}Z_{Y_2}^{\theta_j}(x) + \frac{\partial f_i}{\partial Y_3}Z_{Y_3}^{\theta_j}(x) + \frac{\partial f_i}{\partial Y_4}Z_{Y_4}^{\theta_j}(x).\end{aligned}$$

If we define

$$\mathbf{f}(x) = \begin{bmatrix} f_1(x) & f_2(x) & f_3(x) & f_4(x) \end{bmatrix}^T$$

and

$$\boldsymbol{Z}^{\boldsymbol{\theta_j}}(x) = \begin{bmatrix} Z_{Y_1}^{\theta_j}(x) & Z_{Y_2}^{\theta_j}(x) & Z_{Y_3}^{\theta_j}(x) & Z_{Y_4}^{\theta_j}(x) \end{bmatrix}^T,$$

then for each θ_j, we have a system of four ODEs for $Z_i^{\theta_j}(x)$ given by

$$\frac{d}{dx}\boldsymbol{Z}^{\boldsymbol{\theta_j}}(x) = \frac{\partial}{\partial \theta_j}\mathbf{f}(x) + J(x)\boldsymbol{Z}^{\boldsymbol{\theta_j}}(x), \tag{A.1}$$

where $J(x)$ is the 4×4 Jacobian matrix with entries $J_{ik} = \partial f_i / \partial Y_k$. In particular,

$$J(x) = \begin{bmatrix} 2\frac{Y_4}{Y_3} & 1 & -2\frac{Y_1 Y_4}{Y_3^2} & 2\frac{Y_1}{Y_3} \\ -\alpha\left(\rho(1-2Y_1)\right) - 1) & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -\alpha & 0 & \alpha & 0 \end{bmatrix}.$$

In addition,

$$\frac{\partial \mathbf{f}}{\partial \rho} = \begin{bmatrix} 0 & -\alpha Y_1(1-Y_1) & 0 & 0 \end{bmatrix}^T,$$
$$\frac{\partial \mathbf{f}}{\partial \gamma} = \begin{bmatrix} 0 & 0 & 0 & -\alpha \end{bmatrix}^T,$$
$$\frac{\partial \mathbf{f}}{\partial \alpha} = \begin{bmatrix} 0 & -\rho(1-Y_1)Y_1 + Y_1 & 0 & Y_3 - Y_1 - \gamma \end{bmatrix}^T.$$

Appendix B

To write the total sensitivity of Y_i, we can use a Taylor expansion of $Y_i(\boldsymbol{\theta})$ to obtain

$$\Delta Y_i = Z^{\rho}_{Y_i}(x)\Delta\rho + Z^{\gamma}_{Y_1}(x)\Delta\gamma + Z^{\alpha}_{Y_1}(x)\Delta\alpha + \mathcal{O}\left(\left(\max_j \Delta\theta_j\right)^2\right).$$

Then the relative change in Y_i corresponding to the relative parameter change is

$$\frac{\Delta Y_i}{Y_i} = \left(\frac{\rho}{Y_i}Z^{\rho}_{Y_i}(x)\right)\frac{\Delta\rho}{\rho} + \left(\frac{\gamma}{Y_i}Z^{\gamma}_{Y_i}(x)\right)\frac{\Delta\gamma}{\gamma} + \left(\frac{\alpha}{Y_i}Z^{\alpha}_{Y_i}(x)\right)\frac{\Delta\alpha}{\alpha}.$$

Therefore, the linear sensitivity of Y_i with respect to parameter θ_j is defined as the absolute value of the coefficients,

$$\ell_{\theta_j}(x) = \frac{\theta_j}{Y_i(x)}Z^{\theta_j}_{Y_i}(x).$$

Similar to the linear sensitivity, we can compute the coefficients for quadratic sensitivities; however, we now need to approximate the second order derivatives of W and A with respect to the parameters. After obtaining a quadratic expansion of Y_i, dividing both sides by Y_i, and dividing and multiplying each term on the right-hand side by the appropriate parameters, we have

$$\begin{aligned}\frac{\Delta Y_i}{Y_i} &= \ell_\rho(x)\frac{\Delta\rho}{\rho} + \ell_\gamma(x)\frac{\Delta\gamma}{\gamma} + \ell_\alpha(x)\frac{\Delta\alpha}{\alpha} \\ &\quad + \frac{1}{2}q_{\rho\rho}(x)\left(\frac{\Delta\rho}{\rho}\right)^2 + \frac{1}{2}q_{\gamma\gamma}(x)\left(\frac{\Delta\gamma}{\gamma}\right)^2 + \frac{1}{2}q_{\alpha\alpha}(x)\left(\frac{\Delta\alpha}{\alpha}\right)^2 \\ &\quad + q_{\rho\gamma}(x)\left(\frac{\Delta\rho}{\rho}\right)\left(\frac{\Delta\gamma}{\gamma}\right) + q_{\rho\alpha}(x)\left(\frac{\Delta\rho}{\rho}\right)\left(\frac{\Delta\alpha}{\alpha}\right) + q_{\gamma\alpha}(x)\left(\frac{\Delta\gamma}{\gamma}\right)\left(\frac{\Delta\alpha}{\alpha}\right) \\ &\quad + \mathcal{O}\left(\left(\max_j \Delta\theta_j\right)^3\right),\end{aligned} \tag{B.1}$$

where

$$q_{\theta_k\theta_j}(x) = \frac{\theta_k\theta_j}{Y_i}\frac{\partial^2 Y_i}{\partial\theta_k\partial\theta_j}.$$

Using the work from the computation of linear sensitivities, we can compute these coefficients once we approximate the second order derivatives. Suppose we want to compute the second order derivative of Y_i with respect to, for example, ρ and γ. Then we can approximate this with the following calculation:

$$\frac{\partial^2 Y_i}{\partial\rho\partial\gamma} = \frac{\partial}{\partial\rho}\frac{\partial Y_i}{\partial\gamma} = \frac{\partial Z^{\gamma}_{Y_i}}{\partial\rho} \approx \frac{Z^{\gamma}_{Y_i}|_{\rho=\rho_0+\Delta\rho} - Z^{\gamma}_{Y_i}|_{\rho=\rho_0}}{\Delta\rho}.$$

Appendix C

Let $L(x)$ represent the linear part of the right-hand side of (B.1) and let $Q(x)$ represent the quadratic part of (B.1). That is,

$$L(x) = \ell_\rho(x)\frac{\Delta\rho}{\rho} + \ell_\gamma(x)\frac{\Delta\gamma}{\gamma} + \ell_\alpha(x)\frac{\Delta\alpha}{\alpha},$$

and

$$\begin{aligned} Q(x) = {} & \frac{1}{2}q_{\rho\rho}(x)\left(\frac{\Delta\rho}{\rho}\right)^2 + \frac{1}{2}q_{\gamma\gamma}(x)\left(\frac{\Delta\gamma}{\gamma}\right)^2 + \frac{1}{2}q_{\alpha\alpha}(x)\left(\frac{\Delta\alpha}{\alpha}\right)^2 \\ & + q_{\rho\gamma}(x)\left(\frac{\Delta\rho}{\rho}\right)\left(\frac{\Delta\gamma}{\gamma}\right) + q_{\rho\alpha}(x)\left(\frac{\Delta\rho}{\rho}\right)\left(\frac{\Delta\alpha}{\alpha}\right) + q_{\gamma\alpha}(x)\left(\frac{\Delta\gamma}{\gamma}\right)\left(\frac{\Delta\alpha}{\alpha}\right). \end{aligned}$$

Then we define $|L(x)|$ and $|Q(x)|$ to be the relative linear error and relative quadratic error respectively. For sufficiently small θ_j, we know $\Delta\theta_j/\theta_j \le 1$ for all $\theta_j \in \boldsymbol{\theta}$. Thus,

$$\begin{aligned} |L(x)| &\le |\ell_\rho(x)|\left|\frac{\Delta\rho}{\rho}\right| + |\ell_\gamma(x)|\left|\frac{\Delta\gamma}{\gamma}\right| + |\ell_\alpha(x)|\left|\frac{\Delta\alpha}{\alpha}\right| \\ &\le \sum_{\theta_j\in\boldsymbol{\theta}} |\ell_{\theta_j}(x)|, \end{aligned}$$

and

$$\begin{aligned} |Q(x)| \le {} & \frac{1}{2}|q_{\rho\rho}(x)|\left|\frac{\Delta\rho}{\rho}\right|^2 + \frac{1}{2}|q_{\gamma\gamma}(x)|\left|\frac{\Delta\gamma}{\gamma}\right|^2 + \frac{1}{2}|q_{\alpha\alpha}(x)|\left|\frac{\Delta\alpha}{\alpha}\right|^2 \\ & + |q_{\rho\gamma}(x)|\left|\frac{\Delta\rho}{\rho}\right|\left|\frac{\Delta\gamma}{\gamma}\right| + |q_{\rho\alpha}(x)|\left|\frac{\Delta\rho}{\rho}\right|\left|\frac{\Delta\alpha}{\alpha}\right| + |q_{\gamma\alpha}(x)|\left|\frac{\Delta\gamma}{\gamma}\right|\left|\frac{\Delta\alpha}{\alpha}\right| \\ \le {} & \sum_{\theta_j,\theta_k\in\boldsymbol{\theta}} |q_{\theta_j,\theta_k}(x)|. \end{aligned}$$

Therefore, we consider

$$|L(x)|_{\text{total}} = \sum_{\theta_j\in\boldsymbol{\theta}} |\ell_{\theta_j}(x)|,$$

and

$$|Q(x)|_{\text{total}} = \sum_{\theta_j,\theta_k\in\boldsymbol{\theta}} |q_{\theta_j,\theta_k}(x)|,$$

$$\left|\frac{\Delta Y_i}{Y_i}\right| \le |L(x)|_{\text{total}} + |Q(x)|_{\text{total}}.$$

Acknowledgments

The authors would like to thank an anonymous reviewer whose comments were very valuable.

References

[1] Robert A. Beauregard, *The chaos and complexity of gentrification*, Chap. 3 in Gentrification of the city (Neil Smith and Peter Williams, eds.), Allen & Unwin, 1986, pp. 35–55.
[2] N. Bellomo, A. Bellouquid, J. Nieto, and J. Soler, *Multicellular biological growing systems: hyperbolic limits towards macroscopic description*, Math. Models Methods Appl. Sci. **17** (2007), no. suppl., 1675–1692, DOI 10.1142/S0218202507002431. MR2362760
[3] Stephen Benard and Robb Willer, *A wealth and status-based model of residential segregation*, Mathematical Sociology **31** (2007), no. 2, 149–174.
[4] H. Berestycki and J.-P. Nadal, *Self-organised critical hot spots of criminal activity*, European J. Appl. Math. **21** (2010), no. 4-5, 371–399, DOI 10.1017/S0956792510000185. MR2671615
[5] Yonatan Berman, *Bistable logistic dynamics of wealth inequality*, 2016.
[6] Gary Bridge, Tim Butler, and Loretta Lees, *Mixed communities: Gentrification by stealth?*, Policy Press, 2012.
[7] Elizabeth E. Bruch and Robert D. Mare, *Neighborhood choice and neighborhood change*, American Journal of sociology **112** (2006), no. 3, 667–709.
[8] Karen Chapple and Rick Jacobus, *Retail trade as a route to neighborhood revitalization*, Urban and regional policy and its effects **2** (2009), 19–68.
[9] Camille Zubrinsky Charles, *The dynamics of racial residential segregation*, Annual review of sociology **29** (2003), no. 1, 167–207.
[10] Victor Couture and Jessie Handbury, *Urban revival in America, 2000 to 2010*, Technical report, National Bureau of Economic Research, 2017.
[11] Mark C. Cross and Pierre C. Hohenberg, *Pattern formation outside of equilibrium*, Reviews of modern physics **65** (1993), no. 3, 851.
[12] Samuel Dastrup and Ingrid Gould Ellen, *Linking residents to opportunity: Gentrification and public housing*, Cityscape **18** (2016), no. 3, 87.
[13] Robert P. Dickinson and Robert J. Gelinas, *Sensitivity analysis of ordinary differential equation systems—a direct method*, J. Comput. Phys. **21** (1976), no. 2, 123–143, DOI 10.1016/0021-9991(76)90007-3. MR474829
[14] Y. Dolak and C. Schmeiser, *Kinetic models for chemotaxis: hydrodynamic limits and spatio-temporal mechanisms*, J. Math. Biol. **51** (2005), no. 6, 595–615, DOI 10.1007/s00285-005-0334-6. MR2213630
[15] Lena Edlund, Cecilia Machado, and Maria Micaela Sviatschi, *Bright minds, big rent: gentrification and the rising returns to skill*, Technical report, National Bureau of Economic Research, 2015.
[16] Ingrid Gould Ellen, Keren Mertens Horn, and Davin Reed, *Has falling crime invited gentrification?*, Technical report, Center for Economic Studies, U.S. Census Bureau, 2017.
[17] Radek Erban and Hans G. Othmer, *From individual to collective behavior in bacterial chemotaxis*, SIAM J. Appl. Math. **65** (2004/05), no. 2, 361–391, DOI 10.1137/S0036139903433232. MR2123062
[18] Claude S. Fischer, Gretchen Stockmayer, Jon Stiles, and Michael Hout, *Distinguishing the geographic levels and social dimensions of us metropolitan segregation, 1960–2000*, Demography **41** (2004), no. 1, 37–59.
[19] Richard Fry and Paul Taylor, *The rise of residential segregation by income*, Washington, DC: Pew Research Center (2012), no. 202, 26.
[20] Chris Hamnett, *The blind men and the elephant: The explanation of gentrification*, Transactions of the Institute of British Geographers (1991), 173–189.
[21] A. Hasan, N. Rodríguez, and L. Wong, *Transport and concentration of wealth: Modeling an amenities-based-theory*, Chaos **30** (2020), no. 5, 053110, 16, DOI 10.1063/5.0003767. MR4094692
[22] Derek Hyra, *Commentary: Causes and consequences of gentrification and the future of equitable development policy*, Cityscape **18** (2016), no. 3, 169.

[23] Alexander J. Laurie and Narendra K. Jaggi, *Role of 'vision' in neighbourhood racial segregation: A variant of the schelling segregation model*, Urban Studies **40** (2003), no. 13, 2687–2704.
[24] Loretta Lees, *In the pursuit of difference: Representations of gentrification*, Environment and planning A **28** (1996), no. 3, 453–470.
[25] M. A. Lewis, K. A. J. White, and J. D. Murray, *Analysis of a model for wolf territories*, J. Math. Biol. **35** (1997), no. 7, 749–774, DOI 10.1007/s002850050075. MR1479337
[26] Margot Lutzenhiser and Noelwah R. Netusil, *The effect of open spaces on a home's sale price*, Contemporary Economic Policy **19** (2001), no. 3, 291–298.
[27] Anotida Madzvamuse, Raquel Barreira, and Alf Gerisch, *Cross-diffusion in reaction-diffusion models: analysis, numerics, and applications*, Progress in industrial mathematics at ECMI 2016, Math. Ind., vol. 26, Springer, Cham, 2017, pp. 385–392, DOI 10.1007/978-3-319-63082-3. MR3846615
[28] Douglas S. Massey and Nancy A. Denton, *American apartheid: Segregation and the making of the underclass*, Harvard University Press, 1993.
[29] Daniel Monroe Sullivan and Samuel C. Shaw, *Retail gentrification and race: The case of Alberta Street in Portland, Oregon*, Urban Affairs Review **47** (2011), no. 3, 413–432.
[30] Romans Pancs and Nicolaas J. Vriend, *Schelling's spatial proximity model of segregation revisited*, Journal of Public Economics **91** (2007), no. 1-2, 1–24.
[31] Damaris Rose, *Rethinking gentrification: Beyond the uneven development of Marxist urban theory*, Environment and planning D: Society and Space **2** (1984), no. 1, 47–74.
[32] Thomas C. Schelling, *Dynamic models of segregation*, Journal of mathematical sociology **1** (1971), no. 2, 143–186.
[33] Thomas C. Schelling, *Micromotives and macrobehavior*, WW Norton & Company, 2006.
[34] M. B. Short, M. R. D'Orsogna, V. B. Pasour, G. E. Tita, P. J. Brantingham, A. L. Bertozzi, and L. B. Chayes, *A statistical model of criminal behavior*, Math. Models Methods Appl. Sci. **18** (2008), no. suppl., 1249–1267, DOI 10.1142/S0218202508003029. MR2438215
[35] Austin Troy and J. Morgan Grove, *Property values, parks, and crime: A hedonic analysis in Baltimore, MD*, Landscape and urban planning **87** (2008), no. 3, 233–245.
[36] Jacob L. Vigdor, *Is urban decay bad? is urban revitalization bad too?*, Journal of Urban Economics **68** (2010), no. 3, 277–289.
[37] Richard Voith, *Changing capitalization of CBD-oriented transportation systems: Evidence from Philadelphia, 1970–1988*, Journal of Urban Economics **33** (1993), no. 3, 361–376.
[38] Junfu Zhang, *Residential segregation in an all-integrationist world*, Journal of Economic Behavior & Organization **54** (2004), no. 4, 533–550.
[39] Miriam Zuk, Ariel H. Bierbaum, Karen Chapple, Karolina Gorska, Anastasia Loukaitou-Sideris, Paul Ong, and Trevor Thomas, *Gentrification, displacement and the role of public investment: A literature review*, In Federal Reserve Bank of San Francisco, 2015.

Department of Mathematics, University of California, Davis, Davis, California 95616

Department of Mathematics, University of Utah, Salt Lake City, Utah 84112

Applied Mathematics, University of Colorado Boulder, Boulder, Colorado 80309
Email address: rodrign@colorado.edu

Department of City and Regional Planning, University of North Carolina Chapel Hill, Chapel Hill, North Carolina 27599

Applied Mathematics, University of Colorado Boulder, Boulder, Colorado 80309

Proceedings of Symposia in Applied Mathematics
Volume **80**, 2025
https://doi.org/10.1090/psapm/080/00728

A non-expert's introduction to data ethics for mathematicians

Mason A. Porter

This chapter is dedicated to my current and former PhD students. They mean the world to me.

Abstract. I give a short introduction to data ethics. I begin with some background information and societal context for data ethics. I then discuss data ethics in mathematical-science education and indicate some available course material. I briefly highlight a few efforts—at my home institution and elsewhere—on data ethics, society, and social good. I then discuss open data in research, research replicability and some other ethical issues in research, the tension between privacy and open data and code, and a few controversial studies and reactions to studies. I also discuss ethical principles, institutional review boards, and a few other considerations in the scientific use of human data. I then briefly survey a variety of research and lay articles that are relevant to data ethics and data privacy. I conclude with a brief summary and some closing remarks.

My focal audience is mathematicians, but I hope that this chapter will also be useful to others. I am not an expert about data ethics, and this chapter provides only a starting point on this wide-ranging topic. I encourage you to examine the resources that I discuss and to reflect carefully on data ethics, its role in mathematics education, and the societal implications of data and data analysis. As data and technology continue to evolve, I hope that such careful reflection will continue throughout your life.

Don't say that he's hypocritical
Say rather that he's apolitical
"Once the rockets are up, who cares where they come down?
That's not my department!" says Wernher von Braun.

(Tom Lehrer, *Wernher von Braun*, 1975)

1. Introduction

The use of digital data to examine and help understand human behavior is both powerful and dangerous [**Edi21**]. Every day, it seems like there is a new nightmare with problematic uses of data and algorithms. The use of predictive policing to identify criminal activity can exacerbate existing racial and ethnic

2020 *Mathematics Subject Classification.* Primary 97M70, 97B70; Secondary 97P80, 91C99, 91-11.

Key words and phrases. Data, ethics, social systems, human society, complex systems, mathematics, the mathematics profession.

inequities [**Nat18**], algorithmic social-credit ratings of individuals have frightening dystopian uses [**Wikc**], and other manifestations of "algocracy" (i.e., algorithmic government) have many human ramifications [**ET19**, **OBS+19**, **Has21**, **SMBM23**].

The ever-increasing use of and reliance on "big data" and the accelerating application of tools such as machine learning and artificial intelligence (AI) to make societally consequential decisions can cause very significant harm and exacerbate societal inequities [**O'N16**, **Ree18**]. Problems like algorithmic bias and the collection, measurement, and use of enormous amounts of data about humans and their behavior have significant societal consequences [**WSAO+21**, **LHF+21**, **SVW21**, **Bir21**, **WJ23**]. New technologies like "deep fakes" [**IPA20**], synthetic media in which a person in an existing image or video is replaced with somebody else's likeness,[1] have terrifying potential to cause harm. There are significant risks and potential nefarious uses of tools like generative AI and large language models (LLMs) [**BGMMS21**, **EMFG+22**, **Fer24**]. Modern data analysis also has been accompanied by the reincarnation of pseudosciences, such as physiognomy, which can now be performed at a large scale [**SH22**]. Along with such dangers come enormous potential benefits, and there are nascent efforts to create documents like an "AI Bill of Rights" to help guide the design and deployment of automated systems in a way that protects people [**Off22**].

The situation is already scary. In 2016, Microsoft released a Twitter bot that "learned" from its interactions with other Twitter accounts; in less than a day, it was regularly producing racist tweets [**Vin16**]. In December 2020, Stanford University's use of an algorithm to determine which people would receive the first batches of a COVID-19 vaccine resulted in an inequitable vaccination rollout that prioritized high-ranking doctors over frontline human-facing medical personnel [**Che20**]. The output of algorithms have also led to decisions to cut off medical care [**RH23**]. In 2021, the outgoing Editor-in-Chief of a scientific journal publicly released refereeing data, including the numbers of decisions to accept or reject papers by each of the journal's referees during his tenure [**Sch21**]. In 2023, the LLM ChatGPT invented[2] a sexual-harassment scandal and named a real professor in it [**VO23**] and Microsoft limited its Bing AI chatbot after its unsettling conversations [**Les23**]. In today's world, it often seems necessary to laugh to keep from crying [**Wei21**].

The mathematical, statistical, and computational sciences are interconnected with human communities and society at large [**Cam**, **Eth**, **MC23**, **Mos21**]. Neither our education nor our research occur in a vacuum, and many of our algorithms and other methods are now applied to social systems [**WSAO+21**, **CM23**]. However, it is not traditional in mathematical-science education to discuss data ethics and other ethical considerations [**RSMCM24**, **Mü24**]. What we choose to teach (and choose not to teach) impacts what our mentees do with their education. Consequently, it is crucial for mathematical scientists (and others) to think carefully about ethics and engage with it throughout their careers [**CM24**]. See [**LMP18**] for a short book on ethics and data science, [**FT16**] for a theme journal issue on data ethics, [**Vé21**] for a wide-ranging handbook of digital ethics, and [**MCF22**] for a recent

[1] A light-hearted example is Gollum's precious cover of the song "Nothing Compares to U" [**Dor20**].

[2] It seems that LLMs and generative AI tools often simply make things up.

discussion of the ethics landscape in mathematics and advocacy of a "Hippocratic Oath" for mathematics.

The present chapter, which I intend as an introductory resource about data ethics, is a companion to my oral presentation at the American Mathematical Society Short Course on Mathematical and Computational Methods for Complex Social Systems in January 2021 [**BFPV21**]. My slides and video presentation are available at [**Por21**]. I discuss many of the same key points in my presentation as in the present chapter, although my emphases often differ. I have also learned about more data-ethics resources in the last four years, and certain ethical issues and relevant technologies have gained increased prominence in that time. I encourage you to look at my slides and watch my presentation, and I especially encourage you to look at the resources that I discuss in those slides and in the present chapter. I am not an expert on data ethics—and, to be frank, writing this article has been accompanied by my most serious case of "imposter syndrome" in a long time—and it is important that you look at what actual experts have to say. I will attempt to give some helpful thoughts and pointers to begin a journey in data ethics. It is vital to continue to reflect on data ethics and the societal impact of data and data analysis throughout your life.

In this chapter, I cover a diverse variety of topics. I start with data ethics and education, which is important for almost the entire mathematical-science community. In Section 2, I discuss data ethics in the mathematics community and mathematical-science education. I highlight the importance of curricular data ethics in Section 2.1, and I point to existing educational material in Section 2.2. In Section 3, I describe some of our efforts at University of California, Los Angeles (UCLA) on data ethics and society. I also briefly mention a few efforts by others and point out an important caveat in efforts in "data for social good". In Section 4, I discuss various ethical issues and tensions in research. Some of these topics, such as research replication and appropriately acknowledging others, are directly relevant to everybody in the mathematical-science community. Other topics are most directly relevant to people who use data in their research, although it is still important for others in the mathematical sciences to have some familiarity with them. I discuss research replication and open data and code in Section 4.1, acknowledgements and licensing material for open use in Section 4.2, privacy concerns and its tension with open data and code in Section 4.3, and some controversies in studies and in reactions to studies in Section 4.4. In Section 4.5, I summarize the key ideas of Section 4. In Section 5, I discuss some ethical principles and other considerations in the use of human data in scientific research. In Section 6, I highlight a few research articles about issues in data ethics. I conclude in Section 7.

2. Data ethics, the mathematics community, and mathematical-science education

In this section, I advocate for data ethics in mathematical-science education and point to some existing course materials. Many students who obtain undergraduate or graduate degrees in the mathematical sciences will become data scientists or otherwise will work extensively with data. Therefore, data ethics needs to be a central part of their professional lives, and it is our responsibility as mathematics educators to mentor them to face these ethical challenges with careful reflection, wisdom, and humility.

2.1. The importance of data ethics in the mathematics community and mathematical-science curricula. The mathematical, statistical, and computational sciences are intertwined with human society [**Cam**, **Eth**, **MC23**]. This notion is not new. Throughout his life, educator and civil-rights activist Bob Moses championed mathematical literacy as a civil right for all public-school children, with particular advocacy for those who are most vulnerable [**MWW$^+$23**]. In a 2021 article [**Mos21**] that was published a month after his death, Moses wrote

> *In the 1960s, voting was our organizing tool to demolish Jim Crow and achieve political impact. Since then, for me, it has been algebra. What's math got to do with it?—you ask. Everything, I say.*
>
> *Amidst the planetwide transformation we are undergoing, from industrial to information-age economies and culture, math performance has emerged as a critical measure of equal opportunity.*

Mathematical research also does not occur in a vacuum, especially when we apply our algorithms and other methods to social systems [**CM23**]. It is not a traditional part of curricular education or research education in the mathematical sciences to teach our students and other junior community members about data ethics and other ethical considerations [**RSMCM24**, **Mü24**], but this needs to change.[3] What we choose to teach (and choose not to teach, or choose to mention in only a cursory way) impacts what our students and other mentees do with their education. Other disciplines (e.g., the social, behavioral, and medical sciences) have thought *a lot* more about ethics than mathematics (and allied disciplines), and we should be guided by the best practices that they have developed. Many of these best practices arose in the aftermath of ethically problematic studies, and ethical guidelines have developed as people have tried to learn from past mistakes. As in other arenas, the mathematics community has had major ethical problems (e.g., through various forms of social toxicity), but these problems traditionally have arisen in issues other than the scientific use and abuse of data. This situation has changed; mathematicians are now also facing data-related issues directly in their research. What we learn, teach, and do needs to catch up to this reality.

It is important for mathematical scientists (and others) to think carefully about ethics and engage with it throughout their careers [**CM24**]. We need to incorporate data ethics into the core education in the mathematical, statistical, and computational sciences. We are increasingly using social, animal, and human data (including potentially personal data) in our research. We need to think carefully about when it is appropriate to use such data and when it is not appropriate, and there needs to be systematic training to help mathematical scientists confront these issues. See [**LMP18**] for a short book on ethics and data science, [**FT16**] for a theme journal issue on data ethics, [**Vé21**] for a wide-ranging handbook of digital ethics, [**MCF22**] for a discussion of the ethics landscape in mathematics and advocacy of a "Hippocratic Oath" for mathematics (also see [**D'A22**]), and [**Sku21**] for an article with a useful discussion and pointers to several relevant resources. See [**Ass18**] for the Code of Ethics and Professional Conduct of the Association for Computing

[3]One positive development is that some computer-science conferences now require authors to include ethics statements in their submitted papers.

Machinery (ACM). Buell et al. [**BPT22**] leveraged the ethical-practice standards of the ACM and the American Statistical Association (ASA), which both represent disciplines that are relevant to mathematics, in a survey of mathematicians about ethical standards. See [**TPB24a**, **TPB24b**] for further discussions of the guidelines that they developed for ethical mathematical practice, and see [**WT24**] for associated teaching materials. To help acknowledge issues like data ethics in the mathematical sciences, it is also relevant to include pertinent information on departmental websites. For an example website, which I helped produce at University of Oxford in the aftermath of a study [**TMP12**] by my collaborators and me that used Facebook data, see [**Oxf**].

2.2. Course material on data ethics and society. Many existing courses discuss data ethics and the societal impact of data. Casey Fiesler (Department of Information Science, University of Colorado Boulder) has collected the syllabi of many such courses at [**Fie18**]. A good book on ethics and data science to use as a starting point is [**LMP18**], which the students in UCLA's course on societal impacts of data are asked to read.

Several courses have websites with a lot of useful information about data ethics and related topics. I will highlight a few examples. Matt Salganik (Department of Sociology, Princeton University) has taught a graduate course (from which I drew some material for the present chapter) called *Computational Social Science: Social Research in the Digital Age*. The material for the course's Fall 2016 edition is available at [**Sal16**]. It includes material that appeared later in book form [**Sal17**]. Johan Ugander (Department of Management Science and Engineering, Stanford University) has taught a graduate course called *Data Ethics and Privacy*. See [**Uga20**] for the course website, from which I drew several articles that I mention in Section 6. Johan Ugander's graduate course *Social Algorithms* [**Uga23**] also has excellent resources. Chris Bail (Department of Sociology, Duke University) has taught a graduate course called *Data Science & Society*. See [**Bai22**] for its course materials, schedule, and YouTube videos. A recent book by Chris Wiggins (Department of Applied Physics and Applied Mathematics, Columbia University) and Matthew L. Jones (Department of History, Columbia University) [**WJ23**] is an adaptation of their course *Data: Past, Present, and Future* [**WW24**].

Rachel Thomas (co-founder of FAST.AI and Professor of Practice in the Center for Data Science at Queensland University of Technology) has posted a wealth of resources at [**Tho**]. This material includes a data-ethics course, a data-science blog, a diversity blog, and more. For example, take a look at her collection of short videos on ethics in machine learning [**Tho21**]. The course *Calling Bullshit* by Carl Bergstrom (Department of Biology) and Jevin West (Information School) at University of Washington includes a section on ethics [**BW19**]. This course also includes lots of other valuable material and reading suggestions.

The Cambridge University Ethics in Mathematics Project [**Cam**] also has relevant course material, including a description and recordings of an 8-lecture course called *Ethics for the Working Mathematician*. Other useful material for courses on data ethics are research papers (see Section 6) and discussions of controversies (see Section 4.4).

3. A few efforts related to data ethics, society, and social good

In this section, I briefly discuss a few efforts that are related to data ethics and society. I highlight two recent efforts at University of California, Los Angeles (UCLA), and I briefly mention a few of the many efforts of others. I also highlight the importance of being careful and conscientious in these efforts.

3.1. Two recent initiatives at UCLA.

3.1.1. *An undergraduate course: Societal impacts of data.* I helped design a new undergraduate major in "Data Theory"[4] at UCLA. All students in our Data Theory major must take a new upper-division course, which was designed by Mark Handcock of the Department of Statistics & Data Science, called "Societal Impacts of Data". This course covers a variety of topics—including privacy, algorithmic bias, and many others—through an ethical lens. UCLA's course catalog has the following description of the course:

> *Consideration of impacts that data collected today have upon individuals and society. Rapid increase in scale and types of data collected has impacted commerce and society in new ways. Consideration of economic, social and ethical, legal and political impacts of data, especially that collected on human behavior. Topics include privacy and data protection, intellectual property and confidentiality, sample selection and algorithms, equality and anti-discrimination.*

In a recent offering of the course, [**LMP18**] was used as reading material.

In my view, such courses should be mandatory not only for undergraduates who are majoring in data science (and similar topics), but also for all other students. Because data ethics is crucial for any member of human society, taking a course in it should be a core requirement for literally all students to obtain an undergraduate degree. However, it is especially important for the many students (in mathematics, statistics, computer science, and other subjects) who become data scientists or otherwise work with human data (and nonhuman animal data) in their careers.

3.1.2. *Social-justice data-science postdoctoral scholars.* A new UCLA academic position, which I designed along with Deanna Needell (Department of Mathematics) and Mark Handcock, is a Social-Justice Data-Science (SJDS) postdoctoral scholar. This innovative postdoctoral position, which I hope to see in various forms at academic and other institutions worldwide, is a joint venture of UCLA's mathematics and statistics departments. An SJDS postdoc has two mentors: (1) a faculty member from the Department of Mathematics or the Department of Statistics & Data Science and (2) a faculty member who is a social-justice scholar. We hope to continue to hire SJDS postdocs.

Mathematical, statistical, and computational scientists can accelerate the quantitative study of social-justice issues by harnessing data. They can also leverage recent advances in data science into social justice and activism. Individuals who are trained in other fields (such as sociology) have a long tradition of such involvement in activism. Importantly, mathematicians, statisticians, and computer scientists also have a lot to contribute. Such contributions are not simply a matter

[4]See [**UCLa**] for a description of UCLA's undergraduate major in Data Theory.

of studying abstract problems in mathematical sociology or allied topics. Critically, it requires engagement with social-justice scholars and the communities and other stakeholders that we seek to help. This is what we expect SJDS postdocs to do. Moreover, by sponsoring SJDS scholars as postdocs in mathematics and statistics departments in positions of comparable prestige to our usual postdocs, the mathematical-science community (along with our colleagues in statistics, computer science, and other disciplines) can show students that these paths—whether in academia, in industry, as a data scientist for a nonprofit organization that serves communities, or elsewhere—are available and viable career pathways. It is crucial that we send this message.

3.2. Some other noteworthy efforts. There are many other efforts (which take a variety of forms) on data and society. In Section 2.2, I discussed several courses and resources on data ethics. In this section, I briefly highlight a few initiatives on data and society.

One of these efforts is Mechanism Design for Social Good (MD4SG) [**AG18**, **MD4**], which uses techniques from algorithms, optimization, mechanism design, and disciplinary insights to improve the equity, social welfare, and access to opportunities for historically underserved, disadvantaged, and marginalized communities. Another effort, which was launched by Timnit Gebru, is the Distributed AI Research Institute (DAIR) [**DAI**]. This institute works on community-based and interdisciplinary AI research. AI4All, which was founded by Fei-Fei Li and Olga Russakovsky in 2015, is a nonprofit organization that aims to increase diversity and inclusion in AI education, research, development, and policy [**AI4**].

3.3. An important caveat. There are a diverse variety of both research efforts and practical efforts that fall under the auspices of "data for social good" [**ACX23**]. Harnessing data for social good is both important and laudable, but there are many challenges [**ABK+20**, **BBS22**]. For example, it is vital to engage meaningfully and respectfully with communities, rather than attempting to help people by trying to be a "new sheriff in town", which can be very harmful.

4. Research replication and ethics, open data and code, the tension between privacy and open data and code, and some controversy

4.1. Replication of research. It is crucial to be able to replicate scientific research, and improvement in current practices is necessary to produce reproducible and reliable computational science [**CGH21**]. Obviously, it is necessary to be honest about data and other aspects of research, but professional obligations go far beyond mere honesty. Scholarship needs to be transparent. It is important to precisely explain all details of analysis, implementation, and data cleaning in scholarly works; it is also important to openly provide source code and data. See [**ATG24**] for a set of recommendations for sharing code, and see [**S+25**] for a discussion of guidelines for the management of scientific data. The inclusion of precise explanations and source material is necessary for research dissemination. When a mathematician publishes a theorem and its proof, they also give other people a license to use it freely in their own work. The sharing and portability of knowledge lies at the core of both science and mathematics. Accordingly, it is a professional obligation to share research in a usable form, including by providing

source code and data. Depriving readers of such material is analogous to publishing a theorem statement without a proof or publishing a theorem without permission to use it.

To the extent possible, it is very important to publish the relevant data and code (including code to reproduce all figures) that accompanies a scholarly manuscript.[5] In a manuscript, one also should explicitly and carefully discuss each step in the procedures for data anonymization, cleaning, sampling, and transformation. It is important to be explicit about anything that one does with data so that readers know precisely what choices have been made and can then evaluate whether or not they think that those choices are good ones for the analysis in a manuscript. For example, sampling biases can change the properties of data in fundamental ways [**SWM05**]. Additionally, by providing the original data when possible, others can analyze that data with procedures that involve different choices. There are many choices that scientists make in data analysis—it is impossible not to make such choices—but these choices are a fundamental part of the scientific procedure when conducting research, so it is imperative to inform others of exactly what one has done (with particular highlighting of choices) in any scientific work. They may want to make different choices.

In making data publicly available, posting the output of synthetic models is safer than posting even the safest real-world data (see Section 4.3). When using synthetic data, such as the output of numerical simulations of a differential equation or data that one generates from a random-network model, it is good to publish code to generate the output data (e.g., the examples) that one presents in a manuscript. Publishing user-friendly and open-source code is important for a paper's readers, and it also facilitates the fair evaluation of methods and results. One way to publish code is as supplementary material on a journal website; another way is through repositories such as Bitbucket, GitLab, and GitHub. Posting synthetic data is relevant not only for the output of numerical computations and any other data that one generates, but also even for examples (such as adjacency matrices in a paper about networks) that one constructs by hand in a definition–theorem–proof paper. How relevant it is to include such data depends on the sizes of the examples. Even posting the entries of a 10×10 matrix in a repository saves time for others and reduces transcription errors.

We are all human, and it is easy to forget something or to inadvertently be insufficiently precise about a procedure, so gaps often occur. If somebody e-mails you to ask for a clarification, copy of code (even if poorly commented), or something else, it is important to respond and provide it to them (assuming that it is something that you have the legal and ethical right to provide).

4.2. Acknowledgements, giving credit, and licenses to use and share material. Another part of doing scientific research and presenting it in scholarly works is acknowledging the contributions of others. Naturally, acknowledging contributions includes things like coauthorship and citations of prior work. It also includes things like acknowledging all sources of data, all sources of funding, and thanking people for their useful comments and ideas. In the acknowledgements section of a manuscript, one should include precise details of how one obtained data

[5]In the interest of admitting my own flaws, I note that I have been imperfect in my career about publishing source code with my papers. I am doing this increasingly often, and I seek to improve further.

and how others can also obtain that data (especially if one cannot publish it oneself, perhaps because of privacy considerations or because it is not one's own data to share). It is important to be generous when acknowledging others in manuscripts. If somebody gives useful comments, one should thank them for it (assuming that they want to be thanked).

One should be fair, appropriate, and precise when discussing prior work in a manuscript. It is crucial to give credit where it is due. The research in prior work has heterogeneous levels of mathematical rigor, scientific rigor, and even correctness. How one writes about such work is affected by such things, including some facets that are factual and others that may reflect a variety of opinions. For example, there is a difference in writing that something was "shown" versus "reported" in a prior work. The former wording has a built-in claim, by the author(s) of a manuscript, of the validity of that aspect of that prior work. By contrast, the latter is merely a historical fact (assuming that what one writes is itself accurate).

The increasing prominence of LLMs has brought new ethical concerns to the attribution of proper credit and to the creative process more generally.[6] Scientific journals, funding agencies, professors, and universities are scrambling to develop policies that govern the use of LLMs and other text-generation tools in submissions [**Bra23**]. See [**Con**] for the LLM and AI policy that the Consortium for Mathematics and its Applications (COMAP) has started using for the Mathematical Contest in Modeling (MCM) and their other contests. For UCLA's guidelines for the ethical and "safe" use of AI, see UCLA's new AI website [**UCLb**], which also discusses various AI resources.[7]

Another aspect of open science is the different types of licensing that are available through Creative Commons. See [**Wika**] for a discussion of the different types of Creative Commons licenses. Some licenses allow work to be duplicated for any purpose, and other licenses enforce a variety of use restrictions. I advocate making one's work as open as possible, as others can then use it readily for purposes such as teaching and explaining ideas.

4.3. Privacy concerns and practical considerations for open data and code. There are various tensions and practical considerations with the lofty ideals of openly publishing data and code. For example, one may not be allowed to publish data for privacy reasons or because of nondisclosure agreements. For empirical data, if you have permission to post something (e.g., does the data "belong" to somebody else?) and it does not pose privacy concerns, then it makes sense to post it because doing so promotes good science. Because of privacy issues, one may choose not to publish certain data that one is technically allowed to publish. It is crucial that researchers navigate these issues in a conscientious way. Another issue is that publicly posting usable code and data takes time and energy, and key participants (such as students and other junior researchers) in a project move on to other things,

[6]The use generative AI and LLMs also brings significant ethical concerns about data ownership and copyright infringement, such as through the training data that is used for AI-generated art and text [**Fra22**, **GM23**]. It also can lead to comically disturbing situations, as illustrated by the recent discovery of child sexual-abuse material in a prominent data set (which has been downloaded by many researchers) in the AI community [**Col23**] and by the AI-generated image of a rat with a gigantic penis in a (subsequently retracted) published scientific paper [**Pea24**].

[7]Amidst the very serious concerns about AI, it is also important not to lose sight of its immense promise. For example, see the curated list [**RT24**] (and the announcement of it at [**Tao24**]) of resources for AI in mathematics (e.g., to use AI to assist in mathematical reasoning).

so the team members who are best equipped to do this effectively may no longer be available. In other words, there is sometimes a practical tension between the well-being of one's collaborators and the publishing of code and data.

In Section 4.1, I mentioned that it is important to indicate how one has anonymized data in scholarly works. When is data actually "anonymous", and is it ever possible to "fully" anonymize data? Consider the following scenario, which is discussed in [**Sal16**, **Sal17**]. Suppose that we have a data set of medical records of individuals that includes their names, their home addresses, the zip codes of these addresses, their birth dates, their sexes, their ethnicities, the dates that each individual visited a doctor, the medical diagnoses, the medical procedures, and the prescribed medications. We can try to "anonymize" this data set by removing the names and home addresses of all individuals. We now have a data set of "anonymized" medical records. Suppose that we obtain a data set of voting records and that this data set includes names, home addresses, political-party affiliations, voter-registration dates, zip codes, birth dates, and sexes. Both data sets include the zip codes, birth dates, and sexes of the individuals that are common to the two data sets. One can use such data—namely, data that is common to these two data sets—to deanonymize people in the supposedly "anonymized" data set of medical records [**NS08**]. An infamous example of data deanonymization by combining data sets led to the cancellation of the sequel to the Netflix Prize [**Wikb**].

Given the simultaneous presence of privacy concerns and the desire to produce replicable scientific research, what should one do if the employed data, an algorithm (or part of an algorithm), or something else needs to remain private? This was one key topic of discussion following a publication by Bakshy et al. [**BMA15**], who examined the exposure of different individuals to heterogeneous news and opinions in their Facebook feeds. The paper's authors, who were all Facebook employees at the time, could not reveal how Facebook determined the feeds that individuals saw, so how can others replicate their work to try to evaluate and verify their observations and insights? Which of the insights in [**BMA15**] apply exclusively to Facebook, and which of them also apply to other social-media platforms? In principle, it should be possible to attempt a weaker form of replication of the study's most interesting qualitative results, which are not merely a property of something that is specific to Facebook.

4.4. Controversies in studies and in reactions to studies. Unsurprisingly, some studies that involve human data have been controversial. In other cases, there has been controversy in how authors of studies were treated in the aftermath of their work.

One controversial study, with much ensuing public discussion (see, e.g., [**Luc14**] and many other sources), was an examination of emotional contagions using experiments with Facebook in which user feeds were altered [**KGH14**]. There were angry accusations that the researchers manipulated people's emotions, with additional questioning of subsequent actions by the journal that published the paper. There were also discussions of the procedure to obtain permission to undertake the study in the first place. One key consideration is that there are crucial differences between the ethical procedures for academic and commercial researchers [**Gri16**].

Academic researchers need to obtain approval from an Institutional Review Board (see Section 5.3) before undertaking a study that involves humans, whereas companies like Facebook have publication review boards to approve the publication of a study after it has already been done. Therefore, we know that this study occurred because Facebook concluded that it could be published. By contrast, we do not know about what research is done with our data by Facebook and other entities when an associated document is not placed in the public domain. This leads to an important question: Should academic researchers and companies follow the same rules?

Many technology companies have units that do research on data ethics and related subjects, although that too can lead to controversy. One example is the departure from Google of researcher Timnit Gebru and others who study data ethics [**Wikd**] in the aftermath of a paper that Gebru and her collaborators wrote about the significant risks (including environmental costs, unknown and dangerous biases, and potential uses to deceive people) of LLMs [**Hao20**].

Other research about online social networks has also been controversial. I will mention two examples that have been influential scientifically because of their research findings. One of these examples involved experimental manipulation of feeds by seeding posts on social media with a small number of fake initial upvotes or fake initial downvotes [**AW12**]. In this study, the researchers found that initial upvotes had a persistent effect on the overall positivity of the voting on posts, whereas the initial downvotes were overturned. The other example is the "Tastes, Ties, and Time" study of several waves of Facebook data from students at an Ivy League university in the United States [**LKG+08**]. See [**Per11**] for one discussion of the data-privacy controversy of this study and its associated data set.

4.5. Summary. To summarize some key ideas in Section 4, here are a few things to think about:

- There is tension between open data and personal privacy.
- The use and publication of data, code, and anything else that one reports in a manuscript or discusses with others can be subject to terms-of-service agreements and nondisclosure agreements.
- In what sense can one make one's research replicable if one cannot make all of the associated data (or algorithms or something else) publicly available? There are weaker notions of replication, such as whether or not others can observe similar phenomena in circumstances that are similar but not precisely the same. For instance, in studying human behavior on social media, perhaps certain phenomena are very similar on Facebook and 𝕏 (the social-media platform formerly known as Twitter), but other phenomena are specific to only one of these two social-media platforms.
- Are you comfortable doing research in collaboration with private companies or government entities? Maybe there are some entities with which you are willing to collaborate (perhaps depending on their purposes, goals, and history), but there are others with which you are not willing to collaborate or use data from? If you work with or for such an entity, what is permissible to include in a publication or post online?

5. Ethical principles and other considerations in the scientific use of human data

In this section, I discuss some ethical principles and other considerations in the scientific use of human data.[8] I also discuss Institutional Review Boards (IRBs).[9] Much of my discussion also applies to data from nonhuman animals, but I am focusing on human data in this chapter, so I typically phrase my exposition in human terms.

5.1. Online courses for ethics training. In studies that use human data, it is important to think carefully about ethics and to have formal training in it. A popular choice is the Collaborative Institutional Training Initiative (CITI) program [**CIT**], which offers a variety of courses. For more information, see the website of the UCLA Office of the Human Research Protection Program (OHRPP) [**OHR**]. See Section 5.3 for further discussion of these courses and of IRBs.

5.2. Four key principles. I now enumerate a few critical ethical principles, which are discussed in detail in [**Sal17**, Chapter 6].

In scientific pursuits that involve human data, it is important to do the following:

- be honest and fair (obviously);
- design ethically thoughtful research;
- explain your decisions to others.

Four key principles in research that involves humans are

- respect for persons;
- beneficence;
- justice;
- respect for law and public interest.

These four principles can come into tension with each other, so how do we balance them?

In conducting research with human data, there is a sliding scale: the more your research has the potential to violate personal privacy, the more helpful for humanity its outcome needs to have the potential to be. Four things to ponder with research that involves personal data are the following:

- informed consent;
- understanding and managing informational risk;
- privacy;
- making decisions in the face of uncertainty.

As you design and conduct research, put yourself in the shoes of other people. Think of research ethics as continuous (i.e., there is a sliding scale), rather than as discrete.

[8]Parts of my discussion draw heavily from material in Matt Salganik's course on computational social science [**Sal16**]. See his associated book [**Sal17**] for further discussion.

[9]At some institutions, the backronym for IRB is Internal Review Board.

5.3. Institutional Review Boards (IRBs).

> "Yeah, yeah, but your scientists were so preoccupied over whether or not they could, they didn't stop to think if they *should*."
>
> (stated by the character Ian Malcolm in *Jurassic Park*)

In many situations, it is a professional requirement to obtain permission to undertake a study in the first place. Such permission gives an ethical floor to satisfy; it is not a ceiling. It may be legally and professionally permissible to do something, but it is important to hold oneself to higher standards when it comes to whether or not it is actually the right decision to do it. For example, the limits to informed consent with human data [**LAH+22**] may influence such a decision. Whatever you decide for your own work, make sure that you think carefully about it.

In universities in the United States, a researcher that is working with personal data needs to check with their university's IRB to ensure that they are conducting research in an ethical way. Universities in other countries, private companies, government laboratories, and other organizations often have bodies that are similar to IRBs, but the procedures and especially the specific details are very different [**Gri16**]. A university IRB may inform you that you do not need to submit a formal application for a research project to be approved, or they may inform you that a formal application is necessary. Let your IRB know what data you have (or what data you plan to acquire and how you plan to acquire it) and what you plan to do with it. Different IRBs can rule differently. When an IRB grants permission to undertake a study, they have decided that a proposed project is above the ethical floor and hence that it is permissible to do that project. One's own standards should be higher. (Again, ethical approval is a floor, rather than a ceiling.) In this light, it is worth examining the discussion following a controversial IRB-approved study of "emotional manipulation" in which researchers adjusted user feeds on Facebook [**Luc14**, **KGH14**].

For a more concrete idea about IRBs and conducting ethical research in a university setting, see the materials at [**OHR**]. The requirements to conduct research with human data (and nonhuman animal data) include taking various online training courses, such as those in the CITI program. These courses, which are available online at [**CIT**], are in common use in the United States. The training that is required, expected, and available for research projects that involve human data and other sensitive data is rather different in different countries. Some of the topics that are covered in courses on ethical research are animal care and use, biosafety and biosecurity, human-subject research, information privacy and security, and responsible conduct of research. It is useful to take a variety of these courses even when they are not required.

5.4. Another salient warning. It is important to be cognizant that your research and data can potentially be "weaponized" by other people—who may not care about nuances in research findings and who may interpret unfortunate wording or insufficiently careful exposition in nefarious ways—so you need to be conscientious about the precise wording in your publications and other media. This is particularly relevant in research on social systems and on tools that are applicable to social systems [**Ste22**]. This possibility also makes it particularly crucial to be

open-minded about potential sources of bias in your research. As always, it is important to be ethically thoughtful.

6. Some research and lay articles that are relevant to data ethics and data privacy

There is a wealth of research about data ethics and related topics in computer science, sociology, and other fields [**JIV19**, **LHF+21**, **WSAO+21**, **ABK+20**]. There is no way that I can possibly be exhaustive,[10] so I will highlight a selection of studies to give a taste of existing research. I will also briefly discuss a few articles in blogs, news websites, and other nontechnical venues. Unsurprisingly, much of this research is simultaneously fascinating and concerning [**Exp18**].

6.1. Algorithmic biases and other biases. A key research area is the analysis and mitigation of biases in computational techniques and data analysis.

One central topic is algorithmic bias, in which systematic and repeatable errors result in unfair outcomes, such as privileging one group of people over others [**Bir21**]. For example, such biases have systematically hurt certain racial and demographic groups in predictive policing [**Nat18**]. Algorithmic biases also reinforce negative racial and gender biases in online search-engine results [**Nob18**]. As is now evident, algorithmic biases can have very serious consequences. For example, in 2020, an algorithm for prioritizing COVID-19 vaccinations severely disadvantaged human-facing medical workers [**Che20**].

There is much research on the mitigation of algorithmic biases, such as in the algorithmic hiring of people for jobs [**RBKL20**] and the algorithmic classification of individuals (e.g., for admission to a university) [**DHP+12**]. To encourage transparent reporting, clarify intended use cases, and minimize usage in inappropriate contexts, some researchers have proposed the inclusion of "model cards" to accompany trained machine-learning models [**MWZ+19**]. Such model cards are short documents that give intended use cases; benchmarked evaluations across cultural, demographic, phenotypic, and intersectional groups; and other salient information.

Biases in computer systems, machine learning, and AI go far beyond only algorithmic biases [**FN96**, **HV24**]. See [**MMS+21**, **BHN23**, **CDNSG23**] for reviews of bias and "fairness" in machine learning, and see [**N+20**] for an introductory survey of bias in data-driven AI systems. As advocated in [**KA21**], it is important to go beyond notions of mere algorithmic "fairness" (which focuses on intra-group versus inter-group differences). One must also analyze (1) inequality and the causal impact of algorithms and (2) the distribution of power. It is also important to be cognizant of traps that can beset work on "fair" machine learning in sociotechnical systems [**SBF+19**]. In "fair" machine learning, it is necessary to carefully examine context-specific consequences [**CDNSG23**].

To mitigate biases, it is imperative for researchers to carefully justify their choices of data sets. It is crucial to consider the social context (country, gender, race, and so on) of data. Why is one using a particular data set, and is it appropriately representative for the problem that one is studying [**KDHF21**]? It is common and often convenient to import data sets that were used originally to study one problem for investigations of many other problems, and that can lead to several

[10] I selected some of the highlighted articles, including both research studies and lay articles, from the website for Johan Ugander's course on data ethics and data privacy [**Uga20**].

issues. This is particularly relevant for human data and social data, and it is important to think carefully about whether a data set is appropriate for a given study. Particular facets of a setting can interfere with attempts to generalize a study's insights from its specific use case to other potentially similar situations.

6.2. Human privacy, personal characteristics, and personalization. Other salient research focuses on human privacy. As has been studied thoroughly [**EN16**], people are tracked extensively when they visit websites. User profiles, which encode both characteristics and behavior, are generated through interactions with websites and other digital systems. User profiling is ubiquitous in daily life, and it is also a vast area of research [**PBL24**]. It is possible to infer private traits and attributes from digital records of human behavior [**KSG13**]. One can also steal the identities of visitors to websites [**Nar10**], and trackers can use social-network structure to deanonymize data from browsing the World Wide Web [**SSGN17**]. Companies can exploit the information that they observe (or infer) from website visits. For example, about a decade ago, the online travel agent Orbitz showed higher prices for flights to users of Macintosh computers than to users of other types of computers [**Whi12**]. It is imperative that individuals and other stakeholders have agency in how their personal data are used [**SVW21**], although this can come into tension with the scholarly desire to promote open data (see Section 4). For example, the sharing of data from the continent of Africa has often been driven by non-African stakeholders [**AAB$^+$21**]. Additionally, conventional studies of algorithmic fairness are centered on Western culture; data proxies are different in different cultures, and it is important to consider local context when building data models [**SAH$^+$21**]. See [**LHF$^+$21**, **WSAO$^+$21**] for discussions of access, ethics, and best practices in the algorithmic measurement of human data.

There are many ways to infer the characteristics of people and communities, as well as social and other ties between people, using data analysis. For example, it is possible to infer social ties by examining the geographical proximity of individuals in time and space using offline or online data [**CBC$^+$10**]. Additionally, researchers have used machine learning and Google Street View to estimate the demographic composition of neighborhoods across the United States [**GKW$^+$17**]. Unsurprisingly, there are significant gaps and biases (and associated exclusion issues) in such geographic data [**GD22**]. Moreover, as I discussed in Section 4, the fact that individuals appear in multiple social networks (e.g., multiple social-media platforms and multiple databases) gives considerable ability to identify them or information about them even when they appear to be "anonymous" in those individual networks [**NS08**]. For example, the inference of the sexual orientation of some individuals by combining Netflix data—which was released as part of a public competition to improve Netflix's algorithmic ability to infer user ratings of movies—of movie rentals with movie rating data from the Internet Movie Database (IMDB) led to legal troubles, and Netflix ultimately cancelled a planned sequel competition [**Wikb**]. More recently, researchers have successfully inferred gender using mobile-payment data [**SM23**]. Other researchers have illustrated that one can use interactions to identify people even across long time periods [**CMM$^+$22**]. Tools like data analysis and machine learning can yield crucial and actionable insights (e.g., of racial disparities in police stops [**PSO$^+$20**]), but they need to be applied in a careful and respectful way.

A related issue is personalization [**Uga17**], such as in the targeted advertisements that inundate people on social-media platforms. These advertisements can inform us of desirable products (such as t-shirts, dice, or plushies in my case), but the online personalization of commercial and political advertising can cause serious problems. Infamously, Donald Trump's 2016 presidential campaign employed the company Cambridge Analytica for targeted online political advertising. Additionally, the movie *Straight Outta Compton* was advertised differently to people based on their inferred demographic characteristics [**McA16**]. There is much pertinent research on personalization. For example, there are studies of the effectiveness of spam-based marketing [**KKL$^+$09**], the detection of spam in social-bookmarking websites (such as Pinterest and Digg) [**MCM09**], and other aspects of spam and related online marketing. Users on social-media platforms like Facebook have personalized feeds, and a study that involved the adjustment of such feeds [**KGH14**] led to a major controversy, including with the issue of manipulating people's emotions [**Luc14**].

6.3. AI ethics guidelines, tradeoffs in different ethical values, and theoretical barriers. Guidelines for AI ethics appear to have converged on five key values:[11] transparency, fairness, safety, accountability, and privacy [**JIV19**]. There are inevitable tensions between these ethical values. Academics, public-sector organizations, and private companies emphasize different ethical values, with particularly systematic differences between practitioners and the general public [**JBAO22**].

It is mathematically impossible for an AI system to simultaneously include many parameters, be robust to poisoning (e.g., with fake data) by an adversarial actor, and preserve privacy [**EMFG$^+$22**], so the use of AI will always include tradeoffs between different ethical values even when everything works perfectly. Of course, AI systems don't work perfectly. For example, it was proven recently that one can plant undetectable backdoors into machine-learning models and thereby confuse "adversarially robust" classifiers, demonstrating a major theoretical barrier to the certification of adversarial robustness [**GKVZ22**].

7. Conclusion

The use of digital data to study and analyze humans, create technologies, and enact policies that involve humans is both powerful and dangerous [**Edi21**, **Off22**]. Mathematical scientists have a younger tradition of studying human data than researchers in many other disciplines, such as the social, behavioral, and medical sciences. Most researchers in the mathematical, statistical, and computational sciences have not had research-ethics and data-ethics training. This needs to change. One recent paper has even proposed a "Hippocratic Oath" for the mathematical sciences [**MCF22**]. The mathematical-science community needs to learn from the best practices of other disciplines to ensure that our research is ethical. Other scholarly communities have been considering research ethics a lot longer than we

[11] These values overlap with (but are organized differently from) the five principles to guide the design, use, and deployment of automated systems in the recent blueprint for an AI Bill of Rights from the United States government's Office of Science and Technology Policy (OSTP) [**Off22**]. The OSTP's five principles are (1) safe and effective systems, (2) algorithmic-discrimination protections, (3) data privacy, (4) notice and explanation, and (5) human alternatives, consideration, and fallback.

have, and it is important that we learn from them. As in the social and medical sciences, the mathematical and computational sciences need a robust program of ethics training.

In this chapter, I have discussed many aspects of data ethics in education, publishing, and research. In some of my discussions, I have also touched on issues of data and justice. It is important to distinguish between ethics and justice, including when it comes to data [**Kit14**]. Ethics presupposes laws and social norms, and then one considers what is morally appropriate within those frameworks. Justice is concerned with how to change laws and modify social norms to attain broader equity. Both ethics and justice are core aspects of our interactions with data.

I encourage you to read widely, think about, and discuss how to do research ethically, especially for studies of—or with consequences for—social systems and human data. It is also valuable to read about past research and societal controversies. There have been mistakes in the past (and there continue to be mistakes), and we need to learn from them. We may all set our ethical bars in different places and have different views on different issues, but our scholarship needs to be conscientious and ethically thoughtful. As a reminder, official approval (e.g., from an IRB) to undertake a study is only a lower bound. The ethical bar that one needs to surpass in the design and performance of research is a sliding bar: the more potential for invasion of human privacy (or other potential harm), the more valuable to humanity the potential outcome of a research project has to be.

Be ethically thoughtful.

Acknowledgements

In parts of this chapter, I drew on materials from a class [**Sal16**] that Matt Salganik has taught on computational social science at Princeton University. Johan Ugander suggested several excellent resources and research studies. I thank Jacob Foster, Mark Handcock, Andrei Kleshin, Peter Mucha, Dennis Müller, Deanna Needell, Matt Salganik, Massimo Stella, Alexandria Volkening, Chris Wiggins, Richard Yeh, and an anonymous referee for helpful comments. Finally, I thank my fellow Short Course organizers (Heather Zinn Brooks, Michelle Feng, and Alexandria Volkening) for their continuing support on this project and elsewhere, their collaboration, and their friendship.

References

[AAB+21] Rediet Abebe, Kehinde Aruleba, Abeba Birhane, Sara Kingsley, George Obaido, Sekou L Remy, and Swathi Sadagopan, *Narratives and counternarratives on data sharing in Africa*, Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21), Association for Computing Machinery, New York, NY, USA, 2021, pp. 329–341.

[ABK+20] Rediet Abebe, Solon Barocas, Jon Kleinberg, Karen Levy, Manish Raghavan, and David G. Robinson, *Roles for computing in social change*, Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20), Association for Computing Machinery, New York, NY, USA, 2020, pp. 252–260.

[ACX23] Ahmed Abbasi, Roger H. L. Chiang, and Jennifer Xu, *Data science for social good*, J. Assoc. Inf. Syst. **24** (2023), no. 6, 1439–1458.

[AG18] Rediet Abebe and Kira Goldner, *Mechanism design for social good*, AI Matters **4** (2018), no. 3, 27–34.

[AI4] AI4All, accessed December 26, 2023. `https://ai-4-all.org`

[Ass18] Association for Computing Machinery, *ACM code of ethics and professional conduct, 2018*, accessed December 26, 2023. `https://ethics.acm.org`

[ATG24] Richard J. Abdill, Emma Talarico, and Laura Grieneisen, *A how-to guide for code sharing in biology*, PLoS Biology **22** (2024), no. 9, e3002815.

[AW12] Sinan Aral and Dylan Walker, *Identifying influential and susceptible members of social networks*, Science **337** (2012), no. 6092, 337–341. MR2986000

[Bai22] Christopher A. Bail, *Data science & society*, 2022, Sociology 367S, Spring 2022, Duke University. https://dssoc.github.io

[BBS22] Caroline Buckee, Satchit Balsari, and Andrew Schroeder, *Making data for good better*, PLoS Digital Health **1** (2022), no. 1, e0000010.

[BFPV21] Heather Z. Brooks, Michelle Feng, Mason A. Porter, and Alexandria Volkening, *Short course: Mathematical and computational methods for complex social systems*, 2021. https://zerodivzero.com/short_course/aaac8c66007a4d23a7aa14857a3b778c/titles

[BGMMS21] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell, *On the dangers of stochastic parrots: Can language models be too big?* 🦜, Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21), Association for Computing Machinery, New York, NY, USA, 2021, pp. 610–623.

[BHN23] Solon Barocas, Moritz Hardt, and Arvind Narayanan, *Fairness and machine learning: Limitations and opportunities*, MIT Press, Cambridge, MA, USA, 2023. http://www.fairmlbook.org

[Bir21] Abeba Birhane, *Algorithmic injustice: A relational ethics approach*, Patterns **2** (2021), no. 2, 100205.

[BMA15] Eytan Bakshy, Solomon Messing, and Lada A. Adamic, *Exposure to ideologically diverse news and opinion on Facebook*, Science **348** (2015), no. 6239, 1130–1132. MR3380630

[BPT22] Catherine A. Buell, Victor I. Piercey, and Rochelle E. Tractenberg, *Leveraging guidelines for ethical practice of statistics and computing to develop standards for ethical mathematical practice: A white paper*, arXiv:2209.09311, 2022.

[Bra23] Jeffrey Brainard, *Journals take up arms against AI-written text*, Science **379** (2023), no. 6634, 740–741.

[BW19] Carl T. Bergstrom and Jevin West, *Calling bullshit: Data reasoning in a digital world. Week 10. The ethics of calling bullshit*, 2019, Informatics 270 / Biology 270, University of Washington. https://www.callingbullshit.org/syllabus.html#Ethics

[Cam] The Cambridge University Ethics in Mathematics Project, accessed December 26, 2023. https://www.ethics.maths.cam.ac.uk

[CBC+10] David J. Crandall, Lars Backstrom, Dan Cosley, Siddharth Suri, Daniel Huttenlocher, and Jon Kleinberg, *Inferring social ties from geographic coincidences*, Proceedings of the National Academy of Sciences of the United States of America **107** (2010), no. 52, 22436–22441.

[CDNSG23] Sam Corbett-Davies, Johann D. Gaebler, Hamed Nilforoshan, Ravi Shroff, and Sharad Goel, *The measure and mismeasure of fairness*, Journal of Machine Learning Research **24** (2023), 312. MR4664749

[CGH21] Peter V. Coveney, Derek Groen, and Afons G. Hoekstra, *Reliability and reproducibility in computational science: Implementing validation, verification and uncertainty quantification* in silico, Philosophical Transactions of the Royal Society A **379** (2021), 20200409.

[Che20] Caroline Chen, *Only seven of Stanford's first 5,000 vaccines were designated for medical residents*, ProPublica, December 18, 2020. https://www.propublica.org/article/only-seven-of-stanfords-first-5-000-vaccines-were-designated-for-medical-residents

[CIT] CITI Program, *Explore our courses*, accessed December 26, 2023. https://about.citiprogram.org

[CM23] Maurice Chiodo and Dennis Müller, *Manifesto for the responsible development of mathematical works—A tool for practitioners and for management*, arXiv:2306.09131, 2023.

[CM24] Maurice Chiodo and Dennis Müller, *A field guide to ethics in mathematics*, Notices of the American Mathematical Society **71** (2024), no. 7, 939–947. MR4776111

[CMM+22] Ana-Maria Creţu, Federico Monti, Stefano Marrone, Xiaowen Dong, Michael Bronstein, and Yves-Alexandre de Montjoye, *Interaction data are identifiable even across long periods of time*, Nature Communications **3** (2022), 313.

[Col23] Samantha Cole, *Largest dataset powering AI images removed after discovery of child sexual abuse material*, 404 Media, December 20, 2023. https://www.404media.co/laion-datasets-removed-stanford-csam-child-abuse/

[Con] Consortium for Mathematics and its Applications, *Use of large language models and generative AI tools in COMAP contests*, accessed April 21, 2024. https://www.contest.comap.com/undergraduate/contests/mcm/flyer/Contest_AI_Policy.pdf

[D'A22] Susan D'Agostino, *Do scientists need an AI Hippocratic oath? Maybe. Maybe not.*, Bulletin of the Atomic Scientists, June 9, 2022. https://thebulletin.org/2022/06/do-scientists-need-an-ai-hippocratic-oath-maybe-maybe-not/

[DAI] *Distributed AI Research Institute (DAIR)*, accessed December 26, 2023. https://www.dair-institute.org

[DHP+12] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel, *Fairness through awareness*, Proceedings of the 3rd Innovations in Theoretical Computer Science Conference (ITCS '12), Association for Computing Machinery, New York, NY, USA, 2012, pp. 214–226. MR3388391

[Dor20] Lori Dorn, *A deepfaked Gollum performs a 'precious' cover of Sinead O'Connor's version of 'Nothing Compares 2 U'*, Laughing Squid, December 18, 2020. https://laughingsquid.com/gollum-sinead-oconnor-nothing-compares-2-u/. The music video itself is available at https://www.youtube.com/watch?v=9d9pZi7uZQs.

[Edi21] Editors, *The powers and perils of using digital data to understand human behaviour*, Nature **595** (2021), 149–150.

[EMFG+22] El-Mahdi El-Mhamdi, Sadegh Farhadkhani, Rachid Guerraoui, Nirupam Gupta, Lê-Nguyên Hoang, Rafael Pinot, Sébastien Rouault, and John Stephan, *On the impossible safety of large AI models*, arXiv:2209.15259, 2022.

[EN16] Steven Englehardt and Arvind Narayanan, *Online tracking: A 1-million-site measurement and analysis*, Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, CCS '16, Association for Computing Machinery, New York, NY, USA, 2016, pp. 1388–1401.

[ET19] Zeynep Engin and Philip Treleaven, *Algorithmic government: Automating public services and supporting civil servants in using data science technologies*, The Computer Journal **62** (2019), no. 3, 448–460.

[Eth] Ethical Mathematics, accessed December 26, 2023. https://ethicalmath.com

[Exp18] Explain xkcd, *2072: Evaluating Tech Things*, 2018. https://www.explainxkcd.com/wiki/index.php/2072:_Evaluating_Tech_Things. The original comic (by Randall Munroe) is available at https://xkcd.com/2072/.

[Fer24] Emilio Ferrara, *GenAI against humanity: Nefarious applications of generative artificial intelligence and large language models*, Journal of Computational Social Science **7** (2024), 549–569.

[Fie18] Casey Fiesler, *Tech ethics curricula: A collection of syllabi*, Medium, July 5, 2018. https://cfiesler.medium.com/tech-ethics-curricula-a-collection-of-syllabi-3eedfb76be18

[FN96] Batya Friedman and Helen Nissenbaum, *Bias in computer systems*, ACM Transactions on Information Systems **14** (1996), no. 3, 330–347.

[Fra22] Matthew R. Francis, *The ethics of artificial intelligence-generated art*, SIAM News **56** (2022), no. 09. https://sinews.siam.org/Details-Page/the-ethics-of-artificial-intelligence-generated-art

[FT16] L. Floridi and M. Taddeo, *What is data ethics?*, Philosophical Transactions of the Royal Society A **374** (2016), no. 2083, 20160360.

[GD22] Mark Graham and Martin Dittus, *Geographies of digital exclusion: Data and inequality*, Pluto Press, London, UK, 2022.

[GKVZ22] Shafi Goldwasser, Michael P. Kim, Vinod Vaikuntanathan, and Or Zamir, *Planting undetectable backdoors in machine learning models*, 2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS 2022), IEEE Computer Society, Los Alamitos, CA, USA, 2022, pp. 931–942. MR4537269

[GKW+17] Timnit Gebru, Jonathan Krause, Yilun Wang, Duyun Chen, Jia Deng, Erez Lieberman Aiden, and Li Fei-Fei, *Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States*, Proceedings of the National Academy of Sciences of the United States of America **114** (2017), no. 50, 13108–13113.

[GM23] Michael M. Grynbaum and Ryan Mac, The Times *sues OpenAI and Microsoft over A.I. use of copyrighted work*, The New York Times, December 27, 2023. `https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html`

[Gri16] Peter Grindrod, *Beyond privacy and exposure: Ethical issues within citizen-facing analytics*, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **374** (2016), no. 2083, 20160132.

[Hao20] Karen Hao, *We read the paper that forced Timnit Gebru out of Google. Here's what it says.*, MIT Technology Review, December 4, 2020. `https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru/`

[Has21] Gry Hasselbalch, *Data ethics of power: A human approach in the big data and AI era*, Edward Elgar Publishing, Northampton, MA, USA, 2021.

[HV24] David W. Hogg and Soledad Villar, *Position: Is machine learning good or bad for the natural sciences?*, Proceedings of the 41st International Conference on Machine Learning (ICML '24), Proceedings of Machine Learning Research **235** (2024), 18439–18453.

[IPA20] Institute for Pure and Applied Mathematics, *White paper: Deep fakery*, January 7, 2020. `http://www.ipam.ucla.edu/news/white-paper-deep-fakery/`

[JBAO22] Maurice Jakesch, Zana Buçinca, Saleema Amershi, and Alexandra Olteanu, *How different groups prioritize ethical values for responsible AI*, Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22), Association for Computing Machinery, New York, NY, USA, 2022, pp. 310–323.

[JIV19] Anna Jobin, Marcello Ienca, and Effy Vayena, *The global landscape of AI ethics guidelines*, Nature Machine Intelligence **1** (2019), 389–399.

[KA21] Maximilan Kasy and Rediet Abebe, *Fairness, equality, and power in algorithmic decision-making*, Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21), Association for Computing Machinery, New York, NY, USA, 2021, pp. 576–586.

[KDHF21] Bernard Koch, Emily Denton, Alex Hanna, and Jacob G. Foster, *Reduced, reused and recycled: The life of a dataset in machine learning research*, NeurIPS 2021: Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks, 2021.

[KGH14] Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock, *Experimental evidence of massive-scale emotional contagion through social networks*, Proceedings of the National Academy of Sciences of the United States of America **111** (2014), no. 24, 8788–8790.

[Kit14] Rob Kitchin, *The data revolution: Big data, open data, data infrastructures & their consequences*, Sage Publishing, Thousand Oaks, CA, USA, 2014.

[KKL+09] Chris Kanich, Christian Kreibich, Kirill Levchenko, Brandon Enright, Geoffrey M. Voelker, Vern Paxson, and Stefan Savage, *Spamalytics: An empirical analysis of spam marketing conversion*, Communications of the ACM **52** (2009), no. 9, 99–107.

[KSG13] Michal Kosinski, David Stillwell, and Thore Graepel, *Private traits and attributes are predictable from digital records of human behavior*, Proceedings of the National Academy of Sciences of the United States of America **110** (2013), no. 15, 5802–5805.

[LAH+22] Juniper L. Lovato, Antoine Allard, Randall Harp, Jeremiah Onaolapo, and Laurent Hébert-Dufresne, *Limits of individual consent and models of distributed consent in online social networks*, Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22), Association for Computing Machinery, New York, NY, USA, 2022, pp. 2251–2262.

[Les23] Kif Leswing, *Microsoft limits Bing A.I. chats after the chatbot had some unsettling conversations*, CNBC, February 17, 2023. https://www.cnbc.com/2023/02/17/microsoft-limits-bing-ai-chats-after-the-chatbot-had-some-unsettling-conversations.html

[LHF+21] David Lazer, Eszter Hargittai, Deen Freelon, Sandra González-Bailón, Kevin Munger, Katherine Ognyanova, and Jason Radford, *Meaningful measures of human society in the twenty-first century*, Nature **595** (2021), 189–196.

[LKG+08] Kevin Lewis, Jason Kaufman, Marco Gonzalez, Andreas Wimmer, and Nicholas Christakis, *Tastes, ties, and time: A new social network dataset using Facebook.com*, Social Networks **30** (2008), no. 4, 330–342.

[LMP18] Mike Loukides, Hilary Mason, and DJ Patil, *Ethics and data science*, O'Reilly Media, Inc., Sebastopol, CA, USA, 2018.

[Luc14] Michael Luca, *Were OkCupid's and Facebook's experiments unethical?*, Harvard Business Review, July 29, 2014. https://hbr.org/2014/07/were-okcupids-and-facebooks-experiments-unethical

[Mü24] Dennis Müller, *Situating "ethics in mathematics" as a philosophy of mathematics ethics education*, Ethics and Mathematics Education, Springer, Cham, Switzerland, 2024, pp. 71–87. MR4830879

[MC23] Dennis Müller and Maurice Chiodo, *Mathematical artifacts have politics: The journey from examples to embedded ethics*, arXiv:2308.04871, 2023.

[McA16] Nathan McAlone, *Why 'Straight Outta Compton' had different Facebook trailers for people of different races*, Business Insider, March 16, 2016. https://www.businessinsider.com/why-straight-outta-compton-had-different-trailers-for-people-of-different-races

[MCF22] Dennis Müller, Maurice Chiodo, and James Franklin, *A Hippocratic Oath for mathematicians? Mapping the landscape of ethics in mathematics*, Science and Engineering Ethics **28** (2022), 41.

[MCM09] Benjamin Markines, Ciro Cattuto, and Filippo Menczer, *Social spam detection*, Proceedings of the 5th International Workshop on Adversarial Information Retrieval on the Web (AIRWeb '09), Association for Computing Machinery, New York, NY, USA, 2009, pp. 41–48.

[MD4] Mechanism Design for Social Good, accessed December 31, 2023. https://www.md4sg.com

[MMS+21] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan, *A survey on bias and fairness in machine learning*, ACM Computing Surveys **54** (2021), no. 6, 115.

[Mos21] Bob Moses, *Returning to 'normal' in education is not good enough*, The Imprint, August 24, 2021. https://imprintnews.org/opinion/returning-to-normal-in-education-is-not-good-enough/58069

[MWW+23] Benjamin Moynihan, Robin T. Wilson, Joan Wynne, Lee J. McEwan, Mary M. West, Frank E. Davis, Herbert Clemens, Greg Budzban, Edith Aurora Graf, and Aidan Soguero, *What's math got to do with it?: Bob Moses, algebra, and the movement for civil rights, January 23, 1935–July 25, 2021*, Notices of the American Mathematical Society **70** (2023), no. 2, 258–277. MR4537170

[MWZ+19] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru, *Model cards for model reporting*, Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19), Association for Computing Machinery, New York, NY, USA, 2019, pp. 220–229.

[N+20] Eirini Ntoutsi et al., *Bias in data-driven artificial intelligence systems—An introductory survey*, WIREs Data Mining and Knowledge Discovery **10** (2020), e1356.

[Nar10] Arvind Narayanan, *Cookies, supercookies and ubercookies: Stealing the identity of Web visitors*, 33 Bits of Entropy: The End of Anonymous Data and What to Do About It, February 18, 2010. https://33bits.wordpress.com/2010/02/18/cookies-supercookies-and-ubercookies-stealing-the-identity-of-web-visitors/

[Nat18] National Academies of Sciences, Engineering, and Medicine, *Proactive policing: Effects on crime and communities*, The National Academies Press, Washington, DC, USA, 2018.

[Nob18] Safiya Umoja Noble, *Algorithms of oppression: How search engines reinforce racism*, NYU Press, New York, NY, USA, 2018.

[NS08] Arvind Narayanan and Vitaly Shmatikov, *Robust de-anonymization of large sparse datasets*, 2008 IEEE Symposium on Security and Privacy (sp 2008), 2008, pp. 111–125.

[OBS+19] Osonde A. Osoba, Benjamin Boudreaux, Jessica Saunders, J. Luke Irwin, Pam A. Mueller, and Samantha Cherney, *Algorithmic equity: A framework for social applications*, RAND Corporation, Santa Monica, CA, USA, July 11, 2019. https://www.rand.org/pubs/research_reports/RR2708.html

[Off22] Office of Science and Technology Policy, The White House, *Blueprint for an AI bill of rights: Making automated systems work for the American people*, 2022, accessed March 1, 2025. https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights/

[OHR] OHRPP, *UCLA Office of the Human Research Protection Program*, accessed December 26, 2023. https://ohrpp.research.ucla.edu

[O'N16] Cathy O'Neil, *Weapons of math destruction: How big data increases inequality and threatens democracy*, The Crown Publishing Group, New York, NY, USA, 2016. MR3561130

[Oxf] University of Oxford Mathematical Institute, *Research using data involving humans*, accessed December 26, 2023. https://www.maths.ox.ac.uk/members/policies/data-protection/research-using-data-involving-humans

[PBL24] Erasmo Purificato, Ludovico Boratto, and Ernesto William De Luca, *User modeling and user profiling: A comprehensive survey*, arXiv:2402.09660, 2024.

[Pea24] Jordan Pearson, *Scientific journal publishes AI-generated rat with gigantic penis in worrying incident*, VICE, February 15, 2024. https://www.vice.com/en/article/dy3jbz/scientific-journal-frontiers-publishes-ai-generated-rat-with-gigantic-penis-in-worrying-incident

[Per11] Marc Perry, *Harvard researchers accused of breaching students' privacy*, The Chronicle of Higher Education, July 10, 2011. https://www.chronicle.com/article/harvard-researchers-accused-of-breaching-students-privacy/

[Por21] Mason A. Porter, *Data ethics for mathematicians*, 2021. https://zerodivzero.com/short_course/aaac8c66007a4d23a7aa14857a3b778c/title/181b207c7d2941278be4641ea5fe0e21

[PSO+20] Emma Pierson, Camelia Simoiu, Jan Overgoor, Sam Corbett-Davies, Daniel Jenson, Amy Shoemaker, Vignesh Ramachandran, Phoebe Barghouty, Cheryl Phillips, Ravi Shroff, and Sharad Goel, *A large-scale analysis of racial disparities in police stops across the United States*, Nature Human Behaviour **4** (2020), 736–745.

[RBKL20] Manish Raghavan, Solon Barocas, Jon Kleinberg, and Karen Levy, *Mitigating bias in algorithmic hiring: Evaluating claims and practices*, Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20), Association for Computing Machinery, New York, NY, USA, 2020, pp. 469–481.

[Ree18] Chris Reed, *How should we regulate artificial intelligence?*, Philosophical Transactions of the Royal Society A **376** (2018), no. 2128, 20170360.

[RH23] Casey Ross and Bob Herman, *Denied by AI: How Medicare Advantage plans use algorithms to cut off care for seniors in need*, STAT, March 13, 2023. https://www.statnews.com/2023/03/13/medicare-advantage-plans-denial-artificial-intelligence/

[RSMCM24] Lucy Rycroft-Smith, Dennis Müller, Maurice Chiodo, and Darren Macey, *A useful ethics framework for mathematics teachers*, Ethics and mathematics education: The good, the bad and the ugly (Paul Ernest, ed.), Springer, Cham, Switzerland, 2024, pp. 359–394. MR4830892

[RT24] Talia Ringer and Terrence Tao, *AI for Math resources*, accessed April 21, 2024. https://docs.google.com/document/d/1kD7H4E28656ua8jOGZ934nbH2HcBLyxcRgFDduH5iQ0/edit

[S+25] Herbert M. Sauro et al., *From FAIR to CURE: Guidelines for computational models of biological systems*, arXiv:2502.15597, 2025.

[SAH+21] Nithya Sambasivan, Erin Arnesen, Ben Hutchinson, Tulsee Doshi, and Vinodkumar Prabhakaran, *Re-imagining algorithmic fairness in India and beyond*, Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21), Association for Computing Machinery, New York, NY, USA, 2021, pp. 315–328.

[Sal16] Matthew Salganik, *Computational social science: Social research in the digital age*, Sociology 596, Princeton University, 2016. https://www.princeton.edu/~mjs3/soc596_f2016/

[Sal17] Matthew Salganik, *Bit by bit: Social research in the digital age*, illustrated edition, Princeton University Press, Princeton, NJ, USA, 2017. https://www.bitbybitbook.com

[SBF+19] Andrew D. Selbst, Danah Boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi, *Fairness and abstraction in sociotechnical systems*, Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19), Association for Computing Machinery, New York, NY, USA, 2019, pp. 59–68.

[Sch21] G. William Schwert, *The remarkable growth in financial economics, 1974–2020*, Journal of Financial Economics **140** (2021), no. 3, 1008–1046.

[SH22] Luke Stark and Jevan Hutson, *Physiognomic artificial intelligence*, Fordham Intellectual Property, Media & Entertainment Law Journal **32** (2022), no. 4, 2.

[Sku21] Joe Skufca, *Incorporating ethical discussions in the mathematics classroom*, SIAM News **54** (2021), no. 06. https://www.siam.org/publications/siam-news/articles/incorporating-ethical-discussions-in-the-mathematics-classroom/

[SM23] Ben Stobaugh and Dhiraj Murthy, *Predicting gender and political affiliation using mobile payment data*, arXiv:2302.08026, 2023.

[SMBM23] Vincent J. Straub, Deborah Morgan, Jonathan Bright, and Helen Margetts, *Artificial intelligence in government: Concepts, standards, and a unified framework*, Government Information Quarterly **40** (2023), no. 4, 101881.

[SSGN17] Jessica Su, Ansh Shukla, Sharad Goel, and Arvind Narayanan, *De-anonymizing Web browsing data with social networks*, Proceedings of the 26th International Conference on World Wide Web (WWW '17), International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 2017, pp. 1261–1269.

[Ste22] Janet D. Stemwedel, *Science must not be used to foster white supremacy*, Scientific American, May 24, 2022. https://www.scientificamerican.com/article/science-must-not-be-used-to-foster-white-supremacy/

[SVW21] Jathan Sadowski, Salomé Viljoen, and Meredith Whittaker, *Everyone should decide how their digital data are used—Not just tech companies*, Nature **595** (2021), 169–171.

[SWM05] Michael P. H. Stumpf, Carsten Wiuf, and Robert M. May, *Subnets of scale-free networks are not scale-free: Sampling properties of networks*, Proceedings of the National Academy of Sciences of the United States of America **102** (2005), no. 12, 4221–4224.

[Tao24] Terrence Tao, *Two announcements: AI for Math resources, and erdosproblems.com*, April 19, 2024. https://terrytao.wordpress.com/2024/04/19/two-announcements-ai-for-math-resources-and-erdosproblems-com/

[Tho] Rachel Thomas, Twitter post, August 22, 2020, accessed March 15, 2025. https://twitter.com/math_rachel/status/1297255819965169664

[Tho21] Rachel Thomas, *11 short videos about AI ethics*, fast.ai, August 16, 2021. https://rachel.fast.ai/posts/2021-08-17-eleven-ethics-videos/

[TMP12] Amanda L. Traud, Peter J. Mucha, and Mason A. Porter, *Social structure of Facebook networks*, Physica A **391** (2012), no. 16, 4165–4180.

[TPB24a] Rochelle E. Tractenberg, Victor Piercey, and Catherine A. Buell, *Defining "ethical mathematical practice" through engagement with discipline-adjacent practice standards and the mathematical community*, Science and Engineering Ethics **30** (2024), 15.

[TPB24b] Rochelle E. Tractenberg, Victor Piercey, and Catherine A. Buell, *Proto ethical guidelines for mathematical practice*, 2024. https://osf.io/x5ur9/

[UCLa] University of California, Los Angeles, *Data Theory at UCLA*, accessed December 26, 2023. https://datatheory.ucla.edu

[UCLb] University of California, Los Angeles, *UCLA: Generative AI*, accessed April 21, 2024. https://genai.ucla.edu/

[Uga17] Johan Ugander, *Truth, lies, and an ethics of personalization*, Medium, January 23, 2017. https://medium.com/@jugander/truth-lies-and-an-ethics-of-personalization-e4ccfa7f2b84#.rzap3hm70

[Uga20] Johan Ugander, *Data privacy and data ethics*, Management Science and Engineering 234, Stanford University, 2020. https://web.stanford.edu/group/msande234/cgi-bin/wordpress/

[Uga23] Johan Ugander, *Social algorithms*, Management Science and Engineering 231, Stanford University, 2023. https://msande231.github.io/syllabus

[Vé21] Carissa Véliz (ed.), *The Oxford handbook of digital ethics*, Oxford University Press, Oxford, UK, 2021.

[Vin16] James Vincent, *Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day*, The Verge, March 24, 2016. https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist

[VO23] Pranshu Verma and Will Oremus, *ChatGPT invented a sexual harassment scandal and named a real law prof as the accused*, The Washington Post, April 5, 2023. https://www.washingtonpost.com/technology/2023/04/05/chatgpt-lies/

[Wei21] Zach Weinersmith, *Software*, Saturday Morning Breakfast Cereal, September 3, 2021. https://www.smbc-comics.com/comic/software

[Whi12] Martha C. White, *Orbitz shows higher prices to Mac users*, Time, June 26, 2012. https://business.time.com/2012/06/26/orbitz-shows-higher-prices-to-mac-users/

[Wika] Wikipedia, *Creative Commons license*, accessed March 15, 2025. https://en.wikipedia.org/wiki/Creative_Commons_license

[Wikb] Wikipedia, *Netflix Prize: Cancelled sequel*, accessed December 26, 2023. https://en.wikipedia.org/wiki/Netflix_Prize#Cancelled_sequel

[Wikc] Wikipedia, *Social Credit System*, accessed December 26, 2023. https://en.wikipedia.org/wiki/Social_Credit_System

[Wikd] Wikipedia, *Timnit Gebru: Exit from Google*, accessed December 26, 2023. https://en.wikipedia.org/wiki/Timnit_Gebru#Exit_from_Google

[WJ23] Chris Wiggins and Matthew L. Jones, *How data happened: A history from the age of reason to the age of algorithms*, W. W. Norton & Company, Inc., New York, NY, USA, 2023. MR4795998

[WSAO+21] Claudia Wagner, Markus Strohmaier, Emre Kıcıman Alexandra Olteanu, Noshir Contractor, and Tina Eliassi-Rad, *Measuring algorithmically infused societies*, Nature **595** (2021), 197–204.

[WT24] Stephen M. Walk and Rochelle E. Tractenberg, *Helping students deal with ethical reasoning: The proto-Guidelines for Ethical Practice in Mathematics as a deck of cards*, arXiv:2403.16849, 2024.

[WW24] Madi Whitmann and Chris Wiggins, *Data: Past, present, and future*, History-APMA UN2901, Columbia University, 2024, accessed December 31, 2023. https://data-ppf.github.io

Department of Mathematics, University of California, Los Angeles, Los Angeles, California 90095, USA; Department of Sociology, University of California, Los Angeles, Los Angeles, California 90095, USA; and Santa Fe Institute, Santa Fe, New Mexico, 87501, USA

Email address: mason@math.ucla.edu

Proceedings of Symposia in Applied Mathematics
Volume **80**, 2025
https://doi.org/10.1090/psapm/080/00729

Uncertainty in criminal justice algorithms: Simulation studies of the Pennsylvania Additive Classification Tool

Swarup Dhar, Vanessa Massaro, Darakhshan Mir, and Nathan C. Ryan

Abstract. Much attention has been paid to algorithms related to sentencing, the setting of bail, parole decisions and recidivism while less attention has been paid to carceral algorithms, those algorithms used to determine an incarcerated individual's lived experience. In this paper we study one such algorithm: the Pennsylvania Additive Classification Tool (PACT) that assigns custody levels to incarcerated individuals. We analyze the PACT in ways that criminal justice algorithms are often analyzed: namely, we train an accurate machine learning model for the PACT; we study its fairness across sex, age and race; and we determine which features are most important. In addition to these conventional computations, we propose and carry out some new ways to study such algorithms. In particular, instead of focusing on the outcomes themselves, our approach shifts the attention to the variability in the outcomes. Many carceral algorithms are used repeatedly and there can be a propagation of uncertainty; we develop an approach to describe the propagation of uncertainty using simulation studies and sensitivity analyses. The results from our new approach and the conventional approaches shine light on problematic aspects of the PACT.

1. Introduction

As has been well-established, the use of algorithms in decision making, low- and high-stakes decisions alike, is pervasive in industry and, increasingly, in government. Particular domains where decisions are made using algorithms include online advertising, lending and banking, pretrial detention, to name a few. These decisions are based on predictions which are themselves based on data. A great deal has been written on biases that are manifest in these processes: biases that emerge from sampling issues and measurement error and biases in outcomes. There have been many articles written on various approaches to measuring the bias and, conversely, the fairness of decision making processes in which, often, the details of the algorithm or the predictions that undergird the decision are unknown. See [**12**] for an excellent and insightful review of the various ways bias and fairness have been measured and interpreted, as well as measures of fairness based only on data, and those based on various kinds of models.

2020 *Mathematics Subject Classification.* Primary 11F33, 11F37.

Key words and phrases. Algorithmic fairness, criminal justice, sensitivity analysis, counterfactual fairness, causal models.

An aspect of many of these decision-making algorithms is that they are often reapplied to the same individual at various times. Lum and Isaac [**10**] describe the impact of training a predictive policing algorithm on biased data. They describe a feedback loop in predictive policing of drug crimes: in particular, they observe that using algorithms to determine where to police, results in over-policed communities becoming even more disproportionately over-policed. Ensign, *et al* [**5**] have used Polya urn models to give a mathematical explanation for these feedback loops.

Studying feedback loops is one way to get a handle on the variability in decisions that is generated with the reapplication of predictive algorithms. Another approach to understanding the variability in decisions is to use sensitivity analysis and simulation. Blumstein [**1**] used this approach to model complete criminal justice systems and to understand which variables most influenced the cost and the flow of offenders through the system. A more recent study that carries out a similar analysis but on a larger scale can be found in [**3**]. Moranian *et al* [**13**] carried out a simulation study for juvenile courts to determine which variables influenced the rate of flow most strongly.

In addition to these two particular approaches, there has recently been a large number of simulation studies of the criminal justice system. For example, Cortés and Ghosh propose an agent based model in [**2**]. A recent book by Liu and Eck [**9**] is dedicated to crime simulation using GIS and various mathematical simulation methods.

In this paper, we carry out simulations to understand a particular tool used by the Pennsylvania Department of Corrections (PADOC). This tool is called the Pennsylvania Additive Classification Tool (PACT) and is used by the PADOC to assign a custody level to an incarcerated person. The authors have recently carried out a historical and meta-analysis of the data used by this tool; see [**11**]. There we point out that custody-level determination by PACT is biased by race, that the data is problematic both in how much is missing and in how high its quality is deemed to be despite it containing errors; that the algorithm uses biased input; and that the tool itself has competing goals as it is both supposed to reflect the incarcerated person's rehabilitation and their securitization by the PADOC. The PACT is used for both the initial classification of an incarcerated person into a custody level and the annual reclassification process. We carry out 9 simulation studies of which 6 are about initial classification and the remaining 3 are about repeated reclassification.

Figure 1 summarizes the data and the steps that go into the PACT. Our interest is in the certainty one can reasonably have in the output of a potentially biased tool like the PACT; we describe methods to quantify or describe the uncertainty one should reasonably hold about the tool's output. In particular, we use conventional notions of fairness, simulation studies, sensitivity analysis and counterfactual fairness to quantify the disparate impact the results of the PACT. The data provided to us by the PADOC contained most of these variables. Two exceptions are labeled "Severity of current offense" and "Severity of criminal history": while we were given individual offense gravity score and prior record scores we were not given a single, holistic score to summarize the severity of the offense and the person's criminal history. A third exception is labeled "Stability factors": while a partial list of stability factors is provided in the PADOC documentation, some may have been omitted and there is, again, not a single, holistic variable that measures

the quality and quantity of a person's stability factors. A final exception is labeled "Group affiliation" which the PADOC stopped collecting more than a decade ago.

The variable "Escape history" in Figure 1 warrants some explanation. An escape, according the PADOC, is any occasion where an incarcerated person is not where they are supposed to be at a certain time once they are in the custody of the PADOC. This can range from an actual escape from a carceral institution to not being in their cell at lights out. Since this variable is about how well the incarcerated person adheres to the rules of the PADOC and is dependent on the actual institution in which they are incarcerated, it makes sense for it to appear in both the classification step at the level of the PADOC and the override step at the level of the institution in which they are incarcerated.

The variable "Institutional adjustment" in Figure 1 also warrants some explanation. This variable takes on values from 1 to 4 and reflects the PADOC's sense of how well the incarcerated person will adjust to being incarcerated. This number represents a summary of a person's risk factors as identified in screening interviews, a person's scores on standardized tests, a person's housing needs, etc. A score of 1 means that the person should not have much trouble adjusting while a score of 4 means that they will.

The other variables are self-explanatory and a more detailed description of these variables can be found in Table 1.

In addition to our findings about this particular tool, our unified approach is a way to study and understand the uncertainty of any algorithm whose inner workings are obscured to the public. While much work has been done on bias in outcomes and fairness in decisions, we propose that studying the uncertainty and variability of a model can provide different insights about its impact.

The paper is organized as follows. In Section 2 we describe the PACT in more detail and the data set we are working with. In Section 3 we train a random forest model on our data and then use it to carry out the 6 simulation experiments related to initial classification and the 3 related to repeated reclassification. In Section 4 we carry out fairness analyses, including a measure built around counterfactuals. We conclude with a discussion and description of future directions.

2. The PACT

The PACT is an example of a *carceral* algorithm; that is, an algorithm that determines an incarcerated person's experience during incarceration. Most algorithms studied in the context of criminal justice are related to parole, bail and sentencing. Studying a carceral algorithm like the PACT presents a number of challenges, including gaining access to the data. The PACT was introduced by the PADOC in 1991 and generates a raw score that determines an incarcerated person's custody level, ranging from a custody level of 1 (community corrections) up to 5 (maximum security). The Pennsylvania state documents (see [**14**]) related to PACT tell us that the algorithm "is confidential and not for public dissemination" and so analyzing it fully is impossible. From these same documents we were able to develop the flowchart in Figure 1.

The assignment of custody levels has four main steps. First, the PACT tool is applied to data transmitted to the PADOC and an initial score is derived and then that score is turned into a number from 1 to 5, representing that person's custody level. At that point, the particular prison can decide to override this score either for

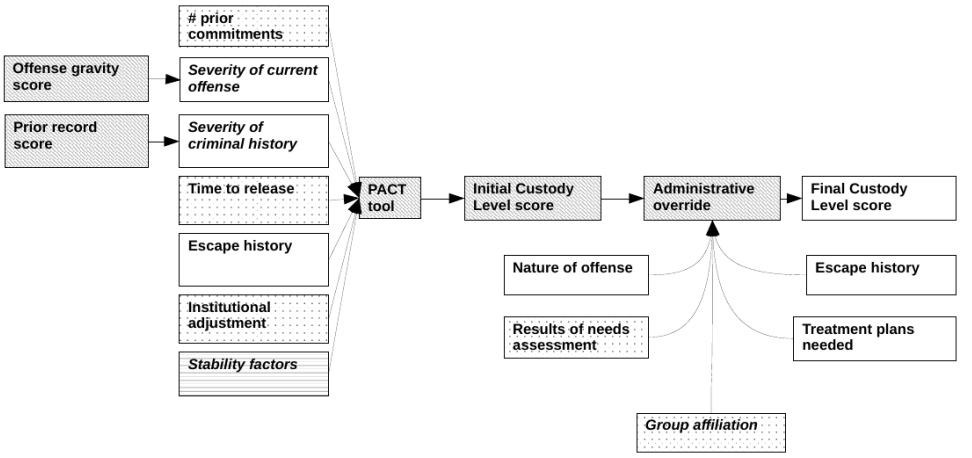


FIGURE 1. A schematic description of how an incarcerated person's custody level score is determined. The process has two main parts: a score determined by the PACT and then an override. Variables in boxes with a background of diagonal lines indicate a lack of transparency about how that variable is calculated; variables in boxes with a background made of little dots are ones where the values are known to have a societal bias; the variable in the box whose background is horizontal lines represents several variables (e.g., marital status, employment status, etc.) that are somewhat arbitrary; and variables in boxes with a white background are ones that are either not known to be biased or whose method of calculation is known. Variables whose labels are written in italics were not explicitly in the data set given to us by the PADOC.

administrative reasons (e.g., number of beds) or for other reasons at the discretion of the particular prison. At the end of each year a similar process of reclassification is carried out. A reclassification score is determined algorithmically and then, once again, the prison can decide to override that score for either administrative or discretionary reasons.

2.1. Summary of the data. In July 2018 we requested a data pull from the PADOC using the Pennsylvania Department of Corrections Research Approach Request Form (RARF). The intent of the request was to study factors that influence parole decisions. The PACT is one such factor and, because of its importance, we decided to study it more closely. We requested data on incarcerated people, including those who have been paroled, who were in the system in 1997, 2002, 2007, 2012, and 2017. We requested variables that were related to parole decisions. Nine months after our initial request we received access to data on more than 280,000 distinct incarcerated people; of those only 146,793 were incarcerated in the years we requested. See Figure 2 for a graph of the distributions of the demographic variables of the people in our data set.

2.2. Descriptions of variables. Most of the variables that we used in our models (see Table 1) are self-explanatory and map unambiguously on to the steps laid out in Figure 1. There are some that merit more explanation and these were discussed above.

In principle (as opposed to in the data), every incarcerated person should have a prior record score and the crimes for which they have been most recently committed should have an offense gravity score. A prior record score is a number from 1 to 4 that indicates a person's criminal history and the number is identified by a statute number in the Pennsylvania Criminal Code. In the data that we have, we are given the statute number but not the person's prior record score and each statute number is often associated to a range of numbers. An offense gravity score is similar but ranges from 1 to 15 and indicates the severity and nature of the crime for which the person has been committed. Like prior record scores, offense gravity scores are not given in the data but a statute is given and the statute, again, has a range of scores associated to it. In both of these cases, we use the maximum scores associated to a statute, as described in Table 1.

Also in Table 1, we take two different approaches to reporting a person's age. For initial classification we use binary variables to describe a person's age but in reclassification we use the numerical age. This is because of how we trained the reclassification models: in our simulations we apply these models yearly and so having a finer handle on a person's age is necessary in order to have as good a model as possible.

Finally, a difference between reclassification and initial classification is that in initial classification a person's institutional adjustment is determined in an opaque way. There is no direct analogue for reclassification and so we use the number of disciplinary reports a person has received, instead. While the number of disciplinary reports is not opaque, it has been shown, in some cases, that the writing of such reports is biased (see, for example, [**16**]).

3. Simulation experiments and sensitivity analysis

In this section we describe simulation experiments that capture the uncertainty of the PACT procedure. We model the PACT procedure without including the override processes for several reasons. The main practical reason is there is a lot of missing data and if we limit our model to those incarcerated people for which we have all the necessary decisions (initial classification and an override and reclassification and an override), there would only be a handful of people left to build and test the model on. A philosophical reason is that we are interested in the workings of the algorithm and not the incidental and subjective contribution that arises in the override process.

3.1. Random forest models for initial classification and reclassification. In our experiments, we use random forests to model both the initial classification and reclassification. Our code is written in Python and uses Scikit-learn [**15**]; our code can be found at [**4**]. The variables we use to train the models are described in Table 1.

A random forest classifier is an ensemble learning method that constructs many decision trees and outputs the class that is selected by the largest number of decision trees. Each decision tree in the random forest is a model of the classification process and is built by splitting the input based on a set of splitting rules based

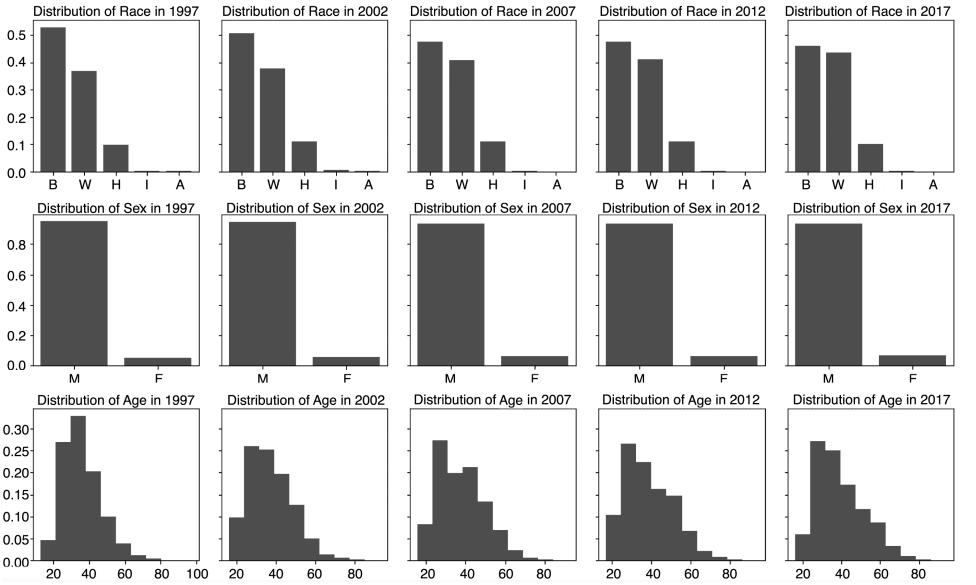


FIGURE 2. Basic demographic data. In the first row, we see the distribution of the race or ethnicity of the incarcerated people in 1007, 2002, 2007, 2012 and 2017. Here B means Black, W means White, H means Hispanic, I means American Indian, and A means Asian and Pacific Islander. The second row gives the distribution of the sex of incarcerated people in the same years. The third row gives the distribution of the ages.

on classification features. There are several ways to measure the quality of a split and in our models we use the Gini impurity, or just impurity. The Gini impurity of a split is the probability of classifying an observation into the wrong class. The lower the impurity, the better the split; that is, the lower the impurity, the lower the likelihood of misclassification.

A random forest classifier is an appropriate choice for this classification task because of some of the advantages inherent to the classifier. First, random forests handle potential interactions well because the splits can (and most likely will) be done on each variable separately and there are likely interactions between many of our variables. Second, random forests handle outliers well since values on the same side of the split are treated equally (e.g., if a split is done on time served and the split is done at 10 years, then people who have been incarcerated for 11 years are treated by the model the same way as people who have been incarcerated 50 years). Third, you often get a high accuracy without having to spend time fine tuning the model's hyperparameters: since the point of this study is not to build the best model but to see the effects of applying the model repeatedly, this advantage was important. We calculated a model's accuracy as the ratio of the number of correct predictions in our prediction set to the total number of predictions in the prediction set; that is, the number of the true positives and true negatives, divided by the sum of the numbers of all positives and negatives.

TABLE 1. A summary of the variables used in the construction of the random forest. The third and fourth columns are, respectively, the importance (as measured by mean decrease in impurity) of the variable in the random forest models we use for initial classification and reclassification; taken together these importance scores also indicate which variables were used in which model.

Variable	Description	IC	RE
gender_female	A binary variable indicating whether the person is identified as female (1) or male (0)	0.010365	0.005438
age_gt_45	A binary variable indicating whether the person is older than 45 years old (1) or not (0)	0.046948	
age_lt_25	A binary variable indicating whether the person is younger than 25 years old (1) or not (0)	0.026485	
age	A quantitative variable used in the reclassification model		0.188189
race_B	A binary variable indicating whether the person is identified as Black (1) or not (0)	0.029642	0.016741
race_A	A binary variable indicating whether the person is identified as Asian of Pacific Islander (1) or not (0)	0.000863	0.000351
race_H	A binary variable indicating whether the person is identified as Hispanic (1) or not (0)	0.014115	0.011085
race_I	A binary variable indicating whether the person is identified as American Indian (1) or not (0)	0.000223	0.000403
race_O	A binary variable indicating whether the person is identified as belonging to a Nonwhite race or ethnicity other than Black, Asian, Hispanic or American Indian (1) or not (0)	0.001427	0.001103
off_1_prs_max	A quantitative variable representing a person's maximum prior record score	0.083801	0.062925
off_1_gs_max	A quantitative variable representing the maximum gravity score a person could have received	0.203287	0.089114
ic_custdy_level	A person's initial custody level upon entering the carceral system		0.038942
prior_commits	The number of times the person has been previously committed to the PADOC	0.148917	0.10117
ic_institut_adj	A person's institutional adjustment score used during initial classification	0.235557	
re_discip_reports	The number of disciplinary reports a person was given since the previous reclassification		0.434796
escape_hist_1	A binary variable representing whether there was an attempted escape	0.018138	0.010878
escape_hist_2	A binary variable representing whether there was a second attempted escape	0.015903	0.010201
escape_hist_3	A binary variable representing whether there was a third attempted escape	0.014793	0.004418
escape_hist_4	A binary variable representing whether there was a fourth attempted escape	0.015923	0.013931
escape_hist_5	A binary variable representing whether there was a fifth attempted escape	0.008083	0.010316
mrt_stat_DIV	A binary variable representing whether the person was divorced at the time of commitment (1) or not (0)	0.033823	
mrt_stat_SEP	A binary variable representing whether the person was separated at the time of commitment (1) or not (0)	0.020496	
mrt_stat_MAR	A binary variable representing whether the person was married at the time of commitment (1) or not (0)	0.028329	
mrt_stat_WID	A binary variable representing whether the person was widowed at the time of commitment (1) or not (0)	0.006492	
employed	A binary variable representing whether the person was employed at the time of commitment (1) or not (0)	0.036389	

The importance of each variable (as measured by mean decrease in impurity) is also listed in Table 1. We observe that for initial classification the most important variables are, in order, the person's institutional adjustment, the gravity score of the offense committed by the person, the number of previous times the person has been committed to the PADOC and the person's prior record score. We observe that a person's race, age and marital status also appear to be important. For reclassification, we find that the number of disciplinary reports a person gets is far and away the most important, with the person's age being the second most important. After these two variables, the prior record score, previous commitments, gravity score are the next most important.

The random forest model for initial classification has an accuracy of 0.79 and the one for reclassification has an accuracy of 0.78.

3.2. Experiments. In this section we describe the experiments we carry out. The goal of these experiments is to explore what happens when a new observation is pushed through the initial classification model. These new observations are very similar to people in the data set but have been changed in particular ways. While we have also trained a model for the reclassification process, in most of this section we report only on the results for the initial classification of people. We do this for two reasons: (1) when we start subsetting the data, we end up with too little data on which to train a reclassification model and (2) we want to focus on the uncertainty at the very beginning of the process.

3.2.1. *Experiment* 1. In this first experiment, we take all people at a fixed custody level, and construct a new sample space. This new sample space consists of the same variables but the values for each variable come from the multiset of the people in the fixed custody level. Then, to sample from this new sample space, we randomly sample a value for each variable from each multiset. An observation generated in this way has values for each variable that someone else in the same custody level has and so, in this sense, this new observation could have been in this data set but differs from people in the data in some number of variables.

In Figure 3 we report changes in initial custody level for 100 new observations in each custody level, generated as described above. We observe that a fairly large fraction of synthetic observations in each custody level would end up in a different custody level when classified with the model we trained at the beginning.

3.2.2. *Experiment* 2. This experiment is very much like Experiment 1, where, again, we fix a custody level. In this experiment we generate a synthetic observation by taking a person in this custody level, randomly choosing a single variable and then choosing a new value for that variable from the multiset of values of all people at that custody level. The synthetic observations, then, will be different from a person in the data in exactly one randomly chosen column. This experiment is more controlled than Experiment 1.

In Figure 4 we observe that in initial classification, there is a lot of variability in how a small change to a person in custody level 5 changes their custody level. For custody levels 2, 3 and 4, there is less variability but there is still a significant number of people for whom a small tweak would send them to a different custody level.

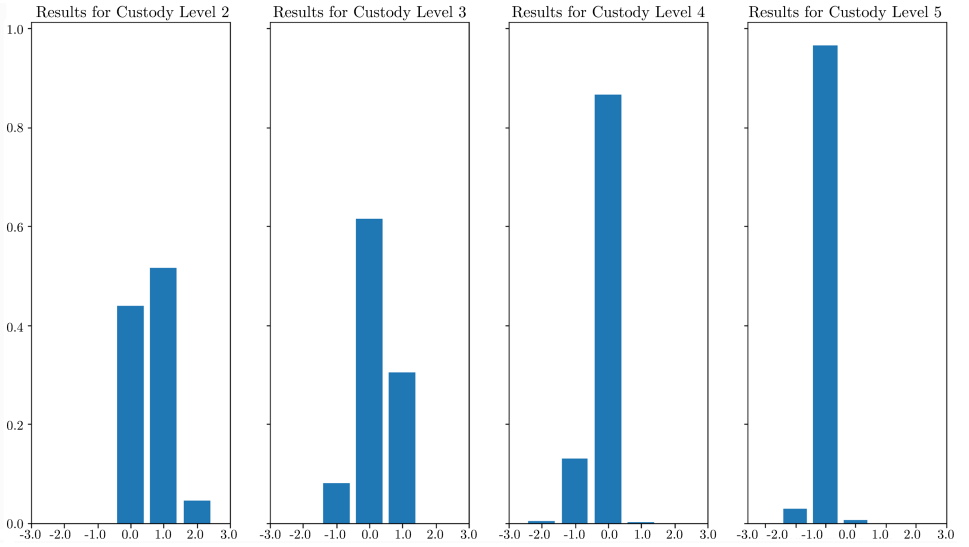


FIGURE 3. Results of Experiment 1 for initial classification. The bar charts report the frequency with which a synthetic observation in a particular custody level changes to a different custody level and by how many custody levels they would change (e.g., -3.0 means the synthetic observation is classified by our random forest model to three custody levels lower than the one the synthetic observation was generated from).

3.2.3. *Experiment* 3. In this third experiment, we proceed as in Experiment 1, but now first stratify for a person's race and then generate the multisets for each race. In particular, we examine the changes in custody level for people identified as being Black and for those identified as being White.

In Figure 5, we observe the following. If the person is Black and in a lower custody level, they are, in this experiment, more likely to be classified to a lower level or to be kept at the same level. On the other hand, if the person is White, they are, in this experiment, more likely to be classified higher or to be kept at the same custody level. Conversely, if the person is at a high custody level, they will be, at least in the context of this experiment, classified to a lower level than they are to be kept at the same level or classified to a higher level. The proportion of Black inmates who would be classified in this experiment to a custody level of 2 or 3 is greater than it is for White inmates.

3.2.4. *Experiment* 4. In this fourth experiment, we proceed as in Experiment 2, but, like in Experiment 3, we now stratify according to a person's race. In particular, we examine the changes in custody level for people identified as Black and for people identified as White.

In Figure 6 we see that when we change one column for Black people in our data set who are at custody level 5, there is very large range of custody levels for the resulting synthetic observation whereas for White people for whom we change a single column, the range is much smaller. We also observe that, unlike in Experiment 3, in this more controlled Experiment 4, we see that people who are identified

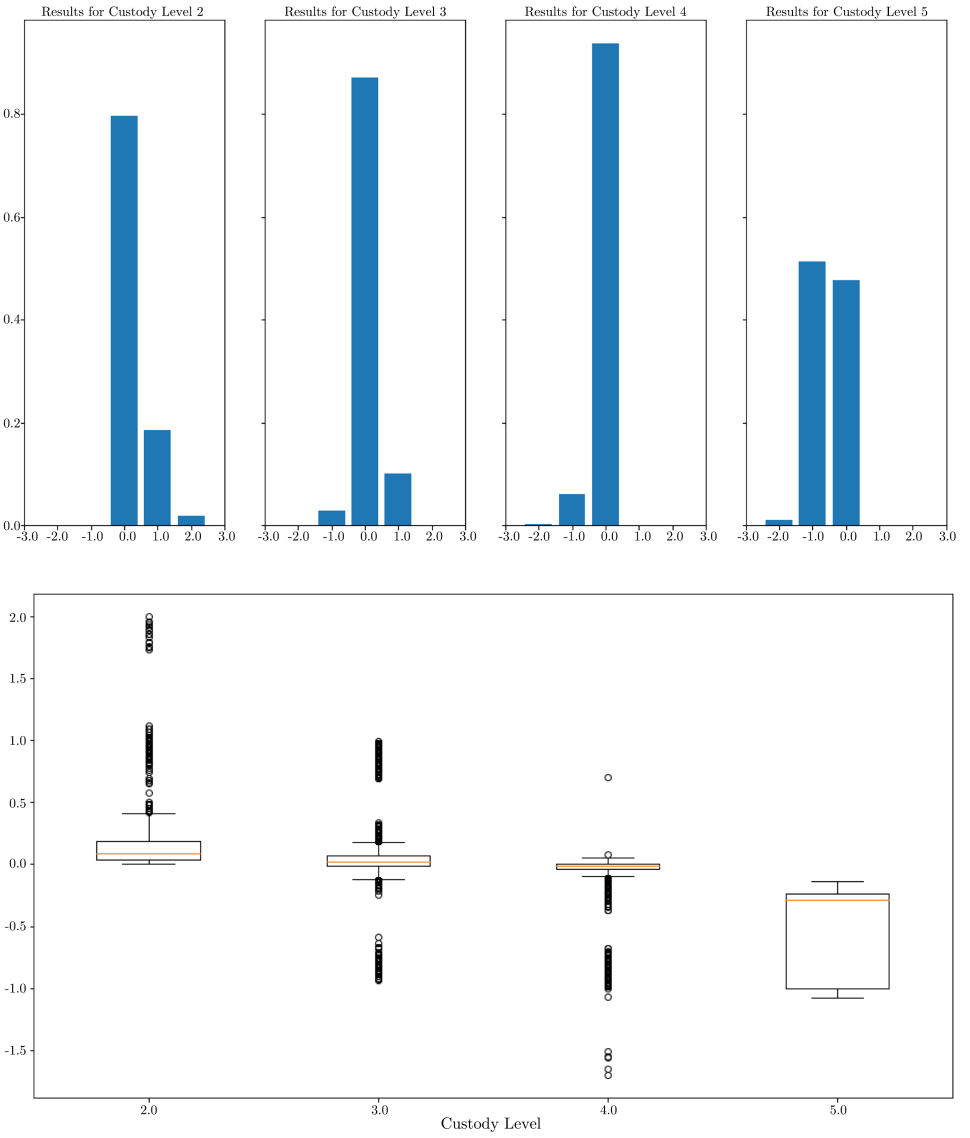


FIGURE 4. Results of Experiment 2 for initial classification. The bar chart reports the frequency with which a synthetic observation in a particular custody level changes to a different custody level and by how many custody levels they would change (e.g., −2.0 means the synthetic observation is classified by our random forest model to three custody levels lower than the one the synthetic observation was generated from). The box plot shows the new custody after changing a single variable for a person and then applying the random forest model for initial classification.

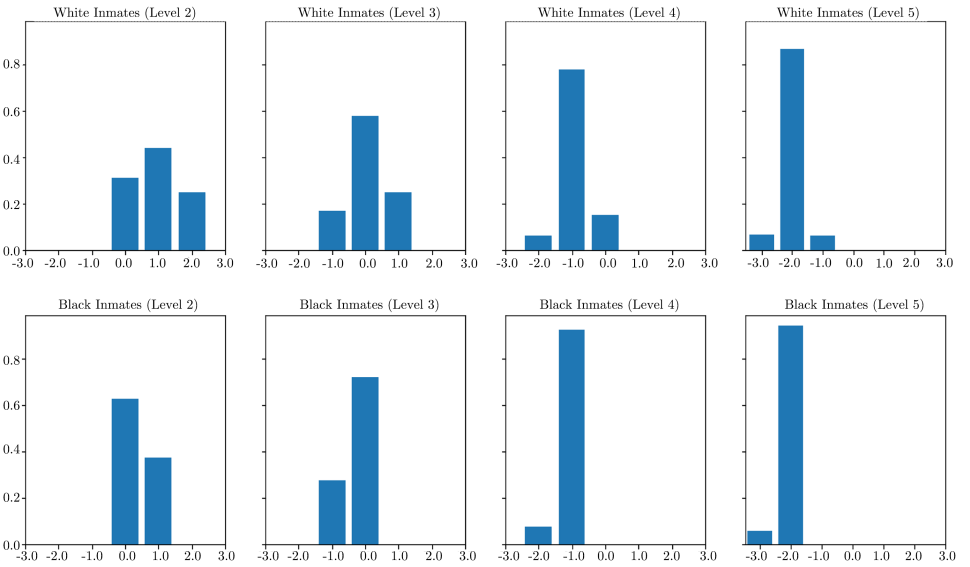


FIGURE 5. Results of Experiment 3 for initial classification for the Black population (bottom) and for the White population (top). The bar charts report the frequency with which a synthetic observation of a particular race and in a particular custody level changes to a different custody level and by how many custody levels they would change (e.g., -3.0 means the synthetic observation is classified by our random forest model to three custody levels lower than the one the synthetic observation was generated from).

as Black and in custody levels 1 through 4, are more likely to be kept at the same custody level or to be moved to a higher custody level as a result of changing one column, as compared to similar people who are identified as White.

3.2.5. *Experiment* 5. The previous four experiments have perturbed observations to make synthetic observations by sampling from the data that we already have. In this fifth experiment, we perturb in a different way. In this experiment we fix all the categorical variables for an observation and generate a new synthetic observation by perturbing the quantitative variables. We do this by calculating the margin of error at a 95% confidence level and then uniformly sampling each quantitative variable from the interval whose radius is equal to the margin of error and whose center is the person's observed value in the data set. For instance, if a person's insitutional adjustment score was 4 and the standard deviation of 100 randomly chosen institutional adjustment scores was 0.12, then we would find the margin of error as

$$t^{\star}_{\alpha/2,df}\, s/\sqrt{n} \approx 1.98 \times 0.12/\sqrt{100} = 0.023\ldots$$

We present this experiment because, while it is trying to capture the same basic idea, it is done in a different way.

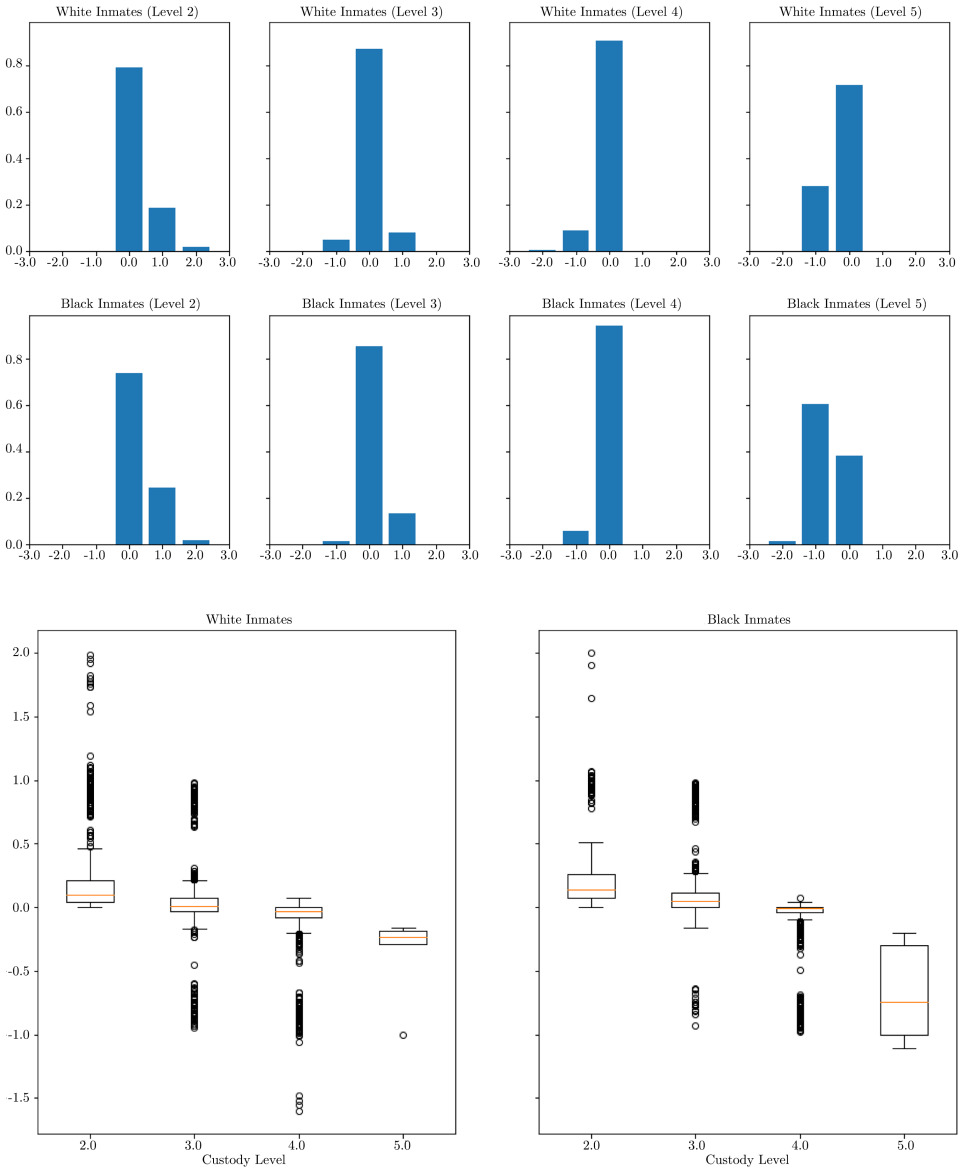


FIGURE 6. Results of Experiment 4 for the Black population and the White population (top). The bar charts report the frequency with which a synthetic observation in a particular custody level changes to a different custody level and by how many custody levels they would change (e.g., −2.0 means the synthetic observation is classified by our random forest model to three custody levels lower than the one the synthetic observation was generated from).

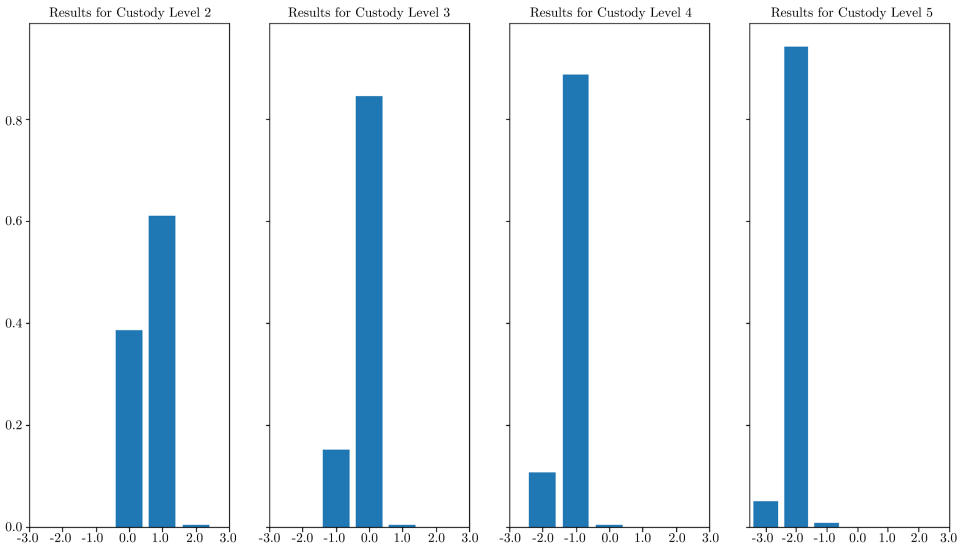


FIGURE 7. Results of Experiment 5 for initial classification. The bar chart reports the frequency with which a synthetic observation with fixed values for their categorical variables and in a particular custody level changes to a different custody level and by how many custody levels they would change (e.g., −3.0 means the synthetic observation is classified by our random forest model to three custody levels lower than the one the synthetic observation was generated from).

In Figure 7 we observe that the results are rather different and that, in fact, in all but custody level 3, we see that the vast majority of synthetic observations would be classified to a different level than the person from whom they were generated was assigned.

3.2.6. *Experiment* 6. Our sixth experiment that measures the uncertainty of the PACT tool (defined to be the variability caused by small perturbations to the data) is a sensitivity analysis in which we determine what effect a 10% increase and decrease to the four most influential variables (namely, the person's prior record score, their offense gravity score, the number of prior commits and their institutional adjustment) has on the predicted custody level.

In Figure 8, we observe some interesting results. First, we note that a small change in a person's prior record score has no apparent impact on the average custody level. Second, we observe that a small change in the number of prior commitments to PADOC a person has also does not have any or much apparent impact on the average custody level. Third, we see that if a person's custody level is low, a ten percent increase in their associated gravity score increases their custody level and, if their custody level is high, a ten percent decrease lowers their custody level. Fourth, we see that, for the most part, institution adjustment has little impact on their custody level except in the case when a person was in custody level 5 and their institutional adjustment score is decreased by 10%. See Table 2 for more details.

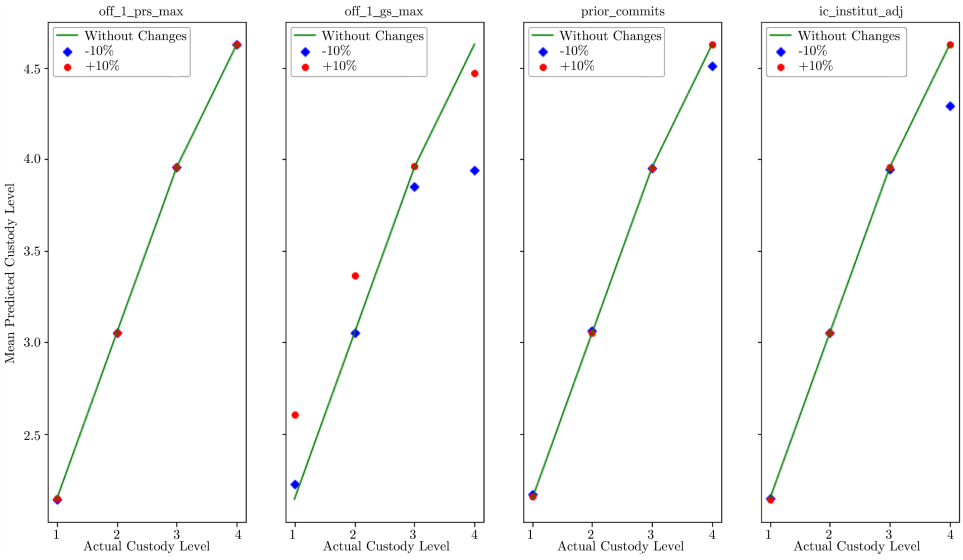


FIGURE 8. Results of Experiment 6 for initial classification. Each panel reports the average custody level before any changes (green line), after a 10% increase in the variable corresponding to that panel (a red dot) and after a 10% decrease in the variable corresponding to that panel (a blue diamond). Each dot or diamond corresponds to a different initial custody level in the data.

TABLE 2. We report the relative percent change imparted on custody level by ten percent decreases or increases on the four most important quantitative variables. An entry in the table that reads $< 0.1\%$ means that the relative percent change was negligible; in particular, that it was between -0.1% and 0.1%.

Variable change	CL 2	CL 3	CL 4	CL 5
10% dec. off_1_prs_max	$< 0.1\%$	$< 0.1\%$	$< 0.1\%$	$< 0.1\%$
10% inc. off_1_prs_max	0.1%	$< 0.1\%$	$< 0.1\%$	$< 0.1\%$
10% dec. off_1_gs_max	3.9%	-0.1%	-2.7%	-13.2%
10% inc. off_1_gs_max	22.1%	10.3%	0.2%	-3.8%
10% dec. prior_commits	0.8%	0.4%	-0.2%	-3.8%
10% inc. prior_commits	0.7%	0.1%	-0.1%	-0.4%
10% dec. ic_institut_adj	0.4%	0.1%	-0.2%	-7.2%
10% inc. ic_institut_adj	$< 0.1\%$	$< 0.1\%$	$< 0.1\%$	$< 0.1\%$

3.3. Repeated reclassification. In this section we describe preliminary work in which we attempt to measure and capture the feedback loops that arise with the periodic reclassification that the PADOC carries out. In this simulation, we sample a fixed number of people in each custody level and apply the reclassification algorithm and then update data related to the passage of time. Since we do not have annual data for any individual person, we increment a person's age and update their previous custody level to be the prediction of the previous application of our random forest model. We fix all the other variables and in this way isolate the effects of time and previous custody level.

We carry out these simulations to get a glimpse of how repeated reclassification works and the impact that it has on individuals and groups (groups of people in the same custody level and groups of people given the same racial identity). We acknowledge the limitations here that arise from the nature of our data (namely that our data is only a snapshot of a single year but we are predicting custody levels over several years), but believe that this could be a first step in empirically studying the feedback loops that have been studied theoretically in [**5**].

In Figure 9 we see the trajectories of 10 individuals starting at each of the four custody levels. This means that we have 10 individuals in, say, custody level 4 and we apply the reclassification model to each of them. If the model predicts they change custody level, we update their custody level and increment their age and, if not, we only update their age. We repeat this process 8 more times and make a plot of the resulting time series. In this particular image we see, for example, that a person who started at custody level 5 is quickly dropped to custody level 3 and then alternates between levels 3 and 4 for the rest of the time.

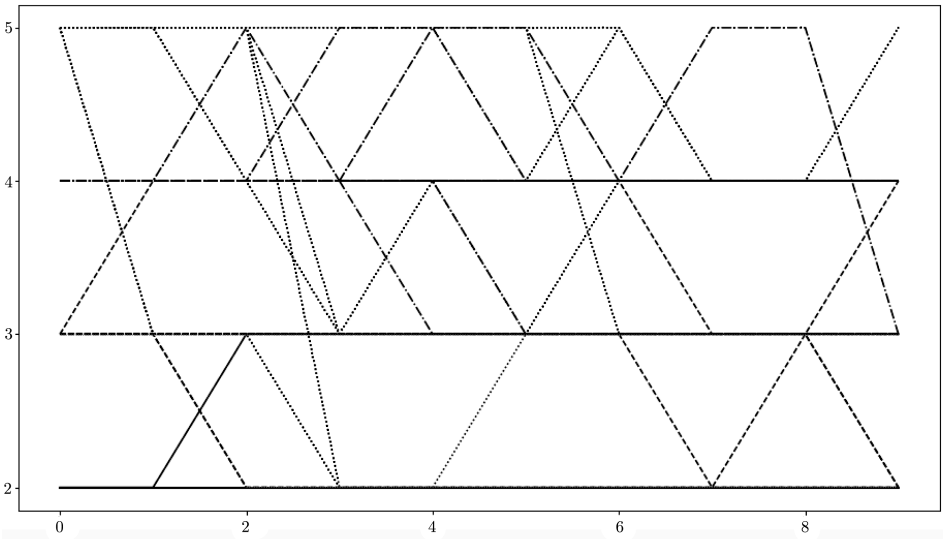


FIGURE 9. The trajectories of 10 individuals at each of the four custody levels. We repeatedly apply the random forest model for reclassification, updating the person's age and previous custody level. Each color represents a different person. While it is hard to follow any individual, it is easy to see that the process is somewhat volatile at the level of individuals.

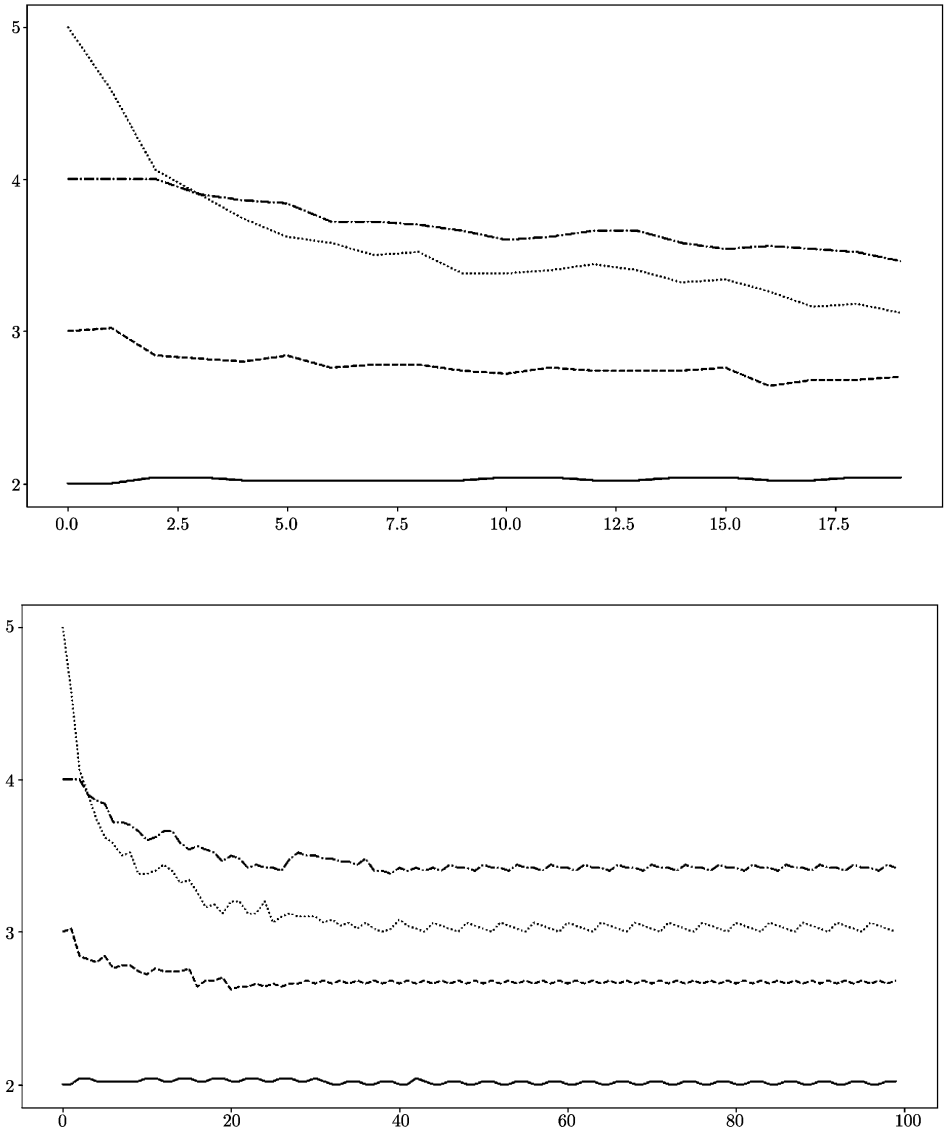


FIGURE 10. The average trajectories of 50 individuals at each of the four custody levels. We repeatedly apply the random forest model for reclassification, updating the person's age and previous custody level and average across the 50 people who started at each custody level. We present both the initial behavior over the first 20 years (top) and the long term behavior (bottom).

Due to the volatility in Figure 9, we also produced Figure 10 in order to understand the aggregate behavior of people who start in each custody level. We know that as people get older, they tend to be classified at a lower level and so it is not surprising that there is a downward trend for the groups other than those who started at custody level 2. What is a little surprising is that in the long run, individuals who start at custody level 5, on average, end up at a lower custody level than those who start at custody level 4. This may be due to how custody level 5 is used by the PADOC and the data we have received (e.g., maybe someone at custody level 5 in our data set is someone who happened to be in solitary confinement when our data was being collected but otherwise is in a lower custody level), but this would require further understanding of how PADOC uses custody levels.

3.3.1. *Differences by race.* In order to study differences in these outcomes by race, we made Figure 11 in which we take a sample of 100 Black incarcerated people in each custody level and 100 White incarcerated people in each custody level, reclassify and then find the average custody level for each sample. In addition to this visualization, we also decided to measure the variability of a person's trajectory in the following way. Every time we reclassify a person, we keep a running tally of the absolute value of how many custody levels they change year to year. So, for example, someone changing from level 3 to 4 would add 1 to the running total and someone changing from 4 to 2 would add 2 to the running total. For each individual, we calculate these totals and then take the average change per year and for each person in the simulation. These numbers are reported in Table 3.

In Figure 11 and in Table 3 we observe two things. First, they both have the overall downward trajectory that we expect because as people age they tend to be classified to a lower custody level. Second, Black people have more volatility in their trajectories. For instance, Black people who start a custody level 2, change custody levels 10 times faster than White people who start at custody level 2. Black people who start at custody level 4 tend to change custody levels 50% faster than White people who start at custody level 4.

TABLE 3. Average weighted number of changes in custody level per person, per year for 100 people starting at each custody level and after applying the reclassification 8 times. We separated it by Black and White people for the sake of comparison.

Initial custody level	Black people	White people
2	0.034	0.003
3	0.087	0.099
4	0.135	0.095
5	0.247	0.220

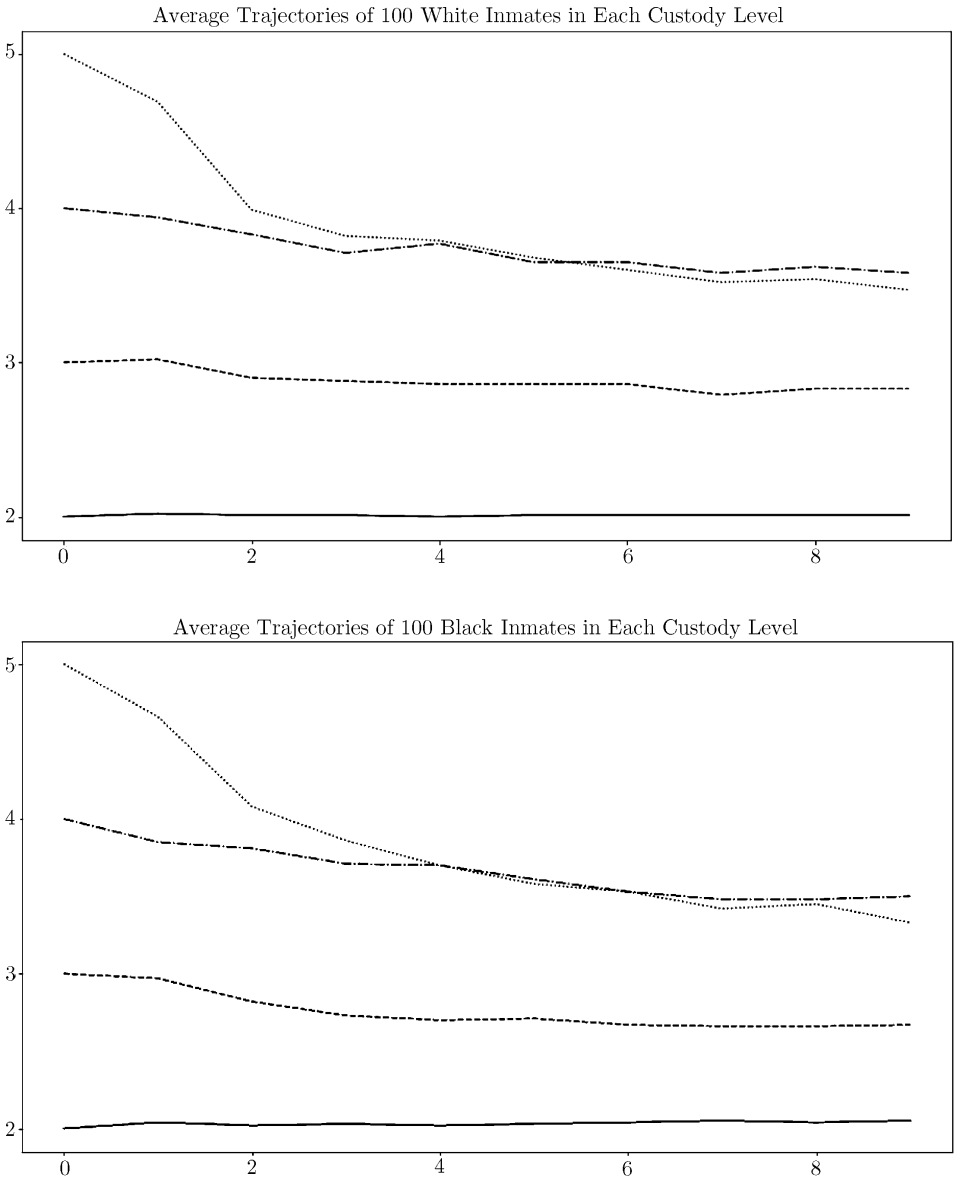


FIGURE 11. The average trajectories for 100 people starting at each custody level and applying the reclassification model 8 times. We separate them by Black and White people for the sake of comparison.

4. Fairness

In this section we discuss the fairness of the data and of the random forest models we have developed. In particular, we ask about the fairness of the outcome, the fairness of the predictions, the fairness of the decision to override an initial classification, conditioned on the protected classes of race, age and gender. We emphasize measures of counterfactual fairness as this is similar in spirit to the simulations we have described above.

4.1. Data-driven fairness of initial classifications and overrides. There are many ways to determine the fairness of algorithms like initial classification via PACT and the more subjective process of determining whether or not a override is deemed to be warranted for a particular person's classification. We start by considering whether or not the two decisions are independent of several protected variables (for us these are age, race, and sex).

We observe in Table 4 the probabilities of being assigned to a high initial custody level does not seem to be independent from whether or not the person is Black, Hispanic, older than 45 years old or female because many of the probabilities are rather different when we toggle between belonging to that protected group and not. For example, the probability of a Black incarcerated person being assigned to a higher custody level is 0.56 while for a non-Black person the probability is 0.28. We also observe in the same table that the institutional adjustment score a person receives is not independent from whether they are Black or whether they are older than 45 years old. Finally, we observe that whether a person gets an override to a higher custody level is not independent from whether they are Black or whether they are female.

TABLE 4. A summary of probabilities of various decisions conditioned on a person belonging to a protected group. The decision $D = 1$ corresponds to a positive answer to the first column and the condition a' is the negation of the condition a.

D	a	$P(D = 1 \mid A = a)$	$P(D = 1 \mid A = a')$
Initial custody level > 3	Black	0.56	0.28
Initial custody level > 3	Hispanic	0.51	0.39
Initial custody level > 3	Age > 45	0.23	0.42
Initial custody level > 3	Female	0.15	0.41
Institutional adjustment > 2	Black	0.54	0.40
Institutional adjustment > 2	Hispanic	0.52	0.47
Institutional adjustment > 2	Age > 45	0.25	0.52
Institutional adjustment > 2	Female	0.45	0.48
Override to a higher custody level	Black	0.54	0.37
Override to a higher custody level	Hispanic	0.42	0.44
Override to a higher custody level	Age > 45	0.51	0.43
Override to a higher custody level	Female	0.27	0.45

TABLE 5. A summary of the probabilities of getting an override to a higher custody ($D = 1$) level conditioned on a person belonging to a protected group ($A = a$) and having a high institutional adjustment score ($B = 1$). The condition a' is the negation of the condition a.

a	$P(D = 1 \mid A = a, B = 1)$	$P(D = 1 \mid A = a', B = 1)$
Black	0.75	0.58
Hispanic	0.69	0.66
Age > 45	0.83	0.65
Female	0.5	0.67

4.2. Fairness of the override process. In addition to the fairness calculations described above, we also want to use the data to determine whether or not an override to a higher custody level was justified by the person having a higher institutional adjustment score. We do this by comparing the probabilities of being given an override to a higher custody level if people have different values for protected variables but the same values for institutional adjustment.

In Table 5, we see that whether a person is given an override to higher custody level is not independent from whether or not a person is Black, whether they are older than 45 years old or whether they are female because the probabilities are rather different when we toggle between belonging to that protected group and not. For example, the probability of a Black incarcerated person with a high institutional adjustment score being assigned to a higher custody level is 0.75 while for a non-Black incarcerated person the probability is 0.58. On the other hand, the decision does appear to be independent from whether or not the person is Hispanic since the probability of a Hispanic incarcerated person with a high institutional adjustment score being assigned to a higher custody level is 0.69 while for a non-Hispanic incarcerated person with a high institutional adjustment score, the probability is 0.68. These probabilities were calculated empirically from the data set and in future work we will explore how to make these calculations more rigorous.

4.3. Algorithmic fairness. In this section we briefly touch on how fair the random forest model is by calculating the statistical conditional parity and the predictive parity of the model. In particular, if $\hat{Y}$ is the random forest prediction for a high or low custody level, we calculated $P(\hat{Y} = 1 \mid A = a)$ and $P(\hat{Y} = 1 \mid A = a')$ and they agreed with the values of $P(D = 1 \mid A = a)$ and $P(D = 1 \mid A = a')$ in the first four columns of Table 4 to two decimal places. So our model performs as fairly as the model run by the PACT. This is perhaps not surprising due to the high accuracy of our model.

4.4. Counterfactual fairness. Intuitively, counterfactual fairness captures the idea that a decision should be fair towards an individual if it is the same in the actual world (the data set) and in a counterfactual world where the individual belongs to a different demographic group. There are several notions of counterfactual fairness including some that, for example, make use of causal models (see, for example, [**8**]). In this paper we consider counterfactual fairness for individuals as described by Sokol, *et al.* in [**17**] and described in more detail below. First, certain variables need to be identified as protected: in our case, we use age (a categorical

variable that identifies a person as being young (< 25 years old), of middle age (between 25 and 45 years old) and as being older (> 45 years old)), race (a categorical variable whose values are "White", "Black" or "Hispanic", since we have subsetted the data accordingly) and sex. Second, we need to identify variables to be used by the classifier. We develop a simpler model and so only include the following variables: 'off_1_prs_max', 'off_1_gs_max', 'prior_commits', and 'ic_instit_adj'.

Next, a classifier and a notion of distance between two points have to be chosen. In this case we train a k-nearest neighbors classifier (classifying whether someone is in a custody level > 3 or not) with the distance being the taxicab distance. In the approach, we start with an individual observation and then conduct a brute force grid search through the space of variables and then classify each point generated in this way. If the grid point and the point that we started with are classified differently and are "close" to each other, the grid point is a counterfactual point.

Table 6 has some examples, one of which we explain here. There was an incarcerated person who was female (1 in the first entry of the initial observation), older than 45, Black, had a prior record score of 0.5, an offense gravity score of 15, 1 prior commitment and an institutional adjustment score of 2. This person was assigned to a low custody level. On the other hand, if this person was between 25 and 45 years old and otherwise the same, they would have been assigned to a high custody level. Also, if this person was Hispanic and otherwise the same, they would have been assigned to a high custody level. In a sample of 500 observations in our data, 107 (21.4%) had counterfactuals with protected variables with values that were close to the observations (taxicab distance ≤ 3) but were classified to the opposite custody level.

TABLE 6. Examples of counterfactuals for particular data points in our data set. The tuple in the first column represents a tuple of variables for the observation for which we are looking for counterfactuals: ('gender_female', 'age_cat', 'race', 'off_1_prs_max', 'off_1_gs_max', 'prior_commits', 'ic_institut_adj'), the second column indicates whether the observation was in a Low or High custody level, the third indicates how the counterfactual differs from the initial observation, the fourth column tells us what custody level the counterfactual is in and the final column tells us the distance between the counterfactual and the initial observation. Sex is coded as 0 for male, and 1 for female; age is coded as 0 for young, 1 for middle age and 2 for older; race is coded as 0 for Black, 1 for Hispanic and 2 for White.

Initial observation	Class	Counterfactual observation	Class	Distance
$(0, 1, 2, 3, 15, 0, 2)$	Low	age_cat: $1 \to 0$	High	1.0
$(0, 1, 1, 1, 12, 1, 2)$	High	race: $1 \to 0$	Low	1.0
		age_cat: $1 \to 0$	Low	1.0
$(1, 2, 2, 0.5, 15, 1, 2)$	Low	age_cat: $2 \to 1$	High	1.0
		race: $2 \to 1$	High	1.0
		gender_female: $1 \to 0$, age_cat: $2 \to 1$	High	2.0
		age_cat: $2 \to 1$, race: $2 \to 1$	High	2.0
		age_cat: $2 \to 1$, race: $2 \to 0$	High	3.0

5. Discussion

Algorithms are playing an ever-increasing role in decision-making, in general, and in criminal justice processes, in particular. The algorithm that we have analyzed, the PACT, is confidential and so the re-creation of the model that we have carried out is necessary to understand it and to critique it. Among other things, our findings underscore problems that arise from the lack of transparency. In an earlier paper [**11**], we note that the PADOC justifies keeping the algorithm confidential because otherwise people would "game" the algorithm. This amounts to an admission that the algorithm is not objective. Moreover, now that we have a fairly accurate model for the PACT, we can determine in which ways it is not objective and what features most influence the custody level assigned by the PACT. A true commitment to fairness and justice would include transparency.

The lack of transparency is seen in our analysis: the most important feature in the random forest model is institutional adjustment. This is a measure of how well a newly incarcerated person is expected to adjust to the particular prison into which they are being committed. There is no published description of how this score is calculated and so a confidential algorithm like the PACT is using opaque variables like institutional adjustment. It is also worth pointing out that other important features (namely, the gravity and prior record scores) are also opaque in the sense that there is no published explanation of how these scores are calculated (here we mean that we know an offense might have a gravity score of 12 but there is no published explanation of why certain crimes have a score of 12 and others have a score of 11 and what a difference of one between two scores means).

In addition to having opaque variables heavily influence the assignment of custody levels, we also find a strong emphasis in the model on variables that are static; that is, on variables that an incarcerated person can do nothing to change while they are incarcerated. For example, we see that the number of prior commitments is important in the classification model that was trained on people some of whom were committed 20 years ago and some of whom were committed a year ago. As another example, the gravity score is based on the initial charge of the offense (this is also documented as having issues with racial bias in charging and sentencing [**7**]), which may have been decades ago for some people.

Demographic or protected variables also lead to very different outcomes and in this sense the PACT has unfair outcomes. For example, a Black person is twice as likely to be given a high custody level than a non-Black person;[1] a younger person is almost twice as likely to be put in a high custody level compared to an older person; and a person identified as male is almost three times more likely to be placed in high custody level than a person identified as female. Similar disparities persist even when we control for a person's institutional adjustment score.

Not only is the PACT unfair in the ways described above, we claim that its unfair in other ways, too. For a fair algorithm, it should be the case that a small change to the input does not change the output very much, especially when those changes are to protected variables.

[1]We point out that the conclusions made from the ratios of the last two columns in Table 4 and 5 are likely more conservative than the reality because the non-Black population also includes the Hispanic population which is also treated unfairly

In Experiment 1, we observed that if we took the existing data set and mixed up the people at each custody level and then reclassified them using the model we trained on the existing data set, almost everybody would change custody levels. This suggests a certain amount of instability to the algorithm. In Experiment 3, we first broke up the data into White people and Black people and then see that White people get sent to all custody levels while Black people are sent exclusively to levels 2 and 3. This suggests that Black people are assigned too high a custody level because if their associated variables were slightly different, they would be assigned to lower custody levels. In Experiments 2 and 4 we randomly choose a single variable and change that to another value in the data set. From Experiment 2, we see that for people in custody level 5 there is a lot of variability of where a slightly different observation could be moved to and, in custody levels 2, 3, and 4, there is a significant number of people for whom a small change would send them to a different custody level. From Experiment 4, the most controlled of the four experiments so far, since we have stratified by custody level and race, and therefore perhaps the most insightful, we see that most people for the most part stay at the same custody level but we see a lot variability for Black people at custody level 5 and a lot of outliers overall. In Experiment 5, we see that bigger changes in the quantitative variables tends to push everyone towards custody levels 3 and 4 and, in Experiment 6, we see that changes in gravity score and institutional adjustment appear to have the greatest impact on custody level.

Additionally, in the counterfactual fairness analysis, we see that there are people for whom there exist synthetic counterfactual data points that are very close to the original person's observation but who are classified differently. The closeness here is just between the person's protected variables and the synthetic observation's protected variables.

While we see some uncertainty in the classification process as described above, we see a lot of uncertainty (in fact, volatility) in the reclassification process. We reiterate that our model for reclassification is not trained on very complete data but, because we have enough people who have been reclassified, the data that was used for their most recent reclassification and the number of years since their initial reclassification, we are fairly confident in its results. What we see is that on an individual level, people are being assigned to different custody levels very often. This might be an example of what geographers have identified as a forced migration [**6**], and will require further study on our part. Prison movement (even within a prison where a custody level determines where the person is housed) is used punitively, emerges in immigration detention centers, and taxes detained people and their families. It contradicts other forms of mobility that make modern, middle-class life as well as the immobility experienced by parolees who must maintain a stable address for long periods of time.

The volatility is also different by race. We see that Black people in custody level 2 change custody level 10 times faster than White people in custody level 2. We see that Black people in custody level 4, change custody level about 50% faster than White inmates. In order to achieve fairness, these numbers should not be as different as they are.

6. Conclusion

In this paper we have approximated a confidential carceral algorithm used by the PADOC to determine an incarcerated person's custody level and, therefore, their lived experience. We identified the most important variables in this model and the fairness of the outcomes both in the data and in the model. We argued for the need to move away from more conventional fairness metrics and towards ways to measuring the uncertainty in these algorithms. By conducting a series of experiments in which the input data was perturbed in various ways and the associated outputs were analyzed, we show that the PACT tool is unfair, is volatile and, ultimately, uncertain. The strength of our conclusions are somewhat tempered by the data that we have. On the one hand, the original data set provided to us by the PADOC has a great deal of missing data and a non trivial amount of wrong data and, on the other hand, our simulations make simplifying assumptions about how the data changes over time. These limitations, though, do not reduce the value of our general approach.

References

[1] Alfred Blumstein and Richard Larson, *Models of a total criminal justice system*, Operations Research **17** (1969), no. 2, 199–232.

[2] Efrén Cruz Cortés and Debashis Ghosh, *A simulation based dynamic evaluation framework for system-wide algorithmic fairness*, Preprint, `arXiv:1903.09209`, 2019.

[3] V. Dabbaghian, P. Jula, P. Borwein, E. Fowler, C. Giles, N. Richardson, A. R. Rutherford, and A. van der Waall, *High-level simulation model of a criminal justice system*, In Theories and Simulations of Complex Social Systems, Springer, 2014, pp. 61–78.

[4] Swarup Dhar and Nathan C. Ryan, *Algorithmic uncertainty of automated decision processes in criminal justice*, 2021, `https://github.com/swarupdhar/algorithmic-uncertainty`.

[5] Danielle Ensign, Sorelle A. Friedler, Scott Neville, Carlos Scheidegger, and Suresh Venkatasubramanian, *Runaway feedback loops in predictive policing*, In Conference on Fairness, Accountability and Transparency, PMLR, 2018, pp. 160–171.

[6] Nick Gill, *Carceral spaces: Mobility and agency in imprisonment and migrant detention*, Routledge, 2016.

[7] Elizabeth Hinton, LeShae Henderson, and Cindy Reed, *An unjust burden: The disparate treatment of Black Americans in the criminal justice system*, Technical report, Vera Institute of Justice, May 2018.

[8] Matt Kusner, Joshua Loftus, Chris Russell, and Ricardo Silva, *Counterfactual fairness*, In Advances in Neural Information Processing Systems (U. Von Luxburg I. Guyon S. Bengio, ed.), Curran Associates, Inc., 2017, `https://proceedings.neurips.cc/paper_files/paper/2017/file/a486cd07e4ac3d270571622f4f316ec5-Paper.pdf`.

[9] Lin Liu and John Eck, *An overview of crime simulation*, Artificial crime analysis systems: Using computer simulations and geographical information systems, IGI Global, 2008, pp. xiv–xxi.

[10] Kristian Lum and William Isaac, *To predict and serve?*, Significance **13** (2016), no. 5, 14–19.

[11] Swarup Dhar, Vanessa Massaro, Darakhshan Mir, and Nathan C. Ryan, *Carceral algorithms and the history of control: an analysis of the Pennsylvania additive classification tool*, Big Data & Society **9** (2022), no. 1, 20539517221094002.

[12] Shira Mitchell, Eric Potash, Solon Barocas, Alexander D'Amour, and Kristian Lum, *Algorithmic fairness: choices, assumptions, and definitions*, Annu. Rev. Stat. Appl. **8** (2021), 141–163, DOI 10.1146/annurev-statistics-042720-125902. MR4243544

[13] Thomas Moranian, Nachum Finger, and Nelson M. Fraiman, *Simulation in criminal justice, a case study of the juvenile court system*, Technical report, Institute of Electrical and Electronics Engineers (IEEE), 1977.

[14] Department of Corrections, *Department of Corrections Procedures Manual: Reception and Classification*, 11.2.1 ed., Harrisburg, PA, 1 2011.

[15] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, and et al., *Scikit-learn: machine learning in Python*, J. Mach. Learn. Res. **12** (2011), 2825–2830. MR2854348
[16] Michael Schwirtz, Michael Winerip, and Robert Gebeloff, *The scourge of racial bias in New York State's prisons*, The New York Times, 2016.
[17] Kacper Sokol, Alexander Hepburn, Rafael Poyiadzi, Matthew Clifford, Raul Santos-Rodriguez, and Peter Flach, *FAT Forensics: A Python toolbox for implementing and deploying fairness, accountability and transparency algorithms in predictive systems*, Journal of Open Source Software **5** (2020), no. 49, 1904.

Department of Mathematics, Washington University in St. Louis, St. Louis, Missouri 63130
Email address: d.swarup@wustl.edu

Department of Geography, Bucknell University, Lewisburg, Pennsylvania 17837
Email address: v.a.massaro@bucknell.edu

Department of Computer Science, Bucknell University, Lewisburg, Pennsylvania 17837
Email address: d.mir@bucknell.edu

Department of Mathematics, Bucknell University, Lewisburg, Pennsylvania 17837
Email address: nathan.ryan@bucknell.edu

Proceedings of Symposia in Applied Mathematics
Volume **80**, 2025
https://doi.org/10.1090/psapm/080/00730

A tutorial on networks of social systems: A mathematical modeling perspective

Heather Z. Brooks

Abstract. This article serves as an introduction to the study of networks of social systems. First, we introduce the reader to key mathematical tools to study social networks, including mathematical representations of networks and essential terminology. We describe several network properties of interest and techniques for measuring these properties. We also discuss some popular generative models of networks and see how the study of these models provides insight into the mechanisms for the emergence of structural patterns. Throughout, we will highlight the patterns that commonly emerge in social networks. The goal is to provide an accessible, broad, and solid foundation for a reader who is new to the field so that they may confidently engage more deeply with the mathematical study of social networks.

Introduction

The study of complex systems is becoming an increasingly important area of inquiry [**BCAB$^+$21**]. With problems in epidemiology, misinformation on social media, voting and decision making, gerrymandering, and social justice appearing front-and-center in recent years, understanding the connections among people and among social entities is perhaps more important than ever.

The study of networks is the study of connectivity. Networks encode the relationships among entities, and through this abstraction, we can study the structure and implications of these relationships. In this article, we will focus our attention specifically on networks that describe social systems. Such networks may describe social interactions and relationships that occur in physical spaces, such as the well-studied Zachary Karate Club network [**Zac77**]. Networks may also be used to describe virtual connections as well, for example, Facebook friendships [**TMP12**]. It is worth noting that we need not limit our study to human social networks, as we observe a variety of interesting networks among other social animals as well (for example, the relationships of penguins at the Kyoto Aquarium [**BF20**], grooming networks [**WBH$^+$18**], social spiders [**FPW21**], and more). At the time of writing, the Colorado Index of Complex Networks [**CTS16**] contains over 2000 network data sets on social systems.

2020 *Mathematics Subject Classification.* Primary 91D30; Secondary 91-10, 05C82.

Key words and phrases. Social networks, mathematical modeling, generative models of networks, network properties.

The author was supported in part by NSF Grant #2109239.

The study of networks is an inherently interdisciplinary field, drawing from physics, biology, computer science, sociology, economics, and beyond. The tools of mathematics and mathematical modeling have an important role to play. In particular, by studying networks through a mathematical lens, we are well-positioned to explore some of the following big-picture questions in social networks:

(1) What network structures are likely to emerge in social systems? How can we interpret these structures?
(2) Why do particular network structures emerge?
(3) How do properties of networks affect the behavior and dynamics of social systems?

In this article, my goal is to introduce the reader to some key mathematical tools to study social networks. In particular, we will focus on two areas: mathematical techniques for measuring network properties and generative models of networks. Along the way, we will highlight patterns that emerge in real social networks and provide you with resources that you can use to guide future study.

1. Mathematical representations of networks

We must first establish a foundation for how to describe and represent social networks mathematically. In this section, we discuss how graphs may be used to represent social networks, introduce some standard matrix representations of graphs, and describe some key terminology and properties. There are many excellent books on networks that describe each of the following properties in more detail. See, for example, [**New18**, **Bul19**, **Jac10**]. These properties will also likely be familiar to those who have encountered graph theory.[1]

1.1. Graphs. We can represent networks with a graph $G = (V, E)$, where V is the set of *vertices* or *nodes* and the number of vertices in the graph is $|V| = n$. $E \subset V \times V$ is the set of *edges* between vertices. Vertices are represented visually with circles, and edges are represented as lines connecting those circles. The interpretation of vertices and edges in a graph to a particular social network setting is an important consideration from a mathematical modeling perspective. Often, vertices are chosen to represent the individuals or entities within a social network, and the edges are chosen to encode interactions, relationships, or other relevant connections between those individuals or entities.

Simple graphs are graphs that contain at most one edge between any distinct pair of vertices and no *self-edges* or *self-loops*, that is, no edges between a vertex and itself (see the example on the left in Figure 1). There are many useful extensions of this framework for the study of social networks. First, in some contexts, it may be helpful to allow self-edges to represent self-interactions. Furthermore, in many social networks, interactions and relationships may not necessarily be reciprocal. For example, one can follow an account on 𝕏 (formerly known as Twitter) without being followed in return. To model this scenario, we often use a *directed graph*, where each edge has a directionality that is represented visually with an arrow. More formally, in a directed graph, the existence of the edge $(1, 2) \in E$ does not

[1]The mathematical study of networks and the study of graph theory are, of course, deeply related. Both fields center around the graph as their primary mathematical object of study, and thus there are many overlaps and shared tools and methods. The difference between these subfields lies primarily in the questions of interest and the motivations behind those questions. A few introductory texts on graph theory include [**Gou12**, **GYA18**, **Wes01**].

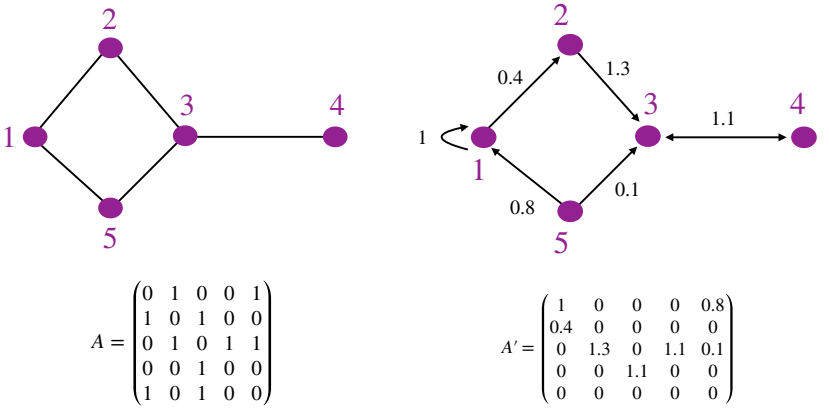


FIGURE 1. Two examples of graphs with their corresponding adjacency matrices. The graph on the left is a simple, undirected graph with 5 nodes and 5 edges. The matrix $\mathbf{A}$ gives its adjacency matrix. The graph on the right is a directed, weighted graph with a self-loop on node 1. This graph has adjacency matrix $\mathbf{A}'$; notice that the edge weights in this graph are encoded within the adjacency matrix.

imply the existence of the edge $(2, 1) \in E$. A graph that is not directed is said to be *undirected*. Figure 1 shows two examples of graphs: one undirected simple graph and one directed, weighted graph with self-loops. It may also be useful to relax the condition that any distinct pair of vertices are connected by at most one edge. To this end, we may choose to model our network with a *multigraph*, where we allow multiple edges (sometimes called *multiedges*) between each pair of nodes. A further generalization of this idea is to allow for each edge in a graph to have a weight, strength, or value (often real-valued); such a graph is known as a *weighted graph*. For example, weighted graphs may be used to encode the distance between two points in a transportation network. In this scenario, one may choose edge weights that are inversely proportional to the geographic distance between the two points.

1.2. Matrix representations of graphs. One reason that graphs are particularly useful mathematical representations of networks is that they can be encoded with matrices. In this section, we will briefly introduce a few of the most commonly used matrix representations of graphs: the adjacency matrix, the incidence matrix, and the graph Laplacian.

1.2.1. *Adjacency matrix.* A graph G with n vertices can be represented by an $n \times n$ matrix $\mathbf{A}$ called the *adjacency matrix*, whose components A_{ij} are defined as follows:

$$A_{ij} = \begin{cases} w_{ij} & \text{if } (j, i) \in E\,, \\ 0 & \text{otherwise}\,, \end{cases}$$

where w_{ij} is the weight of edge (j, i). A simple graph has an adjacency matrix that contains only ones and zeros, with a one in the (i, j)th and (j, i)th component indicating an edge between node i and node j. Thus, if a graph is undirected, its adjacency matrix will be symmetric.

1.2.2. *Incidence matrix.* The adjacency matrix that we discussed in the last section is a matrix representation of a graph that encodes relationships between node pairs. We now briefly introduce an alternate matrix representation called the incidence matrix which encodes relationships between nodes and edges. We say node i is incident to edge j if edge j connects node i to another node in the graph. Let $\mathbf{B}$ denote our incidence matrix. For a simple undirected, unweighted graph, we set $B_{ij} = 1$ if node i is incident to edge j, and 0 otherwise. In this setting, the incidence matrix is related to the adjacency matrix via the relationship $\mathbf{A} = \mathbf{B}\mathbf{B}^T - 2\mathbf{I}_n$, where $\mathbf{I}_n$ is the $n \times n$ identity matrix. As with the adjacency matrix, we may generalize the incidence matrix to represent directed and/or weighted graphs. Given a directed, weighted graph, the incidence matrix $\mathbf{B}$ has entries

$$B_{ij} = \begin{cases} w_{ij} & \text{if edge } j \text{ enters node } i\,, \\ -w_{ij} & \text{if edge } j \text{ leaves node } i\,, \\ 0 & \text{otherwise}\,, \end{cases}$$

where w_{ij} is the weight of edge (j, i). Note that certain authors may flip the sign conventions.

1.2.3. *Graph Laplacian.* The graph Laplacian[2] is defined as

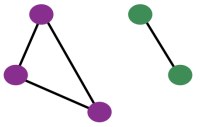
$$\mathbf{L} = \mathbf{D} - \mathbf{A}\,, \tag{1.1}$$

where $\mathbf{A}$ is the adjacency matrix and $\mathbf{D}$ is the diagonal matrix with entries $D_{ii} = \sum_{j=1}^{n} A_{ij}$. See Figure 2 for an example of the graph Laplacian for a graph with 5 nodes. The graph Laplacian arises in many contexts, but it is particularly useful in the study of social networks due to its spectral properties. First, we note that all of the eigenvalues of $\mathbf{L}$ are nonnegative, and the smallest eigenvalue is zero. The algebraic multiplicity of the zero eigenvalue allows us to infer the number of components of a network: a network has c components if and only if $\mathbf{L}$ has c zero eigenvalues (see Section 1.3.3 and Section 2.1 for more information on components). The first nonzero eigenvalue is called the *spectral gap* and can be used as a way to measure how "well-connected" the overall graph is. The associated eigenvector can be used to partition the network. Von Luxburg [**VL07**] provides a tutorial on the graph Laplacian and its applications to clustering problems, which includes proofs of the properties described above.

1.3. Terminology and basic properties. With our mathematical representations of networks in hand, we may now begin to introduce some graph terminology that will help illuminate some basic properties of networks in social systems.

1.3.1. *Degree.* One natural question that arises in the context of social networks is "How many contacts (or interactions, or connections, etc.) does each individual in a network have?" The answer to this question is quantified by the *degree* of a node, that is, is the number of neighbors adjacent to that node. More formally,

[2]There are many variants of the graph Laplacian. The version presented here is a commonly used unnormalized version which is sometimes referred to as the combinatorial graph Laplacian. The way of defining the graph Laplacian—including the choice of normalization—may have impacts on applications, such as consistency in spectral clustering [**VLBB08**].



$$L = \begin{pmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

$$\lambda = \{0,0,2,3,3\}$$

FIGURE 2. An example of a graph and its graph Laplacian L, along with the set of eigenvalues λ of $\mathbf{L}$.

in an undirected network, the degree k_i of node i is easily calculated from the adjacency matrix:

$$k_i = \sum_{j=1}^{n} A_{ij}.$$

In a directed network, this measurement requires somewhat more subtlety. To distinguish between ingoing and outgoing edges of a node, we may define the corresponding notions of *in-degree* and *out-degree*. The in-degree of node i is

$$k_i^{in} = \sum_{j=1}^{n} A_{ij},$$

that is, the number of ingoing edges of node i; the out-degree of node i is

$$k_i^{out} = \sum_{j=1}^{n} A_{ji},$$

the number of outgoing edges of the same node. A familiar example of the use of degree, in-degree, and out-degree can be found in online social media platforms: if we model followerships on $\mathbb{X}$ as a directed network so that the origin and destination of an edge represent the follower and followee, respectively, we can construe the number of accounts one is following as the out-degree and the number of followers of that account as the in-degree. By contrast, if we consider modeling Facebook friendships (which are reciprocal) as an undirected network, we may construe the degree of an account to represent the number of Facebook friends of a particular account.

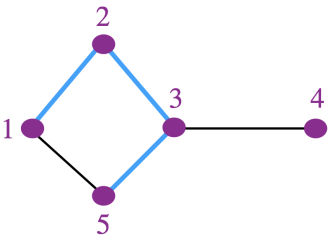


FIGURE 3. An example of a path from node 1 to node 5 is highlighted in blue. This path is of length 3 and consists of the set $\{(1,2),(2,3),(3,5)\}$.

1.3.2. *Paths, walks, and cycles.* Paths and walks can be used to study quantities like the geodesic distance between pairs of nodes; this will be discussed further in Section 2.3. See Figure 3 for an example of a path. The following theorem gives a convenient way to calculate the number of walks between two nodes using the adjacency matrix. This theorem is also useful for finding cycles (as cycles are walks from a node i to itself).

THEOREM 1. *Suppose* $\mathbf{A}$ *is the adjacency matrix corresponding to a graph* G *and* r *is a positive integer. If* $\mathbf{A}^r$ *is the matrix product of* r *copies of* $\mathbf{A}$*, then the number of walks of length* r *between nodes* i *and* j *is the* (i,j)*th component of* $\mathbf{A}^r$*.*

To prove this theorem, proceed inductively on the walk length r. Details of the proof may be found in many introductory texts on networks and graph theory.

1.3.3. *Components and connectivity.* A *component* is a subset of nodes where there exists at least one path between each pair of nodes in the subset such that no other node can be added to the subset and still preserve this property. A network in which all nodes belong to the same single component is *connected.* Some examples are given in Figure 4.

As discussed in Section 1.2.3, the number of components in a graph corresponds to the number of zero eigenvalues in its graph Laplacian. This provides one straightforward method for finding the number of components in a network.

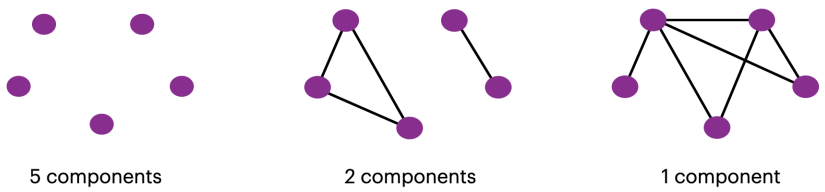


FIGURE 4. A graph with 5 nodes and 5 components (left), a graph with 5 nodes and 2 components (middle), and a graph with 5 nodes and one component (right). The graph on the right is connected.

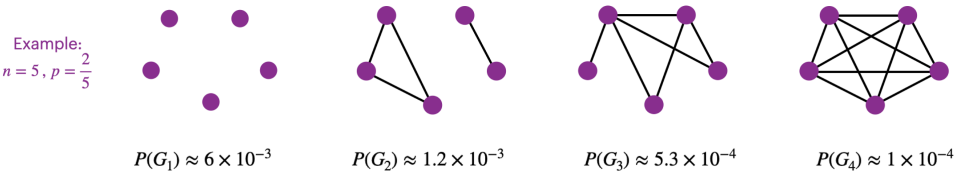


FIGURE 5. Calculating the probabilities $P(G_i)$ of a particular graph realization for a $G(n,p)$ model with $n = 5$ nodes and an independent edge probability of $p = \frac{2}{5}$.

2. Properties of social networks

2.1. Social networks often contain a large component. In many social networks (and, in fact, in undirected networks in general), it is common for one component of the network to contain a large percentage of the nodes—in many cases, upwards of 90% of the nodes [**New18**]. Why might social networks contain a large component? To understand which mechanisms might lead to this property, it is useful to study the emergence of large components in simple generative models of random graphs.

2.1.1. *The $G(n,p)$ model.* Perhaps the simplest random graph model is the $G(n,p)$ *model* (sometimes also known as Erdős–Rényi model). A realization of this model is created by fixing two parameters: n, the number of nodes, and p, the probability of an edge connecting any pair of nodes. Given n and p, we may then generate a network by placing an edge between each distinct pair of nodes with independent probability p. Note that the number of edges m is not fixed—using this strategy, we may generate a network with anywhere between $m = 0$ and $m = \binom{n}{2}$ edges.

Mathematically, it is useful to think of a $G(n,p)$ model as an ensemble of simple networks with n nodes, i.e., a probability distribution over possible graphs in which a particular graph G with m edges appears with probability

$$P(G) = p^m (1-p)^{\binom{n}{2}-m}.$$

This formula holds under the assumption that we can uniquely identify each node and each node pair distinctly and consider the presence or absence of each edge as an independent probabilistic event. See Figure 5 for an example on a small network. Under a different set of assumptions (for example, the observed event is a graph with n unidentifiable nodes, and/or G is a graph with m edges, regardless of location), then the combinatorial calculation must be adjusted accordingly, including accounting for graph isomorphism.

At this point, it is worth mentioning a related (but distinct) class of random graph models known as $G(n,m)$ model.[3] In these models, we create a particular network realization by choosing uniformly at random among the set of all simple graphs with n nodes and m edges. We will not explore $G(n,m)$ models further here; we will refer the interested reader to a networks textbook such as [**New18**] for details on the properties of these generative network models.

Let us return to the $G(n,p)$ model to develop some intuition behind the observed large components in social networks. Do networks generated with the $G(n,p)$

[3]Sometimes the $G(n,m)$ model is also referred to as an Erdős–Rényi model.

model contain a large component? If we ponder this for a few moments, we will realize that the answer to this question must be "it depends." For example, if we consider $n = 1000$ nodes and $p = 0$, then the largest component of any network generated from this model will contain only one node (or 0.1% of the nodes)—certainly not an impressively-sized component! On the other hand, if $n = 100$ and $p = 1$, we will generate a network whose largest component contains all nodes with probability 1. Is there a transition between these two regimes?

To answer this question, rather than looking at the absolute size of the largest component, we will reframe the question as follows: for a given value of p, is there a component whose size grows in proportion to n? (If such a component exists, it is referred to as a *giant component*). Following an argument from Erdős and Rényi, we will show that for the $G(n, p)$ model there is a critical value of p beyond which we expect the emergence of a giant component.

Suppose we let u represent the expected fraction of nodes that do not belong to this giant component. We know that, if node i is not in the giant component, then it must not be connected to any node in the giant component, i.e., every other node $j \neq i$ is either not connected to node i (with probability $1 - p$), or it is connected to i and also not in the giant component (with probability pu). Combining these two observations, we can note that the probability of a node not being connected to the giant component via a particular node j is $1 - p + pu$. Thus, the probability u of a node not being connected to the giant component by any of the $n - 1$ other nodes is

$$u = (1 - p + pu)^{n-1}.$$

We can rewrite the expression above by noting that the probability that any two nodes are adjacent is $p = \frac{\langle k \rangle}{n-1}$, so

$$u = \left(1 - \frac{\langle k \rangle}{n-1}(1-u)\right)^{n-1}, \tag{2.1}$$

$$\Rightarrow \ln u = (n-1)\ln\left(1 - \frac{\langle k \rangle}{n-1}(1-u)\right). \tag{2.2}$$

Rewriting the expression in this way allows us to derive an approximate expression for u by performing a Taylor expansion and neglecting higher-order terms. Doing so yields

$$\ln u \approx -(n-1)\frac{\langle k \rangle}{n-1}(1-u), \tag{2.3}$$

$$\ln u \approx -\langle k \rangle (1-u), \tag{2.4}$$

$$\Rightarrow u \approx e^{-\langle k \rangle (1-u)}. \tag{2.5}$$

Denoting the average fraction of nodes in the giant component by $S = 1 - u$, we obtain the following equation:

$$S = 1 - e^{-\langle k \rangle S}. \tag{2.6}$$

While equation 2.6 does not admit a simple closed form solution, we can use graphical methods to find solutions by looking for intersections of $y = S$ and $y = 1 - e^{-\langle k \rangle S}$.

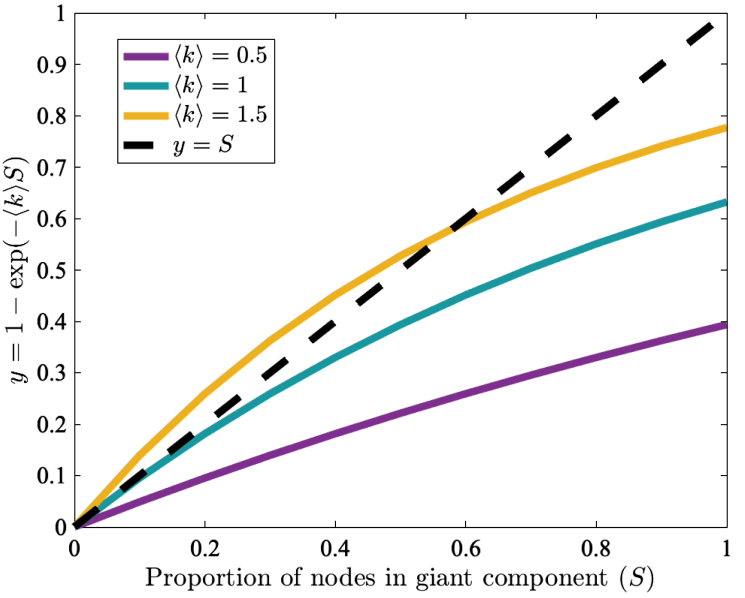


FIGURE 6. The expected proportion of nodes in a giant component (S) for a $G(n,p)$ random graph model is represented graphically with the intersection of $y = S$ (dashed black line) and $y = 1 - e^{-\langle k\rangle S}$. When the mean degree $\langle k\rangle$ is small, the only intersection is at the origin and we do not expect the emergence of a giant component (purple curve). As the mean degree $\langle k\rangle$ becomes large, an additional intersection point appears, signaling the possibility of the emergence of a giant component (yellow curve). The transition point between these regimes occurs at $\langle k\rangle = 1$ (teal curve).

First, we see by inspection that $S = 0$ will always be a solution, that is, it is always possible that a particular network has no giant component. We instead turn our attention to determining when a giant component is possible. When the mean degree $\langle k\rangle$ is small, $S = 0$ is the only solution to equation 2.6, however, for large $\langle k\rangle$, there is an additional intersection point where $S > 0$. For what value of $\langle k\rangle$ does this transition occur? Figure 6 gives us the necessary intuition to answer this question: to obtain two intersection points, it must be the case that $1 - e^{-\langle k\rangle S}$ is growing faster than S at $S = 0$. In particular, the transition from one intersection to two occurs precisely when these two expressions have equal growth rate at $S = 0$. That is,

$$\frac{d}{ds}\left(1 - e^{-\langle k\rangle S}\right)\Big|_{S=0} = 1\,, \tag{2.7}$$

$$\Rightarrow \langle k\rangle e^{-\langle k\rangle S}\Big|_{S=0} = 1\,, \tag{2.8}$$

$$\Rightarrow \langle k\rangle = 1. \tag{2.9}$$

This means that there can be no giant component for $\langle k \rangle < 1$. This calculation confirms our intuition: if the mean degree of a network is less than one, there are (relatively) few edges and it is quite easy for nodes to be disconnected from one another; we might expect that we have many small connected components.

It remains to show that the $S > 0$ solution when $\langle k \rangle > 1$ is observed in practice. Using straightforward probabilistic arguments, one can show that as n grows large, we do in fact expect to see a giant component emerge, and $S \to 1$ as $\langle k \rangle$ grows (that is, the number of nodes in this component approaches the number of nodes in the network).

2.2. Degree distributions of social networks are often heavy-tailed.

2.2.1. *Degree-based centrality measures and degree distributions.* *Centrality measures* quantify the relative importance of nodes in a network. We have already seen one centrality measure in Section 1.3.1: if we suppose that nodes with more connections are more important or influential in our network, then indeed the degree of a node can be used as a centrality measure. Social media networks like 𝕏 provide a good example: we may suppose that an account is more influential than another if it has more followers. It is important to notice that centrality is a relative measure; a node's centrality score only has meaning in comparison with another node.

If k_i is the degree of node i, then the list $\{k_1, k_2, \ldots, k_n\}$ is called the *degree sequence* of a network; if we order this list from largest to smallest then we can see that the node(s) j with the largest k_j has the highest centrality, and we can continue the comparison. This method of calculating a node's centrality has a clear advantage in that it is easy to interpret and calculate: if we let $\mathbf{c}_{deg}$ be the vector containing the centralities of each node, then $\mathbf{c}_{deg} = \mathbf{A}\mathbf{1}$, where $\mathbf{1}$ is the vector of ones.

One might argue that this simple centrality measure misses a key feature of relative importance: perhaps a node's centrality should be based not only on the number of connections it has, but whether or not it is connected to other important nodes (colloquially, "it's who you know"). To incorporate this idea, supposing that the centrality of node i is proportional to the sum of the centralities of its neighbors yields

$$c_i = \frac{1}{r} \sum_{j=1}^{n} \mathbf{A}_{ij} c_j \,, \tag{2.10}$$

where r is a proportionality constant. A quick manipulation of this equation leads us to see that $\mathbf{c}$ is an eigenvector of $\mathbf{A}$ with eigenvalue r. For this reason, this centrality measure is called *eigenvector centrality*. Which eigenvector should we use? Provided that the graph encoded by this adjacency matrix is connected, then an application of the Perron–Frobenius theorem tells us that $\mathbf{A}$ has a unique largest eigenvalue r, and its associated eigenvector contains only positive components.

There are some complications with applying eigenvector centrality for directed graphs or graphs with multiple connected components. The latter could be addressed by calculating centralities for each component separately (thus guaranteeing that the conditions of the Perron–Frobenius theorem are still satisfied). The former requires some more subtle modeling choices: should this centrality measure use in-edges or out-edges to calculate centrality? This will depend on the context of one's problem. It is in the context of directed graphs that we encounter a downside

of eigenvector centrality: it is not difficult to construct networks where every node has zero eigenvector centrality, because (for example) a node with in-degree zero "propagates" its zero centrality throughout the network. Needless to say, eigenvector centrality does not provide a very useful measure in such a situation.

A fix for this is provided by *Katz centrality* [**Kat53**], where we modify the eigenvector centrality by adding a baseline amount $\beta > 0$ to the centrality of each node, that is,

$$\mathbf{c} = \alpha \mathbf{A}\mathbf{c} + \beta \mathbf{1}\,.$$

We are able to solve for $\mathbf{c}$ provided that $\mathbf{I} - \alpha\mathbf{A}$ is invertible, that is, provided that we choose $0 < \alpha < \frac{1}{r}$, where again r is the largest eigenvalue of the adjacency matrix $\mathbf{A}$. For additional details on Katz centrality, see [**BV14**].

One practical consideration when choosing to implement Katz centrality is that if a node with high Katz centrality is connected to many other nodes, all of those adjacent nodes will end up with a high centrality score as well, which may be undesirable in certain contexts. One possible fix for this would be to "dilute" the centrality based on the number of out-edges a node has, that is, the centrality contributed by node i is divided evenly among all of its adjacent nodes j. This modification results in the centrality equation

$$\mathbf{c} = \alpha \mathbf{A}\mathbf{D}^{-1}\mathbf{c} + \beta \mathbf{1}\,, \tag{2.11}$$

where D is the diagonal matrix with elements $D_{ii} = \max(k_i^{out}, 1)$ and β is a constant determined by the choice of α. This is called *PageRank centrality* due to its connection to Google's search algorithms. Again, provided that α is chosen so that $I - \alpha\mathbf{A}\mathbf{D}^{-1}$ is invertible, this will yield a unique centrality ranking for a network's nodes. For PageRank, the choice There is another appealing interpretation of PageRank: Noting that $\mathbf{A}\mathbf{D}^{-1}$ is column stochastic, we can view this matrix as the transition matrix of a random walker traversing edges in our network. The PageRank centrality $\mathbf{c}$ is the stationary distribution for this Markov Process. For more details and applications of PageRank, see [**BV14**, **BL06**, **Gle15**].

We conclude this section by emphasizing that there is no "best" centrality measure: the choice of centrality measure is itself a modeling decision, and should be made with careful consideration of its relevance to the quantities that are desirable to measure for a particular application [**LFH10**]. Indeed, there are many centrality measures based on quantities other than degree: for example, in Section 2.3.3, we will discuss another class of centrality measures based on paths. Before doing so, we will explore degree distributions in social networks. Similar observations may be extended to other degree-based centrality measures.

2.2.2. *Degree distributions and the $G(n,p)$ model.* First, we define the *degree distribution* of a network to be the function $p : \mathbb{N} \to \mathbb{R}$, where $p(k)$ is the number of nodes with degree k divided by the total number of nodes n. In order to understand what we might expect from a degree distribution, we turn again to the simple $G(n,p)$ random graph model introduced in Section 2.1.1. Since these graphs are constructed under the assumption that the probability of any pair of nodes being connected is p, then the probability of a node being connected to k other nodes (out of the possible $n-1$ other nodes in the network) is binomially distributed with probability p: that is, the degree distribution satisfies $p(k) = \binom{n-1}{k} p^k (1-p)^{n-1-k}$. Furthermore, if $n >> 1$ and $p << 1$, the degree distribution is approximately Poisson: $p_k \approx e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$.

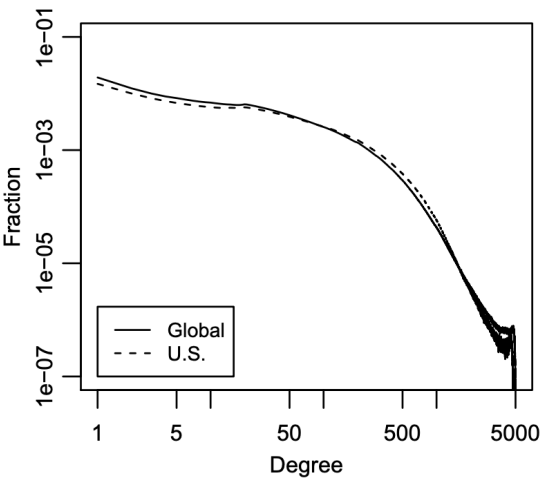


FIGURE 7. From [**UKBM11**]: The degree distribution of Facebook users in 2011. Both globally and within the U.S., this degree distribution has a characteristic "heavy tail": Many users have relatively few friends, while a small proportion of users have a very large number of friends. Note that this distribution has a steep drop-off at 5000 due to an imposed limit on Facebook friends at the time of the study.

2.2.3. *Degree distributions in social networks.* What are the degree distributions observed in real social networks? We will briefly present three case studies (representing a much broader trend) that suggest that degree distributions are often heavy-tailed. That is, in contrast to the degree distributions observed in the $G(n,p)$ model, it is common in real networks that most nodes have a relatively small degree, while a small number of nodes have a very large degree.

In Ugander et al. [**UKBM11**], the authors study the structure of the network formed by users of Facebook, a popular social media platform. In this context, nodes represent users, with an edge occurring between any pair of nodes who are designated as friends on Facebook. This is a notable study in that it contains an impressive $n = 721$ million nodes and 68.7 billion friendship edges. It is interesting to note that the authors report that 99.91% of individuals belong to a single large connected component. The median number of Facebook friends for a user (i.e., the median degree) in this data set is 99. Looking at the degree distribution of this network gives a more detailed picture (Figure 7): most individuals have a relatively small number of friends, while a small subset of users have thousands of friends. This is a characteristic example of a heavy-tailed degree distribution in a social network.

We see heavy-tailed degree distributions continuing to show up in online social media networks today—and not only in the scenario where edges are construed to represent friendships or followerships. To consider a more modern example, in [**TECP20**], the authors study the 'retweet network' of a dataset of messages ('tweets') from the social media network Twitter (now $\mathbb{X}$) with the hashtag #Charlottesville. They create a retweet network by representing accounts as nodes and

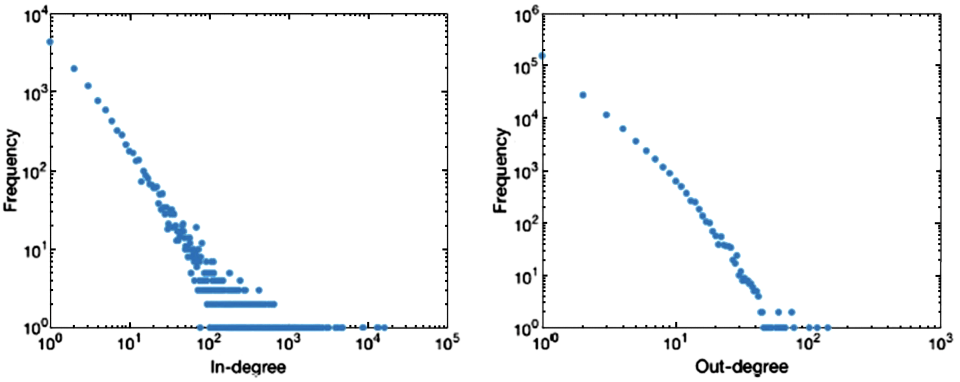


FIGURE 8. From [**TECP20**]:[4]The in-degree distribution (left) and out-degree distribution (right) of a retweet network of Twitter (now 𝕏) messages with the hashtag #Charlottesville. Both degree distributions have a characteristic "heavy tail": that is, a small number of accounts are retweeted much more often than most tweets in the dataset, and similarly a small number of accounts are responsible for a large proportion of retweets. While both distributions share this qualitative feature, we note that the tail of the in-degree distribution is much longer.

using weighted, directed edges to represent the number of times account j retweeted account i (that is, account j shared a message originally posted by account i). Note that accounts do not need to have any followership relationship to be able to retweet. As this network is directed, the authors examine the degree distributions for both in-degree (number of times an account was retweeted) and out-degree (number of times a node posted a retweet). Both distributions again have the characteristic heavy tail seen in other social networks (Figure 8).

Heavy-tailed degree distributions are prevalent in social networks outside of online social media platforms as well. A common object of study is the 'co-authorship' or 'collaboration' network in particular fields, where the nodes represent scientists or authors, and the edges represent whether two individuals have co-authored an article [**AB02**]. In [**New01b**, **New01a**], Newman studies co-authorship networks in physics, biomedical research, and computer science, finding a heavy-tailed degree distribution in each case. In [**BJN$^+$02**], the authors study co-authorship networks of mathematicians and neuroscientists and observe similar results.

As we noted in Section 2.2.2, the $G(n, p)$ random graph model does not produce a network with a heavy-tailed distribution. How might we create a generative model that does reproduce this common feature of social networks? In 1967, Price introduced a so-called cumulative advantage model [**Pri67**, **Pri76**] for citations of journal articles, where new citations are accrued at a rate which is proportional

[4]From "Online reactions to the 2017 'Unite the right' rally in Charlottesville: measuring polarization in Twitter networks using media followership" by Joseph H. Tien, Marisa C. Eisenberg, Sarah T. Cherng, and Mason A. Porter. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (`https://creativecommons.org/licenses/by/4.0/`). Available at `https://link.springer.com/article/10.1007/s41109-019-0223-3/figures/1`.

to the current number of citations. Under these conditions, the distribution of the number of citations in the model is consistent with observations in real-world citations, namely, it follows an *inverse power law* or *Zipf law* (that is, probability of a node having k citations satisfies $P(k) \sim k^{-\alpha}$ for $\alpha > 0$. In this context, Price reported α in the range of 2.5–3). In 1999, Barabási and Albert adapted this idea to the context of undirected networks to create their well-known model of *preferential attachment* [**BA99**]. This generative model can be constructed as follows. Starting with a small number of nodes m_0, add a new node at each time step with $m \leq m_0$ edge stubs. Each of those m edge stubs should then be attached to an existing node i in the network with probability p_{k_i}, where k_i is the degree of node i. If p_{k_i} is proportional to k_i, then we are in the context of preferential attachment: nodes with higher degree are more likely to gain new edges (sometimes colloquially referred to as a 'rich get richer' scenario). After this process is repeated for many steps and a graph is grown via this preferential attachment mechanism, Barabási and Albert showed that the resulting degree distributions again satisfy a power law, a common variant of a heavy-tailed distribution. Networks whose degree distributions satisfy power laws are sometimes called *scale-free*. There has been vigorous debate in recent years over whether the real social networks with heavy-tailed degree distributions are in fact scale-free. While we will not give further details here, curious readers are encouraged to see [**Hol19**] and references therein.

When seeking generative network models to create random networks that satisfy particular degree distributions, we need not limit ourselves to the preferential attachment or $G(n, p)$ models we have thus far described. *Configuration models* are a family of models of random graphs where the degree sequence or degree distribution is fixed (while these models are widespread in the networks literature today, see [**Bol80**] for an early theoretical study). One way to interpret a configuration model is that it's an ensemble of edge-stub matchings, where each matching with a given degree sequence occurs with equal probability (and any other matching outside of the given degree sequence occurs with probability zero). Alternatively, one can think about these models by fixing the degree distribution $p(k)$, and then a particular degree sequence $\{k_i\}$ occurs with probability $\prod_i p(k_i)$. Let us explore this family of models with a concrete example: the ensemble of random graphs with $\{k_i\}$ drawn from a Poisson distribution with mean λ nearly recovers a $G(n, p)$ model for large n. However, these are not identical, as it is important to note that a configuration model that is generated through edge-stub matching may contain self-edges and multi-edges which do not occur in the standard $G(n, p)$ model.

2.3. The "small world" effect: The diameter and/or mean geodesic distance of networks is often (surprisingly) small. In Section 1.3.2, we introduced the notion of paths in graphs. In this section, we will explore the notable properties related to paths in social networks. One natural modeling question to ask is how closely connected any two individuals are within a network. While there are possibly many paths between two nodes in the same connected component, perhaps the most natural way to measure how closely connected two nodes are is to identify the *shortest path* (also known as the *geodesic distance*) between them.

When seeking to understand the topology of a full network, the notion of shortest paths gives some potentially insightful measures. One popular choice is to consider the mean shortest path length (sometimes called the *characteristic path length*), which gives an idea of how many steps it takes on average to connect two

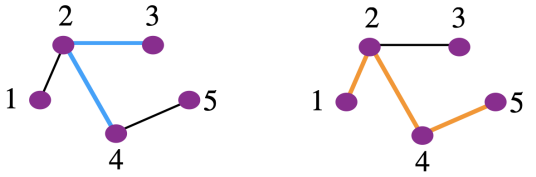


FIGURE 9. An schematic of shortest path and diameter in a network. The length of the shortest path between nodes 3 and 4 is $d_{34} = 2$ (highlighted in blue, left). The diameter of the network is the length of the longest shortest path in the network: this network has a diameter of 3, as $d_{15} = 3$ (highlighted in orange, right) and $d_{35} = 3$.

nodes in a network. If we are instead more interested in the characterizing the extreme behavior (or 'worst case scenarios'), we can use the *diameter* of the network, which is the length of the longest shortest path in our network. Schematics of shortest path and diameter in a small network are shown in Figure 9.

The concept of shortest path or diameter may already be familiar to you. In mathematics, we have a playful quantity called the Erdős number: a scholar who co-authored a paper with Erdős has an Erdős number of 1, a scholar who co-authored a paper with somebody who co-authored a paper with Erdős has an Erdős number of 2, and so on. In fact, we can see that the Erdős number is simply the shortest path length in a co-authorship network between Erdős and another author in the same connected component. Of course, we do not need to limit our collaboration distance to Erdős. MathSciNet has a tool to compute the "collaboration distance" (i.e., the shortest path in the co-authorship network) between any two authors in their database [**AMS**]. Other (less math-centric) versions of this game include Six Degrees of Kevin Bacon (where the goal is find the shortest path between the actor Kevin Bacon to other actors in the 'co-star' network, where edges occur between actors who have appeared in the same film) [**CC98**] and Six Degrees of Wikipedia (where the goal is to find the shortest path between two Wikipedia pages through links) [**Wik**].

Each of the games mentioned in the previous paragraph are amusing examples of the *small world phenomenon*: the diameter (and/or mean geodesic distance) of social networks is often surprisingly small. This concept was popularized in Milgram's message-passing experiment, where participants were asked to try to deliver a letter to an arbitrarily selected person by sharing it with one of their personal acquaintances [**Mil67**]. In revisiting our more modern example of the study of Facebook networks by Ugander et al. [**UKBM11**], we see this play out in this online social network as well. The authors found that 99.6% of users within their largest connected component had a 'hop distance' (i.e., shortest path length) of six or fewer.

A quick back-of-the-envelope calculation can give some intuition as to why the small world phenomenon might occur in the 2011 Facebook friendship networks. Given that the median number of Facebook friends (median degree) is 99, we could estimate that the number of Facebook accounts within path length 6 of our own

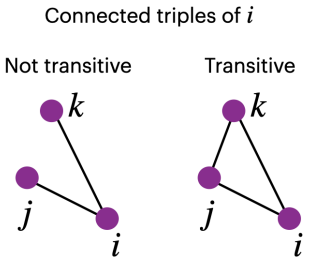


FIGURE 10. A schematic of connected triples of a node i. The example on the left is a connected triple, but is not a transitive triple because there is no edge between nodes j and k. The example on the right is a transitive triple (or triangle), where edges exist between each node pairing of i, j, and k. We can calculate the transitivity by enumerating the transitive triples and the total number of connected triples and taking the ratio of these quantities.

is approximately $99 \times 98 \times 98 \times 98 \times 98 \times 98 = 8.9 \times 10^{11}$, which exceeds the human population at that time (and even more comfortably exceeds the number of Facebook accounts).

While this calculation provides some helpful intuition, we must be quick to point out that we have neglected something very important: we have assumed that the set of our friends and the set of our friend's friends is distinct—surely we must have double-counted some individuals! We explore this feature further in the next section.

2.3.1. *Clustering and the clustering coefficient.* In our quick calculations from the last section, we neglected to account for whether or not two nodes that are adjacent to a node i are also adjacent to each other. The extent to which this is true in a network is called *clustering* or *transitivity*. Colloquially, we might say in social networks that in a network with high clustering, "the friend of my friend is also my friend." One way to measure local clustering for an individual node in a network is by defining the local clustering coefficient C_i, which is the ratio of the number of edges between neighbors of i to the number of possible edges between neighbors of i. The *clustering coefficient* of a network can then be written $C = \langle C_i \rangle$, that is, the mean of the local clustering of all nodes in the network. Another alternative involves calculating the number of "triangles" or "transitive triples" in the graph, i.e., a trio of nodes that are all adjacent to each other (a complete subgraph with 3 nodes). The transitivity for the network can then be defined by taking the ratio of the number of triangles in the graph to the total number of connected triples in the graph and multiplying by 3 (to account for the fact that each triangle gets counted three times as a connected triple). See Figure 10 for a schematic of these quantities.

Real social networks can often exhibit high clustering. In our previous example with collaboration networks [**New01b**], it was shown that two authors have at least a 30% probability of collaborating with each other if they have both collaborated with a third author, and the values of clustering coefficient are higher than expected

even when accounting for papers with 3 or more authors. [**UKBM11**] find a relatively large clustering coefficient in their graph of Facebook friendships (0.14, i.e., on average 14% of all the friend pairs of a median user are friends with each other). The authors of a large study [**MSGL14**] of the Twitter (now 𝕏) followership graph containing 175 million users found a similar result to the Facebook study: nodes in the mutual graph (where an edge exists if two accounts follow each other) with degree 100 also have a clustering coefficient of approximately 0.14.

2.3.2. *Watts–Strogatz model.* As we have explored in the previous two sections, real social networks are often characterized by both small characteristic path lengths and high clustering. Watts and Strogatz set out to create a generative random graph model that could capture both of these features [**WS98**]. Their idea is as follows: starting with a ring lattice of n nodes with k neighbors each, rewire each edge to a different end node with probability p (disallowing duplicate rewirings). This process effectively creates "shortcuts" to decrease path length, while potentially retaining some of the clustering features from the original ring lattice. The parameter p in this model can be tuned to balance these properties: when $p = 0$, we produce a ring lattice (with high clustering, but relatively long path lengths), and when $p = 1$, we recover a $G(n, m)$ model (with low clustering but a small mean shortest path). Watts and Strogatz show that as p is increased from 0, the effect of the shortcuts ensures that mean shortest path decreases rapidly while the mean clustering coefficient remains high until the rewiring probabilities become larger. In particular, for values of p on the order of 0.01, the resulting graphs tend to have relatively low mean shortest path length (roughly 20% of the mean shortest path of the original ring lattice), but with clustering coefficient nearly as high as the original ring lattice. Their 'small world' generative graph model, today often known as the Watts–Strogatz model, remains popular and well-studied.

2.3.3. *Path-based centrality measures.* We conclude this section on path-based properties of real social networks by returning to our earlier discussion about centrality (Section 2.2.1). Given what we now know about the features of shortest paths in social networks, we may be inspired to quantify the relative importance of a node in our network not by its number of connections, but by its path-related features.

One possibility is to consider a node to have high centrality if it is a short 'distance' from many other nodes, i.e., its mean shortest path length to other nodes is small. This is called *closeness centrality*, and one way to write this is to suppose that the centrality is inversely proportional to the mean shortest path to other nodes: If d_{ij} is the shortest path from node i to node j, then the closeness of node i is $c_i = \frac{n}{\sum_{j \neq i} d_{ij}}$. We need to take careful consideration when defining closeness in directed networks or networks with multiple connected components, however. This problem can be solved by considering closeness in each strongly connected components separately (although the scores will not be comparable between components, as the size of the component affects these values). Another option is to redefine closeness in terms of the harmonic mean, so that $c_i = \frac{1}{n-1} \sum_{j \neq i} \frac{1}{d_{ij}}$, allowing that terms where there is no path between two nodes will contribute 0 to a node's centrality (in some sense, this is like supposing d_{ij} is infinite if there is no path between i and j).

Another possibility is to suppose a node is instead important if it lies on many shortest paths between other nodes. For example, such nodes may be important because information or goods would pass through these nodes frequently, and their removal could disrupt paths. There are multiple ways to define such a centrality, which is known as *betweenness centrality*. One straightforward way to calculate this is to sum the number of all shortest paths of a pair of nodes that contain node i, and then normalize by the number of shortest paths between that pair. Summing over all possible pairs in the networks yields a betweenness centrality score.

The centrality scores mentioned in this article are among the most common, but many possible variants abound in the literature. Indeed, you may find that your problem of interest requires the invention of yet a new strategy to define centrality. In selecting a centrality score, it is important to consider the benefits and pitfalls of a given method. Boldi and Vigna [**BV14**] provide a thorough discussion on centrality scores and suggest axioms for centrality measures.

2.4. Assortativity, clustering, and community structure in social networks.

2.4.1. *Assortativity and homophily.* The tendency of people to associate with others whom they perceive to be like themselves is called *homophily*. Supposing that we have a network where nodes can be categorized by different classes, types, or groups, we may like to develop a way to quantify the prevalence of homophily in our network. Informally speaking, a network is said to be *assortative* if a significant fraction of edges are between nodes of the same 'type' and *disassortative* if a significant fraction of edges are between nodes of different 'types.' For example, in a network of romantic relationships of high school students [**BMS04**], the study's authors found this network to be disassortative by gender.

One way to quantify the assortativity of a network is to measure the *modularity* of the network. Suppose that we have a graph $\mathbf{G}$ with a set of node classes or types $\mathcal{C}$. Let $g_i \in \mathcal{C}$ be the type of node i, and m be the number of edges in the network. Then the modularity Q is defined to be

$$Q = \frac{1}{2m} \sum_{i,j} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta_{g_i g_j} , \tag{2.12}$$

where k_i is the degree of node i and $\delta_{(\cdot,\cdot)}$ is the Kronecker delta. Let us quickly deconstruct the meaning of this expression. Notice that the first term will give us the fraction of same-type edges in our network. In the second term, we notice the quantity $\frac{k_i k_j}{2m}$, which will give approximately the expected number of edges between nodes i and j. Thus, the second term gives the expected fraction of edges between all node pairs. In this way, equation 2.12 gives a comparison between the observed fraction of edges between same-type nodes and the expected fraction of edges between same-type nodes, given that they were connected randomly (i.e., via a configuration model as described in Section 2.2.3). If $Q > 0$, then we observe more edges between same-type nodes than would be expected by chance; this corresponds to assortative mixing. If $Q < 0$, then conversely the network is disassortative.

There is an alternative definition of modularity that can provide easier calculation if we don't have information about the degree of individual nodes. First, we can define the fraction of edges that join nodes of type g:

$$e_g = \frac{1}{2m} \sum_{i,j} A_{i,j} \delta_{g_i,g} \delta_{g_j,g} \, .$$

The fraction of "ends" of edges attached to nodes (that is, node stubs) of type g is

$$a_g = \frac{1}{2m} \sum_{i} k_i \delta_{g_i,g} \, .$$

Via some algebraic manipulation, we can rewrite Q from equation 2.12 as

$$Q = \sum_{g} \left(e_g - a_g^2 \right) . \tag{2.13}$$

Depending on the node types or classes under consideration, real-world social networks may be either assortative, disassortative, or neither. However, there is one important feature that is shared across many social networks: they tend to be assortative by degree (i.e., "degree assortative") [**New02**]. This means that in these networks, high-degree nodes are more frequently attached to other high-degree nodes than one would expect were edge stubs connected randomly. In [**New02**], Newman examines a variety of real-world networks and observes degree assortativity in a variety of social networks, including co-authorship networks, film actor collaborations, and connections between business people. This feature is not true of all real-world networks: nonsocial networks such a protein interactions in yeast, synaptic connections in *C. elegans*, and food webs in various aquatic environments are disassortative by degree.

2.4.2. *Community structure.* In Section 2.3.1 and Section 2.4, we have separately considered evidence suggesting that social networks exhibit high clustering and tend to be assortative by degree. In fact, these features of real social networks may arise from the same source: community structure. A graph with *community structure* consists of subgraphs (called "communities") which are more densely connected within those subgraphs than between them (see Figure 11). This somewhat simplistic definition can be made more precise in various ways. For example, one possibility is to define a community via the probabilities of edges existing between nodes within the community versus outside of it; this leads to the notions of *strong community* (each node within a subgraph has a higher probability of connection to other nodes within the subgraph than to any outside of it) and *weak community* (the average probability of connection between nodes within a subgraph is higher than the average probability of connection to nodes in different subgraphs) [**FH16**].

One of the most celebrated examples of the existence of community structure in social networks is the Zachary Karate Club network [**Zac77**]. In his 1977 work, Zachary observed the social interactions of members of a karate club over the course of three years. During this time, members of this club became divided into two factions over a club-related issue. Zachary's insight was that, by characterizing the social interactions of the club members prior to this fission as a network, each faction corresponded to a more densely-connected subgroup of the network—thus, these two social 'communities' (in our modern networks terminology) were in some sense predictive of which side individuals would take on the issue. This network has today become a common benchmark for algorithms for community detection,

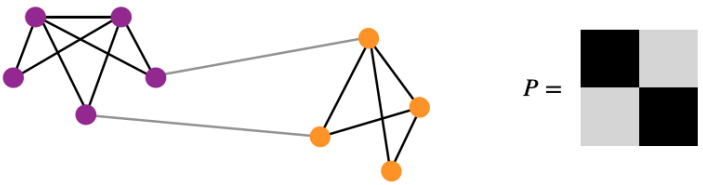


FIGURE 11. A schematic of a small network with community structure. The subgraphs with purple and orange nodes are densely connected (black edges), while there are fewer links connecting nodes between these purple/orange subgraphs (gray edges). Such an example can be generated by a stochastic block model with a matrix P of connection probabilities like the one visualized above: the diagonal blocks P_{ii}, P_{jj} represent higher connection probability between nodes in the same community (black), and the off-diagonal blocks P_{ij}, P_{ji} represent lower connection probability between nodes in different groups (lighter gray).

likely due to the presence of a "ground truth" (i.e., which faction the nodes would ultimately join) and its small size (34 nodes).

More recent examples of social networks with community structure abound. Examples include scientific collaboration networks [**NG04**], where individuals in similar subfields are more densely connected; Facebook friendships of Caltech students, where connections were more dense between students from the same dormitory ('House') [**TMP12**]; and in the retweet network of individuals using a particular hashtag on Twitter (now $\mathbb{X}$), where accounts with similar media-followership behaviors retweet each other much more often [**TECP20**].

Community detection in networks has been and remains a popular area of study, and methods abound to tackle this problem. Many of the topics we have addressed so far in this article form the foundations of these methods: some authors have used centrality-based methods (e.g., using betweenness as in the Girvan–Newman method [**GN02**]) and modularity optimization (e.g., with spectral partitioning via the graph Laplacian as in [**New06**]). To survey community detection methods, benchmarks for community detection, and further generalizations, we refer the reader to [**POM$^+$09**, **FH16**].

At this point, we meet our final class of generative models for this article. *Stochastic block models* are a family of random graph models that capture community structure. Suppose, as before, that we presume each node belongs to a particular group or type; the ith group is denoted g_i. We can then generalize the ideas from the $G(n,p)$ model discussion in Section 2.1.1 by defining the probability of an edge between to nodes depending on their respective types. That is, the probability of an edge between node i and j is $p_{g_i g_j}$. If we define a matrix of connection probabilities $P_{ij} = p_{g_i g_j}$, this creates a block structure in this matrix P (giving this model its name). This versatile model allows us to create random graphs with a variety of interesting features. First, note that the classical $G(n,p)$ model is a special case of this stochastic block model. We can also create a network with community structure (by defining, e.g., $P_{ii}, P_{jj} > P_{ij}, P_{ji}$) or disassortative networks

(with $P_{ij}, P_{ji} > P_{ii}, P_{jj}$). We can even create networks with *core-periphery structure*, where certain nodes (the *core*) have high degree relative to the remaining nodes (the *periphery*). Supposing that group g_i denotes the core nodes, this is achieved by allowing $P_{ii} > P_{ij}, P_{ji} > P_{jj}$. For a visualization of a stochastic block model for generating community structure, see Figure 11.

3. Overview and conclusions

In this article, we have explored several key features of the structure of social networks. We briefly review these properties below.

A large component: It is common in social networks for a high proportion of nodes to be in the same connected component (Section 2.1).

Heavy-tailed degree distribution: Social networks often contain a small number of nodes whose degree is relatively high, and a larger number of nodes with small degree (Section 2.2.3).

Small diameter: In many social networks, the shortest path between any two nodes can be quite small; this is sometimes known as the *small-world property* (Section 2.3).

High level of clustering: If two nodes i and j in a social network are each connected to the same third node k, it is more likely that i and j are also connected to each other (Section 2.3.1).

Assortative by degree: Social networks may be assortative, disassortative, or neither, depending on the node classes or types under consideration. However, one frequently-shared property of social networks is that nodes with high degree are more likely to be connected to other nodes of high degree (Section 2.4).

Community structure: Many social networks contain relatively densely-connected subgraphs called communities, with sparser connections between these communities (Section 2.4.2).

While it is interesting to measure and observe these properties empirically in real-world networks, analyzing generative mathematical models of networks can give insight into the possible mechanisms underlying the commonly-observed structural features listed above. In this article, we described several well-known generative models for networks to help illuminate these properties.

$G(n,p)$ and $G(n,m)$ models: In $G(n,p)$ random graph models, the number of nodes is fixed, with each edge between a pair of nodes occuring with probability p. In $G(n,m)$ random graph models, the number of nodes n and edges m is fixed. These models are simple to generate and are amenable to analysis (Section 2.1.1).

Preferential attachment models: Generative models of preferential attachment such as those by Price and Barabási–Albert give us a way to create a network where the degree distribution of the nodes follows a power law as the number of nodes gets large. This is achieved by attaching new nodes to an existing node with a probability that is proportional to the degree of the existing node (Section 2.2.3).

Configuration models: Configuration models are a family of models where the degree distribution or degree sequence is fixed and a network is generated that satisfies the given distribution. These models are popular choices due to their flexibility, their analytical convenience, and the ability to match degree distributions from real systems (Section 2.2.3).

Watts–Strogatz models: The Watts–Strogatz model is an example of a generative network model that can create networks with high levels of clustering and small shortest path lengths. This model is parameterized via a rewiring probability (Section 2.3.2).

Stochastic block models: Stochastic block models can generalize some of the previously described models to generate networks with community structure by allowing nodes to have heterogeneous connection probabilities between different groups or classes (Section 2.4.2).

The mathematical study of social networks is an exciting and growing field. There is much work to be done both in terms of understanding the applications and in terms of advancing the mathematics of networks [**Por20**]. From a mathematical perspective, there is still much work to be done to understand the properties of time-dependent (or *temporal*) networks, that is, networks that change in time by adding, removing, or changing nodes and/or edges in time [**HS12**, **ML16**]. Another active area of research is the theoretical study and application of dynamical systems on networks [**PG16**]. In such a system, each node has a state that changes in time (perhaps through interactions that are restricted via the network structure, e.g., occurring via edges in the network). There are many interesting properties to study in such systems: transient states, stationary states, stability analyses, bifurcations, and more. *Adaptive* or *co-evolving* network models combine these two ideas by studying networks where the node states and network structure interact and change in time [**SPS$^+$13**, **GDB06**].

There are also several interesting structural generalizations worth noting. Multilayer networks [**KAB$^+$14**, **Por18**, **Fin21**] provide a framework to represent different types of relationships or time-dependent relationships between nodes [**FPW21**]. Research on higher-order networks that encode relationships beyond pairwise relationships includes the study of hypergraphs and simplicial complexes [**BCI$^+$20**, **Bia21**, **BGL16**, **BGHS21**, **MHJ22**], which can be important for understanding structure and dynamics of social systems. The study of dynamical systems on these higher-order structures is still relatively young and there is much work to be done in these areas.

In this tutorial, we provided an overview of foundational topics in understanding structural properties and generative models of social networks. With these foundational techniques in hand, the hope is that the reader feels empowered to continue to dive deeper into the exciting work in the mathematics of social networks.

Acknowledgments

The author gratefully acknowledges her co-organizers Michelle Feng, Alexandria Volkening, and Mason Porter, along with the American Mathematical Society, for their support and dedication in organizing this short course. This article was also informed and improved by the many thoughtful comments and questions from the participants of the 2021 short course. The author also acknowledges helpful reviewers, whose careful comments strengthened and clarified the article.

References

[AB02] Réka Albert and Albert-László Barabási, *Statistical mechanics of complex networks*, Rev. Modern Phys. **74** (2002), no. 1, 47–97, DOI 10.1103/RevModPhys.74.47. MR1895096

[AMS] AMS, *MathSciNet collaboration distance tool*, accessed November 11, 2021. `https://mathscinet.ams.org/mathscinet/freeTools.html?version=2`

[BA99] Albert-László Barabási and Réka Albert, *Emergence of scaling in random networks*, Science **286** (1999), no. 5439, 509–512, DOI 10.1126/science.286.5439.509. MR2091634

[BCAB+21] Joseph B. Bak-Coleman, Mark Alfano, Wolfram Barfuss, Carl T. Bergstrom, Miguel A. Centeno, Iain D. Couzin, Jonathan F. Donges, Mirta Galesic, Andrew S. Gersick, Jennifer Jacquet, et al., *Stewardship of global collective behavior*, Proceedings of the National Academy of Sciences **118** (2021), no. 27, e2025764118.

[BCI+20] 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*, Phys. Rep. **874** (2020), 1–92, DOI 10.1016/j.physrep.2020.05.004. MR4147650

[BF20] Heather Z. Brooks and Michelle Feng, *Penguins of Kyoto multilayer network*, 2020. `https://bitbucket.org/mhfeng/penguins_of_kyoto/src`

[BGHS21] Christian Bick, Elizabeth Gross, Heather A. Harrington, and Michael T. Schaub, *What are higher-order networks?*, SIAM Rev. **65** (2023), no. 3, 686–731, DOI 10.1137/21M1414024. MR4624333

[BGL16] Austin R. Benson, David F. Gleich, and Jure Leskovec, *Higher-order organization of complex networks*, Science **353** (2016), no. 6295, 163–166.

[Bia21] Ginestra Bianconi, *Higher-order networks*, Cambridge University Press, 2021.

[BJN+02] Albert-László Barabási, Hawoong Jeong, Zoltan Néda, Erzsebet Ravasz, Andras Schubert, and Tamas Vicsek, *Evolution of the social network of scientific collaborations*, Phys. A **311** (2002), no. 3-4, 590–614, DOI 10.1016/S0378-4371(02)00736-7. MR1943379

[BL06] Kurt Bryan and Tanya Leise, *The* $25,000,000,000 *eigenvector: the linear algebra behind Google*, SIAM Rev. **48** (2006), no. 3, 569–581, DOI 10.1137/050623280. MR2278443

[BMS04] Peter S. Bearman, James Moody, and Katherine Stovel, *Chains of affection: The structure of adolescent romantic and sexual networks*, American Journal of Sociology **110** (2004), no. 1, 44–91.

[Bol80] Béla Bollobás, *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs*, European J. Combin. **1** (1980), no. 4, 311–316, DOI 10.1016/S0195-6698(80)80030-8. MR595929

[Bul19] Francesco Bullo, *Lectures on network systems*, Kindle Direct Publishing, 2019.

[BV14] Paolo Boldi and Sebastiano Vigna, *Axioms for centrality*, Internet Math. **10** (2014), no. 3-4, 222–262, DOI 10.1080/15427951.2013.865686. MR3259267

[CC98] James J. Collins and Carson C. Chow, *It's a small world*, Nature **393** (1998), no. 6684, 409–410.

[CTS16] Aaron Clauset, Ellen Tucker, and Matthias Sainz, *The Colorado Index of Complex Networks*, 2016, accessed January 20, 2023. `https://icon.colorado.edu`

[FH16] Santo Fortunato and Darko Hric, *Community detection in networks: a user guide*, Phys. Rep. **659** (2016), 1–44, DOI 10.1016/j.physrep.2016.09.002. MR3566093

[Fin21] Kelly R. Finn, *Multilayer network analyses as a toolkit for measuring social structure*, Current Zoology **67** (2021), no. 1, 81–99.

[FPW21] David N. Fisher and Noa Pinter-Wollman, *Using multilayer network analysis to explore the temporal dynamics of collective behavior*, Current Zoology **67** (2021), no. 1, 71–80.

[GDB06] Thilo Gross, Carlos J. Dommar D'Lima, and Bernd Blasius, *Epidemic dynamics on an adaptive network*, Physical Review Letters **96** (2006), no. 20, 208701.

[Gle15] David F. Gleich, *PageRank beyond the web*, SIAM Rev. **57** (2015), no. 3, 321–363, DOI 10.1137/140976649. MR3376760

[GN02] Michelle Girvan and Mark E. J. Newman, *Community structure in social and biological networks*, Proc. Natl. Acad. Sci. USA **99** (2002), no. 12, 7821–7826, DOI 10.1073/pnas.122653799. MR1908073

[Gou12] Ronald Gould, *Graph theory*, Dover Publications, Inc., Mineola, NY, 2012. Corrected reprint of the 1988 original [MR1103114]. MR3114560

[GYA18] Jonathan L. Gross and Jay Yellen, *Graph theory and its applications*, 2nd ed., Discrete Mathematics and its Applications (Boca Raton), Chapman & Hall/CRC, Boca Raton, FL, 2006. MR2181153

[Hol19] Petter Holme, *Rare and everywhere: Perspectives on scale-free networks*, Nature Communications **10** (2019), no. 1, 1–3.

[HS12] Petter Holme and Jari Saramäki, *Temporal networks*, Physics Reports **519** (2012), no. 3, 97–125.

[Jac10] Matthew O. Jackson, *Social and economic networks*, Princeton University Press, Princeton, NJ, 2008. MR2435744

[KAB+14] Mikko Kivelä, Alex Arenas, Marc Barthelemy, James P. Gleeson, Yamir Moreno, and Mason A. Porter, *Multilayer networks*, Journal of Complex Networks **2** (2014), no. 3, 203–271.

[Kat53] Leo Katz, *A new status index derived from sociometric analysis*, Psychometrika **18** (1953), no. 1, 39–43.

[LFH10] Andrea Landherr, Bettina Friedl, and Julia Heidemann, *A critical review of centrality measures in social networks*, Wirtschaftsinformatik **52** (2010), 367–382.

[MHJ22] Raffaella Mulas, Danijela Horak, and Jürgen Jost, *Graphs, simplicial complexes and hypergraphs: spectral theory and topology*, Higher-order systems, Underst. Complex Syst., Springer, Cham, [2022] ©2022, pp. 1–58, DOI 10.1007/978-3-030-91374-8_1. MR4433789

[Mil67] Stanley Milgram, *The small world problem*, Psychology Today **2** (1967), no. 1, 60–67.

[ML16] Naoki Masuda and Renaud Lambiotte, *A guide to temporal networks*, Series on Complexity Science, vol. 4, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2016, DOI 10.1142/q0033. MR3585359

[MSGL14] Seth A. Myers, Aneesh Sharma, Pankaj Gupta, and Jimmy Lin, *Information network or social network? The structure of the Twitter follow graph*, Proceedings of the 23rd International Conference on World Wide Web, 2014, pp. 493–498.

[New01a] Mark E. J. Newman, *Networks*, Oxford University Press, Oxford, 2010. An introduction, DOI 10.1093/acprof:oso/9780199206650.001.0001. MR2676073

[New01b] Mark E. J. Newman, *The structure of scientific collaboration networks*, Proc. Natl. Acad. Sci. USA **98** (2001), no. 2, 404–409, DOI 10.1073/pnas.021544898. MR1812610

[New02] Mark E. J. Newman, *Assortative mixing in networks*, Physical Review Letters **89** (2002), no. 20, 208701.

[New06] Mark E. J. Newman, *Finding community structure in networks using the eigenvectors of matrices*, Physical Review E **74** (2006), no. 3, 036104.

[New18] Mark E. J. Newman, *Networks*, Oxford University Press, Oxford, 2018. Second edition of [MR2676073], DOI 10.1093/oso/9780198805090.001.0001. MR3838417

[NG04] Mark E. J. Newman and Michelle Girvan, *Finding and evaluating community structure in networks*, Physical Review E **69** (2004), no. 2, 026113.

[PG16] Mason A. Porter and James P. Gleeson, *Dynamical systems on networks*, Frontiers in Applied Dynamical Systems: Reviews and Tutorials, vol. 4, Springer, Cham, 2016. A tutorial, DOI 10.1007/978-3-319-26641-1. MR3468887

[POM+09] Mason A. Porter, Jukka-Pekka Onnela, and Peter J. Mucha, *Communities in networks*, Notices Amer. Math. Soc. **56** (2009), no. 9, 1082–1097. MR2568495

[Por18] Mason A. Porter, *What is... a multilayer network?*, Notices Amer. Math. Soc. **65** (2018), no. 11, 1419–1423. MR3838058

[Por20] Mason A. Porter, *Nonlinearity + networks: A 2020 vision*, Emerging Frontiers in Nonlinear Science, Springer, 2020, pp. 131–159.

[Pri67] Derek de Solla Price, *Networks of scientific papers*, Science **149** (1967), no. 3683, 510–515.

[Pri76] Derek de Solla Price, *A general theory of bibliometric and other cumulative advantage processes*, Journal of the American Society for Information Science **27** (1976), no. 5, 292–306.

[SPS+13] Hiroki Sayama, Irene Pestov, Jeffrey Schmidt, Benjamin James Bush, Chun Wong, Junichi Yamanoi, and Thilo Gross, *Modeling complex systems with adaptive networks*, Comput. Math. Appl. **65** (2013), no. 10, 1645–1664, DOI 10.1016/j.camwa.2012.12.005. MR3061729

[TECP20] Mason A. Porter, *Nonlinearity + networks: a 2020 vision*, Emerging frontiers in nonlinear science, Nonlinear Syst. Complex., vol. 32, Springer, Cham, [2020] ©2020, pp. 131–159, DOI 10.1007/978-3-030-44992-6_6. MR4180848

[TMP12] Amanda L. Traud, Peter J. Mucha, and Mason A. Porter, *Social structure of Facebook networks*, Physica A: Statistical Mechanics and its Applications **391** (2012), no. 16, 4165–4180.

[UKBM11] Johan Ugander, Brian Karrer, Lars Backstrom, and Cameron Marlow, *The anatomy of the Facebook social graph*, Preprint, `arXiv:1111.4503`, 2011.

[VL07] Ulrike von Luxburg, *A tutorial on spectral clustering*, Stat. Comput. **17** (2007), no. 4, 395–416, DOI 10.1007/s11222-007-9033-z. MR2409803

[VLBB08] Ulrike von Luxburg, Mikhail Belkin, and Olivier Bousquet, *Consistency of spectral clustering*, Ann. Statist. **36** (2008), no. 2, 555–586. MR2396807

[WBH+18] Nakeya D. Williams, Heather Z. Brooks, Maryann E. Hohn, Candice R. Price, Ami E. Radunskaya, Suzanne S. Sindi, Shelby N. Wilson, and Nina H. Fefferman, *How disease risks can impact the evolution of social behaviors and emergent population organization*, Understanding complex biological systems with mathematics, Springer, 2018, pp. 31–46.

[Wes01] Douglas B. West, *Introduction to graph theory*, Prentice Hall, Inc., Upper Saddle River, NJ, 1996. MR1367739

[Wik] Wikipedia, *Six degrees of Wikipedia*, accessed November 11, 2021. `https://en.wikipedia.org/wiki/Wikipedia:Six_degrees_of_Wikipedia`

[WS98] Duncan J. Watts and Steven H. Strogatz, *Collective dynamics of 'small-world'networks*, Nature **393** (1998), no. 6684, 440–442.

[Zac77] Wayne W. Zachary, *An information flow model for conflict and fission in small groups*, Journal of Anthropological Research **33** (1977), no. 4, 452–473.

Department of Mathematics, Harvey Mudd College, Claremont, California 91711
Email address: `hzinnbrooks@g.hmc.edu`

Proceedings of Symposia in Applied Mathematics
Volume **80**, 2025
https://doi.org/10.1090/psapm/080/00731

Interpreting topology in the context of social science

Michelle Feng

Abstract. Topological data analysis (TDA) is a suite of techniques that allow researchers to analyze the "shape" or "structure" of their data through the lens of topology. Because topology refers to a very specific mathematical notion of shape, researchers who are new to topology may find topological notions to be unintuitive or difficult to use in their own applications. This article aims to give a gentle introduction to thinking in a topological framework. To illustrate the process of turning a social research question into a topological one, some exercises are included for guiding a practitioner from problem ideation to topological computation. Discussion of a hypothetical example is included for demonstration.

Introduction

Topological data analysis (TDA) is a field of applied mathematics which borrows techniques from topology (usually algebraic topology) to analyze "shapes" in a variety of datasets. Persistent homology (PH) is one topological tool which has been widely used in applications. The development of efficient and simple to use software has rapidly expanded the accessibility of PH, allowing researchers to use PH on a broader set of exciting applications. This article will discuss applications of PH to social datasets. I will focus on guiding practitioners through the variety of choices they can make in implementations of PH, explaining what some of the potential effects of those choices might be.

Topology is one field of mathematics which deals with "shape." Loosely, topology is concerned with how things are put together, or connected. When we study an object topologically, we throw away information about its precise shape (for example, curvature, or distance). In losing this information, we're left only with a notion of "structure" which focuses on connection. One way to think about it is to think in terms of building blocks – if we have some very simple building blocks, how can we glue them together to reconstruct our shape?

This article will not be a technical introduction to topological data analysis, but an intuitive one. While some definitions will be included, for the most part, I will be focusing on how to think about topology (in a very specific sense that I personally happen to find useful for working on social applications). There are many wonderful introductory texts about topology [**14**] and TDA [**7**, **22**, **30**] . In this

2020 *Mathematics Subject Classification.* Primary 55N31, 91F99.

article, my goal is to help researchers without a technical background in topology start thinking about how their research questions might be viewed through the framework of topology, and to figure out whether that framework is useful.

The remainder of this article is organized as follows. In Section 1, I give intuition for some topological concepts. In Section 2, I give definitions, building out the intuition from the previous section into mathematical language. Section 3 goes through a short tutorial on synthetic data, giving guidance for setting up a PH computation on a fresh dataset.

1. Intuition

I begin by providing some intuition. This section is meant to help readers with relatively little topological background begin to think topologically. While the frameworks described in this section are by no means the only (or necessarily the "correct") ways to think about topology, I've included observations and motivations that I personally found helpful when first approaching topological material. As with all things, intuition about topology takes time to build, and readers may use or discard observations as they find them useful.

Topology is sometimes described as "rubber sheet geometry" — that is, given an object made out of a rubber sheet, what aspects of its shape would be retained as one bent and stretched and molded it in one's hands? For example, as a child I had a much loved set of novelty dinosaur-shaped rubber bands. If I picked up a triceratops and stretched it, it wouldn't take much to turn it into an oval, or to wrap it around my fingers a few times. Topologically speaking, the triceratops, the oval, and the series of loops were therefore all the same shape (see Figure 1). As long as the rubber band didn't snap, these shapes could morph into and out of each other at my will. A little more mathematically speaking, all of these shapes were S^1 — given some effort, some continuous deformation of the circle could get me to any of these.

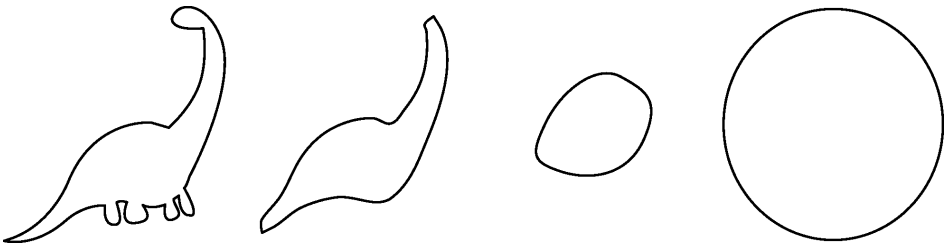


FIGURE 1. Rubber sheet geometry in action: we can continuously deform a line drawing of a brontosaurus into a circle by shrinking its extremities and stretching the curves.

Much of topology is concerned with understanding an object's shape in this very simplified setting. We don't care about the dinosaur horns, or the number of loops I could fit around my fingers. All we care about is the continuous length of the band and the existence of the hole in created in the middle. If we can understand these properties, then we have characterized our rubber band (and all other objects that form one single closed loop which distinguishes inside from outside).

What, then, does this type of rubber sheet geometry mean when applied to a dataset? How do I "stretch" a population, and if I do it, what does that even tell me? It can require some creativity to apply these intuitions to data, but one place to start is to think about how basic topological concepts relate to your data. In the following sections, I give some intuitions for a few topological ideas that I have found useful for thinking about data.

1.1. Nearness and neighborhoods. One of the core ideas used in topology is that of "nearness". In our own lives, we use concepts of nearness frequently. Physical distance has a huge impact on the places we live, the communities we join, and the way we live our lives [**2**, **11**]. In this context, we tend to think of nearness as being something measurable — how many miles away is my local grocery store, or how many minutes will it take me to drive to work? We can also conceptualize nearness in less concrete ways. For example, we might consider nearness in terms of shared ideologies, or social connections. Sometimes, we might find it useful to apply some sort number to these abstract types of nearness (e.g., Erdős numbers). When we measure nearness (usually with a distance metric, or something resembling one) and study the shape of an object equipped with that measurement, we are generally studying something geometric in nature. By contrast, much of topology is based on a more abstract notion of nearness, which doesn't rely on measurable distances.

Nearness in topology can be described in a few different ways, but in this article, I will use the topological concept of a neighborhood. In Section 2, I'll give a more precise definition of a topological space, but here, I will briefly give some guiding intuition. Topologically speaking, a neighborhood tells you when members of a topological space are close together. That is, given a topological space, if I know what all of the neighborhoods in that space are, I can tell you what things in the space are close together. If I want to do something a bit more complex, I can do things like combine neighborhoods to get bigger ones, or shrink neighborhoods. This stretching and shrinking of neighborhoods allows me to identify when some things are closer together than others. That is, if I am a member of a topological space, and you live in my topological neighborhood, then we are neighbors, and therefore, at least somewhat "near" each other. If I expand my topological neighborhood, I might find more individuals who are also near me, but who are not as near me as you are. Of course, I could keep expanding my topological neighborhood to cover the whole space, so that technically, everyone in the space is my neighbor, but there would still be some concept of who my closest neighbors are, simply by looking at the smallest neighborhoods I belong to (and equivalently, I would have some neighbors who are farther than the others, by virtue of only being in a very large neighborhood of myself). In Figure 2, we illustrate some potential neighborhoods of a space.

This neighborhood paradigm gives us one way of building a topological space from data. Given a data set, what does it mean for items in the data set to be close together? What are the neighborhoods of the data? These guiding questions can help practitioners begin to set up a topological framework for analyzing their data.

In some data sets, the answers to these questions may seem obvious. Suppose the data set I'm working with is a map of Los Angeles with some data attached. I could consider the neighborhoods to be, quite literally, the neighborhoods of Los Angeles, as defined by whatever the city government says the boundaries are. I'd have to put in some work to understand what it meant to expand or shrink

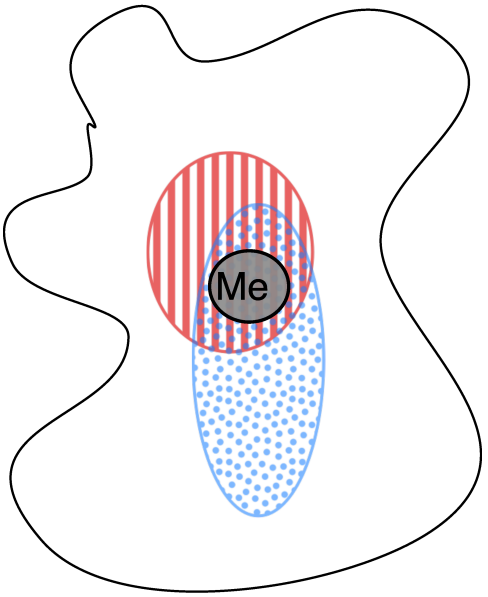
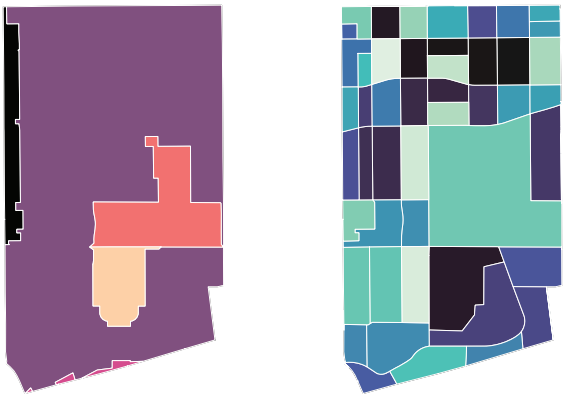


FIGURE 2. Neighborhoods determine nearness. In this figure, the red, blue, and grey neighborhoods contain me, and anything else in those neighborhoods are "close" to me in some sense. Since the grey neighborhood is entirely contained in the other two, my grey neighbors are closer to me than my other red and blue neighbors.

those neighborhoods, but the connection between city neighborhoods and topological neighborhoods doesn't feel like too much of a stretch. But there are a lot of other choices I could make about my topological neighborhoods. I could consider neighborhoods to be epsilon-balls centered at particular addresses, or households that shop at the same grocery stores, or buildings that share the same owners. All of these choices would result in very different ideas of nearness, a different set of neighborhoods (see Figure 3), and likely, different results from the application of standard TDA tools.

1.2. Connectedness. Another useful topological concept is that of connectedness. The definition of connectedness is simple, but not very elucidating — a space is connected if it is not disconnected. Intuitively, this means roughly what you might expect; if a topological space cannot be separated into parts that don't touch each other, then the space is connected. Sometimes (frequently, in my experience with applications), connectedness coincides with path-connectedness. That is, if points A and B are part of the same topological space, then A and B are path-connected if there exists some continuous path through the space that starts at A and ends at B.



(A) (B)

FIGURE 3. Different ways of choosing neighborhoods, using a map of the census tract containing Caltech. In Figure 3a, the census tract is segmented into neighborhoods by shared ZIP code. In Figure 3b, the census tract is segmented into neighborhoods of smaller census blocks. Note here that the visualization doesn't show any methods of shrinking or expanding neighborhoods, as I've said is possible, but instead shows a specific choice of disjoint neighborhoods at a specific scale. In Section 2 I'll talk more about how to shrink and expand neighborhoods.

Sometimes, connectedness can coincide with nearness. For example, if my concept of nearness is "within some physical distance of myself", and my concept of connectedness is "places that I can walk from my current location", then I likely will only be connected to locations which are in a small neighborhood of myself, and I will not be connected to locations which are not.

In other cases, it can be useful to completely separate connectedness and nearness. For example, I certainly live near every person on my block, but I haven't properly met most of them (in my defense, I moved house in the middle of a pandemic, and perhaps more saliently, am also a misanthrope). In this case, if I consider nearness to again be physical distance, and connectedness to be the existence of some social connection, nearness and connectedness have very little to do with each other. I'm almost not connected to all to the people who are nearest me, and many of my connections are very far indeed. Figure 4 shows a visualization of this example.

1.3. Holes. The last concept I will cover in this section is the topological "hole" (alternatively, "void"). As a note, in different areas of topology, a "hole" can refer to slightly different things. In the interest of keeping it vague and intuitive, for the purposes of this section, a hole is some gap in a topological space that prevents the space from being completely solid. In different dimensions, this means different things. That is, a one-dimensional hole is anything you could stick an arbitrarily pen through (e.g., the space in the middle of a hula hoop, or the eye of a needle). Higher-dimensional holes are things you could fill with jam, with the slight caveat that there ought not be some opening the jam can escape through (e.g., the inside

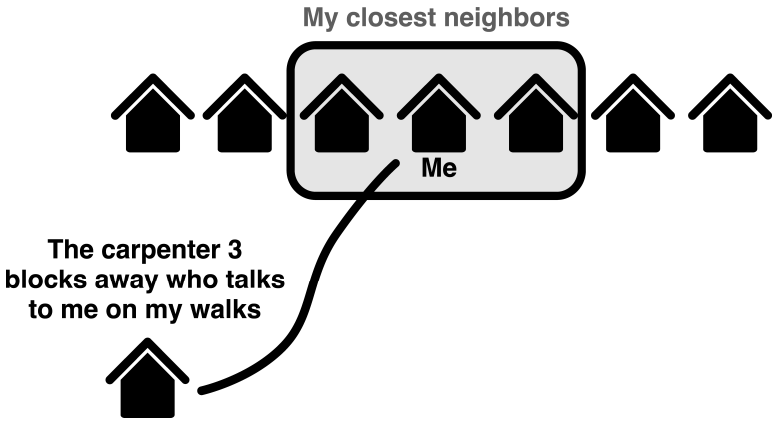


FIGURE 4. Notions of nearness and connectedness don't always coincide. People who are physically near me may not be connected to me in any meaningful sense, for example.

of a donut, or your bedroom with all the doors and windows closed, if you were very ambitious). I'll be a bit more rigorous about holes in Section 2, but for now, hopefully the jam-filling is evocative enough to get us through.

One of the big ideas of topology is that objects can be understood via their holes. Remember what I said before about topology being interested in objects up to stretching and squishing? The idea here is that we can never change the holes of a shape via this kind of stretching or shrinking. If we want to get rid of a hole, we have two options — cutting (imagine slicing a rubber band to remove the hole), or glueing, neither of which is allowed in our rubber sheet geometry paradigm. If we want to make new holes, again we either need to cut one in, or glue some things together. This means that the holes of our topological space are invariant under the types of continuous deformations we allow, and as such, can be used to classify the topology of the space.

For the purposes of data, we can think about these types of holes in a few different ways. The first is to treat them as an invariant, as in the previous paragraph. That is, if we have several different shapes, and we want to be able to tell them apart, we can compute information about their holes, and use that to distinguish them. If we want to tell if shapes are the same, we can also do this using holes, though this gets a bit trickier, and depends on which topological concept of a hole we are using (more on this later).

We can also study the holes in and of themselves. If we understand where the holes are in our dataset, we might be able to interpret them in terms of our data. For example, if our data were examining arctic ice, we might very concretely be able to interpret holes in the topology of our data as holes in ice (with the caveat that we would need to set up a topological space in a way that led to this interpretation). Physical data sets in particular can sometimes have very concrete interpretations of holes[**5**, **15**].

Another way that we can think about topological holes is as obstructions to some type of connectedness. Holes can tell us about ways in which it is impossible to move through a structure. This interpretation can be used in very concrete

ways (e.g., inaccessible areas in a transportation network), or in more abstract ones (e.g., communities which are not connected to each other because of differences in ideology). When studying these topological holes, it's very important to keep in mind the properties of the space they inhabit. The interpretation of holes depends heavily on how you frame your data set as a topological space. Different notions of connectedness or nearness can lead to different interpretations for any holes you observe.

2. Definitions and useful theorems

The concepts discussed in the previous section are far from comprehensive, but, in my view, give a good starting point to practitioners looking to try their hand at TDA. In this section, I will give rigorous mathematical definitions which cover all of the ideas described above, as well as some additional definitions and theorems which appear frequently in TDA software packages. In particular, I will be focusing on persistent homology (PH), a branch of TDA which has been made accessible via a great deal of out of the box software[**8**, **21**, **25**, **26**]. While the concepts discussed above are generally useful ones for thinking about many different types of topological analysis, from this point on the article will largely be guided by what you need to know to get started with PH.

The following sections have been written to be accessible to people with a math background that does not include any topology. I attempt to provide a lot of motivating intuition for unfamiliar readers, but some knowledge of mathematical notation is assumed. References will be provided for readers seeking more in depth material in each of the following sections.

2.1. Topological spaces.

DEFINITION 2.1. Let X be a set. Let $\mathcal{N}$ be a function assigning to each point $x \in X$ a nonempty collection $\mathcal{N}(x)$ of subsets of X. The elements of $\mathcal{N}(x)$ are called *neighborhoods* of x. If the following axioms are satisfied, then X with $\mathcal{N}$ is called a *topological space*.

(1) If N is a neighborhood of x (i.e., $N \in \mathcal{N}(x)$), then $x \in N$
(2) If N is a subset of X and includes a neighborhood of x, then N is a neighborhood of x
(3) The intersection of two neighborhoods of x is a neighborhood of x
(4) Any neighborhood N of x includes a neighborhood M of x such that N is a neighborhood of each point of M

Connecting this definition to the intuitions in Section 1, we see that the function $\mathcal{N}(x)$ is what tells us which subsets of X are a neighborhood of x. The second axiom corresponds to expanding our neighborhoods — if I live in a space, any subset of that space which includes one of my neighborhoods must also be a neighborhood of mine (albeit a potentially bigger neighborhood). The third axiom lets us shrink neighborhoods — if I belong to two different neighborhoods, and take the intersection of those two neighborhoods, I will get another neighborhood which I belong to (potentially smaller). Finally, the last axiom tells me about the relationship between my neighborhood and the neighborhoods of my neighbors. That is, if I belong to a neighborhood N, there is some (potentially smaller) neighborhood M so that N is also a neighborhood of every one of my neighbors in M. Note that

it is possible for me to belong to a set which is not a neighborhood of mine (I'll illustrate more precisely how shortly).

EXAMPLE 2.2. For one mathematical example of a neighborhood, consider the real line $\mathbb{R}$. For any $x \in \mathbb{R}$, we say that N is a neighborhood of x if N contains an open interval around x. We check that the axioms are satisfied as follows.

(1) If N contains an open interval of x, it certainly contains x
(2) If we take a subset of $\mathbb{R}$ that contains a neighborhood of x, then that subset contains an open interval of x, and is therefore a neighborhood of x.
(3) Let N_1 and N_2 be neighborhoods of x. Then N_1 contains some open interval I_1 of x, and similar N_2 contains some open interval I_2 of x. $N_1 \bigcap N_2$ contains $|_1 \bigcap I_2$, which itself is an open interval of x. So $N_1 \bigcap N_2$ is a neighborhood of x.
(4) Let N be a neighborhood of x. Then N contains I, an open interval of x. But I is a neighborhood of x, and for all $y \in I$, I is an open interval containing y. So N contains an open interval of each $y \in I$, and is therefore a neighborhood of each $y \in I$.

Note that with this topology, 0 is in $[0, 1)$, which is not a neighborhood of 0. So points can belong to sets which are not neighborhoods of theirs.

EXAMPLE 2.3. For a more applied example, consider Figure 3.

EXERCISE 2.4. Check that the neighborhood structure described in this example satisfies the axioms in Definition 2.1

The examples given in this section all use Definition 2.1, but there are other equivalent definitions of a topological space. Some of these definitions can be found in [**14**]. In practice, when it comes to data, I find it most elucidating to think about my data in terms of its neighborhoods, which is why I chose to present Definition 2.1, but there are many different ways to conceptualize topological spaces. For practitioners, the most important thing to keep in mind is that having a set of data points does not give you a topological space. Rather, your set (X) must be equipped with some neighborhood function $\mathcal{N}$ in order to be a topological space. TDA computations frequently do not rely on explicitly knowing or stating that neighborhood function, so you may find that for your purposes it isn't necessary to identify your neighborhoods, or check that they fulfill the definition of a topological space. However, it is important to keep in mind that by working with data as a topological space (which we do, when using PH packages, or other tools from TDA), that there is some implicit choice being made in order to abstract our data. When interpreting the results of a TDA computation, it can be helpful to spend some time thinking about what space you're working in.

2.2. Simplicial complexes. While it is useful to know what a topological space is, in practice, in PH it is rare to work with the definition of topological space given above. For ease of computability, it is standard in PH to work with specific types of topological spaces. In this article, we will be working with simplicial complexes. Roughly speaking, a simplicial complex is a topological space made up of triangular building blocks. These building blocks are called simplices. Using these triangular building blocks makes it easier to perform topological computations, so frequently, one works with the building blocks rather than with neighborhoods or

topologies directly. Note that sometimes in PH, practitioners can use other types of complexes (e.g., cubical complexes, cell complexes). The difference between the various types of complexes generally lies in the shape of the building blocks (e.g., rectangular building blocks for cubical complexes, etc). I will focus on simplicial complexes, but the ideas described in this article also apply to other choices of complex. For more detail, see [**7**, **14**].

DEFINITION 2.5. A *k-simplex* is a k-dimensional polytope which is the convex hull of its $k+1$ vertices. A k-simplex is denoted

$$\sigma = (v_0, \ldots, v_k),$$

where the order of the vertices denotes the orientation of the simplex.

For example, a 0-simplex is a point (1 vertex). A 1-simplex is a line segment (2 vertices, and the segment joining them). A 2-simplex is a triangle (3 vertices, the lines joining each pair of them, and the filled in triangle between those lines). A 3-simplex is a tetrahedron, and so on so forth. We can consider a simplex to be the generalization of a triangle to an arbitrary number of dimensions. For a visualization of simplices in different dimensions, see Figure 5.

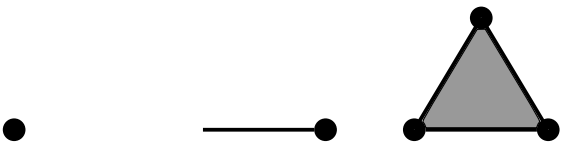


FIGURE 5. From left to right, a 0-simplex (point), 1-simplex (line segmenting connecting two points), and a 2-simplex (3 points connected pairwise by line-segments, with the face between the three line segments filled in).

Note that k-simplices by definition contain smaller simplices, which we call faces.

DEFINITION 2.6. Let $\sigma = (v_0, \ldots, v_k)$ be a k-simplex. Then any subset

$$\tau = \{v_{i_0}, \ldots, v_{i_m}\}$$

of the vertices of σ generates an m-simplex. We call τ an *m-face* of σ.

Roughly speaking, a simplicial complex is a set of simplices which we have glued together along faces (see Figure 6). This allows us to use simplices as building blocks to create much more complicated shapes, as we can chain simplices together along faces of arbitrary dimension. The precise definition of a simplicial complex is as follows.

DEFINITION 2.7. A *simplicial complex* S is a set of simplices that satisfies the following conditions.

(1) Every face of a simplex from S is also in S
(2) The nonempty intersection of any two simplices $\sigma_1, \sigma_2 \in S$ is a face of both σ_1 and σ_2

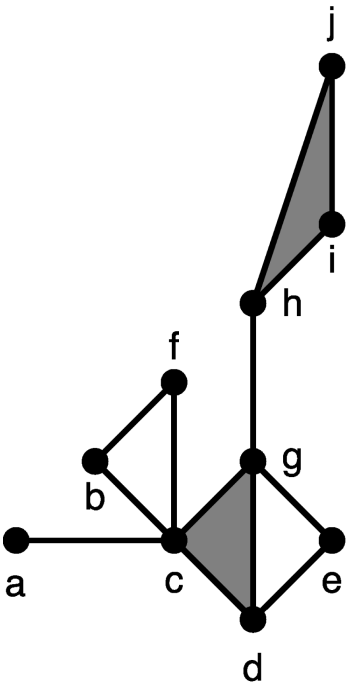


FIGURE 6. A simplicial complex. There are ten 0-simplices, thirteen 1-simplices, and two 2-simplices in this image. As an example of how to compute faces consider the 2-simplex $\tau = \{c, d, g\}$. It has three 1-faces $\{(c, d), (d, g), (g, c)\}$, and three 0-faces $\{c, d, g\}$. If we consider the intersection of the simplices $\tau = \{c, d, g\}$ and $\tau' = \{a, c\}$, the intersection $\tau \cap \tau' = \{c\}$ is a 0-face of both τ and τ'.

Simplicial complexes are useful because they allow us to represent complex objects in a combinatorial fashion. As we will see in the following section, we can compute topological properties of a simplicial complex using combinatorial algorithms. When working with data, the first step is often constructing a simplicial complex on the data set. Technically, all one needs to construct a simplicial complex is some set of vertices, which one can then build up into simplexes that are glued together in some fashion. In a data set based on a point cloud, for example, I could take the points of the point cloud to be our vertices, and then come up of some way of deciding when to put edges between the vertices, triangles between triplets of edges, and so on so forth. That is, I start with 0-simplices, and then add higher-dimensional simplices, checking that I never attempt to put in a k-simplex somewhere where all its faces aren't already part of the complex. In this manner, I would obtain a valid simplicial complex, ready to plug into some computations. I draw an arbitrary simplicial complex based on 10 points in Figure 7.

However, this description of building a simplicial complex is pretty vague. How exactly should I decide when to put in a higher-dimensional simplices? For practitioners, this decision is a very important one. Given a large number of vertices, it is straightforward to see that there are a very large number of possible simplices

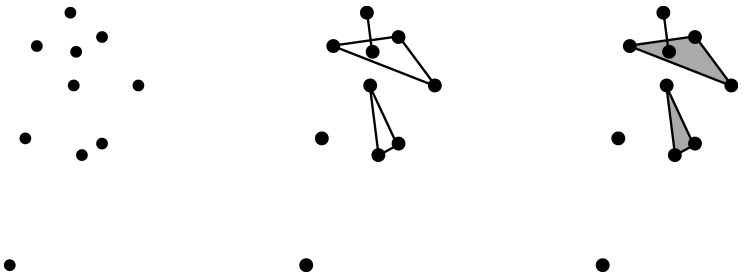


FIGURE 7. Arbitrary construction of a simplicial complex on a point cloud of 10 points. At the left, we start with our 0-simplices. Next, we add 1-simplices arbitrarily. Finally, we add 2-simplices. Note that in this example, all possible 2-simplices have been added to the complex.

that can go or not go into a simplicial complex. How can one distinguish between all of the possible choices of complex? For now, I will leave this question aside until Section 3, but keep in mind that this decision has a tremendous impact on any topological analysis.

2.3. Homology. The primary topological tool used in PH is homology. Intuitively, homology theories associate topological spaces to their holes. Again, in this section, all the definitions I will give are for simplicial homology, but other homology theories exist for other types of complexes. All of them function in the same way — given a topological space of a certain type, what are its holes? For more information on other homology theories see [**14**].

In order to give the full definition of simplicial homology, I will be referring to algebraic concepts (specifically, groups, kernels, and images). For readers unfamiliar with these ideas from algebra, see [**14**]. The general idea of simplicial homology is as follows. Starting with a simplicial complex S, I can find some way of identifying sets of simplices that might surround a hole. Once I've identified these sets, I can check whether each individual hole is filled by other simplices. If not, I observe a hole. There are a lot of definitions in this section, but essentially, these definitions function to build out the tools I need in order to carry out this hole-identifying process.

DEFINITION 2.8. Let S be a simplicial complex. A *simplicial k-chain* is a finite formal sum

$$\sum_{i=1}^{N} c_i \sigma_i,$$

where each c_i is an integer and σ_i is an oriented k-simplex. Each oriented simplex is equation to the negative of the simplex with the opposite orientation. That is,

$$(v_0, \ldots, v_i, \ldots, v_j, \ldots, v_k) = -(v_0, \ldots, v_j, \ldots, v_i, \ldots, v_k).$$

The group of k-chains on S is written C_k. C_k is the free abelian group generated by the k-simplices in S. C_k gives us formalism for working with combinations of simplices, and allows us to define an operator which will be used to detect holes.

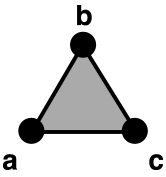


FIGURE 8. The boundary operator applied to this triangle gives us the traversal around the edges of the triangle going clockwise.

DEFINITION 2.9. Let $\sigma = (v_0, \ldots, v_k)$ be an oriented k-simplex, and a basis element of C_k. We define the *boundary operator*

$$\partial_k : C_k \to C_{k-1}$$

by the operation

$$\partial_k(\sigma) = \sum_{i=0}^{k} (-1)^i (v_0, \ldots, \hat{v_i}, \ldots, v_k),$$

where

$$(v_0, \ldots, \hat{v_i}, \ldots, v_k)$$

is the $(k-1)$-face of σ obtained by deleting the i-th vertex.

Consider the example of a triangle (a, b, c), as in Figure 8. Applying the boundary gives

$$\begin{aligned}\partial_2((a,b,c)) &= (-1)^0(b,c) + (-1)^1(a,c) + -1(^2)(a,b)\\ &= (b,c) - (a,c) + (a,b).\end{aligned}$$

The boundary operator is so-named because it allows one to detect boundaries. Given a simplex σ, the boundary operator returns a formal sum of the faces that make up the boundary of σ.

Because of the way we've set up orientation, a direct computation shows us that $\partial^2 = 0$. That is, the boundary of a boundary is nothing. For example, for the figure in Figure 8

$$\begin{aligned}\partial(\partial(a,b,c)) &= \partial((b,c) - (a,c) + (a,b))\\ &= (c-b) - (c-a) + (b-a) = 0.\end{aligned}$$

This observation is the final piece of information needed to detect holes. Suppose that I am looking for a hole shaped like a triangle. First, I will find all the sets of three edges that could potentially surround a triangle. Because $\partial^2 = 0$, this will be precisely the sums $\sigma_0 + \sigma_1 + \sigma_2$ where $\partial(\sigma_0 + \sigma_1 + \sigma_2) = 0$. Then, I observe that if there is a triangle in the middle, there should be some τ such that

$$\partial(\tau) = \sigma_0 + \sigma_1 + \sigma_2.$$

So I have a triangular hole if there is no such τ. I can apply the same logic to holes which surround groups of simplices rather than a single simplex.

Alternatively, what I am doing is looking for elements of C_2 that are in the kernel of ∂_2, but not in the image of ∂_3. That is, the holes are elements of the group

$$H_k(S) = \frac{\ker \partial_k}{\operatorname{im} \partial_{k+1}}.$$

FIGURE 9. For this space, the 0-th Betti number is 2, and the 1-st Betti number is 5.

DEFINITION 2.10. The *k-th homology group* H_k of S is defined to be the quotient

$$H_k(S) = \frac{\ker \partial_k}{\operatorname{im} \partial_{k+1}}.$$

So the homology group gives me the holes in S, but it also gives me a group structure on those holes. This is interesting because it gives me algebraic relations between the holes. In application, these algebraic relationships are rarely used, but the group structure of the homology group is useful for building theory. For practitioners, the rank of $H_k(S)$, called the *k-th Betti number* of S, can be used in a variety of applications. Betti numbers can be used to classify topological spaces, and can also be thought of as the number of times an object can be "cut" while still remaining connected, or the number of "voids" of the correct dimension (see Figure 9). For some applications of Betti numbers, see [**23**, **27**].

As a note, the homological features of a simplicial complex are entirely determined by the construction of the simplicial complex. A k-dimensional hole exists wherever there is some group of k-simplices that are pasted together in a way that could be filled in by $k+1$-simplices, but isn't. That is, if I have a simplicial complex on some data set, I can understand what the k-holes in my data set are by thinking about what the k-simplices and $k+1$-simplices are. This is very useful for interpreting the topological features and homology groups of a simplicial complex constructed from data. While I can treat homology as a black box tool which takes in a topological space and spits out an invariant, I can also use my data to give context and meaning to homological features.

While homology groups of topological spaces are interesting in and of themselves, when it comes to working with data, it can be difficult to figure out which space's homology group to look at. That is, given a data set, I can build a simplicial complex (or other topological space) over it in an arbitrary number of ways. While I could simply pick one method and then compute the homology groups of the chosen construction, the resulting homology groups might be more reflective of my choice of construction than of any underlying topological signals in the data. One way to mitigate this is to look at a collection of constructions and their homology groups, paying attention to which topological features appear across a variety of constructions. If a topological feature appears in many different choices of simplicial complex construction, I can be more confident that the topological feature is reflecting something about the base data set.

2.4. Filtrations. One particularly useful way to examine a set of constructions is to look at a filtered simplicial complex, or filtration.

DEFINITION 2.11. A *filtered simplicial complex* S (or *filtration*) is a collection $\{S_i\}$ of simplicial complexes such that

$$S_0 \subseteq S_1 \subseteq S_2 \subseteq \ldots \subseteq S_n = S. \tag{2.1}$$

I refer to individual S_i as steps of the filtration.

A filtration can be any collection of simplicial complexes that fulfill this nesting property. This means that as with constructing simplicial complexes, there is a great deal of flexibility in constructing a filtration.

The reason that I require each complex in a filtration to be contained in the next goes back to homology computation. As long as this property is satisfied, I can actually compute the homology of every step of the filtration simultaneously. Additionally, I can track topological features in the homology groups of individual steps through the entire filtration. I will leave this aside until Section 2.5, but this tracking ability is a crucial property of filtrations, and will turn out to be very useful later.

It is not difficult to come up with a sequence of simplicial complexes that can satisfy equation 2.1. In the following paragraphs, I will introduce several common choices to help give practitioners some ideas. In practice, familiarity with a data set should guide the choice of filtration. When setting up a filtration, it is useful to keep in mind that elements of the homology group will be entirely dictated by where the simplices are. If I know what the requirements for k-simplices and $k+1$-simplices are, then I can entirely describe elements of H_k based purely on that knowledge. As such, it is wise to leverage any knowledge I have of my data set in choosing where to put the simplices.

2.4.1. *Distance-based constructions.* One of the most common ways to build a filtration is to embed a data set into a metric space (usually $\mathbb{R}^n$), and then to use the metric to determine where the simplices are. One simple way to do this is to look at pairwise distance metrics — if $k+1$ points of the data set are all pairwise within some distance d, then connect those $k+1$ points into a k-simplex as in Figure 10. By increasing the distance d, simplices are added but never taken away, satisfying equation 2.1 and resulting in a filtration. This construction is the Vietoris–Rips construction, which appears widely in application[**28**, **29**].

There are a variety of other ways to construct a filtration based on embeddedness in a metric space, including[**17**, **18**, **24**]. When using these types of constructions, the main consideration is the choice of embedding. These constructions are often convenient for visualization because of the ability to embed the data set and resulting simplicial complex in $\mathbb{R}^n$. They may be inconvenient in situations where it is difficult to decide on a distance metric which describes the relationships between data points well.

2.4.2. *Connection-based constructions.* Another way to build a filtration is to use information about the connectedness of a data set. As an abstract example, suppose I have a data set with individual nodes, where nodes become connected to each other in some known order (consider, for example, a flow network, or a temporal network recording friendships between people over time). Then I can create a filtration where individuals are 0-simplices, and I add further simplices as the individuals become connected to each other. When $k+1$ nodes are all pairwise

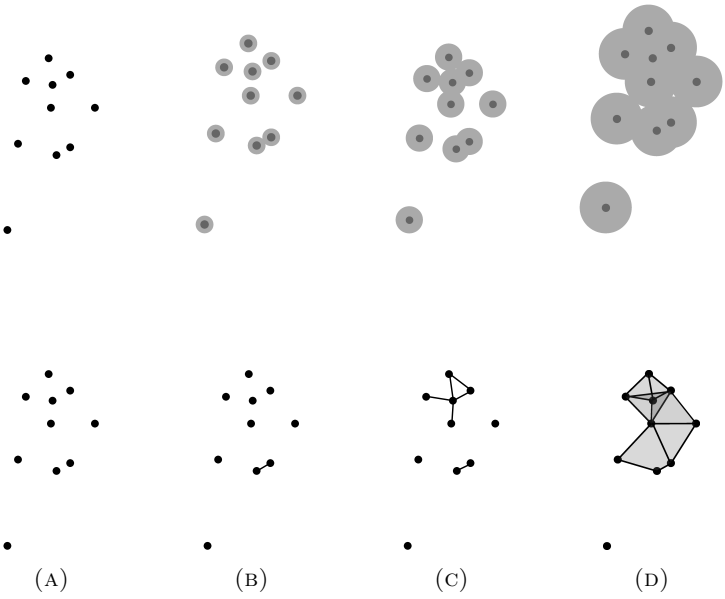


FIGURE 10. An example of how to build a Vietoris–Rips complex. In panel Figure 10a, we display a point cloud above, with associated simplicial complex below. In panels Figures 10b to 10d, we show the point cloud with increasingly large ϵ-balls placed around each point, with the relevant faces between them. Note that in Figure 10d, there are simplices of dimension > 2 in the 5-clique contained at the top part of the image, though these simplices are hard to illustrate.

connected, I add a k-simplex (similar to a Vietoris–Rips complex). This would give me a valid filtration, based on the order of connection.

For an application which uses an idea similar to the one just described, see [**10**]. In this study of precinct networks, one complex (the adjacency complex) is constructed by connecting precincts in order of the strength of their preference for one candidate or another. Precincts with the strongest preferences are added to the simplicial complex first, while precincts with the least strong preferences are added last.

2.4.3. *Function-based constructions.* To generalize, suppose I have a simplicial complex S. Suppose I can define some function $F : S \to \mathbb{R}$ on the simplices of S that satisfies following statement: if $\tau \in S$ is a face of $\sigma \in S$, then $F(\tau) \leq F(\sigma)$. Then I can select some increasing sequence $\epsilon_0 \leq \epsilon_1 \leq \cdots \leq \epsilon_n$ in $\mathbb{R}$ and define a filtration

$$S_0 \subseteq S_1 \subseteq \cdots \subseteq S_n,$$

where

$$S_i = \{\sigma \in S : F(\sigma) \leq \epsilon_i\}.$$

Any way I have of assigning such a function F will give me a valid filtration. Understanding the meaning of topological features in such a filtration relies on knowledge about the simplices of S.

Filtrations are extremely flexible and powerful tools for computing topological features of data. I have already discussed the importance of considering where to put individual simplices, but haven't yet addressed the question of determining steps in a filtration. That is, what difference does it make to put a simplex in S_i rather than S_{i+1}? As long as equation 2.1 is satisfied, does it matter what the differences between the S_i's are? In the next section, I will define a persistent homology and explain how the relationships between individual steps of a filtration can tell us about data.

2.5. Persistent homology. I have now introduced all the tools necessary to define a persistent homology, the primary tool of PH. In Section 2.4, I remarked that given a filtration, it is possible to simultaneously compute the homologies of every step of the filtration. The collection of homologies resulting from such a computation is called the persistent homology of a filtration.

Definition 2.12. Let S be a filtered simplicial complex

$$S_0 \subseteq S_1 \subseteq \cdots \subseteq S_n = S.$$

Then the *m-th persistent homology of S* is the collection of m-th homology groups of each S_i, along with maps

$$H_m(S_0) \to H_m(S_1) \to \cdots \to H_m(S_n),$$

where the maps map every element of $H_m(S_i)$ to some element of $H_m(S_{i+1})$.

For algorithms for computing persistent homologies, see[**22**].

The persistent homologies of a filtration are very powerful because the maps between the individual homology groups $H_m(S_i)$ track homological features through the entire filtration. This means that if I know there is some hole in S_i represented by an element $H_m(S_i)$, I can see where that hole came from in $H_m(S_{i-1})$, and where it goes in $H_m(S_{i+1})$. Crucially, if I observe have some feature $x \in H_m S_i)$, with

$$0 \mapsto x,$$
$$H_m(S_{i-1}) \to H_m(S_i).$$

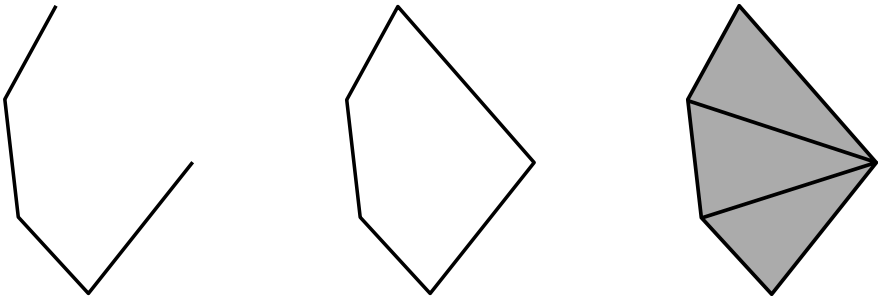
I can conclude that the hole represented by x first appears in S_i. Similarly, if I have

$$x \mapsto 0,$$
$$H_m(S_i) \to H_m(S_{i+1}),$$

then between S_i and S_{i+1}), some number of $m+1$-simplices must have filled in the hole x. These events are generally referred to as the "birth" and "death" of x, and are illustrated in Figure 11.

By tracking the birth and death of each feature, I can characterize changes in the topology as I move through the filtration steps. I can also compute the persistence of each feature using the formula

$$\text{persistence} = \text{death} - \text{birth}.$$



(A) S_0 (B) S_1 (C) S_2

FIGURE 11. Features are born and die. In Figure 11a, there are no features in $H_1(S_0)$. In Figure 11b, there is a single feature in $H_1(S_1)$; we say this feature is born at time 1. Finally, in Figure 11c, the feature from the previous panel has disappeared, since 2-dimensional simplices (colored in grey) have filled the hole. So $H_1(S_2)$ once again has no features, and the feature dies at time 2. Note that if any of the three 2-simplices pictured in Figure 11c were not in S_2, the feature would still exist, and would have no death time.

The persistence of a feature can be interpreted based on the construction of the filtration. For example, if the filtration is based on distance, the persistence of a feature describes the range of distance scales in which the feature appears. That is, features that are more persistent can be seen across a wider range of distance scales. In some applications, this might give a hint that these features are "truer" (if you notice a gap in an image no matter how you squint at it, then you can be sure that gap exists). In other applications, persistence might not say anything about the validity of the features, but instead tell you about the differences in structure at different size scales (imagine zooming in with a microscope, where some features will only visible at small scales, and others only visible at large). Again, the interpretation of the topological information is entirely dependent on the construction of the simplicial filtration, so practitioners should use their judgement to decide what birth, death, and persistence mean in their context.

2.6. Visualizing persistence. There are a large variety of ways to visualize persistence; here, I will introduce two of the most popular, with brief explanations of the strengths of each.

The first visualization I will discuss is the barcode. First introduced in [**12**] , this visualization resembles a horizontal bar graph, with each bar representing one feature. For each feature, a bar is drawn with the left endpoint at the feature's birth time, and the right endpoint at death time. The bars are organized by birth time. Separate barcodes can be drawn for each dimension of the persistent homology. The barcode makes it easy to visually separate features in different dimensions, and the persistence of each feature is clearly visible as the length of the bar. As a result, the eye tends to be drawn to the longest persistence features, emphasizing persistence. For an example of a barcode, see Figure 12.

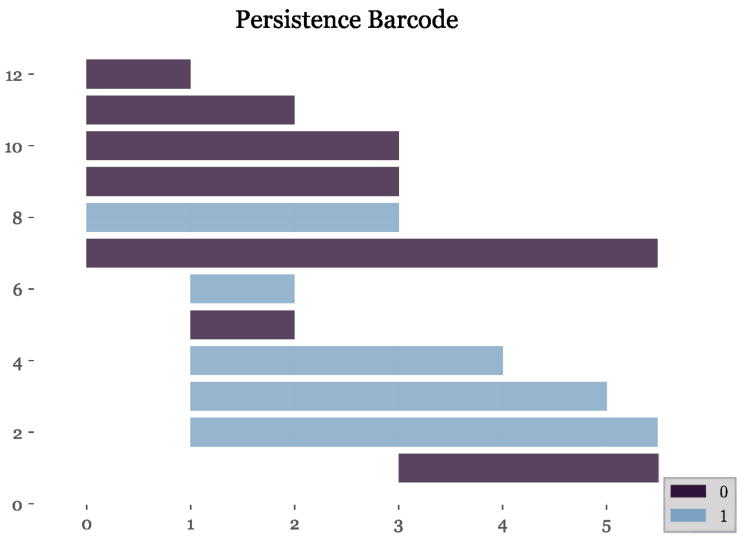


FIGURE 12. A persistence barcode showing both H_0 features and H_1 features. Note that each feature is separately visible, even if the birth and death times coincide with another feature.

Another popular choice of visualization is the persistence diagram (PD). In a PD, features are plotted as points on a scatterplot, with each feature plotted at the coordinate (birth, death). One advantage of PDs is that the distance between PDs can be computed (usually, either bottleneck [**13**] or Wasserstein [**20**] distance is used). This allows practitioners to put PDs into clustering algorithms that require only a pairwise distance matrix for input. Distances between PDs are also fairly well studied and understood [**4**, **19**]. For an example of a PD, see Figure 13.

In addition to the two methods (perhaps the most common) described in this section, there are a variety of other ways of post-processing and visualizing PH, including those discussed in [**1**, **3**]. In application, choice of visualization will depend on what aspects of your PH you wish to emphasize.

3. Tutorial

In this article, I've given a number of (hopefully useful) intuitions for how to think about data as it moves through a PH pipeline. However, the best way to build your own intuition is to practice. In this section, I will walk through a rough version of the thought process I use when setting up a TDA project. If you are so inclined, please feel free to pick your favorite data set and follow along! Other practitioners may approach TDA problems very differently from I do, but if you're not sure where to start, I hope these questions can provide some guidance.

As an example, I'll walk through my own answers to these questions for a hypothetical housing data set. Let's suppose said hypothetical data set contains a map of my neighborhood, with individual addresses, units of housing available at each address, information about whether units are renter or owner occupied, and the monthly rent or mortgage payment.

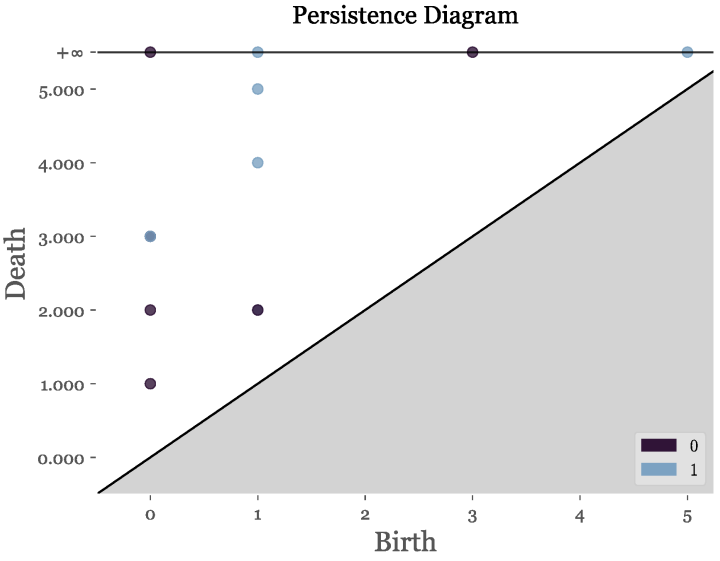


FIGURE 13. A persistence diagram, with the same features as in Figure 12. Here, features with the same birth and death coordinates are plotted on top of each other.

3.1. Setup.

EXERCISE 3.1. Is topological analysis useful for this data set?

- Does the shape of this data provide any useful information?
- When I consider my data set, how much does the geometry matter? Can I squish my data around without losing important information?
- Does connectedness or lack thereof of the data tell me anything interesting? Am I looking for holes, gaps, or obstructions?
- Do I have some sense that the data exists in a space other than the embedding into whatever space the data is currently formatted in?

If the answers to some of these questions are yes, topology might be a reasonable candidate for analyzing the data.

EXAMPLE 3.2. In my hypothetical housing data set, I might observe that the shape of the data is naturally embedded into a 2d landscape. I might also, for example, be interested in affordable housing availability, and areas in which there are noticeable gaps in affordable housing. In addition to information about price, my data set contains renter-occupied vs. owner-occupied categorizations, which might be used to study the data by embedding it in a space that reflects that difference (which may not be reflected on a map).

EXERCISE 3.3. In this data set, what is "nearness"?

- Is this data a point cloud? If yes, how do I interpret individual points? Are they distinct individuals, or are they sampled from something continuous?
- Does the data set come equipped with some existing notion of distance? Is that notion interesting, or irrelevant to the question at hand?
- Which points seem "nearest" each other, and why? Is the "nearness" geometric, or is it based on something else?

These questions help me think about what kind of space my data is living in. Sometimes, that space is some copy of $\mathbb{R}^n$, and the data arrived already embedded in it. In these cases, my interest may be in trying to understand where the gaps in the data are, if the embedded space is sufficiently high-dimensional that I can't look at it directly. Sometimes, the space I'm interested in is somewhat different from the space in which the data is already embedded. When this happens, my topological construction might be difficult to visualize and intuit about, and my interest is in seeing what turns up when I apply topological tools to this other space.

EXAMPLE 3.4. For my hypothetical housing data set, my data can be described by points in the form of addresses. Each point would represent a single address (building). I could also further break down buildings into units, allowing each point to represent a single available unit. For notions of distance, geographic distance between addresses provides one option, which might tell me how far I need to go from a certain block to find available units. I could also think of units as being related by their price, or by their occupancy status (e.g., renter occupied units could be more similar to each other than owner occupied ones).

3.2. Beginning construction.

EXERCISE 3.5. Based on this data set, construct something amenable to topological analysis.

- Is the data a point cloud equipped with a notion of distance? If yes, can I use an out-of-the-box distance-based construction?
- If the data is not equipped with a notion of distance, can I assign one to it that allows me to use a distance-based construction?
- If out-of-the-box isn't an option, what do I want to be my 0-simplices? My 1-simplices? My k-simplices? In what order should my simplices enter the simplicial complex?
- Does it make more sense to replaces triangles with cubes, or with some other type of combinatorial complex?

As discussed in Section 2.2, there are relatively few requirements for constructing a simplicial complex. Out-of-the-box options are very convenient when usable, but if you feel you need something more creative to study the questions you're interested in, this is a good place to spend some time thinking.

EXAMPLE 3.6. In the hypothetical housing data set, I could use an out-of-the-box distance-based construction with all addresses. The resulting features would tell me where there are no available units. If I were to use a distance-based construction but only look at renter-occupied units, I would find areas with no available rental units. If I were to look only at affordable units, I would find gaps in affordable housing availability.

If I instead wanted to construct an ad-hoc complex which started with only affordable units and added units in order of decreasing affordability to the filtration, I would learn about areas which lack affordable housing, and where less affordable housing exists instead.

If I constructed a complex that started with owner-occupied units only and added in rental units to a later step of the filtration, I could find areas where rental units fill gaps in owner-occupied units. This might be interesting if I were studying a neighborhood in which rental units are treated as transitional housing on the way to home ownership, or perhaps in a neighborhood where new rental buildings are replacing owner-occupied housing.

Each of these constructions would tell me something slightly different, and which choice suited my purposes would depend on what I was attempting to study.

3.3. Computations.

Exercise 3.7. Compute a persistent homology. What does it say?

- What were the simplices in my complex? Given that, what are the homological features?
- Are the higher dimensional homological features interesting? Can I interpret them in some easily understandable way?
- What types of connectedness (or lack of connectedness) have I uncovered in my data? Did I learn anything I couldn't have learned from other types of analysis?
- What does persistence mean in this data set? How can I use persistence to filter or further understand the homological features?

In many regards, this step is relatively simple. Most of the interpretation work has already been done in setting up a complex, and the answers to the question above should follow directly from construction choices.

Exercise 3.8. How should the results of the homology computations be communicated?

- Is this the final step, or am I using the homological information as a topological summary to be input into further pipelines?
- What types of visualizations will help me get across the interesting homological features most quickly?
- Without explaining the techniques I used, how can I describe to other researchers the results of my computations?

As a researcher, it's vitally important to be able to communicate results quickly, and to a broad audience of other researchers who may be interested. When working with data, that audience will frequently include people who come from nonmathematical or nontopological backgrounds.

While the outline given in this section is far from the only way to approach doing PH on a fresh data set, I hope that new practitioners might find some helpful guidance in this document.

To see some examples of current research that have made various choices of construction and interpretation to study a social phenomenon, see [**6**, **9**, **10**, **15**, **16**].

4. Conclusion

Topological data analysis is an incredibly flexible tool for studying almost any type of data. When applying methods from TDA to a data set, practitioners have a wide array of choices available to them, all of which will affect the results of their analysis. My hope is that this document can help provide some insight into the differences between various choices, so that researchers interested in TDA can better understand how to apply TDA to their own data and get interpretable results in an appropriate context for their research.

References

[1] Henry Adams, Tegan Emerson, Michael Kirby, Rachel Neville, Chris Peterson, Patrick Shipman, Sofya Chepushtanova, Eric Hanson, Francis Motta, and Lori Ziegelmeier, *Persistence images: a stable vector representation of persistent homology*, J. Mach. Learn. Res. **18** (2017), Paper No. 8, 35. MR3625712

[2] Yair Amichai-Hamburger, Mila Kingsbury, and Barry H. Schneider, *Friendship: An old concept with a new meaning?*, Computers in Human Behavior **29** (2013), no. 1, 33–39.

[3] Peter Bubenik, *Statistical topological data analysis using persistence landscapes*, J. Mach. Learn. Res. **16** (2015), 77–102. MR3317230

[4] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer, *Stability of persistence diagrams*, Discrete Comput. Geom. **37** (2007), no. 1, 103–120, DOI 10.1007/s00454-006-1276-5. MR2279866

[5] Vin de Silva and Robert Ghrist, *Coverage in sensor networks via persistent homology*, Algebr. Geom. Topol. **7** (2007), 339–358, DOI 10.2140/agt.2007.7.339. MR2308949

[6] Moon Duchin, Tom Needham, and Thomas Weighill, *The (homological) persistence of gerrymandering*, Found. Data Sci. **4** (2022), no. 4, 581–622, DOI 10.3934/fods.2021007. MR4622907

[7] Herbert Edelsbrunner and John L. Harer, *Computational topology: An introduction*, American Mathematical Society, Providence, RI, 2010, DOI 10.1090/mbk/069. MR2572029

[8] Brittany Terese Fasy, Jisu Kim, Fabrizio Lecci, and Clément Maria, *Introduction to the R package TDA*, 2014.

[9] Michelle Feng and Mason A. Porter, *Spatial applications of topological data analysis: Cities, snowflakes, random structures, and spiders spinning under the influence*, Physical Review Research **2** (2020), no. 3, 033426.

[10] Michelle Feng and Mason A. Porter, *Persistent homology of geospatial data: a case study with voting*, SIAM Rev. **63** (2021), no. 1, 67–99, DOI 10.1137/19M1241519. MR4209654

[11] Leon Festinger, Stanley Schachter, and Kurt Back, *Social pressures in informal groups: A study of human factors in housing*, 1950.

[12] Robert Ghrist, *Barcodes: the persistent topology of data*, Bull. Amer. Math. Soc. (N.S.) **45** (2008), no. 1, 61–75, DOI 10.1090/S0273-0979-07-01191-3. MR2358377

[13] François Godi, *Bottleneck distance*, GUDHI User and Reference Manual, GUDHI Editorial Board, 2016.

[14] Allen Hatcher, *Algebraic topology*, Cambridge University Press, Cambridge, 2002. MR1867354

[15] Abigail Hickok, Benjamin Jarman, Michael Johnson, Jiajie Luo, and Mason A. Porter, *Persistent homology for resource coverage: a case study of access to polling sites*, SIAM Rev. **66** (2024), no. 3, 481–500, DOI 10.1137/22M150410X. MR4783075

[16] C. J. Carstens and K. J. Horadam, *Persistent homology of collaboration networks*, Math. Probl. Eng., posted on 2013, Art. ID 815035, 7, DOI 10.1155/2013/815035. MR3062959

[17] Clément Jamin, *Tangential complex*, GUDHI User and Reference Manual, GUDHI Editorial Board, 2016.

[18] Siargey Kachanovich, *Witness complex*, GUDHI User and Reference Manual, GUDHI Editorial Board, 2015.

[19] Roland Kwitt, Stefan Huber, Marc Niethammer, Weili Lin, and Ulrich Bauer, *Statistical topological data analysis - A kernel perspective*, Proceedings of the 28th International Conference on Neural Information Processing Systems, NIPS'15, MIT Press, Cambridge, MA, 2015, pp. 3070–3078.
[20] Theo Lacombe, *Wasserstein distance*, GUDHI User and Reference Manual, GUDHI Editorial Board, 2020.
[21] Chris Tralie and Nathaniel Saul, *Scikit-TDA: Topological data analysis for python*, 2019.
[22] Nina Otter, Mason A. Porter, Ulrike Tillmann, Peter Grindrod, and Heather A. Harrington, *A roadmap for the computation of persistent homology*, EPJ Data Science **6** (2017), no. 1, 17.
[23] Pratyush Pranav, Herbert Edelsbrunner, Rien van de Weygaert, Gert Vegter, Michael Kerber, Bernard J. T. Jones, and Mathijs Wintraecken, *The topology of the cosmic web in terms of persistent Betti numbers*, Monthly Notices of the Royal Astronomical Society **465** (2017), no. 4, 4281–4310.
[24] Vincent Rouvreau, *Alpha complex*, GUDHI User and Reference Manual, GUDHI Editorial Board, 2015.
[25] Andrew Tausz, Mikael Vejdemo-Johansson, and Henry Adams, *JavaPlex: A research software package for persistent (co)homology*, Proceedings of ICMS 2014 (Han Hong and Chee Yap, eds.), Lecture Notes in Computer Science 8592, 2014, pp. 129–136.
[26] The GUDHI Project, *GUDHI user and reference manual*, GUDHI Editorial Board, 2015.
[27] Chad M. Topaz, Lori Ziegelmeier, and Tom Halverson, *Topological data analysis of biological aggregation models*, PLoS ONE **10** (2015), no. 5, e0126383.
[28] L. Vietoris, *Über den höheren Zusammenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen* (German), Math. Ann. **97** (1927), no. 1, 454–472, DOI 10.1007/BF01447877. MR1512371
[29] Afra Zomorodian, *Fast construction of the Vietoris-Rips complex*, Computers & Graphics **34** (2010), no. 3, 263–271.
[30] Afra Zomorodian and Gunnar Carlsson, *Computing persistent homology*, Discrete Comput. Geom. **33** (2005), no. 2, 249–274, DOI 10.1007/s00454-004-1146-y. MR2121296

Computing and Mathematical Sciences, California Institute of Technology, Pasadena, California 91125

80 **Heather Z. Brooks, Michelle Feng, Mason A. Porter, and Alexandria Volkening, Editors,** Mathematical and Computational Methods for Complex Social Systems, 2025

79 **Maria Trnkova and Andrew Yarmola, Editors,** 3D Printing in Mathematics, 2023

78 **François Delarue, Editor,** Mean Field Games, 2021

77 **Pablo A. Parrilo and Rekha R. Thomas, Editors,** Sum of Squares: Theory and Applications, 2020

76 **Keenan Crane, Editor,** An Excursion Through Discrete Differential Geometry, 2020

75 **Michael Damron, Firas Rassoul-Agha, and Timo Seppäläinen, Editors,** Random Growth Models, 2018

74 **Jan Bouwe van den Berg and Jean-Philippe Lessard, Editors,** Rigorous Numerics in Dynamics, 2018

73 **Kasso A. Okoudjou, Editor,** Finite Frame Theory, 2016

72 **Van H. Vu, Editor,** Modern Aspects of Random Matrix Theory, 2014

71 **Samson Abramsky and Michael Mislove, Editors,** Mathematical Foundations of Information Flow, 2012

70 **Afra Zomorodian, Editor,** Advances in Applied and Computational Topology, 2012

69 **Karl Sigmund, Editor,** Evolutionary Game Dynamics, 2011

68 **Samuel J. Lomonaco, Jr., Editor,** Quantum Information Science and Its Contributions to Mathematics, 2010

67 **Eitan Tadmor, Jian-Guo Liu, and Athanasios E. Tzavaras, Editors,** Hyperbolic Problems: Theory, Numerics and Applications, 2009

66 **Dorothy Buck and Erica Flapan, Editors,** Applications of Knot Theory, 2009

65 **L. L. Bonilla, A. Carpio, J. M. Vega, and S. Venakides, Editors,** Recent Advances in Nonlinear Partial Differential Equations and Applications, 2007

64 **Reinhard C. Laubenbacher, Editor,** Modeling and Simulation of Biological Networks, 2007

63 **Gestur Ólafsson and Eric Todd Quinto, Editors,** The Radon Transform, Inverse Problems, and Tomography, 2006

62 **Paul Garrett and Daniel Lieman, Editors,** Public-Key Cryptography, 2005

61 **Serkan Hoşten, Jon Lee, and Rekha R. Thomas, Editors,** Trends in Optimization, 2004

60 **Susan G. Williams, Editor,** Symbolic Dynamics and its Applications, 2004

59 **James Sneyd, Editor,** An Introduction to Mathematical Modeling in Physiology, Cell Biology, and Immunology, 2002

58 **Samuel J. Lomonaco, Jr., Editor,** Quantum Computation, 2002

57 **David C. Heath and Glen Swindle, Editors,** Introduction to Mathematical Finance, 1999

56 **Jane Cronin and Robert E. O'Malley, Jr., Editors,** Analyzing Multiscale Phenomena Using Singular Perturbation Methods, 1999

55 **Frederick Hoffman, Editor,** Mathematical Aspects of Artificial Intelligence, 1998

54 **Renato Spigler and Stephanos Venakides, Editors,** Recent Advances in Partial Differential Equations, Venice 1996, 1998

53 **David A. Cox and Bernd Sturmfels, Editors,** Applications of Computational Algebraic Geometry, 1998

52 **V. Mandrekar and P. R.Masani, Editors,** Proceedings of the Norbert Wiener Centenary Congress, 1994, 1997

51 **Louis H. Kauffman, Editor,** The Interface of Knots and Physics, 1996

For a complete list of titles in this series, visit the
AMS Bookstore at **www.ams.org/bookstore/psapmseries/**.

PSAPM 80

Complex Social Systems • Brooks et al., Editors

The spread of memes and misinformation on social media, political redistricting, gentrification in urban communities, pedestrian movement in crowds, and the dynamics of voters are among the many social phenomena that researchers investigate in the field of complex systems. In the study of complex social systems, there is often also societal relevance to improving our understanding of how individuals interact with each other and their environment, giving rise to collective group dynamics.

The mathematical and computational study of complex social systems relies on and motivates the development of methods in many topics, including mathematical modeling, data analysis, network science, and topology and geometry. This volume is a collection of diverse articles about complex social systems. This collection includes both (1) survey and tutorial articles that introduce complex social systems and methods to study them and (2) manuscripts with original research that highlight a variety of mathematical areas and applications.

This book introduces the study of complex social systems to a broad mathematical audience. It will particularly appeal to people who are interested in applied mathematics.

ISBN 978-1-4704-6459-2
9 781470 464592
PSAPM/80

AMS
www.ams.org