Title

Topics in Persistent Homology and Complex Social Systems

Author

Luo, Jiajie

Publication Date

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Topics in Persistent Homology and Complex Social Systems

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Mathematics

by

Jiajie Luo

2024

ABSTRACT OF THE DISSERTATION

Topics in Persistent Homology and Complex Social Systems

by

Jiajie Luo
Doctor of Philosophy in Mathematics
University of California, Los Angeles, 2024
Professor Mason Alexander Porter, Chair

The field of topological data analysis (TDA) uses tools from algebraic topology to capture quantitative structural properties in a data set. Perhaps the most popular tool in TDA is persistent homology (PH), which leverages homology theory to provide insights into the structure of data by quantifying "holes" across different scales. PH has been applied to many domains, including neuroscience, materials science, image processing, and social systems. In this thesis, we study PH in both theory and application.

On the theory side, we build upon the foundations of PH by studying persistence modules (which are algebraic objects that are fundamental to PH) and their interval decompositions. The existence of interval decompositions of a persistence module underlies much of PH. Interval decompositions are guaranteed to exist for persistence modules with coefficients in a field. However, interval decompositions may not exist in more general settings, such as when the coefficients are not in a field. We prove necessary and sufficient conditions for persistence modules with coefficients in a principal ideal domains to have an interval decomposition. We also formulate an algorithm to compute an interval decomposition when one exists.

We then use PH to quantify and assess the accessibility of resources in a geographic area. This allows us to identify regions that have poor resource access. Our work focuses on the accessibility of polling sites across six geographic regions (five cities and Los Angeles County). We adapt traditional approaches in PH to incorporate important factors of polling-site accessibility, including travel times to and from polling sites and waiting times at those sites.

Finally, we discuss the modelling of opinion dynamics on networks. We consider bounded-confidence models (BCMs), which are models of opinion dynamics in which agents are receptive only to agents whose opinions are sufficiently similar (i.e., within a confidence bound). We formulate and analyze two BCMs with adaptive confidence bounds; our BCMs generalize the Hegselmann–Krause and Deffuant–Weisbuch models. Using mathematical analysis and numerical simulations, we demonstrate that our adaptive BCMs exhibit quantitatively and qualitatively different behavior than the associated baseline (i.e., nonadaptive) HW and DW BCMs. This includes fewer major opinion clusters, longer convergence times, and adjacent nodes that converge to the same opinion but are not receptive to each other.

The dissertation of Jiajie Luo is approved.

Christopher R. Anderson

Michael Anthony Hill

Deanna M. Needell

Mason Alexander Porter, Committee Chair

University of California, Los Angeles

2024

To my family.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I lift up my eyes to the hills —
where does my help come from?
My help comes from the LORD,
the Maker of heaven and earth. (Psalm 121:1-2)

And whatever you do, whether in word or deed, do it all in the name of the Lord
Jesus, giving thanks to God the Father through him. (Colossians 3:17)

As I see it, the acknowledgment section is the most important part of a thesis — even more important, might I say, than the "main" contents themselves. It allows us to give thanks and recognize those who have come alongside, while also acknowledging that none of us got where we are now on our own. While attempting to acknowledge everyone who deserves acknowledgment is a probably not feasible, I think the least that I can do is to try my best anyway. To this end, I will make no attempt at brevity.

Starting off, I want to acknowledge and thank my advisor, Mason Porter, whose mentorship played a significant role in my development as a mathematician and a researcher. Throughout my time as your student, you have always been kind, supportive, flexible, and understanding. Not only have you given me countless pointers for references and research directions, you have also given me the space to pursue my own curiosity. Whether it's providing feedback on drafts, forwarding relevant papers/events, or meeting to discuss various things, I appreciate the time, energy, intention, and dedication that you have provided to me throughout the years. I have grown a lot as a mathematician and a researcher during my time at UCLA; much of my growth is attributed to your supervision.

I also want to acknowledge and thank Gregory Henselman-Petrusek, who was my mentor during my internship at the Pacific Northwest National Laboratory, and later, my collabo-

rator in [LH23]. You were always kind and supportive of me, both within the internship and beyond it. You have also taught me a lot about the theoretical ends of persistent homology. I really appreciate all our meetings (the majority of which were after my internship), in which we discussed research, careers, and various other topics.

I want to acknowledge and thank my committee members Chris Anderson, Michael (Mike) Hill, and Deanna Needell. Whether it's through classes you've taught, personal discussions we've had, or a combination of the two, each of you has taught me something that has made me a better mathematician and researcher. More relevantly to this thesis, I want to thank each of you for your input and thoughts on the contents of my research.

I would like to highlight and thank my co-authors and collaborators, who have played significant roles in helping expand the scope of my research. One of the joys of research is the friendships that spark from collaborations such as these. Chapter 3 is adapted from [LH23], which I co-authored with Gregory Henselman-Petrusek; I thank Mason A. Porter and Benjamin Spitz for helpful feedback and discussion. Chapter 4 is adapted from [HJJ23], which I co-authored with Abigail Hickok, Benjamin Jarman, Michael Johnson, and Mason A. Porter; I thank Chris Anderson, Lyndie Chiou, Alex Sherman, and Renata Turkes for helpful comments and discussion. Chapter 5 is adapted from [LLP23], which I co-authored with Grace J. Li and Mason A. Porter; I thank Weiqi Chu, Gillian Grindstaff, Abigail Hickok, and the Networks Journal Club for helpful comments and discussion, as well as Carmela Bernardo, Damon Centola, PJ Lamberson, and Jim Moody for pointers to helpful references. I also want to acknowledge those with whom I have ongoing collaboration that are not in the set of co-authors that I listed above. In no particular order, this includes Sarah Tymochko, Gillian Grindstaff, Weiqi Chu, Yang Yang, Joseph Tien, Anna Nelson, Carlos Martinez Mori, Giulia De Pasquale, Fabiana Ferracina, William Thompson, and Rebecca Hardenbrook.

I acknowledge funding from NSF grant number 1829071, NSF grant number 2136090, and NSF grant number 1922952 through the Algorithms for Threat Detection (ATD) program.

I would like to highlight and acknowledge members of my research group during my time at UCLA. This includes (in no specified order) Abigail Hickok, Kaiyan Peng, Grace Li, Michael Johnson, Sarah Tymochko, Gill Grindstaff, Weiqi Chu, Theo Faust, Filippo Riscica Lizzio, Christine Kling, James Raj, Leah Keating, Sanjukta Krishnagopal, Heather Zinn Brooks, and Phil Chodrow. I've had various helpful and fun conversations with each of you and appreciate your friendship. I also want to make a few extended acknowledgements. Abby, you're an incredibly gifted and talented mathematician. At the same time, you've always been kind, friendly, humble, and down-to-earth. There's a ton that I've learned from you (both through collaborations and beyond) and I look forward to seeing what the future holds for you. Grace, you're easily one of the people who have helped me the most here at UCLA. From entering the days of the basic bootcamp to the present moment, you've been a reliable friend who has been willing to lend a hand in tangible and often nontrivial ways. It's always fun to talk with you, whether it's research-related or anything else.

Beyond those who are directly related with my research, there are quite a number of people I want to name acknowledge here at UCLA. I will begin with those related to qualifying exams (quals). I want to acknowledge my basic-exam study crew, which included Jason Brown, Cecelia Higgins, Grace Li, Andrew Sack, and Yotam Yaniv. Even though the circumstance (we all failed the basic exam coming into graduate school) that brought us together was nontrivially demoralizing, it's encouraging to see that each of us made it through our quals and proceeded to find success in our respective research areas. I also want to acknowledge and thank Bon-Soon Lin, who went above and beyond to help us in our journey with the basic exam. Beyond the basic exam, there are many who went before and alongside me for the other two quals. This includes Abby Hickok and Joel Barnett, who studied with me for the numerical analysis exam; Ben Bowman, who showed me various tricks for the numerical analysis exam; Roman Krutowski, who studied with me for the geometry/topology exam; and Eilon Reisin-Tzur, Ben Spitz, Talon Stark, Matthew Gherman, and Jason Schuchardt, whom I consulted countless times while studying for the geometry/topology exam.

Beyond qualifying exams, I've had the pleasure of meeting and getting to know many great and wonderful people in the mathematics department. This includes a whole bunch of people, including (in no particular order) Dominic Yang, Ben Bowman, Ben Spitz, Talon Stark, Jason Brown, Yotam Yaniv, Jason Schuchardt, Erin George, Ben Johnsrude, Nikita Gladkov, Ezra Thompson, Mohit Bansil, Advika Rajapakse, James Chapman, Roman Krutowski, Matthew Gherman, Bon-Soon Lin, Adam Lott, Eilon Reisin-Tzur, David Soukup, James Leng, Joe Breen, Grace Li, Abby Hickok, Joel Barnett, Joyce Chew, Louis Esser, Zerrin Vural, Hong Kiat Tan, Zi Li Lim, Yizhou Chen, Hedi Xia, Lisang Ding, Haoyang Lyu, Kaiyan Peng, Yushan Han, Joshua Enwright, Zach Baugher, Jas Singh, Rushil Raghavan, Cecelia Higgins, Andrew Sack, Hagen Papenburg, Thierry Laurens, Alec Leng, Evan Davis, Clark Lyons, Victoria Quijano, Rohan Joshi, Nick Boeschert, Nicholas Hu, Raymond Chu, Michael Johnson, Ben Jarman, Matt Kowalski, Jean-Michel Maldague, William Swartworth, Jack Luong, Griffin Pinney, Patrick Hiatt, Jacob Swenberg, Weiyi Liu, Siting Liu, Tingwei Meng, Lingyun Ding, Yang Liu, Sarah Tymochko, Gillian Grindstaff, Eric Hester, Weiqi Chu, Frederick Vu, Bohyun Kim, Kyung Ha, Brian Shin, Brian Lawrence, and probably a bunch of people that I forgot to name (sorry!). Here are a few extended acknowledgements. Dominic, I've always enjoyed our excursions, which includes getting hot chicken, watching me fail at playing your Switch games, and much more. You've always been a reliable and caring friend, which I deeply value. Ben (Spitz) and Talon, we've been friends since our undergraduate days in the College of Creative Studies, and I'm glad that we've been together here at UCLA. There are so many meaningful memories (including the many times that you would visit me at Jesus Burgers on a Friday night) and thoughtful conversations that I've had with the two of you. We've all come a long way, and I hope to keep in touch with you guys in the years ahead. To all the Chinese people above, thank you for letting me speak/practice Chinese with you.

In addition to the folks at UCLA, I want to acknowledge and thank my professors at UC Santa Barbara (UCSB). Starting off, I want to thank Paddy Bartlett, who, as my (first)

advisor in the College of Creative Studies (CCS), played a central role in the building of my mathematical foundations. Thank you for your kindness, patience, and intentionality in building us up as mathematicians; your work has truly gone a long way. In the same vein, I want to thank Maribel Bueno Cachadina, who was (and still is) one of the main mathematics advisors in CCS. Similar to Paddy, you played a foundational role in my mathematical journey (especially in linear algebra), which has helped me tremendously throughout the years. I want to thank Bill Jacob (who is sadly no longer with us), who was my first research advisor and taught me a great deal of algebra. Bill supervised my master's thesis (on abstract Witt rings) back when I was trying to become an algebraist. The way that you taught me algebra has played an important role in my mathematical journey, not only in research but also in teaching. I also want to thank Jeffrey Stopple, who took over as my advisor in CCS after Paddy left for New Zealand. You gave me a lot of good perspective in my mathematical journey, especially pertaining to graduate-school applications. Other professors that I want to acknowledge and thank are Jon McCammond, Darren Long, Daryl Cooper, Peter Garfield, Birge Huisgen-Zimmermann, Ken Goodearl, Stephen Bigelow, Denis Labutin, Karel Casteels, Hector Ceniceros, Chuck Akemann, and Yitang Zhang. Besides professors, I also want to acknowledge Sara Sterphone (who is sadly no longer with us) and Medina Price, along with the staff in both CCS and the UCSB mathematics department, for all of the hard work you do in helping students in your respective programs. Sara and Medina, I appreciate the help that both of you have extended to me, which on numerous occasions, spared me from a lot of stress and annoyance. Both of you had/have (very) difficult and challenging jobs, in which you have performed spectacularly.

I also want to acknowledge friends in both CCS and the mathematics department at UCSB. From CCS, this includes Nicholas Geis, Richard Carini, Landon Settle, Kayla Wright, Chloe Avery, Ben Spitz, Megan Franke, Talon Stark, Kayleigh Adams, Xander Song, and probably other that I forgot to list. From the mathematics department, this includes Garo Sarajian, Christian Bueno, David Wen, Nathan Schley, Qingjing Chen, Aaron Bagheri, Danning Lu,

Junrong Yan, Qiu Yang, Kayla Wright, Micah Pedrick, Seth Althauser, Greg McGrath, Marcos Reyes, Drew Balkin, Yao Xuan, Tom Shifley, Cole Hawkins, Matthew Varble, Andres Martin Rodriguez, Rafael Lainez Reyes, Ethan Robinett, Vijay Higgins, Kelly Pham, Sarafina Ford, Nico Gonzalez, Malik Tuerkoen, Chris Gorman, and probably a ton of people who I (once again) managed to leave out. Here are some extended acknowledgements. Greg, I think it's kind of funny how we first connected through Old School RuneScape (OSRS), which somehow paved the road to you being one of my closest friends in the math department. I've really appreciated and enjoyed the various conversations that we've had, the hospitality you've extended towards me in letting me stay with you and Sarah, and random stuff we did together in OSRS. You've really been a true friend, and I look forward to seeing your next steps. Tom, it was super encouraging having another believer in the department, as well as seeing the ways in which your faith tangibly manifests in your life. I've really appreciated the many conversations we've had, whether it's about math, our relationship with Jesus, or anything else. Megan, in the same vein, I'm thankful that we got to know each other while in CCS. I appreciate your friendship and the fact that there was someone early on in my academic journey with whom I could connect on both faith and mathematics.

While my acknowledgements so far more or less cover those that I've gotten to know through mathematics, they would be far from complete if I were to stop here. Throughout my journey in mathematics, there were many who, while not directly involved in mathematics, walked alongside me throughout the process.

While living in Los Angeles (LA), I've had the good fortune of having some very good roommates. Surui (Tony) Sun and Paul Boone, it's been a pleasure living with the two of you (albeit in different times). Having good roommates is not something to be taken for granted. I'm thankful for our time together.

During my time in LA, I was a part of Epicentre West LA, a non-denominational church that meets in the west side of LA. This church has played a pivotal role in helping me stay

grounded within faith-based community and, in extension, helped strengthen my relationship with God. In no particular order, I want to highlight and acknowledge Steph and Alex Chao, Tiffany and James Yu, Quen and Kathryn Cheng, Freddie Wang and Tiffany Cheuk, Chris Tang and Jackie Cornejo, Bing Ling and Stella Chang, Alex and Leslie Kearsley, Yam and Shirley Seki, Annie Ruan and Lionel Zhang, Brianna Newitt, Sean Chen, Amelia Ino, Erica Drost, Natalie Wing, Lillian Fan, Jess Chee, Grace Gnasigamany, Ricky Medina, Eden Park, Brandon Yan, Stephanie Lim, Anna Reed, Luke Scorziell, and probably a bunch of people that I managed to leave out (yet again). As before, I will give a few extended acknowledgements. Steph (Chao) and Alex, both of you have shown a great deal of love towards me during the time that I've known you. Both of you, time and again, have spent the time to hear how I'm doing, pray for/with me, encourage me, and reach out to me. You've both been incredibly intentional in both words and actions. Tiffany (Yu) and James, I really appreciate both of you for how intentional you've been with me. It's always fun to talk to either of you. You're both really good listeners and it's abundantly clear to me that you care. Quen and Kathryn, it's been a joy getting to know the two of you (albeit in fairly separate settings) and the journey that God has taken you on. It's really amazing and inspiring to see the ways in which your love for God and people manifests in your lives. Freddie and Tiffany (Cheuk), both of you have been really intentional with me (each in your separate ways). I really appreciate the numerous occasions in which you have checked in on me. It's a lot of fun hanging out and hearing the many stories that you tell. Chris and Jackie, it's always kind of funny for me to see how different the two of you are, and yet you two make such a good team. It's been great getting to know the two of you through life group, in which I really appreciate the hospitality that the two of you have extended. Bing and Stella, I enjoy the sincerity that the two of you have, along with the unique sense of humor that you bring. It's really cool to see your trust in God throughout your walk with him; I appreciate the various testimonies that you've given. Bri, I think it's rather fitting how you were my one of the first people I met at Epicentre. You were one of my first friends

at church, and now you're one of my closest friends from church. I really appreciate the numerous conversations we've had, including the things that God has put on your heart. It's incredibly inspiring for me to see your heart for people, not only in words, but also in action and lifestyle. Sean, I feel like if we met at the wrong time in our walks with God, there would have been no chance we would've been friends. Fortunately, we didn't meet at such a time. And now you're definitely one of my closest friends from church. I really love the heart that you have to seek after God and the ways in which you've stepped out to further your relationship with him. I really appreciate the perspectives (many of which are nontrivially different from mine) that you bring to the table and the ways in which you listen to understand. Erica, you've been a really caring friend, and I really appreciate how quick you are to listen. It's evident to me the depths of relationship that you have with God and the ways in which that translates into your daily life. One way or another, I've always felt encouraged by our conversations. Stephanie (Lim), you've been an incredible friend to me and I appreciate all that you've done for me over the last few years. From the late-night sessions after life group to our various after-hour chats in your car, our conversations have always been fruitful and comforting. I especially appreciate your vulnerability in our conversations. Amelia, even though I always joke about how we're opposites of each other, I've come to appreciate the various things that make you who you are. You carry God's presence in really unique ways, and in all the things that make us different, it's encouraging to see the same sort of intimacy that we have with him. Brandon, I really appreciate the generosity and hospitality that you've extended towards me; it's really cool seeing you work those values into your life. I also appreciate the cultural perspectives that you bring and letting me speak/practice Chinese with you. Grace, despite the fact that you have a knack for asking me questions that I usually have no hope to answer in a given time frame, I appreciate the numerous instances in which you checked in on me. I've always enjoyed our conversations when we do get the opportunity, in which it becomes clear the ways that you carry God's heart for people.

I also want to highlight a few non-Epicentre friends. John Lidfors, you were one of my first friends in LA, and I'm glad we've kept in touch. It's been really cool to see you pursue the things that God has put on your heart, and I look forward to what he has in store for you. James and Aissa Craig, I've gotten so much out of your friendship, especially with the intentionality that you've shown. I especially appreciate the many vulnerable conversations that we've shared. David Chabra, I love the heart you have for God and the overall maturity that you possess. Even though I was supposed to be mentoring you, I feel like there are so many ways in which I've learned and felt uplifted by you.

During my time in Santa Barbara, I was a part of Isla Vista Church (IVC). This was the first church in which I've felt at home and the first faith-based community that truly felt like family. Rather than listing a block of people and then doing extended acknowledgements for a small subset, I will exclusively do extended acknowledgements. Starting off, I want to acknowledge and thank Jason and Holly Lomelino, who pastor IVC. If I were to say that the two of you spoke a ton into my life, that would be quite the understatement. I have been encouraged, time and time again, by the faith and trust that both of you have in God and seeing the fruit that comes from it. The two of you love so well and have cultivated a church family that inspires faith and intimacy with Jesus that goes far beyond the city in which you operate. Beyond church and ministry, the way in which the two of you raised your family really models what a house built upon Jesus can look like. I deeply appreciate all of the love, care, and understanding you've extended towards me. Sam and Steph Kim, your friendship throughout the years has meant a lot to me. Sam, you've been one of my most consistent friends ever since I met you. You've poured into me a ton and have been generous to me like no other. You model the father's heart so well in the myriad ways that you care for people. Steph, it's been a joy getting to know you the past few years. It's really inspiring to see the heart you have for people, especially in the work that you do. I look forward to all that God has in store for the two of you! Demis and Azita John, both of you have been incredibly warm and hospitable, both to me and everyone around. The two of you

really exemplify God's heart for people in the way that you live and all that you do. Demis, it's been really encouraging to have someone like you who both understands the context of academia and loves Jesus. Mike and Dia Becchio, I've been encouraged by you two countless times, both in the conversations we share and the way you live. I really appreciate the raw and vulnerable conversations we've had and the testimonies of God's work in and through your family. Ryan Kelleher, I've really appreciated your steadfast friendship throughout the years. We've walked with each other through quite a lot, and it's encouraging to see all that God has done in and through you. You've blessed me a lot in both words and actions.

While the following people were/are not actually a part of IVC, I would like to acknowledge them as well. Josh and Summer Ortiz, I've really valued our long-lasting friendship. Josh, there are so many ways in which God has blessed me through you. In the years that we lived together, you modeled God's love so well through the enormous amounts of grace and love that you extended towards me. Our conversations have always left me feeling uplifted and encouraged. Summer, I've really felt blessed in getting to know you as well. I've been blessed by the many conversations we've had (with and without Josh) and hearing the things that God has put on your heart. Both of you have been endlessly caring and intentional with me, and I really value all of the ways in which God has blessed me through the two of you. Brian and Nancy Thibeault, the two of you were there for me in a slightly strange season in my life; I really appreciate all the ways in which you've been there for me. It's amazing for me to see the two of you live out in faith; you two carry God's presence so well. Joanna Tang, I've always appreciated the conversations that we shared, especially because we approach the world and our relationships with God in such different ways. It's amazing to see the ways in which your faith has driven you through grad school, and I look forward to what God has in store for you in the road ahead.

Before closing, I want to acknowledge my family for the endless love, care, and support that they have extended to me. Mom and Dad, both of you have sacrificed a lot for me (and

Alexis), for which I am beyond thankful. Thank you for always being there as a constant that I can reliably count on. There's no doubt that I couldn't have made it anywhere close to where I am now without you. Alexis, even with our age gap, you're the best sister I could have asked for. You're full of grace and love, and you've always been there for me when it counted. I'm thankful to God for you and look forward to seeing all that he has in store for you in the road ahead. I also want to acknowledge the love and support from my extended family in China. 虽然很久没有见到你们，我非常感谢你们给我的支持。我希望能尽快见到你们。

Lastly and most importantly, thank you Jesus for the grace and strength throughout this journey. I could only have come this far by faith.

VITA

2014–2017 B.S. in Mathematics. College of Creative Studies, University of California, Santa Barbara.

2014–2017 M.A. in Mathematics. University of California, Santa Barbara.

2019–2021 M.A. in Mathematics. University of California, Los Angeles.

CHAPTER 1

Introduction

In this thesis, we will discuss research in persistent homology, complex social systems, and the interface between the two. Our work in persistent homology includes studying the structure of persistence modules, which are algebraic objects that underlie persistent homology. We also use persistent homology to study a complex social system — specifically, to analyze and quantify resource accessibility in geographical regions. Beyond this, our work in complex social systems has largely focused on modeling opinion dynamics. We have formulated and analyzed new "bounded-confidence models," which are models of opinion dynamics that are motivated by the social phenomenon of "selective exposure."

1.1 Persistent Homology

The structure in which data is organized plays an important role in helping us understand it. In recent years, algebraic tools that were designed to study abstract topological spaces have been adapted to describe and analyze the topological structure in data [EH08]. Under the umbrella of topological data analysis (TDA), these tools have played important roles in numerous domains, including neuroscience [Cur17, SGK18, SEN21], materials science [OKT19], image processing [DCS18, PC14], and complex social systems [FP20, FP21, HNP22]. Perhaps the most popular tool in TDA is persistent homology (PH) [Ghr07, ELZ02, OPT17], which uses ideas from homology theory to provide insight into the underlying topological structure of data by identifying and quantifying "holes" across different scales.

PH is built on homology theory, an area of algebraic topology [Hat02] that characterizes a topological space by its holes. For a fixed dimension k and topological space X, the rank of the abelian group $H_k(X;\mathbb{Z})$ — which is the kth "homology group" of X with integer coefficients — is the number of k-dimensional holes in X. A continuous map $f : X \to Y$ between topological spaces gives rise to a homomorphism (between the associated homology groups) that relates holes in X with holes in Y. Notably, given an inclusion $X \subseteq Y$ of topological spaces (viewed as a continuous map) and positive dimension k, the homomorphism (between homology groups) that is associated with the inclusion map provides information on the holes in X that are filled in Y and the holes in Y that are not in X. One can also consider homology with coefficient in a general ring R; in this case, the homology groups are R-modules instead of abelian groups.

To use PH to analyze a data set, one first builds a *filtration*, which is a nested family $\mathcal{K} = \{\mathcal{K}_i\}_{i\in I}$ of topological spaces — in which the indexing set $I \subseteq \mathbb{R}$ is usually an interval or a discrete set of points — that approximates the data across different scales. For example, if a data set is in the form of a discrete set X of points (which is called a *point cloud*) in $\mathbb{R}^n$, one can construct the "Čech filtration" $\{\check{C}_r(X)\}_{r>0}$, which is defined by $\check{C}_r(X) = \bigcup_{x\in X} B_r(x)$. One can use homology to provide information on how the topology changes in a filtration. In particular, by studying the inclusions between spaces in a filtration, along with the associated homomorphisms (which arise from inclusion) on the respective homology groups, one can determine the scales at which holes form and subsequently fill in. This information can be summarized in a "persistence diagram" (PD) [Rob00], which is a multiset $\{(b_i, d_i)\}_{i\in I}$ of points that correspond to the holes in a filtration. See Chapter 2 for more details.

The theory behind PH it built upon "persistence modules" [Oud15, Sch22], which are algebraic objects that arise naturally from PH and capture information about it. A persistence module (which we define in Section 2.3) consists of the homology groups of a filtration and the homomorphisms (which we call "structure maps") between the homology groups that are asso-

ciated with inclusion. Persistence modules can also be described more generally as a functor from a totally-ordered poset category to a module category. An "interval decomposition" of a persistence module provides the information that is recorded in a PD [Oud15, OY23, Pat18]. We will discuss persistence modules, along with their relationship to PH, in Section 2.3.

1.2 Models of Opinion Dynamics

Social interactions play an important role in shaping the opinions of individuals, communities of people, and society at large [BAB21]. An individual's opinion on a topic is often influenced by the people with whom they interact [Jac08], and researchers in many disciplines study such interactions and how they change opinions and actions [Noo20]. In an agent-based model of opinion dynamics, each agent represents an individual and a network encodes which agents are able to interact with each other. Each node (i.e., agent) of a network has an opinion in some opinion space. Studying opinion models allows researchers to examine the evolution of opinions on social networks with time, leading to insights into the spread of ideas [FJ90, JMF15], when communities of individuals reach consensus and when they do not [VBI21], and the formation of "opinion clusters" (i.e., clusters of nodes with similar opinions) [Lor08].

Individuals are often influenced most by people and other sources whose opinions are similar to theirs [CM11]. This phenomenon is encapsulated in a simple form in *bounded-confidence models* (BCMs) [NVT20,HK02,DNA00] of opinion dynamics, in which the nodes of a network have continuous-valued opinions and interacting nodes influence each others' opinions if and only if their opinions are sufficiently similar. A key feature of BCMs is the presence of a "confidence bound," which is a parameter that determines which nodes can influence each other. A node can influence and be influenced by its neighbors only when the difference in their opinions is less than their confidence bound.

The two most popular BCMs are the Hegselmann–Krause (HK) model [Kra00, HK02] and

the Deffuant–Weisbuch (DW) model [DNA00]. Both of these models use discrete time. The HK model updates synchronously; at each time, every node updates its opinions based on the opinions of all of its neighbors. The DW model updates asynchronously; at each time, one selects a dyad (i.e., a pair of adjacent nodes and the edge between them), and the two nodes in the dyad interact and potentially influence each others' opinions. The DW model also has a compromise parameter, which controls how much nodes in a dyad influence each other when they compromise their opinions. In both the HK model and the DW model, the confidence bound is traditionally a single fixed scalar parameter that is the same for all dyads.

1.3 Contributions

In Chapter 3, We formulate a necessary and sufficient condition for which interval decompositions exist for persistence modules with coefficients in a principal ideal domain (PID). Persistence modules with coefficients in a field are guaranteed to have an interval decomposition [Oud15, Gab72]. However, when considering persistence modules with coefficients in a general ring, interval decompositions may not exist (e.g., see the discussion in Section 3.3). We prove that, given a persistence module of pointwise free and finitely-generated modules with coefficients in a PID, an interval decomposition exists if and only if the cokernel of every structure map is free. We also formulate an algorithm to compute an interval decomposition of a persistence module of freely- and finitely-generated modules over a PID when one exists. The formulation of our algorithm is informed and underpinned by the our theoretical result on the existence of interval decompositions.

In Chapter 4, we use TDA to quantify and assess the accessibility of resources in geographical regions. We use PH to identify "holes in coverage," which are regions that have poor access to the resources in question. Our use of PH allows us to consider holes in coverage across all scales, instead of choosing a fixed cutoff distance. To measure the accessibility of resources,

we consider the time cost of resource access — namely, travel and waiting time — to construct our filtration. We focused our work on polling-site access across six geographical regions (five cities and Los Angeles County).

In Chapter 5, we formulate and analyze discrete-time BCMs with heterogeneous and adaptive confidence bounds. We introduce two new models: (1) a BCM with synchronous opinion updates that generalizes the Hegselmann–Krause (HK) model and (2) a BCM with asynchronous opinion updates that generalizes the Deffuant–Weisbuch (DW) model. We analytically and numerically explore our adaptive BCMs' limiting behaviors, including the confidence-bound dynamics, the formation of clusters of nodes with similar opinions, and the time evolution of an "effective graph," which is a time-dependent subgraph of a network with edges between nodes that are currently receptive to each other. For a variety of networks and a wide range of values of the parameters that control the increase and decrease of confidence bounds, we demonstrate numerically that our adaptive BCMs result in fewer major opinion clusters and longer convergence times than the baseline (i.e., nonadaptive) BCMs.

1.4 Organization

In Chapter 2, we briefly present the background for PH that is relevant to this thesis. In Chapter 3, we discuss interval decompositions of persistence modules over a principal ideal domain, which is based on research presented in [LH23]. In Chapter 4, we discuss the use of PH to study resource accessibility, which is based on research presented in [HJJ23]. In Chapter 5, we discuss BCMs with adaptive confidence bounds, which is based on research presented in [LLP23]. In Chapter 6, we give concluding remarks to this thesis.

CHAPTER 2

Background on Persistent Homology

In this section, we give a brief summary of the background for persistent homology (PH). For a more comprehensive background on PH, see [Ghr07, OPT17, DW22]. For background on persistence modules and its relation to PH, see [Oud15, Sch22].

2.1 Homology

Before introducing PH, we provide intuition for homology theory, which is the theoretical foundation on which PH is built. For a comprehensive background and discussion on homology (including the formal definition of a homology group), see [Hat02].

The theory of *homology* is an area of algebraic topology [Hat02] that characterizes a topological space by its "holes." Given a topological space X and a fixed dimension $k > 0$, the rank of the abelian group $H_k(X; \mathbb{Z})$ — which is the kth "homology group" of X with integer coefficients — is the number of k-dimensional holes in X. One can think of k-dimensional holes as follows: 1-dimensional holes are loops that cannot be contracted to a point, 2-dimensional holes are regions bounded by a 2-dimensional surface, and so on. When $k = 0$, the homology group $H_0(X; \mathbb{Z})$ is torsion-free and its rank is the number of connected components.

For example, suppose that X is the torus $S^1 \times S^1$. Because there is only one path component, the zeroth homology group $H_0(X; \mathbb{Z}) \cong \mathbb{Z}$ has rank 1. Because there are two 1-dimensional holes (one for each S^1 in the product), the first homology group $H_1(X; \mathbb{Z}) \cong \mathbb{Z}^2$ has rank 2.

Finally, because there is one region bounded by the torus (which is a 2-dimensional surface), the second homology group $H_2(X;\mathbb{Z}) \cong \mathbb{Z}$ has rank 1.

A continuous map $f : X \to Y$ between topological spaces gives rise to homomorphisms between the associated homology groups. For a fixed dimension k, the homomorphism $f_* : H_k(X;\mathbb{Z}) \to H_k(Y;\mathbb{Z})$ relates k-dimensional holes in X to those in Y. Notably, given an inclusion $X \subseteq Y$ of topological spaces (viewed as a continuous map) and positive dimension k, the homomorphism between homology groups that is associated with the inclusion map provides information on the holes in X that are filled in Y and the holes in Y that are not in X.

Beyond integer coefficients, one can also consider homology with coefficient in a different ring. If one considers homology with coefficients in a ring R, the homology groups are R-modules instead of abelian groups.

2.2 Persistent Homology

2.2.1 Persistent homology overview

When performing PH, one starts with a filtration[1] $\mathcal{K} = \{\mathcal{K}_i\}_{i\in I}$ of topological spaces. We refer to the index i as the *filtration parameter* and specific values $i \in I$ as *filtration-parameter values*. In this setting, one is interested in understanding the filtration-parameter values that holes form and fill in. Homology is a natural tool to understand the holes in a filtration. Taking the nth homology with integer coefficients results in a sequence $\{H_n(\mathcal{K}_i;\mathbb{Z})\}_{i\in I}$ of homology groups and homomorphisms $\{\phi_{i,j} : H_n(\mathcal{K}_i;\mathbb{Z}) \to H_n(\mathcal{K}_i;\mathbb{Z}) | i, j \in I \text{ and } i \leq j\}$ (which we call *structure maps*) between the homology groups that arise from the inclusions in $\mathcal{K}$. In this setting, the homology groups are $\mathbb{Z}$-modules[2] and the structure maps are $\mathbb{Z}$-

[1] Recall the definition of a filtration from Section 1.1.

[2] $\mathbb{Z}$-modules are precisely abelian groups.

module homomorphisms. Because the rank of the homology group $H_n(\mathcal{K}_i;\mathbb{Z})$ is the number of n-dimensional holes in $\mathcal{K}_i$, studying the homology groups and structure maps that arise from $\mathcal{K}$ provides information on the holes in $\mathcal{K}$. In particular, the elements (which we call *homology classes*) of $H_n(\mathcal{K}_i;\mathbb{Z})$ represent n-dimensional holes in $\mathcal{K}_i$. Studying the images of homology classes under the structure maps help identify holes between different filtration-parameter values.

In practice, one usually takes homology with coefficients in a field F. This results in a sequence $\{H_n(\mathcal{K}_i;F)\}_{i\in I}$ of homology groups that are vector spaces over F and structure maps $\{\phi_{i,j} : H_n(\mathcal{K}_i;\mathbb{Z}) \to H_n(\mathcal{K}_i;\mathbb{Z}) | i,j \in I \text{ and } i \leq j\}$ that are F-linear maps between homology groups. The dimension of $H_n(\mathcal{K}_i;F)$ approximates[3] the rank of $H_n(\mathcal{K}_i;\mathbb{Z})$. One can interpret elements of a basis of $H_n(\mathcal{K}_i;F)$ as corresponding to the n-dimensional holes in $\mathcal{K}_i$. As r increases, new homology classes are "born" and existing ones "die." A homology class $\gamma \in H_n(\mathcal{K}_k;F)$ is *born* at filtration-parameter value b if it is not in the image of $\phi_{i,k}$ for any $i < b$. One can interpret the birth of a homology class at b as a hole that forms at that filtration-parameter value in the filtration. Similarly, a homology class $\gamma \in H_n(\mathcal{K}_k;F)$ *dies* at filtration-parameter value d if $\phi_{k,d}(\gamma) = 0$ and $\phi_{k,j}(\gamma) \neq 0$ whenever $j < d$. One can interpret the death of a homology class at d as a hole that fills in at that filtration-parameter value in the filtration.

To identify homology classes between different filtration-parameter values in $\mathcal{K}$, one can choose a basis β_i for each $H_n(\mathcal{K}_i;F)$ such that, given a structure map $\phi_{i,j} : H_n(\mathcal{K}_i;F) \to H_n(\mathcal{K}_j;F)$, we have that $\phi_{i,j}$ maps elements of β_i either to 0 or to elements of β_j in a one-to-one manner. That is, given $v \in \beta_i$, precisely one of the following two statements holds:

1. $f(i \leq j)(v) = 0$;

[3]If $H_n(\mathcal{K}_i;\mathbb{Z})$ is torsion-free, then the dimension of $H_n(\mathcal{K}_i;F)$ and the rank of $H_n(\mathcal{K}_i;\mathbb{Z})$ are equal. This is often the case in practice (see [LTH21, OY23]).

2. $f(i \leq j)(v) \in \beta_j$ and $f(i \leq j)(\tilde{v}) = f(i \leq j)(v)$ implies that $\tilde{v} = v$ for any $\tilde{v} \in \beta_i$.

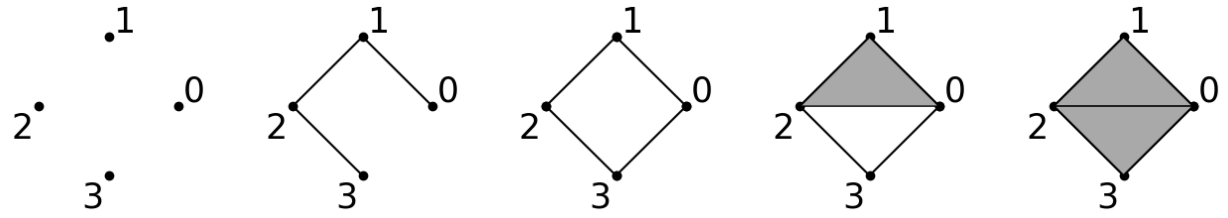
Such a choice of basis may not exist when the coefficient ring for computing homology is not a field (e.g., see Section 3.3). This choice of bases gives a correspondence between homology classes in $\mathcal{K}$ at different filtration-parameter values. For example, given $v \in \beta_i$ and $w \in \beta_j$, one can interpret $\phi_{i,j}(v) = w$ to mean that the hole corresponding to v at filtration-parameter value i can be identified with the hole corresponding to w at filtration-parameter value j. Similarly, one can interpret $\phi_{i,j}(v) = 0$ to mean that the hole corresponding to v fills in at filtration-parameter value j. This choice of basis gives the filtration-parameter values that homology classes in $\mathcal{K}$ are born and die. This is often summarized by a *persistence diagram*, which is a multiset $PD_F^n(\mathcal{K}) = \{(b_k, d_k)\}_k$ of points in $\overline{\mathbb{R}}^2$ such that each point (b_k, d_k) corresponds to a homology class that is born at filtration-parameter value b_k and dies at filtration-parameter value d_k. If $d_k = \infty$, the corresponding hole never dies. For example, the filtration in Figure 2.1 has the associated persistence diagram in Figure 2.2.

2.2.2 Filtration of simplicial complexes

In many applications, one works with a filtration of simplicial complexes. A "simplicial complex" is a set of "simplices" (which are vertices, edges, triangles, tetrahedrons, and higher-dimensional analogues of these objects) that satisfy certain requirements on simplex boundaries and pairwise simplex intersections.

More precisely, simplicial complexes are defined as follows. A *k-simplex* σ is a k-dimensional convex hull of $k+1$ vertices. A *face* of σ is the convex hull of a nonempty subset of the vertices of σ. A *simplicial complex* X is a collection of simplices such that (1) if $\sigma \in X$, then any face of σ is also in X and (2) given $\sigma, \tau \in X$, we have that $\sigma \cap \tau$ is a face of both σ and τ. One often uses a simplicial complex to approximate a topological space.

Given a point cloud $X = \{x_1, \ldots, x_n\}$ in a metric space (M, d), there are several ways to construct a filtration of simplicial complexes that approximates the shape of X. Two of



(a) $\mathcal{K}_0$ (b) $\mathcal{K}_1$ (c) $\mathcal{K}_2$ (d) $\mathcal{K}_3$ (e) $\mathcal{K}_4$

Figure 2.1: An example of a filtration of simplicial complexes. The simplicial complex $\mathcal{K}_i$ has the associated filtration-parameter value i. [This figure appeared originally in [HNP22].]

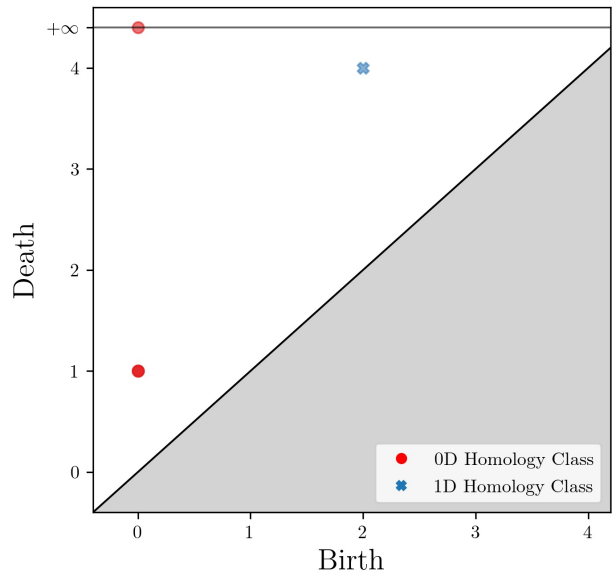


Figure 2.2: The persistence diagram for the 0-dimensional (0D) and 1-dimensional (1D) PH of the filtration in Figure 2.1. [This figure first appeared in [HNP22]]

the most common constructions are the Čech filtration and the Vietoris–Rips (VR) filtration [OPT17]. For $r > 0$, the *Čech complex* $\check{\mathrm{C}}_r(X, M, d)$ at filtration-parameter value r is the simplicial complex that has a simplex with vertices $[x_{i_0}, \ldots, x_{i_k}]$ if the intersection $\bigcap_j B(x_{i_j}, r)$ is nonempty, where $B(x, r) := \{y \in M \mid d(x, y) \leq r\}$. That is, $\check{\mathrm{C}}_r(X, M, d)$ is the nerve of the closed balls $\{B(x_i, r)\}_{x_i \in X}$. By the Nerve Theorem [Bor48], $\check{\mathrm{C}}_r(X, M, d)$ is topologically equivalent (more precisely, it is homotopy-equivalent) to the union $\bigcup_i B(x_i, r)$ of balls in M whenever the balls $B(x_i, r)$ are convex.[4] This implies that $\bigcup_i B(x_i, r)$ and

[4]This condition is satisfied for all r when (M, d) is Euclidean, but it is not always satisfied for non-

$\check{\mathrm{C}}_r(X, M, d)$ have the same homology (i.e., the same set of holes). A *Čech filtration* is a filtration of Čech complexes for increasing filtration-parameter value r. In Figure 2.3, we show an example of a Čech filtration.

In practice, it is uncommon to use Čech filtrations because they are difficult to compute. A *Vietoris–Rips (VR) complex* $\mathrm{VR}_r(X, M, d)$ is an approximation of a Čech complex that is faster to compute because it is only necessary to calculate pairwise distances between points. The VR complex at filtration-parameter value r has a simplex with vertices $[x_{i_0}, \ldots, x_{i_k}]$ if $d(x_{i_j}, x_{i_\ell}) < 2r$ for all j and ℓ. A *VR filtration* is a filtration of VR complexes for a sequence of increasing filtration-parameter values. A VR filtration "approximates" a Čech filtration in the sense that

$$\check{\mathrm{C}}_r(X, M, d) \subseteq \mathrm{VR}_r(X, M, d) \subseteq \check{\mathrm{C}}_{\sqrt{2}r}(X, M, d) \tag{2.1}$$

for all r. The complexes $\mathrm{VR}_r(X, M, d)$ and $\check{\mathrm{C}}_r(X, M, d)$ have the same set of edges for all r.

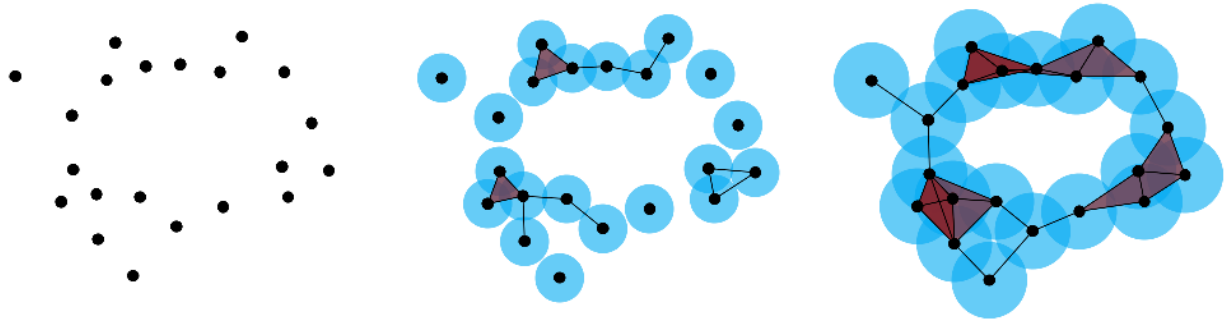


Figure 2.3: Illustration of a Čech filtration for a point cloud X that we sample from an annulus. [This figure appeared originally in [HJJ23]. We generated this figure using [AS11].]

Weighted versions of the Čech and VR filtrations were described in [ACG19]. Given a point cloud $X = \{x_1, \ldots, x_n\}$ in a metric space (M, d) and associated weights $\{w_1, \ldots, w_n\}$, the *radius function* at x_i is

$$r_{x_i}(t) := \begin{cases} -\infty\,, & t < w_i \\ t - w_i\,, & \text{otherwise}\,. \end{cases} \tag{2.2}$$

Euclidean metric spaces.

The closed ball $B(x_i, r_{x_i}(t))$ has no points until time $t = w_i$; at that time, the radius starts growing linearly with t, which is the filtration parameter. The *weighted Čech complex* $\check{\mathrm{C}}_t^{\mathrm{weighted}}(X, M, d, \{w_i\})$ at filtration parameter t is the simplicial complex that has a simplex with vertices $[x_{i_0}, \ldots, x_{i_k}]$ if the intersection $\bigcap_j B(x_{i_j}, r_{x_{i_j}}(t))$ is nonempty. (That is, $\check{\mathrm{C}}_t^{\mathrm{weighted}}(X, M, d, \{w_i\})$ is the nerve of $\{B(x_i, r_{x_i}(t))\}_{x_i \in X}$.) Like the unweighted Čech complex, the weighted Čech complex is homotopy-equivalent to the union $\bigcup_i B(x_i, r_{x_i}(t))$ of balls by the Nerve Theorem whenever the balls $B(x_i, r_{x_i}(t))$ are convex for all x_i. Much like an unweighted Čech complex, a weighted Čech complex requires too much time to compute in practice, so researchers usually instead compute a *weighted VR complex* $\mathrm{VR}_t^{\mathrm{weighted}}(X, M, d, \{w_i\})$. This is the simplicial complex whose vertices are $\{x_i \mid w_i < t\}$ and whose simplices $[x_{i_0}, \ldots, x_{i_k}]$ satisfy $d(x_{i_j}, x_{i_\ell}) + w_{i_j} + w_{i_\ell} < 2t$. The sequence $\{\mathrm{VR}_t^{\mathrm{weighted}}(X, M, d, \{w_i\})\}_t$ for increasing t is the *weighted VR filtration*. Analogously to Equation (2.1), the weighted VR filtration "approximates" the weighted Čech filtration in the sense that

$$\check{\mathrm{C}}_r(X, M, d, \{w_i\}) \subseteq \mathrm{VR}_r(X, M, d, \{w_i\}) \subseteq \check{\mathrm{C}}_{\sqrt{2}r}(X, M, d, \{w_i\}) \tag{2.3}$$

for all r.

In addition to computing a persistence diagram, which tracks the birth and death of homology classes, one can identify the "birth simplex" and "death simplex" of homology classes in a filtration of simplicial complexes. The *birth simplex* is the simplex that creates a homology class, and its *death simplex* is the simplex that fills it in. For example, in Figure 2.1, a 1-dimensional (1D) homology class is born at filtration-parameter value 2. Its birth simplex is the edge with vertices 0 and 3. The same homology class subsequently dies at filtration-parameter value 4. Its death simplex is the triangle with vertices 0, 2, and 3.

2.3 Persistence Modules

Persistent homology yields an algebraic structure called a "persistence module" [Oud15, Sch22]. We briefly discuss the theoretical framework of persistent modules and their relation to persistent homology.

Let R be a ring. A (finitely-indexed) *persistence module* with R-coefficients is a functor $f : \{0, 1, \ldots, m\} \to R\text{-Mod}$, where $\{0, 1, \ldots, m\}$ is a finite totally-ordered poset category and $R-$Mod is the category of R-modules. That is, f can be viewed[5] as an algebraic structure that consists of (1) a finite collection $\{f_i\}_{i=0}^m$ of R-modules and (2) R-homomorphisms $\{f(i \leq j) : f_i \to f_j\}_{0 \leq i \leq j \leq m}$ (which we call *structure maps*[6]) between the aforementioned R-modules such that $f(i \leq i) = \mathrm{id}_{f_i}$ and $f(i \leq j) = f(s \leq j) \circ f(i \leq s)$ for $i \leq s \leq j$.

Given $b, d \in \{0, \cdots, m\}$ such that $0 \leq b < d \leq m$, let $I_R^{b,d}$ denote the persistence module $\{(I_R^{b,d})_i\}_{i=0}^m$ with R-coefficients such that

$$(I_R^{b,d})_i = \begin{cases} R\,, & b \leq i < d \\ 0\,, & \text{otherwise}\,, \end{cases}$$

$$I_R^{b,d}(i \leq j) : (I_R^{b,d})_i \to (I_R^{b,d})_j = \begin{cases} \mathrm{id}_R\,, & b \leq i \leq j < d \\ 0\,, & \text{otherwise}\,. \end{cases}$$

We refer to the persistence modules $I_R^{b,d}$ as *interval modules*. An *interval decomposition* of f

[5]In the category $\{0, \ldots, m\}$, the objects are integers i such that $0 \leq i \leq m$; for each $i, j \in \{0, \ldots, m\}$ such that $i \leq j$, there is a unique morphism (denoted "$i \leq j$") between i and j. As a functor, f assigns an R-module f_i to each $i \in \{0, \cdots, m\}$ and a homomorphism $f(i \leq j) : f_i \to f_j$ to each morphism $i \leq j$.

[6]The structure maps discussed here are generalizations of the structure maps in Section 2.2. The structure maps in Section 2.2 are homomorphisms between homology groups that are associated with the inclusion map between their respective topological spaces, whereas the structure maps defined here can be any homomorphism between R-modules.

is a decomposition of f into interval modules:

$$f \cong \bigoplus_k I_R^{b_k, d_k},$$

where $0 \leq b_k < d_k \leq m$. Another way to view an interval decomposition is to choose a basis β_i for each f_i in a persistence module such that the structure map $f(i \leq j) : f_i \to f_j$ maps elements of β_i either to 0 or to basis elements in β_j in a one-to-one manner. That is, given $v \in \beta_i$, precisely one of the following two statements holds:

1. $f(i \leq j)(v) = 0$;
2. $f(i \leq j)(v) \in \beta_j$ and $f(i \leq j)(\tilde{v}) = f(i \leq j)(v)$ implies that $\tilde{v} = v$ for any $\tilde{v} \in \beta_i$.

The structure theorem of Gabriel [Gab72] is a general result describing the indecomposable isomorphism classes of quiver representations. In the context of persistence modules, it implies the following result.

Theorem 2.1 (Gabriel)**.** *Suppose that R is a field and that f is a persistence module over R. Then*

- *f admits an interval decomposition;*
- *two direct sums of nonzero interval modules are isomorphic — that is, $\bigoplus_{k=1}^{M} I_R^{b_k, d_k} \cong \bigoplus_{\ell=1}^{N} I_R^{b'_\ell, d'_\ell}$ — if and only if $M = N$ and there exists a permutation π such that $(b_1, d_1), \ldots, (b_M, d_M)$ equals $(b'_{\pi(1)}, d'_{\pi(1)}), \ldots, (b'_{\pi(M)}, d'_{\pi(M)})$.*

By Theorem 2.1, when R is a field, every persistence module with coefficients in R admits an interval decomposition. However, interval decompositions for persistence modules with non-field coefficients are not guaranteed to exist.

2.3.1 Relation to persistent homology

Interval decompositions of persistence modules provide a critical foundation to study PH. Given a filtration $\mathcal{K} = \{\mathcal{K}\}_{i=0}^{m}$, one can take the nth homology with coefficients in a field F to obtain a persistence module $H_n(\mathcal{K}; F)$, which consists of (1) a sequence $\{H_n(\mathcal{K}_i; F)\}_{i=0}^{m}$ of vector spaces (these are the *homology groups* of $\mathcal{K}$ in dimension n) over F that correspond to the topological spaces in $\mathcal{K}$ and (2) F-linear maps $\{\phi_i^j : H_n(\mathcal{K}_i; F) \to H_n(\mathcal{K}_j; F)\}_{0 \leq i \leq j \leq m}$ between homology groups (these are the structure maps) that arise from the inclusion maps. Given an interval decomposition

$$H_n(\mathcal{K}; F) \cong \bigoplus_{k=1}^{r} I_F^{b_k, d_k} ,$$

the persistence diagram for $\mathcal{K}$ in dimension n with F-coefficients is the multiset $\mathrm{PD}_n^F(\mathcal{K}) = \{(b_k, d_k)\}_{k=1}^{r}$. The module $H_n(\mathcal{K}; F)$ always admits an interval decomposition, as the coefficient ring is a field. This characterization formalizes the notion of a persistence diagram, which we discussed in Section 2.2. A persistence module can also arise from applying homology to a *persistent topological space*, which is a sequence $X_0 \to \cdots \to X_m$ of topological spaces and continuous maps.

CHAPTER 3

Interval Decomposition of Persistence Modules over a Principal Ideal Domain

In this chapter, we study persistence modules and their interval decompositions. The existence of interval decompositions of a persistence module underlies much theory behind persistent homology (PH) (see Section 2.3). We study persistence modules with coefficients in a principal ideal domain (PID). We formulate and prove necessary and sufficient conditions for the existence of interval modules. Concretely, we show that a persistence module that is pointwise[1] free and finitely-generated over a PID splits as a direct sum of interval modules if and only if every structure map has a free cokernel. We then provide an algorithm that either (1) computes such a decomposition explicitly or (2) verifies that no such decomposition exists. This chapter is adapted[2] from [LH23], which is in collaboration[3] with Gregory Henselman-Petrusek.

Throughout this chapter, we use functor notation to describe persistence modules (see Section 2.3 for details). Recall that a persistence module is a functor $f : \{0, \dots, m\} \to R\text{-Mod}$. In the category $\{0, \dots, m\}$, the objects are integers i such that $0 \leq i \leq m$; for each

[1]The word "pointwise" modifies both "free" and "finitely-generated." In Section 3.1, we define what it means for a persistence module to be pointwise free and finitely-generated.

[2]All figures in this chapter originally appeared in [LH23].

[3]I proved the theoretical results in Sections 3.4, 3.5, and 3.6; formulated and proved the correctness of Algorithm 3.1 in Section 3.7; formulated and proved the correctness of Algorithm 3.2 in Section 3.8 with G. Henselman-Petrusek; and wrote the paper with G. Henselman-Petrusek.

$i, j \in \{0, \ldots, m\}$ such that $i \leq j$, there is a unique morphism (denoted "$i \leq j$") between i and j. As a functor, f assigns an R-module f_i to each $i \in \{0, \ldots, m\}$ and a homomorphism (i.e., structure map) $f(i \leq j) : f_i \to f_j$ to each morphism $i \leq j$.

This chapter proceeds as follows. We introduce our problem and summarize our contributions in Section 3.1. We discuss related works in Section 3.2. We relate our work to the independence of persistence diagrams (PDs) with respect to field choice in Section 3.3. We present and briefly prove the uniqueness of interval decompositions, when they exist, in Section 3.4. We present a necessary and sufficient condition for the existence of interval decompositions for pointwise free and finitely-generated persistence modules over PIDs in Section 3.5, and we prove necessity and sufficiency in Sections 3.5 and 3.6, respectively. We then provide a simple algorithm to compute an interval decomposition (when it exists), with a high-level description in Section 3.7 and a detailed description in the language of matrix algebra in Section 3.8. We conclude and discuss our results in Section 3.9. We give information about the Smith-normal-form factorization of matrices and other relevant matrix-algebra facts in Section 3.A.

3.1 Introduction

Persistence modules are algebraic objects that are central to PH. Under certain conditions, a persistence module admits an interval decomposition, in which it decomposes into indecomposable pieces called "interval modules." Interval decompositions are rich in information. For example, they provide the information for PDs (see Section 2.2).

The literature on PH has focused traditionally on homology groups with coefficients in a field because persistence modules with field coefficients are guaranteed to decompose into interval modules (see Theorem 2.1). However, there is an increasing amount of research on PH with different coefficients (e.g., the ring of integers) [Pat18, GM23, PS23]. A fundamental question that one can ask is how to determine whether or not a persistence module not decomposes

into interval modules.

We answer this question for any persistence module f that is *pointwise free and finitely-generated* over a principle ideal domain R. Concretely, this condition entails that $f : \{0, \ldots, m\} \to R$-Mod is a persistence module in which f_a is a free and finitely-generated R-module for all $a \in \{0, \ldots, m\}$. For any persistence module f that satisfies this condition, we have Theorem 3.1, which is the main computational result of this chapter.

Theorem 3.1. *Let $f : \{0, \ldots, m\} \to R$-Mod be a persistence module that is pointwise free and finitely-generated over R. There is a formal procedure (see Algorithm 3.2) to determine whether f admits an interval decomposition and, if so, to explicitly construct such a decomposition. The procedure has finite (respectively, polynomial) time if matrices over the coefficient ring R can be multiplied, inverted, and placed in Smith normal form in finite (respectively, polynomial) time.*

The procedure relies on and is informed by Theorem 3.2, which is the main theoretical result of this chapter.

Theorem 3.2. *Let $f : \{0, \ldots, m\} \to R$-Mod be a persistence module that is pointwise free and finitely-generated over R. The persistence module f splits into a direct sum of interval modules if and only if the cokernel of every structure map $f(a \leq b) : f_a \to f_b$ is free.*

Theorem 3.2 offers structural insights into PH for both theoretical and applied researchers. Theorists gain access to a wider variety of tools in settings in which interval decompositions exist. For example, if a persistence module f of free abelian groups decomposes as a direct sum of interval modules, then the PD of $f \otimes F$ is identical for every choice of coefficient field F (see Lemma 3.4). Practitioners gain both computational power and analytical insight into persistence modules with non-field coefficients. For example, the ability to reason with persistence modules with PID coefficients (and, in particular, $\mathbb{Z}$-coefficients) is important to several forms of dimension reduction that involve circular and projective coordinates

(see [DV09, Per18, Per20, SGB23]).

3.2 Related Works

The study of interval decompositions belongs to a large body of literature dedicated to understanding the isomorphism invariants of persistence modules [ELZ02, CZ09]. Patel [Pat18] generalized the notion of PDs to persistence modules of the form $f : \{0, \dots, m\} \to C$, where C is a symmetric monoidal category with images or (under a slightly different construction) where C is an abelian category. These ideas were extended recently using Galois connections [GM23] and Möbius homology [PS23]. An important class of examples comes from 2-parameter persistence [BL23]; specifically, consider functors from the poset category of $\{0, \dots, m\} \times \{0, \dots, m\}$ into C. These objects are used in applications [LW15, CDH22, BL24] and can be modeled as persistence modules of persistence modules (i.e. functors $\{0, \dots, m\} \to D$, where D is the category[4] of persistence modules with values in C).

Obayashi and Yoshiwaki [OY23] studied the dependence of PDs on the choice of field in the setting of filtrations (i.e., nested sequences of topological spaces). They introduced conditions on the homology of a filtration that are necessary and sufficient for the associated PDs to be independent of the choice of field[5] (see Section 3.3). They produced an algorithm to verify field-choice independence by adapting a standard algorithm to compute PDs for simplex-wise filtrations.[6] They also conducted numerical experiments that demonstrate empirically that PDs rarely depend on the choice of field. Li et al. [LTH21] reported similar empirical results.

[4]Concretely, this is the category whose objects are functors of the form $f : \{0, \dots, m\} \to C$.

[5]We will prove that these conditions are equivalent to the condition that the associated persistence module with $\mathbb{Z}$-coefficients splits as a direct sum of interval submodules (see Theorem 3.3).

[6]A *simplex-wise filtration* on a simplicial complex $K = \{\sigma_1, \dots, \sigma_m\}$ is a nested sequence of sub-simplicial complexes $\emptyset = \mathcal{K}_0 \subseteq \cdots \subseteq \mathcal{K}_m = K$ such that $\mathcal{K}_p = \{\sigma_1, \dots, \sigma_p\}$ for all $p \in \{0, \dots, m\}$.

There remain several gaps in the current literature. Although PDs have been defined for general families of functors $\{0, \ldots, m\} \to C$ [Pat18], it is only known that interval decompositions exist when C is a category of finite-dimensional vector spaces (see Theorem 2.1). Our results concerning the fundamental structure of persistence modules have ramifications in PH. In particular, we complement the results of [OY23] by linking field-independence to the algebraic structure (concretely, the indecomposable factors) of a persistence module, rather than the topology of an underlying simplicial complex. Such a perspective is important, as it helps one to relate topological ideas involving filtrations with algebraic ideas involving the corresponding persistence modules. By better understanding persistence modules, we also gain a better grasp of their relation to PH and its generalizations.

3.3 Persistent Diagrams and Independence of Field Choice

Theorem 3.2 allows us to deduce that the PD of a filtered topological space is independent of the choice of coefficient field if and only if the associated persistence module over $\mathbb{Z}$ splits as a direct sum of interval submodules (subject to an extra homological condition that we state in Theorem 3.3). Our proof relies on a result of Obayashi and Yoshiwaki [OY23, Theorem 1.9] that guarantees the equivalence of conditions 3 and 4 in Theorem 3.3.

Theorem 3.3. *Let $\mathcal{K} = \{\mathcal{K}_i\}_{i=0}^m$ be a filtration of topological spaces. Suppose that the homology groups $H_{k-1}(\mathcal{K}_a; \mathbb{Z})$ and $H_k(\mathcal{K}_a; \mathbb{Z})$ are free for every a such that $0 \leq a \leq m$. Then the following are equivalent.*

1. *The persistence module $H_k(\mathcal{K}; \mathbb{Z})$ splits as a direct sum of interval submodules.*
2. *For all a and b such that $0 \leq a \leq b \leq m$, the cokernel of the structure map $\phi_a^b : H_k(\mathcal{K}_a; \mathbb{Z}) \to H_k(\mathcal{K}_b; \mathbb{Z})$ is free.*
3. *(Obayashi and Yoshiwaki) For all a and b such that $0 \leq a \leq b \leq m$, the relative homology group $H_k(\mathcal{K}_b, \mathcal{K}_a)$ is free.*

4. *(Obayashi and Yoshiwaki) The persistence diagram $PD_k^F(\mathcal{K})$ of dimension k is independent of the choice of coefficient field F.*

Proof. Obayashi and Yoshiwaki [OY23, Theorem 1.9] showed that conditions 3 and 4 are equivalent. Theorem 3.2 gives the equivalence between 1 and 2. It remains to show the equivalence between 2 and 3.

We have the following long exact[7] sequence for relative homology[8]:

$$\cdots \to H_k(\mathcal{K}_a;\mathbb{Z}) \xrightarrow{\phi_a^b} H_k(\mathcal{K}_b;\mathbb{Z}) \xrightarrow{j_*} H_k(\mathcal{K}_b,\mathcal{K}_a;\mathbb{Z}) \xrightarrow{\partial} H_{k-1}(\mathcal{K}_a;\mathbb{Z}) \to \cdots ,$$

from which we can extract the following short exact sequence

$$0 \to \operatorname{coker}(\phi_a^b) \to H_k(\mathcal{K}_b,\mathcal{K}_a;\mathbb{Z}) \to \operatorname{Im}(\partial) \to 0\,.$$

Suppose that $\operatorname{coker}(\phi_a^b)$ is free. Because $\operatorname{Im}(\partial)$ is free (it is a subgroup of the free abelian group $H_{k-1}(X_a;\mathbb{Z})$), this short exact sequence splits, which implies that $H_k(\mathcal{K}_b,\mathcal{K}_a;\mathbb{Z}) \cong \operatorname{coker}(\phi_a^b) \oplus \operatorname{Im}(\partial)$. Because $\operatorname{coker}(\phi_a^b)$ is free, $H_k(\mathcal{K}_b,\mathcal{K}_a;\mathbb{Z})$ must also be free.

Conversely, if $H_k(\mathcal{K}_b,\mathcal{K}_a;\mathbb{Z})$ is free, then $\operatorname{coker}(\phi_a^b)$ must also be free, because it injects into a free abelian group. □

Remark 3.1. *In a prior version of [OY23] (which is available on arXiv [OY20]), Obayashi and Yoshiwaki gave a constructive proof of the equivalence of conditions 1 and 3 in the special case in which $\{\mathcal{K}_i\}_{i=1}^m$ is a simplex-wise filtration on a simplicial complex [OY20, Lemma 2].*

[7] A sequence $M_1 \xrightarrow{\eta_1} M_2 \xrightarrow{\eta_2} \cdots \to M_n$ of modules is *exact* if $\operatorname{Im}(\eta_j) = \ker(\eta_{j+1})$ for all j. If it precisely of the form $0 \to A \to B \to C \to 0$, we say that it is a *short exact sequence*. Otherwise, it is a *long exact sequence*.

[8] Such a long exact sequence exists for any pair of nested topological spaces. The key detail in our discussion is that the sequence is exact; the precise definitions of j_* and ∂ are not important. See [Hat02, Section 2.1] for a detailed exposition, including the definitions of j_* and ∂.

We illustrate the relation between the field-choice independence of a PD of a filtration and the existence of an interval decomposition of the corresponding persistence module with $\mathbb{Z}$ coefficients. Consider the following persistent topological space $\mathcal{K}$:

$$\bullet \hookrightarrow S^1 \xrightarrow{f} S^1 \hookrightarrow \mathbb{C}\,,$$

where $\bullet$ is a single point, $S^1 = \{z \in \mathbb{C} |\ \|x\| = 1\}$, and $f(z) = z^2$. By taking the first homology over a field F, we obtain the following persistence module[9]:

$$0 \to F \xrightarrow{\cdot 2} F \to 0\,.$$

If $\mathrm{char}(F) = 2$, the multiplication-by-2 map is the zero map, so we can write

$$(0 \to F \xrightarrow{\cdot 2} F \to 0) \cong (0 \to F \to 0 \to 0) \oplus (0 \to 0 \to F \to 0)\,.$$

Otherwise, the multiplication-by-2 map is an isomorphism and the persistence module in question is an interval module. Therefore, when indexing from 0, the corresponding PD $\mathrm{PD}_1^F(\mathcal{K})$ for first homology depends on the characteristic of the coefficient field:

$$\mathrm{PD}_1^F(\mathcal{K}) = \begin{cases} \{(1,2),(2,3)\}\,, & \mathrm{char}(F) = 2\,, \\ \{(1,3)\}\,, & \text{otherwise}\,. \end{cases}$$

When taking the first homology of $\mathcal{K}$ with integer coefficients, the corresponding persistence module is

$$0 \to \mathbb{Z} \xrightarrow{\cdot 2} \mathbb{Z} \to 0\,,$$

which does not decompose into interval modules.

[9]The map $F \xrightarrow{\cdot 2} F$ denotes the multiplication-by-2 map, which is given by $x \mapsto 2x$.

3.4 Uniqueness of Interval Decompositions

In this section, we prove that interval decompositions of persistence modules that are pointwise free and finitely-generated over a PID are unique when they exist.

We fix a PID R and a persistence module f that is pointwise free and finitely-generated over R; we use the notation R and f for the rest of this chapter. We also assume, without loss of generality, that $f_0 = f_m = 0$.

Lemma 3.4. *Let $f : \{0, \dots, m\} \to R\text{-}Mod$ be a persistence module that is pointwise free and finitely-generated over R. Suppose that f admits an interval decomposition*

$$f \cong \bigoplus_k I_R^{b_k, d_k} .$$

Let $F = \mathrm{Frac}(R)$ be the field of fractions of R, and consider the persistence module $f \otimes F$, which is the persistence module over F that is given by

$$\begin{aligned} (f \otimes F)_a &= f_a \otimes F , \\ (f \otimes F)(a \leq b) &= f(a \leq b) \otimes \mathrm{id}_F . \end{aligned}$$

Then $\bigoplus_k I_F^{b_k, d_k}$ is an interval decomposition of $f \otimes F$.

Proof. Observe that $I_R^{b,d} \otimes F = I_F^{b,d}$. Because a tensor product distributes over a direct sum,

$$\begin{aligned} f \otimes F &\cong \left(\bigoplus_k I_R^{b_k, d_k} \right) \otimes F \\ &= \bigoplus_k (I_R^{b_k, d_k} \otimes F) \\ &= \bigoplus_k I_F^{b_k, d_k} , \end{aligned}$$

as desired. □

Theorem 3.5. *Let $f : \{0, \ldots, m\} \to R\text{-}Mod$ be a persistence module that is pointwise free and finitely-generated over R, and suppose f admits an interval decomposition. Then the associated multiset of intervals is unique.*

Proof. Let $F = \mathrm{Frac}(R)$ be the field of fractions of R. Suppose that $\bigoplus_{k=1}^{r} I_R^{b_k,d_k}$ and $\bigoplus_{\ell=1}^{r'} I_R^{b'_\ell,d'_\ell}$ are both interval decompositions of f. By Lemma 3.4, both $\bigoplus_{k=1}^{r} I_F^{b_k,d_k}$ and $\bigoplus_{\ell=1}^{r'} I_F^{b'_\ell,d'_\ell}$ are interval decompositions of $f \otimes F$. Theorem 2.1 guarantees that interval decompositions of persistence modules over F are unique up to permutation, so $r = r'$ and there exists a permutation π such that $(b_k, d_k) = (b'_{\pi(k)}, d'_{\pi(k)})$ for every k. □

Theorem 3.5 extends to persistence modules with coefficients in any integral domain.

3.5 Proof of Theorem 3.2: Necessity

We may now prove the main theoretical result, which we repeat here for ease of reference.

Theorem 3.2. *Let f be a persistence module that is pointwise free and finitely-generated over R. Then f splits into a direct sum of interval modules if and only if the cokernel of every structure map $f(a \leq b)$ is free.*

The condition that $f(a \leq b)$ has a free cokernel is equivalent to the following well-known conditions:

- The module f_b splits as a direct sum $I \oplus C$ for some submodule $C \subseteq f_b$, where I is the image of $f(a \leq b)$.
- In the language of the Smith normal form (see Section 3.A), the map $f(a \leq b)$ has unit elementary divisors. That is, if A is the matrix representation of $f(a \leq b)$ with respect to some pair of bases for f_a and f_b, and if $SAT = D$ is the Smith normal form

of A, then the nonzero diagonal entries of D are units.[10]

The proof of Theorem 3.2 has two halves: necessity and sufficiency. Necessity is straightforward and is established in Lemma 3.6. Sufficiency requires substantially greater effort and is established in Section 3.6.

Lemma 3.6 (Necessity)**.** *If f splits into a direct sum of interval modules, then the cokernel of every structure map $f(a \leq b)$ is free.*

Proof. Suppose that we can write $f = h^1 \oplus h^2 \oplus \cdots \oplus h^r$, where each h^k is an interval module. Fix a and b such that $0 \leq a \leq b \leq m$, and let $N = \{k : h_b^k \neq 0\}$ be the indices of the interval modules that are nonzero at b. Let $X = \{k \in N : h_a^k \neq 0\}$ (i.e., the set of indices of intervals including $[a, b]$) and $Y = \{k \in N : h_a^k = 0\}$ (i.e., the set of indices that are 0 at a). In particular, $\bigoplus_{k \in X} h_b^k = \mathrm{Im}(f(a \leq b))$, because X is precisely the set of indices for the intervals that contain $[a, b]$.

We note the following:

$$\begin{aligned} f_b &= \bigoplus_{k=1}^{r} h_b^k \\ &= \bigoplus_{k \in N} h_b^k \\ &= \bigoplus_{k \in X} h_b^k \oplus \bigoplus_{k \in Y} h_b^k \\ &= \mathrm{Im}(f(a \leq b)) \oplus \bigoplus_{k \in Y} h_b^k \,. \end{aligned}$$

Because f_b is free and submodules of free modules are free, $\mathrm{Coker}(f(a \leq b)) = f_b / \mathrm{Im}(f(a \leq b)) \cong \bigoplus_{k \in Y} h_b^k$ is free. □

[10] A "unit" is a ring element that has a multiplicative inverse.

3.6 Proof of Theorem 3.2: Sufficiency

We now build the framework to prove sufficiency. That is, we prove the following theorem.

Theorem 3.7 (Sufficiency)**.** *Let f be a persistence module that is pointwise free and finitely-generated over R, and suppose that the cokernel of every structure map $f(a \leq b)$ is free. Then f splits into a direct sum of interval modules.*

To prove Theorem 3.7, we proceed as follows. For each module f_a, we will study the "saecular lattice" [GH21] of f_a. The saecular lattice (see Definition 3.7) is a submodule lattice that consists of submodules generated by all images into and all kernels out of f_a. We will use this lattice of submodules to prove the existence of certain subspace complements, which will allow us to leverage the free-cokernel assumption to build structured families of bases through incremental extensions. These bases will yield a constructive proof of existence for interval decompositions. We use these ideas to formulate Algorithm 3.2 in Section 3.8.

Before building the necessary machinery for the proof of Theorem 3.7, we first provide intuition by illustrating the key ideas of our proof. See Figures 3.1 and 3.2 for illustrations.

For a non-negative integer n, let $\mathbf{n} = \{1, \ldots, n\}$. This definition holds even when $n = 0$; in this case, $\mathbf{0} = \emptyset$. We note that f is indexed by $\{0, 1, \ldots, m\}$, which can also be written as $\mathbf{m} \cup \{0\}$. Because $f_0 = f_m = 0$, every map of form $f(0 \leq p)$ or $f(p \leq m)$ is a zero map.

Definition 3.1. *Fix $a \in \mathbf{m}$, and let $x, y \in \mathbf{m} \cup \{0\}$. Define*

$$\mathrm{Ker}[a, y] = \begin{cases} \ker(f(a \leq y)), & a \leq y \\ 0, & \textit{otherwise} \end{cases}$$

and

$$\mathrm{Im}[x, a] = \begin{cases} \mathrm{Im}(f(x \leq a)), & x \leq a \\ f_a, & \textit{otherwise}. \end{cases}$$

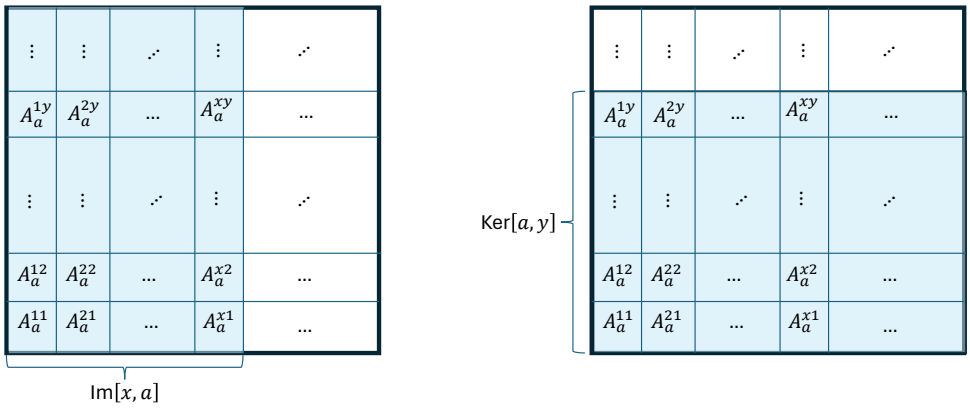


Figure 3.1: Given f_a, we illustrate how the summands given by $\{A_a^{ij}\}_{1\leq i,j\leq m}$ form (left) $\mathrm{Im}[x,a]$ and (right) $\mathrm{Ker}[a,y]$.

Note that $\mathrm{Ker}[a,m] = f_a$ because $f_m = 0$ and that $\mathrm{Ker}[a,0] = 0$ because $0 \leq a$. Thus, for each a, we have a nested sequence of submodules $0 = \mathrm{Ker}[a,0] \subseteq \cdots \subseteq \mathrm{Ker}[a,m] = f_a$ beginning with 0 and ending with f_a that are the kernels of the maps $f(a \leq x)$. Similarly, the submodules $\mathrm{Im}[x,a]$ form a nested sequence $0 = \mathrm{Im}[0,a] \subseteq \cdots \subseteq \mathrm{Im}[m,a] = f_a$ of submodules beginning with 0 and ending with f_a that are the images of the maps $f(x \leq a)$.

For each f_a, we construct a family $\{A_a^{ij}\}_{1\leq i,j\leq m}$ of submodules such that[11]

$$\bigoplus_{j\leq y} A_a^{ij} = \mathrm{Ker}[a,y]\,,$$
$$\bigoplus_{i\leq x} A_a^{ij} = \mathrm{Im}[x,a]\,.$$

In the context of PH, one can view A_a^{ij} as a submodule of f_a consisting of homology classes that are born at i and die at j. Note that $A_a^{ij} = 0$ when $a \notin [i,j)$. We construct our family of submodules $\{A_a^{ij}\}_{1\leq i,j\leq m}$ for each f_a such that when fixing $a, b \in \{0,\ldots,m\}$ with $a \leq b$,

[11]See Figure 3.1 for an illustration.

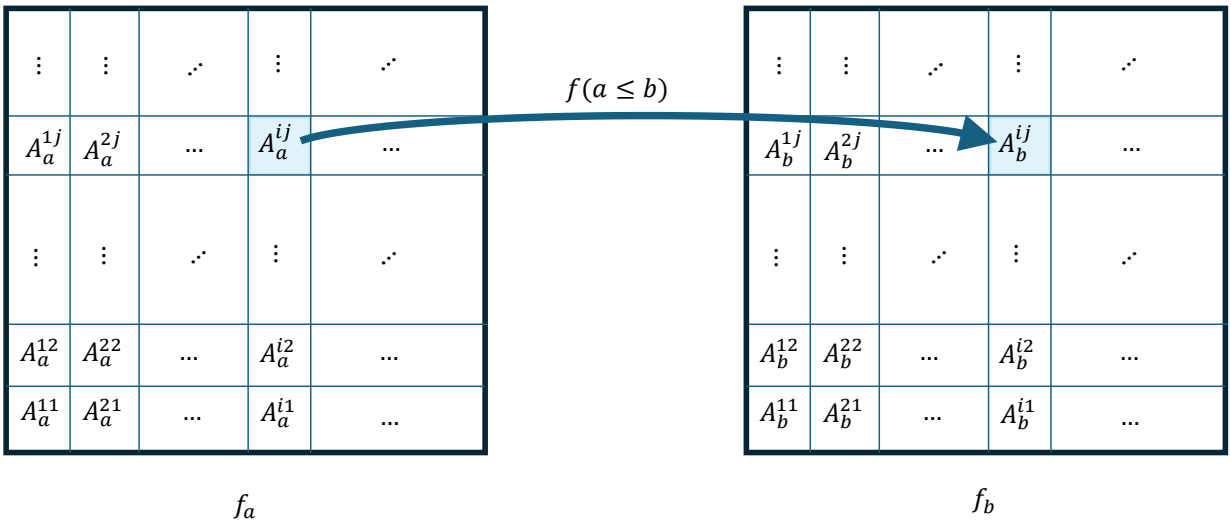


Figure 3.2: Given a and b with $a \leq b$, we illustrate that $f(a \leq b)(A_a^{ij}) \subseteq A_b^{ij}$.

we have [12]

$$f(a \leq b)(A_a^{ij}) \subseteq A_b^{ij}\,.$$

In this setting, the restricted map $f(a \leq b)|_{A_a^{ij}} : A_a^{ij} \to A_b^{ij}$ is an isomorphism when $a, b \in [i, j)$. If a or b lies outside $[i, j)$, then A_a^{ij} or A_b^{ij} is 0, so the map between them is a zero map.

This construction allows us to identify A_a^{ij} with A_b^{ij} whenever both of these modules are nonzero. Because a basis of A_a^{ij} yields a basis of A_b^{ij} (and vice versa), we can start with an appropriate basis on the left end of our persistence module and work our way right. That is, we can use a basis of f_a to construct a suitable basis of f_b. This yields a collection of bases (one for each f_a) that forms an interval decomposition.

3.6.1 A brief overview of lattice theory

This section uses many ideas from lattice theory. We provide a brief overview of the relevant lattice-theoretic ideas. For a more comprehensive background on lattice theory, see [Bir95,

[12] See Figure 3.2 for an illustration

Gra12].

Definition 3.2. *A* lattice $(\mathcal{L}, \leq)$ *is a poset that is closed under greatest lower bound and least upper bound. That is, every* $a, b \in \mathcal{L}$ *has a greatest lower bound and least upper bound in* $\mathcal{L}$*. The* join *of* a *and* b *is their greatest lower bound, which we denote* $a \wedge b$*. The* meet *of* a *and* b *is their least upper bound, which we denote* $a \vee b$*.*

An important lattice to consider is the order lattice of submodules. Given an R-module M, the *order lattice of submodules* (which we also call the "submodule lattice") of M is the set of submodules of M ordered by inclusion, where the meet and join operations are intersection and sum, respectively. We denote the order lattice of submodules of M by $\mathrm{Sub}(M)$.

An important property of submodule lattices is that they are modular (see [Gra12, Theorem 2.2.6(a)]).

Definition 3.3. *A lattice* $\mathcal{L}$ *is* modular *if, for any* $\ell_1, \ell_2, \ell_3 \in \mathcal{L}$ *such that* $\ell_1 \leq \ell_3$*, we have* $\ell_1 \vee (\ell_2 \wedge \ell_3) = (\ell_1 \vee \ell_2) \wedge \ell_3$*.*

We now review other relevant definitions.

Definition 3.4. *Let* $\mathcal{L}$ *be a lattice. We say that a subset* $\mathcal{S} \subseteq \mathcal{L}$ *is a* sublattice *if* $\mathcal{S}$ *is closed under meet and join.*

Definition 3.5. *Let* $\mathcal{L}$ *and* $\mathcal{M}$ *be lattices. A function* $\Lambda : \mathcal{L} \to \mathcal{M}$ *is a* lattice homomorphism *if* Λ *preserves meet and join. That is, for all* $a, b \in \mathcal{L}$*, we have* $\Lambda(a \wedge b) = \Lambda(a) \wedge \Lambda(b)$ *and* $\Lambda(a \vee b) = \Lambda(a) \vee \Lambda(b)$*.*

Definition 3.6. *A lattice* $\mathcal{C}$ *is a* chain *if it is totally ordered. That is, for every* $a, b \in \mathcal{C}$*, we have either* $a \leq b$ *or* $b \leq a$*.*

3.6.2 Saecular submodule lattices and homomorphisms

Definition 3.7. *The* saecular lattice *of f_a is the sublattice of* $\mathrm{Sub}(f_a)$ *generated by all submodules of the form* $\mathrm{Im}[x, a]$ *and* $\mathrm{Ker}[a, y]$. *We denote this sublattice by* $\mathrm{ImKer}(f_a)$.

Given $a \leq b$, we can relate submodules in $\mathrm{ImKer}(f_a)$ and $\mathrm{ImKer}(f_b)$ by pushing forward (i.e., applying $f(a \leq b)$ to submodules in $\mathrm{ImKer}(f_a)$) and pulling back (i.e., applying $f(a \leq b)^{-1}$ to submodules in $\mathrm{ImKer}(f_b)$). Lemmas 3.8, 3.9, and 3.10 express these relationships.

Lemma 3.8. *Fix $a, b \in \mathbf{m}$ such that $a \leq b$. The following hold for any $x, y \in \mathbf{m} \cup \{0\}$:*

$$f(a \leq b)(\mathrm{Im}[x, a]) = \mathrm{Im}[x, b] \cap \mathrm{Im}[a, b]\,, \tag{3.1}$$

$$f(a \leq b)(\mathrm{Ker}[a, y]) = \mathrm{Ker}[b, y] \cap \mathrm{Im}[a, b]\,. \tag{3.2}$$

Proof. We first show identity (3.1). Let L and R, respectively, denote the left-hand side and right-hand sides of the identity (3.1).

Suppose first that $x \leq a$. Then $\mathrm{Im}[x, b] \subseteq \mathrm{Im}[a, b]$, so

$$\begin{aligned} R &= \mathrm{Im}[x, b] \\ &= f(x \leq b)(f_x) \\ &= (f(a \leq b) \circ f(x \leq a))(f_x) \\ &= f(a \leq b)(\mathrm{Im}[x, a]) \\ &= L\,. \end{aligned}$$

Now suppose that $a \leq x \leq b$. Then $\mathrm{Im}[x, a] = f_a$, so $L = f(a \leq b)(f_a) = \mathrm{Im}[a, b]$. Because $\mathrm{Im}[a, b] \subseteq \mathrm{Im}[x, b]$, the right-hand side is also $\mathrm{Im}[a, b]$.

Finally, suppose that $a \leq b \leq x$. Then the left-hand side is again $f(a \leq b)(f_a) = \mathrm{Im}[a, b]$. Because $\mathrm{Im}[x, b] = f_b$, the right-hand side is also $\mathrm{Im}[a, b]$.

We now show identity (3.2). Let L and R, respectively, denote the left-hand side and right-hand sides of identity (3.2).

Suppose first that $y \leq a \leq b$. Then $\mathrm{Ker}[a, y]$ and $\mathrm{Ker}[b, y]$ are 0, by definition, and $0 = L = R$.

Now suppose that $a \leq y \leq b$. Then $\mathrm{Ker}[a, y] \subseteq \mathrm{Ker}[a, b]$, so $0 = f(a \leq b)(\mathrm{Ker}[a, y]) = L$. The right-hand side of (3.2) also vanishes, because $\mathrm{Ker}[b, y] = 0$, by definition.

Finally, suppose $a \leq b \leq y$. Then $L \subseteq \mathrm{Ker}[b, y]$, because $f(b \leq y)(f(a \leq b)(\mathrm{Ker}[a, y])) = f(a \leq y)(\mathrm{Ker}[a, y]) = 0$. Because $L \subseteq \mathrm{Im}[a, b]$ as well, it follows that $L \subseteq R$. To prove the opposite containment, suppose that $z' \in L$. Then $z' = f(a \leq b)(z)$ for some $z \in f_a$, which implies that $0 = f(b \leq y) \circ f(a \leq b)(z) = f(a \leq y)(z)$. Thus, $z \in \mathrm{Ker}[a, y]$, so $z' \in f(a \leq b)(\mathrm{Ker}[a, y]) = L$. Because z' is arbitrary, it follows that $R \subseteq L$. The desired conclusion follows. □

Lemma 3.9. *Fix $a, b \in \mathbf{m}$ such that $a \leq b$. The following holds for any $x, y \in \mathbf{m} \cup \{0\}$:*

$$f(a \leq b)\left(\mathrm{Im}[x, a] \cap \mathrm{Ker}[a, y]\right) = \mathrm{Im}[x, b] \cap \mathrm{Ker}[b, y] \cap \mathrm{Im}[a, b]\,. \tag{3.3}$$

Proof. First, consider $a \leq x$. In this case, $\mathrm{Im}[x, a] = f_a$ and $\mathrm{Im}[a, b] \subseteq \mathrm{Im}[x, b]$. Our task therefore reduces to showing that $f(a \leq b)(\mathrm{Ker}[a, y]) = \mathrm{Ker}[b, y] \cap \mathrm{Im}[a, b]$, which we proved in Lemma 3.8.

Now suppose that $y \leq b$. Then $f(a \leq b)(\mathrm{Ker}[a, y]) = 0$, so $f(a \leq b)(\mathrm{Im}[x, a] \cap \mathrm{Ker}[a, y]) = 0$. Therefore, the left-hand side of (3.3) vanishes. Because $\mathrm{Ker}[b, y] = 0$, by definition, the right-hand side of (3.3) also vanishes. This proves the desired result for $y \leq b$.

This leaves $x < a \leq b < y$. Because $x < a$, we have $\mathrm{Im}[x, b] \subseteq \mathrm{Im}[a, b]$, so we only need to show that

$$f(a \leq b)(\mathrm{Im}[x,a] \cap \mathrm{Ker}[a,y]) = \mathrm{Im}[x,b] \cap \mathrm{Ker}[b,y]\,. \tag{3.4}$$

Applying $f(a \leq b)$ to each term in the intersection on the left-hand side of (3.4) yields a subset of the corresponding term in the intersection on the right-hand side of (3.4), so the left-hand side is a subset of the right-hand side. For the other inclusion, fix an arbitrary element of $\mathrm{Im}[x,b] \cap \mathrm{Ker}[b,y]$, which can be expressed in the form $f(x \leq b)(\alpha)$ for some α. We note that $f(x \leq b)(\alpha) = f(a \leq b) \circ f(x \leq a)(\alpha)$, so it is sufficient to show that $f(x \leq a)(\alpha) \in \mathrm{Ker}[a,y]$. This holds because $f(a \leq y) \circ f(x \leq a)(\alpha) = f(b \leq y) \circ f(a \leq b) \circ f(x \leq a)(\alpha) = 0$. Thus, we see that $f(x \leq a)(\alpha) \in \mathrm{Im}[x,a] \cap \mathrm{Ker}[a,y]$, so $f(x \leq b)(\alpha) \in f(x \leq b)(\mathrm{Im}[x,a] \cap \mathrm{Ker}[a,y])$. □

Lemma 3.10. *Fix $a, b \in \mathbf{m}$ such that $a \leq b$. The following holds for any $x, y \in \mathbf{m} \cup \{0\}$:*

$$f(a \leq b)^{-1}(\mathrm{Im}[x,b] + \mathrm{Ker}[b,y]) = \mathrm{Im}[x,a] + \mathrm{Ker}[a,y] + \mathrm{Ker}[a,b]\,. \tag{3.5}$$

Proof. Throughout this proof, let L and R denote the left-hand side and right-hand side of Equation (3.5), respectively.

We consider three cases: (i) $a \leq x$, (ii) $y \leq b$, and (iii) $x \leq a \leq b \leq y$.

We first consider (i). In this case, the right-hand side is simply f_a, because $\mathrm{Im}[x,a] = f_a$. If $b \leq x$, then $\mathrm{Im}[x,b] = f_b$, which implies that the left-hand side is $f(a \leq b)^{-1}(\mathrm{Im}[x,b] + \mathrm{Ker}[b,y]) = f(a \leq b)^{-1}(f_b) = f_a$, as desired. Otherwise, with $a \leq x \leq b$, we see that

$$\begin{aligned} f(a \leq b)^{-1}(\mathrm{Im}[x,b] + \mathrm{Ker}[b,y]) &\supseteq f(a \leq b)^{-1}(\mathrm{Im}[x,b]) \\ &\supseteq f(a \leq b)^{-1}(\mathrm{Im}[a,b]) \\ &= f_a \end{aligned}$$

so the left-hand side is f_a as well.

Now consider (ii). In this case, $\mathrm{Ker}[b, y] = 0$ and $\mathrm{Ker}[a, y] = 0$. Therefore, we only need to show that

$$f(a \leq b)^{-1}(\mathrm{Im}[x, b]) = \mathrm{Im}[x, a] + \mathrm{Ker}[a, b]\,.$$

We subdivide (ii) into three subcases: $b \leq x$, $a \leq x \leq b$, and $x \leq a$.

If $b \leq x$, we see that $\mathrm{Im}[x, b] = f_b$ and $\mathrm{Im}[x, a] = f_a$; because $f(a \leq b)^{-1}(f_b) = f_a$, the desired expression holds.

If $a \leq x \leq b$, then the right-hand side is f_a, because $\mathrm{Im}[x, a] = f_a$. We note that

$$\begin{aligned} f(a \leq b)^{-1}(\mathrm{Im}[x, b]) &\supseteq f(a \leq b)^{-1}(\mathrm{Im}[a, b]) \\ &= f_a\,, \end{aligned}$$

so the left-hand side is f_a, and the desired expression again holds.

Consider $x \leq a$. First, we show that $L \supseteq R$. If we take $f(x \leq a)(w) \in \mathrm{Im}[x, a]$, then $f(a \leq b) \circ f(x \leq a)(w) = f(x \leq b)(w) \in \mathrm{Im}[x, b]$. We see that $\mathrm{Ker}[a, b] \subseteq f(a \leq b)^{-1}(\mathrm{Im}[x, b])$, so $L \supseteq R$. To see that $L \subseteq R$, we note that if $z \in f(a \leq b)^{-1}(\mathrm{Im}[x, b])$, then $f(a \leq b)(z) = f(x \leq b)(\alpha)$ for some $\alpha \in f_x$. Because $f(x \leq b) = f(a \leq b) \circ f(x \leq a)$, we see that $f(a \leq b)(z) = f(a \leq b) \circ f(x \leq a)(\alpha)$, which implies that $z = f(x \leq a)(\alpha) + \zeta$ for $\zeta \in \mathrm{Ker}[a, b]$, as desired.

Finally, we consider (iii). In this situation, $\mathrm{Ker}[a, b] \subseteq \mathrm{Ker}[a, y]$, so our desired expression reduces to

$$f(a \leq b)^{-1}(\mathrm{Im}[x, b] + \mathrm{Ker}[b, y]) = \mathrm{Im}[x, a] + \mathrm{Ker}[a, y]\,.$$

We first show that $L \supseteq R$. Indeed, if we take $f(x \leq a)(w) \in \mathrm{Im}[x, a]$, then $f(a \leq b) \circ f(x \leq a)(w) = f(x \leq b)(w) \in \mathrm{Im}[x, b]$. Additionally, if we take $z \in \mathrm{Ker}[a, y]$, then $f(a \leq b)(z) \in \mathrm{Ker}[b, y]$, as $f(b \leq y) \circ f(a \leq b)(z) = f(a \leq y)(z) = 0$. Thus, $L \supseteq R$.

To see that $L \subseteq R$, we take $z \in f(a \leq b)^{-1}(\mathrm{Im}[x, b] + \mathrm{Ker}[b, y])$. Note that $f(a \leq b)(z) = f(x \leq b)(\alpha) + \zeta$ for $\alpha \in f_x$ and $\zeta \in \mathrm{Ker}[b, y]$. As $f(x \leq b) = f(a \leq b) \circ f(x \leq a)$, the right-hand side is equal to $f(a \leq b) \circ f(x \leq a)(\alpha) + \zeta$. Applying $f(b \leq y)$ to both sides then gives $f(a \leq y)(z) = f(a \leq y) \circ f(x \leq a)(\alpha)$. This implies that $z = f(x \leq a)(\alpha) + \xi$ for $\xi \in \mathrm{Ker}[a, d]$, as desired.

As the above cases are now all settled, we are done. □

Having now established some basic properties of the saecular lattice, we introduce a second lattice.

Definition 3.8. *Let n_1 and n_2 be non-negative integers. Define $\mathbf{n_1 n_2} := \mathbf{n_1} \times \mathbf{n_2}$ as the product poset of $\mathbf{n_1}$ and $\mathbf{n_2}$. That is, given $(p, q), (p', q') \in \mathbf{n_1 n_2}$, we have that $p' \leq p$ and $q' \leq q$ if and only if $(p', q') \leq (p, q)$. A subset $S \subseteq \mathbf{n_1 n_2}$ is* down-closed *if $(p, q) \in S$ implies that $(p', q') \in S$ for all $(p', q') \leq (p, q)$. We write* $\mathrm{DownSets}(\mathbf{n_1 n_2})$ *for the poset of* down-closed subsets *of $\mathbf{n_1 n_2}$, ordered under inclusion. The poset* $\mathrm{DownSets}(\mathbf{n_1 n_2})$ *forms a lattice, where the meet and join operations are intersection and union, respectively.*

The DownSet and saecular lattices can be related with the following theorem, which is a consequence of a result in Birkhoff [Bir95].

Theorem 3.11. *For each $a \in \mathbf{m}$, there exists a (unique) lattice homomorphism*

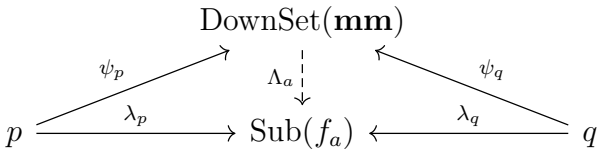
$$\Lambda_a : \mathrm{DownSet}(\mathbf{mm}) \to \mathrm{ImKer}(f_a)$$

such that $\mathbf{xm} \mapsto \mathrm{Im}[x, a]$ and $\mathbf{my} \mapsto \mathrm{Ker}[a, y]$.

Proof. This proof uses the modularity of submodule lattices.

Consider the lattice $\mathrm{Sub}(f_a)$ of submodules of f_a. Let p and q be the chains $p_0 \leq \cdots \leq p_m$ and $q_0 \leq \cdots \leq q_m$, respectively. Define lattice homomorphisms $\lambda_p : p \to \mathrm{Sub}(f_a)$ and $\lambda_q : q \to$

$\mathrm{Sub}(f_a)$ by $\lambda_p(p_x) = \mathrm{Im}[x, a]$ and $\lambda_q(q_y) = \mathrm{Ker}[a, y]$, respectively. By [Bir95, Section III.7, Theorem 9], there exists a unique lattice homomorphism $\Lambda_a : \mathrm{DownSet}(\mathbf{mm}) \to \mathrm{Sub}(f_a)$ such that the following diagram commutes:



In this diagram, $\psi_p : p \to \mathrm{DownSet}(\mathbf{mm})$ and $\psi_q : q \to \mathrm{DownSet}(\mathbf{mm})$ are defined by $\psi_p(p_x) = \mathbf{xm}$ and $\psi_q(q_y) = \mathbf{my}$, respectively. Moreover, the codomain of Λ_a can be restricted to $\mathrm{ImKer}f_a$ because the images of λ_p and λ_q are both in $\mathrm{ImKer}f_a$. We therefore have a unique lattice homomorphism $\Lambda_a : \mathrm{DownSet}(\mathbf{mm}) \to \mathrm{ImKer}(f_a)$, as desired. □

Definition 3.9. *We call* $\Lambda_a : \mathrm{DownSet}(\mathbf{mm}) \to \mathrm{ImKer}(f_a)$ *the* saecular homomorphism *correstricted to* a.

The saecular homomorphisms allow us to define a pair of commutative diagrams, which we introduce in Theorems 3.12 and 3.13. These diagrams have great utility; for example, we will use them to show that the direct image operator $H \mapsto f(a \leq b)(H)$ and the inverse image operator $G \mapsto f(a \leq b)^{-1}(G)$ induce homomorphisms between saecular lattices.

Theorem 3.12. *Fix* $a, b \in \mathbf{m}$ *such that* $a \leq b$. *The following diagram commutes:*

$$S \longmapsto S \cap \mathbf{am}$$

$$\begin{array}{ccccccc}
\mathbf{xm} & & \mathbf{my} & & \mathrm{DownSets}(\mathbf{mm}) & \xrightarrow{g} & \mathrm{DownSets}(\mathbf{mm}) \\
\downarrow & & \downarrow & & \downarrow{\scriptstyle \Lambda_a} & & \downarrow{\scriptstyle \Lambda_b} \\
\mathrm{Im}[x, a] & & \mathrm{Ker}[a, y] & & \mathrm{ImKer}(f_a) & \xrightarrow{h} & \mathrm{ImKer}(f_b)
\end{array}
\qquad
\begin{array}{cc}
\mathbf{xm} & \mathbf{my} \\
\downarrow & \downarrow \\
\mathrm{Im}[x, b] & \mathrm{Ker}[b, y]
\end{array}$$

$$H \longmapsto f(a \leq b)(H)$$

Proof. We first consider sets of the form $\mathbf{cm} \cap \mathbf{md}$. On one hand, Lemma 3.9 implies that

$$\mathbf{cm} \cap \mathbf{md} \xmapsto{\Lambda_a} \mathrm{Im}[c,a] \cap \mathrm{Ker}[a,d] \xmapsto{h} \mathrm{Im}[c,b] \cap \mathrm{Ker}[b,d] \cap \mathrm{Im}[a,b]\,.$$

On the other hand,

$$\mathbf{cm} \cap \mathbf{md} \xmapsto{g} \mathbf{cm} \cap \mathbf{md} \cap \mathbf{am} \xmapsto{\Lambda_b} \mathrm{Im}[c,a] \cap \mathrm{Ker}[b,d] \cap \mathrm{Im}[a,b]\,.$$

Therefore, $h \circ \Lambda_a(\mathbf{cm} \cap \mathbf{md}) = \Lambda_b \circ g(\mathbf{cm} \cap \mathbf{md})$. All sets in DownSets[$\mathbf{mm}$] can be formed as unions of sets of the form $\mathbf{cm} \cap \mathbf{md}$. Because g, Λ_a, and Λ_b are lattice homomorphisms and h preserves join (because the image of unions of sets is equal to the union of images of sets), it follows that $h \circ \Lambda_a$ and $\Lambda_b \circ g$ agree on every set in Downsets($\mathbf{mm}$), so the diagram commutes. □

We have a similar commutative diagram for $f(a \leq b)^{-1}$.

Theorem 3.13. *Fix $a, b \in \mathbf{m}$ such that $a \leq b$. The following diagram commutes:*

$$S \cup \mathbf{ma} \longleftarrow S$$

$$\begin{array}{ccccccc}
\mathbf{xm} & & \mathbf{my} & & \mathrm{DownSets}(\mathbf{mm}) & \xleftarrow{g} \mathrm{DownSets}(\mathbf{mm}) & & \mathbf{xm} & & \mathbf{my} \\
\downarrow & & \downarrow & & \downarrow{\scriptstyle \Lambda_a} & \downarrow{\scriptstyle \Lambda_b} & & \downarrow & & \downarrow \\
\mathrm{Im}[x,a] & & \mathrm{Ker}[a,y] & & \mathrm{ImKer}(f_a) & \xleftarrow{h} \mathrm{ImKer}(f_b) & & \mathrm{Im}[x,b] & & \mathrm{Ker}[b,y]
\end{array}$$

$$f(a \leq b)^{-1}(H) \longleftarrow H$$

Proof. First, we consider sets of the form $\mathbf{cm} \cup \mathbf{md}$. On one hand,

$$\mathbf{cm} \cup \mathbf{md} \xmapsto{g} \mathbf{cm} \cup \mathbf{md} \cup \mathbf{ma} \xmapsto{\Lambda_a} \mathrm{Im}[c,a] + \mathrm{Ker}[a,d] + \mathrm{Ker}[a,b]\,.$$

On the other hand, by Lemma 3.10, applying Λ_b followed by h yields

$$\mathbf{cm} \cup \mathbf{md} \xmapsto{\Lambda_b} \mathrm{Im}[c,b] + \mathrm{Ker}[b,d] \xmapsto{h} \mathrm{Im}[c,a] + \mathrm{Ker}[a,d] + \mathrm{Ker}[a,b]\,.$$

Thus, for sets of the form $\mathbf{cm} \cup \mathbf{md}$, we see that $\Lambda_a \circ g = h \circ \Lambda_b$.

Any set in DownSets($\mathbf{mm}$) can be written as an intersection of sets of the form $\mathbf{cm} \cup \mathbf{md}$. Because g, Λ_a, and Λ_b are lattice homomorphism and h preserves meet (because the preimage of an intersection is equal to the intersection of preimages), we see that $\Lambda_a \circ g = h \circ \Lambda_b$ for every set in DownSets($\mathbf{mm}$). □

Remark 3.2. *Theorems 3.12 and 3.13 imply that the function h (which is given by applying $f(a \leq b)$ and $f(a \leq b)^{-1}$, respectively) in both diagrams are lattice homomorphisms. This is because, in each diagram, the functions g, Λ_a, and Λ_b lattice homomorphisms and Λ_a and Λ_b are surjective.*

3.6.3 Complements in the saecular submodule lattices

In this section, we show that if the cokernel of $f(a \leq b)$ is free for all a and b such that $0 \leq a \leq b \leq m$, then for any f_c and $A, B \in \mathrm{ImKer}(f_c)$ such that $A \subseteq B$, the cokernel of the inclusion $A \subseteq B$ (i.e., the quotient B/A) is free. This implies that whenever we have an inclusion $A \subseteq B$ between elements of a saecular lattice, there exists a complementary submodule C such that $A \oplus C = B$.

Definition 3.10. *We say that C* complements *$A \subseteq B$ (or that C is a* complement *of A in B) whenever $A \oplus C = B$. In this situation, we say that A* admits *(or* has*) a complement in B.*

Remark 3.3. *Complements are not unique in general. For example, let $B = \mathbb{Z}^2$ and $A = \{(a,0)|a \in \mathbb{Z}\}$. Consider $C_1 = \{(0,b)|b \in \mathbb{Z}\}$ and $C_2 = \{(b,b)|b \in \mathbb{Z}\}$. Both C_1 and C_2 are complements of A in B.*

In Section 3.6.4, we will use these complements to explicitly construct interval decompositions.

For convenience, we adopt the notation

$$\mathrm{Coker}[x, a] = f_a / \mathrm{Im}[x, a] \,.$$

For the remainder of Section 3.6, we assume that the cokernel of every structure map $f(a \leq b)$ is free. Thus,

$$f_b \cong \mathrm{Im}[a, b] \oplus \mathrm{Coker}[a, b] \,.$$

Theorem 3.14. *Fix $a, b \in \mathbf{m}$ such that $a \leq b$. For any $x, y \in \mathbf{m} \cup \{0\}$, the submodule $\mathrm{Ker}[a, y] + \mathrm{Im}[x, a] \subseteq f_a$ has a complement in f_a.*

Proof. First note that if $y \leq a$, then $\mathrm{Ker}[a, y] + \mathrm{Im}[x, a] = \mathrm{Im}[x, a]$, which we know has a complement in f_a that is isomorphic to $\mathrm{Coker}[x, a]$.

If $a \leq x$, then $\mathrm{Ker}[a, y] + \mathrm{Im}[x, a] = f_a$ has a trivial complement $C = 0$ in f_a.

We now need to prove the result for $x \leq a \leq y$. Because f_a is free, it is sufficient to show that the quotient $\frac{f_a}{\mathrm{Ker}[a,y]+\mathrm{Im}[x,a]}$ is free.

Let C be a complement of $\mathrm{Im}[a, y]$ in f_y. Then we have a chain of isomorphisms:

$$\begin{aligned}
\mathrm{Coker}[x, y] &\cong \frac{f_y}{\mathrm{Im}[x, y]} \\
&= \frac{\mathrm{Im}[a, y] \oplus C}{\mathrm{Im}[x, y]} \\
&\cong \frac{\mathrm{Im}[a, y]}{\mathrm{Im}[x, y]} \oplus C && (3.6) \\
&\cong \frac{f_a}{\mathrm{Ker}[a, y] + \mathrm{Im}[x, a]} \oplus C \,. && (3.7)
\end{aligned}$$

The last two isomorphisms (i.e., (3.6) and (3.7)) must be proved.

The isomorphism (3.6) holds because $\mathrm{Im}[x, y] \subseteq \mathrm{Im}[a, y]$.

For the isomorphism (3.7), recall that $x \leq a \leq y$ and that the following hold by definition:

- $\mathrm{Ker}[a, y] = \ker(f(a \leq y))$;
- $\mathrm{Im}[x, a] = \mathrm{Im}(f(x \leq a)) = f(x \leq a)(f_a)$;
- $\mathrm{Im}[x, y] = \mathrm{Im}(f(x \leq y)) = f(x \leq y)(f_x) = f(a \leq y)(f(x \leq a)(f_x))$;
- $\mathrm{Im}[a, y] = \mathrm{Im}(f(a \leq y)) = f(a \leq y)(f_a)$.

We now show the following chain of isomorphisms:

$$\frac{\mathrm{Im}[a, y]}{\mathrm{Im}[x, y]} \cong \frac{f(a \leq y)(f_a)}{f(a \leq y)(f(x \leq a)(f_x))} \tag{3.8}$$

$$\cong \frac{f(a \leq y)(f_a)}{f(a \leq y)\Big(f(x \leq a)(f_x) + \ker(f(a \leq y))\Big)} \tag{3.9}$$

$$\cong \frac{f_a / \ker(f(a \leq y))}{\Big(f(x \leq a)(f_x) + \ker(f(a \leq y))\Big) / \ker(f(a \leq y))} \tag{3.10}$$

$$\cong \frac{f_a}{f(x \leq a)(f_x) + \ker(f(a \leq y))} \tag{3.11}$$

$$\cong \frac{f_a}{\mathrm{Ker}[a, y] + \mathrm{Im}[x, a]}\,. \tag{3.12}$$

The isomorphism (3.9) is a follows from the inequality

$$f(a \leq y)(f(x \leq a)(f_x)) = f(a \leq y)\Big(f(x \leq a)(f_x) + \ker(f(a \leq y))\Big)\,.$$

The isomorphism (3.10) is a result of the first isomorphism theorem.[13] The isomorphism

[13]The first isomorphism theorem states that given an R-module homomorphism $\psi : M_1 \to M_2$, we have the isomorphism $\psi(M_1) \cong M_1 / \ker(\psi)$. See [DF04, Section 10.2, Theorem 4.1].

in (3.11) is a result of the third isomorphism theorem.[14] This establishes the isomorphism (3.7). The isomorphism (3.7) shows that $\frac{f_a}{\mathrm{Ker}[a,y]+\mathrm{Im}[x,a]}$ can be viewed as a submodule of $\mathrm{Coker}[x,y]$, which is free by hypothesis. Because submodules of free modules over PIDs are free, $\frac{f_a}{\mathrm{Ker}[a,y]+\mathrm{Im}[x,a]}$ is free. Therefore, the inclusion $\mathrm{Ker}[a,y]+\mathrm{Im}[x,a] \subseteq f_a$ has a free cokernel, so the submodule $\mathrm{Ker}[a,y]+\mathrm{Im}[x,a]$ has a complement in f_a, as desired. □

Using Theorem 3.14, as well as some lemmas that we state and prove shortly, we have the following result.

Theorem 3.15. *Let $a \in \mathbf{m}$ and $L, H \in \mathrm{ImKer}(f_a)$ such that $L \subseteq H$. Then L has a complement in H.*

We prove Theorem 3.15 using the results of Lemmas 3.16, 3.17, 3.18, and 3.19.

Remark 3.4. *In our proof of Lemma 3.16, we will use the fact that a finitely-generated module over a PID is free if and only if it is torsion-free.*

Lemma 3.16. *Let G be a free and finitely-generated module over R, and let $A \subseteq G$. Then the following are equivalent:*

(1) A has a complement in G;

(2) G/A is torsion-free;

(3) if $mz \in A$ for $z \in G$ and nonzero $m \in R$, then $z \in A$.

Proof. The equivalence between (1), (2), and (3) follows from standard properties of finitely-generated modules over a PID.

To see that (1) implies (2), we note that if $G = A \oplus B$, then $G/A \cong B$, which is free (and

[14]The third isomorphism theorem states that given an R-module M with submodules $M_1, M_2 \subseteq M$ such that $M_1 \subseteq M_2 \subseteq M_3$, we have the isomorphism $M/M_2 \cong \frac{M/M_1}{M_2/M_1}$. See [DF04, Section 10.3, Theorem 4.1].

therefore torsion-free), as it is a submodule of a free module over a PID.

We now show that (2) implies (1). Let $\phi : G \to G/A$ denote the projection map. Let $\{\tilde{g}_i\}_{i\in I}$ be a basis of G/A (we can choose such a basis because G/A is torsion-free and therefore free), and choose $g_i \in \phi^{-1}(\tilde{g}_i)$. Let $B \subseteq G$ be the submodule generated by $\{g_i\}_{i\in I}$. We claim that $A \oplus B = G$. Any nontrivial linear combination of the elements g_i yields a nontrivial linear combination of elements $\tilde{g}_i$ under ϕ; this linear combination is nonzero because the $\tilde{g}_i$ are linearly independent. Therefore, no nontrivial linear combination of elements g_i lies in $\ker(\phi) = A$, so $A \cap B = 0$. We now show that $A + B = G$. Take $g \in G$, and write $\phi(g) = \sum_{i\in I} c_i\tilde{g}_i = \sum_{i\in I} c_i\phi(g_i)$ for some $c_i \in R$. From this, we see that $g = \sum_{i\in I} c_i g_i + a$ for $a \in \ker(\phi) = A$, as desired.

We now show that (2) implies (3). Suppose that $mz \in A$ with nonzero $m \in R$ and $z \in G$. If $z \in G \setminus A$, then $z + A \in G/A$ is nonzero, but $mz + A = 0$. This implies that $z + A$ is in the torsion of G/A, which is a contradiction.

We now show that (3) implies (2). If G/A has torsion, then there is a nonzero $z + A$ such that $mz + A = 0$ for a nonzero $m \in R$. This implies that $z \in G \setminus A$ and that $mz \in A$. This shows the desired result by contrapositive, thus completing the proof. □

Lemma 3.17. *If the submodules $H, K \subseteq f_a$ each admit complements, then so does $H \cap K$.*

Proof. Suppose that $mz \in H \cap K$ for nonzero $m \in R$ and $z \in f_a$. Because $mz \in H$ and H admits a complement, we see by Lemma 3.16 that $z \in H$. Similarly, $z \in K$, which implies that $z \in H \cap K$. Therefore, by Lemma 3.16, $H \cap K$ has a complement in f_a. □

Lemma 3.18. *If $L \subseteq H \subseteq f_a$ and L admits a complement in f_a, then L admits a complement in H.*

Proof. By Lemma 3.16, it is enough to show that H/L is free. Because $H/L \subseteq f_a/L$ and

f_a/L is free, we must have that H/L is free as well. □

Lemma 3.19. *Every element of the saecular lattice* $\mathrm{ImKer}(f_a)$ *can be expressed as an intersection of submodules of the form* $\mathrm{Im}[c,a] + \mathrm{Ker}[a,d]$ *for some choice of* $c, d \in \mathbf{m} \cup \{0\}$.

Proof. Submodules of the form $\mathrm{Im}[c,a] + \mathrm{Ker}[a,d]$ are precisely the image of subsets of the form $\mathbf{cm} \cup \mathbf{md}$ under Λ_a. Because any set in $\mathrm{DownSets}(\mathbf{mm})$ can be written as an intersection of sets of the form $\mathbf{cm} \cup \mathbf{md}$ (noting that c or d can be 0) and Λ_a is a lattice homomorphism, it follows that any submodule of f_a can be written as an intersection of submodules of the form $\mathrm{Im}[c,a] + \mathrm{Ker}[a,m]$. □

Lemmas 3.16, 3.17, 3.18, 3.19, and Theorem 3.14 allow us to prove Theorem 3.15.

Proof of Theorem 3.15. By Lemma 3.19, every $L \in \mathrm{ImKer}(f_a)$ can be represented as an intersection of submodules of the form $\mathrm{Im}[c,a] + \mathrm{Ker}[a,d]$. By Theorem 3.14 and Lemma 3.17, L admits a complement in f_a. By Lemma 3.18, L has a complement in H, as desired. □

3.6.4 Existence of interval decompositions

We now use the saecular lattice (see Definition 3.7) of each f_a and the existence of complements in the saecular lattice (see Theorem 3.15) to prove the existence of interval decompositions. We once again assume that all structure maps $f(a \leq b)$ have free cokernel.

Our argument proceeds as follows. First, we use the results of Section 3.6.3 to obtain a direct-sum decomposition $f_a \cong \bigoplus_{i,j} A_a^{ij}$ for all a, where each A_a^{ij} is a complement in the saecular lattice (see Theorem 3.22). We then modify the summands in each decomposition so that $f(a \leq b)(A_a^{ij}) \in \{A_b^{ij}, 0\}$ whenever $a \leq b$ (see Lemma 3.24). The resulting family of summands A_a^{ij} then naturally combines to form an interval decomposition of f.

Lemma 3.20. *Suppose that we have a diagram of inclusions of down-closed sets in* **mm***:*

$$\begin{array}{ccc} Y & \hookrightarrow & T \\ \uparrow & & \uparrow \\ X & \hookrightarrow & S\,, \end{array}$$

where[15] $Y = X \sqcup \{(i,j)\}$ *and* $T = S \sqcup \{(i,j)\}$*. Let* A_a^{ij} *be a complement of* $\Lambda_a(X)$ *in* $\Lambda_a(Y)$*. Then* A_a^{ij} *is also a complement of* $\Lambda_a(S)$ *in* $\Lambda_a(T)$*.*

Proof. Note that A_a^{ij} is a submodule of $\Lambda_a(T)$. Moreover, $Y \cap S = X$ and $Y \cup S = T$. From this, we have

$$\begin{aligned} \frac{\Lambda_a(Y)}{\Lambda_a(X)} &= \frac{\Lambda_a(Y)}{\Lambda_a(Y \cap S)} \\ &= \frac{\Lambda_a(Y)}{\Lambda_a(Y) \cap \Lambda_a(S)} \\ &\cong \frac{\Lambda_a(Y) + \Lambda_a(S)}{\Lambda_a(S)} \\ &= \frac{\Lambda_a(Y \cup S)}{\Lambda_a(S)} \\ &= \frac{\Lambda_a(T)}{\Lambda_a(S)}\,. \end{aligned} \tag{3.13}$$

The isomorphism in (3.13) arises from the second isomorphism theorem.[16] We therefore have the commutative diagram

$$\begin{array}{ccc} & & \Lambda_a(Y)/\Lambda_a(X) \\ & \nearrow & \cong\downarrow \\ A_a^{ij} & \longrightarrow & \Lambda_a(T)/\Lambda_a(S)\,, \end{array}$$

so the map $A_a^{ij} \to \Lambda_a(T)/\Lambda_a(S)$ is injective. Therefore, $\Lambda_a(S) \cap A_a^{ij} = 0$. Moreover, $\Lambda_a(S) +$

[15]Given a set $\mathcal{S}$ and subsets $\mathcal{A}_1, \mathcal{A}_2 \subseteq \mathcal{S}$, the set $\mathcal{S}$ is a disjoint union of $\mathcal{A}_1$ and $\mathcal{A}_2$ (which we write $\mathcal{A}_1 \sqcup \mathcal{A}_2 = \mathcal{S}$) if $\mathcal{A}_1 \cup \mathcal{A}_2 = \mathcal{S}$ and $\mathcal{A}_1 \cap \mathcal{A}_2 = \emptyset$.

[16]The second isomorphism theorem states that given an R-module M and submodules $S, T \subseteq M$, there is an isomorphism $\psi : S/(S \cap T) \to (S+T)/T$ given by $\psi(a + (S \cap T)) = a + T$. See [DF04, Section 10.2, Theorem 4.2].

$A_a^{ij} = \Lambda_a(T)$, because

$$\begin{aligned}\Lambda_a(S) + A_a^{ij} &= \Lambda_a(X) + A_a^{ij} + \Lambda_a(S) \\ &= \Lambda_a(Y) + \Lambda_a(S) \\ &= \Lambda_a(Y \cup S) \\ &= \Lambda_a(T) \ .\end{aligned}$$

As a result, A_a^{ij} is a complement of $\Lambda_a(S)$ in $\Lambda_a(T)$, as desired. □

For each f_a, the submodule A_a^{ij} in Lemma 3.20 can identified with $(i, j) \in \mathbf{mm}$. In fact, the set of A^{ij} are summands of f_a. One can also use these summands of f_a to construct any submodule in $\mathrm{ImKer}(f_a)$.

Theorem 3.20 tells us that every complement of $\Lambda_a(X) \subseteq \Lambda_a(Y)$ is also a complement of $\Lambda(S) \subseteq \Lambda(T)$ when X and Y are small enough with respect to inclusion (i.e., satisfy the relation in Lemma 3.20).

For the remainder of Section 3.6, we fix a family of complements, with X and Y as small as possible. That is, for each $a, i, j \in \mathbf{m}$, choose a direct-sum decomposition

$$\Lambda_a(Y) = \Lambda_a(X) \oplus A_a^{ij} \tag{3.14}$$

in which Y and X are defined $Y = \mathbf{ij} = \mathbf{im} \cap \mathbf{mj}$ and $X = Y \setminus \{(i, j)\}$.

Our next technical result is Lemma 3.21. Viewed at a high level, this result states that when a space of the form $\mathrm{Im}[x, a]$ and a space of the form $\mathrm{Ker}[a, y]$ are each large enough to contain A_a^{ij}, then so is their intersection. However, decrementing x and y by a single unit can make the image and kernel spaces small enough so that they do not include any nonzero element of A_a^{ij}, even when we take the sum of the two spaces.

Lemma 3.21. *For each* $(i,j) \in \mathbf{mm}$*, we have*

$$0 = A_a^{ij} \cap \big(\mathrm{Im}[i-1,a] + \mathrm{Ker}[a,j-1]\big)\,, \tag{3.15}$$

$$A_a^{ij} = A_a^{ij} \cap \big(\mathrm{Im}[i,a] \cap \mathrm{Ker}[a,j]\big) \qquad \textit{(equivalently, } A_a^{ij} \subseteq \mathrm{Im}[i,a] \cap \mathrm{Ker}[a,j]\,\textit{.)} \tag{3.16}$$

Proof. Define down-closed sets $X \subseteq Y$ and $S \subseteq T$ such that

$$\begin{aligned} X &= Y \setminus \{(i,j)\} = Y \cap S\,, \\ Y &= \mathbf{ij} = \mathbf{im} \cap \mathbf{mj}\,, \\ S &= \{(p,q) \in \mathbf{mm} : p < i \text{ or } q < j\} = (\mathbf{i}-\mathbf{1})\mathbf{m} \cup \mathbf{m}(\mathbf{j}-\mathbf{1})\,, \\ T &= S \cup \{(i,j)\} = S \cup Y\,. \end{aligned}$$

Then A_a^{ij} complements $\Lambda_a(S) \subseteq \Lambda_a(T)$, where $\Lambda_a(S) = \mathrm{Im}[i-1,a]+\mathrm{Ker}[a,j-1]$, so Equation (3.15) follows. Likewise, A_a^{ij} complements $\Lambda_a(X) \subseteq \Lambda_a(Y)$, where $\Lambda_a(Y) = \mathrm{Im}[i,a]\cap\mathrm{Ker}[a,j]$. Therefore, $A_a^{ij} \subseteq \mathrm{Im}[i,a] \cap \mathrm{Ker}[a,j]$, so Equation (3.16) follows. □

Theorem 3.22. *For every down-closed subset* $S \subseteq \mathbf{mm}$*, we have*

$$\Lambda_a(S) = \bigoplus_{(i,j)\in S} A_a^{ij}\,.$$

Proof. Choose a sequence of down-closed sets $\emptyset = S_0 \subseteq S_1 \subseteq \cdots \subseteq S_n = S$ such that each set difference $S_p \setminus S_{p-1}$ is a singleton $\{(i_p,j_p)\}$. The submodule $A_a^{i_p,j_p}$ complements $\Lambda_a(S_{p-1}) \subseteq \Lambda_a(S_p)$ for all p, by Lemma 3.20. Therefore,

$$\Lambda_a(S) = \bigoplus_p A_a^{i_p,j_p} = \bigoplus_{(i,j)\in S} A_a^{ij}.$$

□

In the special case where f is a persistence module from PH, one can view each A_a^{ij} as homology classes that are born at i and die at j. This interpretation is valid by Lemma 3.21, because $A_a^{ij} \subseteq \mathrm{Ker}[a, j]$, $A_a^{ij} \cap \mathrm{Ker}[a, j-1] = 0$, $A_a^{ij} \subseteq \mathrm{Im}[i, a]$, and $\mathrm{Im}[i-1, a] \cap A_a^{ij} = 0$. Lemmas 3.23 and 3.24 provide additional context for this interpretation. They also provide the necessary ingredients to prove Theorem 3.7.

Lemma 3.23. *If $a \notin [i, j)$, then $A_a^{ij} = 0$.*

Proof. On one hand, if $a < i$, then $\mathrm{Im}[i, a] = f_a = \mathrm{Im}[i-1, a]$. Because we also have $\mathrm{Im}[i-1, a] \cap A_a^{ij} = 0$, it follows that $A_a^{ij} = 0$.

On the other hand, $a \geq j$ implies that $\mathrm{Ker}[a, j] = 0$. Because $A_a^{ij} \subseteq \mathrm{Ker}[a, j]$, it follows that $A_a^{ij} = 0$. □

Lemma 3.24. *Suppose that $a, b \in [i, j)$, where $a \leq b$. Define $Y = \mathbf{ij}$ and $X = Y \setminus \{(i, j)\}$. Then*

$$\Lambda_b(Y) = \Lambda_b(X) \oplus f(a \leq b)(A_a^{ij}) \,.$$

That is, the image $f(a \leq b)(A_a^{ij})$ satisfies the same condition (namely, Equation (3.14)) that we impose on A_b^{ij} with b in place of a. Furthermore, $f(a \leq b)$ restricts to an isomorphism $A_a^{ij} \to f(a \leq b)(A_a^{ij})$.

Proof. Define $X'\Lambda_a(X)$, $Y'\Lambda_a(Y)$, $X''\Lambda_b(X)$, and $Y''\Lambda_b(Y)$.

Let g be the map $S \mapsto S \cap \mathbf{am}$, and let h be the direct image-operator $H \mapsto f(a \leq b)(H)$. By hypothesis $Y' = X' \oplus A_a^{ij}$. First, we wish to show that $Y'' = X'' \oplus h(A_a^{ij})$.
Because $g(X) = X$ and $g(Y) = Y$, it follows from Theorem 3.12 that $h(X') = (h \circ \Lambda_a)(X) = (\Lambda_b \circ g)(X) = X''$ and $h(Y') = (h \circ \Lambda_a)(Y) = (\Lambda_b \circ g)(Y) = Y''$. Additionally, because the kernel of $f(a \leq b)|_{Y'}$ equals $\mathrm{Ker}[a, b] \cap Y' = \Lambda_a(\mathbf{mb}) \cap \Lambda_a(Y) = \Lambda_a(\mathbf{mb} \cap Y)$, which is a

submodule of $\Lambda_a(X) = X'$, we have

$$Y'' = f(a \leq b)(Y') = f(a \leq b)(X' \oplus A_a^{ij}) = h(X') \oplus h(A_a^{ij}) = X'' \oplus h(A^{ij}),$$

as desired. In addition, because the kernel of $f(a \leq b)|_{Y'}$ is a submodule of X', it follows that $f(a \leq b)$ restricts to an isomorphism $A_a^{ij} \to h(A_a^{ij})$. The desired conclusion follows. $\square$

To prove Theorem 3.7, we first introduce the notion of generalized interval modules. We will show that they decompose as a direct sum of interval modules.

Definition 3.11. *Let R be a commutative ring. A* generalized interval module *over R is a persistence module $J^{ij} : \{0, 1, \ldots, m\} \to R\text{-}Mod$ such that $J_a^{ij} = 0$ for $a \notin [i, j)$ and $J^{ij}(a \leq b) : J_a^{ij} \to J_b^{ij}$ is an isomorphism whenever $a, b \in [i, j)$.*

Proposition 3.25. *Let J^{ij} be a generalized interval module. Then J^{ij} decomposes as a direct sum of interval modules.*

Proof. Choose any basis β_i of J_i^{ij}. For $a \notin [i, j)$, let $\beta_a = \emptyset$; for $a \in (i, j)$, let $\beta_a = J^{ij}(i \leq a)(\beta_i)$.

It suffices to show that, for all a, we have that β_a is a basis of J_a^{ij} and that $J^{ij}(a \leq b)$ maps $\beta_a \setminus \ker(J^{ij}(a \leq b))$ injectively into β_b for all $a \leq b$.

When a or b are outside $[i, j)$, this is trivial, as one of β_a or β_b must be empty. When $i \leq a \leq b < j$, we note that $J^{ij}(a \leq b)$ is an isomorphism. Therefore, each β_a for $a \in [i, j)$ is a basis and $J^{ij}(a \leq b)$ is a bijection between β_a and β_b, which satisfies our desired claim. $\square$

We conclude this section by proving Theorem 3.7, thus completing the proof of Theorem 3.2.

Theorem 3.7. It suffices to show that f decomposes as a direct sum of generalized interval modules, which decompose into interval modules by Proposition 3.25. By Lemmas 3.23 and

3.24, we can assume without loss of generality that

$$A_a^{ij} = \begin{cases} f(i \le a)(A_i^{ij})\,, & a \in [i,j)\,, \\ 0\,, & \text{otherwise}\,. \end{cases}$$

Additionally, Lemma 3.24 implies that $f(a \le b)|_{A_a^{ij}} : A_a^{ij} \to A_b^{ij}$ is an isomorphism whenever $a, b \in [i,j)$ and is the zero map otherwise. It therefore follows that, for each i and j, there is a persistence submodule A^{ij} such that $(A^{ij})_a = A_a^{ij}$ and $A^{ij}(a \le b) = f(a \le b)|_{A_a^{ij}}$. Thus, A^{ij} is a generalized interval module. For each a, we have

$$\begin{aligned} f_a &= \bigoplus_{ij} A_a^{ij} \qquad \text{(by Theorem 3.22)} \\ &= \bigoplus_{ij} (A^{ij})_a\,. \end{aligned}$$

Additionally, for each a and b such that $a \le b$, we have

$$\begin{aligned} f(a \le b) &= \bigoplus_{i,j} f(a \le b)|_{A_a^{ij}} \\ &= \bigoplus_{i,j} A^{ij}(a \le b)\,. \end{aligned}$$

Therefore, f is a direct sum of the generalized interval modules A^{ij}, as desired. □

3.7 Algorithm for Computing an Interval Decomposition

In this section, we provide a simple presentation of our algorithm to compute interval decompositions for persistence modules that are pointwise free and finitely-generated over R. We provide a detailed description in the language of matrix algebra in Section 3.8.

Let $f : \{0\} \cup \mathbf{m} \to R\text{-Mod}$ be a persistence module that is pointwise free and finitely-generated over R. Recall that $f_0 = f_m = 0$ by convention (see Section 3.6). Our task is to

define a sequence of bases $\beta_i \subseteq f_i$ such that $f(i \leq j)$ maps $\beta_i \setminus \mathrm{Ker}[i,j]$ injectively into β_j for all $i \leq j$.

Suppose that the cokernel of $f(i \leq j)$ is free for each $i, j \in \{0, \ldots, m\}$ such that $i \leq j$. For each $i \leq j$, let K_j^i be a complement of

$$\Big(\mathrm{Im}[i-1,i] + \mathrm{Ker}[i,j-1]\Big) \cap \mathrm{Ker}[i,j]$$

in $\mathrm{Ker}[i,j]$. That is,

$$\Big(\Big(\mathrm{Im}[i-1,i] + \mathrm{Ker}[i,j-1]\Big) \cap \mathrm{Ker}[i,j]\Big) \oplus K_j^i = \mathrm{Ker}[i,j]\,.$$

Such complements exist by Theorem 3.15. Let γ_j^i be a basis of K_j^i.

We now present Algorithm 3.1, which computes an interval decomposition.

Algorithm 3.1 Computation of an interval decomposition.

Require: A persistence module $f : \{0\} \cup \mathbf{m} \to R\text{-Mod}$ that is pointwise free and finitely-generated over R. For each $i, j \in \{0\} \cup \mathbf{m}$ such that $i \leq j$, we require the map $f(i \leq j)$ to have a free cokernel.

Ensure: Returns an interval decomposition

1: $\beta_0 \leftarrow \emptyset$
2: **for** $i = 1, \ldots, m-1$ **do**
3: $\quad \beta_i \leftarrow f(i-1 \leq i)(\beta_{i-1})$
4: $\quad \beta_i \leftarrow \beta_i \setminus \{0\}$
5: $\quad$ **for** $j = i+1, \ldots, m$ **do**
6: $\quad\quad \beta_i \leftarrow \beta_i \cup \gamma_j^i$
7: **return** $\beta_1, \ldots, \beta_{m-1}$

Theorem 3.26. *Let $\beta_1, \cdots, \beta_{m-1}$ be bases that are returned by Algorithm 3.1. Each β_i is a basis of f_i that contains a basis of each* $\mathrm{Ker}[i,j]$ *and* $\mathrm{Im}[k,i]$. *Moreover, the bases $\beta_i \subseteq f_i$ yield an interval decomposition of f.*

Proof. We first show that each basis β_i that Algorithm 3.1 returns contains a basis of each

$\mathrm{Ker}[i,j]$ and $\mathrm{Im}[k,i]$. We proceed by induction.

Let $i = 1$. Because $f_0 = 0$, we have that $f(0 \leq 1)(\beta_0) = \emptyset$. For each $j = 1, \ldots, m$ in the inner loop (see line 5 of Algorithm 3.1), we append to β_1 a basis γ_j^1 of a complement K_j^1 of $\mathrm{Ker}[1, j-1]$ in $\mathrm{Ker}[1,j]$. This process gives a basis of $\mathrm{Ker}[1,m] = f_1$ (because $f_m = 0$). Additionally, note that $\mathrm{Im}[0,1] = 0$ and $\mathrm{Im}[k,1] = f_1$ for $k \geq 1$. Whether $\mathrm{Im}[k,1] = 0$ (when $k = 0$) or $\mathrm{Im}[k,1] = f_1$ (when $k \geq 1$), we can a find subset of $S \subseteq \beta_1$ (either $\emptyset$ or β_1) such that S spans $\mathrm{Im}[k,1]$. Thus, β_1 is a basis of f_1 that contains a basis of every $\mathrm{Im}[i,j]$ and $\mathrm{Im}[k,i]$. This completes the base case.

Now suppose that $\beta_1, \ldots, \beta_{i-1}$ are bases of $f_1, \ldots, f_{i-1}$, respectively, that contain bases of the relevant images and kernels. We will show that β_i is a basis of f_i that contains a basis of each kernel and image.

We first show that $f(i-1 \leq i)(\beta_{i-1}) \setminus \{0\}$ is linearly independent (as a multiset). To see this, we first note that β_{i-1} is a basis of f_{i-1} and contains a basis of each $\mathrm{Ker}[i-1,j]$. Consider $\kappa_{i-1} = \beta_{i-1} \cap \mathrm{Ker}[i-1,i]$ and $\kappa_{i-1}^{\perp} = \beta_{i-1} \setminus \kappa_{i-1}$. We note that $\mathrm{Ker}(f(i-1 \leq i)) = \mathrm{span}(\kappa_{i-1})$, which implies that $f(i-1 \leq i)(\kappa_{i-1}^{\perp})$ (which is precisely $f(i-1 \leq i)(\beta_{i-1}) \setminus \{0\}$) is linearly independent. Moreover, $f(i-1 \leq i)(\beta_{i-1}) \setminus \{0\}$ spans $\mathrm{Im}[i-1,i]$, so it forms a basis of $\mathrm{Im}[i-1,i]$

For each $j = i+1, \ldots, m$ in the inner loop (see line 5 of Algorithm 3.1), adding the elements of γ_j^i to β_i preserves linear independence. This is the case because the union of linearly independent sets of complementary spaces is linearly independent. In fact, on the jth iteration, we are extending β_i from a basis of $\mathrm{Im}[i-1,i] + \mathrm{Ker}[i,j-1]$ to a basis of $\mathrm{Im}[i-1,i] + \mathrm{Ker}[i,j]$. After the $j = m$ iteration of the inner loop, we have a basis of $\mathrm{Im}[i-1,i] + \mathrm{Ker}[i,m] = \mathrm{Ker}[i,m] = f_i$ because $f_m = 0$. By construction, β_i contains a basis of every $\mathrm{Ker}[i,j]$.

We now show that β_i contains a basis of every $\mathrm{Im}[k,i]$ by induction on i. The base case

$i = 0$ is trivial. Additionally, we only need to consider $k < i$ because $\mathrm{Im}[k, i] = f_i$ for $k \geq i$. The inductive hypothesis implies that β_{i-1} contains a basis of $\mathrm{Im}[k, i-1]$ for each $k < i$. Let $\iota_k^{i-1} \subseteq \beta_{i-1}$ be a basis of $\mathrm{Im}[k, i-1]$. Then $f(i-1 \leq i)(\iota_k^{i-1})$ is a spanning set of $f(i-1 \leq i)(\mathrm{Im}[k, i-1]) = \mathrm{Im}[k, i]$. Because $f(i-1 \leq i)(\iota_k^{i-1}) \setminus \{0\} \subseteq \beta_i$, we see that β_i contains a spanning set (in fact, a basis, because β_i is a basis of f_i) of $\mathrm{Im}[k, i]$. Thus, β_i contains a basis of each $\mathrm{Im}[k, i]$, as desired. This concludes the argument that β_i contains a basis of each submodule of the form $\mathrm{Ker}[i, j]$ and $\mathrm{Im}[k, i]$.

We still need to show that the bases $\beta_1, \ldots, \beta_{m-1}$ decompose f into a direct sum of interval submodules. That is, we need to verify that for $i, j \in \{0, \ldots, m\}$ such that $0 < i < j < m$, the structure map $f(i \leq j)$ (1) sends $\beta_i \cap \mathrm{Ker}[i, j]$ to 0 and (2) injectively maps $\beta_i \setminus \mathrm{Ker}[i, j]$ into β_j. The first claim (1) holds by Definition 3.1. The second claim (2) holds because $f(i \leq j)(\beta_i) \setminus \{0\}$ is linearly independent (as a multiset) and because $f(i \leq j)(\beta_i) \setminus \{0\} \subseteq \beta_j$, by construction. □

3.8 Matrix Algorithm

In this section, we translate Algorithm 3.1 into the language of matrix algebra. We describe the algorithm in Section 3.8.1, prove that this algorithm is correct in Section 3.8.2, and provide a complexity bound in Section 3.8.3. We give a review of relevant facts on Smith-normal-form factorization in Section 3.A.

For simplicity, assume that each module f_a is a copy of R^d, for some d, where R is the ring of coefficients. We identify the elements of R^d with length-d column vectors, and we let F_a denote the matrix such that $f(a \leq a+1)(v) = F_a v$ for all $v \in f_a$. Given matrices $M_1, M_2, \cdots, M_r$ of the same height, we denote $[M_1|M_2|\cdots|M_r]$ as the matrix obtained by concatenating $M_1, M_2, \cdots, M_r$.

Recall that $f_0 = f_m = 0$ by convention (see Section 3.6). As in Section 3.7, our task is to

define a sequence of bases $\beta_i \subseteq f_i$ such that $f(i \leq j)$ maps $\beta_i \setminus \mathrm{Ker}[i,j]$ injectively into β_j for all $i \leq j$.

Definition 3.12. *Let $A \in M_{r,s}(R)$ be given. We say that the* span *of A is the span of its columns. If $r \geq s$ and A has unit elementary divisors, we say that $B \in M_{r,r-s}(R)$ is a* complement *of A if $[A|B]$ is invertible.*

Definition 3.13. *A* kernel filtration matrix *at f_a is an invertible matrix*

$$X_a = [X_a^{a+1}| \cdots |X_a^n]$$

with column submatrices $X_a^{a+1}, \ldots, X_a^n$ such that

$$\textsc{Colspace}([X_a^{a+1}| \cdots |X_a^j]) = \mathrm{Ker}[a,j]$$

for all j.

We write $(X_a^j)^{-1}$ for the row submatrix of X_a^{-1} such that $(X_a^j)^{-1}X_a^j = I$ is the identity matrix. That is, the row indices of $(X_a^j)^{-1}$ in X_a^{-1} equal the column indices of X_a^j in X_a.

3.8.1 The algorithm

We give a self-contained description of the matrix algorithm (see Algorithm 3.2).

Algorithm 3.2 Interval decomposition via matrix factorization.

Require: A persistence module $f : \{0, \dots, m\} \to R\text{-Mod}$ such that each f_i equals R^k for some k. Concretely, this module is encoded by a sequence of matrices F_i representing the maps $f(i \leq i+1)$ for each integer i such that $1 \leq i < m-1$. We do not require f to admit an interval decomposition.

Ensure: Returns an interval decomposition of f, if such a decomposition exists. Otherwise returns a certificate that no such decomposition exists.

1: Let $\beta_i = \emptyset$ for $i = 1, \dots, m$
2: Let Y_0 be the (unique and trivial) kernel filtration matrix[17] at f_0
3: **for** $i = 1, \dots, m-2$ **do**
4: Compute a kernel filtration matrix $X_i = [X_i^{i+1} | \cdots | X_i^m]$ at f_i via repeated Smith-normal-form factorization (see Proposition 3.33)
5: **for** $j = i+1, \dots, m-1$ **do**
6: Obtain a Smith-normal-form factorization of $A_i^j := (X_i^j)^{-1} F_{i-1} Y_{i-1}^j$
7: **if** this factorization yields a complement B_i^j of A_i^j (see Proposition 3.32) **then**
8: Define $Y_i^j = [X_i^j B_i^j | F_{i-1} Y_{i-1}^j]$
9: **else**
10: STOP; the persistence module f does not split into interval submodules
11: Let β_i be the set of columns of $Y_i = [Y_i^{i+1} | \cdots | Y_i^m]$
12: **return** $\beta_1, \dots, \beta_m$

When matrix multiplication and the computation of a Smith normal form are polynomial-time procedures with respect to matrix dimensions, it is readily checked that Algorithm 3.2 is polynomial time with respect to matrix dimension and persistence-module length. We compute the complexity of Algorithm 3.2 (see Proposition 3.31) relative to the complexity of matrix multiplication, matrix inversion, and computing a Smith normal form.

Remark 3.5. *We can perform the computations of the inner-loop iterations (see line 5) of Algorithm 3.2 in parallel. That is, for any fixed $i \in \{1, \cdots, m-2\}$, we can compute each A_i^j and its Smith-normal-form factorization for $j \in \{i+1, \cdots, m-1\}$ in parallel.*

[17] Because the rank of f_0 is 0, any basis of f_0 is empty. We view the kernel filtration matrix at f_0 as a degenerate 0×0 matrix.

3.8.2 Correctness of Algorithm 3.2

We now prove that Algorithm 3.2 is correct.

We say that the matrices $Y_1, \ldots, Y_i$ *cohere* if $Y_k = [Y_k^{k+1}| \cdots |Y_k^m]$ is a kernel filtration matrix at f_k for all $k < i$. By construction of Y_k, this implies that for $k < i-1$, the following conditions hold:

- the set of columns of Y_k^{k+1} maps to 0 under F_k;
- the set of columns of Y_k^j maps injectively into the set of columns of Y_{k+1}^j under F_k for $j > k+1$.

Therefore, the bases $\beta_1, \ldots, \beta_m$, which are given by the columns of the matrices $Y_1, \ldots, Y_m$, decompose f into interval modules if $Y_1, \ldots, Y_m$ cohere.

Lemma 3.27. *Suppose that i satisfies $1 \leq i \leq m-1$. If A_p^q admits a complement for every p and q such that $q > p$ and $p \leq i$, then $Y_1, \ldots, Y_i$ cohere.*

Proof. We will show that $Y_1, \ldots, Y_i$ cohere for all i. We proceed by induction on i.

First, suppose that $i = 1$. Because Y_0 is a degenerate matrix, each A_1^j is also degenerate so we can view A_1^j as a matrix of size $d \times 0$ for some d. It follows that any $d \times d$ invertible matrix B_1^j is a complement of A_1^j. Because $X_1 = [X_1^2| \cdots |X_1^m]$ is a kernel filtration matrix by hypothesis and the column space of X_1^j equals that of $X_1^j B_1^j$, the matrix $Y_1 = [X_1^2 B_1^2| \cdots |X_1^m B_1^m]$ is also a kernel filtration matrix. This establishes the base case $i = 1$.

Now suppose that the desired conclusion holds for $i < k$, and consider $i = k$. The matrix Y_{i-1} is a kernel filtration matrix. Therefore, given a kernel filtration matrix X_i at f_i, for all j such that $i+1 \leq j \leq m-1$, the columns of $F_{i-1} Y_{i-1}^j$ lie in the column space of $[X_i^{i+1}| \cdots |X_i^j]$.

Therefore, the matrix $X_i^{-1}F_{i-1}Y_{i-1}$ has the following block structure:

$$X_i^{-1}F_{i-1}Y_{i-1} = \begin{bmatrix} 0 & A_i^{i+1} & * & \cdots & * \\ 0 & 0 & A_i^{i+2} & \cdots & * \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & A_i^m \end{bmatrix}. \tag{3.17}$$

Consider the matrix

$$M = \begin{bmatrix} A_i^{i+1} & B_i^{i+1} & * & 0 & \cdots & * & 0 \\ 0 & 0 & A_i^{i+2} & B_i^{i+2} & \cdots & * & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & A_i^m & B_i^m \end{bmatrix},$$

where B_i^j is a complement of A_i^j. The matrix M is invertible because each matrix $[A_i^k|B_i^k]$ is invertible. Moreover, the product $X_iM = Y_i$ is a kernel filtration matrix because X_i is a kernel filtration matrix and M is block-upper-triangular with invertible diagonal blocks. The desired conclusion follows. □

Lemma 3.28. *Suppose i satisfies $1 \leq i \leq m-1$. If Y_{i-1} is a kernel filtration matrix, then the columns of A_i^j in Equation* (3.17) *are linearly independent for $j > i$.*

Proof. Suppose that $j > i$, and let $C = \{k+1, k+2, \ldots, k+p\}$ denote the set of column indices in the matrix (3.17) that contain the submatrix A_i^j.

Seeking a contradiction, suppose that the columns of A_i^j are linearly dependent. Then there exists a nonzero vector $\hat{v}$ such that $A_i^j\hat{v} = 0$. Define a vector v by padding $\hat{v}$ with 0 entries to align the nonzero entries with the columns that contain A_i^j in the matrix (3.17). That is,

we define v such that

$$v_q = \begin{cases} \hat{v}_\ell\,, & q \in C \text{ and } q = k + \ell \\ 0\,, & \text{otherwise}\,. \end{cases}$$

By inspection of the block structure of (3.17) and recalling that $A_i^j \hat{v} = 0$, it follows that $F_{i-1}Y_{i-1}v$ lies in the column space of $[X_i^{i+1}|\cdots|X_i^{j-1}]$, which equals $\mathrm{Ker}[i, j-1]$. Consequently, $Y_{i-1}v$ lies in $\mathrm{Ker}[i-1, j-1]$. However, because Y_{i-1} is a kernel filtration matrix at f_{i-1}, the space $\mathrm{Ker}[i-1, j-1]$ equals the column space of $[Y_{i-1}^i|\cdots|Y_{i-1}^{j-1}]$. Therefore, the columns of Y_{i-1} are linearly dependent. This contradicts the hypothesis that the columns of Y_{i-1} form a basis of f_{i-1}. The desired conclusion follows. □

Lemma 3.29. *If Algorithm 3.2 terminates because some A_i^j lacks a complement, then f does not split as a direct sum of interval modules.*

Proof. Let i and j be the first indices such that A_i^j lacks a complement. For each $p < i$, the matrix A_p^q has a complement. For each $q < j$, the matrix A_i^q has a complement; let B_i^q be a complement of A_i^q.

Consider the matrix

$$N = \begin{bmatrix} A_i^{i+1} & B_i^{i+1} & * & 0 & \cdots & * \\ 0 & 0 & A_i^{i+2} & B_i^{i+2} & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & A_i^j \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Because each submatrix of the form $[A_i^k|B_i^k]$ is invertible, the span of the product X_iN contains the span of $[X_i^{i+1}|\cdots|X_i^{j-1}]$, which equals $\mathrm{Ker}[i, j-1]$ because X_i is a kernel filtration matrix.

Because every preceding matrix A_p^q admits a complement, by Lemma 3.27, the matrix Y_{i-1} is a kernel filtration matrix at f_{i-1}. The columns of the matrix $[Y_{i-1}^i|\cdots|Y_{i-1}^j]$ therefore span the subspace $\mathrm{Ker}[i-1,j]$ of f_{i-1}. Consequently, the columns of the product matrix $F_{i-1}[Y_{i-1}^i|\cdots|Y_{i-1}^j]$ span a subspace belonging to the saecular sublattice of f_i, because the direct image operator restricts to a homomorphism of saecular lattices (see Theorem 3.12). This image subspace equals the span of $F_{i-1}[Y_{i-1}^{i+1}|\cdots|Y_{i-1}^j]$, because $F_{i-1}Y_{i-1}^i = 0$. Every column of X_iN is a column of $F_{i-1}[Y_{i-1}^{i+1}|\cdots|Y_{i-1}^j]$ or a column in the span of $[X_i^{i+1}|\cdots|X_i^{j-1}]$, so X_iN spans a subspace S in the saecular lattice at f_i.

Because A_i^j lacks a complement, at least one of two conditions must hold; either (i) A_i^j has a non-unit elementary divisors or (ii) the columns of A_i^j are linearly dependent. Because Y_{i-1} is a kernel filtration matrix, the columns of A_i^j are linearly independent, by Lemma 3.28. Therefore, A_i^j has a non-unit elementary divisor. It follows that the quotient f_i/S has torsion, which implies that S does not admit a complement in f_i. This implies that f does not split as a direct sum of interval modules, by Theorems 3.2 and 3.15. □

Theorem 3.30. *Let $f : \{0,\cdots,m\} \to R\text{−Mod}$ be a persistence module that is pointwise free and finitely-generated. If f has an interval decomposition, then Algorithm 3.2 returns an interval decomposition. If f has no interval decomposition, then Algorithm 3.2 certifies that no interval decomposition exists.*

Proof. Algorithm 3.2 terminates either at line 10 or at line 12. If f has an interval decomposition, then by Lemma 3.29, Algorithm 3.2 does not terminate at line 10. This implies that Algorithm 3.2 terminates at line 12. By Lemma 3.27, Algorithm 3.2 yields a set of bases $\beta_1,\ldots,\beta_m$ that gives an interval decomposition of f. If f does not have an interval decomposition, then by Lemma 3.27, Algorithm 3.2 does not terminate at line 12. So, Algorithm 3.2 must terminate at line 10, which certifies that f does not have an interval decomposition. □

3.8.3 Complexity of Algorithm 3.2

We now provide a worst-case upper bound on the computational complexity of Algorithm 3.2. The upper bound depends on the computational complexity of multiplying two matrices, inverting a matrix, and computing the Smith normal form of a matrix. We denote the complexity of these operations on $d \times d$ matrices by M_d, R_d, and S_d, respectively. Because the complexity of these operations depends on the choice of R, we express the complexity of Algorithm 3.2 in terms of M_d, R_d, and S_d.

Proposition 3.31. *Let $f : \{0, \dots, m\} \to R\text{-}Mod$ be a persistence module that is pointwise free and finitely-generated, and let $d = \max_i(\text{rank}(f_i))$. Let M_d be the complexity of matrix multiplication of two $d \times d$ matrices, R_d be the complexity of inverting a $d \times d$ matrix, and S_d be the complexity of computing the Smith normal form of a $d \times d$ matrix. Algorithm 3.2 has worst-case complexity $O(mK_{m,d})$, where $K_{m,d} = \max(mM_d, mS_d, R_d)$.*

Proof. Consider the ith iteration of the outer loop (see line 3) of Algorithm 3.2. In this iteration, we must compute one kernel filtration matrix at f_i and run once through the inner loop (see line 5).

We first account for the computation of the kernel filtration matrix at f_i. To do this, we follow the procedure in Proposition 3.33. For each step $k = i, \dots, m-1$, we take the matrix product $F_{k-1} \cdots F_i T_i^+ \cdots T_{k-2}^+$, multiply by F_k on the left, and multiply by T_{k-1}^- on the right to obtain the matrix product $F_k \cdots F_i T_i^+ \cdots T_{k-2}^+ T_{k-1}^-$. We then compute the Smith normal form. Each step thus requires two matrix multiplications and one Smith-normal-form factorization. Because the dimension of each F_i is at most $d \times d$, the dimensions of matrices F_k, T_{k-1}^+, T_{k-1}^-, and $F_{k-1} \cdots F_i T_i^+ \cdots T_{k-2}^+$ is at most $d \times d$. Lastly, we need to compute $F_k \cdots F_i T_i^+ \cdots T_{k-2}^+ T_{k-1}^+$, which is needed for step $k+1$. This requires an additional matrix multiplication, where we multiply by T_{k-1}^+ instead of by T_{k-1}^- on the right. In total, we require the computation of three matrix multiplications and one Smith normal form. The

total computational cost is thus at most $3M_d + S_d$. We iteratively perform these operations $m - i$ times, so in total the computation of a kernel filtration matrix requires at most $(m - i)(3M_d + S_d)$ arithmetic operations.

We now account for the computations of the inner loop (see line 5 of Algorithm 3.2). By construction, each matrix A_i^j is a submatrix of $X_i^{-1}F_{i-1}Y_{i-1}$. Thus, the combined cost of computing every matrix A_i^j is at most the cost of one matrix inversion and two matrix multiplications (of matrices with sizes of at most $d \times d$). In addition, for each $j = i+1, \ldots, m-1$, we must compute (1) a complement B_i^j of A_i^j (if one exists), which requires the computation of one Smith normal form, and (2) a matrix Y_i^j, which requires two matrix multiplications (see line 8 of Algorithm 3.2). Thus, in total, we require one matrix inversion, $2+2(m-i-1) = 2(m-i)$ matrix multiplications, and $m - i - 1$ Smith-normal-form factorizations. In total, the inner loop computation requires at most $2(m-i)M_d + (m-i-1)S_d + R_d$ arithmetic operations.

Overall, each iteration $i = 1, \ldots, m-2$ of the outer loop (see line 3) of Algorithm 3.2 requires at most of at most $5(m-i)M_d + (2m-2i-1)S_d + R_d$ arithmetic operations. The worst-case complexity of Algorithm 3.2 therefore is at most $O(mK_{m,d})$, where $K_{m,d} = \max(mM_d, mS_d, R_d)$. □

Remark 3.6. *Proposition 3.31 provides a worst-case complexity of Algorithm 3.2, but we believe that a typical computation is much faster. We did not calculate the typical number (e.g., average-case complexity) of required arithmetic operations that is required by Algorithm 3.2.*

3.9 Conclusion

We showed that a finitely-indexed persistence module that is pointwise free and finitely-generated over a PID splits as a direct sum of interval submodules if and only if the cokernel

of every structure map is free (see Theorem 3.2). One direction of this equivalence is simple to prove (see Section 3.5), but the other requires significant machinery. We studied the saecular lattice of each module in the persistence module, and we constructed a compatible direct-sum decomposition. This gives the necessary ingredients to build an interval decomposition. We constructed the bases that comprise our interval decomposition by relating corresponding direct summands in different modules to each other.

We gave a concise algorithm that returns either an interval decomposition or a certificate that no such decomposition exists. This algorithm has two variants. The first, Algorithm 3.1, is primarily combinatorial; it exposes the underlying algebra in a simple format. It works by iteratively extending a family of bases with new elements that are compatible with the saecular lattice. The second, Algorithm 3.2, translates this procedure into the language of matrix algebra; it requires no special machinery beyond the Smith normal form, and it offers natural possibilities for parallelization.

3.9.1 Future directions

A theoretical direction to explore is to extend our work to the setting of persistence modules of general R-modules, rather than restricting ourselves to free modules. This requires us to generalize the notion of "interval decomposition" to the setting of persistence modules valued over modules with torsion. One can then determine the conditions on persistence modules for the existence of interval decompositions. One can perhaps take inspiration from [Pat18] for insights into generalized persistence.

It also would be interesting to generalize our work from the setting of finitely-indexed persistence modules (for which the domain category is a finite totally-ordered poset category) to continuously-indexed persistence modules (for which the domain category is a continuous-valued totally-ordered poset category, such as $[0, 1]$ or $\mathbb{R}$). We expect that many of the ideas from the present chapter can be adapted to this more general setting, although the treatment

may require extra care with respect to limits (specifically, upper continuity of the saecular lattice; see [GH21]).

It is natural to ask whether the methods explored in this chapter can also be applied in the setting of zig-zag persistence modules. A persistence module consists of a sequence $\{M_i\}_{i=0}^{m}$ of R-modules with maps $\phi_i : M_i \to M_{i+1}$. Zig-zag persistence modules [CS10] are generalizations of persistence modules that consist of a sequence $\{M_i\}_{i=1}^{r}$ of R-modules in which the maps ϕ_i can go in either direction (i.e., either $\phi_i : M_i \to M_{i+1}$ or $\phi_i : M_{i+1} \to M_i$). When the ring of coefficients is a field, Gabriel's theorem [Gab72] guarantees that any zig-zag persistence module admits an interval decomposition. It would be interesting to consider the conditions under which zig-zag persistence modules of free and finitely-generated modules over a PID admit interval decompositions.

3.A Details on the Smith Normal Form

The decomposition procedure in Algorithm 3.2 requires no special machinery except the Smith normal form. In this section, we provide a cursory review of the relevant facts about the Smith normal form. See [DF04] for a detailed introduction.

We say that a matrix S with entries in a PID R is *invertible* if there exists a matrix S^{-1} with entries in R such that $SS^{-1} = S^{-1}S$ is the identity.

A Smith-normal-form factorization of a matrix $A \in M_{m,n}(R)$ is an equation $SAT = D$, where S and T are invertible and D is a diagonal matrix of form $\mathrm{diag}(\alpha_1, \ldots, \alpha_r, 0, \ldots 0)$ in which α_i divides α_{i+1} for all $i < r$. We refer to $\alpha_1, \ldots, \alpha_r$ as the *elementary divisors* of A. We say that A has *unit elementary divisors* if its elementary divisors are all units.[18]

We refer to the column submatrix of T that consists of the first r columns as the *positive*

[18]Recall that a unit is a ring element with a multiplicative inverse.

part of T; we denote it by T^+. The submatrix that consists of the remaining $m-r$ columns is the *negative* part of T; we denote it by T^-. These matrices are complementary in the sense of Definition 3.14.

Definition 3.14. *Let $A \in M_{p,q}(R)$, where $p \geq q$ and A has unit elementary divisors. We say that $B \in M_{p,p-q}(R)$ is a* complement *of A if $[A|B]$ is invertible.*

Remark 3.7. *Let $A \in M_{m,n}(R)$, and let $SAT = D$ be a Smith-normal-form factorization. Then AT^+ has linearly-independent columns and $AT^- = 0$.*

We define the positive part of S^{-1} to be the submatrix $(S^{-1})^+$ composed of the first r columns and the negative part of S^{-1} to be the submatrix $(S^{-1})^-$ composed of the last $m-r$ columns. These submatrices bear a special relationship to the cokernel of A, as formalized in Proposition 3.32. The proof consists of straightforward calculations in matrix algebra.

Proposition 3.32. *Fix a matrix $A \in M_{m,n}(R)$, and let $SAT = D$ be a Smith-normal-form factorization of A. Then the following are equivalent:*

1. *The homomorphism represented by A has a free cokernel.*
2. *The diagonal entries of D are units or 0.*

When these conditions hold,

1. *the image of A is the column space of $(S^{-1})^+$;*
2. *the column space of $(S^{-1})^-$ is a complement of the image space of A.*

The Smith normal form also has a useful relationship with nested kernels. Given a persistence module $f : \{0, \ldots, m\} \to R\text{-Mod}$, recall that F_a is the matrix representation of the map $f(a \leq a+1)$.

Proposition 3.33. *Given $a \in \{0, \ldots, m\}$, let $S_a F_a T_a$ be a Smith-normal-form factorization*

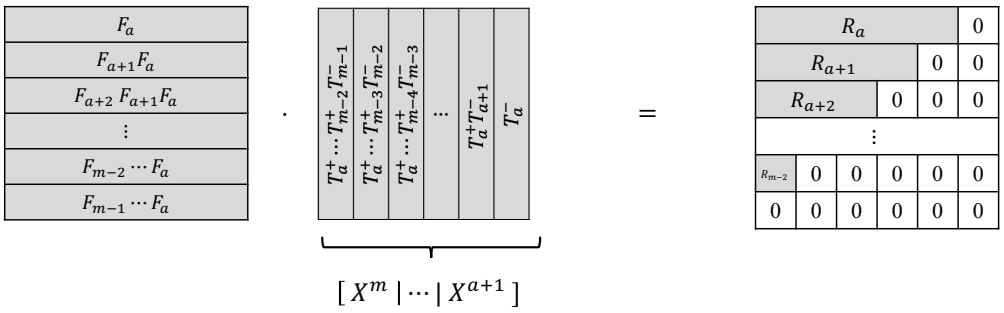


Figure 3.3: A product of matrices with block structure. The matrix R_k has linearly independent columns and can be expressed in the form (3.20). Observe that $R_{m-1} = 0$ because $F_{m-1} = 0$.

of F_a, and let S_k and T_k be inductively defined such that

$$S_k\Big(F_k \cdots F_a T_a^+ \cdots T_{k-1}^+\Big)T_k \tag{3.18}$$

is a Smith-normal-form factorization of the product matrix $F_k \cdots F_a T_a^+ \cdots T_{k-1}^+$ for all $k > a$. If $X^k := T_a^+ \cdots T_{k-1}^+ T_k^-$, then the columns of $[X^{a+1}|\cdots|X^k]$ form a basis of the kernel of $f(a \le k)$ for all $k > a$.

Proof. For any k, we have that $F_k \cdots F_a \cdot [X^k|\cdots|X^{a+1}] = 0$. This is because, by construction, $F_r \cdots F_a X_r = F_r \cdots F_a T_a^+ \cdots T_{r-1}^+ T_r^- = 0$ for all r satisfying $a < r \le k$. This implies that the columns of $[X^k|\cdots|X^{a+1}]$ are in the kernel of $F_k \cdots F_a$.

A straightforward calculation yields[19]

$$[X^m|\cdots|X^{a+1}] = T_a \cdot \mathrm{diag}(T_{a+1}, \mathrm{id}) \cdots \mathrm{diag}(T_m, \mathrm{id})\,, \tag{3.19}$$

where $\mathrm{diag}(T_k, \mathrm{id})$ denotes the block-diagonal matrix with T_k in the upper-left corner and an identity matrix of appropriate size in the lower-right corner to form a matrix of the same size as T_a. Therefore, $[X^m|\cdots|X^{a+1}]$ is invertible because it is a product of invertible matrices.

[19]Observe that $T_{m-1}^- = T_{m-1}$ because F_{m-1} is the zero map.

Consequently, for any k, the matrix $[X^k|\cdots|X^{a+1}]$ has linearly independent columns.

We still need to show that the columns of $[X^k|\cdots|X^{a+1}]$ span the kernel of $F_k\cdots F_a$. Let

$$\tilde{F} = \begin{bmatrix} F_a \\ F_{a+1}F_a \\ \vdots \\ F_{m-2}\cdots F_a \\ F_{m-1}\cdots F_a \end{bmatrix}$$

and consider the product $\tilde{F}[X^m|\cdots|X^{a+1}]$, which is of the form in Figure 3.3.

For each k, we can write

$$R_k = F_k\cdots F_a T_a^+\cdots T_k^+[\ T_{k+1}^+\cdots T_{m-2}^+T_{m-1}^- \mid T_{k+1}^+\cdots T_{m-3}^+T_{m-2}^- \mid \cdots \mid T_{k+1}^-\ ]. \tag{3.20}$$

This implies that the columns of $[X^k|\cdots|X^{a+1}]$ span a submodule L of the kernel (which we denote K) of $F_k\cdots F_a$. Because $F_k\cdots F_a T_a^+\cdots T_k^+$ and
$[\ T_{k+1}^+\cdots T_{m-2}^+T_{m-1}^- \mid T_{k+1}^+\cdots T_{m-3}^+T_{m-2}^- \mid \cdots \mid T_{k+1}^-\ ]$ have linearly independent columns, the submatrix R_k also has linearly independent columns. Therefore, by counting dimensions, we have $\dim(L) = \dim(K)$. Therefore, $L = K$ if and only if K/L is torsion-free. Because $[X^k|\cdots|X^{a+1}]$ admits a complement — namely, $[X^m|\cdots|X^{k-1}]$ — we see that $c\cdot v \in L$ implies that $v \in L$ for every nonzero scalar $c \in R$. Therefore, K/L is torsion-free, so $K = L$. Because the columns of $[X^k|\cdots|X^{a+1}]$ are linearly independent, they form a basis of K. $\square$

CHAPTER 4

Persistent Homology for Resource Coverage: A Case Study of Access to Polling Sites

In this chapter, we use persistent homology (PH) to study the accessibility of resources that are distributed in a geographical region. We investigate the accessibility of polling sites across six geographical regions (five cities and Los Angeles County) in the United States. The information from PH provides summary statistics of the distribution of polling sites. We use these statistics to infer which regions have poor polling-site access. This chapter is adapted[1] from [HJJ23], which was led jointly by Abigail Hickok, Benjamin Jarman, Michael Johnson and me and was co-authored with Mason A. Porter.[2]

This chapter proceeds as follows. We introduce the goals of our work in Section 4.1, describe our approach in Section 4.2, present our persistence diagrams (PDs) and analyze them Section 4.3, and conclude and discuss implications, limitations, and potential future directions of our work in Section 4.4. Our code is available at `https://bitbucket.org/jerryluo8/coveragetda/src/main/`.

[1] All figures in this chapter originally appeared in [HJJ23].

[2] I co-designed the methodology with A. Hickok, B. Jarman, and M. Johnson; computed the distance matrices (as described in Section 4.2.1) with B. Jarman and M. Johnson; computed the persistence diagrams in Section 4.3; and wrote the paper with all co-authors.

4.1 Introduction

The geographical distribution of resources such as polling sites (i.e., locations where people vote), hospitals, COVID-19 vaccination sites, Department of Motor Vehicles (DMV) locations, and Planned Parenthood clinics is a major factor in the equitability of access to those resources. Consequently, given the locations of a set of resource sites, it is important to quantify their geographical coverage and to identify underserved geographical regions (i.e., "holes in coverage").

A naive approach to quantifying resource coverage is to consider the geographical distances from resource sites by simply calculating the percentage of people who reside within some cutoff distance D of the nearest resource site. This naive approach is common in policy. For example, in March 2021, United States President Joseph Biden announced a goal to ensure that at least 90% of the adult U.S. population is within 5 miles (i.e., $D = 5$ miles) of a COVID-19 vaccination site [The21]. As another example, it is required by Indian law that 100% of voters live within 2 km of a polling site [SSH19] (i.e., $D = 2$ km). However, such an approach poses at least two issues:

(1) it requires choosing an arbitrary cutoff distance D; and

(2) using only geographical distance fails to account for many other factors, such as population density and the availability (and facility) of public transportation, that affect the ease of access to a resource.

These issues severely limit the utility of this naive approach.

In the present chapter, we use topological data analysis (TDA) to study holes in resource coverage. As we discussed in Chapter 2, one of the main tools in TDA is PH, which uses ideas from algebraic topology to (1) identify clusters and holes in a data set and (2) measure their persistences at different scales. We use PH to analyze data in the form of a point

cloud in a metric space (M, d).[3] In this chapter, X is a collection of resource sites, with specified latitudes and longitudes, and $M = \mathbb{R}^2$ with a non-Euclidean distance function d (see Section 4.2). Given a point cloud X and a scale parameter $r > 0$, one can consider the *r-coverage* $C_r := \bigcup_{i=1}^{n} B(x_i, r)$. As the scale parameter r grows, holes arise and subsequently fill in. PH tracks the formation and disappearance of these holes. When a point cloud is a collection of resource sites, one can interpret holes that persist for a large range of r as holes in coverage. Our TDA approach gives a way to measure and evaluate how equitably a resource is distributed geographically.

Our approach using PH addresses both of the issues (see points (1) and (2)) of the naive approach that we discussed above. First, PH eliminates the need to choose an arbitrary cutoff distance because one can study holes in coverage at all scales. Second, instead of employing geographical distance as a metric, we construct a distance function d that is based on travel times. We also incorporate the waiting time at each resource site by constructing a weighted Vietoris–Rips (VR) filtration (see Section 2.2.2). In a city with a high population density or a poor transportation system, the time that is spent waiting at or traveling to a resource site can be a much higher barrier to access than geographical distance [GS03, HK05]. We estimate waiting times using Global Positioning System (GPS) ping data from mobile phones at the resource sites, and we estimate travel times using street-network data, per capita car-ownership data, and the Google Maps application programming interface (API) [Goo]. Using these estimates, we construct a weighted VR filtration. We weight vertices by our estimates of waiting times, and we define the distance between two vertices to be the estimated round-trip travel time between them. Because the weighted VR filtration is stable, small errors in our estimates cause only small errors in the resultant PH [ACG19].

In the present chapter, we examine polling sites as a case study of using PH to study

[3]One can weaken the requirement that d is a metric. In this chapter, we use a distance function d that is not technically a metric because it does not satisfy the triangle inequality.

the coverage of resource sites. We restrict our attention to six cities: Atlanta, Chicago, Jacksonville (in Florida), Los Angeles[4], New York City (NYC), and Salt Lake City. We use these cities in part because data about them (e.g., car-ownership data) is widely available. Additionally, these cities differ considerably in their demographics and infrastructures, and we can thus compare a variety of different types of cities. Atlanta and NYC are both infamous for long waiting times at polling sites, especially in non-White neighborhoods [Fow20,Kan19]. In 2020, some counties in the Atlanta metropolitan area had a mean of 3,600 voters per polling site; the number of polling sites had been cut statewide in Georgia by 10% since 2013 [Fow20]. In NYC, each polling site had a mean of 4,173 voters in 2018. As a comparison, in 2004, Los Angeles County and Chicago had only an estimated 1,300 and 725 voters per polling site, respectively [Kan19]. However, Los Angeles is infamous for its traffic [Sch21], which can affect voters' travel times to polling sites. Los Angeles and Chicago also differ in the quality of their public transportation, which also affects travel times to polling sites. In our investigation, we seek both to compare the coverage of polling sites in our six focal cities and to identify underserved areas within each city.

4.1.1 Related work

One can use tools from geography to study resource accessibility. Pearce et al. [PWB06] used a geographical-information-systems (GIS) approach to examine the accessibility of community resources and how it affects health. Hawthorne and Kwan [HK12] used a GIS approach and a notion of perceived distance to measure healthcare inequality in low-income urban communities. Brabyn and Barnett [BB04] illustrated that there are regional variations in geographical accessibility to general-practitioner doctors in New Zealand and that these regional variations depend on how one measures accessibility.

Another motivation for our study of resource-site coverage is the related problem of sensor

[4]For Los Angeles, we actually study Los Angeles County. We discuss the reasons for this choice in Section 4.2.5.

coverage. Given a set S of sensors in a domain $\Omega \subseteq \mathbb{R}^2$, one seeks to determine if every point in Ω is within sensing range of at least one sensor in S. Typically, each sensor has a fixed, uniform sensing radius r_s. In this case, the problem is equivalent to determining whether or not the domain Ω is covered by balls of radius r_s around each $s \in S$. In [SG07, SG06], de Silva and Ghrist gave homological criteria for sensor coverage. Approaches to study sensor coverage that use computational geometry (specifically, ones that involve the Voronoi diagram of S and the Delauney triangulation of S) were discussed in [LWF03, MKP01].

Our problem is also a coverage problem, but there are important differences. The key conceptual difference is that we consider neighborhoods whose sizes depend on a filtration parameter, rather than neighborhoods with a fixed radius. Additionally, we do not seek to determine whether or not balls of any particular radius cover a domain; instead, our goal is to quantify the coverage at all choices of radius and to determine how the holes in coverage evolve as we increase the filtration parameter. Another difference between the present chapter and sensor-coverage problems is that our point cloud represents a set of resource sites (in particular, polling sites), rather than a set of sensors. In a sensor network, pairwise communication between sensors can play a role in whether or not the sensors are fully "connected" to each other (in a graph-theoretic sense) and in determining whether or not a domain is covered [ZH05]. By contrast, communication between resource sites does not play a role in access to those resource sites.

Several studies include applications of PH to geospatial data [CJ23]. Feng and Porter [FP21] developed two methods to construct filtrations—one that uses adjacency structures and one that uses the level-set method [OF03] of front propagation—and applied their approaches to examine geospatial distributions of voting results in the 2016 United States presidential election. They identified "political islands" (i.e., precincts that voted more heavily for a candidate than its surrounding precincts). In [FP20], Feng and Porter used their approaches to study various spatial networks. Stolz et al. [SHP16] used PH to examine the geospatial

distribution of voting results in the "Brexit" referendum. Hickok et al. [HNP22] used PH to study geospatial anomalies in COVID-19 case-rate data (see also [FHP22]) and vaccination-rate data. Cocoran and Jones [CJ23] used PH to perform (1) a point-pattern analysis of pubs across different cities in the United Kingdom (UK) and (2) a spatiotemporal analysis of rainfall in the UK.

4.2 Our Construction of Weighted VR Complexes

For each city, we construct a weighted VR filtration (see Section 2.2.2) in which the point cloud $X = \{x_i\}$ is the set of polling sites in $\mathbb{R}^2$ and the weight w_i of a point x_i is an estimate of the waiting time at the corresponding polling site. Instead of computing a weighted VR filtration with respect to Euclidean distance, we define a distance function that estimates the mean amount of time that it takes to travel to and from a polling site. With respect to this distance function, the union $\bigcup_i B(x_i, r_{x_i}(t))$ (see 2.2) is the set of points y such that the estimated time for an individual at y to vote (including waiting time and travel time[5] in both directions) is at most t. The weighted Čech complex $\check{\mathrm{C}}_t^{\mathrm{weighted}}(X, \mathbb{R}^2, d, \{w_i\})$ is an approximation of $\bigcup_i B(x_i, r_{x_i}(t))$. When the balls $B(x_i, r_{x_i}(t))$ are convex, the weighted Čech complex is homotopy-equivalent to $\bigcup_i B(x_i, r_{x_i}(t))$, so these two complexes have the same homology (i.e., the same set of holes). The weighted VR complex $\mathrm{VR}_t^{\mathrm{weighted}}(X, \mathbb{R}^2, d, \{w_i\})$ is an approximation of the weighted Čech complex.

We construct our distance function as follows. Let x and y be two polling sites. We estimate the expected time for an individual to travel from x to y and back to be

$$\begin{aligned}\tilde{d}(x, y) := C(Z(x)) \min\{t_{\mathrm{car}}(x, y), t_{\mathrm{pub}}(x, y), t_{\mathrm{walk}}(x, y)\} \\ + [1 - C(Z(x))] \min\{t_{\mathrm{pub}}(x, y), t_{\mathrm{walk}}(x, y)\},\end{aligned}$$

[5]Incorporating information (such as waiting times) other than travel times is sensible both in principle and in practice. In our computational experiments, using only travel times yields results that differ drastically from those that we present in this chapter.

where $Z(x)$ is the zip code that includes x (a polling site), $C(Z(x))$ is an estimate of the fraction of voting-age people in $Z(x)$ who can travel by car to a polling site, and $t_{\text{car}}(x, y)$, $t_{\text{pub}}(x, y)$, and $t_{\text{walk}}(x, y)$ are estimates of the expected travel times from x to y and back by car, public transportation, and walking, respectively. We calculate $C(Z(x))$ by dividing an estimate of the number of personal vehicles in $Z(x)$ by an estimate of the voting-age population in $Z(x)$ (see Section 4.2.3). We discuss how we calculate t_{car}, t_{pub}, and t_{walk} in Section 4.2.1.

Our definition of $\tilde{d}(x, y)$ captures the cost (in time) to vote. In particular, $\tilde{d}(x, y)$ is an estimate of the mean travel time for an individual who resides in zip code $Z(x)$ to travel from x to y and back. We assume that all individuals choose the fastest mode of transportation that is available to them. Therefore, individuals who can travel by car choose the fastest option between driving, taking public transportation, and walking. Their travel time is $\min\{t_{\text{car}}(x, y), t_{\text{pub}}(x, y), t_{\text{walk}}(x, y)\}$. Likewise, we assume that individuals who do not have access to travel by car choose the fastest option between taking public transportation and walking. Their travel time is $\min\{t_{\text{pub}}(x, y), t_{\text{walk}}(x, y)\}$. Our estimate of the fraction of a population that has a car is $C(Z(x))$, so the fraction without a car is $1-C(Z(x))$. Therefore, $\tilde{d}(x, y)$ is the estimated mean time for an individual who resides in zip code $Z(x)$ to travel from x to y and back.

The function $\tilde{d}(x, y)$ is not symmetric (i.e., $\tilde{d}(x, y) \neq \tilde{d}(y, x)$) because $C(Z(x)) \neq C(Z(y))$. However, we need a symmetric function to construct a weighted VR filtration. To construct a symmetric distance function that is based on $\tilde{d}(x, y)$, we define the distance between x and y to be a weighted average of $\tilde{d}(x, y)$ and $\tilde{d}(y, x)$, where we determine the weights from the populations of the zip codes that include x and y. More precisely, we define the distance between x and y to be

$$d(x, y) := \frac{1}{P}[P_{Z(x)}\tilde{d}(x, y) + P_{Z(y)}\tilde{d}(y, x)]\,, \tag{4.1}$$

where $P_{Z(x)}$ and $P_{Z(y)}$ are the populations of zip codes $Z(x)$ and $Z(y)$, respectively, and $P := P_{Z(x)} + P_{Z(y)}$ is the sum of the populations of $Z(x)$ and $Z(y)$. With respect to this distance function, the ball $B(x, r)$ is the set of points y such that the expected time for an individual to travel back and forth between x and y is at most r, where the individual to start randomly at x or y with probabilities that are weighted by the populations of their associated zip codes. Although our distance function is not technically a metric (because it does not satisfy the triangle inequality), we can still construct a weighted VR filtration using the definition in Section 2.2.2.

4.2.1 Estimating travel times

To compute our distance function (see 4.1), we need to estimate the pairwise travel times by car, public transportation, and walking between each pair of polling sites. We measure these times in minutes.

We estimate the time that it takes to walk between each pair of polling sites by using street networks, which are available through the OpenStreetMap tool [Ope21], for each of our cities. Using OpenStreetMap, we calculate a shortest path (by geographical distance) between each pair of polling sites. In Figure 4.1, we show an example of a shortest path between two polling sites in Atlanta.

Let $L(x, y)$ denote the length (which we measure in meters) of a shortest path (by geographical distance) between polling sites x and y. Our estimate of the walking time (in minutes) from x to y and back is $t_{\text{walk}}(x, y) := 2L(x, y)/v_{\text{walk}}$, where $v_{\text{walk}} = 85.2$ meters per minute is an estimate of the mean walking speed of an adult human [BBH06].

To estimate the travel times by car and public transportation, we use the Google Maps Distance Matrix API [Goo]. Because of budgetary constraints (and the cost of five dollars per thousand API queries), we use this API to estimate only the travel times between each polling site and its 25 geographically closest polling sites. We refer to these sites as a polling

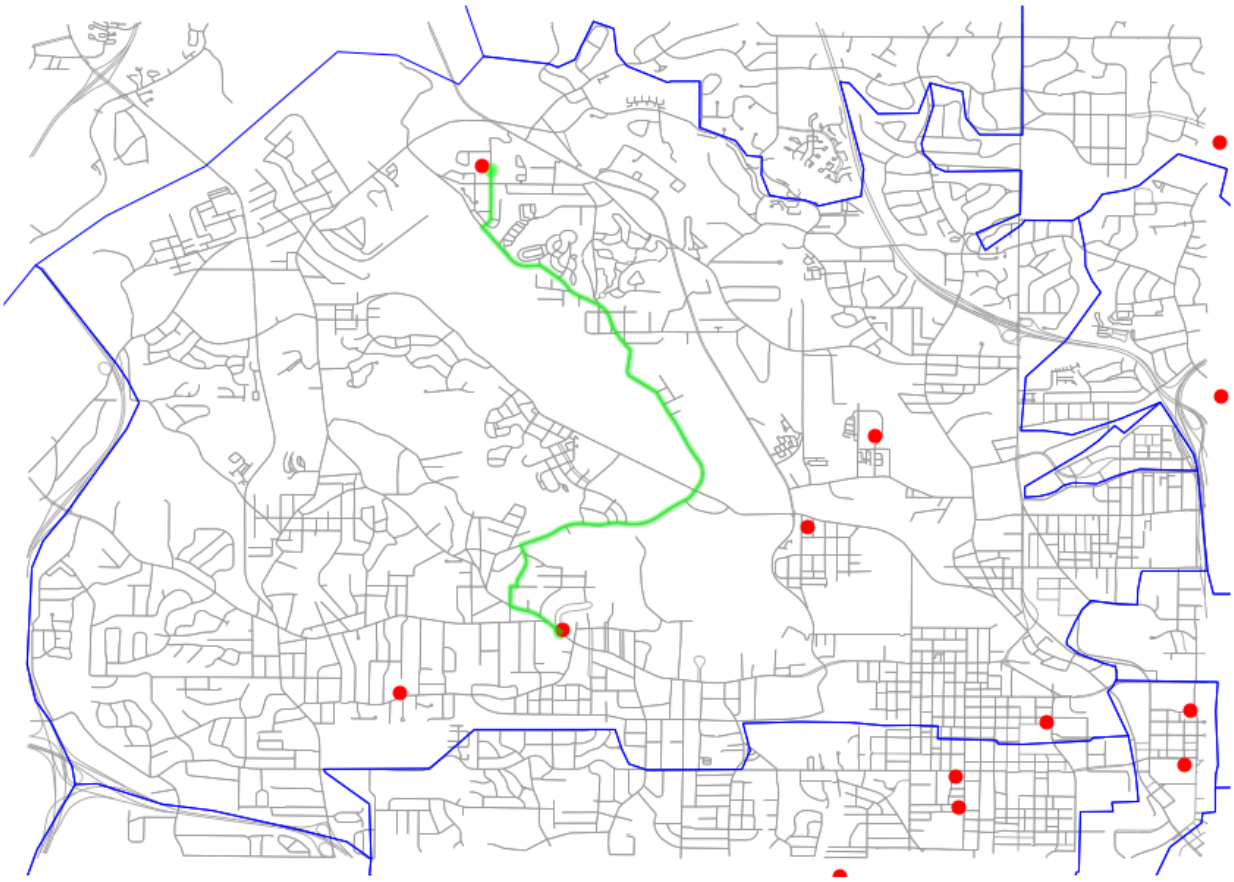


Figure 4.1: A shortest path (by geographical distance) between two polling sites in zip code 30314 in Atlanta.

site's 25 nearest neighbors.

For each of the 25 nearest neighbors, we separately calculate both the time from a polling site to each neighbor and the time to a polling site from each neighbor. These two travel times are often different because of different traffic conditions or other factors. We estimate the remaining pairwise travel times as follows. Let G be the unweighted, undirected graph whose vertices are the polling sites and whose edges connect each vertex to its 25 nearest neighbors.[6] Let G_{car} and G_{pub} be the weighted, directed graphs whose vertices and edges[7]

[6]The relation of being one of a vertex's 25 nearest neighbors is not symmetric. Therefore, the degrees of some vertices are larger than 25.

[7]We view each undirected edge (x_i, x_j) of G as a bidirectional edge, and we include both of the associated

are those of G and whose weights are the travel times (by car and public transportation, respectively) that we compute using the Google Maps API. The weight of the directed edge from vertex x to vertex y is the travel time from x to y. Therefore, the weight of the edge from x to y may differ from the weight of the edge from y to x. For any two polling sites x and y, let the travel times $\tilde{t}_{\text{car}}(x, y)$ and $\tilde{t}_{\text{pub}}(x, y)$ be the length of a shortest weighted path from x to y in the graphs G_{car} and G_{pub}, respectively. The corresponding symmetrized travel times $t_{\text{car}}(x, y)$ and $t_{\text{pub}}(x, y)$ are

$$t_{\text{car}}(x, y) := \tilde{t}_{\text{car}}(x, y) + \tilde{t}_{\text{car}}(y, x)\,,$$
$$t_{\text{pub}}(x, y) := \tilde{t}_{\text{pub}}(x, y) + \tilde{t}_{\text{pub}}(y, x)\,.$$

4.2.2 Estimating waiting times

Our weighted VR filtrations have weights at each vertex (i.e., polling site) that are given by an estimate of the mean time that a voter spends (i.e., the mean waiting time) at that polling site. In a nationwide study of waiting times at polling sites during the 2016 U.S. presidential election [CHP19], Chen et al. used smartphone data of hundreds of thousands of voters to estimate waiting times. They also examined potential relationships between waiting times and racial demographics.

We construct our waiting-time estimates using the congressional district-level estimates in [CHP19] (see their Table C.2). For each polling site x, we use the mean of the waiting-time estimates for each congressional district that overlaps with the zip code $Z(x)$ that contains x. This averaging procedure yields estimates of waiting times at the zip-code level. (We transform our waiting-time data to the zip-code level because the rest of our data is at the zip-code level.)

---

directed edges in the directed graphs G_{car} and G_{pub}.

4.2.3 Estimates of demographic information

We obtain estimates of demographic data at the zip-code level from 2019 five-year American Community Survey data [US]. We use voting-age population data from their Table ACSDT5Y2019.B29001 and vehicle-access data from their Table ACSDT5Y2019.B25046.

4.2.4 Polling-site zip codes

Much of our data is at the zip-code level, and we treat a polling site's zip code as representative of its local area. Certain polling sites (predominantly government buildings) have their own zip codes, despite their populations of 0. We adjust the zip codes of such polling sites to match the zip codes of the directly surrounding areas.

4.2.5 Treatments of our cities

The city of Atlanta does not include the suburbs of the Atlanta metropolitan area, so we use the entire area that is served by the Atlanta Regional Commission.

Chicago's boundary is not convex (especially in the northwest), so we include all areas of all zip codes, even when only a small portion of the zip code lies within the city of Chicago.

Because of the oddly-shaped city boundaries of Los Angeles, which include several holes, we use the entirety of Los Angeles County (except for its islands).

Because of the disconnected nature of New York City, we subdivide it into three regions (Queens and Brooklyn, Manhattan and the Bronx, and Staten Island) and treat each region separately. We then combine our results for the three regions into a single presentation. For example, we combine the PDs into a single PD for all of New York City.

See the file "readme.txt" in our repository
`https://bitbucket.org/jerryluo8/coveragetda/src/main/` for more information about the city boundaries.

4.3 Results

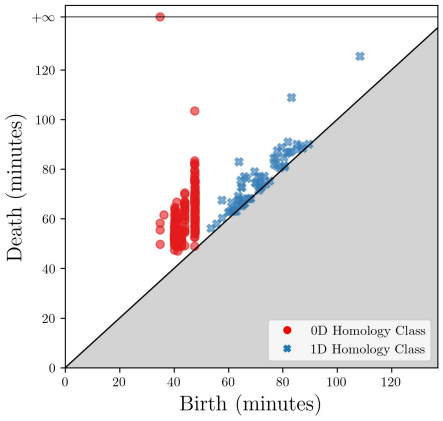
We compute the PH of the weighted VR filtrations of Section 4.2 for Atlanta, Chicago, Jacksonville, Los Angeles County, New York City, and Salt Lake City. We show their PDs in Figure 4.2. We examine both 0D and 1D homology classes. The 0D homology classes represent holes between different connected regions of coverage, and the 1D homology classes represent holes in coverage that are bounded by closed paths. A homology class that dies at filtration-parameter value t represents a hole in coverage that persists until time t. An individual who lives in a hole in coverage that dies at time t needs t minutes (including both waiting time at a polling site and travel time back and forth to the site) to cast a vote.

In our analysis, we emphasize homology-class death values. We view homology-class birth values as largely irrelevant to our application. A homology-class birth value indicates the filtration-parameter value at which a coverage hole materializes. We use birth values only in the following way. If the death value divided by the birth value (i.e., the "death/birth ratio") of a homology class is very small (i.e., it is close to 1), then it is possible that this class is merely an artifact of using a VR approximation of a Čech complex. We thus focus on homology classes whose death/birth ratios are at least 1.05.[8] Beyond this, we use only the homology-class death values and death simplices.

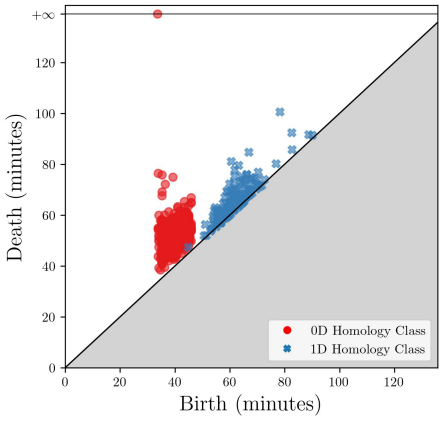
Larger homology-class death values suggest that a city may have worse coverage, and a wider distribution of death values suggests that there may be more variation in polling-site accessibility within a city. In Figure 4.3, we show a box plot of the distribution of homology-class death values for each city. In Table 4.1, we show the medians and variances of the 0D and 1D homology-class death values for each city.

We compare the coverages of the cities by examining the death values in the PDs. For

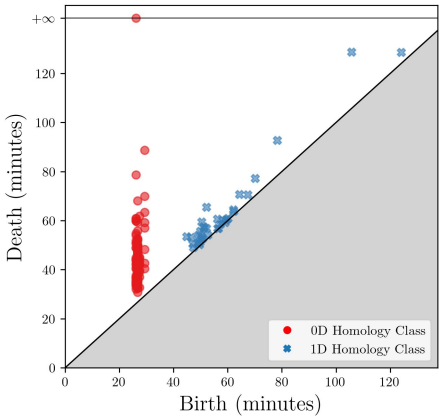
[8]Interested readers can explore thresholds other than 1.05 by using our data, which is available at `https://bitbucket.org/jerryluo8/coveragetda/src/main/`. We describe the data in detail in the "readme.txt" file.



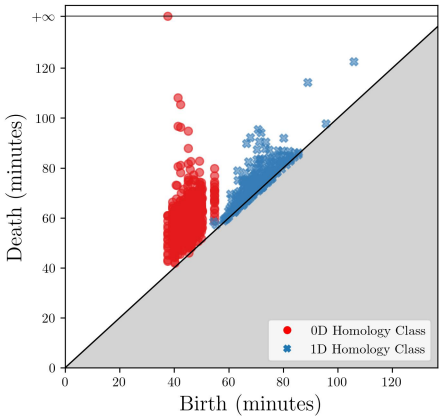
(a) Atlanta



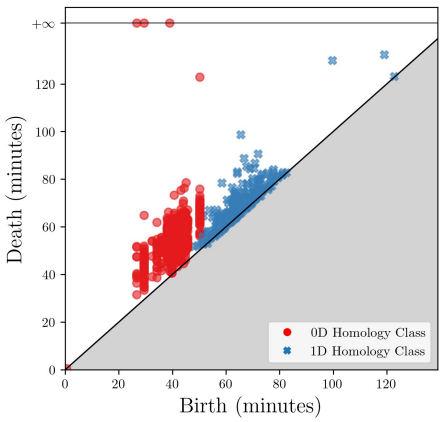
(b) Chicago



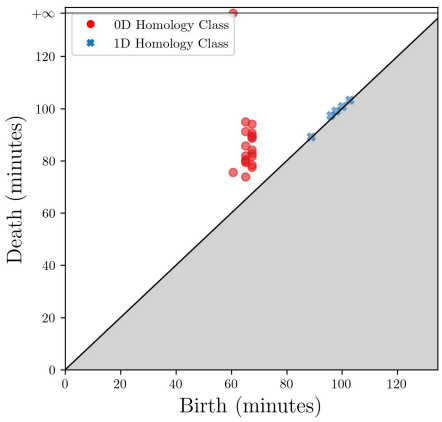
(c) Jacksonville (Florida)



(d) Los Angeles County



(e) New York City



(f) Salt Lake City

Figure 4.2: The PDs for each city for the PH of our weighted VR complexes.

City	Homology Dimension	Median (minutes)	Variance (minutes)
Atlanta	0	59.9	75.4
	1	77.1	150.8
Chicago	0	53.1	30.2
	1	66.3	59.7
Jacksonville (Florida)	0	42.8	75.7
	1	57.5	394.4
Los Angeles County	0	59.6	53.3
	1	76.1	84.6
New York City	0	65.1	49.2
	1	82.9	207.1
Salt Lake City	0	82.8	37.3
	1	N/A	N/A

Table 4.1: The medians and variances of the homology-class death values for each city. (As we discussed in the main text, we consider Los Angeles County, rather than only the city of Los Angeles.) We consider homology classes whose death/birth ratio is at least 1.05. Salt Lake City has no such 1D homology classes.

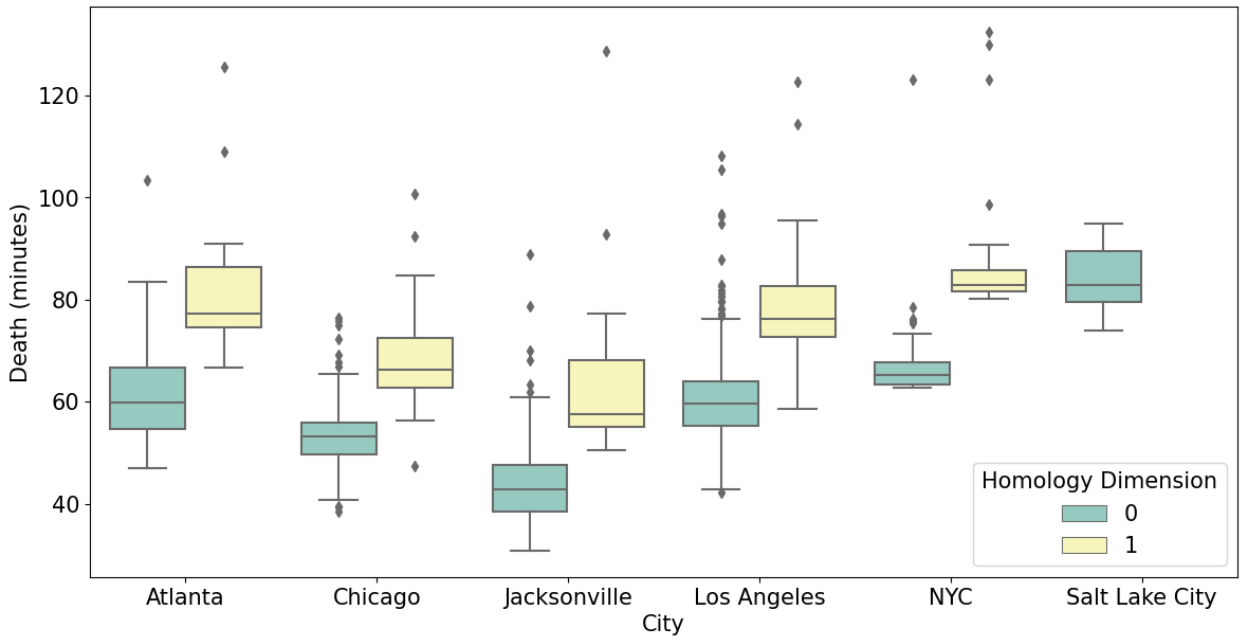
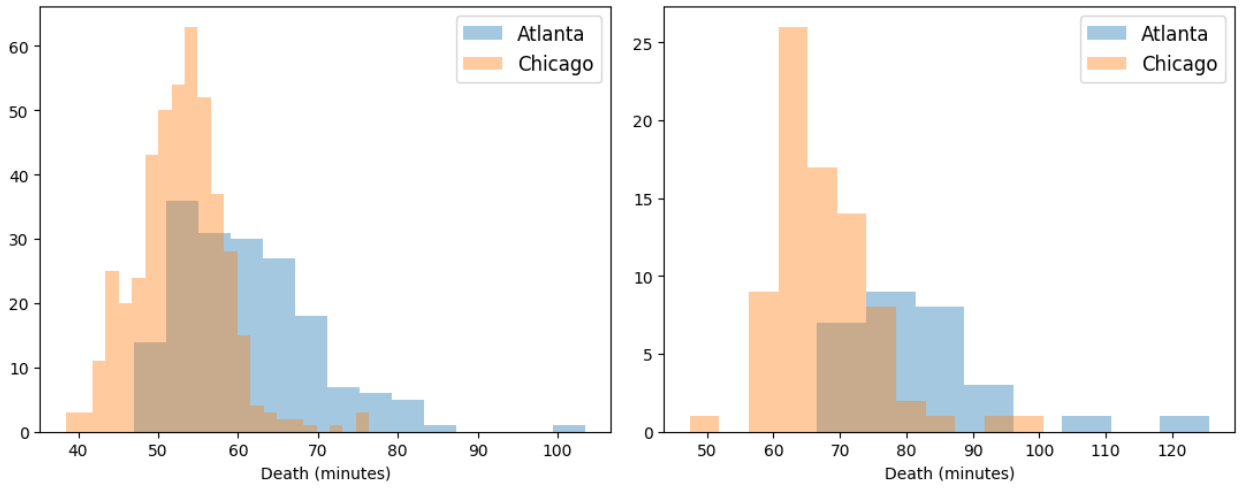


Figure 4.3: Box plots of the death values of the 0D and 1D homology classes for each city. We only consider homology classes whose death/birth ratio is at least 1.05. Salt Lake City has no such 1D homology classes.



(a) 0D homology classes

(b) 1D homology classes

Figure 4.4: Histograms of the death values of the 0D and 1D homology classes for Atlanta and Chicago. We only consider homology classes whose death/birth ratio is at least 1.05.

example, in the PDs for Atlanta and Chicago in Figure 4.2, we see that Atlanta's homology classes tend to die later than Chicago's homology classes. We also see this in the box plots Figure 4.3 and in Figure 4.4, in which we plot the distributions of death values for Atlanta and Chicago. Our PDs and visualizations of summary statistics suggest that Chicago has better polling-site coverage than Atlanta.

We use the death simplices to locate and visualize holes in polling-site coverage. We interpret the death simplex of a homology class as the "epicenter" of an associated coverage hole because the death simplex represents the last part of the hole to be covered. The death simplex of a 0D homology class is an edge between two polling sites; there is a hole in coverage between those two sites. Similarly, the death simplex of a 1D homology class is a triangle that is the convex hull of three polling sites; there is a hole in coverage between those three sites. In Figures 4.5 and 4.6, we show the death simplices with the largest death

values[9] for the 0D and 1D homology classes,[10] respectively. For example, consider panels (a) and (b) of Figure 4.5 and Figure 4.6, in which we show the 0D and 1D homology-class death simplices for Atlanta and Chicago. The areas of lowest coverage (i.e., the areas that have the death simplices with the largest death values) in Atlanta tend to be in the southwest, whereas the areas of lowest coverage in Chicago tend to be in the northwest and southeast. There is one 1D homology class in Atlanta that has a significantly larger death filtration value than the other classes in Atlanta and any of the classes in Chicago. This homology class represents a hole in coverage in southwest Atlanta (see Figure 4.6a).

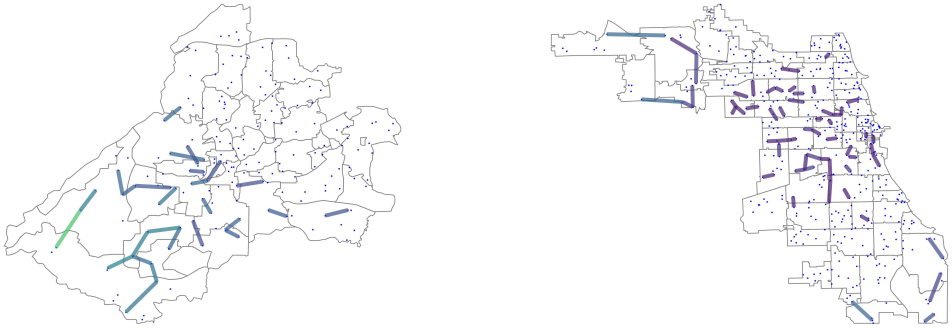
4.4 Conclusions and Discussion

4.4.1 Summary

We showed that PH is a helpful approach to study accessibility and equitability. It allows one to examine holes in resource coverage with respect to an appropriate choice of "distance," which one constructs to incorporate important features of a problem of interest. The distance can be based on geography, time, or something else. In the present chapter, we used PH to study and quantify holes in polling-site coverage in six United States cities (technically, in five cities and in Los Angeles County). For each city, we constructed a filtration in which a homology class that dies at time t represents a geographical region in which it takes t minutes to cast a vote (including both travel time and waiting time). We interpreted the

[9]More precisely, for each city and each homology dimension (0 and 1), we show the death simplices whose death values have a z-score of at least 1. We calculate the z-scores as follows. Let d be the death value of a p-dimensional homology class (where $p = 0$ or $p = 1$) for city C. The z-score of d is $z = (d - \mu_{C,p})/\sigma_{C,p}$, where $\mu_{C,p}$ and $\sigma_{C,p}$ are the mean and standard deviation of the distribution of death values of the p-dimensional homology classes for city C.

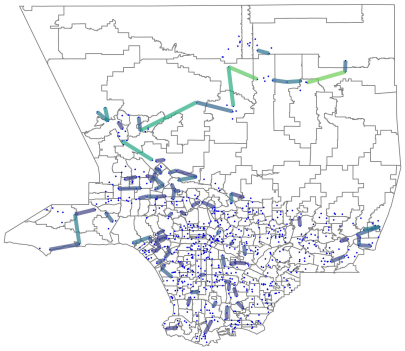
[10]In Figure 4.6, in which we show the death simplices of the 1D homology classes, some of the polling sites appear to be covered by death simplices whose vertices are other polling sites. At least two factors may contribute to this. One factor is that our measure of distance is not a Euclidean metric, even though we plot the death simplices in Figure 4.6 as Euclidean triangles. The Euclidean triangles can sometimes cover polling sites that are not among its vertices, but geodesic triangles may not cover those polling sites. Another possibility is that a polling site x has such a long waiting time that it does not show up in the filtration until after the homology class whose death simplex includes x has already died.



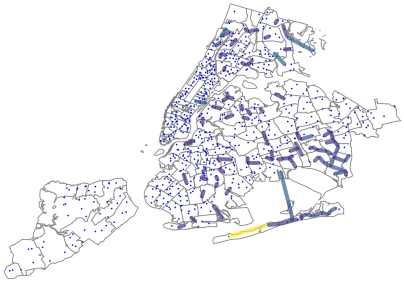
(a) Atlanta

(b) Chicago

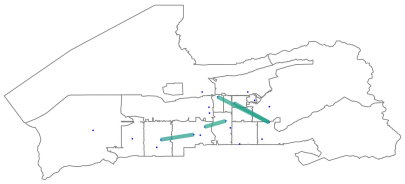
(c) Jacksonville (Florida)



(d) Los Angeles County



(e) New York City



(f) Salt Lake City

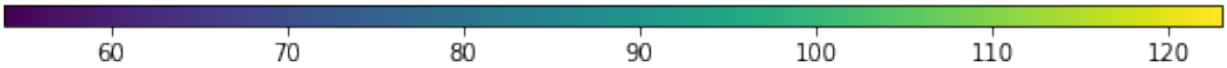
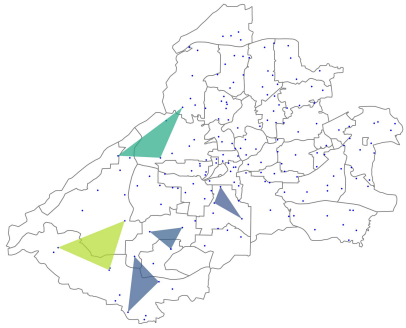
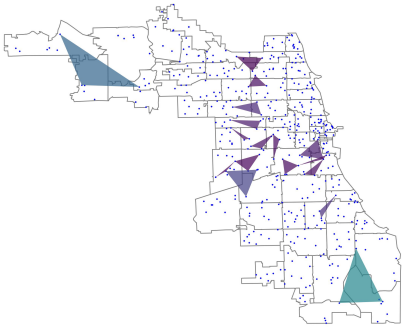


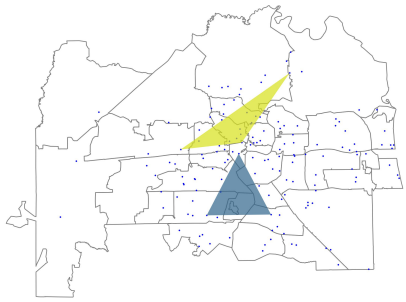
Figure 4.5: Death simplices with the largest death values for the 0D homology classes. The colors correspond to the death values (in minutes). We only consider homology classes whose death/birth ratio is at least 1.05.



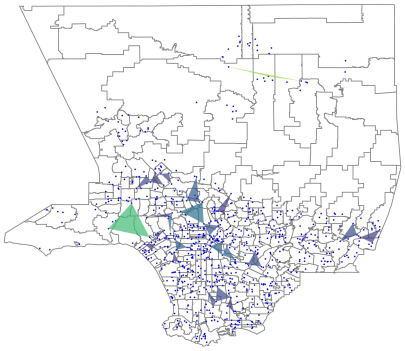
(a) Atlanta



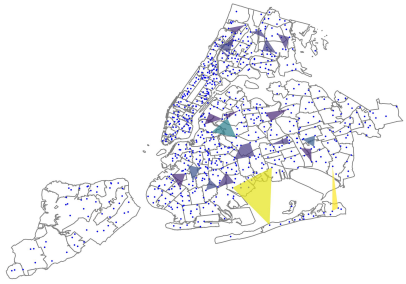
(b) Chicago



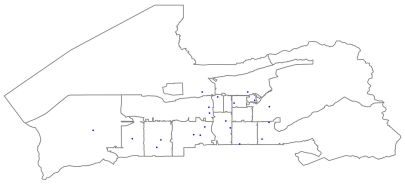
(c) Jacksonville (Florida)



(d) Los Angeles County



(e) New York City



(f) Salt Lake City

Figure 4.6: Death simplices with the largest death values for the 1D homology classes. The colors correspond to the death values (in minutes). We only consider homology classes whose death/birth ratio is at least 1.05.

death simplex of a homology class as the location of the corresponding hole in resource coverage. The information in the PH allowed us both to compare the accessibility of voting across our chosen cities and to determine the locations of the coverage holes within each city.

A key benefit of PH is that it enabled us to identify holes in polling-site coverage at all time scales. It also allowed us to use a distance that we designed for the problem at hand, rather than merely using geographical distance, which does not capture important factors in resource accessibility [BB06]. We based our distance function on estimates of travel time, which is more reasonable and accurate than geographical distance for capturing resource accessibility [PWB06].

4.4.2 Limitations

To conduct our study, we needed to estimate a variety of quantities (see Section 4.2), including travel times, waiting times, and demographic information. We also made several simplifications because of computational and monetary constraints. We now discuss some issues that are important to address before attempting to incorporate our approach into policy-making.

One limitation of our study is our estimation of travel times. As we discussed in Section 4.2.1, we computed travel times using the Google Maps API. Because of monetary constraints, we only computed a subset of the relevant travel times and used a graph-based estimate to determine the others. Additionally, we computed each travel time between polling sites only once. Computing more precise estimates of travel times is important to better capture the accessibility of polling sites. One way to do this is to compute travel times between the same two polling sites multiple times across different days and times of day and take an average. Such additional computations can also help yield estimates of best-case and worst-case scenarios.

Another limitation of our study is the granularity of our data. As we discussed in Sec-

tion 4.2.2, our waiting-time data is at the scale of congressional districts. Because there is heterogeneity in the waiting times at different polling sites in the same congressional district, it is important to obtain finer-grained data for the waiting times at polling sites. Having finer-grained waiting times (e.g., if possible, procuring an estimated waiting time for each polling site) would improve our ability to capture voting accessibility.

We also made several topological approximations. We worked with a weighted VR filtration, which approximates a weighted Čech filtration, which in turn approximates the nested set $\{\bigcup_i B(x_i, r_{x_i}(t))\}_{t\in\mathbb{R}}$ of spaces, where $\{x_i\}$ is a set of polling sites and $r_{x_i}(t)$ is the radius function that we defined in Section 4.2. The nested set of spaces is directly relevant to our application, as the holes in $\bigcup_i B(x_i, r_{x_i}(t))$ are the true holes in polling-site coverage. We made our approximations, which are standard in TDA and are well-justified [OPT17] (see our discussion in Section 2.2.2), to reduce computational cost. However, the convexity condition of the Nerve Theorem, which justifies the approximation of $\bigcup_i B(x_i, r_{x_i}(t))$ by a weighted Čech complex, is not guaranteed to be satisfied for all times t. The Nerve Theorem implies that the weighted Čech complex is homotopy-equivalent to $\bigcup_i B(x_i, r_{x_i}(t))$ whenever the balls $B(x_i, r_{x_i}(t))$ are convex. This condition always holds in Euclidean space, but it is not guaranteed to hold in the space that we defined in Section 4.2.[11] Homotopy-equivalence is important because homotopy-equivalent spaces have the same homology and thus have the same set of holes.

Finally, our approach only detects holes in the convex hull of a set of resource sites. Although this may be inconsequential if resource sites are sufficiently spread out geographically, it can

[11] Although our space is not Euclidean, it is still reasonable to assume that it is approximately *locally* Euclidean. That is, for each polling site x, there is a constant $a > 0$ such that if y is sufficiently close to x, then $d(x, y) \approx a \cdot d_E(x, y)$, where $d(x, y)$ is defined by (4.1) and $d_E(x, y)$ is the Euclidean distance. This approximation holds because car-ownership rates and traffic conditions do not vary much within a sufficiently small neighborhood. We verified empirically that our distance function is approximately locally Euclidean by showing that, for each polling site x, there is a strong linear correlation between the pairwise distances $d(x, y)$ and the pairwise Euclidean distances $d_E(x, y)$ when y is sufficiently close to x. Because our distance function is approximately locally Euclidean, sufficiently small balls (with respect to our distance function) behave like Euclidean balls, so the Nerve Theorem is applicable for sufficiently small filtration values.

be problematic if the resource sites are overly concentrated near a few locations. One way to address this issue is to incorporate city boundaries into the construction of the filtrations. This would help capture holes in coverage in regions that lie outside the convex hull of the resource sites, and it would also help identify the filtration-parameter value t at which an entire city is covered by the balls $B(x_i, r_{x_i}(t))$.

4.4.3 Future work

As we discussed in Section 4.4.2, we made several topological approximations of our mathematical object of interest, which is the nested set $\{\bigcup_i B(x_i, r_{x_i}(t))\}_{t\in\mathbb{R}}$ of spaces. Instead of using a weighted VR filtration, one can construct a more direct approximation of $\{\bigcup_i B(x_i, r_{x_i}(t))\}_{t\in\mathbb{R}}$. One can first discretize a city by imposing a grid on it. For each point on such a grid, one can then construct the filtered cubical complex that is induced by the travel time to the nearest polling site. However, this is much more computationally expensive than our approach, and it would also entail many more travel-time queries (which cost money) than in the present chapter.[12]

It is also important to incorporate city boundaries into the construction of filtrations. One way to do this is as follows. Let $x_1, \ldots, x_n$ denote the resource sites, and let $y_1, \ldots, y_m$ denote the points that one obtains by discretizing a city boundary. One can extend our distance function (4.1) by defining[13]

$$d(x_i, y_j) := \frac{2}{P}[P_{Z(x_i)}\tilde{d}(x_i, y_j) + P_{Z(y_j)}\tilde{d}(y_j, x_i)]\,, \tag{4.2}$$

[12]Our distance function (4.1) is symmetric, but recall that it is not a metric because it does not satisfy the triangle inequality. Therefore, we cannot use techniques such as distance transforms and level-set propagation to reduce the computational complexity of calculating the filtration $\{\bigcup_i B(x_i, r_{x_i}(t))\}_{t\in\mathbb{R}}\}$.

[13]The factor of 2 arises from the fact that x_i is a resource site but y_j is not a resource site.

where P, Z, and $\tilde{d}$ are the same as in the distance function (4.1) and

$$d(y_i, y_j) = \begin{cases} 0\,, & y_i \text{ and } y_j \text{ are adjacent points of the discretized city boundary} \\ \infty\,, & \text{otherwise}\,. \end{cases} \tag{4.3}$$

At each filtration-parameter value, the simplicial complex that one constructs using the distance function (4.1) with the extensions (4.2) and (4.3) includes both the points that one obtains by discretizing the boundary and the edges that connect adjacent boundary points. The largest death value is then the filtration-parameter value t that corresponds to the time at which an entire city is covered by the balls $\{B(x_i, r_{x_i}(t))\}$ (i.e., when there are no longer any holes in coverage).

In this chapter, we used death simplices to locate holes in coverage, but other approaches are also possible. For example, by calculating minimal generators [LTH21], one can identify representative cycles that encircle holes. The topological pipeline "hyperTDA" was introduced recently [BYM22] to analyze the structure of minimal generators by constructing a hypergraph, calculating hypergraph centrality measures, and employing community detection. This approach may provide insights into the spatial structure of minimal generators. Another potentially viable approach is to use decorated merge trees (DMTs) [CHM22] to locate holes in coverage. DMTs allow one to match holes with associated clusters of points.

Although we explored a specific case study (namely, the accessibility of polling sites), it is also relevant to conduct similar investigations for other resources, such as public parks, hospitals, vaccine distribution centers, grocery stores, Planned Parenthood clinics, and DMV locations. One can use similar data to construct a filtration, although it may be necessary to modify the choices of distance and weighting. One can also use ideas from mobility theory [BBG18] to help construct suitable distances and weightings. For example, all DMV offices offer largely the same services, so it seems reasonable to assume that people will go to their nearest office. Therefore, in a study of DMV accessibility, it seems appropriate to use

travel time as a distance function, just as we did in our analysis of polling sites. However, in other applications, it is not reasonable to use travel time alone as a distance function. For example, different grocery stores may offer different products at different prices, so travel time alone may not be appropriate as a choice of distance function. Additionally, although waiting time is a significant factor for investigating the coverage of polling sites, there are many applications for which it does not make sense to incorporate waiting time. For example, the time that is spent in a public park or recreation center is typically not a barrier to access. In applications in which waiting time is not an accessibility factor, it seems more appropriate to use a standard VR filtration than a weighted VR filtration. With salient modifications (such as those that we described in this subsection and in Section 4.4.2), we can apply our approach to many other types of resource sites.

CHAPTER 5

Bounded-Confidence Models of Opinion Dynamics with Adaptive Confidence Bounds

In this chapter, we study bounded-confidence models (BCMs) with adaptive confidence bounds. As we discussed in Section 1.2, a BCM is a model of opinion dynamics on a network such that agents (represented by network nodes) have continuous-valued opinions and are only receptive to neighbors whose opinions are sufficiently similar (i.e., they differ by less than a "confidence bound"). We formulate and analyze two new BCMs that incorporate distinct, time-dependent confidence bounds for each dyad (i.e., pair of adjacent nodes). One of our models generalize the Hegselmann–Krause (HK) model (see Section 5.2.1) and the other generalizes the Deffuant–Weisbuch (DW) model (see Section 5.2.3). This chapter is adapted[1] from [LLP23], which was led jointly by Grace J. Li and me and is co-authored with Mason A. Porter.[2]

This chapter proceeds as follows. In Section 5.1, we introduce the context and motivation for the formulation of our BCMs. In Section 5.2, we introduce our adaptive-confidence BCMs and discuss associated baseline BCMs. In Section 5.3, we give theoretical guarantees for our adaptive-confidence BCMs. We describe the specifications of our numerical simulations

[1] All figures in this chapter originally appeared in [LLP23].

[2] I contributed the main idea for this project (incorporating adaptive confidence bounds in BCMs); formulated our adaptive BCMs with G. J. Li; formulated and proved Theorem 5.2, Theorem 5.6, Theorem 5.8, and Theorem 5.11; and wrote the paper with all co-authors.

in Section 5.4 and the results of our numerical simulations in Section 5.5. In Section 5.6, we summarize our main results and discuss possible avenues of future work. In Section 5.A and Section 5.B, we prove the results of Section 5.3.2. We present additional numerical results for our adaptive-confidence HK model in Section 5.C, and we present additional numerical results for our adaptive-confidence DW model in Section 5.D. Our code and plots are in our code repository, which is available at `https://gitlab.com/graceli1/Adaptive-Confidence-BCM`.

5.1 Introduction and Motivation of Our Models

An individual's opinion on a topic is often influenced by the people with whom they interact [Jac08], and researchers in many disciplines study such interactions and how they change opinions and actions [Noo20]. Individuals are often influenced most by people and other sources whose opinions are similar to theirs [CM11]. This phenomenon is encapsulated in a simple form in BCMs [NVT20, HK02, DNA00], in which the nodes of a network have continuous-valued opinions and interacting nodes influence each others' opinions if and only if their opinions are sufficiently similar. A key feature of BCMs is the presence of a confidence bound, which determines which nodes can influence each other. A node can only influence and be influenced by its neighbors when the difference between their opinions is less than their confidence bound.

The two most popular BCMs are the HK model [Kra00, HK02] and the DW model [DNA00]. In the HK and DW models, the confidence bound is a constant scalar value that is shared by all dyads (i.e., pairs of adjacent nodes). In the present chapter, we formulate and study adaptive-confidence BCMs that generalize the HK and DW models by incorporating distinct, time-dependent confidence bounds for each dyad. The confidence bounds in our adaptive-confidence BCMs change after nodes interact with each other. These changes highlight the idea that the quality of an interaction between individuals can affect how much they trust

each other [GAP13,CS20,LMM17]. For example, in online marketplaces, trust between users depends on their past experiences with each other and on the reported experiences of other users in reputation systems [SNP13, RKK07, RKZ00]. The word "trust" can have different meanings in different disciplines; one interpretation is that trust represents an expectation about future behavior [SNP13]. Rather than considering trust, our BCMs use a notion of "receptiveness," which encodes the willingness of an individual to consider the future opinions of another individual. When two nodes interact with each other, their mutual receptiveness changes. See [BCF07, XLC17, NGW23] for other opinion models with interaction-influenced receptiveness.

In our adaptive-confidence BCMs, when two nodes successfully compromise their opinions in an interaction (i.e., they have a "positive interaction"), they become more receptive to each other. Likewise, when two nodes interact but do not change their opinions (i.e., they have a "negative interaction"), they become less receptive to each other. When nodes i and j interact and influence each others' opinions (i.e., their current opinion difference is smaller than their current confidence bound), we increase their confidence bound c_{ij}. When nodes i and j interact and do not influence each others' opinions (i.e., their current opinion difference is at least as large as their current confidence bound), we decrease their confidence bound c_{ij}. In our adaptive-confidence BCMs, each dyad has a distinct confidence bound and interactions are symmetric (i.e., either both nodes influence each other or neither node influences the other). One can interpret the increase of a dyadic confidence bound in our BCMs as nodes becoming more receptive to nodes with whom they compromise, and one can interpret the decrease of a dyadic confidence bound as nodes becoming less receptive to nodes with whom they do not compromise. When nodes in our BCMs have a negative interaction, they adapt their dyadic confidence bounds, but their opinions stay the same. Other researchers have considered BCMs with "repulsion," in which the opinions of interacting nodes with sufficiently different opinions move farther apart from each other [AC15, HDJ08, KF23].

In the present chapter, we study the time evolution and long-term behaviors of our adaptive-confidence HK and DW models. We examine the formation of "limit opinion clusters" (i.e., sets of nodes that converge to the same opinion), the dynamics of the confidence bounds, and the convergence rate of the opinions. We simulate our models on various networks (see Section 5.4.1) and study the time evolution of their associated "effective graphs," which are time-dependent subgraphs of a network with edges only between nodes that are receptive to each other (see Section 5.3). We show numerically that our adaptive-confidence BCMs tend to have less "opinion fragmentation"[3] than their associated baseline (i.e., nonadaptive) BCMs. In our numerical simulations, we study "final" opinion clusters (see Section 5.4.2) to approximate limit opinion clusters. We demonstrate numerically that the connected components of the final effective graphs in our BCMs can have more complicated structures than those of the baseline BCMs.

5.1.1 Related work

There has been much research on standard (i.e., nonadaptive) HK and DW models on networks through numerical simulations [MVP18, For04, Lor06, HK02] and both heuristic analytical arguments and mathematically rigorous proofs [BKR03, Lor05, Lor08, HK02]. The DW model was studied initially on both a fully-mixed population (i.e., a complete network) and on a square-lattice network [DNA00], and the HK model was studied initially only on a fully-mixed population [HK02]. Subsequently, the DW and HK models have been studied on a variety of networks [For05, MVP18, SFH21]. See [NVT20, Noo20, BAP24] for reviews of research on the standard DW and HK models and their generalizations.

Many researchers have generalized the HK and DW models by incorporating heterogeneity into the confidence bounds. Lorenz [Lor09] extended these BCMs so that each node has

[3] In our study, "opinion fragmentation" signifies the existence of at least two "major" opinion clusters, which include more than 1% of the nodes of a network. In Section 5.4.3, we give more detail about how we define opinion fragmentation and major opinion clusters.

its own confidence bound, which can result in asymmetric influence and opinion updates. Using numerical simulations, Lorenz demonstrated that these BCMs are more likely than the baseline BCMs to reach a consensus state when there are both open-minded and close-minded nodes (which have large and small confidence bounds, respectively). By analyzing the heterogeneous-confidence DW model of [Lor09] on a complete graph, Chen et al. [CSM20] proved almost-sure convergence of opinions for certain parameter values and derived sufficient conditions for the nodes of a network to eventually reach a consensus. In a related work, Chen et al. [CSD20] examined a heterogeneous HK model with "environmental noise" (e.g., from media sources) and showed that heterogeneous confidence bounds in this setting can yield larger differences in node opinions in the infinite-time limit. Su et al. [SGW17] examined the heterogeneous-confidence HK model of [Lor09] and proved that at least some nodes of a network converge to a steady-state opinion in finite time.

Researchers have also incorporated edge-based heterogeneities in the confidence bounds of BCMs. Etesami [Ete19] examined an HK model on networks with time-independent edge-heterogeneous confidence bounds and proved that their model is Lyapunov stable. Shang [Sha14] studied a DW model in which each edge has a confidence bound that takes a value from an independent and identically distributed Poisson process. They derived sufficient conditions for consensus to occur almost surely for a one-dimensional lattice graph.

Other generalizations of BCMs and related opinion models generalize the model parameters by making them time-dependent or adaptive. Weisbuch et al. [WDA02] studied a generalized DW model in which each node has a heterogeneous, time-dependent confidence bound that is proportional to the standard deviation of the opinions that that node observed in all prior interactions. They also considered a variant of their model that places more weight on the observed opinions from recent interactions. Deffuant et al. [DAW02] examined a DW model with "relative agreement". In their model, each node has an uncertainty parameter that determines (1) whether it and the node with which it interacts influence each other and

(2) the amount by which they influence each other. A node changes both its opinion and its uncertainty when it is influenced by another node. Bagnoli et al. [BCF07] considered a BCM on complete graphs in which each pair of adjacent nodes (i.e., each dyad) has an associated time-dependent affinity value (which determines whether or not they can influence each other) that depends on the magnitude of their opinion difference. Chacoma and Zanette [CZ15] examined opinion and confidence changes in a questionnaire-based experiment, and they then proposed an agent-based opinion model based on the results of their experiment. Their model is not a BCM, but it does incorporate a notion of time-dependent confidence between nodes. Bernardo, Vasca, and Iervolino [VBI21, BVI22] developed variants of the HK model in which nodes have individual, time-dependent confidence bounds[4] that depend on the opinions of neighboring nodes. In their models, nodes adapt their confidence bounds through a heterophilic mechanism (i.e., they seek neighboring nodes whose opinions differ from theirs). By contrast, in our models, nodes do not actively seek neighbors with different opinions. Instead, their mutual receptiveness increases when their opinions are sufficiently close to each other.

In this chapter, we incorporate adaptivity into the confidence bounds of BCMs, but one can instead incorporate adaptivity in the network structures of BCMs [KB08a, KB08b, DSC17, KFP23].[5] Kozma and Barrat [KB08a, KB08b] modified the DW model to allow rewiring of "discordant" edges, which occur between nodes whose opinions differ from each other by more than the confidence bound. In their model, rewired edges connect to new nodes uniformly at random. Recently, Kan et al. [KFP23] generalized this model by including both a confidence bound and an opinion-tolerance threshold, with discordant edges occurring between nodes whose opinions differ by more than that threshold. They incorporated opinion homophily into the rewiring probabilities, so nodes are more likely to rewire to nodes with more similar opinions. They observed in numerical simulations that it is often harder to achieve consensus

[4]The confidence bounds update with time in different ways in the models in [VBI21] and [BVI22].

[5]See the reviews [SBL23, BGK23] for discussions of various notions of adaptivity in dynamical systems.

in their adaptive DW model than in an associated baseline DW model.

There has been much theoretical development of models of opinion dynamics, and it is important to empirically validate these models [Vaz22, GOD21]. Some researchers have used questionnaires [CZ15, VMG16, TFM16] or data from social-media platforms [Koz22a, Koz23] to examine how opinions change in controlled experimental settings. Another approach is to develop models of opinion dynamics that infer model parameters [CLP24, Koz22b] or opinion trajectories [MDB20] from empirical data. There are many challenges to developing and validating models of opinion dynamics that represent real-world situations [Mas19, BAB21], but mechanistic modeling is valuable, as it (1) forces researchers to clearly specify relationships and assumptions during model development and (2) provides a framework to explore complex social phenomena [Vaz22, HL15].

5.2 Baseline and Adaptive BCMs

We extend the DW and HK models by introducing adaptive confidence bounds. For both the HK and DW models, which we study on networks, we first present the baseline BCM and then introduce our adaptive-confidence generalization of it. The nodes in our BCMs represent agents that have opinions that lie in the closed interval $[0, 1]$. Let $G = (V, E)$, where V is the set of nodes and E is the set of edges, denote a time-independent, unweighted, and undirected graph without self-edges or multi-edges. The edges in the set E specify which pairs of nodes can interact with each other at each discrete time t. Let $N = |V|$ denote the number of nodes of the graph (i.e., network), $x_i(t)$ denote the opinion of node i at time t, and $\vec{x}(t)$ denote the vector of the opinions of all nodes at time t (i.e., the entry $[\vec{x}(t)]_i = x_i(t)$). We denote the edge that is attached to adjacent nodes i and j by (i, j).

5.2.1 The Hegselman–Krause (HK) model

The baseline HK model [Kra00, HK02] is a discrete-time synchronous BCM on a time-independent, unweighted, and undirected graph $G = (V, E)$ with no self-edges or multi-edges.[6] At each time t, we update the opinion of each node i by calculating

$$x_i(t+1) = |I(i, x(t))|^{-1} \sum_{j \in I(i,x(t))} x_j(t)\,, \tag{5.1}$$

where[7] $I(i, x(t)) = \{i\} \cup \{j \mid |x_i(t) - x_j(t)| < c \text{ and } (i, j) \in E\} \subseteq \{1, 2, \ldots, N\}$. The *confidence bound* c controls the "open-mindedness" of nodes to different opinions. We say that adjacent nodes i and j are *receptive* to each other at time t if their opinion difference is less than the confidence bound c (i.e., $|x_i(t) - x_j(t)| < c$). Accordingly, $I(i, x(t))$ is node i itself along with all adjacent nodes to which i is receptive.[8] The confidence bound c in the baseline HK model is homogeneous (i.e., the confidence bound is the same for all dyads) and time-independent.

5.2.2 Our HK model with adaptive confidence bounds

Our HK model with adaptive confidence bounds is similar to the baseline HK model with update rule (5.1), but now each edge $(i, j) \in E$ has a dyadic confidence bound $c_{ij}(t) \in [0, 1]$ that is time-dependent and changes after each interaction between the nodes in that dyad. We refer to this model as our *adaptive-confidence HK model*. Instead of a fixed confidence bound, there is an initial confidence bound $c_0 \in (0, 1)$ and we initialize all of the confidence

[6] The HK model was examined initially on a fully-mixed population [HK02], but we use its extension to networks (see e.g., [For05, PFT18, SFH21]) as our "baseline HK model".

[7] In [Kra00, HK02], $I(i, x(t)) = \{i\} \cup \{j \mid |x_i(t) - x_j(t)| \leq c \text{ and } (i, j) \in E\}$. We use a strict inequality to be consistent with the strict inequality in the DW model.

[8] An alternative interpretation is that each node has a self-edge. We do not use this interpretation.

bounds[9] to $c_{ij}(0) = c_0$ for each edge $(i, j) \in E$. There is also a confidence-increase parameter $\gamma \in [0, 1]$ and a confidence-decrease parameter $\delta \in [0, 1]$, which control how much $c_{ij}(t)$ increases and decreases, respectively, after each interaction.

At each time t, we update the opinion of each node i by calculating

$$x_i(t+1) = |I(i, x(t))|^{-1} \sum_{j \in I(i,x(t))} x_j(t)\,, \tag{5.2}$$

where[10] $I(i, x(t)) = \{i\} \cup \{j \mid |x_i(t) - x_j(t)| < c_{ij}(t) \text{ and } (i, j) \in E\} \subseteq \{1, 2, \ldots, N\}$. Adjacent nodes i and j are receptive to each other at time t if their opinion difference is less than their dyadic confidence bound c_{ij} (i.e., $|x_i(t) - x_j(t)| < c_{ij}(t)$). At each time, we also update each confidence bound c_{ij} by calculating

$$c_{ij}(t+1) = \begin{cases} c_{ij}(t) + \gamma(1 - c_{ij}(t))\,, & \text{if } |x_i(t) - x_j(t)| < c_{ij}(t) \\ \delta c_{ij}(t)\,, & \text{if } |x_i(t) - x_j(t)| \geq c_{ij}(t)\,. \end{cases} \tag{5.3}$$

That is, if the opinion difference between nodes i and j is smaller than their confidence bound at time t, their associated dyadic confidence bound c_{ij} increases; otherwise, their dyadic confidence bound decreases. Larger values of γ correspond to sharper increases in the receptiveness between nodes when nodes compromise their opinions. Smaller values of δ correspond to sharper drops in the receptiveness between nodes when nodes interact but do not compromise.

Because $c_0 \in (0, 1)$ and $\gamma, \delta \in [0, 1]$, the update rule (5.3) preserves $c_{ij}(t) \in (0, 1)$ for each

[9] When $c_0 = 0$, nodes are never receptive to their neighbors (i.e., $c_{ij}(t) = 0$ for all adjacent nodes i and j at all times t). When $c_0 = 1$, all nodes are always receptive to all of their neighbors (i.e., $c_{ij}(t) = 1$ for all adjacent nodes i and j at all times t). We do not examine these values of c_0.

[10] Although Equations (5.1) and (5.2) look the same, they use different definitions of the quantity $I(i, x(t))$. Equation (5.1) has a homogeneous and time-independent confidence bound, whereas Equation (5.2) has heterogeneous and adaptive confidence bounds.

edge (i, j) and all times t. If $(\gamma, \delta) = (0, 1)$, then $c_{ij}(t) = c_0$ for all t and all edges $(i, j) \in E$. That is, the confidence bounds are homogeneous and time-independent, so our adaptive-confidence HK model reduces to the baseline HK model.

5.2.3 The Deffuant–Weisbuch (DW) model

The baseline DW model [DNA00] is a discrete-time asynchronous BCM on a time-independent, unweighted, and undirected graph $G = (V, E)$ with no self-edges or multi-edges. At each time t, we choose an edge $(i, j) \in E$ uniformly at random. If nodes i and j are receptive to each other (i.e., if the opinion difference $|x_i(t) - x_j(t)|$ is less than the confidence bound c), we update the opinions of these nodes by calculating

$$\begin{aligned} x_i(t+1) &= x_i(t) + \mu(x_j(t) - x_i(t))\,, \\ x_j(t+1) &= x_j(t) + \mu(x_i(t) - x_j(t))\,, \end{aligned} \tag{5.4}$$

where $\mu \in (0, 0.5]$ is the compromise parameter[11]. Otherwise, the opinions x_i and x_j remain the same. At a given time t, we do not update the opinions of any nodes other than i and j. The confidence bound c in our baseline model is homogeneous (i.e., the confidence bound is the same for all dyads) and time-independent. As in the HK model, the confidence bound c controls the open-mindedness of nodes to different opinions. The compromise parameter μ indicates how much nodes adjust their opinions when they interact with a node to whom they are receptive. When $\mu = 0.5$, two interacting nodes that are receptive to each other precisely average their opinions; when $\mu \in (0, 0.5)$, interacting nodes that are receptive to each other move towards each others' opinions, but they do not adopt the mean opinion. Unlike in the HK model, the asynchronous update rule (5.4) of the DW model incorporates only pairwise opinion updates.

[11] Alternatively, one can consider $\mu \in (0, 1)$ as in Meng et al. [MVP18], although this is an uncommon choice. When $\mu > 0.5$, nodes "overcompromise" when they change their opinions; they overshoot the mean opinion and change which side of the mean opinion they are on.

5.2.4 Our DW model with adaptive confidence bounds

We refer to our DW model with adaptive confidence bounds as our *adaptive-confidence DW model*. As in the baseline DW model, there is a compromise parameter $\mu \in (0, 0.5]$. As in our adaptive-confidence HK model, we initialize the confidence bounds in our adaptive-confidence DW model to be $c_{ij}(0) = c_0$, where $c_0 \in (0, 1)$ is the initial confidence bound.[12] There again is a confidence-increase parameter $\gamma \in [0, 1]$ and a confidence-decrease parameter $\delta \in [0, 1]$, which control how much $c_{ij}(t)$ increases and decreases, respectively, after each interaction.

At each time t, we select an edge $(i, j) \in E$ uniformly at random. If nodes i and j are receptive to each other (i.e., if $|x_i(t) - x_j(t)| < c_{ij}(t)$), we update the opinions of nodes i and j using the DW update rule (5.4). Otherwise, the opinions x_i and x_j remain the same. We also update the dyadic confidence bound c_{ij} using Equation (5.3). That is, if the opinions of nodes i and j differ by less than their current dyadic confidence bound at time t, the confidence bound increases; otherwise, it decreases. The update rules preserves $c_{ij}(t) \in (0, 1)$ for each edge (i, j) and all times t. All other opinions and confidence bounds remain the same. Our adaptive-confidence DW model reduces to the baseline DW model when $(\gamma, \delta) = (0, 1)$.

5.3 Theoretical Results

We now discuss some theoretical guarantees of our BCMs.

As a consequence of Theorem 5.1 (which we state shortly), the opinion of each node in our BCMs converges to some limit value. We define the *limit opinion* x^i of node i as $\lim_{t\to\infty} x_i(t)$.

[12] As in our adaptive-confidence HK model, when $c_0 = 0$, nodes are never receptive to any of their neighbors (i.e., $c_{ij}(t) = 0$ for all adjacent nodes i and j at all times t). When $c_0 = 1$, nodes are always receptive to all of their neighbors (i.e., $c_{ij}(t) = 1$ for all adjacent nodes i and j at all times t). We do not examine these values of c_0.

We say that nodes i and j are in the same *limit opinion cluster* if

$$\lim_{t\to\infty} x_i(t) = \lim_{t\to\infty} x_j(t)\,. \tag{5.5}$$

Therefore, Equation (5.5) gives an equivalence relation on the set of nodes; the limit opinion clusters are the equivalence classes.

Let $G = (V, E)$ be a time-independent, unweighted, and undirected graph without self-edges or multi-edges. We study our BCMs on such graphs. A graph G in a BCM has an associated time-dependent *effective graph* $G_{\text{eff}}(t)$, which is a subgraph of G with edges only between nodes that are receptive to each other at time t.[13] That is,

$$\begin{aligned} G_{\text{eff}}(t) &= (V, E_{\text{eff}}(t))\,, \\ E_{\text{eff}}(t) &= \{(i,j) \in E \text{ such that } |x_i(t) - x_j(t)| < c_{ij}(t)\}\,. \end{aligned} \tag{5.6}$$

Consider the following theorem, which was stated and proved by Lorenz [Lor05].

Theorem 5.1. *Let $\{A(t)\}_{t=0}^{\infty} \in \mathbb{R}_{\geq 0}^{N\times N}$ be a sequence of row-stochastic matrices. Suppose that each matrix satisfies the following properties:*

(1) The diagonal entries of $A(t)$ are positive.

(2) For each $i, j \in \{1, \ldots, N\}$, we have that $[A(t)]_{ij} > 0$ if and only if $[A(t)]_{ji} > 0$.

(3) There is a constant $\alpha > 0$ such that the smallest positive entry of $A(t)$ for each t is larger than α.

[13] Other researchers have referred to the effective graph as a "confidence graph" [BAP24], a "communication graph" [BBC13], and a "corresponding graph" [YDH14].

Given times t_0 and t_1 with $t_0 < t_1$, let

$$A(t_0, t_1) = A(t_1 - 1) \times A(t_1 - 2) \times \cdots \times A(t_0)\,. \tag{5.7}$$

If conditions (1)–(3) are satisfied, then there exists a time t' and pairwise-disjoint classes $\mathcal{I}_1 \cup \cdots \cup \mathcal{I}_p = \{1, \ldots, N\}$ such that if we reindex the rows and columns of the matrices in the order $\mathcal{I}_1, \ldots, \mathcal{I}_p$, then

$$\lim_{t\to\infty} A(0,t) = \begin{bmatrix} K_1 & & 0 \\ & \ddots & \\ 0 & & K_p \end{bmatrix} A\,(0,t')\ , \tag{5.8}$$

where each K_q, with $q \in \{1, 2, \ldots, p\}$, is a row-stochastic matrix of size $|\mathcal{I}_q| \times |\mathcal{I}_q|$ whose rows are all the same.

As stated in [Lor05], Theorem 5.1 guarantees that the opinion of each node convergences to a limit opinion in the baseline HK and DW models. Because the node opinions in our adaptive-confidence HK and DW models update in the same way as in the corresponding baseline BCMs (see (5.1) and (5.4), respectively), it follows that the node opinions in our models also converge to a limit opinion.

5.3.1 Adaptive-confidence HK model

5.3.1.1 Confidence-bound analysis

In Theorem 5.2, we give our main result about the behavior of the confidence bounds (which update according to (5.3)) in our adaptive-confidence HK model.

Theorem 5.2. *In our adaptive-confidence HK model (with update rules* (5.2) *and* (5.3)*) with parameters $\gamma \in (0, 1]$ and $\delta \in [0, 1)$, the dyadic confidence bound $c_{ij}(t)$ of each pair of adjacent nodes, i and j, converges either to* 0 *or to* 1*. Furthermore, if i and j are in different*

limit opinion clusters, then $c_{ij}(t)$ converges to 0.

We prove Theorem 5.2 by proving Lemma 5.3, Lemma 5.4 and Lemma 5.5, which we state shortly. Because $c_{ij}(t) \in [0, 1]$, Lemma 5.4 gives convergence (because an eventually monotone[14] sequence in $[0, 1]$ must converge). By Lemma 5.5, we then have convergence either to 0 or to 1. Furthermore, by Lemma 5.4, if nodes i and j are in different limit opinion clusters, then $c_{ij}(t)$ must eventually be strictly decreasing and hence must converge to 0. However, if adjacent nodes i and j are in the same limit opinion cluster, then c_{ij} does not necessarily converge to 1. In fact, as we discuss in Section 5.5, our numerical simulations suggest that it is possible for the confidence bound of adjacent nodes in the same limit opinion cluster to instead converge to 0.

In Theorem 5.2, Lemma 5.4, and Lemma 5.5, we consider our adaptive-confidence HK model with parameters $\gamma \in (0, 1]$ and $\delta \in [0, 1)$. These parameter restrictions preclude the baseline HK model (which is equivalent to our adaptive-confidence HK model with $(\gamma, \delta) = (0, 1)$). However, in Lemma 5.3, we consider $\gamma \in [0, 1]$ and $\delta \in [0, 1]$. Therefore, Lemma 5.3 also applies to the baseline HK model, so we use it in our proof of Theorem 5.7 for the baseline HK model.

Lemma 5.3. *Consider our adaptive-confidence HK model (with update rules* (5.2) *and* (5.3)*) with parameters $\gamma \in [0, 1]$ and $\delta \in [0, 1]$. There is a time T such that no adjacent nodes i and j in different limit opinion clusters (i.e., $x^i \neq x^j$) are receptive to each other (i.e., $|x_i(t) - x_j(t)| < c_{ij}(t)$) at any time $t \geq T$.*

Proof. Consider a pair of adjacent nodes, i and j, that are in different limit opinion clusters (i.e., $x^i \neq x^j$). Let d be 1 more than the largest degree of a node of the graph G; that is,

[14]We say that a discrete time series $a(t)$ is "eventually monotone increasing" (respectively, "eventually monotone decreasing") if there exists a time T such that $a(t + 1) \geq a(t)$ (respectively, $a(t + 1) \leq a(t)$) for all $t \geq T$. Additionally, we say that $a(t)$ is "eventually strictly increasing" (respectively, "eventually strictly decreasing") if there exists a time T such that $a(t + 1) > a(t)$ (respectively, $a(t + 1) < a(t)$) for all times $t \geq T$.

$d = 1 + \max_{i \in V} \deg(i)$. Choose T such that the inequalities

$$|x_k(t) - x^k| < \frac{1}{4d} \min_{x^m \neq x^n} |x^m - x^n|, \tag{5.9}$$

$$|x_k(t) - x_k(t')| < \frac{1}{4d} \min_{x^m \neq x^n} |x^m - x^n|. \tag{5.10}$$

hold for each node k and for all $t' > t \geq T$.

We claim that nodes i and j are not receptive to each other at any time $t \geq T$. Suppose the contrary. There then must exist some time $t \geq T$ and adjacent nodes i and j with $x^i \neq x^j$ and $|x_i(t) - x_j(t)| < c_{ij}(t)$. Fix such a value of t and choose a node i that gives the smallest limit opinion value x^i such that there is a neighboring node j with $x^j \neq x^i$ and $|x_i(t) - x_j(t)| < c_{ij}(t)$.

For this node i, let $q = |I(i, x(t))| \leq d$. Because of our choice of x^i, we have

$$\frac{1}{q} \left| \sum_{\substack{j \in I(i,x(t)) \\ j \neq i}} (x^i - x^j) \right| = \frac{1}{q} \sum_{\substack{j \in I(i,x(t)) \\ j \neq i}} (x^j - x^i) \geq \frac{1}{d} \min_{x^m \neq x^n} |x^m - x^n|. \tag{5.11}$$

Using (5.9), we obtain

$$\begin{aligned}\frac{1}{q}\left|\sum_{\substack{j\in I(i,x(t))\\ j\neq i}}(x^i-x^j)\right| &\leq \frac{1}{q}\sum_{\substack{j\in I(i,x(t))\\ j\neq i}}\left|x_i(t)-x^i\right|+\frac{1}{q}\left|\sum_{\substack{j\in I(i,x(t))\\ j\neq i}}(x_i(t)-x_j(t))\right| \\ &\qquad +\frac{1}{q}\sum_{\substack{j\in I(i,x(t))\\ j\neq i}}\left|x_j(t)-x^j\right| \\ &< 2\left(\frac{q-1}{q}\right)\left(\frac{1}{4d}\right)\min_{x^m\neq x^n}|x^m-x^n|+\frac{1}{q}\left|\sum_{\substack{j\in I(i,x(t))\\ j\neq i}}(x_i(t)-x_j(t))\right| \\ &< \frac{1}{2d}\min_{x^m\neq x^n}|x^m-x^n|+\frac{1}{q}\left|\sum_{\substack{j\in I(i,x(t))\\ j\neq i}}(x_i(t)-x_j(t))\right|. \end{aligned} \tag{5.12}$$

Combining (5.11) and (5.12) yields

$$\frac{1}{q}\left|\sum_{\substack{j\in I(i,x(t))\\ j\neq i}}(x_i(t)-x_j(t))\right| > \frac{1}{2d}\min_{x^m\neq x^n}|x^m-x^n|. \tag{5.13}$$

Using the HK opinion-update rule (5.2) and the inequality (5.10), we also have

$$\frac{1}{q}\left|\sum_{\substack{j\in I(i,x(t))\\ j\neq i}}(x_i(t)-x_j(t))\right| = |x_i(t+1)-x_i(t)| < \frac{1}{4d}\min_{x^m\neq x^n}|x^m-x^n|. \tag{5.14}$$

The relations (5.13) and (5.14) cannot hold simultaneously, so nodes i and j are not receptive to each other at any time $t \geq T$. □

Lemma 5.4. *In our adaptive-confidence HK model (with update rules* (5.2) *and* (5.3)*) with parameters* $\gamma \in (0,1]$ *and* $\delta \in [0,1)$*, the dyadic confidence bound* $c_{ij}(t)$ *of each pair of adjacent nodes,* i *and* j*, is eventually either strictly increasing or strictly decreasing. That is, there*

is a time T such that exactly one of the inequalities $c_{ij}(t_1) < c_{ij}(t_2)$ and $c_{ij}(t_1) > c_{ij}(t_2)$ holds for all times $t_2 > t_1 \geq T$. Furthermore, if nodes i and j are in different limit opinion clusters (i.e., if $x^i \neq x^j$), then $c_{ij}(t)$ is eventually strictly decreasing.

Proof. We first consider c_{ij} for adjacent nodes i and j that are in different limit opinion clusters (i.e., $x^i \neq x^j$). By Lemma 5.3, there is a time T such that nodes i and j are mutually unreceptive at all times $t \geq T$. Consequently, $c_{ij}(t)$ cannot increase at any time $t \geq T$ and must be monotone decreasing.

Because the adaptive-confidence HK model updates synchronously and the initial confidence bound $c_0 \in (0, 1)$, each confidence bound c_{ij} must change at each time t. That is, for all pairs of adjacent nodes i and j and all times t, we have $c_{ij}(t+1) \neq c_{ij}(t)$. Consequently, for all adjacent nodes i and j in distinct limit opinion clusters and for all times $t \geq T$, we have that c_{ij} is strictly decreasing (i.e., $c_{ij}(t+1) < c_{ij}(t)$).

Now consider adjacent nodes i and j that are in the same limit opinion cluster (i.e., $x^i = x^j$). Choose $T > 0$ such that

$$|x_k(t) - x^k| < \frac{\gamma}{2} \tag{5.15}$$

for each node k and all $t \geq T$. We claim that there exists some $T_{ij} \geq T$ such that the dyadic confidence bound c_{ij} is either strictly decreasing (i.e., $c_{ij}(t+1) < c_{ij}(t)$) or strictly increasing (i.e., $c_{ij}(t+1) > c_{ij}(t)$) for all times $t \geq T_{ij}$.

If c_{ij} is strictly decreasing for all $t \geq T$, we choose $T_{ij} = T$. If c_{ij} is not strictly decreasing for all $t \geq T$, there must exist some time $T_{ij} \geq T$ at which $|x_i(T_{ij}) - x_j(T_{ij})| < c_{ij}(T_{ij})$; therefore, $c_{ij}(T_{ij}+1) > c_{ij}(T_{ij})$. Without loss of generality, let T_{ij} be the earliest such time. We will show by induction that $c_{ij}(t+1) > c_{ij}(t)$ for all $t \geq T_{ij}$. By assumption, this inequality holds for the base case $t = T_{ij}$.

Suppose that $c_{ij}(t+1) > c_{ij}(t)$ for some value of $t \geq T_{ij}$. We must then also have $|x_i(t) -$

$x_j(t)| < c_{ij}(t)$ and

$$c_{ij}(t+1) = c_{ij}(t) + \gamma(1 - c_{ij}(t)) \geq \gamma\,. \tag{5.16}$$

By the inequality (5.15), we have that

$$|x_k(t') - x_{k'}(t')| \leq |x_k(t') - x^k| + |x^k - x^{k'}| + |x^{k'} - x_{k'}(t')| < \gamma \tag{5.17}$$

for each node pair k and k' with $x^k = x^{k'}$ and all times $t' \geq T$. Because $t + 1 > T_{ij} \geq T$, it follows that

$$|x_i(t+1) - x_j(t+1)| < \gamma < c_{ij}(t+1)\,, \tag{5.18}$$

so $c_{ij}(t+2) > c_{ij}(t+1)$. Consequently, by induction, if c_{ij} increases at $t = T_{ij}$, then c_{ij} is strictly increasing (i.e., $c_{ij}(t+1) > c_{ij}(t)$) for all $t \geq T_{ij}$. Therefore, there exists some time T_{ij} such that c_{ij} is either strictly decreasing or strictly increasing for all $t \geq T_{ij}$.

In summary, we have shown that c_{ij} is eventually strictly decreasing for all adjacent nodes i and j in different limit opinion clusters. Additionally, for all adjacent nodes i and j in the same limit opinion cluster, we have shown that c_{ij} is eventually either strictly decreasing or strictly increasing. □

Lemma 5.5. *In our adaptive-confidence HK model (with update rules* (5.2) *and* (5.3)*), let* $\gamma \in (0, 1]$ *and* $\delta \in [0, 1)$*. Suppose that* $c^{ij} = \lim_{t\to\infty} c_{ij}(t)$ *exists. It then follows that either* $c^{ij} = 0$ *or* $c^{ij} = 1$*.*

Proof. Given $\epsilon > 0$, choose a time T so that the inequalities

$$|c_{ij}(t) - c^{ij}| < \epsilon/2\,, \tag{5.19}$$

$$|c_{ij}(t_1) - c_{ij}(t_2)| < \frac{1}{2}\left(\min\{1 - \delta, \gamma\}\right)\epsilon \tag{5.20}$$

hold for all times $t, t_1, t_2 \geq T$. Fix some time $t \geq T$. It must be the case that either

$$c_{ij}(t+1) = \delta c_{ij}(t) \tag{5.21}$$

or

$$c_{ij}(t+1) = c_{ij}(t) + \gamma(1 - c_{ij}(t))\,. \tag{5.22}$$

Suppose first that $c_{ij}(t+1) = \delta c_{ij}(t)$. In this case, we claim that $c^{ij} = 0$. To verify this claim, first note that $c_{ij}(t) - c_{ij}(t+1) = (1-\delta)c_{ij}(t)$. Because $c_{ij}(t) - c_{ij}(t+1) < \frac{1}{2}(1-\delta)\epsilon$, we see that $c_{ij}(t) < \epsilon/2$. Therefore,

$$\begin{aligned} 0 \leq c^{ij} &\leq |c^{ij} - c_{ij}(t)| + |c_{ij}(t)| \\ &< \epsilon/2 + \epsilon/2 \\ &= \epsilon\,, \end{aligned}$$

which implies that $c^{ij} = 0$.

Now suppose that $c_{ij}(t+1) = c_{ij}(t) + \gamma(1 - c_{ij}(t))$. Note that $c_{ij}(t+1) - c_{ij}(t) = \gamma(1 - c_{ij}(t)) < \frac{1}{2}\gamma\epsilon$, which implies that $1 - c_{ij}(t) < \epsilon/2$. Additionally,

$$\begin{aligned} 0 \leq 1 - c^{ij} &\leq |1 - c_{ij}(t)| + |c_{ij}(t) - c^{ij}| \\ &< \epsilon/2 + \epsilon/2 \\ &= \epsilon\,, \end{aligned}$$

which implies that $c^{ij} = 1$.

Therefore, it follows that either $c^{ij} = 0$ or $c^{ij} = 1$. □

5.3.1.2 Effective-graph analysis

In this section, we discuss the convergence of effective graphs in our adaptive-confidence HK model (see Theorem 5.6) and the baseline HK model (see Theorem 5.7). Our proofs of convergence employ some results from Section 5.3.1.1.

Theorem 5.6. *In our adaptive-confidence HK model with parameters $\gamma \in (0, 1]$ and $\delta \in [0, 1)$, the effective graph $G_{\text{eff}}(t)$ is eventually constant with respect to time. That is, there is some time T such that $G_{\text{eff}}(t) = G_{\text{eff}}(T)$ for all times $t \geq T$. Moreover, all of the edges of the* limit effective graph $\lim_{t\to\infty} G_{\text{eff}}(t)$ *are between nodes in the same limit opinion cluster.*

Proof. By Lemma 5.4, we can choose a time T such that each dyadic confidence bound c_{ij} is either strictly increasing or strictly decreasing for all times $t \geq T$.

For $t \geq T$, if $c_{ij}(t)$ is strictly decreasing, then we necessarily have that $|x_i(t) - x_j(t)| \geq c_{ij}(t)$ for all $t \geq T$, so $(i, j) \notin E_{\text{eff}}(t)$ for all $t \geq T$. If $c_{ij}(t)$ is strictly increasing, then $|x_i(t) - x_j(t)| < c_{ij}(t)$ for all $t \geq T$, so $(i, j) \in E_{\text{eff}}(t)$ for all $t \geq T$. Therefore, the set $E_{\text{eff}}(t)$ of edges of the effective graph is constant for all $t \geq T$, so the effective graph is constant for $t \geq T$.

For nodes i and j in different limit opinion clusters (i.e., $x^i \neq x^j$), Lemma 5.4 guarantees that the confidence bound c_{ij} is strictly decreasing for all times $t \geq T$. Therefore, $|x_i(t) - x_j(t)| \geq c_{ij}(t)$ for all $t \geq T$, so $(i, j) \notin E_{\text{eff}}(t)$ for all $t \geq T$. □

Theorem 5.6 states that all edges of a limit effective graph are between nodes in the same limit opinion cluster. However, the edges between nodes in the same limit opinion cluster do not have to exist in the limit effective graph. As we will discuss in Section 5.4.3 and Section 5.5, our numerical simulations suggest that our adaptive-confidence BCMs can have adjacent nodes in the same limit opinion cluster whose associated dyadic confidence bound converges to 0. The associated edge is thus not in the limit effective graph.

Theorem 5.7 guarantees that the effective graphs in the baseline HK model converge in the limit $t \to \infty$. Unlike in our adaptive-confidence HK model, all edges between nodes in the same limit opinion cluster in the baseline HK model must exist in the limit effective graph. Therefore, the limit opinion values in the baseline HK model fully determine the structure of the limit effective graph.

Theorem 5.7. *In the baseline HK model (with update rule* (5.1)*), the effective graph* $G_{\text{eff}}(t) = (V, E_{\text{eff}}(t))$ *is eventually constant with respect to time. Moreover, the edge* $(i, j) \in E$ *exists in the limit effective graph if and only if this edge is between two nodes in the same limit opinion cluster (i.e.,* $x^i = x^j$*).*

Proof. We first consider adjacent nodes, i and j, that are in different limit opinion clusters (i.e., $x^i \neq x^j$). By Lemma 5.3, because our adaptive-confidence HK model with $\gamma = 0$ and $\delta = 1$ reduces to the baseline HK model, there exists a time T_1 such that nodes i and j are not receptive to each other (i.e., $|x_i(t) - x_j(t)| \geq c$) at any time $t \geq T_1$. Therefore, the edge $(i, j) \notin E_{\text{eff}}(t)$ at any time $t \geq T_1$.

Now consider adjacent nodes, i and j, that are in the same limit opinion cluster (i.e., $x^i = x^j$). Choose a time T_2 such that $|x_k(t) - x^k| < c/2$ for each node k and all times $t \geq T_2$. For all $t \geq T_2$, we then have

$$|x_i(t) - x_j(t)| \leq |x_i(t) - x^i| + |x^i - x^j| + |x^j - x_j(t)| < c/2 + 0 + c/2 = c\,.$$

Therefore, at any time $t \geq T_2$, nodes i and j are receptive to each other and the edge $(i, j) \in E_{\text{eff}}(t)$. By taking $T = \max\{T_1, T_2\}$, for any time $t \geq T$, we have that $(i, j) \notin E_{\text{eff}}(t)$ for all edges (i, j) with $x^i \neq x^j$ and that $(i, j) \in E_{\text{eff}}(t)$ for all edges (i, j) with $x^i = x^j$. □

5.3.2 Adaptive-confidence DW model

In this section, we discuss our theoretical results for the confidence bounds and effective graphs in our adaptive-confidence DW model. Both the baseline DW model and our adaptive-confidence DW model are asynchronous and stochastic. At each discrete time, we uniformly randomly select one pair of adjacent nodes to interact. Because of the stochasticity in the baseline and adaptive-confidence DW models, our theoretical results for them are in an "almost sure" sense. By contrast, our theoretical results (see Section 5.3.1) are deterministic for the baseline and adaptive HK models.

5.3.2.1 Confidence-bound analysis

We now give our main result about the behavior of the confidence bounds in our adaptive-confidence DW model (which has the update rules (5.4) and (5.3)). This result mirrors the main result for our adaptive-confidence HK model in Section 5.3.1.1.

Theorem 5.8. *In our adaptive-confidence DW model (with update rules* (5.4) *and* (5.3)*)with parameters* $\gamma \in (0, 1]$ *and* $\delta \in [0, 1)$*, the dyadic confidence bound* $c_{ij}(t)$ *converges either to* 0 *or to* 1 *almost surely. Moreover, if nodes* i *and* j *are in different limit opinion clusters (i.e.,* $x^i \neq x^j$*), then* $c_{ij}(t)$ *converges to* 0 *almost surely.*

We prove Theorem 5.8 by proving Lemma 5.9 and Lemma 5.10, which we state shortly and prove in Section 5.A.1. Because our adaptive-confidence DW model updates asynchronously, Lemma 5.9 guarantees eventual monotone increase or decrease of the confidence bounds. This result differs from the eventual strict increase or decrease in the confidence bounds in our adaptive-confidence HK model (see Lemma 5.4). (See Footnote 14 for our usage of the terms "eventual monotone increase" (and decrease) and "eventual strict increase" (and decrease).) Because $c_{ij}(t) \in [0, 1]$, Lemma 5.9 implies convergence (because an eventually monotone sequence in $[0, 1]$ must converge). By Lemma 5.10, we have almost sure convergence to 0 or to 1. Moreover, if nodes i and j are in different limit opinion clusters, then Lemma 5.9

guarantees that $c_{ij}(t)$ is monotone decreasing. By Lemma 5.10, $c_{ij}(t)$ thus almost surely converges to 0.

Lemma 5.9. *In our adaptive-confidence DW model (with update rules* (5.4) *and* (5.3)*) with parameters* $\gamma \in (0, 1]$ *and* $\delta \in [0, 1)$*, the dyadic confidence bound* $c_{ij}(t)$ *of each pair of adjacent nodes,* i *and* j*, is eventually monotone increasing or monotone decreasing. That is, there is a time* T *such that exactly one of the inequalities* $c_{ij}(t_1) \leq c_{ij}(t_2)$ *and* $c_{ij}(t_1) \geq c_{ij}(t_2)$ *holds for all times* $t_2 > t_1 \geq T$*.*

Furthermore, if nodes i *and* j *are in different limit opinion clusters, then* $c_{ij}(t)$ *is eventually monotone decreasing.*

Lemma 5.10. *Consider our adaptive-confidence DW model (with update rules* (5.4) *and* (5.3)*) with parameters* $\gamma \in (0, 1]$ *and* $\delta \in [0, 1)$*. Suppose that* $c_{ij} = \lim_{t\to\infty} c_{ij}(t)$ *exists. It then follows that, almost surely, either* $c_{ij} = 0$ *or* $c_{ij} = 1$*.*

5.3.2.2 Effective-graph analysis

We now present Theorem 5.11, which is our main result about effective graphs in our adaptive-confidence DW model. In Section 5.A.2, we present its proof, which uses results from Section 5.3.2.1.

Theorem 5.11. *In our adaptive-confidence DW model (with update rules* (5.4) *and* (5.3)*) with parameters* $\gamma \in (0, 1]$ *and* $\delta \in [0, 1)$*, the effective graph* $G_{\text{eff}}(t)$ *almost surely eventually has edges only between nodes of the same limit opinion cluster. That is, there is almost surely some time* T *such that* $(i, j) \in E_{\text{eff}}(t)$ *implies that* $x^i = x^j$ *for all* $t \geq T$*.*

Unlike in our adaptive-confidence HK model, $\lim_{t\to\infty} G_{\text{eff}}(t)$ may not exist in our adaptive-confidence DW model. When the limit does exist, we refer to $\lim_{t\to\infty} G_{\text{eff}}(t)$ as the *limit effective graph.*

For completeness, we now state Theorem 5.12, which guarantees the almost-sure convergence

of the effective graphs as $t \to \infty$ in the baseline DW model. In Section 5.B, we prove Theorem 5.12 by first proving Lemma 5.13 and Lemma 5.14. Because the baseline DW model has a time-independent confidence bound c, our proof of Theorem 5.12 uses different ideas than our proof of Theorem 5.11.

Theorem 5.12. *Consider the baseline DW model (with update rule* (5.4)*). Almost surely, the effective graph $G_{\text{eff}}(t)$ is eventually constant with respect to time. That is, there is almost surely a time T such that $G_{\text{eff}}(t) = G_{\text{eff}}(T)$ for all times $t \geq T$.*
Furthermore, suppose that the limit effective graph $\lim_{t\to\infty} G_{\text{eff}}(t)$ exists. If adjacent nodes i and j have the same limit opinion (i.e., if $x^i = x^j$), then the edge (i, j) is in the limit effective graph. Additionally, if the edge (i, j) is in the limit effective graph, then $x^i = x^j$ almost surely.

5.4 Details of Our Numerical Simulations

We now discuss the details of our numerical simulations of our adaptive-confidence HK and DW models.

5.4.1 Network structures

We first simulate our adaptive-confidence HK and DW models on complete graphs to better understand their behaviors. We subsequently examine how different network structures affect those behaviors. We simulate our adaptive-confidence HK model on synthetic networks that we generate using random-graph models, and we simulate both adaptive-confidence BCMs on networks from empirical data. Because of computational limitations, we consider larger networks for the adaptive-confidence HK model than for the adaptive-model DW model.

We simulate our adaptive-confidence HK model on a complete graph, $G(N, p)$ Erdős–Rényi (ER) random graphs, and two-community stochastic-block-model (SBM) random graphs. In each case, we consider graphs with 1000 nodes. We also simulate our adaptive-confidence

HK model on social networks from the FACEBOOK100 data set [RKM11, TMP12].

A $G(N, p)$ ER graph has N nodes and independent probability p of an edge between each pair of distinct nodes [New18]. When $p = 1$, this yields a complete graph. We consider $G(N, p)$ graphs with $p \in \{0.1, 0.5\}$ to vary the sparsity of the graphs while still yielding connected graphs for our simulations. All of the ER graphs in our simulations are connected.

To determine how a network with an underlying community structure affects the dynamics of our adaptive-confidence HK model, we consider undirected SBM networks [New18] with a 2×2 block structure in which each block corresponds to an ER graph. To construct these SBMs, we partition a network into two sets of nodes; one set (which we denote by A) has 75% of the nodes, and the other set (which we denote by B) has the remaining 25% of the nodes. Our choice is inspired by the two-community SBM that was considered in [KP20]. We define a symmetric edge-probability matrix

$$P = \begin{bmatrix} P_{\mathrm{AA}} & P_{\mathrm{AB}} \\ P_{\mathrm{AB}} & P_{\mathrm{BB}} \end{bmatrix}, \tag{5.23}$$

where P_{AA} and P_{BB} are the probabilities of an edge between two nodes within the sets A and B, respectively, and P_{AB} is the probability of an edge between a node in set A and a node in set B. In our simulations, $P_{\mathrm{AA}} = P_{\mathrm{BB}} = 1$ and $P_{\mathrm{AB}} = 0.01$.

In addition to synthetic networks, we also simulate our adaptive-confidence HK model on several real-world networks. For each network, we use the largest connected component[15] (LCC). In Table 5.1, we give the numbers of nodes and edges in the LCCs of these networks, which are social networks from the FACEBOOK100 data set [RKM11, TMP12]. In each of these networks, the nodes are the Facebook pages of individuals at a university and the edges encode Facebook "friendships" between individuals in a one-day snapshot of the network from

[15] A connected component [New18] of an undirected network G is a maximal subgraph with a path between each pair of nodes.

fall 2005 [RKM11, TMP12]. The numbers of nodes in the LCCs of the examined Facebook networks range from 962 to 14,917.

For our adaptive-confidence DW model, we examine a complete graph and one real-world network. We simulate our adaptive-confidence DW model on a 100-node complete graph, which is one tenth of the size of the complete graph that we consider for our adaptive-confidence HK model. We use this smaller size because of computational limitations. Our simulations of our adaptive-confidence DW model on a 100-node complete graph frequently reach our "bailout time" (see Section 5.4.2 and Table 5.5) for small initial confidence bounds. We also simulate our adaptive-confidence DW model on the LCC of the real-world NetScience network of coauthorships between researchers in network science [New06].

Table 5.1: The real-world networks on which we simulate our adaptive-confidence BCMs. For each network, we use the largest connected component and indicate the numbers of nodes and edges in that component.

Network	Number of Nodes	Number of Edges	Model
NetScience	379	914	DW
Reed	962	18,812	HK
Swarthmore	1657	61,049	HK
Oberlin	2920	89,912	HK
Pepperdine	3440	152,003	HK
Rice	4083	184,826	HK
UC Santa Barbara	14,917	482,215	HK

5.4.2 Simulation specifications

In Table 5.2, we indicate the values of the model parameters that we examine in our simulations of our BCMs. The BCM parameters are the confidence-increase parameter γ, the confidence-decrease parameter δ, the initial confidence bound c_0, and (for the adaptive-confidence DW model only) the compromise parameter μ. For both our HK and DW models, the parameter pair $(\gamma, \delta) = (0, 1)$ corresponds to the associated baseline BCM.

Our BCM simulations include randomness from the initial opinions of the nodes and from the

specific networks in random-graph ensembles. The adaptive-confidence DW model also has randomness from the selection of nodes at each time step. We use Monte Carlo simulations to mitigate the effects of noise. For each parameter set of a random-graph model (i.e., the ER and SBM graphs), we generate 5 graphs. Additionally, for each graph, we generate 10 sets of initial opinions uniformly at random and reuse these sets of opinions for all BCM parameter values.

Table 5.2: The BCM parameter values that we examine in simulations of our adaptive-confidence BCMs. We consider more parameter values for complete graphs than for the other networks. We consider all of the indicated values for complete graphs, and we consider values without the asterisk (*) for the ER, SBM, and real-world networks.

Model	**BCM Parameters**
Adaptive-Confidence HK	$\gamma \in \{0, 0.0001^*, 0.0005^*, 0.001, 0.005, 0.01, 0.05, 0.1^*\}$
	$\delta \in \{0.01^*, 0.1^*, 0.5, 0.9, 0.95, 0.99, 1\}$
	$c_0 \in \{0.02, 0.03, \ldots, 0.19, 0.20, 0.30, 0.40, 0.50\}$
Adaptive-Confidence DW	$\gamma \in \{0.1, 0.3, 0.5^*\}$
	$\delta \in \{0.3^*, 0.5, 0.7^*\}$
	$c_0 \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$
	$\mu \in \{0.1, 0.3, 0.5\}$

* We consider these parameter values only for complete graphs.

For our numerical simulations, we need a stopping criterion, as it can potentially take arbitrarily long for nodes to reach their limit opinions in a BCM simulation. We consider the effective graph $G_{\text{eff}}(t)$, which we recall (see Section 5.3) is the subgraph of the original graph G with edges only between nodes that are receptive to each other at time t. In our simulations, each of the connected components of $G_{\text{eff}}(t)$ is an "opinion cluster" $K_r(t)$ at time t. Our stopping criterion checks that the maximum difference in opinions between nodes in the same opinion cluster is less than some tolerance. That is,

$$\max\{|x_i(t) - x_j(t)| \text{ such that } i, j \in K_r(t) \text{ for some } r\} < \text{tolerance}\,. \tag{5.24}$$

We use a tolerance value of 1×10^{-6} for our adaptive-confidence HK model. Because of

computational limitations, we use a tolerance value of 0.02 for our adaptive-confidence DW model. We refer to the time T_f at which we reach our stopping criterion as the "convergence time" of our simulations. In our simulations, the "final effective graph" is the effective graph at the convergence time (i.e., the time T_f that a simulation reaches our stopping criterion). We refer to the connected components of the final effective graph as the "final opinion clusters" of a simulation; they approximate the limit opinion clusters.

Our theoretical results about effective graphs inform our stopping criterion. Theorem 5.6 and Theorem 5.7 give theoretical guarantees for our adaptive-confidence HK model and the baseline HK model, respectively, that eventually the only edges of an effective graph are those between adjacent nodes in the same limit opinion cluster. Theorem 5.11 and Theorem 5.12 give similar but weaker guarantees for our adaptive-confidence DW model and the baseline DW model. Consequently, if one of our simulations runs for sufficiently many time steps, its final opinion clusters are a good approximation of the limit opinion clusters. However, instead of imposing a set number of time steps for our simulations, we use a tolerance value as a proxy to determine a "sufficient" number of time steps.

The final and limit opinion clusters in our models may not be the same, as our choice of tolerance values can lead to simulations stopping before we can determine their limit opinion clusters. Additionally, if two distinct connected components in an effective graph converge to the same limit opinion value, the nodes in those connected components are in the same limit opinion cluster. However, the sets of nodes that constitute these connected components are distinct final opinion clusters. In practice, our simulations are unlikely to have distinct opinion clusters that converge to the same opinion in the infinite-time limit. Therefore, for small tolerance values, our final opinion clusters are a good approximation of the limit opinion clusters.

To ensure that our simulations stop after a reasonable amount of time, we use a bailout time of 10^6 time steps. Our simulations of the adaptive-confidence HK model in the present

chapter never reach this bailout time. However, our simulations of the adaptive-confidence DW model frequently reach the bailout time for small values of c_0. See Section 5.D.1 and Table 5.5.

5.4.3 Quantifying model behaviors

In our numerical simulations, we investigate the convergence time and characterize the final opinions. To examine the convergence time, we record the number T_f of time steps that it takes for simulations to reach our stopping criterion. To characterize opinions, we determine whether there is consensus or opinion fragmentation (which we define shortly), quantify the opinion fragmentation using Shannon entropy, and examine the numbers of nodes and edges in each opinion cluster.

In models of opinion dynamics, it is common to investigate whether or not nodes eventually reach a consensus (i.e., arrive at the same opinion) [NVT20]. In practice, to determine whether a simulation reaches a consensus state, we use notions of "major" and "minor" clusters. Consider a 1000-node network in which 998 nodes have one steady-state opinion but the remaining 2 nodes have a different steady-state opinion. In applications, it does not seem appropriate to characterize this situation as opinion polarization or fragmentation. Therefore, we use notions of major and minor opinion clusters [Lor08, LAZ04], which we characterize in an ad hoc way. We define a "major" opinion cluster as a final opinion cluster with strictly more than 1% of the nodes of a network. A final opinion cluster that is not a major cluster is a "minor" cluster. We say that a simulation that results in one major cluster yields a "consensus" state and that a simulation that results in at least two major clusters yields "opinion fragmentation" (i.e., a "fragmented" state).[16] We track the numbers and sizes of all major and minor clusters, and we use all clusters (i.e., both major and minor clusters)

[16]In studies of opinion dynamics, it is common to use the term "fragmentation" to refer to situations with three or more opinion clusters and to use the term "polarization" to refer to situations with exactly two opinion clusters. In the present chapter, it is convenient to quantify any state other than consensus as a fragmented state.

to quantify opinion fragmentation.

There are many ways to quantify opinion fragmentation [BGS16,MRU22,AWA22]. We distinguish situations in which the final opinion clusters (major or minor) are of similar sizes from situations in which these clusters have a broad range of sizes. Following Han et al. [HFQ20], we calculate Shannon entropy to quantify opinion fragmentation.[17] At some time t, suppose that there are R opinion clusters, which we denote by $K_r(t)$ for $r \in \{1, 2, \ldots, R\}$. We refer to the set $\{K_r(t)\}_{r=1}^{R}$, which is a partition of the set of nodes of a network, as an "opinion-cluster profile". The fraction of nodes in opinion cluster $K_r(t)$ is $|K_r(t)|/N$, where N is the number of nodes of a network. The Shannon entropy $H(t)$ of an opinion-cluster profile is

$$H(t) = -\sum_{r=1}^{R} \frac{|K_r(t)|}{N} \ln\left(\frac{|K_r(t)|}{N}\right) . \tag{5.25}$$

We calculate $H(T_f)$, which is the Shannon entropy of the opinion-cluster profile of final opinion clusters at convergence time. Computing Shannon entropy allows us to use a scalar value to quantify the distribution of opinion-cluster sizes, with larger entropies indicating greater opinion fragmentation. The Shannon entropy is larger when there are more opinion clusters. Additionally, for a fixed number R of opinion clusters, the Shannon entropy is larger when the opinion clusters are evenly sized than when the sizes are heterogeneous. When comparing two opinion-cluster profiles, we consider both the numbers of major clusters and the Shannon entropies. When there are sufficiently few minor clusters, we expect the number of major clusters to follow the same trend as the Shannon entropy.

To examine the structure of final opinion clusters, we study the properties of final effective graphs. Our theoretical results allow the possibility that some adjacent nodes converge to the same opinion without being mutually receptive. We observe this phenomenon in our numerical simulations. (See Section 5.5 for more discussion.) To quantify this behavior, we

[17] See [LP23] for another study that followed [HFQ20] by calculating Shannon entropy to help quantify opinion fragmentation in a BCM.

calculate a weighted average of the fractions of edges (which we call the "weighted-average edge fraction") that are in each opinion cluster of the final effective graph. In an opinion-cluster profile $\{K_i(t)\}_{r=1}^{R}$, let $E(r)$ denote the set of edges of the original graph G between nodes in opinion cluster r and let $E_{\text{eff}}(t,r)$ denote the set of edges of the effective graph $G_{\text{eff}}(t)$ that are in opinion cluster r. That is,

$$E(r) = \{(i,j) \in E \text{ such that } i,j \in K_r(t)\}\,,$$
$$E_{\text{eff}}(t,r) = \{(i,j) \in E_{\text{eff}}(t) \text{ such that } i,j \in K_r(t)\}\,.$$

The weighted average of the fractions of edges (i.e., the weighted-average edge fraction) that are in the effective graph for each opinion cluster is

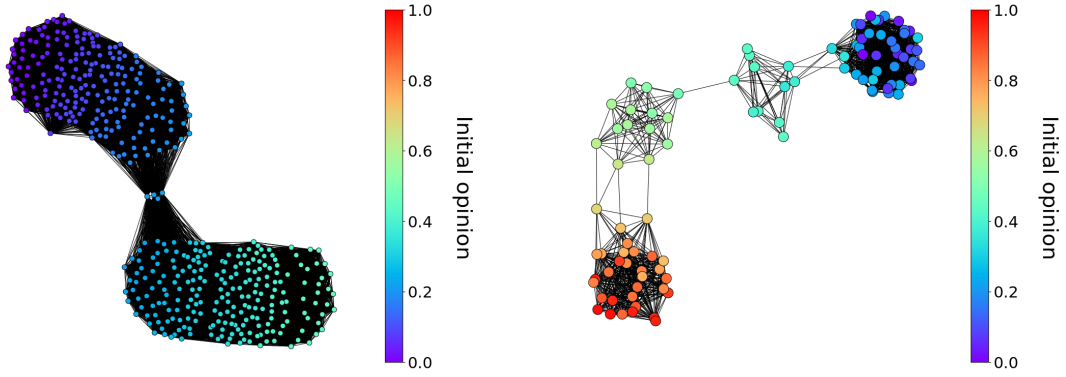
$$W(t) = \sum_{\substack{r=1 \\ E(r)\neq 0}}^{R} \left(\frac{|K_r(t)|}{N-\ell}\right)\left(\frac{|E_{\text{eff}}(t,r)|}{|E(r)|}\right)\,, \tag{5.26}$$

where ℓ is the number of isolated nodes of the effective graph.[18] An isolated node is an opinion cluster with $E(r) = 0$. We are interested in the value of $W(t)$ at the convergence time T_f. Therefore, we calculate the weighted-average edge fraction $W(T_f)$. If each opinion cluster of an effective graph has all of its associated original edges of G, then $W = 1$. The value of W is progressively smaller when there are progressively fewer edges between nodes in the same opinion cluster of the effective graph. In Figure 5.1, we show examples of effective graphs with $W(T_f) < 1$.

5.5 Results of Our Numerical Simulations

We now present the results of numerical simulations of our adaptive-confidence BCMs. We consider various values of the BCM parameters, which are the initial confidence bound c_0, the

[18]If every component of an effective graph is an isolated node (i.e, $N = \ell$), then one can take either $W(t) = 0$ or $W(t) = 1$. In our simulations, this situation never occurred.



(a) One final opinion cluster in a simulation of our adaptive-confidence HK model that does not reach consensus on a 1000-node complete graph with $\gamma = 0.001$, $\delta = 0.5$, and $c_0 = 0.1$.

(b) The final effective graph in a simulation of our adaptive-confidence DW model that reaches consensus on a 100-node complete graph with $\gamma = 0.1$, $\delta = 0.5$, $c_0 = 0.1$, and $\mu = 0.1$.

Figure 5.1: Examples of final effective graphs with $W(T_f) < 1$. We color the nodes by their initial opinion values.

confidence-increase parameter γ, the confidence-decrease parameter δ, and (for the adaptive-confidence DW model only) the compromise parameter μ. We use the BCM-parameter values in Table 5.2, including the values that correspond to the baseline models (i.e., $(\gamma, \delta) = (0, 1)$). Our code and plots are available in our code repository.

As we described in Section 5.4.3, for both of our adaptive-confidence BCMs, we examine the number of major clusters (which we use to determine whether a simulation reaches a consensus state or a fragmented state), the number of minor clusters, the Shannon entropy $H(T_f)$ (see Equation (5.25)), the weighted-average edge fraction $W(T_f)$ (see Equation (5.26)), and the convergence time T_f. When the Shannon entropy and the number of major clusters follow similar trends, we only show results for the number of major clusters, as it is easier to interpret than the entropy. To avoid drowning readers with too much repetition, we include some of our plots of our adaptive-confidence HK model and adaptive-confidence DW model in Section 5.C and Section 5.D, respectively. Furthermore, we do not show plots for all of our numerical results; the omitted plots are available in our code repository.

Our simulation results and theoretical results about effective graphs complement each other. In Theorem 5.2, we proved for our adaptive-confidence HK model that all dyadic confidence bounds converge either to 0 or to 1. We also proved that the dyadic confidence bounds for node pairs in different limit opinion clusters must converge to 0. However, we have not proven whether or not the dyadic confidence bounds for nodes in the same limit opinion cluster converge to 1, so it is possible for such confidence bounds to converge to 0. (We first mentioned this point in Section 5.3.1.1.) If a dyadic confidence bound convergences to 0, then the corresponding edge is absent in the limit effective graph (which is guaranteed to exist by Theorem 5.6). Our numerical simulations suggest that a final opinion cluster can include adjacent nodes whose dyadic confidence bound converges to 0. In particular, in many simulations, we observe that the weighted-average edge fraction $W(T_f) < 1$, which corresponds to absent edges of the final effective graph between nodes that are in the same final opinion cluster. For our adaptive-confidence DW model, we prove analogous theoretical results (see Theorem 5.8 and Theorem 5.11) and we again observe simulations with $W(T_f) < 1$.

5.5.1 Adaptive-confidence HK model

5.5.1.1 Summary of our simulation results

For our adaptive-confidence HK model, all of our numerical simulations reach a consensus state for $c_0 \geq 0.3$. We show our simulation results for $c_0 \in \{0.02, 0.03, \ldots, 0.20\}$; we include the results for the other examined values of c_0 (see Table 5.2) in our code repository. We examine the numbers of major and minor clusters, the Shannon entropy $H(T_f)$ (see Equation (5.25)), the weighted-average edge fraction $W(T_f)$ (see Equation (5.26)), and the convergence time T_f. We plot each of these quantities versus the initial confidence bound c_0. For each value of the confidence-increase parameter γ, we generate one plot; each plot has one curve for each value of the confidence-decrease parameter δ. Each point in our plots is a mean of our numerical simulations for the associated values of the BCM parameter set (γ, δ, and

Table 5.3: Summary of the observed trends in our adaptive-confidence HK model. Unless we note otherwise, we observe these trends for the complete graph, all examined random-graph models, and all examined real-world networks.

Quantity	Trends
Convergence Time	• For fixed values of the initial confidence bound c_0, our adaptive-confidence HK model tends to converge more slowly than the baseline HK model.
	• When our simulations reach a consensus state, for fixed values of c_0 and the confidence-increase parameter γ, our model converges faster when the confidence-decrease parameter $\delta = 1$ than when $\delta \leq 0.9$. For $\delta \in \{0.95, 0.99\}$, the convergence time transitions from the $\delta \leq 0.9$ behavior to the $\delta = 1$ behavior as we increase c_0.
Opinion Fragmentation	• Our adaptive-confidence HK model yields consensus for $\gamma \geq 0.05$.
	• For fixed values of c_0, our adaptive-confidence HK model tends to yield fewer major clusters than the baseline HK model. When we fix the other BCM parameters, the number of major clusters decreases as either (1) we decrease δ or (2) we increase γ.
	• For our synthetic networks, we observe that the trends in Shannon entropy match the trends in the numbers of major clusters and that our adaptive-confidence HK model tends to yield less opinion fragmentation than the baseline HK model.*
	• For the baseline HK model, as we increase c_0, the number of major clusters tends to decrease. In our adaptive-confidence HK model, for simulations without consensus and for sufficiently large γ, the number of major clusters first increases and then decreases as we increase c_0.
$W(T_f)$	• When $\delta = 1$, both our adaptive-confidence HK model and the baseline HK model yield $W(T_f) = 1$.
	• For fixed γ and c_0, as we increase δ, the weighted-average edge fraction $W(T_f)$ tends to decrease.
	• For simulations that reach a consensus state, we observe two qualitative behaviors for $W(T_f)$: for $\delta \leq 0.9$, we observe that $W(T_f) < 1$ and that there is a seemingly linear relationship between $W(T_f)$ and c_0; for $\delta = 1$, we observe that $W(T_f) = 1$. Additionally, for $\delta \in \{0.95, 0.99\}$, the behavior of $W(T_f)$ transitions from the $\delta \leq 0.9$ behavior to the $\delta = 1$ behavior as we increase c_0.

* For the FACEBOOK100 networks, we do not observe this trend, seemingly because of the large numbers of minor clusters (which are incorporated in our calculation of Shannon entropy in (5.25)) for these networks.

c_0). We also show one standard deviation from the mean. For our simulations on a complete graph and the FACEBOOK100 networks, each point in our plots is a mean of 10 simulations (from 10 sets of initial opinions). For our simulations on $G(N, p)$ ER random graphs and SBM random graphs, each point in our plots is a mean of 50 simulations (from 5 random graphs that each have 10 sets of initial opinions). We include all plots — including those that we do not present in the present section, in Section 5.5.1, or in Section 5.C — in our code repository. In Table 5.3, we summarize the trends that we observe in our simulations.

In all of our simulations of our adaptive-confidence HK model, we observe that $\gamma \geq 0.001$ results in fewer major clusters than in the baseline HK model. For a fixed initial confidence bound c_0, our adaptive-confidence HK model tends to yield fewer major opinion clusters and less opinion fragmentation as either (1) we increase γ for fixed δ or (2) we decrease δ for fixed γ. Intuitively, one expects larger values of γ to encourage consensus because a larger γ entails a larger increase in a dyadic confidence bound after a positive interaction. Less intuitively, smaller values of δ, which entail a larger decrease in a dyadic confidence bound after a negative interaction, also seem to encourage consensus. In our adaptive-confidence HK model, we update opinions synchronously, with each node interacting with of all its neighboring nodes at each time. When two adjacent nodes are mutually unreceptive, their dyadic confidence bound decreases. Given the synchronous updates in our adaptive-confidence HK model, we hypothesize that small values of δ yield a faster decrease than large values of δ in the dyadic confidence bound between mutually unreceptive nodes. For small values of δ, pairs of nodes may quickly become mutually unreceptive and remain mutually unreceptive. Meanwhile, individual nodes can be receptive to (and thus average) fewer "conflicting" opinions[19], possibly aiding in reaching a consensus.

In the baseline HK model, the number of major opinion clusters tends to decrease as we

[19]When the neighbors to which a node is receptive have large differences in opinions with each other, we say that that node is receptive to "conflicting" opinions.

increase c_0. For intermediate values of γ (e.g., $\gamma \in \{0.005, 0.01\}$ for a complete graph; see panels (E) and (F) in Section 5.5.1.2), we do not observe this trend in our adaptive-confidence HK model. Instead, as we increase c_0, we observe an increase and then a decrease in the number of major clusters; unlike for the baseline HK model, small values of c_0 tend to promote more consensus (i.e., it tends to yield fewer major clusters). For smaller values of c_0, it seems that nodes tend to be receptive to fewer nodes early in a simulation, so fewer opinions can influence them. Therefore, for these values of c_0, nodes that are mutually receptive may quickly approach a consensus when γ is sufficiently large. Our calculations of the weighted-average edge fraction $W(T_f)$ support this hypothesis. For sufficiently large γ, small values of c_0 yield small values of $W(T_f)$, indicating that final opinion clusters tend to have many pairs of mutually unreceptive nodes.

5.5.1.2 A complete graph

We now discuss the simulations of our adaptive-confidence HK model on a complete graph. We summarize the observed trends in Table 5.3.

In Section 5.5.1.2, we observe for a 1000-node complete graph that our adaptive-confidence HK model yields fewer major clusters (i.e., it encourages more consensus) than the baseline HK model for a wide range of BCM parameter values. Our adaptive-confidence HK model always reaches a consensus state for $\gamma \geq 0.05$. In our simulations that do not reach consensus, we tend to observe progressively more major clusters and more opinion fragmentation as either (1) we decrease γ for fixed δ and c_0 or (2) we increase δ for fixed γ and c_0. For the baseline HK model and our adaptive-confidence HK model with small values of γ (specifically, $\gamma \in \{0.0001, 0.0005, 0.001\}$), the number of major opinion clusters tends to decrease as we increase c_0. We do not observe this trend for larger values of γ (specifically, $\gamma \in \{0.005, 0.01\}$). Instead, for these values of γ, small values of c_0 tend to promote more consensus. For example, simulations always reach a consensus state when $\gamma = 0.01$ and $c_0 \leq 0.08$. At the end of Section 5.5.1.1, we discussed our hypothesis behind this observation.

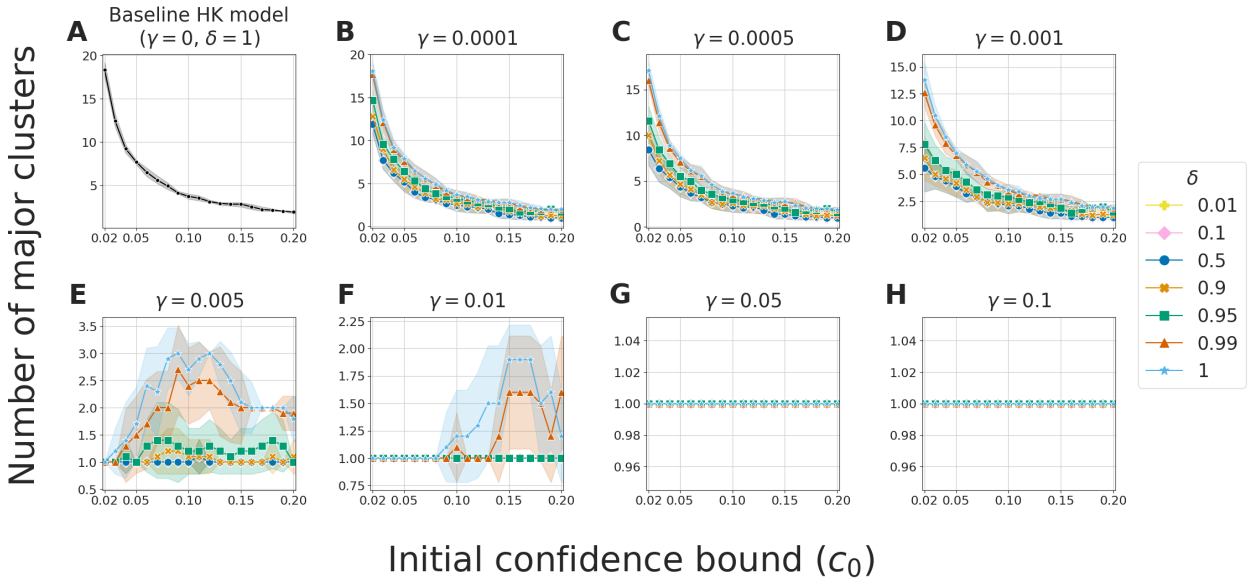


Figure 5.2: The numbers of major clusters in simulations of our adaptive-confidence HK model on a 1000-node complete graph for various combinations of the BCM parameters γ, δ, and c_0. In this and subsequent figures, we plot the mean value of our simulations for each set of BCM parameters. The bands around each curve indicate one standard deviation around the mean values. For clarity, in this figure and in subsequent figures, the vertical axes of different panels have different scales.

We observe very few minor clusters in our simulations of our adaptive-confidence HK model on a 1000-node complete graph. For each value of the BCM parameter set (γ, δ, c_0), the mean number of minor clusters in our 10 simulations is bounded above by 1. Because there are few minor clusters, the Shannon entropy (which accounts for both major clusters and minor clusters; see Equation (5.25)) and the number of major clusters follow similar trends for a 1000-node complete graph.

In Section 5.5.1.2, we show $W(T_f)$ (see Equation (5.26)), which is the weighted average of the fractions of edges in the opinion clusters of the final effective graph. When $\delta = 1$, both our adaptive-confidence HK model and the baseline HK model yield $W(T_f) = 1$. This indicates that all final opinion clusters (i.e., the connected components of the effective graph at time T_f) are complete subgraphs (i.e., cliques). By contrast, in our adaptive-confidence

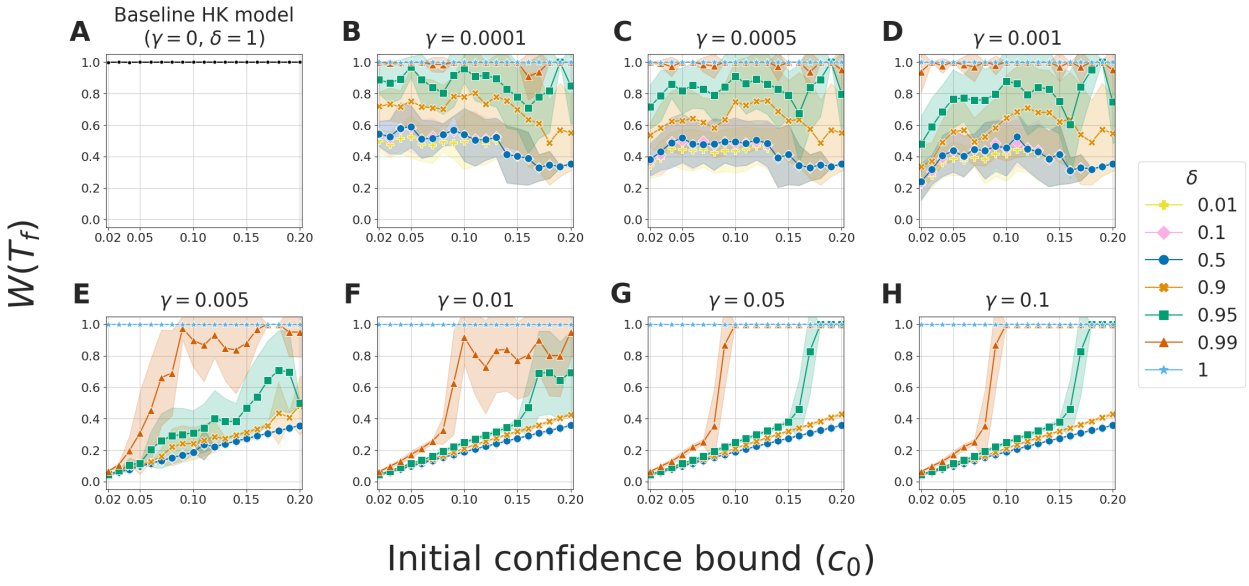


Figure 5.3: The weighted-average edge fraction $W(T_f)$ (see Equation (5.26)) in simulations of our adaptive-confidence HK model on a 1000-node complete graph for various combinations of the BCM parameters γ, δ, and c_0.

HK model, when $\delta < 1$ and for a wide range of the other BCM parameters, we observe that $W(T_f) < 1$. This indicates that some nodes that are adjacent in the graph G and in the same final opinion cluster do not have an edge between them in the final effective graph $G_{\text{eff}}(T_f)$. The nodes in these dyads are thus not receptive to each other (and hence do not influence each other's opinions), but they nevertheless converge to the same opinion. When δ is sufficiently small (specifically, $\delta \leq 0.9$), we observe that our adaptive-confidence HK model can reach a consensus with $W(T_f) < 1$. For sufficiently large values of γ (specifically, $\gamma \geq 0.05$), even though the nodes in some dyads are not receptive to each other, most nodes (at least 99% of them, based on our definition of major cluster) still converge to the same final opinion and hence reach a consensus.

For fixed values of γ and c_0, we observe that $W(T_f)$ tends to decrease as we decrease δ. For $\gamma \in \{0.05, 0.1\}$, our simulations always reach a consensus state. In these simulations, for each fixed δ, we observe that $W(T_f)$ appears to increase monotonically with respect to c_0.

Additionally, for these values of γ, we observe a transition in $W(T_f)$ as a function of δ. For $\delta \leq 0.9$, we observe that $W(T_f) < 1$ and that there is a seemingly linear relationship between $W(T_f)$ and c_0. When $\delta = 1$, we observe that $W(T_f) = 1$. For $\delta \in \{0.95, 0.99\}$, the behavior of $W(T_f)$ transitions from the $\delta \leq 0.9$ behavior to the $\delta = 1$ behavior as we increase c_0. This transition between behaviors occurs for smaller c_0 for $\delta = 0.99$ than for $\delta = 0.95$.

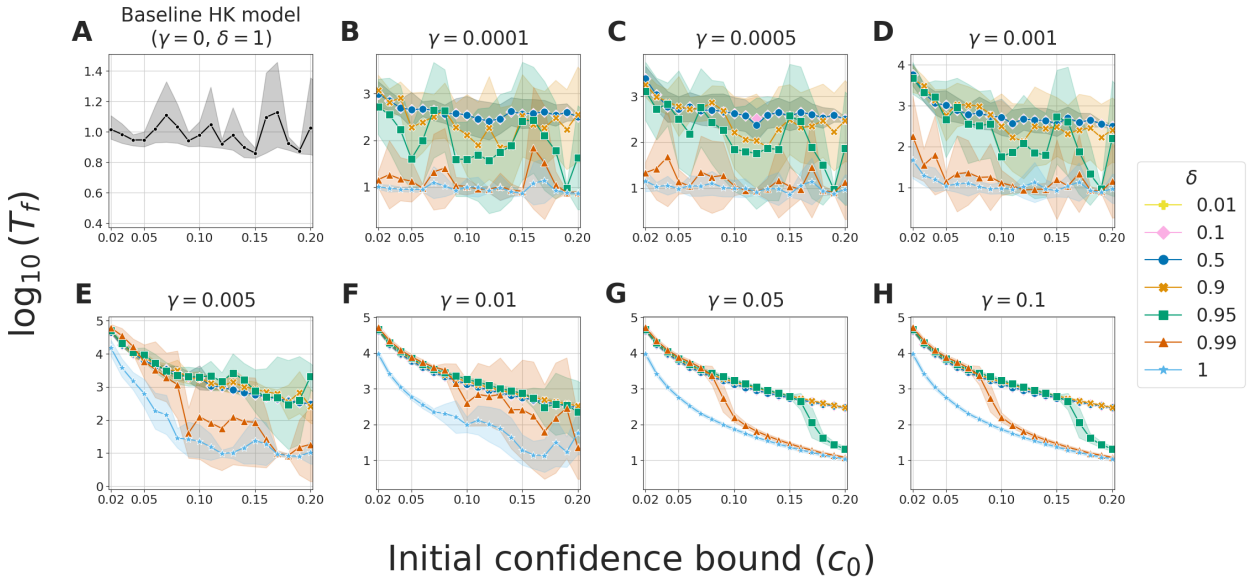


Figure 5.4: The convergence times (in terms of the number of time steps) on a logarithmic scale in simulations of our adaptive-confidence HK model on a 1000-node complete graph for various combinations of the BCM parameters γ, δ, and c_0.

In Section 5.5.1.2, for fixed c_0, we observe that our adaptive-confidence HK model takes longer to converge than the baseline HK model. For a 1000-node complete graph and fixed BCM parameters (i.e., γ, δ, and c_0), we observe that the logarithm $\log_{10}(T_f)$ of the convergence time T_f for our adaptive-confidence HK model can be up to 4 more than the logarithm of the convergence time for the baseline HK model. That is, the convergence time can be as much as 10^4 times larger. The convergence time tends to increase as either (1) we increase γ for fixed δ and c_0 or (2) we decrease δ for fixed γ and c_0. For large values of γ (as is especially evident for $\gamma \in \{0.05, 0.1\}$), the convergence time decreases with c_0. As with $W(T_f)$, for these values of γ, we observe a transition in the convergence time as a function of δ. In

Section 5.5.1.2, the curves of $\log_{10}(T_f)$ versus c_0 for $\delta \leq 0.9$ overlay each other and indicate larger convergence times than the curve for $\delta = 1$. The curves for $\delta = 0.95$ and $\delta = 0.99$ transition from the $\delta \leq 0.9$ behavior to the $\delta = 1$ behavior as we increase c_0. By contrast, for the baseline HK model and for our model with small values of γ, we observe no clear pattern between the convergence time and initial confidence bound. When our adaptive-confidence HK model reaches a consensus state, we observe from the behavior of $W(T_f)$ (see Section 5.5.1.2) and the convergence times (see Section 5.5.1.2) in our simulations that there is qualitative transition in the model behavior as we vary δ. We are not aware of previous discussions of similar transitions in variants of the HK model.

5.5.1.3 Erdős–Rényi (ER) and two-community stochastic-block-model (SBM) graphs

We now discuss our simulations of our adaptive-confidence HK model on $G(N, p)$ ER random graphs and two-community SBM random graphs. As in our simulations on the complete graph, we observe the trends in Table 5.3. We briefly discuss how our observations for these random graphs differ from our observations for the complete graph. We give additional details about our ER simulations in Section 5.C.1, and we give additional details about our SBM simulations in Section 5.C.2.

For fixed BCM parameters (namely, γ, δ, and c_0), we tend to observe fewer major opinion clusters for $G(1000, 0.1)$ graphs than for the 1000-node complete graph. Additionally, for small initial confidence bounds (specifically, $c_0 \leq 0.04$), we observe more minor clusters for the $G(1000, 0.1)$ graphs than the $G(1000, 0.5)$ graphs and the 1000-node complete graph. (For $G(1000, 0.1)$ graphs, once we take the mean for each BCM parameter set, we sometimes observe as many as 20 minor clusters.) The expected mean degree of a $G(N, p)$ ER graph is $p(N - 1)$ [New18]. Therefore, for small probability p, we expect more nodes to have small degrees. For small initial confidence bounds, we hypothesize that many nodes with small degrees quickly disconnect to form minor opinion clusters in the effective graph.

For our two-community SBM graphs, for fixed BCM parameters, the numbers of major clusters and Shannon entropies are similar to those for the complete graph. Each of our SBM graphs consists of two complete graphs that are joined by a small number of edges (see Section 5.4.1) to yield a two-community structure. It seems that this two-community structure does not significantly impact the simulation results of our adaptive-confidence HK model.

5.5.1.4 FACEBOOK100 university networks

We now discuss our simulations of our adaptive-confidence HK model on FACEBOOK100 networks (see Section 5.4.1) [RKM11, TMP12]. We show plots of the number of major clusters and Shannon entropy for the UC Santa Barbara network. In Section 5.C.3, we show a plot of the number of major clusters for Reed College. In our code repository, we include all plots for our simulations on FACEBOOK100 networks (including plots for the other examined quantities and the other four universities in Table 5.1).

The six FACEBOOK100 networks (see Table 5.1) mostly exhibit the same trends.[20] Except for the trends in Shannon entropy, we observe the same trends (see Table 5.3) for the FACEBOOK100 networks that we observed for the synthetic networks. For the FACEBOOK100 networks, most of the final opinion clusters for both our adaptive-confidence HK model and the baseline HK model are minor opinion clusters. Our simulations on the UC Santa Barbara network yield more minor clusters than our simulations on the other FACEBOOK100 networks; when $c_0 = 0.02$ and $\delta \leq 0.9$, the UC Santa Barbara network has more than 4000 minor clusters. Our calculation of Shannon entropy (see (5.25)) includes contributions from minor opinion clusters. Therefore, because of the large numbers of minor clusters for the FACEBOOK100 networks, the Shannon entropy and numbers of major opinion clusters fol-

[20]The Reed College network is a notable exception. For small initial confidence bounds $c_0 \leq 0.04$ and fixed values of the BCM parameters, it tends to have more major clusters and larger Shannon entropies than the other five FACEBOOK100 networks. We hypothesize that this observation, which we discuss further in Section 5.C.3, arises from finite-size effects.

low different trends. For these networks, they thus give complementary views of opinion fragmentation.

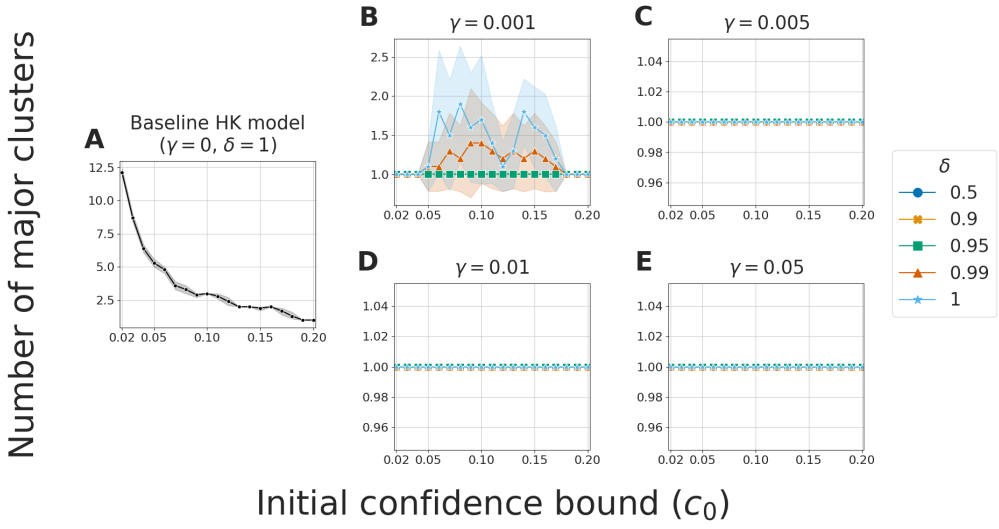


Figure 5.5: The numbers of major clusters in simulations of our adaptive-confidence HK model on the UC Santa Barbara network for various combinations of the BCM parameters γ, δ, and c_0.

In Figure 5.5, we observe for the UC Santa Barbara network that our adaptive-confidence HK model always yields consensus (the number of major clusters is exactly 1) when $\gamma \geq 0.005$. For these values of γ, the Shannon entropy (see Figure 5.6) tends to decrease as we increase c_0 for fixed values of γ and δ. This trend occurs because the number of minor opinion clusters also tends to decrease as we increase c_0 for fixed values of γ and δ. This observation contrasts with our simulations of our adaptive-confidence HK model on synthetic networks (see Sections 5.5.1.2 and 5.5.1.3), for which we observed that the Shannon entropy follows similar trends as the number of major clusters as we vary one of γ, δ, or c_0 while fixing the other BCM parameters. We believe that one reason for this difference is that the FACEBOOK100 networks have many small-degree nodes, which allow more minor opinion clusters to form.

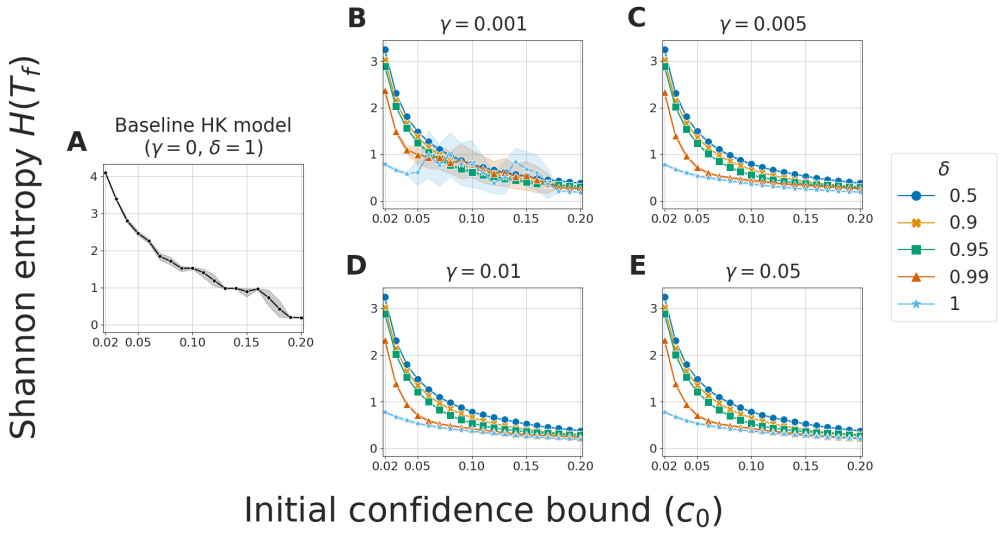


Figure 5.6: The Shannon entropies in simulations of our adaptive-confidence HK model on the UC Santa Barbara network for various combinations of the BCM parameters γ, δ, and c_0.

5.5.2 Adaptive-confidence DW model

We now briefly discuss our simulations of our adaptive-confidence DW model on a 100-node complete graph and the NetScience network [New06]. In Table 5.4, we summarize the trends that we observe in these simulations. Because of the long computation times, we consider much smaller graphs and fewer BCM parameter values for our adaptive-confidence DW model than we did for our adaptive-confidence HK model. Notably, the value of the compromise parameter μ (which is in the DW models but is not the HK models) affects our simulation results.

Our adaptive-confidence DW model yields some of the same trends that we obtained in our adaptive-confidence HK model. One of these trends is that, for both the 100-node complete graph and the NetScience network, the weighted-average edge fraction $W(T_f) < 1$ for some BCM parameters. Another trend is that our adaptive-confidence DW model has longer convergence times than the baseline DW model in our simulations on the 100-node complete graph. Additionally, for the 100-node complete graph, we observe less opinion fragmentation

as we increase γ. We do not observe these latter two trends in our simulations on the NetScience network. In Section 5.D, we give a detailed discussion of the results of our simulations of our adaptive-confidence DW model.

Table 5.4: Summary of the observed trends in our adaptive-confidence DW model.

Quantity	Trends
Convergence Time	• For the complete graph, for fixed values of the compromise parameter μ and initial confidence bound $c_0 \leq 0.3$, our adaptive-confidence DW model tends to converge more slowly than the baseline DW model. • For the NetScience network, for fixed values of μ and c_0, our adaptive-confidence DW model and the baseline DW model have similar convergence times.
Number of Major Clusters	• For the complete graph, when we fix the other BCM parameters, we (1) tend to observe fewer major clusters as we increase the confidence-increase parameter γ and (2) observe little effect on the numbers of major clusters when we vary the confidence-decrease parameter δ. • For the complete graph, for a fixed value of $c_0 \leq 0.3$, our adaptive-confidence DW model yields fewer major clusters when $\mu = 0.1$ than when $\mu \in \{0.3, 0.5\}$. The baseline DW model does not have this behavior. • For the NetScience network, for a fixed value of c_0, our adaptive-confidence DW model yields at least as many major clusters as the baseline DW model. For this network, μ has little effect on the number of major clusters.
$W(T_f)$	• The baseline DW model always yields $W(T_f) = 1$. Our adaptive-confidence DW model also yields $W(T_f) = 1$ for the complete graph with $c_0 \geq 0.4$ and the NetScience network with $c_0 \in \{0.8, 0.9\}$. • When $W(T_f) < 1$, for fixed values of the parameters γ, δ, and c_0, decreasing μ tends to also decrease $W(T_f)$ for both the complete graph and the NetScience network.

5.6 Conclusions and Discussion

5.6.1 Summary and discussion of our results

We developed two bounded-confidence models (BCMs) — a synchronously-updating one that generalizes the Hegselmann–Krause (HK) model and an asynchronously-updating one that generalizes the Deffuant–Weisbuch (DW) model — with adaptive confidence bounds. The confidence bounds in our adaptive-confidence BCMs are distinct for each dyad (i.e., pair of adjacent nodes) of a network and change when nodes interact with each other. One can interpret the changes in confidence bounds as changes in receptiveness between nodes. We demonstrated that incorporating time-dependent, adaptive confidence bounds in our BCMs yields a variety of interesting behaviors, such as adjacent nodes that converge to the same limit opinion but are eventually unreceptive to each other.

For both our adaptive-confidence HK model and our adaptive-confidence DW model, we proved convergence properties for the dyadic confidence bounds and the limiting behaviors of effective graphs, which track which nodes of a network are able to influence each other. We demonstrated using numerical simulations that our BCMs have fewer major opinion clusters and take longer to converge than the associated baseline BCMs. See Table 5.3 for a summary of the trends in our adaptive-confidence HK model, and see Table 5.4 for a summary of the trends in our adaptive-confidence DW model.

The results of our numerical simulations of our adaptive-confidence BCMs complement our theoretical results (which informed the stopping criteria in our computations). For our adaptive-confidence HK model, we proved that (1) all dyadic confidence bounds must converge either to 0 or to 1 and (2) the dyadic confidence bounds between nodes in different limit opinion clusters must converge to 0 (see Theorem 5.2). For our adaptive-confidence DW model, we proved that analogous results hold almost surely (see Theorem 5.8). However, in both of our adaptive-confidence BCMs, the dyadic confidence bounds between nodes in

the same limit opinion cluster do not necessarily converge to 1, as it is possible for them to instead converge to 0. Indeed, when the confidence-decrease parameter $\delta < 1$, our numerical simulations of both of our adaptive-confidence BCMs demonstrate for a wide range of the other BCM parameter values that some dyads in the same final opinion cluster have confidence bounds that converge to 0. Although the nodes in these dyads are unreceptive to each other, they still converge to the same opinion. The nodes in these dyads do not have an edge between them in the final effective graph, so the final opinion clusters (i.e., the connected components of the final effective graph) in our BCMs can have a richer structure than those in the baseline BCMs.

5.6.2 Future work

Our investigation lays groundwork and provides a point of comparison for the study of more complicated adaptive-confidence mechanisms in BCMs. Future investigations of adaptive-confidence BCMs include establishing additional theoretical guarantees, examining and validating such BCMs in sociological contexts, and generalizing these models in various ways.

It is worthwhile to further explore the theoretical guarantees of our adaptive-confidence BCMs. We showed (see Theorem 5.11) that, almost surely, the effective graph in our adaptive-confidence DW model eventually only has edges between nodes in the same limit opinion cluster. However, unlike for our adaptive-confidence HK model (see Theorem 5.6), we did not prove any guarantee that the effective graph of our adaptive-confidence DW model is eventually constant (not even almost surely). Further theoretical analysis of our adaptive-confidence DW model can help strengthen knowledge of its properties, including the structural properties of the limit effective graphs.

It is also relevant to analytically and numerically study the mutual receptiveness of nodes in our BCMs when they reach a consensus state. In our numerical simulations of our adaptive-confidence BCMs, when the confidence-decrease parameter $\delta < 1$, some adjacent nodes in

the same final opinion cluster are eventually not receptive to each other. More specifically, our numerical simulations suggest that some adjacent nodes can converge to the same limit opinion without having an edge between them in the limit effective graph. One can explore this behavior of our BCMs and determine how the model parameters influence the existence of edges between adjacent nodes with the same limit opinion in limit effective graphs.

It is also important to consider how the behaviors of our BCMs connect to real-life social situations. One can interpret the opinion values in our models as representing outwardly expressed opinions, which may differ from internally held beliefs [Kur95]. The achievement of a "consensus" can represent agents arriving at the same outwardly expressed behavior or decision, rather than achieving an actual agreement of their internal values [Hor62]. Researchers have studied models with both internal and expressed opinions [Noo20, CLY20, HLJ21], and one can incorporate such considerations into adaptive-confidence BCMs.

In our adaptive-confidence BCMs, adjacent agents that are unreceptive to each other's opinions can still interact with each other. Alternatively, a pair of agents can eventually stop interacting with each other — effectively changing the network structure — after repeated negative interactions. Researchers have modeled such ideas, along with network restructuring to consider new social interactions, using adaptive networks with edge rewiring [KFP23, PRM22]. A possible area of further study is the investigation of which models effectively have "mediator" nodes that assist in bringing together the opinions of agents that are unreceptive to each other or no longer interact. If there are such mediator nodes, one can examine whether or not they share common characteristics or are identifiable from network structure and initial agent opinions.

There are many possible areas to explore in the study of adaptive opinion models. In research on opinion dynamics, it is important to incorporate network adaptivity, which provides fertile ground for theoretical, computational, and empirical investigations of opinion dynamics.

5.A Proofs of Our Theoretical Results for Our Adaptive-Confidence DW Model

We now prove the results for our adaptive-confidence DW model that we presented in Section 5.3.2.

5.A.1 Proofs of our confidence-bound results

We first prove Lemma 5.9, which states that each confidence bound $c_{ij}(t)$ is eventually monotone.

Proof of Lemma 5.9. We first consider $c_{ij}(t)$ for adjacent nodes, i and j, that are in different limit opinion clusters (i.e., $x^i \neq x^j$). Choose a time T such that the inequalities

$$|x_k(t) - x^k| < \frac{1}{4} \min_{x^m \neq x^n} |x^m - x^n|, \tag{5.27}$$

$$|x_k(t) - x_k(t')| < \frac{\mu}{4} \min_{x^m \neq x^n} |x^m - x^n| \tag{5.28}$$

hold for each node k and all times $t' > t \geq T$.

We claim that $c_{ij}(t)$ is monotone decreasing (i.e., $c_{ij}(t+1) \leq c_{ij}(t)$) for all $t \geq T$. Note that

$$|x^i - x^j| \geq \min_{x^m \neq x^n} |x^m - x^n|. \tag{5.29}$$

By the triangle inequality and (5.27), we have

$$\begin{aligned} |x^i - x^j| &\leq |x^i - x_i(t)| + |x_i(t) - x_j(t)| + |x_j(t) - x^j| \\ &< \frac{1}{2} \min_{x^m \neq x^n} (|x^m - x^n|) + |x_i(t) - x_j(t)|. \end{aligned}$$

Rearranging terms and using (5.29) yields

$$|x_i(t) - x_j(t)| > \frac{1}{2} \min_{x^m \neq x^n} |x^m - x^n| . \tag{5.30}$$

Suppose that $c_{ij}(t)$ increases (i.e., $c_{ij}(t+1) > c_{ij}(t)$) at time $t \geq T$. This implies that $x_j(t+1) = x_j(t) + \mu(x_i(t) - x_j(t))$, which in turn implies that

$$|x_j(t+1) - x_j(t)| = \mu |x_i(t) - x_j(t)| . \tag{5.31}$$

By (5.28), we have

$$|x_j(t+1) - x_j(t)| < \frac{\mu}{4} \min_{x^m \neq x^n} |x^m - x^n| , \tag{5.32}$$

and (5.31) and (5.32) together imply that

$$|x_i(t) - x_j(t)| < \frac{1}{4} \min_{x^m \neq x^n} |x^m - x^n| . \tag{5.33}$$

The inequalities (5.30) and (5.33) cannot hold simultaneously. Therefore, any interactions between nodes i and j for times $t \geq T$ must result in a decrease of c_{ij}. Consequently, for all adjacent nodes i and j from distinct limit opinion clusters, c_{ij} is monotone decreasing (i.e., $c_{ij}(t+1) \leq c_{ij}(t)$) for all $t \geq T$.

Now consider adjacent nodes, i and j, that are in the same limit opinion cluster (i.e., $x^i = x^j$). Choose a time $T > 0$ such that

$$|x_k(t) - x^k| < \frac{\gamma}{2} \tag{5.34}$$

for each node k and all times $t \geq T$. We claim that there is some time $T_{ij} \geq T$ such that either $c_{ij}(t)$ is monotone decreasing (i.e., $c_{ij}(t+1) \leq c_{ij}(t)$) or it is monotone increasing (i.e., $c_{ij}(t+1) \geq c_{ij}(t)$) for all $t \geq T_{ij}$.

If $c_{ij}(t)$ is monotone decreasing for all $t \geq T$, choose $T_{ij} = T$. If $c_{ij}(t)$ is not monotone

decreasing for all $t \geq T$, there must exist a time $T_{ij} \geq T$ at which $|x_i(T_{ij}) - x_j(T_{ij})| < c_{ij}(T_{ij})$. This implies that

$$c_{ij}(T_{ij}+1) = c_{ij}(T_{ij}) + \gamma(1 - c_{ij}(T_{ij})) \geq \gamma\,. \tag{5.35}$$

We claim that $c_{ij}(t)$ only increases or remains constant for times $t \geq T_{ij}$. By (5.34), we have

$$|x_k(t) - x_{k'}(t)| \leq |x_k(t) - x^k| + |x^k - x^{k'}| + |x^{k'} - x_{k'}(t)| < \gamma \tag{5.36}$$

for each node pair k and k' with $x^k = x^{k'}$ and all times $t \geq T$. Therefore,

$$|x_i(t) - x_j(t)| < \gamma \tag{5.37}$$

for all times $t \geq T_{ij} \geq T$, which implies that subsequent interactions between nodes i and j increase $c_{ij}(t)$ (because $c_{ij}(t) \geq \gamma$). Consequently, if $c_{ij}(t)$ increases at a certain time $T_{ij} \geq T$, then it subsequently either increases or remains constant. If $c_{ij}(t)$ never increases after time T, then by definition it is eventually monotone decreasing. This implies that $c_{ij}(t)$ is either eventually monotone increasing (i.e., $c_{ij}(t_2) \geq c_{ij}(t_1)$ for all $t_2 > t_1 \geq T$) or eventually monotone decreasing. □

We now prove Lemma 5.10, which states that if $c_{ij}(t)$ converges, then its limit $c^{ij} = \lim_{t\to\infty} c_{ij}(t)$ is either 0 or 1 almost surely.

Proof of Lemma 5.10. We prove this lemma using an argument that is similar to the one that we used to prove Lemma 5.5. Unlike in an HK model, adjacent nodes in a DW model need not interact with each other at each discrete time. In fact, it is possible (although it occurs with probability 0) that there exists a pair of adjacent nodes that only interact a finite number of times.

Given $\epsilon > 0$, choose a time T such that the inequalities

$$|c_{ij}(t) - c^{ij}| < \epsilon/2\,, \tag{5.38}$$

$$|c_{ij}(t_1) - c_{ij}(t_2)| < \frac{1}{2}\,(\min\{1-\delta, \gamma\})\,\epsilon \tag{5.39}$$

hold for all times $t, t_1, t_2 \geq T$. Suppose that we choose the adjacent nodes i and j to interact at some time $t \geq T$. It then follows that either $c_{ij}(t+1) = \delta c_{ij}(t)$ or $c_{ij}(t+1) = c_{ij}(t) + \gamma(1 - c_{ij}(t))$.

Suppose that $c_{ij}(t+1) = \delta c_{ij}(t)$. In this case, we claim that $c^{ij} = 0$. To verify this claim, note that $c_{ij}(t) - c_{ij}(t+1) = (1-\delta)c_{ij}(t)$. We know from (5.39) that $c_{ij}(t) - c_{ij}(t+1) < \frac{1}{2}(1-\delta)\epsilon$, so $c_{ij}(t) < \epsilon/2$. Therefore, with (5.38), we obtain

$$\begin{aligned} 0 \leq c^{ij} &\leq |c^{ij} - c_{ij}(t)| + |c_{ij}(t)| \\ &< \epsilon/2 + \epsilon/2 \\ &= \epsilon\,, \end{aligned}$$

which implies that $c^{ij} = 0$.

Now suppose that $c_{ij}(t+1) = c_{ij}(t) + \gamma(1 - c_{ij}(t))$. Rearranging terms yields $c_{ij}(t+1) - c_{ij}(t) = \gamma(1 - c_{ij}(t)) < \frac{1}{2}\gamma\epsilon$, which implies that $1 - c_{ij}(t) < \epsilon/2$. Therefore,

$$\begin{aligned} 0 \leq 1 - c^{ij} &\leq |1 - c_{ij}(t)| + |c_{ij}(t) - c^{ij}| \\ &< \epsilon/2 + \epsilon/2 \\ &= \epsilon\,, \end{aligned}$$

which implies that $c^{ij} = 1$.

Consequently, if nodes i and j interact infinitely often, c^{ij} must be either 0 or 1. By the Borel–

Cantelli lemma, nodes i and j interact infinitely many times with probability 1. Therefore, it is almost surely the case that either $c^{ij} = 0$ or $c^{ij} = 1$. □

5.A.2 Proof of the effective-graph theorem for our adaptive-confidence DW model

We now prove Theorem 5.11, which is our main result about effective graphs in our adaptive-confidence DW model. It states that, almost surely, an effective graph in our adaptive-confidence DW model eventually has edges only between adjacent nodes in the same limit opinion cluster.

Proof of Theorem 5.11. By Theorem 5.8, for adjacent nodes i and j that are in different limit opinion clusters, $c_{ij}(t)$ almost surely converges to 0. Therefore, almost surely, there is some T_1 such that

$$c_{ij}(t) < \frac{1}{2} \min_{x^m \neq x^n} |x^m - x^n| \tag{5.40}$$

for all times $t \geq T_1$. We also choose T_2 such that

$$|x_k(t) - x^k| < \frac{1}{4} \min_{x^m \neq x^n} |x^m - x^n| \tag{5.41}$$

for each node k and all times $t \geq T_2$.

Let $T = \max\{T_1, T_2\}$, and fix adjacent nodes i and j that are in different limit opinion clusters. The time T exists almost surely because T_1 exists almost surely. For all times $t \geq T$, the inequality (5.41) implies that

$$\begin{aligned} |x^i - x^j| &\leq |x_i(t) - x^i| + |x_i(t) - x_j(t)| + |x_j(t) - x^j| \\ &\leq \frac{1}{2} \min_{x^m \neq x^n} |x^m - x^n| + |x_i(t) - x_j(t)| \,. \end{aligned}$$

Because $\min_{x^m \neq x^n} |x^m - x^n| \leq |x^i - x^j|$, it follows that

$$\min_{x^m \neq x^n} |x^m - x^n| \leq \frac{1}{2} \min_{x^m \neq x^n} |x^m - x^n| + |x_i(t) - x_j(t)| \,. \tag{5.42}$$

Therefore, with (5.40), we obtain

$$c_{ij}(t) < \frac{1}{2} \min_{x^k \neq x^{k'}} |x^k - x^{k'}| \leq |x_i(t) - x_j(t)| \,.$$

That is, $|x_i(t) - x_j(t)| \geq c_{ij}(t)$ for all $t \geq T$, so the edge (i, j) is not in the effective graph at time t for all $t \geq T$. Therefore, the only edges in the effective graph for times $t \geq T$ are between nodes in the same limit opinion cluster. □

The effective-graph theorem for our adaptive-confidence DW model (see Theorem 5.11) is weaker than that for our adaptive-confidence HK model (see Theorem 5.6). In particular, we are unable to conclude for the adaptive-confidence DW model that the effective graph is eventually constant (or even almost surely eventually constant). The obstruction to obtaining such a guarantee is the stochasticity that arises from the asynchronous opinion updating of the adaptive-confidence DW model. In particular, consider adjacent nodes i and j in the same limit opinion cluster. By Lemma 5.9, there exists a time T such that c_{ij} is monotone for all times $t \geq T$. Suppose that c_{ij} is monotone decreasing for all $t \geq T$. If nodes i and j interact at time $t \geq T$, then $c_{ij}(t)$ decreases and the edge (i, j) is not in the effective graph (i.e., $(i, j) \notin E_{\text{Eff}}(t)$). However, if nodes i and j do not interact at time t, it is possible that edge $(i, j) \in E_{\text{Eff}}(t)$.(By contrast, for our adaptive-confidence HK model, which updates synchronously, the existence of a time T such that c_{ij} is strictly decreasing for all times $t \geq T$ implies that the edge $(i, j) \notin E_{\text{Eff}}(t)$ for times $t \geq T$.) Suppose that $t_1, t_2, \ldots$ are successive times at which nodes i and j interact after time T. For each s, it can be the case that both the inequality $|x_i(t_s) - x_j(t_s)| \geq c_{ij}(t_s)$ holds and there is a time $\tilde{t}_s$ between t_s and t_{s+1} such that $|x_i(\tilde{t}_s) - x_j(\tilde{t}_s)| < c_{ij}(\tilde{t}_s) = c_{ij}(t_s)$. That is, between each pair of interaction

times t_s and t_{s+1}, the opinions of nodes i and j can (because of other adjacent nodes) first become close enough so that the difference between their opinions is less than their confidence bound and then subsequently become sufficiently far apart so that the difference between their opinions exceeds their confidence bound. In this situation, the effective graph is not eventually constant.

Although the example in the previous paragraph may seem pathological, it is unclear whether and how frequently such situations can occur. There also may be other scenarios in which an effective graph is not eventually constant. This issue does not arise in the proof of Theorem 5.6 because the nodes in each dyad interact at every time in the adaptive-confidence HK model.

5.B Proof of the Effective-Graph Theorem for the Baseline DW Model

We now prove Theorem 5.12, which is our convergence result for effective graphs in the baseline DW model. To do this, we first prove Lemma 5.13 and Lemma 5.14.

Lemma 5.13. *Consider the baseline DW model (with update rule* (5.4)*). There is a time T_1 and there is almost surely a time T_2 such that the following statements hold for all adjacent nodes i and j.*

(1) If $|x^i - x^j| < c$, then $|x_i(t) - x_j(t)| < c$ and the edge (i, j) is in the effective graph for all times $t \geq T_1$.

(2) If $|x^i - x^j| > c$, then $|x_i(t) - x_j(t)| > c$ and the edge (i, j) is not in the effective graph at any time $t \geq T_1$.

(3) If $|x^i - x^j| = c$, then $|x_i(t) - x_j(t)| \geq c$ and the edge (i, j) is not in the effective graph at any time $t \geq T_2$.

Proof. Consider adjacent nodes i and j, and let $\Delta_{ij} = |x^i - x^j|$ denote the difference between their opinions.

We first consider the case in which $\Delta_{ij} \neq c$. Choose a time T_{ij} such that

$$|x_k(t) - x^k| < \frac{1}{2}|c - \Delta_{ij}| \tag{5.43}$$

for node $k \in \{i, j\}$ and all times $t \geq T_{ij}$.

Suppose that $\Delta_{ij} < c$. For all times $t \geq T_{ij}$, we have

$$\begin{aligned} |x_i(t) - x_j(t)| &\leq |x_i(t) - x^i| + |x^i - x^j| + |x_j(t) - x^j| \\ &< 2\left(\frac{1}{2}\right)(c - \Delta_{ij}) + \Delta_{ij} \\ &= c\,. \end{aligned}$$

Therefore, the edge (i, j) is in the effective graph for all $t \geq T_{ij}$.

Now suppose that $\Delta_{ij} > c$. Without loss of generality, let $x^i > x^j$. For all times $t \geq T_{ij}$, we have

$$\begin{aligned} x_i(t) - x_j(t) &> \left(x^i - \frac{1}{2}\,|c - \Delta_{ij}|\right) - \left(x^j + \frac{1}{2}\,|c - \Delta_{ij}|\right) \\ &= (x^i - x^j) - |c - \Delta_{ij}| \\ &= \Delta_{ij} - \Delta_{ij} + c \\ &= c\,. \end{aligned}$$

Therefore, the edge (i, j) is not in the effective graph at any time $t \geq T_{ij}$.

If there are no adjacent nodes i and j with $|x^i - x^j| \neq c$, then let $T_1 = 0$. Otherwise, let

$$T_1 = \max_{(i,j)\in E} \{T_{ij} \text{ such that } |x^i - x^j| \neq c\}\,. \tag{5.44}$$

We have shown that statements (1) and (2) hold for all times $t \geq T_1$.

We now consider the case $\Delta_{ij} = c$. Without loss of generality, let $x^i > x^j$. Choose a time $\tilde{T}_{ij}$ so that

$$|x^k - x_k(t)| < \frac{\mu c}{2(1+2\mu)} \tag{5.45}$$

for node $k \in \{i, j\}$ and all times $t \geq \tilde{T}_{ij}$. We will show that, almost surely, there are a finite number of times $t \geq \tilde{T}_{ij}$ such that $|x_i(t) - x_j(t)| < c$. Suppose on the contrary that there is a sequence $t_1, t_2, \ldots$ of times such that $t_k \geq \tilde{T}_{ij}$ and $|x_i(t_k) - x_j(t_k)| < c$ for all k. At each time t, nodes i and j interact with probability $1/|E| > 0$, where $|E|$ is the number of edges in the graph. Therefore, with probability 1, nodes i and j interact at some time $t_k \geq \tilde{T}_{ij}$ with $|x_i(t_k) - x_j(t_k)| < c$. Nodes i and j compromise their opinions at time t_k, so the inequality (5.45) implies that

$$\begin{aligned}
|x_i(t+1) - x_i(t)| = \mu|x_i(t) - x_j(t)| &\geq \mu\left[\left(x^i - \frac{\mu c}{2(1+2\mu)}\right) - \left(x^j + \frac{\mu c}{2(1+2\mu)}\right)\right] \\
&= \mu\left[c - 2\left(\frac{\mu c}{2(1+2\mu)}\right)\right] \\
&= \frac{\mu c[(1+2\mu) - \mu]}{1+2\mu} \\
&= (1+\mu)\frac{\mu c}{1+2\mu} \\
&> \frac{\mu c}{1+2\mu}\,.
\end{aligned} \tag{5.46}$$

From the inequality (5.45), we have

$$|x_i(t+1) - x_i(t)| \leq |x_i(t+1) - x^i| + |x^i - x_i(t)| < 2\left(\frac{\mu c}{2(1+2\mu)}\right) = \frac{\mu c}{1+2\mu}\,,$$

which cannot hold simultaneously with inequality (5.46). Therefore, with probability 1, there are a finite number of times $t \geq \tilde{T}_{ij}$ such that $|x_i(t) - x_j(t)| < c$. Consequently, there almost surely exists some time $T_{ij} \geq \tilde{T}_{ij}$ such that $|x_i(t) - x_j(t)| \geq c$ and the edge (i, j) is not in the effective graph for any $t \geq T_{ij}$.

If there are no adjacent nodes i and j with $|x^i - x^j| \neq c$, then let $T_2 = 0$. Otherwise, let

$$T_2 = \max_{(i,j)\in E} \{T_{ij} \text{ such that } |x^i - x^j| = c\}, \tag{5.47}$$

where T_2 exists almost surely because each T_{ij} exists almost surely. We have shown that statement (3) holds for all times $t \geq T_2$ if T_2 exists. □

Lemma 5.14. *For adjacent nodes i and j with $|x^i - x^j| < c$, we have that $x^i = x^j$ almost surely.*

Proof. Fix adjacent nodes i and j with $|x^i - x^j| < c$, and let $\Delta_{ij} = |x^i - x^j|$ denote the distance between their opinions. Without loss of generality, let $x^i > x^j$. We want to show that $\Delta_{ij} = 0$ almost surely. Suppose instead that $\Delta_{ij} > 0$. Fix ϵ so that $0 < \epsilon < \min\{\frac{1}{4}(c - \Delta_{ij}), \frac{\Delta_{ij}}{2(1+1/\mu)}\}$ and choose T_{ij} so that

$$|x^k - x_k(t)| < \epsilon \tag{5.48}$$

for each node k and all times $t \geq T_{ij}$.

By the Borel–Cantelli lemma, there is almost surely some time $t \geq T_{ij}$ at which nodes i and j interact. The inequality $\epsilon < \frac{1}{4}(c - \Delta_{ij})$ implies that

$$\begin{aligned} |x_i(t) - x_j(t)| &\leq |x_i(t) - x^i| + |x^i - x^j| + |x^j - x_j(t)| \\ &< \frac{1}{4}(c - \Delta_{ij}) + \Delta_{ij} + \frac{1}{4}(c - \Delta_{ij}) = \frac{1}{2}(\Delta_{ij} + c) \\ &< c, \end{aligned}$$

so nodes i and j are receptive to each other at time t. Consequently, if they interact at time t, they update their opinions and

$$\begin{aligned} x_j(t+1) &= x_j(t) + \mu[x_i(t) - x_j(t)] \\ &\geq x_j(t) + \mu[x^i - \epsilon - (x^j + \epsilon)] \\ &= x_j(t) + \mu(\Delta_{ij} - 2\epsilon) \\ &\geq (x^j - \epsilon) + \mu(\Delta_{ij} - 2\epsilon) \\ &> x^j + \epsilon\,, \end{aligned} \tag{5.49}$$

where the last inequality holds because $\epsilon < \frac{\Delta_{ij}}{2(1+1/\mu)}$, which we rearrange to obtain $2\epsilon < \mu(\Delta_{ij} - 2\epsilon)$. The inequality (5.48) implies that

$$|x^j - x_j(t+1)| < \epsilon\,, \tag{5.50}$$

which cannot hold simultaneously with the inequality (5.49). Therefore, if $0 < x^i - x^j < c$, then nodes i and j cannot interact at times $t \geq T_{ij}$. However, by the Borel–Cantelli lemma, nodes i and j almost surely interact infinitely often. Consequently, $0 < x^i - x^j < c$ with probability 0. Therefore, we almost surely have $x^i = x^j$. □

We now use Lemma 5.13 and Lemma 5.14 to prove Theorem 5.12.

Proof of Theorem 5.12. There is a time T_1 such that statements (1) and (2) of Lemma 5.13 hold, and there is almost surely a time T_2 such that statement (3) of Lemma 5.13 holds. Therefore, there is almost surely a time $T = \max\{T_1, T_2\}$ such that all three statements (1)–(3) of Lemma 5.13 hold for all times $t \geq T$. Consequently, the edges of the effective graph satisfy $E_{\text{Eff}}(t) = E_{\text{Eff}}(T)$ for all $t \geq T$. The effective graph is thus eventually constant with respect to time for all $t \geq T$.

Suppose that the limit effective graph $\lim_{t\to\infty} G_{\text{eff}}(t)$ exists. For adjacent nodes i and j with the same limit opinion (i.e., $x^i = x^j$), we know that $|x^i - x^j| = 0 < c$. By statement (1) of Lemma 5.13, there thus exists a time T_1 such that the edge (i, j) is in the effective graph for all times $t \geq T_1$. Therefore, the edge (i, j) is in the limit effective graph.

Now suppose that the edge (i, j) is in the limit effective graph. We seek to show that $x^i = x^j$ almost surely. Because the edge (i, j) is in the limit effective graph, there exists a time $\tilde{T}$ such that $(i, j) \in E_{\text{Eff}}(t)$ for all times $t \geq \tilde{T}$. Consequently, by statement (2) of Lemma 5.13, it cannot be the case that $|x^i - x^j| > c$. Therefore, by statement (3) of Lemma 5.13, it almost surely cannot be the case that $|x^i - x^j| = c$. Consequently, we almost surely have $|x^i - x^j| < c$. By Lemma 5.14, it is almost surely the case that $x^i = x^j$. □

5.C Additional Results and Discussion of Our Numerical Simulations of Our Adaptive-Confidence HK Model

In this section, we show additional numerical results for our adaptive-confidence HK model on ER graphs (see Section 5.C.1), two-community SBM graphs (see Section 5.C.2), and the Reed College network (see Figure 5.9). We again examine the numbers of major and minor clusters, the Shannon entropy $H(T_f)$ (see Equation (5.25)), the weighted-average edge fraction $W(T_f)$ (see Equation (5.26)), and the convergence time T_f. We consider the values of the BCM parameters (namely, the confidence-increase parameter γ, confidence-decrease parameter δ, and initial confidence bound c_0) in Table 5.2. Each point in our plots is a mean of our numerical simulations for the associated values of the BCM parameter set (γ, δ, c_0). All plots, including those that we do not include in this section, are available in our code repository.

5.C.1 Erdős–Rényi graphs

We now discuss additional results of our simulations of our adaptive-confidence HK model on $G(N, p)$ ER random graphs. We generate 5 ER random graphs for each value of $p \in \{0.1, 0.5\}$. Each point in our plots is a mean of 50 simulations (from 5 random graphs that each have 10 sets of initial opinions). For fixed BCM parameters (namely, γ, δ, and c_0), our results for $G(1000, 0.5)$ graphs are more similar than those for $G(1000, 0.1)$ graphs to our results for the 1000-node complete graph.

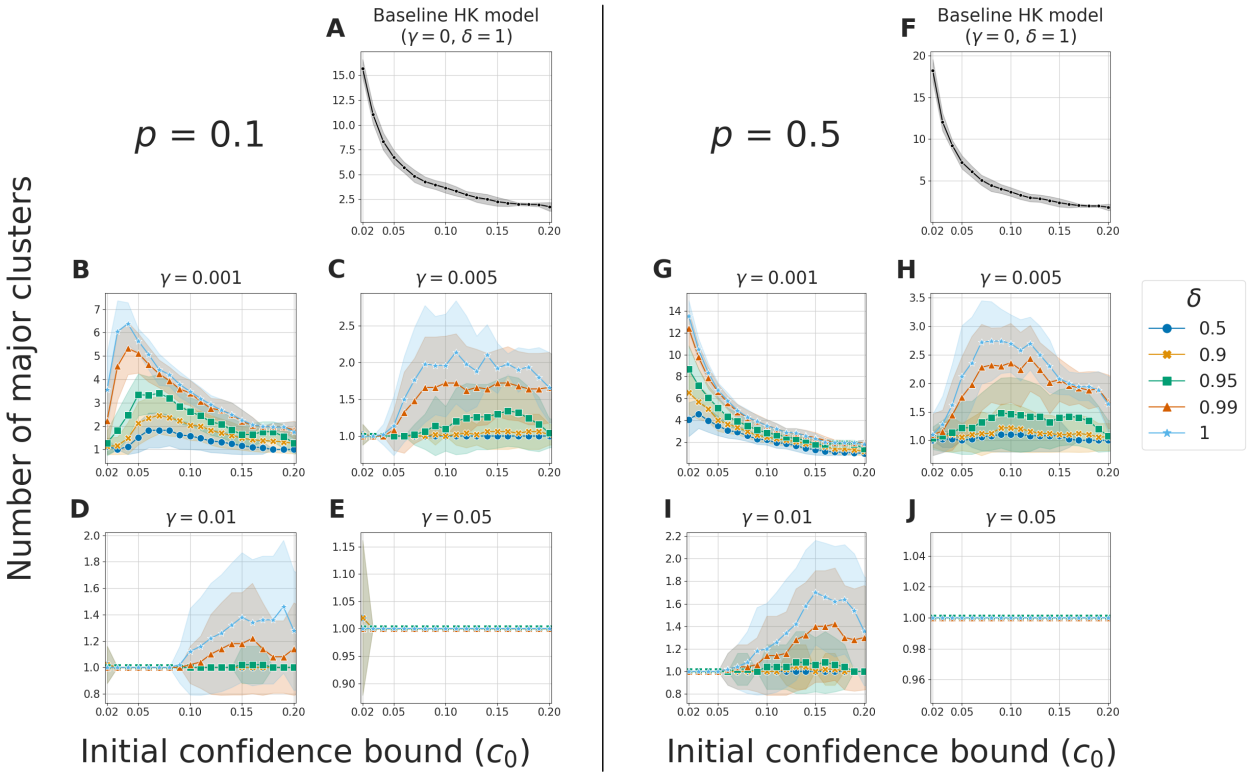


Figure 5.7: The numbers of major clusters in simulations of our adaptive-confidence HK model on $G(1000, p)$ ER random graphs with (A–E) $p = 0.1$ and (F–J) $p = 0.5$ for various combinations of the BCM parameters γ, δ, and c_0.

In Section 5.C.1, we show the numbers of major clusters in our simulations of our adaptive-confidence HK model on ER graphs. For a complete graph and $G(1000, 0.5)$ graphs, for fixed values of γ and δ, the number of major clusters tends to decrease as we increase c_0. For the

$G(1000, 0.1)$ graphs, when $\gamma = 0.001$, small values of c_0 tend to yield few major clusters. As we increase c_0, we observe an initial increase in the number of major clusters followed by a decrease in that number. By contrast, for the 1000-node complete graph, small values of c_0 tend to yield the most major clusters. As we discussed in Section 5.5.1.3, $G(1000, 0.1)$ graphs have more small-degree nodes than the complete graph. These small-degree nodes can easily form minor opinion clusters, especially for small values of c_0. We hypothesize that these minor clusters form quickly in a simulation and that the nodes in them quickly become unreceptive to the other nodes of a network. It is thus possible that the nodes that are not in these minor clusters become receptive to fewer neighbors with conflicting opinions.[21]

As we discussed in Section 5.5.1.3, for small values of c_0, our simulations of the adaptive-confidence HK model on the $G(1000, 0.1)$ graphs yield more minor clusters than our simulations on the $G(1000, 0.5)$ graphs and the 1000-node complete graph. Nevertheless, although Shannon entropy (see Equation (5.25)) accounts for minor clusters, we still observe that it follows similar trends as the number of major clusters for both $p = 0.1$ and $p = 0.5$. Specifically, the Shannon entropy tends to increase as either (1) we decrease γ for fixed δ and c_0 or (2) we increase δ for fixed γ and c_0.

For our simulations on ER graphs with both $p = 0.1$ and $p = 0.5$, we observe the convergence-time trends in Table 5.3. For fixed values of γ, δ, and c_0, the mean convergence time for $p = 0.1$ is at least as long as that for $p = 0.5$. Unlike for the complete graph, the ER graphs do not have a clear trend in the dependence of the convergence time either on γ (with fixed δ and c_0) or on δ (with fixed γ and c_0). As with the 1000-node complete graph, our fastest convergence times for ER graphs typically occur for $\delta = 1$. For fixed γ and c_0, we often observe that the convergence time increases as we decrease δ. However, we do not always observe this trend; for some values of γ and c_0, smaller values of δ yield faster convergence

[21]When the neighbors to which a node is receptive have very different opinions, recall (see Footnote 19) that that node is receptive to "conflicting" opinions., resulting in more consensus (i.e., fewer major opinion clusters).

than larger values of δ.

5.C.2 Two-community stochastic-block-model graphs

We now discuss additional results of our simulations of our adaptive-confidence HK model on two-community SBM graphs. Each of our SBM graphs consists of two complete graphs that are joined by a small number of edges (see Section 5.4.1). This yields a two-community structure.

In Figure 5.8, we show the numbers of major clusters in our simulations on SBM graphs. For fixed values of the BCM parameters (namely, γ, δ, and c_0), these simulations yield similar numbers of major clusters as in our simulations on the 1000-node complete graph (see Section 5.5.1.2) and $G(1000, 0.5)$ ER graphs (see Section 5.C.1). We observe few minor clusters; for each BCM parameter set, the mean number of minor clusters is bounded above by 3. Consequently, the Shannon entropy and the number of major clusters follow similar trends.

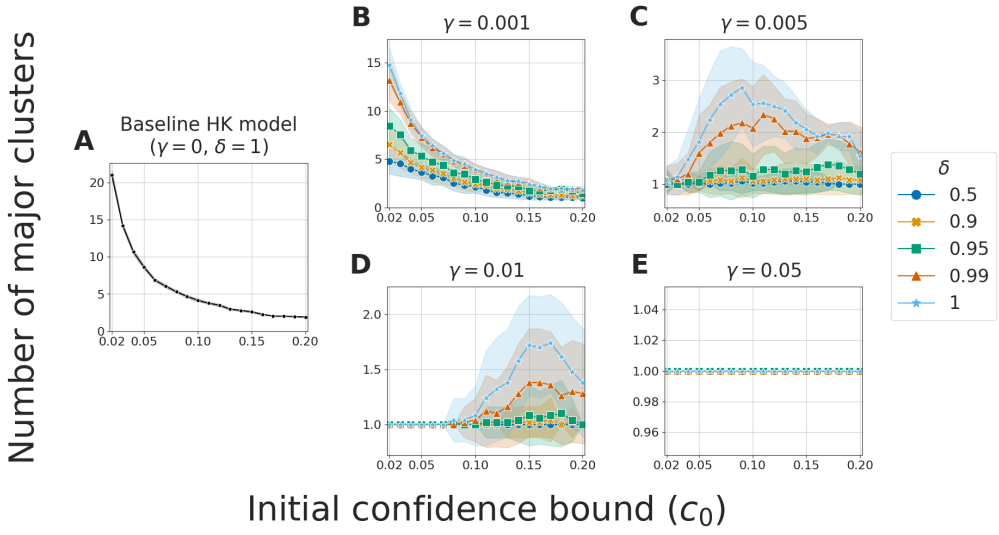


Figure 5.8: The numbers of major clusters in simulations of our adaptive-confidence HK model on 1000-node SBM random graphs with connection probabilities $p_{aa} = p_{bb} = 1$ and $p_{ab} = 0.01$ for various combinations of the BCM parameters γ, δ, and c_0.

The convergence times in our simulations on SBM graphs follow the trends in Table 5.3. For

fixed values of γ and c_0, we do not observe a clear trend in how the convergence time changes as we vary δ. One commonality between the SBM graphs, the ER graphs, and the complete graph is that $\delta = 1$ gives the fastest convergence times. For a wide range of fixed values of γ and δ, we also observe that the convergence time tends to decreases as we increase c_0 for both our adaptive-confidence HK model and the baseline HK model.

5.C.3 Number of major clusters in simulations on the Reed College network

In our simulations of our adaptive-confidence HK model on the FACEBOOK100 networks, the Reed College network (see Figure 5.9) has more major opinion clusters than the other universities for very small initial confidence bounds c_0 (specifically, $c_0 \in \{0.02, 0.03, 0.04\}$). This difference may arise from the small size of the Reed College network in concert with our definition of major cluster. For example, a final opinion cluster with 20 nodes is a major cluster for the Reed College network (which has 962 nodes in its LCC), but an opinion cluster of that size is a minor cluster for the UC Santa Barbara network (which has 14,917 nodes in its LCC).

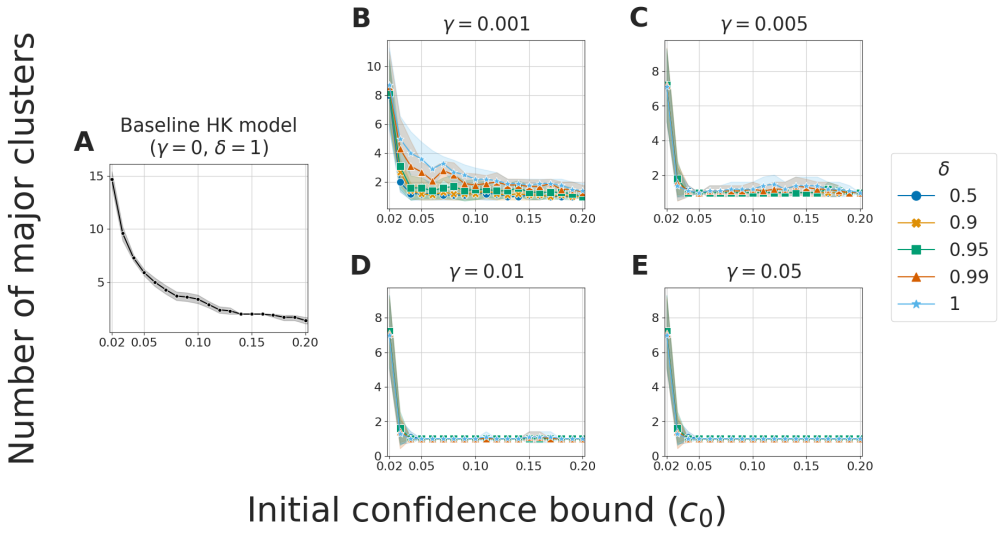


Figure 5.9: The numbers of major clusters in simulations of our adaptive-confidence HK model on the Reed College network for various combinations of the BCM parameters γ, δ, and c_0.

5.D Additional Results and Discussion of Our Numerical Simulations of Our Adaptive-Confidence DW Model

We now further discuss our simulations of our adaptive-confidence DW model. We simulate this model a 100-node complete graph and the NetScience network. We simulate our adaptive-confidence DW model with the values of the BCM parameters (namely, the initial confidence bound c_0, the confidence-increase parameter γ, the confidence-decrease parameter δ, and the compromise parameter μ) in Table 5.2. See Table 5.4 for a summary of the trends for the two networks.

We explore the dependence of the numbers of major and minor clusters, the Shannon entropy $H(T_f)$ (see Equation (5.25)), the weighted-average edge fraction $W(T_f)$ (see Equation (5.26)), and the convergence time T_f on the initial confidence bound c_0. For each value of (γ, δ), we generate one plot; each plot has one curve for each value of the compromise parameter μ. Each point in our plots is the mean of 10 numerical simulations (from 10 sets of initial opinions) with one BCM parameter set $(\gamma, \delta, c_0, \mu)$. We also show one standard deviation from the mean. All plots, including those that we do not present in this section, are available in our code repository.

5.D.1 A complete graph

We first discuss our simulations of our adaptive-confidence DW model on a 100-node complete graph. In the present section, we show plots of the numbers of major opinion clusters (see Figure 5.10) and the weighted-average edge fractions $W(T_f)$ (see Figure 5.11).

Our adaptive-confidence DW model tends to converge more slowly than both the baseline DW model and our adaptive-confidence HK model. Our simulations of our adaptive DW model often reach the bailout time, particularly for small values of c_0 and μ. In Table 5.5, we indicate the numbers of simulations that reach the bailout time. In some simulations,

Table 5.5: Summary of the numbers of simulations of our adaptive-confidence DW model that reach the bailout time of 10^6 time steps. For each combination of the BCM parameters (γ, δ, c_0, and μ), we run 10 simulations, which each have a different set of initial opinions. In each table entry, the focal number is the number of simulations that reach the bailout time and the number in parentheses is the number of those simulations for which we are also unable determine the final opinion clusters. We run our simulations with $(\gamma, \delta) = (0.1, 0.5)$ to convergence (i.e., without a bailout time); for those simulations, we do not track the number of opinion clusters at the bailout time.

		Number of simulations that reach bailout (number of simulations for which we are also unable to determine the final opinion clusters)					
		$\mu = 0.1$			$\mu = 0.3$		$\mu = 0.5$
		$c_0 = 0.1$	$c_0 = 0.2$	$c_0 = 0.3$	$c_0 = 0.1$	$c_0 = 0.2$	$c_0 = 0.1$
$\gamma = 0.1$	$\delta = 0.3$	9 (7)	2 (2)	1 (1)	0	0	0
	$\delta = 0.5$	8	1	0	1	0	0
	$\delta = 0.7$	9 (6)	2 (2)	1 (1)	2 (0)	0	0
$\gamma = 0.3$	$\delta = 0.3$	9 (5)	0	0	2 (1)	0	2 (0)
	$\delta = 0.5$	8 (7)	0	0	2 (2)	0	0
	$\delta = 0.7$	7 (4)	0	0	5 (3)	2 (1)	0
$\gamma = 0.5$	$\delta = 0.3$	9 (6)	0	0	2 (2)	1 (0)	0
	$\delta = 0.5$	8 (4)	0	0	2 (1)	0	0
	$\delta = 0.7$	6 (4)	0	0	7 (4)	0	1 (1)

despite reaching the bailout time, we are still able to identify the final opinion clusters. However, the maximum difference in the opinions of the nodes in these clusters is not within our tolerance value (see (5.24)) of 0.02 for our adaptive-confidence DW model. In those instances, we still use the cluster information to calculate the numbers of major and minor opinion clusters, the Shannon entropy $H(T_f)$, and the weighted-average edge fraction $W(T_f)$. For our simulations of our adaptive-confidence DW model with $(\gamma, \delta) = (0.1, 0.5)$, we run each simulation to convergence (i.e., until we reach the stopping condition that we described in Section 5.4.2). We plot the results of these simulations in Figure 5.10E and Figure 5.11B. Although some simulations reach the bailout time, the information that we are able to obtain about the opinion clusters (from both the simulations that we run to convergence and the simulations that reach the bailout time) give us confidence in the trends in Table 5.4.

In Figure 5.10, we observe for a wide range of BCM parameter values that our adaptive-

confidence DW model yields fewer major clusters (i.e., it encourages more consensus) than the baseline DW model. When $c_0 \geq 0.5$, our adaptive-confidence DW model and the baseline DW model always reach consensus. For fixed values of γ, δ, and c_0, when our adaptive-confidence DW model does not reach consensus, decreasing the compromise parameter μ tends to result in fewer major clusters. By contrast, μ has little effect on the number of major clusters in the baseline DW model. Increasing γ with the other BCM parameters (i.e., δ, c_0, and μ) fixed also tends to result in fewer major clusters. Changing δ with the other parameters fixed has little effect on the number of major clusters. In fact, changing δ with the other parameters fixed appears to have little effect on any of the computed quantities, so we show results only for $\delta = 0.5$ in our subsequent figures. In our code repository, we include plots for the other examined values of δ.

We observe very few minor clusters in our simulations of our adaptive-confidence DW model on the 100-node complete graph. For each BCM parameter set $(\gamma, \delta, c_0, \mu)$, the mean number of minor clusters in our 10 simulations is bounded above by 1. Consequently, the number of major clusters and Shannon entropy follow similar trends. Overall, in our simulations on the 100-node complete graph, our adaptive-confidence DW model encourages more consensus than the baseline DW model and this difference between these two models becomes more pronounced for larger values of the confidence-increase parameter γ and smaller values of the compromise parameter μ.

In Figure 5.11, we show the weighted-average edge fraction $W(T_f)$ (see (5.26)). The baseline DW model always has $W(T_f) = 1$. By contrast, for sufficiently small initial confidence values c_0, our adaptive-confidence DW model yields $W(T_f) < 1$. For $\mu = 0.1$ and small c_0 (specifically, $c_0 \leq 0.3$), our adaptive-confidence DW model can reach consensus with $W(T_f) < 1$. As in our adaptive-confidence HK model (see our discussion in Section 5.5.1.2), this observation indicates that some adjacent nodes in the same final opinion cluster are not receptive to each other.

For fixed values of $c_0 \leq 0.3$ and μ, our adaptive-confidence DW model tends to converge more slowly than the baseline DW model. Additionally, when we fix the other BCM parameters (i.e., γ, δ, and μ), the convergence time tends to increase as we decrease c_0. As we showed in Table 5.5, for small values of c_0 (specifically, $c_0 \in \{0.1, 0.2\}$), more simulations reach the bailout time as we decrease μ. In both our adaptive-confidence DW model and the baseline DW model, $\mu = 0.1$ yields longer convergence times than $\mu \in \{0.3, 0.5\}$ for fixed values of γ, δ, and c_0.

5.D.2 Network of network-scientist coauthorships

We now discuss our simulations of our adaptive-confidence DW model on the NetScience network [New06], which is a network of network scientists with unweighted and undirected edges that encode paper coauthorships.

For the NetScience network and fixed values of c_0 and μ, our adaptive-confidence DW model tends to have at least as many major opinion clusters (see Figure 5.12) and minor opinion clusters (see Figure 5.13) as the baseline DW model. In Figure 5.13, we see for $c_0 \leq 0.5$ that both our adaptive-confidence DW model and the baseline DW model yield many more minor clusters for the NetScience network than for the 100-node complete graph. For values of c_0 that are near the transition between consensus and opinion fragmentation (specifically, $c_0 \in \{0.3, 0.4, 0.5\}$), our adaptive-confidence DW model yields noticeably more major clusters and minor clusters than the baseline DW model. The transition between consensus and fragmentation appears to occur for a larger threshold in our adaptive-confidence DW model than in the baseline DW model. For the NetScience network (and unlike for the 100-node complete graph), changing the value of μ with the other BCM parameters fixed appears to have little effect on the numbers of major and minor opinion clusters.

For the NetScience network and fixed values of c_0 and μ, our adaptive-confidence DW model has convergence times that are similar to those of the baseline DW model. All of our

simulations of our adaptive-confidence DW model on the NETSCIENCE network converge before reaching the bailout time. We obtain the longest convergence times for $c_0 = 0.3$. By contrast, for the 100-node complete graph, the convergence time increases as we decrease c_0 and many simulations reach the bailout time for $c_0 \in \{0.1, 0.2\}$. In both our adaptive-confidence DW model and the baseline DW model, $\mu = 0.1$ yields longer convergence times than $\mu \in \{0.3, 0.5\}$ for fixed values of γ, δ, and c_0. We do not observe a clear trend in how the convergence time changes either as a function of γ (with fixed δ, c_0, and μ) or as a function of δ (with fixed γ, c_0, and μ).

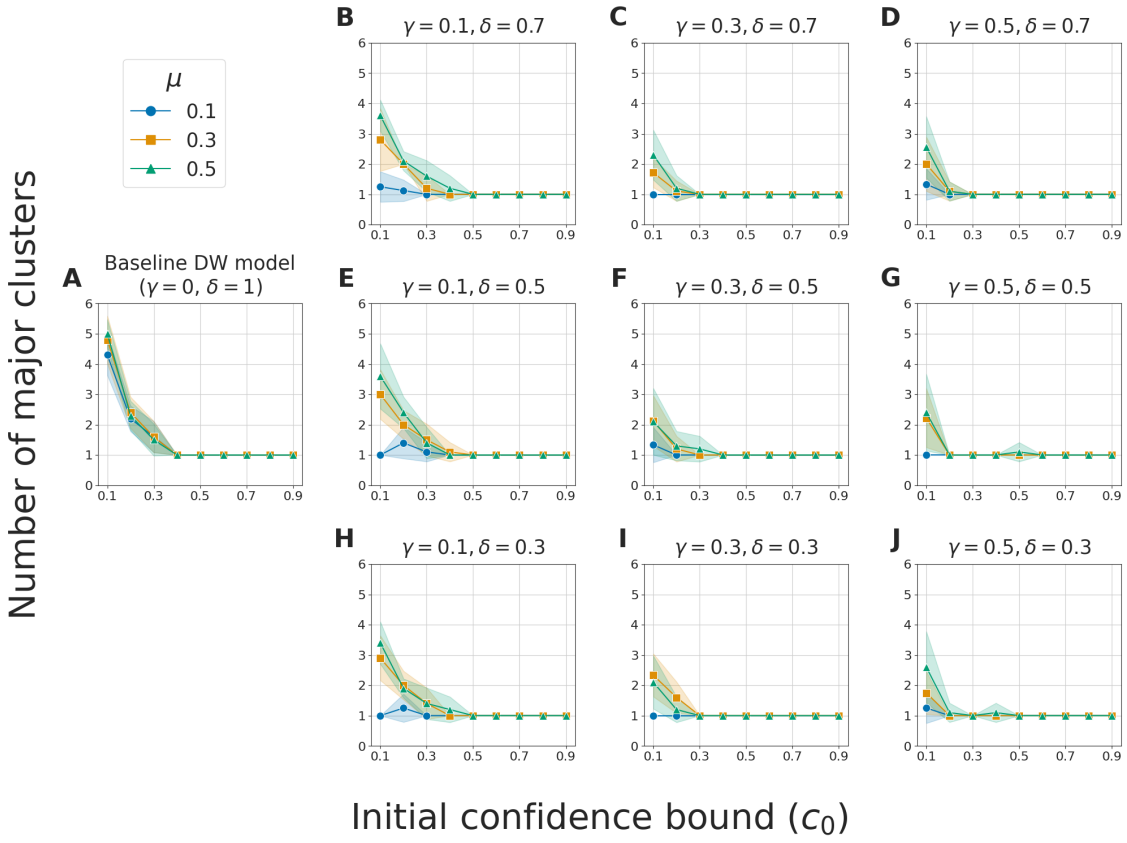


Figure 5.10: The numbers of major clusters in simulations of (A) the baseline DW model and (B–J) our adaptive-confidence DW model on a 100-node complete graph for various combinations of the BCM parameters γ, δ, c_0, and μ. In this figure and subsequent figures, we do not use simulations in which we are unable to determine the final opinion clusters (see Table 5.5) to calculate the means and standard deviations. In (E), in which we show our simulations with $(\gamma, \delta) = (0.1, 0.5)$, we run all of our simulations to convergence (i.e., we ignore the bailout time) and use all of our simulations to calculate the mean numbers of major opinion clusters.

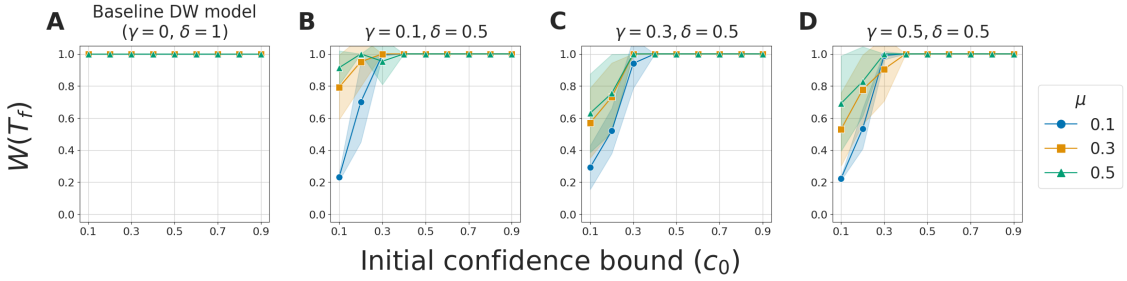


Figure 5.11: The weighted-average edge fraction $W(T_f)$ (see Equation (5.26)) in simulations of (A) the baseline DW model and (B–D) our adaptive-confidence DW model on a 100-node complete graph for various combinations of the BCM parameters γ, δ, c_0, and μ. In (E), in which we show our simulations with $(\gamma, \delta) = (0.1, 0.5)$, we run all of our simulations to convergence (i.e., we ignore the bailout time) and use the resulting final opinion clusters.

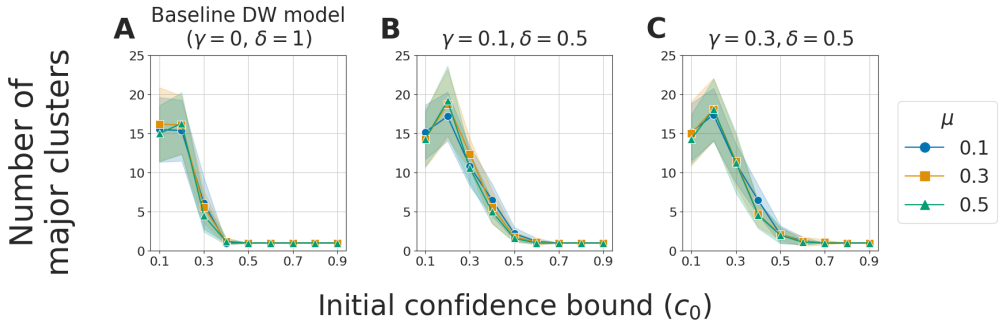


Figure 5.12: The numbers of major clusters in simulations of (A) the baseline DW model and (B, C) our adaptive-confidence DW model on the NetScience network for various combinations of the BCM parameters γ, δ, c_0, and μ.

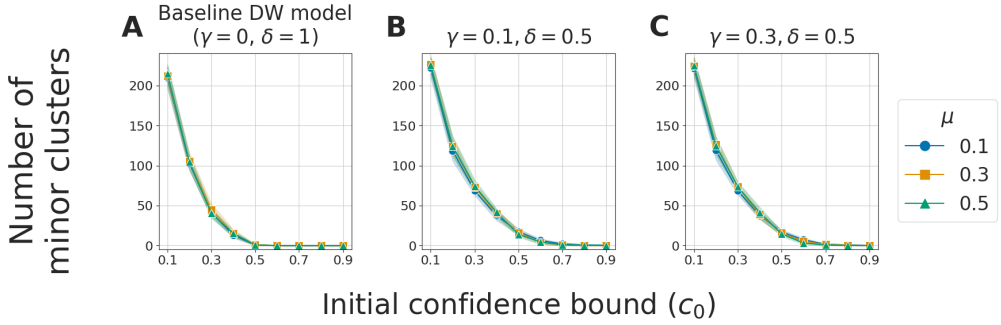


Figure 5.13: The numbers of minor clusters in simulations of (A) the baseline DW model and (B, C) our adaptive-confidence DW model on the NetScience network for various combinations of the BCM parameters γ, δ, c_0, and μ.

CHAPTER 6

Conclusion

In this thesis, we explored persistent homology (PH), complex social systems, and the interface between the two.

We began by discussing persistence modules over a principal ideal domain (PID) and conditions under which interval decompositions exist. In the context of PH, an interval decomposition of a persistence module gives important topological information to compute the persistence diagram (PD) of a filtration (see Section 2.3). We proved that a persistence module that is pointwise free and finitely-generated over a PID has an interval decomposition if and only if its structure maps have free cokernels. This theoretical result informs our formulation of Algorithm 3.2, which takes, as input, a persistence module that is pointwise free and finitely-generated over a PID and outputs an interval decomposition (if one exist). Our work provides structural insights into persistence modules over non-field coefficients. It also provides insights into the computation of PH — namely, the conditions under which a PD is independent of field choice.

Following our theoretical contributions in PH, we studied an application of PH to resource coverage in geographical regions. A key advantage of using PH is that it allows us to consider resource coverage across all scales, eliminating the need to choose a fixed cutoff distance. We investigated polling-site access in six geographical regions (five cities and Los Angeles County) as a case study. For each region, we constructed a weighted Vietoris–Rips (VR) complex based on travel times to and from and waiting times at polling sites. By

computing PH, we quantitatively identified holes in coverage (i.e., regions that have poor resource access). Our work demonstrates the benefits of using PH to assess and quantify the accessibility of resources.

Finally, we formulated and analyzed new discrete-time bounded-confidence models (BCMs) of opinion dynamics with heterogeneous and adaptive confidence bounds. In our BCMs, individuals who have positive interactions are more receptive to each other in future interactions, and individuals who have negative interactions are less receptive to each other in future interactions. We extended the traditional Hegselmann–Krause (HK) and Deffuant–Weisbuch (DW) BCMs by incorporating parameters that control the increase and decrease of confidence bounds. We analytically explored the properties and behavior of our BCMs, including the confidence-bound dynamics, the formulation of opinion clusters, and the time-evolution of an effective graph. For a variety of networks and a wide range of values of our BCM parameters, we numerically demonstrated that our adaptive BCMs result in fewer major opinion clusters and longer convergence times than the baseline HK and DW BCMs.

In conclusion, we studied PH and complex social systems, both individually and together. We established new insights into the theoretical foundations of PH, introduced an application of PH to study complex social systems that have a geospatial component, and studied opinion dynamics by formulating new BCMs. There are many opportunities to advance PH and complex social systems — both individually and in tandem — through theoretical and computational exploration.

REFERENCES

[AC15] Meysam Alizadeh and Claudio Cioffi-Revilla. "Activation regimes in opinion dynamics: Comparing asynchronous updating schemes." *J. Artif. Soc. Soc. Simul.*, **18**(3): 8, 2015.

[ACG19] Hirokazu Anai, Frédéric Chazal, Marc Glisse, Yuichi Ike, Hiroya Inakoshi, Raphaël Tinarrage, and Yuhei Umeda. "DTM-based filtrations." In Gill Barequet and Yusu Wang, editors, *35th International Symposium on Computational Geometry (SoCG 2019)*, volume 129 of *Leibniz International Proceedings in Informatics (LIPIcs)*, 58:1–58:15, Dagstuhl, Germany, 2019. Schloss Dagstuhl—Leibniz-Zentrum fuer Informatik.

[AS11] Henry Adams and Jan Segert. "Simplicial complex filtration demonstrations in Mathematica.", 2011. Available at `https://www.math.colostate.edu/~adams/research/` (accessed 2 May 2022).

[AWA22] Johnathan A. Adams, Gentry White, and Robyn P. Araujo. "Mathematical measures of societal polarisation." *PLoS ONE*, **17**(10): e0275283, 2022.

[BAB21] 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, Albert B. Kao, Rachel E. Moran, Pawel Romanczuk, Daniel I. Rubenstein, Kaia J. Tombak, Jay J. Van Bavel, and Elke U. Weber. "Stewardship of global collective behavior." *Proc. Natl. Acad. Sci. USA*, **118**(27): e2025764118, 2021.

[BAP24] Carmela Bernardo, Claudio Altafini, Anton Proskurnikov, and Francesco Vasca. "Bounded confidence opinion dynamics: A survey." *Automatica*, **159**: 111302, 2024.

[BB04] Lars Brabyn and Ross Barnett. "Population need and geographical access to general practitioners in rural New Zealand." *New Zealand Med. J.*, **117**(1199): U996, 2004.

[BB06] Lars Brabyn and Paul Beere. "Population access to hospital emergency departments and the impacts of health reform in New Zealand." *Health Inform. J.*, **12**(3): 227–237, 2006.

[BBC13] Arnab Bhattacharyya, Mark Braverman, Bernard Chazelle, and Huy L. Nguyen. "On the convergence of the Hegselmann-Krause system." In *Proceedings of the 4th Conference on Innovations in Theoretical Computer Science*, ITCS '13, 61–66, New York, NY, USA, 2013. Association for Computing Machinery.

[BBG18] Hugo Barbosa, Marc Barthelemy, Gourab Ghoshal, Charlotte R. James, Maxime Lenormand, Thomas Louail, Ronaldo Menezes, José J. Ramasco, Filippo Simini, and Marcello Tomasini. "Human mobility: Models and applications." *Physics Rep.*, **734**: 1–74, 2018.

[BBH06] Raymond C. Browning, Emily A. Baker, Jessica A. Herron, and Rodger Kram. "Effects of obesity and sex on the energetic cost and preferred speed of walking." *J. Appl. Physiol.*, **100**(2): 390–398, 2006.

[BCF07] Franco Bagnoli, Timoteo Carletti, Duccio Fanelli, Alessio Guarino, and Andrea Guazzini. "Dynamical affinity in opinion dynamics modeling." *Phys. Rev. E*, **76**: 066105, 2007.

[BGK23] Rico Berner, Thilo Gross, Christian Kuehn, Jürgen Kurths, and Serhiy Yanchuk. "Adaptive dynamical networks." *Phys. Rep.*, **1031**: 1–59, 2023. Adaptive dynamical networks.

[BGS16] Aaron Bramson, Patrick Grim, Daniel J. Singer, Steven Fisher, William Berger, Graham Sack, and Carissa Flocken. "Disambiguation of social polarization concepts and measures." *J. Math. Sociology*, **40**(2): 80–111, 2016.

[Bir95] Garrett Birkhoff. *Lattice Theory.* Number 25 in Colloquium Publications. American Mathematical Society, Providence, RI, USA, 3rd edition, 1995.

[BKR03] E. Ben-Naim, P. L. Krapivsky, and S. Redner. "Bifurcations and patterns in compromise processes." *Phys. D*, **183**(3): 190–204, 2003.

[BL23] Magnus Bakke Botnan and Michael Lesnick. "An introduction to multiparameter persistence." In Aslak Bakke Buan, Henning Krause, and Øyvind Solberg, editors, *EMS Series of Congress Reports*, volume 19, 77–150. EMS Press, 2023.

[BL24] Andrew J. Blumberg and Michael Lesnick. "Stability of 2-parameter persistent homology." *Found. Comput. Math.*, **24**: 385–427, 2024.

[Bor48] Karol Borsuk. "On the imbedding of systems of compacta in simplicial complexes." *Fund. Math.*, **35**(11): 217–234, 1948.

[BVI22] Carmela Bernardo, Francesco Vasca, and Raffaele Iervolino. "Heterogeneous opinion dynamics with confidence thresholds adaptation." *IEEE Trans. Control Netw. Syst.*, **9**(3): 1068–1079, 2022.

[BYM22] Agnese Barbensi, Hee Rhang Yoon, Christian Degnbol Madsen, Deborah O. Ajayi, Michael P. H. Stumpf, and Heather A. Harrington. "Hypergraphs for multiscale cycles in structured data." Preprint arXiv:2210.07545, 2022.

[CDH22] Yu-Min Chung, Sarah Day, and Chuan-Shen Hu. "A multi-parameter persistence framework for mathematical morphology." *Sci. Rep.*, **12**(1): 6427, 2022.

[CHM22] Justin Curry, Haibin Hang, Washington Mio, Tom Needham, and Osman Berat Okutan. "Decorated merge trees for persistent homology." *J. App. Comp. Topol.*, **6**: 371–428, 2022.

[CHP19] M. Keith Chen, Kareem Haggag, Devin G. Pope, and Ryne Rohla. "Racial disparities in voting wait times: Evidence from smartphone data." Working Paper 26487, National Bureau of Economic Research, 2019. Available at SSRN: `https://ssrn.com/abstract=3492890`.

[CJ23] Padraig Corcoran and Christopher B. Jones. "Topological data analysis for geographical information science using persistent homology." *Int. J. Geogr. Inf. Sci.*, **37**(3): 1–34, 2023.

[CLP24] Weiqi Chu, Qin Li, and Mason A. Porter. "Inference of interaction kernels in mean-field models of opinion dynamics." *SIAM Appl. Math.*, **84**(3): 1096–1115, 2024.

[CLY20] Chun Cheng, Yun Luo, and Changbin Yu. "Consensus for expressed and private opinions under self-persuasion." *IFAC-PapersOnLine*, **53**(2): 2483–2488, 2020. 21st IFAC World Congress.

[CM11] Daniel Chandler and Rod Munday. *A Dictionary of Media and Communication.* Oxford University Press, Oxford, UK, 2011.

[CS10] Gunnar Carlsson and Vin de Silva. "Zigzag persistence." *Found. Comput. Math.*, **10**(4): 367–405, 2010.

[CS20] Ginny Seung Choi and Virgil Henry Storr. "Market interactions, trust and reciprocity." *PLoS ONE*, **15**(5): e0232704, 2020.

[CSD20] Ge Chen, Wei Su, Songyuan Ding, and Yiguang Hong. "Heterogeneous Hegselmann–Krause dynamics with environment and communication noise." *IEEE Trans. Automat. Control*, **65**(8): 3409–3424, 2020.

[CSM20] Ge Chen, Wei Su, Wenjun Mei, and Francesco Bullo. "Convergence properties of the heterogeneous Deffuant–Weisbuch model." *Automatica*, **114**: 108825, 2020.

[Cur17] Carina Curto. "What can topology tell us about the neural code?" *Bull. Amer. Math. Soc.*, **54**(1): 63–78, 2017.

[CZ09] Gunnar Carlsson and Afra Zomorodian. "The theory of multidimensional persistence." *Discrete Comput. Geom.*, **42**(1): 71–93, 2009.

[CZ15] Andrés Chacoma and Damián H. Zanette. "Opinion formation by social influence: From experiments to modeling." *PLoS ONE*, **10**(10): e0140406, 2015.

[DAW02] Guillaume Deffuant, Frédéric Amblard, Gérard Weisbuch, and Thierry Faure. "How can extremism prevail? A study based on the relative agreement interaction model." *J. Artif. Soc. Soc. Simul.*, **5**(4): 1, 2002.

[DCS18] Luis Diaz-Garcia, Giovanny Covarrubias-Pazaran, Brandon Schlautman, Edward Grygleski, and Juan Zalapa. "Image-based phenotyping for identification of QTL determining fruit shape and size in American cranberry (*Vaccinium macrocarpon* L.)." *PeerJ*, **6**: e5461, 2018.

[DF04] David Steven Dummit and Richard M. Foote. *Abstract Algebra*. Wiley, Hoboken, NJ, 3rd edition, 2004.

[DNA00] Guillaume Deffuant, David Neau, Frederic Amblard, and Gérard Weisbuch. "Mixing beliefs among interacting agents." *Adv. Complex Syst.*, **03**(01n04): 87–98, 2000.

[DSC17] Michela Del Vicario, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, and Walter Quattrociocchi. "Modeling confirmation bias and polarization." *Sci. Rep.*, **7**(1): 40391, 2017.

[DV09] Vin De Silva and Mikael Vejdemo-Johansson. "Persistent cohomology and circular coordinates." In *Proceedings of the twenty-fifth annual symposium on Computational geometry*, 227–236, Aarhus Denmark, 2009. Association for Computing Machinery.

[DW22] Tamal K. Dey and Yusu Wang. *Computational Topology for Data Analysis*. Cambridge University Press, Cambridge, UK, New edition, 2022.

[EH08] Herbert Edelsbrunner and John Harer. "Persistent homology — A survey." *Surveys on Discrete and Computational Geometry*, **453**: 257–282, 2008.

[ELZ02] Herbert Edelsbrunner, D. Letscher, and A. Zomorodian. "Topological persistence and simplification." *Discrete Comput. Geom.*, **28**(4): 511–533, 2002.

[Ete19] S. Rasoul Etesami. "A simple framework for stability analysis of state-dependent networks of heterogeneous agents." *SIAM J. Control Optim.*, **57**(3): 1757–1782, 2019.

[FHP22] Michelle Feng, Abigail Hickok, and Mason A. Porter. "Topological data analysis of spatial systems." In Federico Battiston and Giovanni Petri, editors, *Higher-Order Systems*, Understanding Complex Systems, chapter 16, 389–399. Springer

International Publishing, Cham, Switzerland, 2022.

[FJ90] Noah E. Friedkin and Eugene C. Johnsen. "Social influence and opinions." *J. Math. Sociology*, **15**(3–4): 193–206, 1990.

[For04] Santo Fortunato. "Universality of the threshold for complete consensus for the opinion dynamics of Deffuant et al." *Int. J. Mod. Phys. C*, **15**(09): 1301–1307, 2004.

[For05] Santo Fortunato. "On the consensus threshold for the opinion dyanamics of Krause–Hegselmann." *Internat. J. Modern Phys. C*, **16**(2): 259–270, 2005.

[Fow20] Stephan Fowler. "Why do nonwhite Georgia voters have to wait in line For hours? Too few polling places." NPR, 2020. Available at `https://www.npr.org/2020/10/17/924527679/why-do-nonwhite-georgia-voters-have-to-wait-in-line-for-hours-too-few-polling-pl`.

[FP20] Michelle Feng and Mason A. Porter. "Spatial applications of topological data analysis: Cities, snowflakes, random structures, and spiders spinning under the influence." *Phys. Rev. Res.*, **2**: 033426, 2020.

[FP21] Michelle Feng and Mason A. Porter. "Persistent homology of geospatial data: A case study with voting." *SIAM Rev.*, **63**(1): 67–99, 2021.

[Gab72] Peter Gabriel. "Unzerlegbare darstellungen I." *Manuscripta Math.*, **6**: 71–103, 1972.

[GAP13] Jennifer L. Glanville, Matthew A. Andersson, and Pamela Paxton. "Do social connections create trust? An examination using new longitudinal data." *Social Forces*, **92**(2): 545–562, 2013.

[GH21] Robert Ghrist and Gregory Henselman-Petrusek. "Saecular persistence.", 2021. Preprint arXiv:2112.04927.

[Ghr07] Robert Ghrist. "Barcodes: The persistent topology of data." *Bull. Amer. Math. Soc.*, **45**(1): 61–76, 2007.

[GM23] Aziz Burak Gulen and Alexander McCleary. "Galois connections in persistent homology.", 2023. Preprint arXiv:2201.06650.

[GOD21] Mirta Galesic, Henrik Olsson, Jonas Dalege, Tamara van der Does, and Daniel L. Stein. "Integrating social and cognitive aspects of belief dynamics: Towards a unifying framework." *J. R. Soc. Interface*, **18**: 20200857, 2021.

[Goo] Google Developers. "Distance Matrix API." Available at `https://`

developers.google.com/maps/documentation/distance-matrix (accessed 4–7 November 2021).

[Gra12] Marco Grandis. *Homological Algebra: The Interplay of Homology with Distributive Lattices and Orthodox Semigroups*. World Scientific, Singapore, 2012.

[GS03] James Gimpel and Jason Schuknecht. "Political participation and accessibility of the ballot box." *Political Geogr.*, **22**: 471–488, 2003.

[Hat02] Allen Hatcher. *Algebraic Topology*. Cambridge University Press, Cambridge, UK, 2002.

[HDJ08] Sylvie Huet, Guillaume Deffuant, and Wander Jager. "A rejection mechanism in 2D bounded confidence provides more conformity." *Adv. Complex Syst.*, **11**(04): 529–549, 2008.

[HFQ20] Wenchen Han, Yuee Feng, Xiaolan Qian, Qihui Yang, and Changwei Huang. "Clusters and the entropy in opinion dynamics on complex networks." *Phys. A*, **559**: 125033, 2020.

[HJJ23] 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.", 2023. Preprint arXiv:2206.04834 (SIAM Rev., in press).

[HK02] Rainer Hegselmann and Ulrich Krause. "Opinion dynamics and bounded confidence models, analysis and simulation." *J. Artif. Soc. Soc. Simul.*, **5**(3): 2, 2002.

[HK05] Moshe Haspel and H. Gibbs Knotts. "Location, location, location: Precinct placement and the costs of voting." *J. Politics*, **67**(2): 560–573, 2005.

[HK12] Timothy L Hawthorne and Mei-Po Kwan. "Using GIS and perceived distance to understand the unequal geographies of healthcare in lower-income urban neighbourhoods." *Geogr. J.*, **178**(1): 18–30, 2012.

[HL15] Petter Holme and Fredrik Liljeros. "Mechanistic models in computational social science." *Front. Phys.*, **3**: 78, 2015.

[HLJ21] Jian Hou, Wenshan Li, and Mingyue Jiang. "Opinion dynamics in modified expressed and private model with bounded confidence." *Phys. A*, **574**: 125968, 2021.

[HNP22] Abigail Hickok, Deanna Needell, and Mason A. Porter. "Analysis of spatial and spatiotemporal anomalies using persistent homology: Case studies with COVID-19 data." *SIAM J. Math. Data Sci.*, **4**(3): 1116–1144, 2022.

[Hor62] Irving Louis Horowitz. "Consensus, conflict and cooperation: A sociological inventory." *Social Forces*, **41**(2): 177–188, 1962.

[Jac08] Matthew O. Jackson. *Social and Economic Networks*. Princeton University Press, Princeton, NJ, USA, 2008.

[JMF15] Peng Jia, Anahita MirTabatabaei, Noah E. Friedkin, and Francesco Bullo. "Opinion dynamics and the evolution of social power in influence networks." *SIAM Rev.*, **57**(3): 367–397, 2015.

[Kan19] Henry Kanengiser. "In New York, where you live can determine how hard it is to vote." Available at `https://citylimits.org/2019/04/25/nyc-polling-place-shortage-inequality/`, 2019.

[KB08a] Balazs Kozma and Alain Barrat. "Consensus formation on adaptive networks." *Phys. Rev. E*, **77**: 016102, 2008.

[KB08b] Balazs Kozma and Alain Barrat. "Consensus formation on coevolving networks: Groups' formation and structure." *J. Phys. A Math. Theor.*, **41**(22): 224020, 2008.

[KF23] Claudia Kann and Michelle Feng. "Repulsive bounded-confidence model of opinion dynamics in polarized communities." Preprint arXiv:2301.02210, 2023.

[KFP23] Unchitta Kan, Michelle Feng, and Mason A. Porter. "An adaptive bounded-confidence model of opinion dynamics on networks." *J. Complex Net.*, **11**(1): cnac055, 2023.

[Koz22a] Ivan V. Kozitsin. "Formal models of opinion formation and their application to real data: Evidence from online social networks." *J. Math. Sociology*, **46**(2): 120–147, 2022.

[Koz22b] Ivan V. Kozitsin. "Linking theory and empirics: A general framework to model opinion formation processes." *Sci. Rep.*, **12**: 5543, 2022.

[Koz23] Ivan V. Kozitsin. "Opinion dynamics of online social network users: A micro-level analysis." *J. Math. Sociology*, **47**(1): 1–41, 2023.

[KP20] Yacoub H. Kureh and Mason A. Porter. "Fitting in and breaking up: A nonlinear version of coevolving voter models." *Phys. Rev. E*, **101**: 062303, 2020.

[Kra00] Ulrich Krause. "A discrete nonlinear and non-autonomous model of consensus." In Saber N. Elaydi, Jerry Popenda, and Jerry Rakowski, editors, *Communications in Difference Equations: Proceedings of the Fourth International Conference on Difference Equations*, 227–236. CRC Press, Amsterdam, The Netherlands, 2000.

[Kur95] Timur Kuran. *Private Truths, Public Lies*. Harvard University Press, Cambridge, MA, USA, 1995.

[LAZ04] M. F. Laguna, Guillermo Abramson, and Damián H. Zanette. "Minorities in a model for opinion formation." *Complexity*, **9**(4): 31–36, 2004.

[LH23] Jiajie Luo and Gregory Henselman-Petrusek. "Interval decomposition of persistence modules over a principal ideal domain.", 2023. Preprint arXiv:2310.07971.

[LLP23] Grace J. Li, Jiajie Luo, and Mason A. Porter. "Bounded-confidence models of opinion dynamics with adaptive confidence bounds.", 2023. Preprint arXiv:2303.07563.

[LMM17] Diandian Li, Liang Meng, and Qingguo Ma. "Who deserves my trust? Cue-elicited feedback negativity tracks reputation learning in repeated social interactions." *Front. Hum. Neurosci.*, **11**: 307, 2017.

[Lor05] Jan Lorenz. "A stabilization theorem for dynamics of continuous opinions." *Phys. A*, **355**(1): 217–223, 2005.

[Lor06] Jan Lorenz. "Consensus strikes back in the Hegselmann–Krause model of continuous opinion dynamics under bounded confidence." *J. Artif. Soc. Soc. Simul.*, **9(1)**(1): 8, 2006.

[Lor08] Jan Lorenz. "Continuous opinion dynamics under bounded confidence: A survey." *Int. J. Mod. Phys. C*, **18**(12): 1819–1838, 2008.

[Lor09] Jan Lorenz. "Heterogeneous bounds of confidence: Meet, discuss and find consensus!" *Complexity*, : 43–52, 2009.

[LP23] Grace J. Li and Mason A. Porter. "Bounded-confidence model of opinion dynamics with heterogeneous node-activity levels." *Phys. Rev. Res.*, **5**: 023179, 2023.

[LTH21] Lu Li, Connor Thompson, Gregory Henselman-Petrusek, Chad Giusti, and Lori Ziegelmeier. "Minimal cycle representatives in persistent homology using linear programming: An empirical study with user's guide." *Front. Artif. Intell.*, **4**: 681117, 2021.

[LW15] Michael Lesnick and Matthew Wright. "Interactive visualization of 2-d persistence modules." Preprint arXiv:1512.00180v1, 2015.

[LWF03] Xiang-Yang Li, Peng-Jun Wan, and O. Frieder. "Coverage in wireless ad hoc sensor networks." **52**(6): 753–763, 2003.

[Mas19] Michael Mäs. "Challenges to simulation calidation in the social sciences. A critical

rationalist perspective." In Claus Beisbart and Nicole J. Saam, editors, *Computer Simulation Validation: Fundamental Concepts, Methodological Frameworks, and Philosophical Perspectives*, 857–879. Springer International Publishing, Cham, Switzerland, 2019.

[MDB20] Corrado Monti, Gianmarco De Francisci Morales, and Francesco Bonchi. "Learning opinion dynamics from social traces." In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '20, 764–773, New York, NY, USA, 2020. Association for Computing Machinery.

[MKP01] S. Meguerdichian, F. Koushanfar, M. Potkonjak, and M. B. Srivastava. "Coverage problems in wireless ad-hoc sensor networks." In *Proceedings IEEE INFOCOM 2001. Conference on Computer Communications. Twentieth Annual Joint Conference of the IEEE Computer and Communications Society (Cat. No.01CH37213)*, volume 3, 1380–1387, 2001.

[MRU22] Christopher Musco, Indu Ramesh, Johan Ugander, and R. Teal Witter. "How to quantify polarization in models of opinion dynamics." In *Proceedings of the 17th International Workshop on Mining and Learning with Graphs (MLG)*, 2022.

[MVP18] X. Flora Meng, Robert A. Van Gorder, and Mason A. Porter. "Opinion formation and distribution in a bounded-confidence model on various networks." *Phys. Rev. E*, **97**: 022312, 2018.

[New06] M. E. J. Newman. "Finding community structure in networks using the eigenvectors of matrices." *Phys. Rev. E*, **74**: 036104, 2006.

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

[NGW23] Andrew Nugent, Susana N. Gomes, and Marie-Therese Wolfram. "On evolving network models and their influence on opinion formation." *Phys. D*, **456**: 133914, 2023.

[Noo20] Hossein Noorazar. "Recent advances in opinion propagation dynamics: A 2020 survey." *Eur. Phys. J. Plus*, **135**(6): 521, 2020.

[NVT20] Hossein Noorazar, Kevin R. Vixie, Arghavan Talebanpour, and Yunfeng Hu. "From classical to modern opinion dynamics." *Int. J. Mod. Phys. C*, **31**(7): 2050101, 2020.

[OF03] Stanley J. Osher and Ronald Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*. Springer-Verlag, Heidelberg, Germany, 2003.

[OKT19] Yohei Onodera, Shinji Kohara, Shuta Tahara, Atsunobu Masuno, Hiroyuki Inoue, Motoki Shiga, Akihiko Hirata, Koichi Tsuchiya, Yasuaki Hiraoka, Ippei Obayashi, Koji Ohara, Akitoshi Mizuno, and Osami Sakata. "Understanding diffraction patterns of glassy, liquid and amorphous materials via persistent homology analyses." *J. Ceram. Soc. Jpn.*, **127**(12): 853–863, 2019.

[Ope21] OpenStreetMap contributors. "OpenStreetMap.", 2021. Version 1.1.1 (accessed 19–22 January 2022).

[OPT17] N. Otter, M. A. Porter, U. Tillmann, P. Grindrod, and H. A. Harrington. "A roadmap for the computation of persistent homology." *EPJ — Data Sci.*, **6**: 17, 2017.

[Oud15] Steve Y. Oudot. *Persistence Theory: From Quiver Representations to Data Analysis*, volume 209 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, USA, 2015.

[OY20] Ippei Obayashi and Michio Yoshiwaki. "Field choice problem in persistent homology.", 2020. Preprint arXiv:1911.11350v3.

[OY23] Ippei Obayashi and Michio Yoshiwaki. "Field choice problem in persistent homology." *Discrete Comput. Geom.*, **70**(3): 645–670, 2023.

[Pat18] Amit Patel. "Generalized persistence diagrams." *J. Appl. Comput. Topol.*, **1**(3): 397–419, 2018.

[PC14] Jose A. Perea and Gunnar Carlsson. "A Klein-bottle-based dictionary for texture representation." *Int. J. Comput. Vis.*, **107**(1): 75–97, 2014.

[Per18] Jose A. Perea. "Multiscale projective coordinates via persistent cohomology of sparse filtrations." *Discrete Comput. Geom.*, **59**(1): 175–225, 2018.

[Per20] Jose A. Perea. "Sparse circular coordinates via principal $\mathbb{Z}$-bundles." In Nils A. Baas, Gunnar E. Carlsson, Gereon Quick, Markus Szymik, and Marius Thaule, editors, *Topological Data Analysis*, Abel Symposia, 435–458, Cham, Switzerland, 2020. Springer International Publishing.

[PFT18] Rohit Parasnis, Massimo Franceschetti, and Behrouz Touri. "Hegselmann–Krause dynamics with limited connectivity." In *2018 IEEE Conference on Decision and Control (CDC)*, 5364–5369, 2018. ISSN: 2576-2370.

[PRM22] Valentina Pansanella, Giulio Rossetti, and Letizia Milli. "Modeling algorithmic bias: Simplicial complexes and evolving network topologies." *Appl. Netw. Sci.*, **7**(1): 57, 2022.

[PS23] Amit Patel and Primoz Skraba. "Möbius homology.", 2023. Preprint arXiv:2307.01040.

[PWB06] Jamie Pearce, Karen Witten, and Phil Bartie. "Neighbourhoods and health: A GIS approach to measuring community resource accessibility." *J. Epidemiol. Community Health*, **60**(5): 389–395, 2006.

[RKK07] Sini Ruohomaa, Lea Kutvonen, and Eleni Koutrouli. "Reputation management survey." In *The Second International Conference on Availability, Reliability and Security (ARES'07)*, 103–111, Vienna, Austria, 2007.

[RKM11] Veronica Red, Eric D. Kelsic, Peter J. Mucha, and Mason A. Porter. "Comparing community structure to characteristics in online collegiate social networks." *SIAM Rev.*, **53**(3): 526–543, 2011.

[RKZ00] Paul Resnick, Ko Kuwabara, Richard Zeckhauser, and Eric Friedman. "Reputation systems." *Commun. ACM*, **43**(12): 45–48, 2000.

[Rob00] Vanessa Robins. *Computational Topology at Multiple Resolutions: Foundations and Applications to Fractals and Dynamics*. Ph.D. Thesis, University of Colorado at Boulder, USA, 2000.

[SBL23] Jakub Sawicki, Rico Berner, Sarah A. M. Loos, Mehrnaz Anvari, Rolf Bader, Wolfram Barfuss, Nicola Botta, Nuria Brede, Igor Franović, Daniel J. Gauthier, Sebastian Goldt, Aida Hajizadeh, Philipp Hövel, Omer Karin, Philipp Lorenz-Spreen, Christoph Miehl, Jan Mölter, Simona Olmi, Eckehard Schöll, Alireza Seif, Peter A. Tass, Giovanni Volpe, Serhiy Yanchuk, and Jürgen Kurths. "Perspectives on adaptive dynamical systems." *Chaos*, **33**(7): 071501, 2023.

[Sch21] Rachel Schnalzer. "Traffic is terrible again. Here's how to get it closer to spring 2020 levels." LA Times, 2021. Available at `https://www.latimes.com/business/story/2021-07-22/los-angeles-traffic-congestion-commute-pandemic`.

[Sch22] Hal Schenck. *Algebraic Foundations for Applied Topology and Data Analysis*. Springer International Publishing, Cham, Switzerland, 2022.

[SEN21] Bernadette J. Stolz, Tegan Emerson, Satu Nahkuri, Mason A Porter, and Heather A Harrington. "Topological data analysis of task-based fMRI data from experiments on schizophrenia." *J. Phys. Complex.*, **2**(3): 035006, 2021.

[SFH21] Hendrik Schawe, Sylvain Fontaine, and Laura Hernández. "When network bridges foster consensus. Bounded confidence models in networked societies." *Phys. Rev. Res.*, **3**: 023208, 2021.

[SG06] Vin de Silva and Robert Ghrist. "Coordinate-free coverage in sensor networks with controlled boundaries via homology." *Int. J. Robot. Res.*, **25**(12): 1205–1222, 2006.

[SG07] Vin de Silva and Robert Ghrist. "Coverage in sensor networks via persistent homology." *Alg. Geom. Topology*, **7**(1): 339–358, 2007.

[SGB23] Luis Scoccola, Hitesh Gakhar, Johnathan Bush, Nikolas Schonsheck, Tatum Rask, Ling Zhou, and Jose A. Perea. "Toroidal coordinates: Decorrelating circular coordinates with lattice reduction." In Erin W. Chambers and Joachim Gudmundsson, editors, *39th International Symposium on Computational Geometry (SoCG 2023)*, volume 258 of *Leibniz International Proceedings in Informatics (LIPIcs)*, 57:1–57:20, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[SGK18] Ann E. Sizemore, Chad Giusti, Ari Kahn, Jean M. Vettel, Richard F. Betzel, and Danielle S. Bassett. "Cliques and cavities in the human connectome." *J. Comput. Neurosci.*, **44**(1): 115–145, 2018.

[SGW17] Wei Su, YaJuan Gu, Sha Wang, and YongGuang Yu. "Partial convergence of heterogeneous Hegselmann–Krause opinion dynamics." *Sci. China Technol. Sci.*, **60**(9): 1433–1438, 2017.

[Sha14] Yilun Shang. "An agent based model for opinion dynamics with random confidence threshold." *Commun. Nonlin. Sci. Numer. Simul.*, **19**(10): 3766–3777, 2014.

[SHP16] Bernadette J. Stolz, Heather A. Harrington, and Mason A. Porter. "The topological "shape" of Brexit." *Preprint arXiv:1610.00752*, 2016.

[SNP13] Wanita Sherchan, Surya Nepal, and Cecile Paris. "A survey of trust in social networks." *ACM Comput. Surv.*, **45**(4): 47, 2013.

[SSH19] Simon Scarr, Manas Sharma, and Marco Hernandez. "Roads, boats and elephants: How India mobilised a million polling stations." Reuters, 2019. Available at `https://graphics.reuters.com/INDIA-ELECTION-STATIONS/010092FY33Z/index.html`.

[TFM16] Károly Takács, Andreas Flache, and Michael Mäs. "Discrepancy and disliking do not induce negative opinion shifts." *PLoS ONE*, **11**(6): e0157948, 2016.

[The21] The White House. "FACT SHEET: President Biden Announces 90% of the Adult U.S. Population will be Eligible for Vaccination and 90% will have a Vaccination Site Within 5 Miles of Home by April 19." Available at `https://www.whitehouse.gov/briefing-room/statements-releases/2021/`

03/29/fact-sheet-president-biden-announces-90-of-the-adult-u-s-population-will-be-eligible-for-vaccination-and-90-will-have-a-vaccination-site-within-5-miles-of-home-by-april-19/, 2021.

[TMP12] Amanda L. Traud, Peter J. Mucha, and Mason A. Porter. "Social structure of Facebook networks." *Phys. A*, **391**(16): 4165–4180, 2012.

[US] U.S. Census Bureau. "American Community Survey, 2015–2019 Estimates." Available at https://data.census.gov/ (accessed 7 November 2021).

[Vaz22] Federico Vazquez. "Modeling and analysis of social phenomena: Challenges and possible research directions." *Entropy*, **24**(4): 491, 2022.

[VBI21] Francesco Vasca, Carmela Bernardo, and Raffaele Iervolino. "Practical consensus in bounded confidence opinion dynamics." *Automatica*, **129**: 109683, 2021.

[VMG16] Corentin Vande Kerckhove, Samuel Martin, Pascal Gend, Peter J. Rentfrow, Julien M. Hendrickx, and Vincent D. Blondel. "Modelling influence and opinion evolution in online collective behaviour." *PLoS ONE*, **11**(6): e0157685, 2016.

[WDA02] Gérard Weisbuch, Guillaume Deffuant, Frédéric Amblard, and Jean-Pierre Nadal. "Meet, discuss, and segregate!" *Complexity*, **7**(3): 55–63, 2002.

[XLC17] Fei Xiong, Yun Liu, and Junjun Cheng. "Modeling and predicting opinion formation with trust propagation in online social networks." *Commun. Nonlinear Sci. Numer. Simul.*, **44**: 513–524, 2017.

[YDH14] Yuecheng Yang, Dimos V. Dimarogonas, and Xiaoming Hu. "Opinion consensus of modified Hegselmann–Krause models." *Automatica*, **50**(2): 622–627, 2014.

[ZH05] Honghai Zhang and Jennifer C. Hou. "Maintaining sensing coverage and connectivity in large sensor networks." *Ad Hoc & Sensor Wireless Netw.*, **1**: 89–124, 2005.