Title

Feedback Driven Dynamics in Socio-Algorithmic Systems

Author

Curmei, Mihaela

Publication Date

2024

Peer reviewed|Thesis/dissertation

Feedback Driven Dynamics in Socio-Algorithmic Systems

By

Mihaela Curmei

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering- Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Associate Professor Benjamin Recht, Chair
Assistant Professor Nika Haghtalab
Associate Professor Murat Arcak
Assistant Professor Yixin Wang

Summer 2024

Feedback Driven Dynamics in Socio-Algorithmic Systems

Abstract

Feedback Driven Dynamics in Socio-Algorithmic Systems

by

Mihaela Curmei

Doctor of Philosophy in Engineering- Electrical Engineering and Computer Sciences

University of California, Berkeley

Associate Professor Benjamin Recht, Chair

Algorithmic decision systems are an ubiquitous part of daily life, operating in dynamic environments where user interactions and feedback loops complicate predictability, reliability, and utility. This thesis explores the multifaceted dynamics between users and algorithmic decision systems, addressing both immediate impacts and long-term implications of their interactions. By examining single-user interactions and extending to broader social network and market dynamics, we aim to uncover trade-offs and limitations in these feedback settings. This comprehensive study provides insights into these interactions, emphasizing the importance of considering both individual and collective behaviors in system design and evaluation.

Part I focuses on user-recommender interactions, providing insights into how users and recommendation systems affect each other. We introduce a user-centric notion of agency in algorithmic recommendations, focusing on the feasible outcomes of one-step interactions between users and systems. By proposing an evaluation procedure based on stochastic reachability, we quantify the maximum probability of recommending a target piece of content to a user under allowable strategic modifications. This framework allows us to detect biases and systemic limitations in content discovery with minimal assumptions about user behavior. Transitioning from single-step interactions, we explore how recommendations influence users in multi-step closed-loop dynamics, requiring modeling of user behavior. We develop psychologically grounded dynamic preference models to capture classic psychological effects such as Mere Exposure, Operant Conditioning, and Hedonic Adaptation. Simulation-based studies show these models manifest distinct behaviors, informing system design and allowing for critical evaluations based on psychological plausibility.

Part II broadens the scope to consider the interactions of multiple users and multiple algorithmic decision-makers, examining complex social network and market dynamics. First, we investigate social networks where user interactions are mediated by link recommendation algorithms, examining the interplay of multiple users and relationships within evolving networks. Using an extended Jackson-Rogers model, we evaluate how link recommendations

influence network evolution over time, revealing delayed and indirect effects on network structures. Additionally, we examine markets with multiple algorithmic decision-makers, analyzing dynamics where users allocate their participation among services to minimize individual risk, while services update their models to reduce risk based on current user populations. Termed risk-reducing dynamics, this class includes common model updates such as gradient descent and multiplicative weights. Our findings indicate that repeated myopic updates with multiple learners result in market segmentation as the only stable outcome. We argue that specialization is an emergent property of competition, which alleviates typical concerns of representation disparity.

To my family

Contents

Acknowledgments

Completing a PhD is a journey filled with challenges and triumphs, and I would not have reached this point without the incredible support and mentorship I've received at Berkeley.

First and foremost, I am deeply grateful to my advisor, Professor Ben Recht. His encouragement to pursue diverse interests and his independent thinking have profoundly influenced my research. His mentorship has helped me find my voice and conviction in my work. Ben taught me to ensure something makes sense in the simplest settings before adding complexity. His approach of slow and thoughtful work, rather than rushing to follow trends, has shielded me from external pressures and allowed me to follow my curiosity.

I am also thankful to my committee members, Professors Nika Haghtalab, Murat Arcak, and Yixin Wang. Their insightful feedback has sharpened my thinking and focused my research. I admire Nika for her pure approach to research and her technical prowess. Her detailed knowledge of the intersection of Machine Learning and Economics and her enthusiasm for her work are truly inspiring. My interest in competition dynamics began in Murat's Population Games class, and I often revisit his class notes, which are a work of art in themselves. Yixin, a longtime mentor and collaborator, has shaped my understanding of causal inference and social network dynamics.

What I love most about Berkeley is its open and collaborative nature. I am grateful to Professors Mike Jordan and Jacob Steinhardt for welcoming me to their lab meetings and treating me like an honorary group member. Their meetings sparked many research ideas and projects, some of which I will discuss in this thesis. Through TA-ing their Data 102 course, I discovered my talent for teaching.

This thesis would not have been possible without the help of many fantastic collaborators. Much of my work stems from conversations and collaborations with Sarah Dean, who introduced me to the concept of reachability in recommendation systems and has been a role model in numerous ways. Her articulate research and technically beautiful projects are truly inspiring. I admire her work ethic and ability to manage everything, even staying up late with me to meet a deadline after a grueling day of job interviews. Another significant influence is Andy Haupt, a longtime collaborator who has challenged me to articulate the normative implications of my work. Together, we have written proofs, developed software, created marketing campaigns, ran surveys, and learned a lot about translating research to the real world.

I had the fortune to learn about privacy from Walid Krinchene and Li Zhang during a summer internship at Google Research. Their wealth of knowledge and attention to detail were invaluable. I am grateful to Karl Krauth, Sarah Dean, Wenshuo Guo, and Alex Wei for teaching me to write good research code and be systematic in answering questions. Working on household Covid transmission with Andrew Ilyas, Jacob Steinhardt, and Owain Evans taught me much about probabilistic modeling while addressing a timely issue. Mentoring undergraduate students Paris Zhang and Shangen Lu honed my mentoring skills, and I appreciate the work we did together. My friend Ana Stoica's work on recommendations in networks has also greatly influenced me.

I am fortunate to have collaborated with many others, including Lilian Ratliff, Jamie Morgenstern, Maryam Fazel, Daniel Jiang, Avinand Bose and Serena Wang, who provided insights into game theory, optimization, fairness, and dynamical systems. My research unexpectedly led me to build web applications, databases, and browser extensions. I am thankful to AI Forensics, particularly Francois Marie de Jovencel and Marc Faddoul, for their technical expertise and research vision. Collaborating with Deb Raji has shaped my thinking about the role of data quality in model performance and allowed me to gain hands-on experience with deep learning models. I want to express my gratitude to the entire Modest Yachts group, with its many iterations, and the broader FML community. I appreciate the stimulating conversations I've had over the years with Paula Gradu, Sara Fridovich Keil, Holly Jackson, Jessica Dai, Chris Harshaw, John Miller, Juanky Perdomo, Chloe Hsu, Tijana Zrnic, Esther Rolf, Meena Jagadeesan, Anastasios Angelopoulos and Frances Ding. I am also grateful to the amazing EECS staff, including Shirley, Ria, Naomi, Jean, and Kosta, for smoothing out many bureaucratic and technical bumps for me.

Looking back, my Berkeley journey would not have been possible without Professors Georgina Hall and Amir Ali Ahmadi. They inspired me to pursue graduate school and instilled in me a love for optimization. Our work on SOS optimization remains some of the most technically sound research I've done. I also thank Mike Hopcroft, my former manager at Microsoft, who encouraged me to pursue my PhD.

My friends and family have been the key to my happiness and sanity throughout grad school. I am grateful to Suzanne and Stace Wright for not only renting me a room for a year but also being my adoptive family during the Covid shutdowns. Moving into the HIP house, a 60-person coop, was one of the best decisions I made. I discovered a wonderful, diverse community and developed a passion for extravagant eight-course meals served at scale. I leave Berkeley with as much cooking knowledge as technical knowledge. I want to thank my friends Annie, Andrei, Muireann, Chenling, Neil, Hoai, Rudy, Brandon, Alma, Yev, Elena, Orr, Nam, Meesh, Adam, Storm, Kartik, Parneet, Miles, Luke and Jae for their support and camaraderie. A special thanks to my dear friend Nicoleta Copaci for our enduring friendship and being the best travel buddy.

Lastly, I am beyond grateful to my fiancé, Sanjit Singh Batra, for his unwavering support and love. Meeting you has been the biggest inflection point in my life and I look forward to our future together. I am also grateful to Jatin Batra, Jagjit Batra, Harleen, Rohit, and Jasnit for welcoming me with open arms in your family.

Most importantly, to my parents, Vladimir and Rodica Curmei whose boundless love and selfless sacrifice have been my greatest source of strength. *Va iubesc enorm!*

Chapter 1

Introduction

We live in a deeply interconnected world, where algorithmic decisions mediate many of our interactions. The news we consume, the jobs we apply for, and the connections we form are often curated and ranked by algorithms, transforming a deluge of options into a discernible set. This is incredibly useful, as navigating the vast scale of the web would be impossible without such algorithms. The tools that underpin the filtering, curation, and ranking of information are based on principles of statistical machine learning, focused on achieving good predictive performance on held-out validation sets. However, this approach misses a critical aspect: machine learning algorithms are not just predictive, but also interventional, as they impact the environment they were designed to predict. Similarly, people are active actors that influence and are influenced by algorithmic decisions.

These dynamics create challenges when designing and evaluating algorithmic systems. For instance, in news recommendations, people's preferences and opinions are influenced by the news they consume. In this context, a recommender system might achieve high levels of engagement not by accurately predicting preferences but by shaping them to be more predictable. As a result recommendation systems have been implicated in propagating biases and pernicious feedback loops, leading to opinion polarization, extremism, bias and other negative outcomes [Rib+20; FCF20; Jia+19; CSE18].

The goal of this thesis is to move beyond the static paradigm to better understand the evolving and interconnected nature of algorithmic decision-making. We do so by expanding along both the axis of time and the complexity of stakeholder interactions (see Figure 1.1). In Part I of this thesis, we discuss the dynamics of interactions and impacts between users and algorithmic systems, using recommendation systems as a motivating example. In Chapter 2 we frame the question of user impact in terms of reachability, a concept from Control Theory that characterizes the state spaces a system can evolve into, under a set of allowable actions. This concept encapsulates a user's ability to steer a system, providing a tool to audit content accessibility and discovery, and to differentiate between unavoidable system biases and those due to intrinsic user choices. Further complicating this is the notion that user preferences are constant and that recommendation algorithms simply discover these preferences. Increasingly, there is interest in understanding how recommendation systems

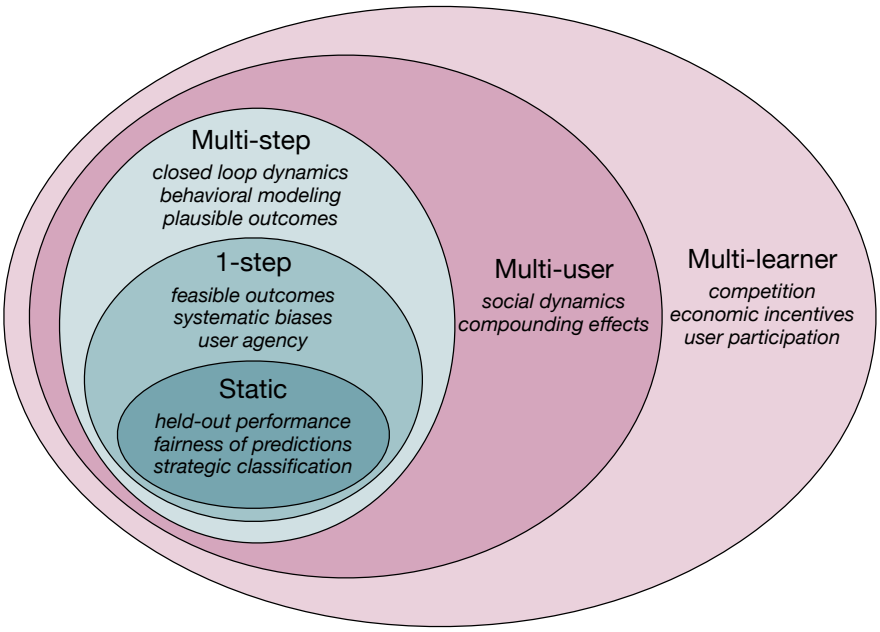


Figure 1.1: Diagram illustrating the expanded scope of this thesis along two axes: time and stakeholder interactions.

shape users [Ado+13; Car+21; Car+23]. This involves modeling user behavior in closed-loop systems, moving from feasibility—what universes are possible—to probability, which requires specific behavioral modeling. In Chapter 3, we explore ways to build behavioral models grounded in psychological evidence and use simulations to test their plausibility.

Having examined the interaction between a single learner and a single user over one time step in Chapter 2 and in closed loop over multiple steps in Chapter 3, in Part II we broaden the scope to consider dynamics embedded in larger social contexts such as networks and markets. First, we consider how algorithmic decision-making operates within existing social networks, which have their own dynamics. In Chapter 4, we investigate how link recommendations impact social networks, highlighting the complex interaction between natural and algorithmic dynamics. We demonstrate the evaluation challenges posed by these dynamics, illustrating delayed and indirect effects of recommendations on network structures. Finally, in the broadest lens, we consider feedback-driven dynamics at the ecosystem level, where besides multiple users, there are also multiple self-interested learners. In Chapter 5, we analyze a broad class of decision dynamics, where participants incrementally improve based on available information. We characterize these dynamics and show that market segmentation is the only stable equilibrium.

In the remainder of this chapter, we introduce and motivate the settings for the chapters to follow.

1.1 How do users and recommendation systems impact each other?

Despite being the most ubiquitous instance of algorithmic feedback loops in practice, the development of recommendation systems has historically focused on improving accuracy metrics to optimize user experiences. However, mere predictive accuracy is insufficient to achieve optimal performance both for individual users and the broader ecosystem. This realization has led to the exploration of additional evaluation metrics such as novelty, serendipity, diversity, and coverage [Her+04; CVW11]. While introducing and measuring these alternative metrics is motivated by the desire to understand the impact of algorithmic systems, they have largely centered on observational biases, examining the outcomes of single rounds of recommendations without considering user behaviors. This perspective overlooks the interactive nature of recommendation systems and their long-term impacts. Addressing these gaps, we propose a shift towards understanding recommendation systems as interventional tools that influence and are influenced by user behaviors over time.

The canonical model we use in this part of the thesis represents both users and items as vectors in a shared embedding space. Let user u at time t have embedding $\hat{\mathbf{p}}_u^t$, and item i be represented by an item embedding $\hat{\mathbf{v}}_i^t$, both vectors in $\mathbb{R}^d$. This assumption creates a similarity structure that is the basis for scoring unseen user-item pairs. These representations are part of the algorithmic modeling and are chosen to perform well on held-out datasets. Once a good model is selected, the resulting scores are used for ranking and curating recommendations. On top of the scores, there is typically an item selection policy that aims to balance exploitation of known and liked content with exploration of new and promising alternatives.

As recommendations (i.e. interventions) are made and new data is collected, recommendation systems typically incorporate new data points and update the representations of users and items to better reflect user preferences. When a person submits a new rating, that person's data alone might not significantly impact the item representation. For example, if a single user rates "Mad Max 1" as 5/5 and "Mad Max 2" as 1/5, it might have an insignificant impact on item representation, but it would significantly impact how the recommender represents the user, thus affecting future recommendations. Given this, our analysis of the dynamics between a user and a recommendation system assumes that item representations are fixed over time and that the dynamics are due to changes in how users are represented. We generally assume that $\hat{\mathbf{p}}_u^t = f(\hat{\mathbf{p}}_u^{t-1}, \texttt{rec}_t, a_t)$, where the representation of the user is a function of the prior representation, the recommendation decision, and the decision of the user.

In Chapter 2, we begin by studying one round of interaction between a user and a recommendation system (e.g., the recommender recommends five items, and the user rates them). In this chapter we aim to answer which states in the preference space are reachable by the user: i.e., $\hat{\mathbf{p}}$ such that $\exists a \in \mathcal{A} : f(\hat{\mathbf{p}}_u^{t-1}, \mathbf{rec}, a) = \hat{\mathbf{p}}$. This perspective allows us to discuss feasible outcomes of user-recommendation system dynamics, measure user agency and systematic biases without explicitly modeling user behaviors. While this framing resembles

the well-studied setting of strategic classification [Har+16], it differs in two important ways. First, strategic classification is concerned primarily with the sensitivity of static classifiers to adversarial modification of user features at test time, whereas reachability considers the sensitivity of the model update process, i.e., the model dynamics, with respect to strategic actions. The other difference is normative perspective on algorithmic sensitivity; in our case, it is seen as positive as it captures user agency, while it is viewed negatively in strategic classification as it captures the brittleness and manipulability of the algorithm.

In Chapter 3, we expand from single time step to multi-step closed-loop dynamics. Although we could still frame it in the context of feasibility, in the long term, we are more concerned with what are plausible outcomes. Hence, we need to model user behavior, which is a challenging task mathematically. First, we need a model of how users make decisions given recommendations, and second, how recommendations influence the underlying human decision-making. Here again, the canonical assumption is that true underlying preferences exist, which are typically assumed to lie in the same Euclidean space as the user-item representations, i.e. the *true (and unknown) preference* of user u at time t can be captured by $\mathbf{p}_u^t \in \mathbb{R}^d$. The rating model typically assumed is a factor model, i.e., $r_{ui}^t = \langle \mathbf{p}_u^t, \mathbf{v}_u^t \rangle$. The influence of the recommender on the user's true preference is similarly assumed to be Markovian, where $\mathbf{p}_u^t = g(\mathbf{p}_u^{t-1}, \texttt{rec}_t)$. This basic model is shared by a considerable body of recent literature on societal impacts of recommendation systems [Jia+19; Kal+21; Pas+21]. Given the widespread use of this modeling assumption, we are interested in making these model developments more systematic and realistic. We propose doing so by modeling explicit behavioral effects grounded in psychological evidence and testing whether they create plausible outcomes. This creates a system for refining, critiquing, and validating behavioral models.

1.2 How do feedback driven dynamics impact ecosystems and markets?

In the previous part, we focused on the feedback loop between a single user and a single algorithmic decision-making system. The feedback loop between a person and an algorithmic system does not occur in isolation; it is influenced by existing social and economic structures. Conversely, these user-learner dynamics shape broader social and economic phenomena. For instance, recommendation systems have been found to alter social network structures by reducing degrees of separation [Raj+22] or reinforcing partisan views [Cin+22].

The impact of algorithmic systems on society is challenging to assess because these systems are embedded in dynamic social contexts. Consider, for instance, the community of researchers who form a social network through paper citations. Without tools like Google Scholar, these networks evolved based on academic affiliations and co-authorships. The explosion of research, particularly in fields like Machine Learning, necessitates tools that filter and rank relevant articles, altering the dynamics of academic networks. In Chapter 4, we study how algorithmic interventions in the form of link recommendations impact the dynamics

of the underlying social networks. Rather than focusing on specific algorithms, we address the methodological challenges of measuring the impact of algorithmic interventions. These challenges arise because algorithmic recommendations interact with the natural evolution of networks. For example, a researcher who typically cites papers within their citation network (triadic closure) might receive a highly relevant algorithmic recommendation outside of this local network. Even a single algorithmic link can significantly impact the researcher's local neighborhood, connecting them to previously unknown authors, and affecting the future organic evolution of the network. The main contribution of this chapter is to illustrate the complex dynamics between the underlying natural evolution and algorithmic decisions. We introduce the concept of indirect and delayed impacts to highlight the challenges posed by feedback loop dynamics embedded in broader dynamical systems such as social networks.

Having considered multiple users and the dynamics between them, in Chapter 5, we further expand on the number of stakeholders by considering multiple algorithmic decision-makers, whose interactions are mediated by competition for the same pool of users. Here, user behavior is driven by rational economic choices, such as selecting the best music recommendation service among Spotify, YouTube Music, Pandora, and Apple Music. As users shift to services that deliver better recommendations, the resulting distribution shifts prompt learners to update their models. These updates further influence user participation, creating a feedback loop. We analyze these dynamics and show that market segmentation is the only stable outcome. Unlike the single-learner scenario, where participation dynamics can cause representation disparity, multiple learners lead to specialization, potentially maximizing social welfare.

Finally, integrating complex dynamics in the study and development of algorithmic systems requires perspectives from Control Theory, Behavioral Psychology, Causal Inference, and Economics. This thesis introduces tools and techniques from each domain to advance understanding in this area.

Part I

How do users and recommendation systems impact each other?

Chapter 2

Measuring user agency via stochastic reachability

In this chapter we study a user-centric notion of agency with regard to algorithmic recommendations which quantifies the possible outcomes of a 1-step interaction between the user and the system. We propose an evaluation procedure based on stochastic reachability to quantify the maximum probability of recommending a target piece of content to an user for a set of allowable strategic modifications. This framework allows us to compute an upper bound on the likelihood of recommendation with minimal assumptions about user behavior. Stochastic reachability can be used to detect biases in the availability of content and diagnose systemic limitations in the opportunities for discovery granted to users. We show that reachability can be computed efficiently as a convex program for a variety of practical settings, and further argue that reachability is not inherently at odds with accuracy. We demonstrate evaluations of recommendation algorithms trained on large datasets of explicit and implicit ratings. Our results illustrate how recommendation system design and user interventions impact reachability and how these effects can be distributed unevenly.

This chapter is based on the paper "Quantifying Availability and Discovery in Recommender Systems via Stochastic Reachability"[CDR21] written in collaboration with Sarah Dean and Benjamin Recht.

2.1 Background

The development of recommendation systems has historically focused on improving user engagement and satisfaction by improving accuracy metrics. However, a good predictive accuracy is not sufficient to achieve good performance from both the perspective of the individual experience and the perspective of the broader ecosystem. More recently, the literature on recommendation systems has proposed a variety of other metrics for evaluation, including notions of novelty, serendipity, diversity, and coverage [Her+04; CVW11]. Empirical

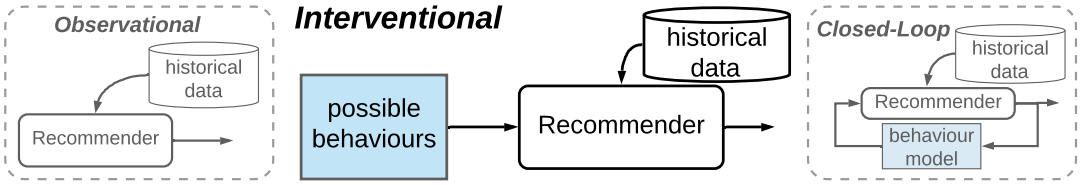


Figure 2.1: Conceptual framings of recommendation systems consider user behaviors to varying degrees. In this work we focus on evaluating interventional properties.

investigations have found evidence of popularity and demographic bias in domains such as movies, music, books, and hotels [Abd+19; Eks+18a; Eks+18b; Jan+15]. Alternative metrics are useful both for diagnosing biases and as objectives for post hoc mitigating techniques such as calibration [Ste18] and reranking [SJ18]. A inherent limitation of these approaches is that they focus on *observational* bias induced by recommendations, i.e. examining the result of a single round of recommendations without considering individuals' behaviors. While certainly useful, they fail to provide a more comprehensive understanding of the interactive nature of recommendation systems.

The behavior of recommendation systems over time and in *closed-loop* is still an open area of study. It is difficult to definitively link observational evidence of radicalization [Rib+20; FCF20] to proprietary recommendation algorithms. Empirical studies of human behavior find mixed results on the relationship between recommendation and content diversity [Ngu+14; FGR16]. Simulation studies [CSE18; Yao+21; Kra+20] and theoretical investigations [DGL13] shed light on phenomena in simplified settings, showing how homogenization, popularity bias, performance, and polarization depend on assumed user behavior models. Even ensuring accuracy in sequential dynamic settings requires contending with closed-loop behaviors. Recommendation algorithms must mitigate biased sampling in order to learn underlying user scoring models, using causal inference based techniques [Sch+16; Yan+18] or by balancing exploitation and exploration [Kaw+15; MGP15]. Reinforcement Learning algorithms face these challenges while considering a longer time horizon [Che+19; Ie+19], implicitly using data to exploit user behavior.

Our work eschews behavior models in favor of an *interventional* framework that considers a variety of possible user actions (see Figure 2.1. Giving users control over their recommendations has been found to have positive effects, while reducing agency has negative effects [Har+15; Luk+21]. The formal perspective we take on agency and access in recommender systems was first introduced by [DRR20], and is closely related to a body of work on recourse in consequential decision making [USL19; Kar+20].

Contributions

- **Formulation:** In Section 2.2 we introduce stochastic reachability which measures a granular notion of user agency by quantifying the *possible* outcomes of a 1-step interaction between a user and recommendation system. We further propose a series of aggregate metrics based on stochastic reachability to diagnose the *availability* of content and *discovery* possibilities for individuals.
- **Method:** We show in Section 2.4 that we can efficiently compute stochastic reachability by solving a convex optimization problem for a broad class of relevant recommenders.
- **Analysis:** In Section 2.5, we draw connections between the stochastic and deterministic settings. This perspective allows us to describe the relationship between agency and stochasticity and further to argue that there is not an inherent trade-off between reachability and model accuracy.
- **Audit:** Finally in Section 2.6 , we present an audit of recommendation systems using a variety of datasets and recommender designs. We explore how design decisions influence reachability and the extent to which biases in the training datasets are propagated.

2.2 Stochastic Recommenders

Setting

We consider systems composed of n individuals as well as a collection of m pieces of content. For consistency with the recommender systems literature, we refer to individuals as users, pieces of content as items, and expressed preferences as ratings. We will denote a rating by user u of item i as $r_{ui} \in \mathcal{R}$, where $\mathcal{R} \subseteq \mathbb{R}$ denotes the space of values which ratings can take. For example, ratings corresponding to the percentage of a video watched would have $\mathcal{R} = [0, 1]$ while discrete star ratings would have $\mathcal{R} = \{1, 2, 3, 4, 5\}$. The number of *observed ratings* will generally be much smaller than the total number of possible ratings, and we denote by $\Omega_u \subseteq \{1, \dots, m\}$ the set of items seen by the user u. The goal of a recommendation system is to understand the preferences of users and recommend relevant content.

In this work, we focus on the common setting in which recommenders are the composition of a *scoring function* ϕ with *selection rule* π (Figure 2.2). The scoring function models the preferences of users. It is constructed based on historical data (e.g. observed ratings, user/item features) and returns a score for each user and item pair. For a given user u and item i, we denote $s_{ui} \in \mathbb{R}$ to be the associated score, and for user u we will denote by $\mathbf{s}_u \in \mathbb{R}^m$ the vector of scores for all items. A common example of a scoring function is a machine learning model which predicts future ratings based on historical data.

We will focus on the way that scores are updated after a round of user interaction. For example, if a user consumes and rates several new items, the recommender system should update the scores in response. Therefore, we parameterize the score function by an update

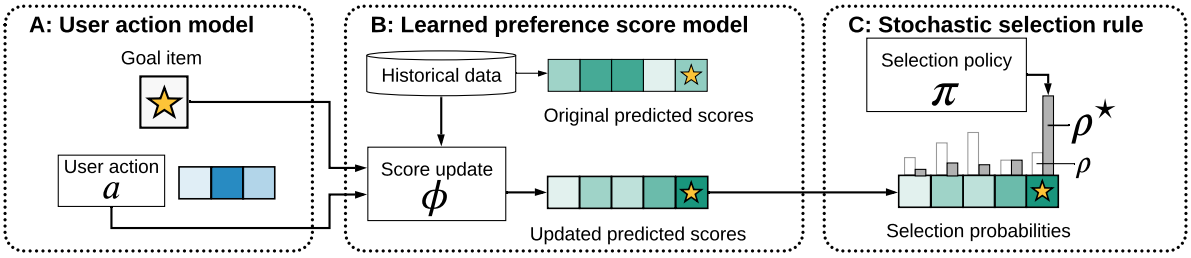


Figure 2.2: We audit recommender systems under a user action model (A), learned scoring model (B), and stochastic selection rule (C).

rule, so that the new score vector is $\mathbf{s}_u^+ = \phi_u(\mathbf{a})$, where $\mathbf{a} \in \mathcal{A}_u$ represents actions taken by user u and $\mathcal{A}_u$ represents the set of all possible actions. Thus ϕ_u encodes the historical data, the scoring model class, and the update algorithm. The action space $\mathcal{A}_u$ represents possibilities for system interaction, encoding for example limitations due to user interface design. We define the form of the score update function and discuss the action space in more detail in Section 2.4.

The selection rule π is a policy which, for given user u and scores $\mathbf{s}_u$, selects one or more items from a set of specified *target items* $\Omega_u^t \subseteq \{1, \ldots, m\}$ as the next recommendation. The simplest selection rule is a top-1 policy, which is a deterministic rule that selects the item with the highest score for each user. A simple stochastic rule is the ε-greedy policy which with probability $1 - \varepsilon$ selects the top scoring item and with probability ε chooses uniformly from the remaining items. Many additional approaches to recommendation can be viewed as the composition of a score function with a selection policy. This setting also encompasses implicit feedback scenarios, where clicks or other behaviors are defined as or aggregated into "ratings."

We are primarily interested in stochastic policies which select items according to a probability distribution on the scores $\mathbf{s}_u$ parametrized by a exploration parameter. Policies of this form are often used to balance exploration and exploration in online or sequential learning settings. A stochastic selection rule recommends an item i according to $\mathbb{P}(\pi(\mathbf{s}_u, \Omega_u^t) = i)$, which is 0 for all non-target items $i \notin \Omega_u^t$. For example, to select among items that have not yet been seen by the user, the target items are set as $\Omega_u^t = \Omega_u^c$ (recalling that Ω_u denotes the set of items seen by the user u and the superscript $\cdot^c$ denotes the complement of a set). Deterministic policies are a special case of stochastic policies, with a degenerate distribution.

Stochastic policies have been proposed in the recommender system literature to improve diversity [Chr+15] or efficiently explore in a sequential setting [Kaw+15]. By balancing exploitation of items with high predicted ratings against explorations of items with lower predictions, preferences can be estimated so that future predicted ratings are more accurate. However, we decidedly do not take a perspective based on accuracy. Rather than supposing that users' reactions are predictable, we consider a perspective centered on agency and access.

2.3 Defining Reachability

First defined in the context of recommendations by [DRR20], an item i is *deterministically reachable* by a user u if there is some allowable modification to the user's ratings $\mathbf{r}_u$ that causes item to be recommended. Allowable modifications can include history edits, such as removing or changing ratings of previously rated items. They can also include future looking modifications which assign ratings to a subset of unseen items.

In the setting where recommendations are made stochastically, we define an item i to be ρ *reachable* by a user u if there is some allowable action $\mathbf{a}$ such that the updated probability that item i is recommended after applying action $\mathbf{a}$ is at least ρ, i.e.:

$$\mathbb{P}(\pi(\phi_u(\mathbf{a}), \Omega_u^t) = i) \geq \rho$$

The maximum ρ reachability for a user-item pair is defined as the solution to the following optimization problem:

$$\rho^\star(u, i) = \max_{\mathbf{a} \in \mathcal{A}_u} \ \mathbb{P}(\pi(\phi_u(\mathbf{a}), \Omega_u^t) = i). \tag{2.1}$$

We will also refer to $\rho^\star(u, i)$ as "max reachability." for the user-item pair.

For example, in the case of ε-greedy policy, $\rho^\star(u, i) = 1 - \varepsilon$ if item i is deterministically reachable by user u, and is $\varepsilon/(|\Omega_u^t| - 1)$ otherwise.

By measuring the maximum achievable probability of recommending an item to a user, we are characterizing a granular metric of *access* within the recommender system. It can also be viewed as an upper bound on the likelihood of recommendation with minimal assumptions about user behavior. It may be illuminating to contrast this measure with a notion of expected reachability. Computing expected reachability would require specifying the distribution over user actions, which would amount to modelling human behavior. In contrast, max reachability requires specifying only the constraints arising from system design choices to define $\mathcal{A}_u$ (e.g. the user interface). By computing max reachability, we focus our analysis on the design of the recommender system, and avoid conclusions which are dependent on behavioral modelling choices.

Two related notions of user agency with respect to a target item i are *lift* and *rank gain*. The lift measures the ratio between the maximum achievable probability of recommendation and the baseline:

$$\lambda^\star(u, i) = \frac{\rho^\star(u, i)}{\rho_0(u, i)} \tag{2.2}$$

where the baseline $\rho_0(u, i)$ is defined to capture the default probability of recommendation in the absence of strategic behavior, e.g. $\mathbb{P}\left(\pi\left(\mathbf{s}_u, \Omega_u^t\right) = i\right)$.

The rank gain for an item i is the difference in the ranked position of the item within the original list of scores $\mathbf{s}_u$ and its rank within the updated list of scores $\mathbf{s}_u^+$.

Lift and rank gain are related concepts, but ranked position is combinatorial in nature and thus difficult to optimize for directly. They both measure agency because they compare

the default behavior of a system to its behavior under a strategic intervention by the user. Given that recommenders are designed with personalization in mind, we view the ability of users to influence the model in a positive light. This is in contrast to much recent work in robust machine literature where strategic manipulation is undesirable.

Diagnosing System Limitations

The analysis of stochastic reachability can be used to audit recommender systems and diagnose systemic biases from an interventional perspective (see Figure 2.1).

Unlike studies of observational bias, reachability analysis takes into account system interactivity. However, unlike studies of closed-loop bias, there is no dependence on a user behavioral model. Because max reachability considers the best case over possible actions, it isolates structural biases from those caused in part by user behavior.

Max reachability is a metric defined for each user-item pair, and disparities across users and items can be detected through aggregations. Aggregating over target items gives insight into a user's ability to discover content, thus detecting users who have been "pigeonholed" by the algorithm. Aggregations over users can be used to compare how the system makes items available for recommendation. We define the following user- and item-based aggregations:

$$D_u = \sum_{i \in \Omega_u^t} \frac{\mathbf{1}\{\rho_{ui} > \rho_t\}}{|\Omega_u^t|}, \quad A_i = \frac{\sum_u \rho_{ui} \mathbf{1}\{i \in \Omega_u^t\}}{\sum_u \mathbf{1}\{i \in \Omega_u^t\}} \tag{2.3}$$

The user discovery D_u is the proportion of target items that have a high chance of being recommended, as determined by the threshold ρ_t. A natural threshold is the better-than-uniform threshold, $\rho_t = 1/|\Omega_u^t|$, recalling that Ω_u^t is the set of target items. When $\rho_{ui} = \rho_0(u, i)$, baseline discovery counts the number of items that will be recommended with better-than-uniform probability and is determined by the spread of the recommendation distribution. When $\rho_{ui} = \rho^\star(u, i)$, discovery counts the number of items that a user *could* be recommended with better-than-uniform probability in the best case. Low best-case discovery means that the recommender system inherently limits user access to content.

The item availability A_i is the average likelihood of recommendation over all users who have item i as a target. It can be thought of as the chance that a uniformly selected user will be recommended item i. When $\rho_{ui} = \rho_0(u, i)$, the baseline availability measures the prevalence of the item in the recommendations. When $\rho_{ui} = \rho^\star(u, i)$, availability measures the prevalence of an item in the best case. Low best-case availability means that the recommender system inherently limits the distribution of a given item.

2.4 Computing Reachability

In this section, we consider a broad class of recommender system designs for which the max reachability problem can be efficiently solved via convex optimization.

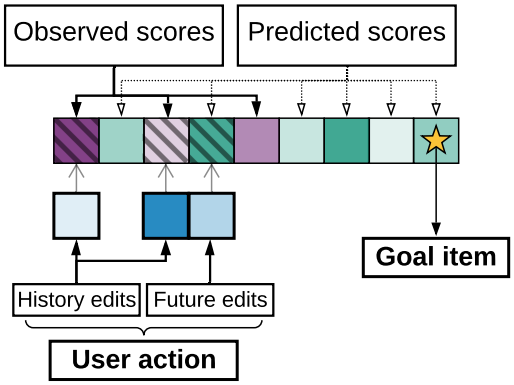


Figure 2.3: User action space: The shade represents the magnitude of historical (purple) or predicted (green) rating. The *action items* are marked with diagonal lines; they can be strategically modified to maximize the recommendation probability of the *goal item* (star). The value of the user action is shaded in blue.

User action model

We suppose that users interact with the system through expressed preferences, and thus actions are updates to the vector $\mathbf{r}_u \in \mathcal{R}^m$, a sparse vector of observed ratings. For each user, the action model is based on distinguishing between *action* and *immutable* items.

Let $\Omega_u^{\mathcal{A}}$ denote the set of action items for which the ratings can be strategically modified by the user u. Then the action set $\mathcal{A}_u = \mathcal{R}^{|\Omega_u^{\mathcal{A}}|}$ corresponds to changing or setting the value of these ratings (Figure 2.3 provides an illustration). The action set should be defined to correspond to the interface through which a user interacts with the recommender system. In practice, most commercial recommendation systems surface only "in the next" items, thus limiting the user action space. Other designs more mindful of promoting reachability could include a display view of "previously viewed" or "discovery" items

The updated rating vector $\mathbf{r}_u^+ \in \mathcal{R}^m$ is equal to $\mathbf{r}_u$ at the indices corresponding to immutable items and equal to the action $\mathbf{a}$ at the action items. Note the partition into action and immutable is distinct from earlier partition of items into observed and unobserved; action items can be both seen (history edits) and unseen (future reactions), as illustrated in Figure 2.2 (A). For the reachability problem, we will consider a set of target items Ω_u^t that does not intersect with the action items $\Omega_u^{\mathcal{A}}$. Depending on the specifics of the recommendation setting, we may also require that it does not intersect with the previously rated items Ω_u.

We remark that additional user or item features used for scoring and thus recommendations could be incorporated into this framework as either mutable or immutable features. The only computational difficulty arises when mutable features are discrete or categorical.

Scoring function

The recommender model is composed of a scoring function ϕ and a selection function π. We consider *affine score update functions* where for each user, scores are determined by an affine function of the action: $\mathbf{s}_u^+ = \phi_u(\mathbf{a}) = B_u\mathbf{a} + \mathbf{c}_u$ where $B_u \in \mathbb{R}^{m\times|\Omega_u^{\mathcal{A}}|}$ and $\mathbf{c}_u \in \mathbb{R}^m$ are model parameters determined in part by historical data. Such a scoring model arises from a variety of scoring rules, as shown below:

Example 2.4.1 (Matrix Factorization with Gradient Descent)**.** Matrix factorization (MF) models compute scores as rating predictions so that $S = PQ^\top$, where $P \in \mathbb{R}^{n\times d}$ and $Q \in \mathbb{R}^{m\times d}$ are respectively user and item factors for some latent dimension d. They are learned via the optimizating the loss on the observed ratings:

$$\mathcal{L}(P, Q) = \sum_u \sum_{i\in\Omega_u} \|\mathbf{p}_u^\top \mathbf{q}_i - r_{ui}\|_2^2 \,.$$

Given a user action sets $\Omega_u^{\mathcal{A}}$ and corresponding strategic ratings the objective function has an additional action term:

$$\mathcal{L}^{\mathcal{A}}(P, Q) = \sum_u \sum_{i\in\Omega_u^{\mathcal{A}}} \|\mathbf{p}_u^\top \mathbf{q}_i - r_{ui}\|_2^2.$$

The one-step update rule for a user factor under a gradient descent (GD) minimization scheme with step size α, is:

$$\mathbf{p}_u^+ = \mathbf{p}_u - \alpha\nabla\mathcal{L}_{\mathbf{p}_u}^{\mathcal{A}}(Q) = \mathbf{p}_u - \alpha \sum_{i\in\Omega_u^{\mathcal{A}}} \left(\mathbf{q}_i\mathbf{q}_i^\top \mathbf{p}_u - \mathbf{q}_i r_{ui}\right) ,$$

Notice that the score update is affine in the action items. Therefore, we have an affine score function:

$$\mathbf{s}_u^+ = \phi_u(\mathbf{a}) = Q\mathbf{p}_u^+ = Q\left(\mathbf{p}_u - \alpha Q_{\mathcal{A}}^\top Q_{\mathcal{A}}\mathbf{p}_u - \alpha Q_{\mathcal{A}}^\top \mathbf{a}\right)$$

where we define $Q_{\mathcal{A}} = Q_{\Omega_u^{\mathcal{A}}} \in \mathbb{R}^{|\Omega_u^{\mathcal{A}}|\times d}$. Therefore,

$$B_u = -\alpha Q Q_{\mathcal{A}}^\top, \quad \mathbf{c}_u = Q\left(\mathbf{p}_u - \alpha Q_{\mathcal{A}}^\top Q_{\mathcal{A}}\mathbf{p}_u\right) \,.$$

Example 2.4.2 (Neighborhood based models)**.** Item-neighborhood models compute scores as rating predictions by a weighted average, with:

$$s_{ui} = \frac{\sum_{j\in\mathcal{N}_i} w_{ij} r_{uj}}{\sum_{j\in\mathcal{N}_i} |w_{ij}|}$$

where w_{ij} are weights representing similarities between items and $\mathcal{N}_i$ is a set of indices of previously rated items in the neighborhood of item i. Regardless of how the neighborhood is determined and how the are computed, the predicted scores are a linear function of observed scores: $\mathbf{s}_u = W\mathbf{r}_u$.

Since the neighborhood and the similarity weights are not impacted by user actions the score updates take the form:

$$\phi_u(\mathbf{a}) = W\mathbf{r}_u^+ = \underbrace{W\mathbf{r}_u}_{\mathbf{c}_u} + \underbrace{WE_{\Omega_u^{\mathcal{A}}}}_{B_u}\mathbf{a}$$

where $E_{\Omega_u^{\mathcal{A}}}$ selects rows of W corresponding to action items.

In both examples, the action matrices can be decomposed into two terms. The first is a term that depends only on the recommender model (e.g. item factors Q or weights W), while the second is dependent on the user action model (e.g. action item factors $Q_{\mathcal{A}}$ or action selector $E_{\Omega_u^{\mathcal{A}}}$).

For simplicity of presentation, the examples above leave out model bias terms, which are common in practice. Incorporating these model biases changes only the definition of the affine term in the score update expression as seen in the additional examples below:

Example 2.4.3 (Biased MF-SGD)**.** Biased matrix factorization models [KB15] compute scores as rating predictions with

$$s_{ui} = \mathbf{p}_u^\top \mathbf{q}_i + f_u + g_i + \mu$$

$P \in \mathbb{R}^{n\times d}$ and $Q \in \mathbb{R}^{m\times d}$ are respectively user and item factors for some latent dimension d, $\mathbf{f} \in \mathbb{R}^n$ and $\mathbf{g} \in \mathbb{R}^m$ are respectively user and item biases, and $\mu \in \mathbb{R}$ is a global bias.

The parameters are learned via the regularized optimization

$$\min_{P,Q,\mathbf{f},\mathbf{g},\mu} \sum_u \sum_{i\in\Omega_u} \|\mathbf{p}_u^\top \mathbf{q}_i + f_u + g_i + \mu - r_{ui}\|_2^2 + \lambda\|P\|_F^2 + \lambda\|Q\|_F^2 \,.$$

Under a gradient descent minimization scheme [Kor08; Zin03] with step size α, the one-step update rule for a user factor is:

$$\mathbf{p}_u^+ = \mathbf{p}_u - \alpha \sum_{i\in\Omega_u^{\mathcal{A}}} (\mathbf{q}_i\mathbf{q}_i^\top \mathbf{p}_u + \mathbf{q}_i(f_u + g_i + \mu) - \mathbf{q}_i r_{ui}) - \alpha\lambda\mathbf{p}_u \,.$$

User bias terms can be updated in a similar manner, but because the user bias is equal across items, it does not impact the selection of items.

Notice that this expression is affine in the mutable ratings. Therefore, we have an affine score function:

$$\phi_u(\mathbf{a}) = Q\mathbf{p}_u^+ = Q\left((1-\alpha\lambda)\mathbf{p}_u - \alpha Q_{\mathcal{A}}^\top(Q_{\mathcal{A}}\mathbf{p}_u + \mathbf{g}_{\mathcal{A}} + (\mu + f_u)\mathbf{1}) + \alpha Q_{\mathcal{A}}^\top\mathbf{a}\right)$$

where we define $Q_{\mathcal{A}} = Q_{\Omega_u^{\mathcal{A}}} \in \mathbb{R}^{|\Omega_u^{\mathcal{A}}|\times d}$ and $\mathbf{g}_{\mathcal{A}} = \mathbf{g}_{\Omega_u^{\mathcal{A}}} \in \mathbb{R}^{|\Omega_u^{\mathcal{A}}|}$. Therefore,

$$B_u = \alpha Q Q_{\mathcal{A}}^\top, \quad \mathbf{c}_u = Q\left((1+\lambda)\mathbf{p}_u - \alpha Q_{\mathcal{A}}^\top(Q_{\mathcal{A}}\mathbf{p}_u + \mathbf{g}_{\mathcal{A}} + (\mu + f_u)\mathbf{1})\right) \,.$$

Example 2.4.4 (Biased MF-ALS)**.** Rather than a gradient descent minimization scheme, we may instead update the model with an alternating least-squares (ALS) strategy [Zho+08]. In this case, the update rule is:

$$\begin{aligned}\mathbf{p}_u^+ &= \arg\min_{\mathbf{p}} \sum_{i \in \Omega_u^{\mathcal{A}} \cap \Omega_u} \|\mathbf{p}^\top \mathbf{q}_i + f_u + g_i + \mu - r_{ui}\|_2^2 + \lambda \|\mathbf{p}\|_2^2 \\ &= (Q_u^\top Q_u + \lambda I)^{-1}(Q^\top \mathbf{r}_u + Q_{\mathcal{A}}^\top(\mathbf{g}_{\mathcal{A}} + (\mu + f_u)\mathbf{1}) + Q_{\mathcal{A}}^\top \mathbf{a})\end{aligned}$$

where we define $Q_u = Q_{\Omega_u^{\mathcal{A}} \cap \Omega_u}$. Similar to in the SGD setting, this is an affine expression, and therefore we end up with the affine score parameters

$$B_u = Q(Q_u^\top Q_u + \lambda I)^{-1} Q_{\mathcal{A}}^\top, \quad \mathbf{c}_u = Q(Q_u^\top Q_u + \lambda I)^{-1}(Q^\top \mathbf{r}_u + Q_{\mathcal{A}}^\top(\mathbf{g}_{\mathcal{A}} + (\mu + f_u)\mathbf{1}))\,.$$

Example 2.4.5 (Biased Item-KNN)**.** Biased neighborhood models [DK11] compute scores as rating predictions by a weighted average, with

$$s_{ui} = \mu + f_u + g_i + \frac{\sum_{j \in \mathcal{N}_i} w_{ij}(r_{uj} - \mu - f_u - g_i)}{\sum_{j \in \mathcal{N}_i} |w_{ij}|}$$

where w_{ij} are weights representing similarities between items, $\mathcal{N}_i$ is a set of indices which are in the neighborhood of item i, and $\mathbf{f}, \mathbf{g}, \mu$ are bias terms. Regardless of the details of how these parameters are computed, the predicted scores are an affine function of observed scores:

$$\mathbf{s}_u = W\mathbf{r}_u - W(\mathbf{g} + (\mu + f_u)\mathbf{1}) + \mathbf{g} + (\mu + f_u)\mathbf{1}$$

where we can define

$$W_{ij} = \begin{cases} \frac{w_{ij}}{\sum_{j \in \mathcal{N}_i} |w_{ij}|} & j \in \mathcal{N}_i \\ 0 & \text{otherwise} \end{cases}$$

Therefore, the score updates take the form:

$$\phi_u(\mathbf{a}) = \underbrace{W(\mathbf{r}_u - \mathbf{g} + (\mu + f_u)\mathbf{1}) + \mathbf{g} + (\mu + f_u)\mathbf{1})}_{\mathbf{c}_u} + \underbrace{WE_{\Omega_u^{\mathcal{A}}}}_{B_u}\mathbf{a}$$

where $E_{\Omega_u^{\mathcal{A}}}$ selects rows of W corresponding to action items.

Example 2.4.6 (SLIM and EASE)**.** SLIM [NK11] and EASE [Ste19] are two other popular models where user-item scores are computed as

$$s_{ui} = \mathbf{w}_i^\top \mathbf{r}_u$$

for $\mathbf{w}_i$ the row vectors of a weight matrix W. For SLIM, the sparse weights are computed as

$$\begin{aligned}\min_W \ & \frac{1}{2}\|R - RW\|_F^2 + \frac{\beta}{2}\|W\|_F^2 + \lambda\|W\|_1 \\ \text{s.t. } & W \geq 0, \operatorname{diag}(W) = 0\end{aligned}$$

For EASE, the weights are computed as

$$\begin{aligned} \min_{W} \ & \frac{1}{2}\|R - RW\|_F^2 + \lambda\|W\|_F^2 \\ \text{s.t. } & \operatorname{diag}(W) = 0 \end{aligned}$$

In both cases, the score updates take the form

$$\phi_u(\mathbf{a}) = \underbrace{W\mathbf{r}_u}_{\mathbf{c}_u} + \underbrace{WE_{\Omega_u^{\mathcal{A}}}}_{B_u}\mathbf{a}\,.$$

Item selection rule

We now turn to the selection component of the recommender, which translates the score $\mathbf{s}_u$ into a probability distribution over target items. The stochastic policy we consider throughout this chapter is the soft-max selection:

Definition 2.4.7 (Soft-max selection policy)**.** For $i \in \Omega_u^t$, the probability of item selection is given by:

$$\mathbb{P}(\pi_\beta(\mathbf{s}_u, \Omega_u^t) = i) = \frac{e^{\beta s_{ui}}}{\sum_{j\in\Omega_u^t} e^{\beta s_{uj}}}\,.$$

This form of stochastic policy samples an item according to a Boltzmann distribution defined by the predicted scores (Figure 2.2 part C). Distributions of this form are common in machine learning applications, and are known as Boltzmann sampling in reinforcement learning or online learning settings [Wei+17; Ces+17]. The choice of this distribution is further motivated by the principle of maximum entropy from Statistical Mechanics which states that the soft-max distribution maximizes entropy (diversity) for a given expected value (mean rating).

Convex Optimization

We now show that under affine score update models and soft-max selection rules, the maximum stochastic reachability problem can be solved by an equivalent convex problem. First notice that for a soft-max selection rule with parameter β, we have that:

$$\log\left(\mathbb{P}(\pi_\beta(\mathbf{s}_u, \Omega_u^t) = i)\right) = \beta s_{ui} - \operatorname*{LSE}_{j\in\Omega_u^t}\left(\beta s_{uj}\right)$$

where LSE is the log-sum-exp function.

Maximizing stochastic reachability is equivalent to minimizing its negative log-likelihood. Letting $\mathbf{b}_{ui}$ denote the i-th row of the action matrix B_u and substituting the form of the score update rule, we have the equivalent optimization problem:

$$\min_{\mathbf{a}\in\mathcal{A}_u} \operatorname*{LSE}_{j\in\Omega_u^t}\left(\beta(\mathbf{b}_{uj}^\top\mathbf{a} + c_{uj})\right) - \beta(\mathbf{b}_{ui}^\top\mathbf{a} + c_{ui}) \tag{2.4}$$

If the optimal value to (2.4) is $\gamma^\star(u,i)$, then the optimal value for (2.1) is given by $\rho^\star(u,i) = e^{-\gamma^\star(u,i)}$. The objective in (2.4) is convex because log-sum-exp is a convex function, affine functions are convex, and the composition of a convex and an affine function is convex. Therefore, whenever the action space $\mathcal{A}_u$ is convex, so is the optimization problem. The size of the decision variable scales with the dimension of the action, while the objective function relies on a matrix-vector product of size $|\Omega_u^t| \times |\mathcal{A}_u|$. Being able to solve the maximum reachability problem quickly is of interest, since auditing an entire system requires computing $\rho^\star$ for many user and item pairs.

2.5 Geometry of Reachability

In this section, we explore the connection between stochastic and deterministic reachability to illustrate how both randomness and agency contribute to discovery as defined by the max reachability metric. We then argue by example that it is possible to design scoring models that guarantee deterministic reachability, and that doing so does not induce accuracy trade-offs.

Connection to Deterministic Recommendation

We now explore how the soft-max style selection rule is a relaxation of top-1 recommendation. For larger values of β, the selection rule distribution becomes closer to the deterministic top-1 rule. This also means that the stochastic reachability problem can be viewed as a relaxation of the top-1 reachability problem.

In stochastic settings it is relevant to inquire the extent to which randomness impacts discovery and availability. In the deterministic setting, the reachability of an item to a user is closely tied to agency—the ability of a user to influence their outcomes. The addition of randomness induces exploration, but not in a way that is controllable by users. In the following result, we show how this trade-off manifests in the max reachability metric itself.

Proposition 2.5.1. *Consider the stochastic reachability problem for a β-soft-max selection rule as $\beta \to \infty$. Then if an item i is top-1 reachable by user u, $\rho^\star(u,i) \to 1$. In the opposite case that item i* is not *top-1 reachable, we have that $\rho^\star(u,i) \to 0$.*

Proof. Define

$$\gamma_\beta(\mathbf{a}) = \operatorname*{LSE}_{j \in \Omega_u^t} (\beta \phi_{uj}(\mathbf{a})) - \beta \phi_{ui}(\mathbf{a})$$

and see that $\rho_{ui}(\mathbf{a}) = e^{-\gamma_\beta(\mathbf{a})}$. Then we see that

$$\lim_{\beta \to \infty} \frac{1}{\beta} \gamma_\beta(\mathbf{a}) = \max_{j \notin \Omega_u} (\phi_{uj}(\mathbf{a})) - \phi_{ui}(\mathbf{a})$$

yields a top-1 expression. If an item i is top-1 reachable for user u, then there is some $\mathbf{a}$ such that the above expression is equal to zero. Therefore, as $\beta \to \infty$, $\gamma^\star \to 0$, hence $\rho^\star \to 1$.

In the opposite case when an item *is not* top-1 reachable we have that $\gamma^\star \to \infty$, hence $\rho^\star \to 0$. □

This connection yields insight into the relationship between max reachability, randomness, and agency in stochastic recommender systems. For items which are top-1 reachable, larger values of β result in larger $\rho^\star$, and in fact the largest possible max reachability is attained as $\beta \to \infty$, i.e. there is no randomness. On the other hand, if β is too large, then items which are not top-1 reachable will have small $\rho^\star$. There is some optimal finite $\beta \geq 0$ that maximizes $\rho^\star$ for top-1 unreachable items. Therefore, we see a delicate balance when it comes to ensuring access with randomness.

Viewed in another light, this result says that for a fixed $\beta \gg 1$, deterministic top-1 reachability ensures that $\rho^\star$ will be close to 1. We explore this perspective in the next section.

Reachability Without Sacrificing Accuracy

Specializing to affine score update models, we now highlight how parameters of the recommender and action models play a role in determining max reachability. Building on the connection to deterministic reachability, we make use of results about model and action space geometry from [DRR20]. We recall the definition of the convex hull.

Definition 2.5.2 (Convex hull)**.** The *convex hull* of a set of vectors $\mathcal{V} = \{\mathbf{v}_i\}_{i=1}^n$ is defined as

$$\operatorname{conv}(\mathcal{V}) = \left\{ \sum_{i=1}^n w_i \mathbf{v}_i \mid \mathbf{w} \in \mathbb{R}_+^n, \quad \sum_{i=1}^n w_i = 1 \right\} .$$

A point $\mathbf{v}_j \in \mathcal{V}$ is a *vertex* of the convex hull if

$$\mathbf{v}_j \notin \operatorname{conv}(\mathcal{V} \setminus \{\mathbf{v}_j\}) .$$

Proposition 2.5.3. *If $\mathbf{b}_{ui}$ is a vertex on the convex hull of $\{\mathbf{b}_{uj}\}_{j \in \Omega_u^t}$ and actions are real-valued, then $\rho_{ui}^\star \to 1$ as $\beta \to \infty$.*

Proof. We begin by showing that if $\mathbf{b}_{ui}$ is a vertex on the convex hull of $\mathcal{B} = \{\mathbf{b}_{uj}\}_{j \in \Omega_u^t}$, then item i is top-1 reachable. This argument is similar to the proof of Results 1 and 2 in [DRR20].

Item i is top-1 reachable if there exists some $\mathbf{a} \in \mathbb{R}^{|\Omega_u^A|}$ such that $\mathbf{b}_{ui}^\top \mathbf{a} + c_{ui} \geq \mathbf{b}_{uj}^\top \mathbf{a} + c_{uj}$ for all $j \neq i$. Therefore, top-1 reachability is equivalent to the feasibility of the following linear program

$$\begin{aligned} &\min\ 0^\top \mathbf{a} \\ &\text{s.t. } D_{ui}\mathbf{a} \geq \mathbf{f}_{ui} \end{aligned}$$

where D_{ui} has rows given by $\mathbf{b}_{ui} - \mathbf{b}_{uj}$ and $\mathbf{f}_{ui}$ has entries given by $c_{uj} - c_{ui}$ for all $j \in \Omega_u^t$ with $j \neq i$. Feasibility of this linear program is equivalent to boundedness of its dual:

$$\begin{aligned} &\max\ \mathbf{f}_{ui}^\top \lambda \\ &\text{s.t. } D_{ui}^\top \lambda = 0, \quad \lambda \geq 0. \end{aligned}$$

We now show that if $\mathbf{b}_{ui}$ is a vertex on the convex hull of $\mathcal{B}$, then the dual is bounded because the only feasible solution is $\lambda = 0$. To see why, notice that

$$D_{ui}^\top \lambda = 0 \iff \mathbf{b}_{ui} \sum_{\substack{j \in \Omega_u^t \\ j \neq i}} \lambda_j = \sum_{\substack{j \in \Omega_u^t \\ j \neq i}} \lambda_j \mathbf{b}_{uj}$$

If this expression is true for some $\lambda \neq 0$, then we can write

$$\mathbf{b}_{ui} = \sum_{\substack{j \in \Omega_u^t \\ j \neq i}} w_j \mathbf{b}_{uj}, \quad w_j = \frac{\lambda_j}{\sum_{\substack{j \in \Omega_u^t \\ j \neq i}} \lambda_j} \implies \mathbf{b}_{ui} \in \operatorname{conv}\left(\mathcal{B} \setminus \{\mathbf{b}_{ui}\}\right) .$$

This is a contradiction, and therefore it must be that $\lambda = 0$ and therefore the dual is bounded and item i is top-1 reachable.

To finish the proof, we appeal to Proposition 2.5.1 to argue that since item i is top-1 reachable, then $\rho_{ui}^\star \to 1$ as $\beta \to \infty$. □

This result highlights how the geometry of the score model determines when it is preferable for the system to have minimal exploration, from the perspective of reachability.

We now consider whether relevant geometric properties of the model are predetermined by the goal of accurate prediction. Is there a tension between ensuring reachability and accuracy? We answer in the negative by presenting a construction for the case of matrix factorization models. Our result shows that the item and user factors (P and Q) can be slightly altered such that all items become top-1 reachable at no loss of predictive accuracy. The construction expands the latent dimension of the user and item factors by one and relies on a notion of sufficient richness for action items.

Definition 2.5.4 (Rich actions)**.** For a set of item factors $\{\mathbf{q}_j\}_{j=1}^m$, let $C = \max_j \|\mathbf{q}_j\|_2$. Then a set of action items $\Omega_u^{\mathcal{A}} \subseteq \{1, \ldots, m\}$ is *sufficiently rich* if the vertical concatenation of their item factors with a norm factor is full rank:

$$\operatorname{rank}\left(\begin{bmatrix} \mathbf{q}_i^\top & \sqrt{C^2 - \|\mathbf{q}_i\|_2^2} \end{bmatrix}_{i \in \Omega_u^{\mathcal{A}}} \right) = d + 1 .$$

Notice that this can only be true if $|\Omega_u^{\mathcal{A}}| \geq d + 1$. Further note that the norm factor makes all terms have the same norm C and such are at the boundary of the ℓ_2 ball with radius C in $\mathbb{R}^{d+1}$. Thus, richness is guaranteed for any $d + 1$ action items, as long as the corresponding $\mathbf{q}_j$ are unique.

Proposition 2.5.5. *Consider the MF model with user factors $P \in \mathbb{R}^{n \times d}$ and item factors $Q \in \mathbb{R}^{m \times d}$. Further consider any user u with a sufficiently rich set of at least $d + 1$ action items and real-valued actions. Then there exist $\widetilde{P} \in \mathbb{R}^{n \times d+1}$ and $\widetilde{Q} \in \mathbb{R}^{m \times d+1}$ such that $PQ^\top = \widetilde{P}\widetilde{Q}^\top$ and under this model, $\rho^\star(u, i) \to 1$ as $\beta \to \infty$ for all target items $i \in \Omega_u^t$.*

Proof. Let C be the maximum row norm of Q and define $\mathbf{v} \in \mathbb{R}^m$ satisfying $v_i^2 = C^2 - \|\mathbf{q}_i\|_2^2$. Then we construct modified item and user factors as

$$\widetilde{Q} = \begin{bmatrix} Q & \mathbf{v} \end{bmatrix}, \quad \widetilde{P} = \begin{bmatrix} P & \mathbf{0} \end{bmatrix}.$$

Therefore, we have that $\widetilde{P}\widetilde{Q}^\top = PQ^\top$ and by construction, each row of $\widetilde{Q}$ has norm C, so each $\widetilde{\mathbf{q}}_i$ is on the boundary of the ℓ_2 ball in $\mathbb{R}^{d+1}$. For an arbitrary user u, the score model parameters are given by $\widetilde{\mathbf{b}}_{ui} = \widetilde{Q}_{\mathcal{A}}\widetilde{\mathbf{q}}_i$. We show by contradiction that as long as the action items are sufficiently rich, each $\widetilde{\mathbf{b}}_{ui}$ is a vertex on the convex hull of $\{\widetilde{\mathbf{b}}_{uj}\}_{j=1}^n$. Supposing this is not the case for an arbitrary i,

$$\widetilde{\mathbf{b}}_{ui} = \sum_{\substack{j=1 \\ j\neq i}}^{n} w_j \widetilde{\mathbf{b}}_{uj} \iff \widetilde{Q}_{\mathcal{A}}\widetilde{\mathbf{q}}_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} w_j \widetilde{Q}_{\mathcal{A}}\widetilde{\mathbf{q}}_i \implies \widetilde{\mathbf{q}}_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} w_j \widetilde{\mathbf{q}}_i$$

where the final implication follows because the fact that $\widetilde{Q}_{\mathcal{A}}$ is full rank (due to richness) implies that $\widetilde{Q}_{\mathcal{A}}^\top \widetilde{Q}_{\mathcal{A}}$ is invertible. This is a contradiction, and therefore we have that each $\widetilde{\mathbf{b}}_{ui}$ must be a vertex on the convex hull of $\{\widetilde{\mathbf{b}}_{uj}\}_{j=1}^n$.

Finally, we appeal to Proposition 2.5.3 to argue that $\rho^\star(u, i) \to 1$ as $\beta \to \infty$ for all target items $i \in \Omega_u^t$. □

The existence of such a construction demonstrates that there is not an unavoidable trade-off between accuracy and reachability in recommender systems.

2.6 Audit Demonstration

Datasets

We evaluate[1] max ρ reachability in settings based on three popular recommendation datasets: MovieLens 1M (`ML1M`) [HK15], LastFM 360K [Cel10] and MIcrosoft News Dataset (MIND) [Wu+20]. Table 2.1 provides summary statistics:

MovieLens 1 Million `ML1M` is a dataset of 1 through 5 explicit ratings of movies, containing data of 6040 unique users for 3706 unique movies. There are a total of 1000209 ratings (4.47% rating density). The original data is accompanied by additional user attributes such as age, gender, occupation and zip code. It was downloaded from Group Lens[2] via the RecLab [Kra+20] interface[3] without any additional pre-processing. Figure 2.4 illustrates descriptive statistics for the `ML1M` dataset.

[1] Reproduction code available at `github.com/modestyachts/stochastic-rec-reachability`

[2] `https://grouplens.org/datasets/movielens/1m/`

[3] `https://github.com/berkeley-reclab/RecLab`

Table 2.1: Summary statistics of the audit datasets. The last two rows are the test set accuracy for matrix factorization and item neighborhood scoring models.

Data set	ML 1M	LastFM 360K	MIND
Users	6040	13698	50000
Items	3706	20109	247
Ratings	1000209	178388	670773
Density (%)	4.47%	0.065%	5.54%
LibFM rmse	0.716	1.122	0.318
KNN rmse	0.756	1.868	-

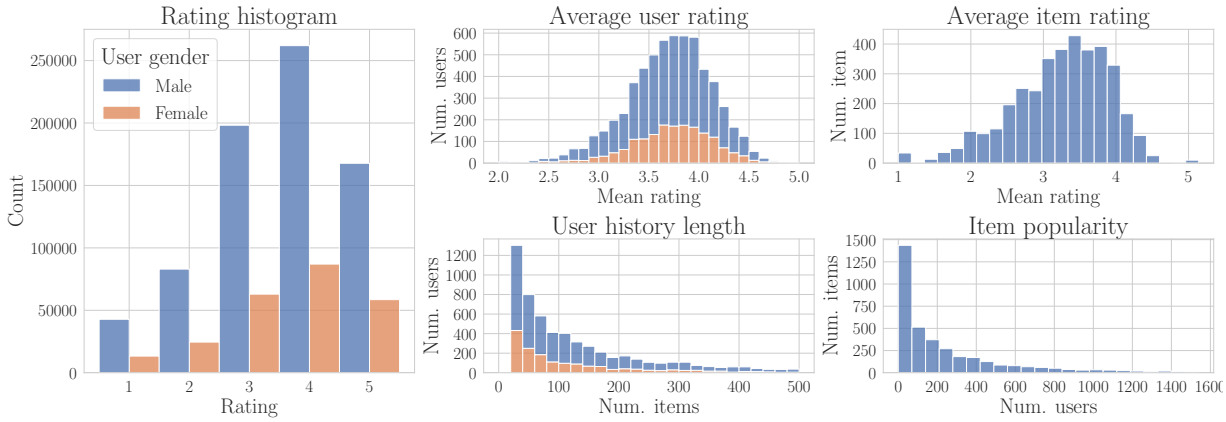


Figure 2.4: Descriptive statistics for the MovieLens 1M dataset split by user gender (28.3% female). The mean ratings of both users and items are roughly normally distributed while user's history length and item popularity display power law distributions.

LastFM 360K LastFM is an implicit rating dataset containing the number of times a user has listened to songs of an artist. We used the version of the LastFM dataset[4] preprocessed by [Sha+20]. For computational tractability, we select a random subset of 10% of users and 10% artists yielding 13698 users, 20109 items and 178388 ratings (0.056% rating density). The item ratings are not explicitly expressed by users as in the MovieLens case. For a user u and an artist i we define implicit ratings $r_{ui} = \log(\#\text{listens}(u, i) + 1)$. This data is accompanied by artist gender, an item attribute. Figure 2.5 illustrates the descriptive statistics for the LastFM dataset.

[4] https://zenodo.org/record/3964506#.XyE5N0FKg5n

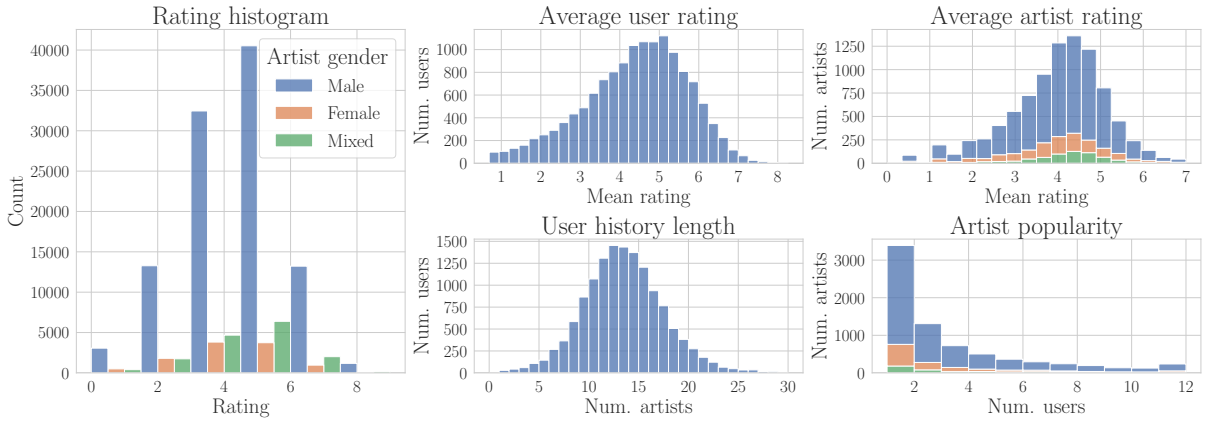


Figure 2.5: Descriptive statistics for the LastFM dataset split by artist gender (over 54% of artists have unknown gender, 36% are male, 6.5% are female and 3.5% are mixed gender). Unlike ML 1M, for the LastFM dataset the user history lengths are normally distributed around a mean of around 12 artists.

MIcrosoft News Dataset (MIND) MIND is an implicit rating dataset containing clicks and impressions data collected from logs of the Microsoft News website [5]. We use the `MIND-small` dataset[6], which contains behaviour log data for 50000 randomly sampled users. There are 42416 unique news articles, spanning 17 categories and 247 subcategories. We transform news level click data into subcategory level aggregation and define the rating associated with a user-subcategory pair as a function of the number of times that the user clicked on news from that subcategory: $r_{ui} = \log(\#\text{clicks}(u, i) + 1)$. The resulting aggregated dataset contains 670773 ratings (5.54% rating density). Figure 2.6 illustrates descriptive statistics for the MIND dataset.

Scoring models

We consider two scoring models: one based on matrix factorization (MF) as well as a neighborhood based model (KNN). We use the LibFM SGD implementation [Ren12] for the MF model and use the item-based k-nearest neighbors model implemented by [Kra+20]. For each dataset and recommender model we perform grid search for progressively finer meshes over the tunable hyper-parameters of the recommender. We use recommenders implemented by the RecLab library. For each dataset and recommender we evaluate hyperparameters on a 10% split of test data. The best hyper-parameters for each setting are presented in Table 2.2.

LibFM We performed hyper-parameter tuning to find suitable learning rate and regularization parameter for each dataset. Following [Dac+21] we consider `lr` $\in (0.001, 0.5)$ as the

[5] https://microsoftnews.msn.com/
[6] https://msnews.github.io/

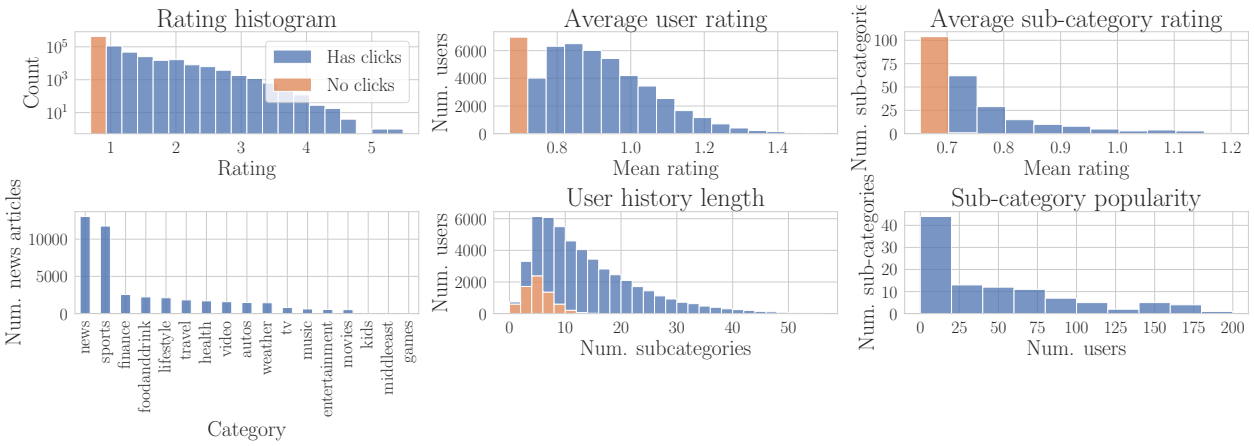


Figure 2.6: Descriptive statistics for the MIND dataset: The orange bars correspond to either user or items that have been displayed but have not clicked/ have not been clicked on. Unlike ML 1M and LastFM, the MIND ratings have strongly skewed distribution, with most user-subcategory ratings corresponding to users clicking on a small number of articles from the sub-category. There is a long tail of higher ratings that corresponds to most popular subcategories. The leftmost plot illustrates the unequal distribution of news articles across categories. The same qualitative behaviour holds for sub-categories.

range of hyper-parameters for the learning rate and $\texttt{reg} \in (10^{-5}, 10^{0})$ for the regularization parameter. In all experimental settings we follow the setup of [RZK19] and use 64 latent dimensions and train with SGD for 128 iterations.

KNN We perform hyperparameter tuning with respect to neighborhood size and shrinkage parameter. Following [Dac+21] we consider the range $(5, 1000)$ for the neighborhood size and $(0, 1000)$ for the shrinkage parameter. We tune KNN only for the `ML1M` dataset.

All experiments were performed on a 64 bit desktop machine equipped with 20 CPUs (Intel(R) Core(TM) i9-7900X CPU @ 3.30GHz) and a 62 GiB RAM. Average run times for training an instance of each recommender can be found in Table 2.2.

Computing Reachability

To compute reachability, it is further necessary to specify additional elements of the recommendation pipeline: the user action model, the set of target items, and the soft-max selection parameter.

ML 1M We compute max stochastic reachability for the LibFM and KNN scoring model. We consider three types of user action spaces: *History Edits*, *Future Edits*, and *Next K* in which users can strategically modify the ratings associated to K randomly chosen items from their history, K randomly chosen items from that they have not yet seen, or the top-K

	LibFM			
Dataset	Learning Rate	Regularization	Test RMSE	Run time (s)
ML 1M	0.0112	0.0681	0.716	2.76 ± 0.32
LastFM	0.0478	0.2278	1.122	0.78 ± 0.13
MIND	0.09	0.0373	0.318	3.23 ± 0.37
	KNN			
Dataset	Neighborhood size	Shrinkage	Test RMSE	Run time (s)
ML 1M	100	22.22	0.756	0.34 ± 0.07

Table 2.2: Tuning results: Optimal parameters for LibFM and KNN along with corresponding test performance and average run times.

unseen items according to the baseline scores of the scoring model. For each of the action spaces we consider $K \in \{5, 10, 20\}$.

We perform reachability experiments on a random 3% subset of users (176). For each choice of scoring model, action space type and action space size we sample for each user 500 random items that have not been previously rated and are not action items. For each user-item pair we compute reachability for a range of stochasticity parameters $\beta \in \{1, 2, 4, 10\}$. Note that across all experimental settings we compute reachability for the same subset of users, but different subsets of randomly selected target items.

We use the ML 1M dataset to primarily gain insights in the role that scoring models, item selection stochasticity and strategic action spaces play in determining the maximum achievable degree of stochastic reachability in a recommender system.

LastFM We run reachability experiment for LibFM recommender with *Next* $K = 10$ action model and stochasticity parameter $\beta = 2$. We compute $\rho^\star$ values for 100 randomly sampled users and 500 randomly sampled items from the set of non-action items (target items can include previously seen items). Unlike the ML 1M dataset, the set of target items is shared among all users.

MIND We run reachability experiments for LibFM recommender with *Next* $K = 10$ action model and stochasticity parameter $\beta = 2$. We compute reachability for all items and users.

Conic Program Implementation

The optimization problem in (2.4) is convex, and we solve it as a conic optimization problem using the MOSEK Python API under an academic license [ApS19]. We reformulate (2.4) as

Num. actions	ML 1M (LibFM)	ML 1M (KNN)	LastFM	MIND
K = 5	0.82 ± 0.04	9.8 ± 3.4	-	-
K = 10	0.87 ± 0.04	10.2 ± 6.1	4.91 ± 0.32	0.44 ± 0.01
K= 20	0.91 ± 0.05	11.4 ± 6.8	-	-

Table 2.3: Reachability run times (in seconds) for a user-item pair. Due to internal representation of action spaces as matrices the runtime dependence on the dimension of the action space is fairly modest. We do not observe significant run time differences between different types of action spaces. We further add multiprocessing functionality to parallelize reachability computations over multiple target items.

an optimization over the exponential cone:

$$\begin{aligned} \min_{t,\mathbf{a},\mathbf{u}} \quad & t - \beta(\mathbf{b}_{ui}^\top \mathbf{a} + c_{ui}) \\ \text{s.t.} \quad & \mathbf{a} \in \mathcal{A}_u, \quad \sum_{j \in \Omega_u^t} u_j \leq 1, \\ & \left(u_j, 1, \beta(\mathbf{b}_{uj}^\top \mathbf{a} + c_{uj}) - t\right) \in \mathcal{K}_{exp} \quad \forall \quad j \in \Omega_u^t \end{aligned} \tag{2.5}$$

The parameters B_u and $\mathbf{c}_u$ are computed for each user based on the scoring model. For the LibFM model, we consider user updates with $\alpha = 0.1$ and $\lambda = 0$. Average run times for computing reachability of a user-item pair in various settings can be found in Table 2.3.

2.7 Results

Impact of Recommender Pipeline

We begin by examining the role of recommender pipeline components: stochasticity of item selection, user action models, and choice of scoring model. All presented experiments in this section use the `ML1M` dataset.

These experiments show that more stochastic recommendations correspond to higher average max reachability values, whereas more deterministic recommenders have a more disparate impact, with a small number of items achieving higher $\rho^\star$. We also see that the impact of the user action space differs depending on the scoring model. For neighborhood based scoring models, strategic manipulations to the history are most effective at maximizing reachability, whereas manipulations of the items most likely to be recommended next are ineffective.

Role of stochasticity We investigate the role of the β parameter in the item selection policy. Figure 2.7 illustrates the relationship between the stochasticity of the selection policy

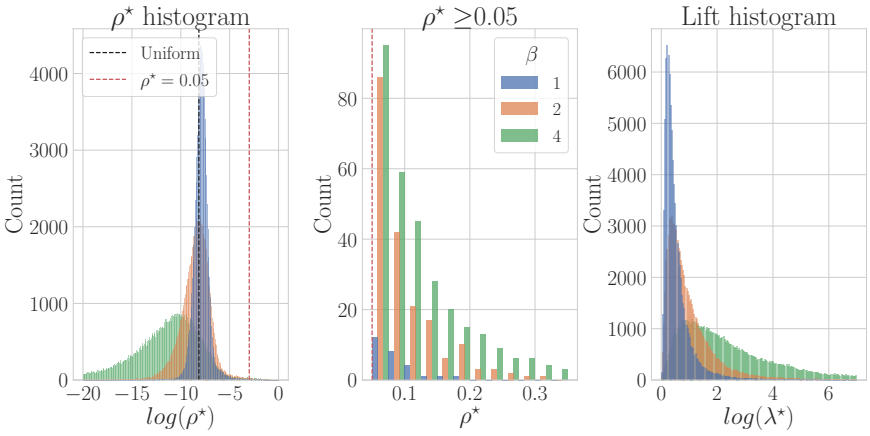


Figure 2.7: Left: Histogram of log max reachability values for $\beta = [1, 2, 4]$. Black dotted line denotes $\rho^\star$ for uniformly random recommender. Center: Histogram of $\rho^\star > 0.05$ (red dotted line). Right: Histogram of log-lifts. Reachability evaluated on `ML1M` for $K = 5$ Random Future action space and a LibFM model.

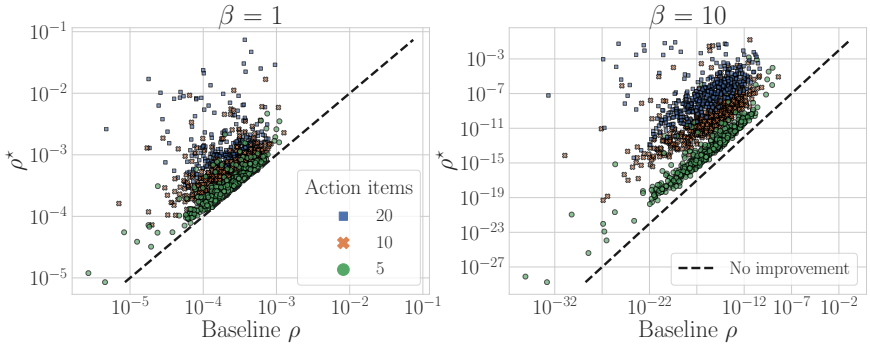


Figure 2.8: Log scale scatterplot of $\rho^\star$ values against baseline ρ for $K \in [5, 10, 20]$. Colors indicate action space size K. We compare low (left) and high (right) stochasticity. Reachability evaluated on `ML1M` for Random Future action space and a LibFM model.

and max reachability. There are significantly more target items with better than random reachability for low values of β. However, higher values of β yield more items with high reachability potential ($> 5\%$ likelihood of recommendation). These items are typically items that are top-1 or close to top-1 reachable. While lower β values provide better reachability on average and higher β values provide better reachability at the "top", higher β uniformly out-performs lower β values in terms of the lift metric. This suggests that larger β corresponds to more user agency, since the relative effect of strategic behavior is larger. However, note that for very large values of β, high lift values are not so much the effect of improved reachability as they are due to very low baseline recommendation probabilities.

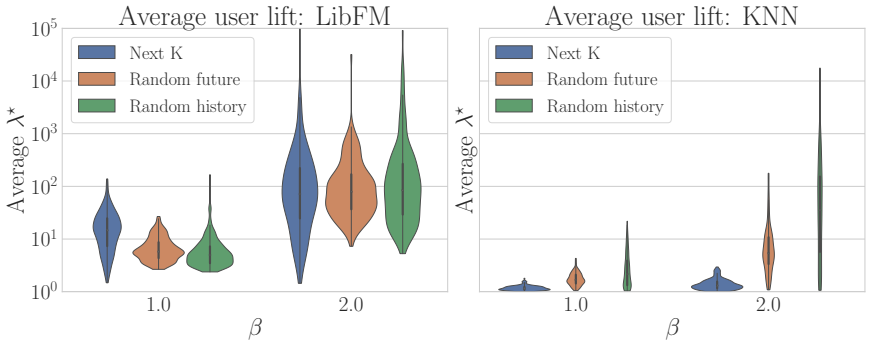


Figure 2.9: The distribution of average lifts (a notion of agency) over users. Colors indicate different user action spaces for LibFM (left) and KNN (right) on `ML1M`.

Role of user action model We now consider different action space sizes. In Figure 2.8 we plot max reachability for target items of a particular user over varying levels of selection rule stochasticity and varying action space sizes. Larger action spaces correspond to improved item reachability for all values of β. However, increases in the number of action items have a more pronounced effect for larger β values.

While increasing the size of the action space uniformly improves reachability, the same cannot be said about the type of action space. For each user, we compute the average lift over target items as a metric for user agency in a recommender (Figure 2.9). For LibFM, the choice of action space does not strongly impact the average user lift, though *Next K* displays more variance across users than the other two. However, for Item KNN, there is a stark difference between *Next K* and and random action spaces.

Role of scoring model As Figure 2.9 illustrates, a system using LibFM provides more agency on average than one using KNN. We now consider how this relates to properties of the scoring models. First, consider the fact that for LibFM, there is higher variance among user-level average lifts observed for *Next K* action space compared with random action spaces. This can be understood as resulting from the user-specific nature of *Next K* recommended items. On the other hand, random action spaces are user independent, so it is not surprising that there is less variation across users.

In a neighborhood-based model users have leverage to increase the ρ reachability only for target items in the neighborhood of action items. In the case of KNN, the next items up for recommendation are in close geometrical proximity to each other. This limits the opportunity for discovery of more distant items for *Next K* action space. On the other hand, the action items are more uniformly over space of item ratings in random action models, thus contributing to much higher opportunities for discovery. Additionally, see that *History Edits* displays higher lift values than other action spaces due to the fact that for this action space the strategic ratings are a larger proportion of the total ratings.

dataset	model	prevalence	baseline availability	max availability
ML1M	libfm	0.346280	0.827492	0.501316
ML1M	knn	0.346280	0.949581	0.942986
mind	libfm	0.863992	0.825251	0.435212
lastfm	libfm	0.133318	0.671101	0.145949

Table 2.4: Spearman's rank correlation between item popularity and item prevalance, item popularity and baseline availability ($\beta = 2$) and item popularity and max availability (*Next* K action space where $K = 10$).

For completeness we present full experimental results over different datasets, scoring models, selection rules and action spaces in Figure 2.15, Figure 2.16 and Figure 2.17.

Bias in Movie, Music, and News Recommendation

We futher compare aggregated stochastic reachability with properties of the user and the items to investigate bias. We aggregate baseline and max reachability to compute user-level metrics of discovery and item-level metrics of availability. The audit demonstrates popularity bias for items with respect to baseline availability. This bias persists in the best case for neighborhood based recommenders and is thus unavoidable, whereas it could be mitigated for MF recommenders. User discovery aggregation reveals inconclusive results with weak correlations between the length of users' experience and their ability to access content.

Popularity bias

To systematically study the popularity bias, we compute the Spearman rank-order correlation coefficient to measure the presence of a monotonic relationship between popularity (as measured by average rating) and availability (either in the baseline or best-case scenario as defined in in (2.3)). We also compute the correlation between popularity and prevalence in the dataset, as measured by the number of ratings.

For instance in the case of movie recommendation (ML1M dataset) for both LibFM and KNN models, the baseline availability displays a correlation with item popularity, with Spearman's rank-order correlations of $r_s = 0.83$ (LibFM) and $r_s = 0.95$ (KNN) respectively (see Table 2.4. This suggests that as recommendations are made and consumed, more popular items will be recommended at disproportionate rates.

In terms of best-case availability, KNN displays a similar correlation between availability and popularity ($r_s = 0.94$). This suggests that KNN recommenders display unavoidable systematic biases meaning that regardless of user actions popularity bias will propagate in time through recommendations. This does not necessarily hold for LibFM, where the best

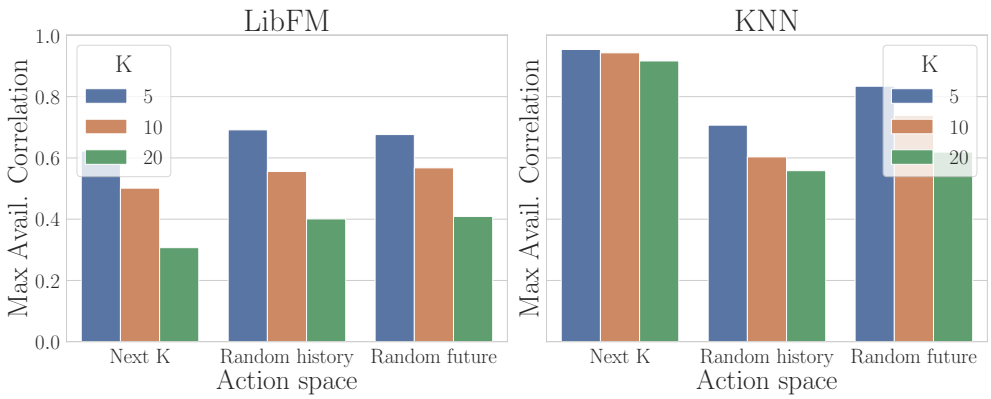


Figure 2.10: Comparison of Spearman's correlation between item popularity and max availability for different action spaces and models. Reachability evaluated on `ML1M` with $\beta = 2$.

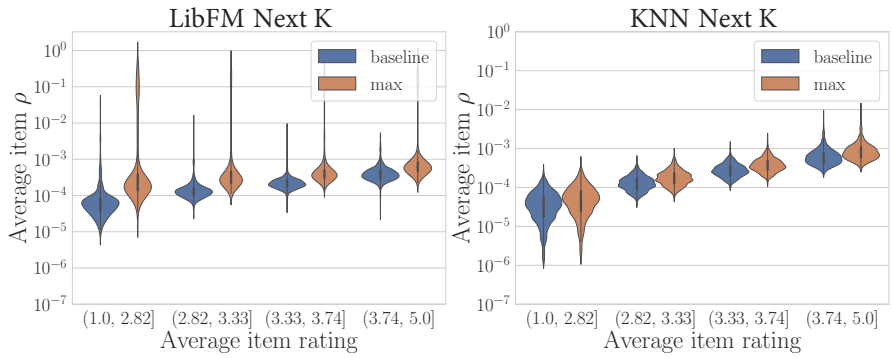
case availability is less clearly correlated with popularity ($r_s = 0.50$), which means that any biases observed in deployment are due in part to user behavior.

The impact of user action spaces is shown in Figure 2.10, which plots the correlation between popularity and best case availability for different action spaces. For comparison, the correlation between popularity and baseline availability is just over 0.8 for all of these settings. Increasing the action space (higher K) uniformly reduces correlations between availability and popularity. The type of action space does not play a major role for LibFM, whereas for KNN the availability bias is dependent on the action space. The *Random History* action space allows for considerable decreases in unavoidable popularity bias. The findings are further corroborated by Figure 2.11

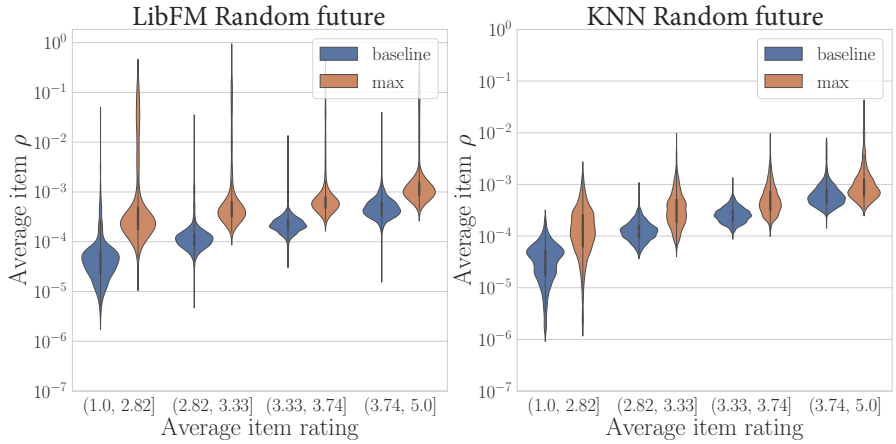
Lower correlation for LibFM best case availability holds in the additional settings of music artist and news recommendation as shown in Figure 2.12.

Experience bias

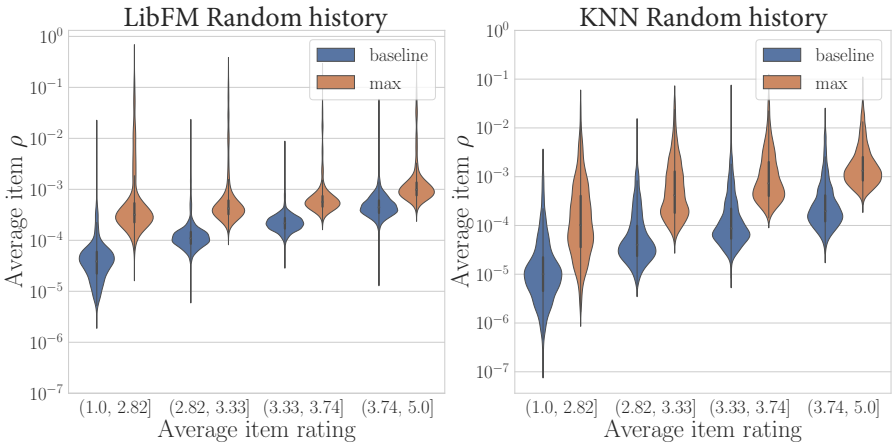
To investigate the impact of experience on user discovery we similarly compute the Spearman rank-order correlation coefficient to measure the presence of a monotonic relationship between user experience (as measured by number of items rated) and discovery (either in the baseline or max case as defined in in (2.3)). We investigate experience bias by considering how the discovery metric changes as a function of the number of different items a user has consumed so far. We observe correlation values of varying sign across datasets and models, and none are particularly strong (Table 2.5). Figure 2.13 illustrates that experience is weakly correlated with baseline discovery for movie recommendation ($r_s = 0.48$), but not so much for news recommendation ($r_s = 0.05$). The best case discovery is much higher, meaning that users have the opportunity to discover many of their target items. However, the weak correlation with experience remains for best case discovery of movies ($r_s = 0.53$).



(a) *Next K* action space



(b) *Random Future* action space



(c) *Random History* action space

Figure 2.11: Comparison of baseline and best case availability of content, across four popularity categories for LibFM (left) and KNN (right) scoring models with soft-max selection policy parameterized by $\beta = 2$. Reachability evaluated on `ML1M` for *Next K Random Future* and *Random History*, action space with $K = 10$.

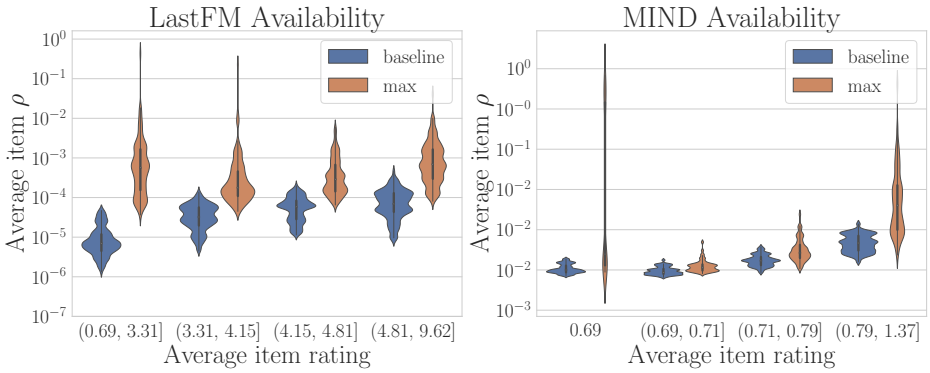


Figure 2.12: Comparison of baseline and best case availability of content for four popularity categories for LastFM (left) and MIND (right) with *Next 10* actions, LibFM model, and $\beta = 2$.

dataset	model	baseline discovery	max discovery
`ML1M`	libfm	0.475777	0.530359
`ML1M`	knn	0.206556	-0.031929
mind	libfm	0.050961	0.112558
lastfm	libfm	-0.084130	-0.089226

Table 2.5: Spearman's rank correlation with experience for *Next K* with $K = 10$ and $\beta = 2$.

Gender bias

Finally, we investigate gender bias. We compare discovery across user gender for `ML1M` and availability across artist gender for LastFM (Figure 2.14). We do not observe any trends in either baseline or max values.

2.8 Discussion

In this chapter, we generalize reachability as first defined by [DRR20] to incorporate stochastic recommendation policies. We show that for linear scoring models and soft-max item selection rules, max reachability can be computed via a convex program for a range of user action models. Due to this computational efficiency, reachability analysis can be used to audit recommendation algorithms. Our experiments illustrate the impact of system design choices and historical data on the availability of content and users' opportunities for discovery, highlighting instances in which popularity bias is inevitable regardless of user behavior.

The reachability metric provides an upper bound for discovery and availability within a recommendation system. While it has the benefit of making minimal assumptions about

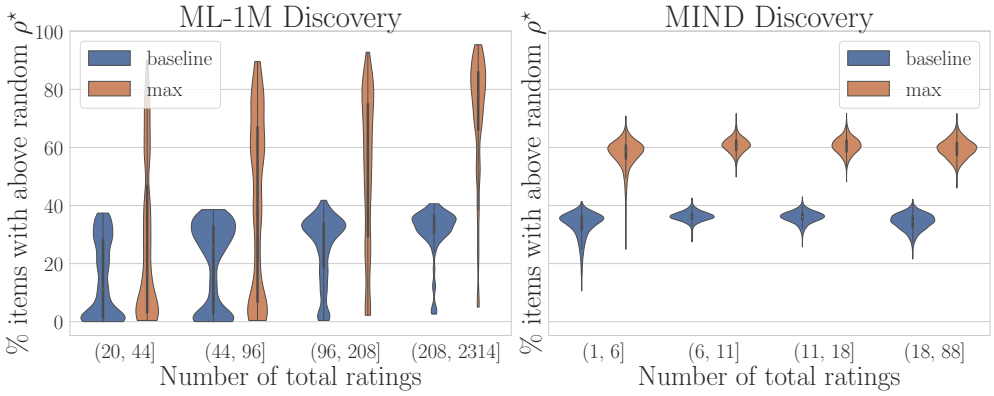
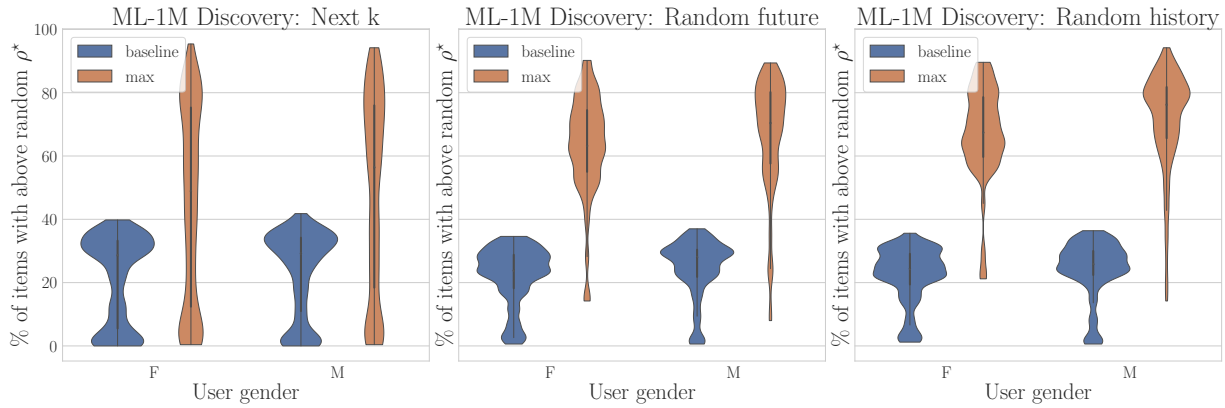


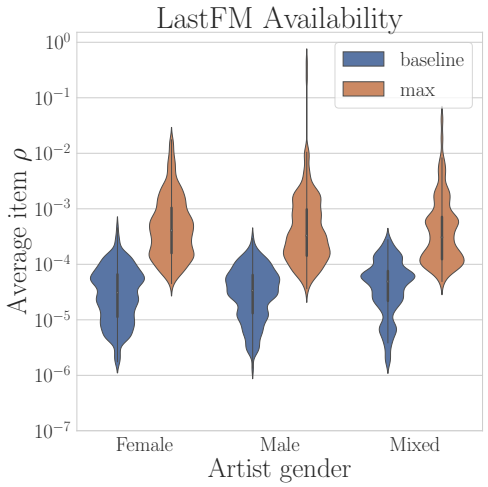
Figure 2.13: Comparison of baseline and best case fraction of items with better than random $\rho^\star$, grouped across four levels of user history length. Reachability evaluated on `ML1M` (left) and MIND (right) for *Next 10* action space, $\beta = 2$, and LibFM model.

user behavior, the drawback is that it allows for perfectly strategic behaviors that would require users to have full knowledge of the internal structure of the model. The results of a reachability audit may not be reflective of probable user experience, and thus reachability acts as a necessary but not sufficient condition.

Nonetheless, reachability audit can lead to actionable insights by identifying inherent limits in system design. They allow system designers to assess potential biases before releasing algorithmic updates into production. Moreover, as reachability depends on the choice of action space, such system-level insights might motivate user interface design: for example, a sidebar encouraging users to re-rate K items from their history. Furthermore, our results on the lack of a trade-off between accuracy and reachability are encouraging. Minimum one-step reachability conditions could be efficiently incorporated as a design requirement into learning algorithms for scoring models.



(a) Baseline and maximum discovery across user gender evaluated on `ML1M` dataset for different action spaces ($K = 10$



(b) Availability across artist gender evaluated

Figure 2.14: Side by side comparison of baseline and maximum discovery across user gender and availability across artist gender (rightmost panel). Reachability evaluated for LibFM model with soft-max parameter $\beta = 2$.

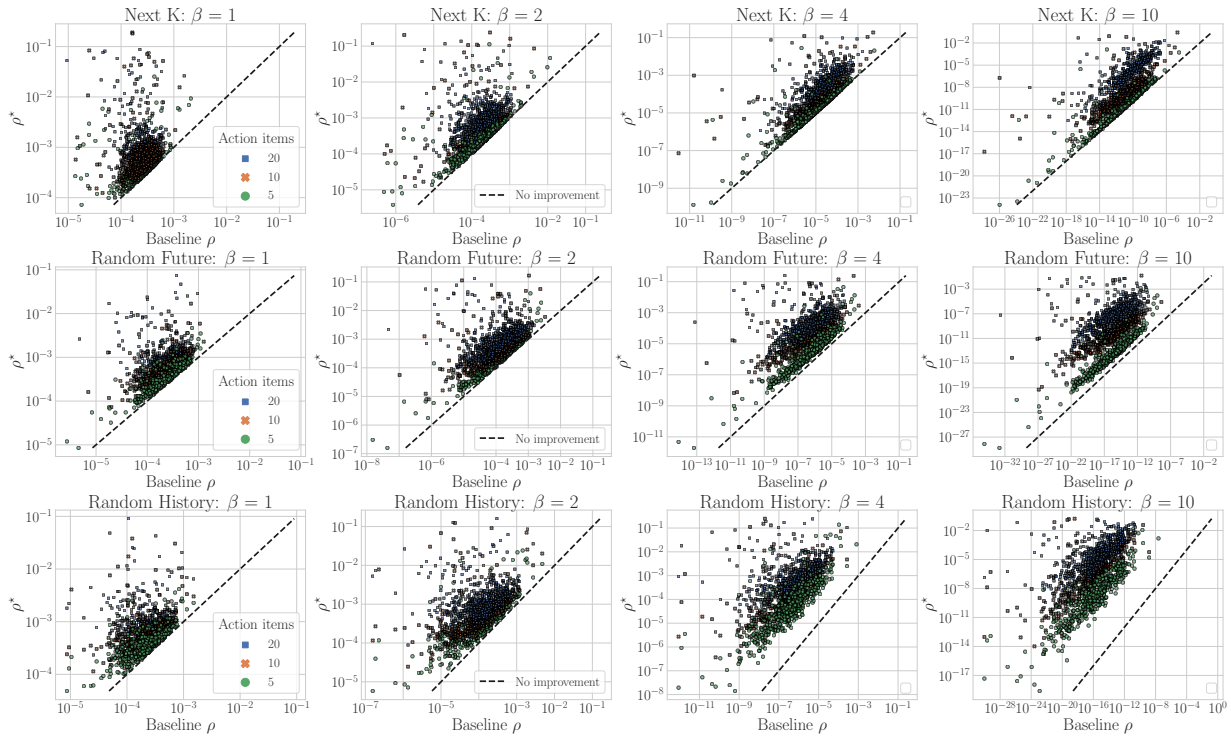


Figure 2.15: Log scale scatterplots of $\rho^\star$ against baseline ρ evaluated for the LibFM scoring model.

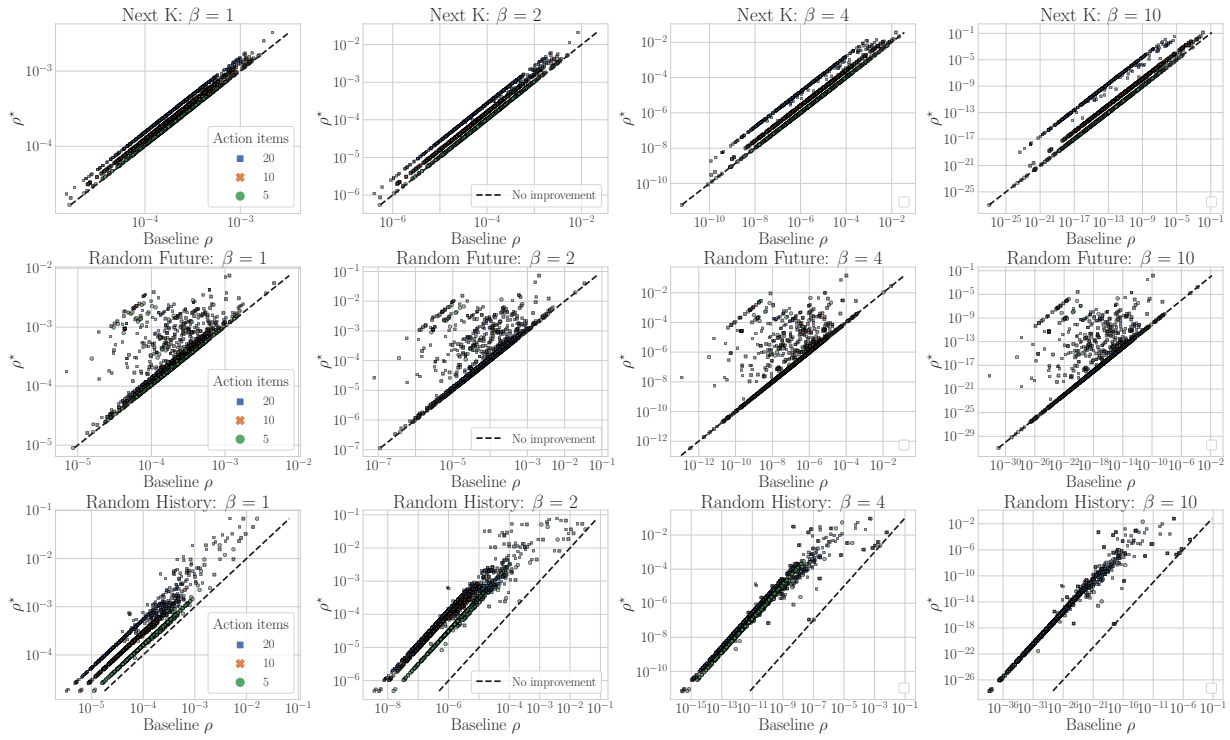


Figure 2.16: Log scale scatterplots of $\rho^\star$ against baseline ρ evaluated for the KNN scoring model.

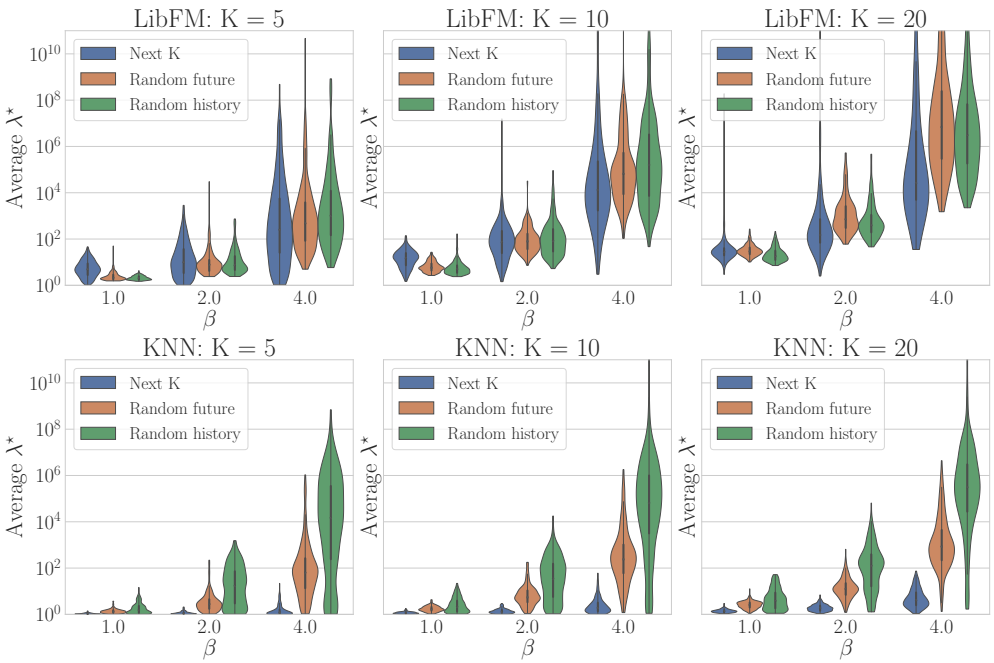


Figure 2.17: Side by side comparison of average user lifts for LibFM (top row)and KNN(bottom row).

Chapter 3

Modeling psychologically-grounded preference shifts in content recommenders

In this chapter, we transition from single-step interactions to multi-step closed-loop dynamics, which require modeling user behavior. While the benefit of the single-step reachability approach lies in its minimal behavioral assumptions, the drawback is that its results rely on the implicit assumption that users are perfectly strategic with full knowledge of the internal model structure. And such, the results of a reachability audit may not accurately reflect realistic user experiences.

We begin with the premise that modeling the influence of recommendations on people's preferences must be grounded in psychologically plausible models. We contribute a methodology for developing grounded dynamic preference models and demonstrate this method with three classic effects from the psychology literature: Mere-Exposure, Operant Conditioning, and Hedonic Adaptation. We conduct simulation-based studies to show that the psychological models manifest distinct behaviors that can inform system design. This has two direct implications for dynamic user modeling in recommendation systems. First, the methodology we outline is broadly applicable for psychologically grounding dynamic preference models. It allows us to critique recent contributions based on their limited discussion of psychological foundation and their implausible predictions. Second, we discuss implications of dynamic preference models for the evaluation and design of recommendation systems.

This chapter is based on the paper "Towards psychologically-grounded dynamic preference models"[Cur+22] written in collaboration with Andreas Haupt, Benjamin Recht and Dylan Hadfield-Menell.

3.1 Background

In much of recommendation systems research, preferences are implicitly or explicitly assumed to be static throughout time and not altered by recommendation. While useful in many contexts, this assumption has drawbacks. For example, static preferences in user models

may account for the poor generalization performance of offline metrics in predicting deployed recommendation performance [RSZ16; Gar+14; Jeu19; Kra+20]. Furthermore, dynamic preferences have been central to arguments of negative impacts of recommendation, such as polarization, filter bubbles, and extremism [RPF18; Jia+19; GLH19; CSE18].

This chapter considers the design and validation of dynamic user preference models. We argue that behavioral models should be designed with, at least, two desiderata.

First, they should be grounded in experimental psychological evidence. A wealth of research in psychology identifies several such behavioral patterns, or *psychological effects*. In what follows we propose a methodology to leverage these results and design psychologically plausible dynamic preference models whose predictions are compatible with findings in psychology. We demonstrate our method by designing models to capture three established findings about the formation and dynamics of preferences: *Mere Exposure*, [HTW13; Nor02; FSA07; CC02], where simply experiencing a stimulus (e.g., a piece of content) tends to make humans view it more positively; *Operant Conditioning* [FAF11; Hel47; Fox17], where preferences shift towards actions that are associated to "positive feedback" (e.g., consuming "better than expected" content); and *Hedonic Adaptation* [CIR15; NM08; YG15], where satisfaction levels (e.g., with content) return to a baseline after a period of time.

Second, models should be verified in simulation for plausibility. For any given effect, there are many ways it could be formalized mathematically. Simulations of user-recommender dynamics can help with these modeling choices. Behavioral models should produce plausible predictions across a range of recommender system designs and initial conditions. For example, a model where behaviour shifts to consuming a single type of content with probability close to one is unlikely to match real observations.

Having a combination of predictions from simulations and background from psychology feeds back into improving user modeling and hence recommendation system evaluation and design. This may help address the on-/offline gap in evaluation as well as increase the credibility of the theoretical study of societal impacts of recommendation.

Contributions

- **Case studies:** We consider three case studies of behavioral effects due to consumption established in psychological research.[1] In Section 3.2 we review the relevant literature

[1]While the models we present are classical in psychology, we do not claim that these are models will fit empirical data in recommendation systems well. Our mathematical formalization of theories is particularly parsimonious by formulating preference shifts as changes to vector of user factors. This parsimony, however, leads to simplifications of several potentially psychologically relevant factors, which we do not claim to exhaustively model. First, our models do not differentiate between exposure to and consumption of content. Furthermore, they assume that preferences are purely evolving instead of formed, i.e. users have preferences for even unseen content. In addition, the models are formulated abstractly, as opposed to for a concrete application domain, suppressing features that might make different psychological effects particularly salient in different domains. Our models are exemplary for the application of a methodology, and will need empirical validation before used to model interactions in a real systems. The selection of good models, we claim, however, benefits from inclusion of psychological grounding and extensive testing.

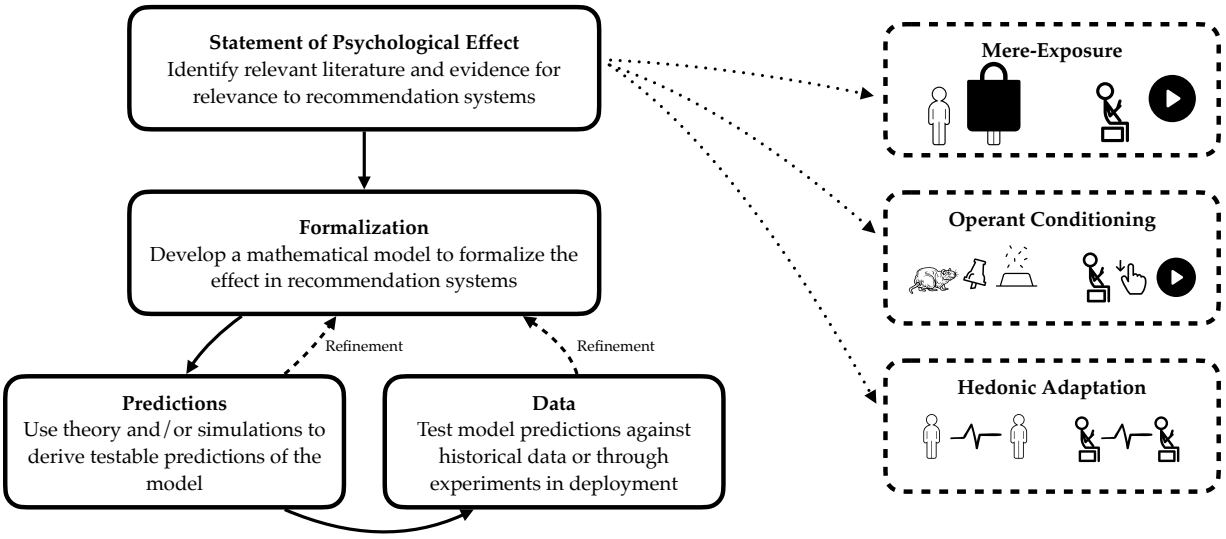


Figure 3.1: The methodology derived in this article. First, formulate a psychological theory that should ground a particular effect. Then, formalize it, and discipline it with predictions derived from simulations and data from deployed recommendation systems. This article exemplifies this methodology for three effects: Mere-Exposure Effect, which we introduce using Goetzinger's [Zaj68] classical experiment, Operant Conditioning, which we introduce using Skinner's experiments [Ski38], and hedonic adaptation, which we introduce using Brickman's study of happiness after extreme life events [Bri71]. Vignettes on the right depict these classical experiments.

of these effects with a focus on their applicability to recommendation systems. In Section 3.3 we propose a modeling framework to mathematically formalize these effects. In Section 3.4, using simulations, we characterize qualitative properties of our proposed models and derive testable predictions that can be used to test the validity of a particular mathematical formalization independently of the applicability of a behavioral effect. Finally, we discuss observational and experimental approaches to validating psychologically-plausible dynamic preference models.

- **Methodology:** We synthesize the case studies into a broadly applicable methodology for psychologically grounding dynamic preference models in Section 3.5. We propose a multi-step procedure which first calls for explicitly stating and providing evidence in support of a psychological effect. Following this, a designer iterates by proposing a mathematical formalization for the behavioral model, deriving testable predictions and comparing them against data relevant to the recommendation context. The proposed methodology is depicted in Figure 3.1.

- **Analysis:** Finally, we discuss two implications of this study for recommendation system evaluation and design. First, in Section 3.5 we critique recent contributions

based on their limited discussion of psychological foundations and their implausible predictions. Second, in Section 3.6 we demonstrate in an example that recommendation system metrics such as engagement and diversity are unable to capture desirable recommendation system performance in the presence of dynamic user preferences.

Related Work

Evaluation First we relate to literature of recommendation system evaluation. Changes in preference in response to recommendations can affect the evaluation of recommendation systems. The main design paradigm of recommendation systems is to use offline train-test splits to design and select recommendations. Several papers argue that offline metrics are poor indicators of online performance [Jeu19; Gar+14; Bee+13; Sun+20]. Often, the lack of external validity of the current evaluation methodology is attributed to the unmodeled feedback dynamics between users and recommenders [RSZ16; Kra+20], further motivating the need to study dynamic preference models.

Societal Impacts Well-founded dynamic preference models can help resolve the apparent contradictions between findings on the social impacts of recommendation systems. On the one hand, several studies claim that the dynamic interaction of recommendation systems with users can lead to polarization [DGL13; RPF18], filter bubbles [Par11; GLH19], homogenization [CSE18], echo chambers [Noo+20; DZ21; Jia+21], and extremism [Mun19; Rib+20; Jia+19]. However, empirical audits find that algorithmically recommended content is more diverse than natural consumption [Ngu+14], and that real systems do not exhibit the strong extremism or polarization effects implied by theoretical models [MG21; LZ19; Hos+20; Tom+21; Lev21]. The study [Car+21] points out that recommendation systems might lead to undesired changes in preferences, and proposes to design for *safe preference shifts*, which are preference trajectories that are deemed "desirable".

Dynamic Preference Models Some of the existing dynamic preference models assume that preferences change independently of recommendations. Examples are Dynamic Poisson Factorization (DPF) [Cha+15; Hos+18], which assumes that users and content items have latent representations that evolve as Gaussian random walks; and [Kor09], which models behavioral changes as external concept drifts which affect item quality and average user rating levels. Other works consider the feedback loop between recommendation systems and users directly. [RPF18] models 1-dimensional opinion dynamics on a sphere, [Jia+19] and [Kal+21] propose models of behavior shift due to content exposure and content consumption, respectively. Each of these theoretical contributions predicts polarization and extremism of user preferences.

Psychology-Informed Recommendation Recent surveys [JJ21; Lex+21] cover ways in which recommenders incorporate and account for psychological effects. The work of [Lex+21]

focuses mainly on ways in which affective, cognitive, and personality factors impact user engagement, and it does not cover the feedback loop between recommendation systems and user preferences. [JJ21] reviews psychologically-informed recommenders for behavioral priming and digital nudging; it points out a lack of research into psychological foundations of dynamic preference models.

3.2 Psychological Effects

This section introduces three psychological effects, Mere-Exposure, Operant Conditioning, and Hedonic Adaptation. For each effect, we will first introduce a classical example of the effect, give general references, and then provide experiments most relevant to recommendation systems. While we find strong connections for the Mere-Exposure and Hedonic Adaptation effects, we will introduce a third effect, Operant Conditioning, whose experimental evaluation allows for a less straightforward connection to recommendation systems. As a running example in this section, we will consider Alice, a user that initially does not like sports content but it is exposed to it by a recommendation system.

Mere-Exposure

Mere-Exposure or the familiarity effect says that humans tend to like more what they are exposed to more often. A classical experiment, cited in [Zaj68], is the Black Bag Experiment. In 1968 at Oregon State University, C. Goetzinger let a person fully covered with a black bag participate in a course. Other students in the course were hostile at first, but later became friendly towards the person covered with the bag. In a recommendation system context, Mere-Exposure may mean that reactions to content become more favorable after exposure to and/or consumption of similar content. In the case of Alice, whose initial preferences are not favoring sports, a Mere-Exposure effect would predict that with repeated exposure or consumption she becomes more familiar with sports content and starts to appreciate it more.

Mere-Exposure effects are well-established in psychology. [Bor89] conducts a meta-study of 200 studies of research in the first three decades following its introduction in [Zaj68]. Much of the research on Mere-Exposure effects is using experimental setups in which subjects are exposed to non-meaningful stimuli, such as Japanese characters for non-Japanese speakers [Dec+09]. While this allows to control for prior exposure to the content, such research is not directly meaningful for understanding the Mere-Exposure effect in recommendation systems.

Most relevant to recommendation systems is Mere-Exposure research on advertisement and audiovisual content. An illustrative example of this kind is [HTW13, Experiment 1]. In it the experimenter repeatedly shows users images of fictitious, but plausible, products (e.g., a smoke filter) of different aspect ratios. Users' reported rating of attractiveness of aspect ratios of products increased significantly with the number of times it has had been shown to the user. On average each additional exposure led to a 0.2 points increase on a 7-point scale. (The experiments in [Nor02; FSA07] for the exposure to banner ads and [CC02] for

exposure to pictures of fashion designs made similar findings.) The study found as well that "conspicuous" exposure, i.e. exposure that draws attention to the unfamiliarity of the product, leads to a smaller Mere-Exposure effect (0.03 points increase per additional exposure).

Operant Conditioning

Operant Conditioning is the effect that beings tend to engage more in activities that are associated with positive stimuli (positive reinforcement) and avoid activities associated with negative stimuli (negative reinforcement). Evidence for this effect comes primarily from animal studies. In a classical experiment from [Ski38], B.F. Skinner puts a hungry rat into box containing a food dispenser and a button. Food is released whenever the button is pressed. If the rat does not press the button, it gets a small electric shock. As an observation of this experiment, the rat presses the button more and more frequently. Translated to recommendation systems, Operant Conditioning predicts higher engagement with content that was "surprisingly good" and less that is associated with content that was "surprisingly bad". If the initially sports-averse Alice reacts according to Operant Conditioning, depending on her baseline level, she might be underwhelmed by the sports content, and dislikes it more after being exposed to it, or her baseline is very low, in which case she might start liking it.

Operant Conditioning is well-documented in Behavioral Psychology, both in animals and humans. Skinner [Ski38; FS57] conducted several studies in animals which finds evidence for Operant Conditioning. To our knowledge the first study of Operant Conditioning in humans is [Ful49] which conditioned a young man with a developmental disorder to raise his arm. We refer the reader to the monograph [CHH+07, Chapter 11-13] for a treatment of Operant Conditioning and the following behavioral movement in psychology.

While much of the work on Operant Conditioning is conducted in animal experiments, work that most closely resembled a recommendation systems is in consumer psychology. [Fox04] (see surveys [Fox10; Fox17] for follow-up work) introduced the Behavioral Perspectives Model, which classifies consumer choices into several "reinforcers" which might depend on the product-dependent, social, or monetary (shopping online is often costly) consequences of making consumption decisions. As an exemplary experiment, [FAF11] considers the effect of externally provided reinforcers (e.g. shipping cost, shipping duration, or price) on consumer choice among two different online shops. [FAF11] report that a significant fraction of the subjects followed the (positive) reinforcements set by the experimenter. We note that an interpretation of Operant Conditioning for dynamic *preference* models is challenging. On the one hand, behavioralism views behavior purely as a black box, which is why the effect is framed around behavior, not preferences. Our translation to preferences requires assuming that Operant Conditioning may lead to changes in preferences. Second, the definition of baseline we will consider in the quantitative model below depends on past preferences, which is closer to literature on adaptation, e.g. the quantitative model of [Hel47], than of Operant Conditioning.

Hedonic Adaptation

Hedonic Adaptation is the effect that after some time, any change in happiness fades, and humans return to a baseline level. In a classical study [Bri71], P. Brickman asked lottery winners and paraplegics for their happiness and how they expect their happiness level to be in a year, finding that major life events had negligible effect on their happiness. Hedonic Adaptation means that engagement with content returns to a baseline level after some time. If Alice's "baseline self" does not like sports content, she will return to this baseline irrespective of the content recommended—which she also would if her "baseline" self likes sports.

[Kah+99, ch. 16] reviews many earlier findings on the human adaptation to repeated exposure to noise (inconclusive evidence) as well as incarceration and increased income (supporting evidence).

In Hedonic Adaptation literature most relevant to recommendation systems, studies consumption scenarios. [CI20]'s experiment lets students choose a sticker and attach it to an everyday object. Eliciting self-reported happiness with the sticker, the study finds a 4.5 point decline on a 100-point scale of reported happiness with a sticker in a 3-day interval. [YG15] finds a loss between 1.75 point ("low sentimental value") to 7.75 point ("high sentimental value") decrease on a 100-point scale for Google Image search result shown to users for 6 short intervals of 10 seconds. [NM08] plays songs and asks for reports of happiness with the song at different points. Still on a 100-point scale, the preference is reduced by 15 points from 10 seconds into the song to one minute into the song. Finally, [CIR15] showed paintings to subjects in three 15-second exposure intervals. Exposure led to a reduction in happiness with the painting of about 12 points on a 100-point scale.

All of these studies have relatively short exposure times. However, content types are comparable to many contemporary recommendation systems, and demonstrated effects on self-reported happiness are quite substantial.

3.3 Formalizations of Behavioral Models

In this section, we propose mathematical formalizations of Mere-Exposure, Operant Conditioning, and Hedonic Adaptation. We start with our basic notation.

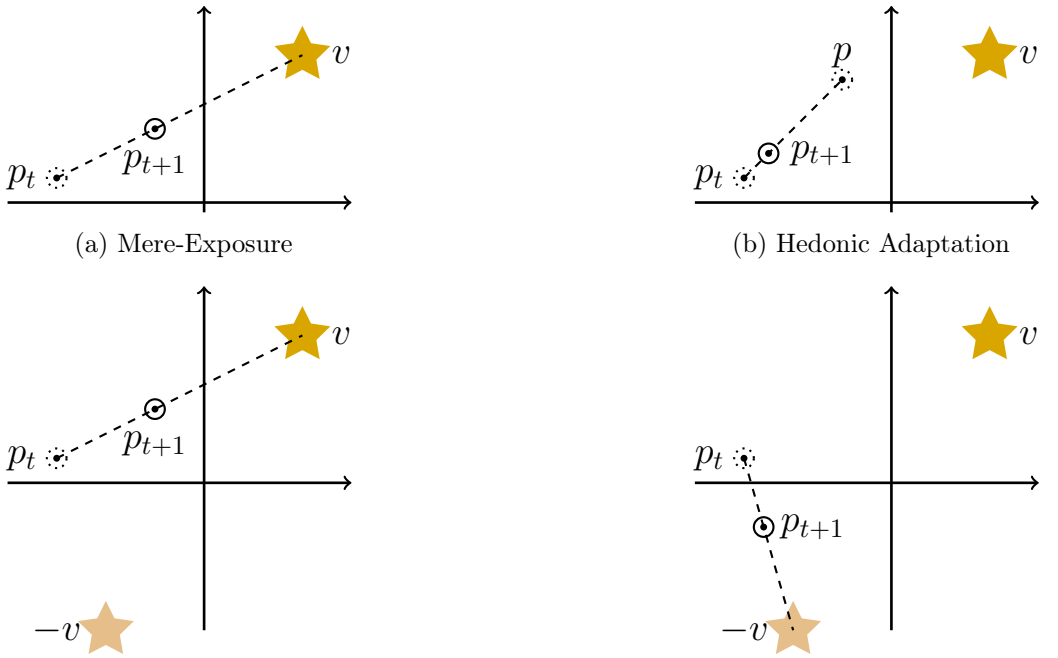
At each round, a recommendation system recommends one of N pieces of content, each with an associated d-dimensional item vector. Throughout, we assume that the item vectors are fixed.[2] Since the item representations are known and fixed, it is without loss to consider a single user at a time. Denote by $\mathbf{p}_t \in \mathbb{R}^d$ the preference vector of the user at time t. The user reacts to a piece of content with item vector $\mathbf{v} \in \mathbb{R}^d$ according to a rating function, which we assume to be linear, as is common in collaborative filtering, $r(\mathbf{p}_t, \mathbf{v}) \coloneqq \langle \mathbf{p}_t, \mathbf{v} \rangle + \varepsilon_t \in \mathbb{R}$ for

[2]This assumption is restrictive, but is justified in content-based recommenders, e.g., when content is featurized in terms of topic, genre, length, political inclination, etc., or high-dimensional content embeddings of images, videos or text based on supervised or unsupervised learning. The assumption that item vectors are fixed also holds approximately in Collaborative Filtering settings if item representations are updated less frequently than user representations.

independent noise ε_t, on which we will make additional distributional assumptions below. At each time step t, the recommender chooses a piece of content with associated latent representation $\mathbf{v}_t$.

In contrast to static recommendation system models, we model user preferences change due to content exposure and consumption: $\mathbf{p}_{t+1} - \mathbf{p}_t = f(\texttt{Hist}_t)$ for some user-item history $\texttt{Hist}_t = (\mathbf{p}_1, \mathbf{v}_1, \mathbf{p}_2, \mathbf{v}_2, \ldots, \mathbf{p}_t, \mathbf{v}_t)$, and a potentially random function f. For a static user, $f(\texttt{Hist}_t)$ is constant 0.[3]

Next, we propose quantitative models for the psychological effects introduced in Section 3.2.



(a) Mere-Exposure

(b) Hedonic Adaptation

(c) Operant Conditioning (positive)

(d) Operant Conditioning (negative)

Figure 3.2: Updates in preference space for Mere-Exposure, Operant Conditioning and Hedonic Adaptation. Mere-Exposure moves preference vectors a constant fraction of the distance towards the recommended content; Operant Conditioning either moves towards, or away, depending on the direction and magnitude of *surprise*. Hedonic Adaptation leads to convergence towards a baseline rating.

[3]Note that our model does not differentiate between the effects of exposure to and consumption of content. A more general model would model consumption probabilities, and allow for separate effect strengths of exposure and consumption.

Mere-Exposure

For Mere-Exposure, we consider the linear model

$$\mathbf{p}_{t+1} - \mathbf{p}_t = \gamma f_{\text{ME}}(\mathbf{v}_t, \mathbf{p}_t) = \gamma(\mathbf{v}_t - \mathbf{p}_t) \tag{3.1}$$

for some $\gamma \in [0, 1]$, compare Figure 3.2a. Whenever an element $\mathbf{v}_t$ is shown, the preference vector moves a γ-fraction of the line between $\mathbf{u}_t$ and $\mathbf{v}_t$. This captures the idea that whenever users are exposed to content, this makes them like this content more.

Our model is parameterised by a single quantity that determines how much a user moves into a certain direction. The strength of preference movement might depend, given the psychology literature reviewed above, on the conspicuousness of content exposure.

Operant Conditioning

The space of possible models that capture Operant Conditioning is large, and we consider a particular parameterization, highlighting some of the qualitative features of positive and negative reinforcement. The model we consider here captures that users will adjust their reaction to content that surprises them:

$$\mathbf{p}_{t+1} - \mathbf{p}_t = \gamma f_{\text{OC}}(\texttt{Hist}_t) = \gamma|\texttt{surp}(\texttt{Hist}_t)|(\texttt{sgn}(\texttt{surp}(\texttt{Hist}_t))\mathbf{v}_t - \mathbf{p}_t)$$

for $\gamma \in [0, 1]$, compare Figure 3.2c for the case of positive reinforcement, $\texttt{sgn}(\texttt{surp}(\texttt{Hist}_t)) = 1$, and Figure 3.2d for the case of negative reinforcement, $\texttt{sgn}(\texttt{surp}(\texttt{Hist}_t)) = -1$. The magnitude of the preference shift is scaled by the size of a *surprise* term, and the direction of the change is determined by the sign of the surprise. If the surprise is positive then the preference moves in the direction of the item v. Conversely, if the surprise is negative then the preference moves towards $-\mathbf{v}$.

The surprise is a function of difference between a baseline level of engagement and the current rating. We model the expected engagement as an discounted average of historical ratings, and use an arctan function to map to the range $[-1, 1]$. This yields the surprise term of the form:

$$\texttt{surp}(\texttt{Hist}_t) := \arctan\left(\frac{\sum_{\tau=1}^{t-1} \delta^\tau r_{t-\tau}}{\sum_{\tau=1}^{t-1} \delta^\tau} - r_t\right), \delta \in [0, 1].$$

The choice of an exponential decay is motivated in psychology and neuroscience, compare, e.g., habituation [Mar09].

Hedonic Adaptation

For hedonic adaptation, we propose a fairly simple model: a linear drift towards a constant *baseline preference vector* $\mathbf{p}$:

$$\mathbf{p}_{t+1} - \mathbf{p}_t = \gamma f_{\text{HA}}(\mathbf{p}_t) = \gamma(\mathbf{p} - \mathbf{p}_t)$$

for $\gamma \in [0, 1]$, compare Figure 3.2b. The update moves the user towards a (fixed) baseline preference $\mathbf{p} \in \mathbb{R}^d$. Note that this behavioral shift is irrespective of the recommended content $\mathbf{v}_t$.

A shared property of all proposed models is that they do not lead to arbitrarily large preference vectors.

Proposition 3.3.1. *Let* $f \in \text{conv}(f_{OC}, f_{HA}, f_{ME}, 0)$*; i.e* $\mathbf{p}_{t+1} - \mathbf{p}_t = \gamma_{ME} f_{ME}(\mathbf{p}_t, \mathbf{v}_t) + \gamma_{OC} f_{OC}(\mathbf{p}_t, \mathbf{v}_t) + \gamma_{HA} f_{HA}(\mathbf{p}_t)$ *where* $\gamma_{ME} + \gamma_{OC} + \gamma_{HA} \leq 1$*. Then for any sequence of recommendations,* $\mathbf{p}_t$ *is bounded.*

Proof. Denote $-V = \{-\mathbf{v} | \mathbf{v} \in V\}$ the negations of item factors. Note that the following statement, which we prove by induction in $t \in \mathbb{N}$, is sufficient to prove boundedness:

Claim. $\mathbf{p}_t \in \text{conv}(\{\mathbf{p}_0\} \cup V \cup -V)$.

For $t = 0$, the claim is trivially satisfied. Assume hence that $\mathbf{p}_t \in \text{conv}(\{\mathbf{p}_0\} \cup V \cup -V)$. As for any sets A, B, $\text{conv}(\text{conv}(A)) = \text{conv}(A)$ and $A \subseteq B \implies \text{conv}(A) \subseteq \text{conv}(B)$, we have,

$$a \in \text{conv}(A) \text{ and } A \subseteq \text{conv}(B) \implies a \in \text{conv}(B). \tag{3.2}$$

Applying this statement for $a \coloneqq \mathbf{p}_{t+1}$, $A \coloneqq \{\mathbf{v}_t, \mathbf{p}_t\}$ and $B \coloneqq (\{\mathbf{p}_0\} \cup V \cup -V)$. $A \subseteq \text{conv}(B)$ holds by the induction hypothesis. $a \in \text{conv}(A)$ holds by the the definition of Mere-Exposure, Operant Conditioning and Hedonic Adaptation, as well as $\texttt{surp} \in [-1, 1]$. This concludes the induction step. Hence, the claim, and hence boundedness, holds. □

3.4 Testing Plausibility in Simulation

This section presents simulation results for the formalizations $f_{\text{OC}}, f_{\text{ME}}$, and f_{HA} of Operant Conditioning, Mere-Exposure, and Hedonic Adaptation. We use our results to compare against psychological evidence.

Experimental Setup

User behavior We sample N i.i.d item vectors from a multivariate normal distribution, $\mathbf{v}_i \sim \mathcal{N}(0, \sigma I_d)$. The initial user preference $\mathbf{p}_0$ is sampled from the same distribution.

At time t, the recommender selects an item $\mathbf{v}_t$ based on estimated scores $s_{it} = \langle \mathbf{u}_t, \mathbf{v}_i \rangle, i = 1, 2, \ldots, N$. The user observes the recommended item and responds with a rating $r(\mathbf{p}_t, \mathbf{v}_t) = \langle \mathbf{p}_t, \mathbf{v}_t \rangle + \varepsilon$. As a result of exposure to and/or consumption of the content the user preference updates to $\mathbf{p}_{t+1} = \mathbf{p}_t + \gamma f(\mathbf{p}_t, \mathbf{v}_t) + \varepsilon'_t$, where $\varepsilon'_t \sim \mathcal{N}(0, \sigma' I_d)$ is zero mean stochastic noise applied to the preference dynamic.

Preference Estimation We make recommendations based on estimated user preferences, $\mathbf{u}_t$. We initialize the estimate with a random multivariate normal vector: $\mathbf{u}_0 \sim \mathcal{N}(0, \sigma I_d)$. Given a recommended item i, and observed rating r_{it} we update the representation of the user preference estimate $\mathbf{u}_t$ according to Online Gradient Descent (OGD) [Haz+16] for the loss function $\ell_t(u) = \frac{1}{2}\left((r_t - \langle u, \mathbf{v}_i\rangle)^2 + \eta\|u\|^2\right)$, leading to an update

$$\mathbf{u}_{t+1} = \mathbf{u}_t - \alpha\nabla_{\mathbf{u}_t}\ell_t(\mathbf{u}_t) = (1-\alpha\eta)\mathbf{u}_t + \alpha(r_{it} - s_{it})\mathbf{v}_i. \tag{3.3}$$

The hyperparameter $\eta > 0$ is a *learning rate*.

We repeat our experimental setup in the *oracle* model in which the recommender has direct access to preferences. We find that the qualitative insights from this section hold both in the oracle and in the estimation model; thus the dynamic patterns that we observe are primarily driven by behavioral shift rather than by estimation error.

Item Selection We consider three baseline recommenders and softmax selection rule with with different temperatures:

1. Baselines: The `uniform` selection chooses an item uniformly at random from the set of items. The `constant` selection repeatedly selects the same item for all recommendation rounds. `greedy` selection picks the item with the maximum predicted score $i_t^* := \arg\max_i s_{ti}$, breaking ties randomly.

2. `softmax` selection: Given predicted scores $\{s_{it}\}_{i=1}^N$ and the estimate of the preference vector $\mathbf{u}_t$, a `softmax` selection rule with temperature β selects item i with probability[4]

$$\mathbb{P}[i_t^* = i] \propto \exp\left(\frac{\beta}{\|\mathbf{u}_t\|}s_{it}\right).$$

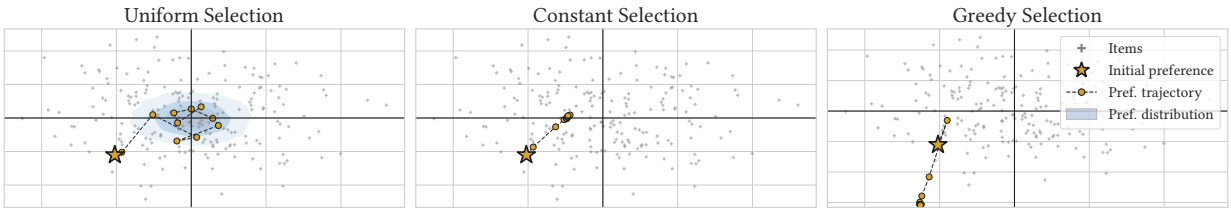
In each of the following illustrations of preference trajectories, we will depict the long-term preference distribution using a cloud, while the first moves of preferences are depicted by connected dots.

Qualitative Behaviors and Testable Predictions

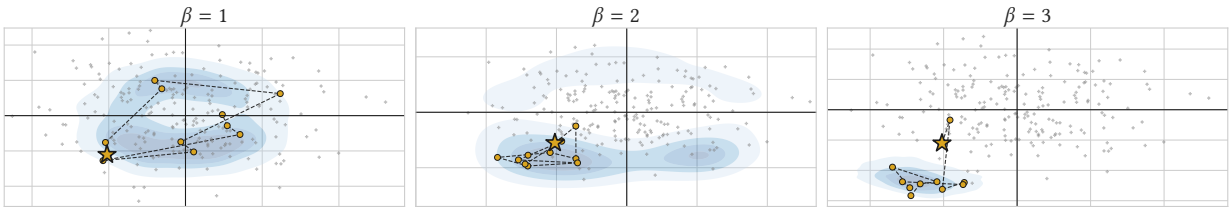
Mere-Exposure

Under this dynamic, users move in the direction of the recommended content, irrespective of how highly they rate it. Figure 3.3 displays user trajectories in a 2-dimensional preference space for f_{ME} with $\gamma_{ME} = 0.1$. In the case of `uniform` recommendations, the preferences converge to a ball centered at the origin. For the `greedy` and `constant` baseline recommenders, the user preference converges to the latent representation of the item that is repeatedly

[4]We scale the β parameter by the norm of the estimated preference to maintain the expected engagement between consecutive rounds constant.



(a) **Baseline recommenders**: `uniform`, `constant` and `greedy` selection rules



(b) `softmax` **recommenders**: $\beta = 1, 2, 3$

Figure 3.3: **Preference Trajectory** ($d = 2$, $N = 1000$): Mere-Exposure dynamic $\gamma_{ME} = 0.1$. Note preferences stay stochastic and lead to a long-term distribution around the origin for the `uniform`. `constant` and `greedy` converge to a repeatedly recommended item.

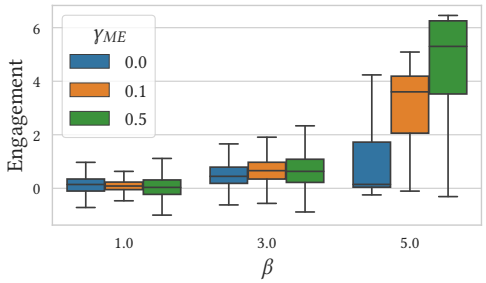
recommended. With the `softmax` selection rule, we observe that the long-term distribution of preferences over time traverses the item space. Under selection rules that favor exploration, for instance `softmax`($\beta = 1$), the preference vector moves faster around the item space, yet stays closer to the origin, compared to selection policies that favor expected engagement; e.g. `softmax`($\beta = 3$).

Figure 3.4 shows how the β parameter affects the engagement of the user, the magnitude of their preference and the the diversity of their content consumption, operationalized as entropy of the distribution of recommended content. We first observe that engagement may increase due to Mere-Exposure (see the increase of engagement for constant β and increasing γ_{ME}). We further note that a recommender with $\beta = 5$ has a very high engagement in particular with high Mere-Exposure. This might raise the question of whether engagement is a valid metric for users with dynamic preference. We will consider this question further in Section 3.6.

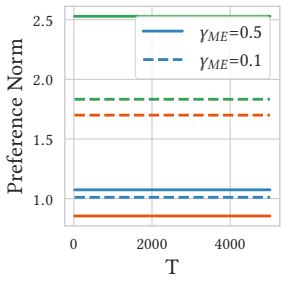
Here we derive the testable prediction that for `softmax` selection rules and Mere-Exposure dynamics the estimates of the preference vector stay relatively constant over time, and the magnitude of the preference increases with the temperature parameter β.

Operant Conditioning

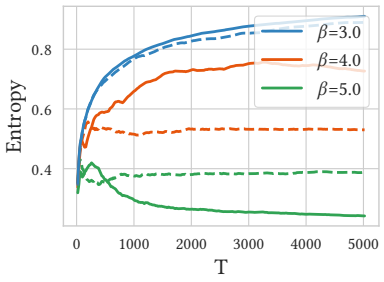
When the preference shift is governed by Operant Conditioning, we observe that the norm of the preference vector oscillates. This phenomenon can be explained by the type of reinforcement that the Operant Conditioning f_{OC} induces. It is illustrative to analyze the



(a) Large γ_{ME} leads to higher engagement for high β.



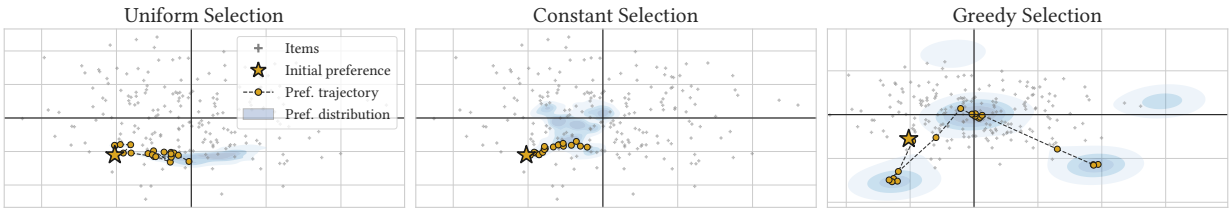
(b) Magnitude over time.



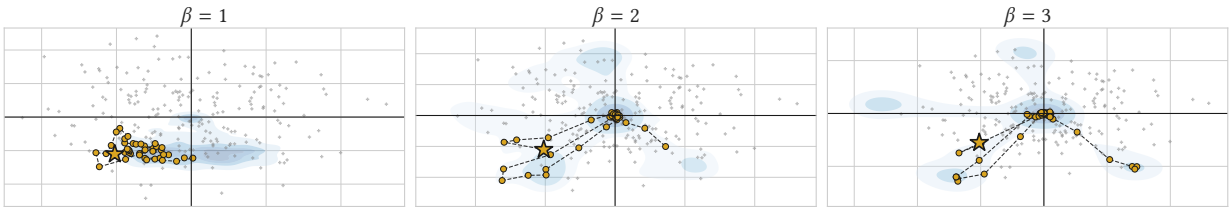
(c) Entropy over time.

Figure 3.4: **Higher Dimensions ($d = 8$, $N = 5000$):** Dependence of user engagement, preference magnitude, and diversity of consumption for different `softmax` β parameters. (a) A higher-β softmax leads to higher engagement, which is exaggerated by stronger Mere-Exposure. (b) A very high-β `softmax` (5.0) leads to preferences of high norm. (c) The entropy of consumed content may be high even for moderately high-β `softmax` (4.0).

behavior for the `greedy` recommender. In the beginning, the user is served a recommendation for which she will respond positively. Indeed the surprise term is positive since the expected engagement is 0, and thus the user preference will shift in the direction of the item. With the preference moving towards the item, its score increases and the `greedy` selection picks it again. The positive reinforcement of preferences from the previous round ensures that the surprise term is still positive; leading to further amplifications in the preference of the direction of the item. Eventually, given the increases in the expected engagement from the previous round, the expected surprise goes to 0. At this time any noise in the response can make the surprise term negative and thus sending the preference vector in the direction of $-\mathbf{v}$. As $-\mathbf{v}$ is nearly diametrically opposed to the preference at this time, even a small negative surprise would considerably decrease the magnitude of the preference vector. However, since the movement is confined to the direction of the original preference vector, the `greedy` selection will keep recommending it. As the preference decreases in magnitude, so does the engagement of the user. However, as expected engagement is a lagging metric, the surprise term becomes even more negative, thereby creating a downwards spiral which ends with the user getting completely bored and losing their preference ($\mathbf{p}_t \approx 0$). After enough time-steps



(a) **Baseline Recommenders**: `uniform`, `constant` and `greedy` selection rules



(b) `softmax` **Recommenders**: $\beta = 1, 2, 3$

Figure 3.5: **Preference Trajectory** ($d = 2$, $N = 1000$): Operant Conditioning, $\gamma_{OC} = 0.1$. Preferences oscillate for `constant` and `greedy` selection.

the historical expectation decreases enough such that some other direction (by random chance) has a positive surprise term, commencing again the amplification of the preference in that direction.[5]

The `softmax` selection rules are less extreme version of the `greedy` recommender, yet many of them still show oscillatory patterns. The period and amplitude of these oscillation depends on the `softmax` parameter β and on the decay parameter in the Operant Conditioning update, δ. The larger the β, the larger is the amplitude of the preference swings and the shorter is the period, compare Figure 3.6.

The oscillations seen in the simulations are testable predictions of Operant Conditioning (note that Figure 3.6 shows estimated scores, not unobserved user preferences). The review of psychological literature in Section 3.2 did not show evidence of such oscillatory patterns in consumption under Operant Conditioning, which makes testing this model with data on a deployed recommender particularly important.

Hedonic Adaptation

Pure Hedonic Adaptation leads to convergence to the baseline point, and does so independently of the recommendation policy. In combination with other effects, Hedonic Adaptation limits the amount by which other effects are perceived. For example when combined with Mere-Exposure (Figure 3.7a) Hedonic Adaptation provides a strong drift towards the baseline preference; and thus the user preference moves "less" within item space. In joint dynamic

[5]It might stabilize at high values of engagement if preference updates are noiseless.

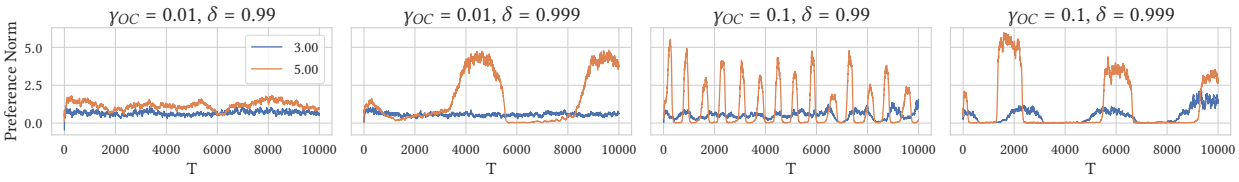
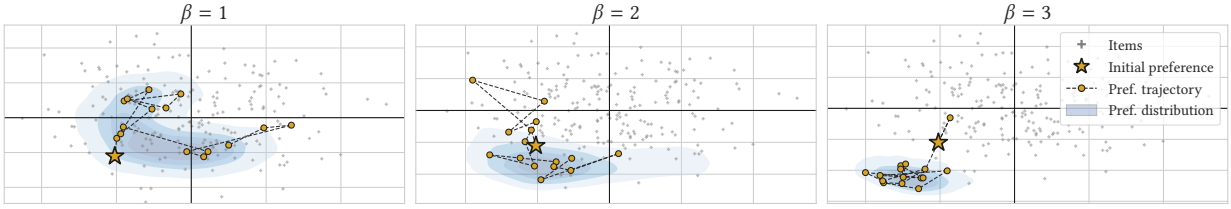
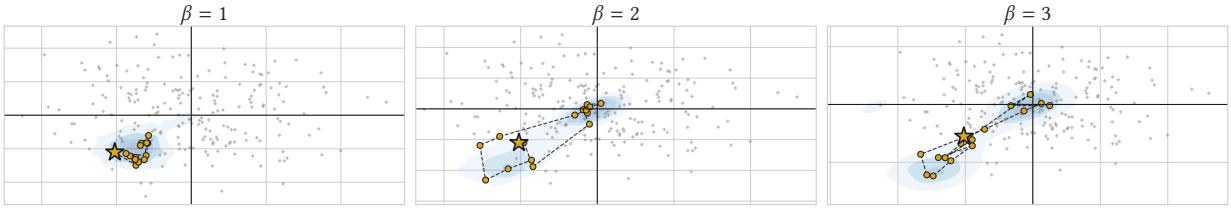


Figure 3.6: **Magnitude of Scores in Higher Dimensions ($d = 8$, $N = 5000$)**: Across behavioral models, higher-engagement selection policies (high β) correspond to more extreme oscillations. When OC effects are large ($\gamma_{OC} = 0.1$), the discount factor δ has a significant impact on the period of oscillations. Lower δ corresponds to a recency bias, where older ratings play a diminished role in forming baseline expectations for engagement, and thus lead to oscillatory patterns of higher frequency.



(a) Mere-Exposure with Hedonic Adaptation: `softmax` recommenders



(b) Operant Conditioning with Hedonic Adaptation; `softmax` recommenders

Figure 3.7: **Preference Trajectory ($d = 2$, $N = 1000$):** Hedonic Adaptation $\gamma_{HA} = 0.01$. Hedonic adaptation biases the trajectories and long-term preference distribution in the direction of the long-term preference distribution.

with Operant Conditioning (Figure 3.7b), user preferences still oscillate, but the oscillations are limited to the direction of the baseline preference. In both cases, Hedonic Adaptation biases, but does not overwhelm the dynamics observed for Mere-Exposure (moving through preference space) and Operant Conditioning (oscillations). Biases towards part of the item space are testable predictions of f_{HA} in combination with other behavioral models.

Perfect preference estimation

We repeat our experimental setup in the *oracle* model in which the recommender has direct access to preferences. We find that the qualitative insights from this section hold both in the oracle and in the estimation model; thus the dynamic patterns that we observe are primarily driven by behavioral shift rather than by estimation error.

Figure 3.8 shows the user preference dynamics for no estimation error. The effects we observe mostly align with what was observed earlier: Mere-Exposure dynamics leads to circling in preference space. Where the magnitude of the preference increases with the `softmax`temperature parameter and the "speed" of the dynamic decreases when β increases. Operant Conditioning displays similar oscillatory patterns as in the case when recommendations are served based on estimated scores. Mixed dynamics with Hedonic Adaptation show similar patterns, biasing the preference trajectories towards the initial preference.

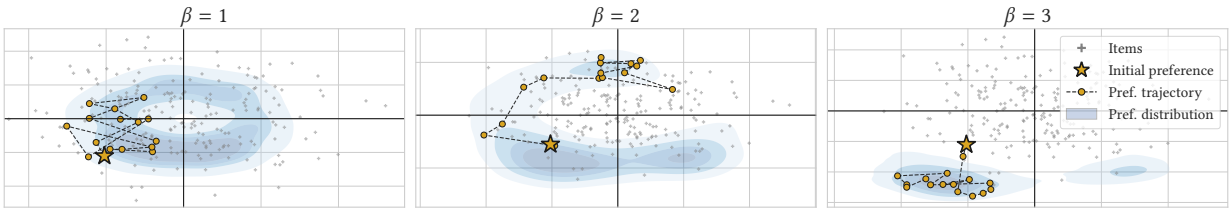
Further we consider how behavioral models interact by considering a preference shift driven by more than one behavioral effect. Figure 3.9 display the preference trajectory of the combined Mere-Exposure and Operant Conditioning dynamics. This experiments allows us to understand the relative strengths of the ME and OC effects. In Figure 3.9a the trajectory qualitatively resembles Mere-Exposure. In Figure 3.9b we half the strength of the ME effect which yields a preference trajectory with the oscillatory trademarks of Operant Conditioning.

3.5 Using Testable Hypotheses to Critique Dynamic Preference Models

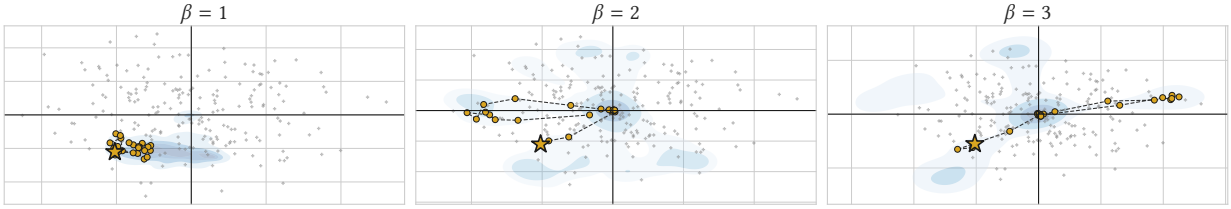
The testable predictions of Mere-Exposure and Operant Conditioning are quite distinct when interacting with recommenders. ME predicts that user ratings will be fairly constant even if recommended content changes over time; whereas OC predicts oscillatory patterns in the ratings. These predictions may or may not be in line with findings in psychology or observed in deployed recommendation systems. Both types of models check allow for further refinement of user-behavioral models.

First, the qualitative behaviors found may be challenged, potentially motivating other quantitative models. The oscillation pattern we observed for our quantitative model of Operant Conditioning, f_{OC} is, to our knowledge, not known in psychology. This might rely on the fact that `softmax` recommendation does not resemble stimuli typically studied in psychology, or might point to a weakness of the model we proposed. Validating such models on deployed system can help select or reject formalizations of models.

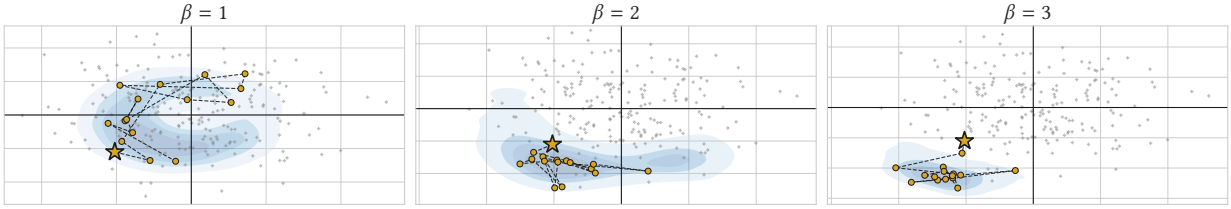
Refinements of functional forms are also possible. Under f_{OC}, negative reinforcement is stronger than positive reinforcement, as our experiments in this section showed. User preferences in recommendation systems might not exhibit this asymmetry, and motivate new models for Operant Conditioning, e.g., decreasing the slope of the surprise when surprise is negative. Similarly, the fact that the preference lingers around 0 for several time periods



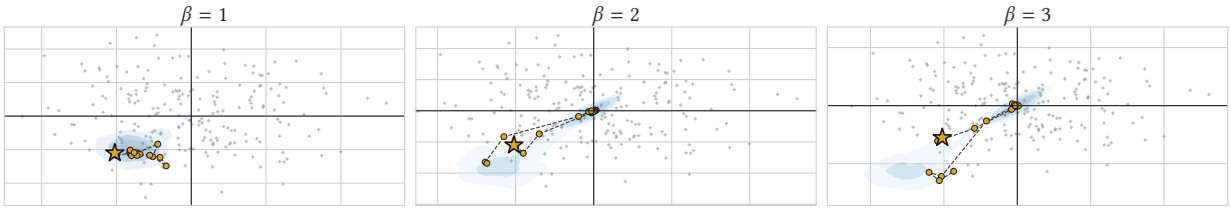
(a) Mere-Exposure



(b) Operant Conditioning

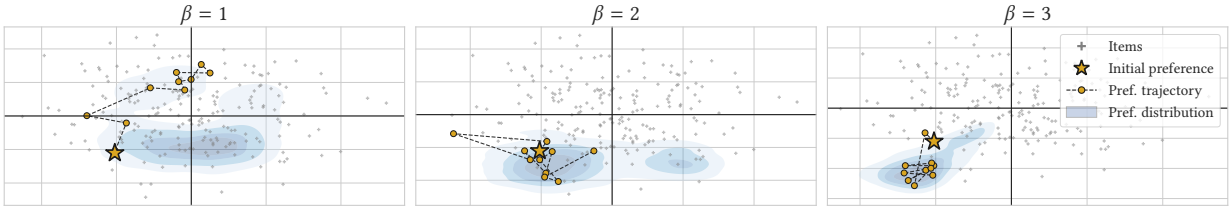


(c) Mere-Exposure and Hedonic Adaptation

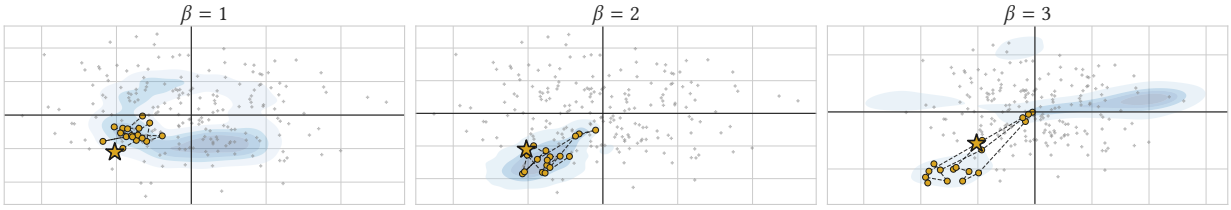


(d) Operant Conditioning and Hedonic Adaptation

Figure 3.8: **Oracle Model:** Preference dynamics with no estimation error



(a) $\gamma_{ME} = \gamma_{OC} = 0.1$



(b) $\gamma_{ME} = 0.05$, $\gamma_{OC} = 0.1$

Figure 3.9: **Mere-Exposure and Operant Conditioning**: Preference trajectories

suggests that our formulation of the user surprise overemphasizes early exposure. This suggests further refinement of our model by reducing the discount factor δ.

Synthesized Methodology for Psychologically Grounded Dynamic Preference Models

In Figure 3.1 we synthesize the following steps for grounding dynamic preference models in psycholagical evidence through the lense of plausibility.

Statement Declare a psychological effect and review relevant psychology literature (Section 3.2);

Formalization Formalize the effect within a recommendation system model (Section 3.3);

Predictions Inspect properties of the model using theory, simulations, or a combination of the two to derive testable predictions (Section 3.4);

Data Test the predictions of the models against historical and/or interventional data in a deployed recommendation system;

Predictions and data may be used to refine the formalization chosen for a particular effect. In addition to the application of the proposed methodology to modelling concrete effects, it may be used to reconsider some models proposed in the recent literature on dynamic preference models in recommendation systems. Next, we give three examples of such a discussion.

Example 3.5.1. In [Pas+21] the authors consider and evaluate several dynamic user models. The authors write "an exponentially weighted moving average [...] distribution is obtained within each time window". The paper considers an update $u_{t+1} = (1-\gamma)u_t + \gamma v_t$, where the item vectors v are modeled as unit vectors with respect to the music genre, and the user factor is a probability vector encoding the likelihood that a user will consume content from that genre. While this model is functionally equivalent to Mere-Exposure, [Pas+21] presents the model in a purely mathematical description. Our review on Mere-Exposure might allow, for example, to make a model context-adaptive: If users are detected as listening in the background (conspicuous Mere-Exposure), γ_{ME} is increased.

Example 3.5.2. The authors of [Car+21] assume a user model based on a *theory of chosen preferences* [Ber+21]. The authors of [Car+21] describe [Ber+21]'s (metacognitive) model as: "on a high-level, at each timestep users choose their next-timestep preferences to be more 'convenient' ones—ones which users expect to lead them to higher engagement value." While [Ber+21] gives examples of how their models explain several behavioral effects (conformism, closed-mindedness, sour grapes—the psychological effect that things that are unattainable are less liked), among others, [Car+21] does not discuss whether this effect, and the particular cognitive model, is relevant in a recommendation system. Explicit psychological models would have allowed to identify parameter ranges for which users following [Ber+21]'s theory resemble Mere Exposure, or another psychological effect.

Example 3.5.3. Having structural models allows critiquing the precise formulation of a psychological effect. [Jia+19; Kal+21] study preference dynamics in closed loop feedback with recommendations and argue theoretically and via simulations that recommendation systems lead to amplification of preferences and consequently to radicalization of users. In [Jia+19] the authors model the effects that repeated recommendation of an item have on preferences. They propose a model which bears similarity to our Mere-Exposure model and concludes that recommendations lead to unbounded preferences. As unbounded interest in a certain type of content is not a plausible prediction, one can critique the proposed user drift dynamics. [Kal+21] proposes a model akin Operant Conditioning and argue for the extremization of user preferences by proving divergence of the preference estimates. The structural model of preference update is formalized in such a way that the estimated preference vectors impact the true preferences directly, which might not capture the correct causal relationship between recommendations and user preferences.

3.6 Evaluation and Design for Dynamic Users

In this section we consider how dynamic preference models may affect recommendation system evaluation metrics and design. This is based on the observation in Section 3.4 that `softmax` recommenders were able to attain high levels of both engagement and diversity. In dynamic settings we illustrate with an example that a recommendation algorithm that improves both engagement and diversity might have unintended consequences.

We will measure engagement as the average rating over time and diversity as the entropy of the normalized counts of the consumed items. When user have static preferences the `softmax` recommender is known to make the optimal trade-off between the expected ratings and the entropy of item selection probabilities, compare, e.g., [Jay57].

Theorem 3.6.1. *For any finite set of items $\mathbf{v}_i, i = 1, 2, \ldots, N$ with scores $\{s_i\}_{i=1}^N \in \mathbb{R}^N$, the recommendation distribution $\{\pi_i\}_{i=1}^N \in \Delta^N$ that maximizes entropy, $H(\pi) := -\sum_{i=1}^M \pi_i \ln(\pi_i)$ subject to an lower bound on the expected score, $\mathbf{E}_{s\sim\pi}[s] \geq a$, is `softmax`$(\beta)$ for some $\beta \in \mathbb{R}$.*

Proof. Consider the following optimization problem:

$$\max_{\pi\in\Delta^N} -\sum_{i=1}^N \pi_i \ln(\pi_i)$$
$$\text{s.t.} \sum_{i=1}^N \pi_i s_i = a.$$

The corresponding Lagrangian is:

$$\mathcal{L}(\pi, \alpha, \mu) = -\sum_{i=1}^N \pi_i \ln \pi_i + \alpha \left(\sum_{i=1}^N \pi_i - 1\right) + \beta \left(\sum_{i=1}^N \pi_i s_i - a\right),$$

The stationarity conditions implies:

$$-\ln \pi_i - 1 + \alpha + \beta s_i = 0, \tag{3.4}$$

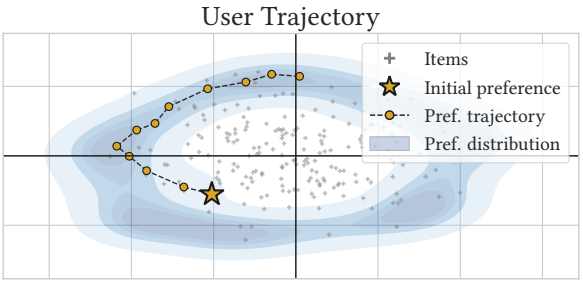
Rearranging (3.4) yields $\pi_i \propto e^{\beta s_i}$. The value of α is such that $\exp(1-\alpha) = \sum_{i=1}^N \exp(\beta s_i)$ and β is such that the feasibility constraint $\sum_{i=1}^N \pi_i s_i = a$ is satisfied. □

We show that the optimality of `softmax` no longer holds in dynamic settings by proposing an algorithm which deliberately limits availability of content, yet outperformes `softmax` both in terms of diversity and engagement.

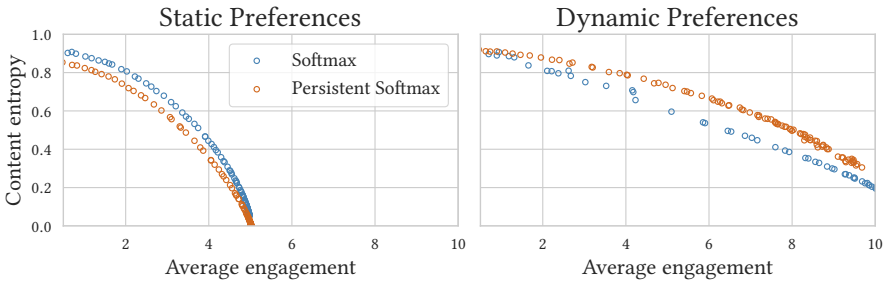
The `softmax` with Momentum As observed in Section 3.3, Mere-Exposure users tend to circle around, and a recommender might use this behavior to move beyond the Pareto frontier of what is possible with a `softmax` recommender. Figure 3.10 shows the application of a recommendation policy that "nudges" user preferences to shift particularly strongly. The recommendation policy used, and benchmarked against the statically optimal `softmax` policy for different temperatures is

$$\mathbb{P}[\mathbf{v}_{t+1} = \mathbf{v}] \propto e^{\beta\langle \mathbf{v}, \mathbf{u}_t\rangle} \, \mathbb{1}_{\langle \hat{\mathbf{u}}_t - \hat{\mathbf{u}}_{t-1}, \mathbf{v}\rangle > 0}, \tag{3.5}$$

where $\mathbb{1}$ is the indicator function. One might call (3.5) a *persistent* recommender or a *recommender with momentum*, that only recommends content that is in the half of the space that the user moved to in the last two directions. This recommender, by deliberately changing user preferences, allows for higher entropy of the consumed content.



(a) Preference trajectory ($\gamma_{ME} = 0.1$)



(b) Engagement vs. diversity

Figure 3.10: **`Persistent softmax` recommendation ($d = 2$, $N = 1000$):** (a) Preference trajectory with Mere-Exposure dynamic $\gamma_{ME} = 0.1$. The 'clockwise' bias of the recommendations leads to faster 'circling' in preference space. (b) Trade-offs between engagement and diversity for static preferences ($\gamma_{ME} = 0$) and dynamic preferences $\gamma_{ME} = 0.1$. `Persistent softmax` is sub-optimal for static preferences but performs strictly better in both content diversity and engagement when preferences are dynamic.

Recommendation System Evaluation with Dynamic Preference Models We see that the persistent softmax recommender seems to be preferable to vanilla softmax on both engagement and diversity dimensions. However, this points to a potential gap between the goal of diversity and its operationalization as consumption entropy. The recommendation system deliberately changes users' preferences to increase diversity of consumed content, which arguably is a undesirable property. Hence, metrics for recommendation systems assuming static users, such as ratings for engagement and entropy for diversity might not capture long-term recommendation system health. In practice, algorithms that seems to improve both in terms of engagement and diversity of consumption might be more opaque than the "manipulative" recommender considered here, which requires making explicit trade-offs in metrics.

3.7 Discussion

This chapter studies psychologically-grounded dynamic preference models for three behavioral effects relevant to recommendation systems: Mere-Exposure, Operant Conditioning and Hedonic Adaptation. Mere-Exposure is the effect that the more familiar something is, the more it is liked. Operant Conditioning captures the effect that preference for something increase when associated to positive surprise, and decrease when the surprise is negative with respect to a baseline expectation. Lastly, Hedonic Adaptation is the effect that after long enough, all happiness will return to a base level. Next, we formalize quantitative models for these models and find that they have qualitatively different testable predictions. We review psychological and experimental evidence relating to each effect and proposed how to refine said models. We generalize our approach into a broadly applicable methodology for grounding dynamic preference models. We conclude by showing that recommender systems evaluation is impacted by dynamic user models. We highlight two areas for future work.

Estimation We give examples of plausible models describing the effect of consumption of content on user preferences and we further study the qualitative properties of these models. The statistically efficient and scalable estimation of user models from empirical data is an important step for future work. Here we distinguish between estimating the strength of posited behavioral effects in benchmark datasets and designing online experimental setups. The case of statistical estimation from historical feedback sequences of users falls under the category of (in the case of Mere-Exposure and Hedonic Adaptation, linear) dynamical systems learning, the collaborative estimation of item factors might imply significant additional complications. In such cases, the item factor estimates might be biased due to the dynamics of the user. Tensor completion [Liu+12] and recent work in the econometrics of state dependence[Tor19] are promising directions for such estimation. The contextual factors affecting the size of dynamic effects, e.g. conspicuousness for Mere Exposure [Kih87], is another important dimension for estimation. The design of online experiments requires either access to real content recommendation systems or careful use of small scale proxy user studies that can test given behavioral hypotheses.

Evaluation and Design Our discussion in Section 3.6 presented an example where in the presence of dynamic user models a recommender with potentially undesirable properties led to higher engagement and diversity proxies than `softmax`, which is provably optimally trading off these two metrics for a static user. Evaluating recommendation systems and designing for metrics hence crucially depends on the behavioral models employed. Further studies that propose and investigate recommendation system metrics for recommender-user systems with dynamic preferences are needed to meaningfully estimate, and hence design recommendation systems.

Part II

How do feedback driven dynamics impact ecosystems and markets?

Chapter 4

Delayed and indirect effects of link recommendations

Building upon our exploration of multi-step closed-loop dynamics and psychologically grounded user behavior models in Chapter 3, we now shift our focus from individual user interactions to the broader context of social networks where user interactions are mediated via link recommendation algorithms. This chapter transitions us from examining single-user dynamics to the complex interplay of multiple users and relationships within evolving social networks, highlighting the intricate feedback loops at play.

Evaluating the impacts of link recommendations on social networks presents significant challenges, as prior studies have been constrained by limited settings. Observational studies struggle with causal questions, and naive A/B tests often yield biased results due to unaccounted network interference. Additionally, many evaluations rely on static network models, overlooking potential feedback loops between link recommendations and organic network evolution. In this chapter, we extend the renowned Jackson-Rogers model to investigate how link recommendations influence network evolution over time. Our findings reveal surprising delayed and indirect effects on the structural properties of networks. Furthermore, we show that the effects of recommendations can persist in networks, in part due to their indirect impacts on natural dynamics even after recommendations are turned off.

This chapter is based on the paper "Delayed and Indirect Impacts of Link Recommendations"[Zha+23] written in collaboration with Han Zhang, Shangen Lu and Yixin Wang.

4.1 Background

Link recommendation algorithms such as Facebook's "People You May Know", Twitter's "Who to Follow" and LinkedIn's "Recommended for You" have an ever-increasing influence on the evolution of social networks, with some accounts crediting to algorithmic recommendations over 50% of links in social networks [Tea22]. This can cause downstream effects on information

flow, opinion dynamics, and resource allocation. For instance, recommendations can be a polarizing force increasing network segregation, which in turn may re-inforce opinion echo-chambers [Cin+22] or restrict access to information and resources for less connected communities [DGL13; ACF22; SLL21]. At the same time, recommendation systems can also surface "long-range" connections between nodes that would not otherwise be exposed to each other, and thus, increase network integration by promoting and maintaining diverse links [Raj+22; DSV19]. Given the ubiquity of algorithmic recommendations on social media, studying their impacts is seeing increased academic and regulatory interest. However, such studies are challenging for a variety of reasons, such as lack of normative framing [DGM10] and limited access to large-scale platforms. In this work, we explore a more foundational evaluation challenge stemming from the fact that social networks have underlying dynamics that interfere with algorithmic recommendations.

Existing real-world evaluations of link recommendation algorithms rely on A/B tests (experimental) or longitudinal data (observational). However, both experimental and observational evaluations can yield misleading conclusions. The validity of A/B tests relies on the Stable Unit Treatment Value Assumption (SUTVA) [IR15; Gui+15], which is violated in case of network interference. Observational studies may fail to assess causal impacts, as longitudinal evaluation lacks counterfactual measurements on what the network evolution would have been without algorithmic recommendations. These challenges have motivated the use of simulation-based evaluations of link recommendations. In simulation studies, one can make explicit network modeling assumptions and then evaluate the impact of recommenders. Despite a growing number of works in this space [SRC18; Cin+22; Fab+20; Fab+22; Fer+22; SLL21], existing simulation-based evaluation falls short of providing insights into the mechanisms through which recommendations impact social networks. Existing simulation studies primarily consider static networks, and thus do not take into account feedback loops between link recommendation and organic network evolution.

In dynamic settings, the main challenge is to measure impacts relative to a "baseline" network. Given this, existing works often measure the effects relative to the initial networks; they rarely measure relative effects to a plausible counterfactual based on the natural evolution of the graph without link recommendations. Such evaluations can lead to qualitatively wrong conclusions. For instance, [Abe+22] shows that triadic closure – the most common type of friend-of-friend recommendations – can reduce segregation with respect to the initial network before intervention. However, triadic closure can, at the same time, increase network segregation in relative terms with respect to a natural evolution dynamic which assumes the addition of random edges.

Here we consider *dynamic* impacts of link recommendations through simulations. We find that link recommendations can have surprising delayed and indirect effects on the structural properties of networks. For instance, we find that the short-term effects of friend-of-friend recommendations can be qualitatively different from long-term effects. In the short term they can alleviate the degree inequality, but in the long term they can increase the degree inequality. Moreover, we demonstrate that the effects of link recommendations can persist in the network even after recommendations have been discontinued. This phenomenon is

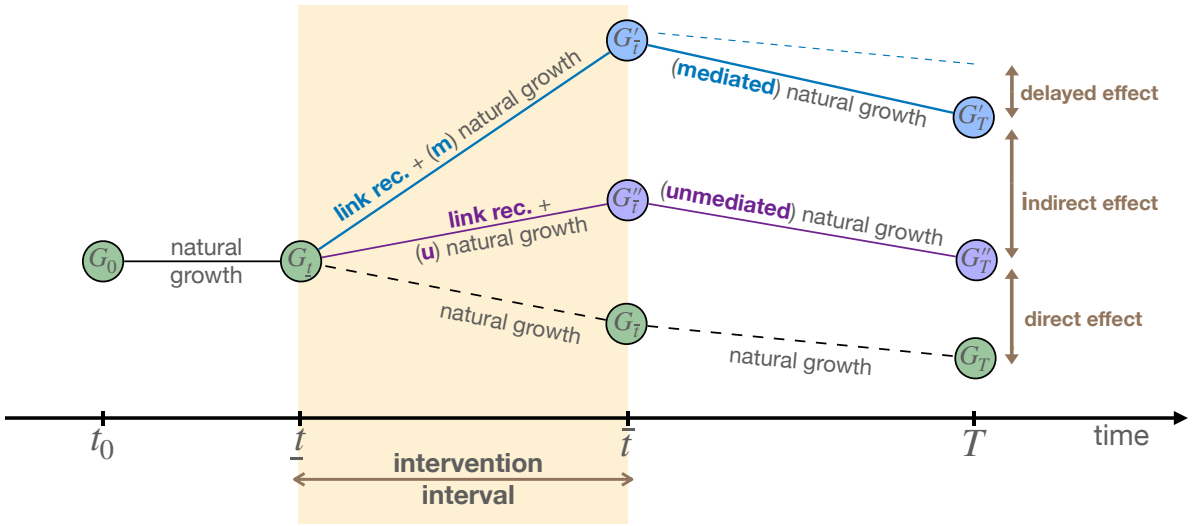


Figure 4.1: Delayed and indirect effects. The image displays two counterfactual evolutions of the network. The solid blue line represents the trajectory for an intervention interval $[\underline{t}, \bar{t}]$ in which the full network receives algorithmic recommendations. The dashed black trajectory corresponds to a counterfactual network evolution without recommendations. The total causal effect of recommendation at time t, Effect_t, can be computed as the difference between the counterfactual trajectories at time t (solid blue and dashed black line). The delayed effect at some time $T \geq \bar{t}$ is the difference: $\text{Effect}_T - \text{Effect}_{\bar{t}}$(note that dashed blue line is parallel to the dashed black line). The solid purple trajectory illustrates the counterfactual evolution of the network in which the indirect effects are removed. The difference between the purple and dashed black curve captures the direct effects and the difference between the blue and purple line captures the indirect effects.

due to the fact that recommendations impact the network in two ways: directly through the creation of algorithmic edges, and indirectly by altering the natural growth dynamics. Furthermore, indirect effects amplify the direct effects, contributing to the persistent impact of link recommendation algorithms. A stylized illustration of the indirect and delayed effects can be found in Figure 4.1.

Contributions

- **Modeling:** In Section 4.2 we propose a dynamic network formation model which extends the Jackson-Rogers model [JR07] to incorporate algorithmic recommendations. Unlike the classic model, our dynamic model includes not only the original two phases —dubbed "meeting strangers" and "meeting friends"— but also a third phase of "meeting recommendations." Additionally, we consider latent node representations, enabling us to model community structure and node activity levels in a flexible manner.

- **Methodology:** We monitor the progression of network metrics over different intervention windows. We compare the immediate impacts observed during intervention with the delayed impacts observed after the intervention has ended. Further, we measure the indirect impacts that recommendations have on network properties. To tease out the indirect impact, we compare the observed network evolution with a counterfactual baseline network that discounts the influence of recommendations on natural growth (Section 4.3).
- **Experimental findings:** Our study reveals diverse qualitative patterns for delayed and indirect effects; we find that different duration of the intervention and/or different times of measurement can lead to drastically different conclusions about *"How do recommendations impact networks?"*(Section 4.4). Furthermore, we find that indirect effects can be substantial; they can significantly amplify the impact of recommendations and persist even after the intervention has ended (Section 4.5).

Related Work

Simulation studies Simulation studies typically analyze the impact of link recommendations on static networks. Some works [Kar+18; Fab+20; Esp+22; SRC18] focus on a single round of recommendations and examine observed edges, finding that homophily — the preference for within-group links — leads to exposure bias toward the more homophilous group, even if it's a minority. Similarly, [Fab+22; Fer+22] find that homophily and node degree are strong indicators of which nodes will receive disproportionate visibility. Other works [Fab+22; Fer+22] make explicit behavioral assumptions on how nodes accept link recommendations and consider cumulative effects of link recommendations over multiple rounds. They reveal algorithmic amplification of biases via "rich get richer" effects and increases in observed homophily over time. A shortcoming of existing simulation-based evaluations is that they implicitly assume that the addition of algorithmic edges is the only change in the network. Conversely, we make explicit modeling assumptions about the underlying network dynamics. This evaluation setup allows us to measure the impact of link recommendations with respect to counterfactual natural evolution. Furthermore, by emphasizing underlying temporal dynamics we can pose more subtle evaluation questions such as: "How does the effect of algorithmic intervention fade over time once recommendations are stopped?" or "How do algorithmic recommendations bias the underlying network growth?"

Platform studies There is limited publicly available research evaluating link recommendation algorithms on real social networking platforms. In one experimental study [DGM10], several recommendation algorithms were compared on IBM's SocialBlue network and found to reduce group homophily. The study also revealed that friend-of-friend recommendations had the highest rate of acceptance but the lowest level of edge activity. On the other hand, [Raj+22] found that recommending more distant connections or "weak ties" through LinkedIn's "People You May Know" algorithm led to higher transmission of job opportunities.

In observational settings, a longitudinal study comparing the Twitter network before and after the introduction of the "Who To Follow" recommender in 2010 [SSG16] found that while recommendations increased the number of connections for all users, the highest gains were achieved by the most popular nodes. Conversely, a study comparing links formed naturally and links formed via recommendations on Flickr and Tumblr [AB17] found that the recommended links were more diverse and less biased towards popular users. So far, existing observational platform studies provide a limited understanding of the underlying mechanisms and lack access to counterfactual network evolution. A/B tests, on the other hand, may produce wrong estimates of the effects due to network interference. Our counterfactual simulations allow us to articulate in stylized settings the source of bias in both longitudinal and A/B evaluations.

Theoretical investigations. The impacts of friend-of-friend recommendations in homophilous networks were studied theoretically in a number of works. Under choice homophily, which captures the setting when nodes preferentially accept recommendations to within-group nodes, [SRC18; Asi+20] show that recommendations lead to more exposure gains for the homophilous group which further exacerbates homophily. In contrast, [Abe+22] shows that, when the closure of triangles via friend-of-friend recommendations is not biased in favor of in-group edges, recommendations can in fact improve network integration. In our work, we investigate friend-of-friend recommendations and their impacts on network segregation and show that their effects further depend on the length of intervention as well as the time of measurement.

Delayed algorithmic impacts. Temporal dynamics associated with algorithmic interventions have been previously studied in the context of fairness in Machine Learning. These works showcase broad settings where algorithmic interventions designed to improve fairness [Liu+18; DGM10; Akp+22], robustness [Mil+19] or diversity [Cur+22] in the short term, lead to the opposite effect in the long run. Our work uncovers similar surprising temporal dynamics in the case of link recommendations.

4.2 Dynamic Network Model

To illustrate temporal complexities of evaluating link recommendation algorithms in a dynamic setting, we consider a stylized network evolution model. We propose an extension to the classic Jackson-Rogers (JR) network evolution model [JR07]. The JR model has been validated empirically and shown to display proprieties of real network such as decreasing diameter over time, increased edge densification and emergence of community structure [JR07; RF17]. Our extension adds an optional recommendation phase to model the feedback loop between natural network evolution and algorithmic recommendations.

Setup

Let $G^t = (V^t, E^t)$ be an undirected network where V^t and E^t are the set of nodes and edges at time t. Nodes are characterized by their group identity g_i and latent representation $\mathbf{v}_i \in \mathbb{R}^d$. We assume group identities are static over time and latent representations are sampled from a group-specific multivariate normal distribution $\mathbf{v}_i \sim \mathcal{D}_{g_i}$. At each time step, the network evolves via new nodes which first connect "organically" the existing nodes. Further at each step, the network evolves "algorithmically" via connections mediated by a link recommender.

Natural growth

Similar to the classic JR model, upon the arrival of a new node, the network evolves in two phases: "Meeting Strangers" and "Meeting Friends". In the "Meeting Strangers" phase, the new node makes connections with existing nodes at random. Unlike the classic model where connections are made with a fixed probability, we model connections probabilities based on latent representation of the nodes $\mathbf{v}_i$. This additional modeling assumption allows us to consider various community structures. In the "Meeting Strangers" phase, N_s candidate nodes are sampled uniformly from the existing network, among which the arriving node might make zero, one or more connections. Specifically, the arriving node i connects with each candidate node j with probability proportional to the inner product of their respective latent embeddings: $p_{i,j} = \sigma(\langle \mathbf{v}_i, \mathbf{v}_j \rangle)$ where $\sigma(\cdot)$ is a scaled and translated sigmoid function[1].

In the subsequent "Meeting Friends" phase, the incoming node i makes additional connections based on neighborhood proximity. Here N_f candidate nodes are sampled from the set of nodes at distance 2 (neighbors of neighbors) from node i. Node i connects with each candidate node j with constant probability. Optionally, we model attrition effects by considering node departures from the network with a hazard function [2] that increases with the age of the node.

Algorithmic intervention

We introduce algorithmic intervention as a third "meeting recommendations" phase, whereby nodes in the network receive link recommendation. This phase applies not only to incoming nodes but also to existing nodes in the network. Upon receiving a recommendation, nodes accept it according to a behavioral model. We consider neighborhood and affinity-based recommendations. The prototypical neighborhood recommendation is the friend-of-friend (`FoF`) recommendation where candidate nodes are selected uniformly from the set of nodes at distance 2: $\mathbb{P}(\mathbf{rec}_i = j) = \frac{\mathbf{1}(dist(i,j)=2)}{\sum_{j'} \mathbf{1}(dist(i,j')=2)}$. Conversely, affinity-based (`Latent`) recommendations utilize the latent structure rather than local neighborhood structure to make recommendations. Specifically, the `Latent` recommender computes affinity scores for all the

[1] We consider $\sigma(x) = \frac{1}{1+e^{-ax+b}}$ where we set a and b to match desired average linkage probabilities.

[2] $h(\mathbf{a}) = cd^{\mathbf{a}} + k$ where $\mathbf{a}$ is the age of the node and c, d, k are tuned to model mean and variance of lifespans in the network.

nodes in the network in terms of inner products and recommends candidates with probability proportional to the scores: $\mathbb{P}(\mathbf{rec}_i = j) \propto e^{\beta s_j}$, where $s_j = \langle \mathbf{v}_i, \mathbf{v}_j \rangle$ is the similarity between node i and node j's embeddings. High values of softmax temperature parameter β correspond to more deterministic recommenders. Lower values of β lead to more random recommendations and qualitatively capture the effects of estimation noise for learning-based recommenders.

Behavioral models

We consider two behavioral models for accepting link recommendations: the constant probability baseline and the embedding-based probability where nodes accept new links based on proximity in latent space. Embedding-based probabilities model more granular notions of choice homophily [Asi+20]. We consider an additional behavioral option in which upon acceptance of a new recommended link, an edge is removed at random from the set of existing edges. This option is in line with several works which model recommendations as a *rewiring* [Asi+20; Fab+21; SLL21] process that does not impact the average degree.

Algorithm 1 Simulating network evolution

input : initial G_0; time-steps T; hazard function;
communities: prevalence c_g, latent distribution $\mathcal{D}_g$;
natural growth: N_s, N_f, dist-2 connect prob p_2;
intervention: window $[\underline{t}, \overline{t}]$, `recommender`, `behavior`;

for t *in* $1 \dots T$ **do**
  Natural growth sample incoming node i according to group prevalence c_g and latent distribution $\mathcal{D}_g$ strangers $\leftarrow$ samples N_s nodes **for** s *in strangers* **do**
    add edge $i - s$ with probability $\propto \langle \mathbf{v}_i, \mathbf{v}_s \rangle$
  friends $\leftarrow$ samples N_f neighbors of neighbor nodes **for** f *in friends* **do**
    add edge $i - f$ with probability p_2
  if $t \in [\underline{t}, \overline{t}]$ **then**
    Algorithmic intervention **for** *node* j *in treatment group* $G_t^{treatment}$ **do**
      candidate = `recommender`(j, G_t) **if** `behavior`$(j,$ candidate$) =$ accept **then**
        add edge $j-$candidate
  nodes to remove $\leftarrow$ hazard function

4.3 Evaluating Effects of Recommenders on Structural Metrics

We analyze two simple baseline recommendation algorithms: a neighborhood-based and an affinity-based algorithm. Recommendation algorithms deployed in practice use both neighborhood and affinity information, with potentially additional learning components. Here, however, to understand the underlying mechanisms behind evaluation phenomena in

dynamic networks, we opt for simplicity in the choice of recommendation algorithms. We define the evaluation procedure with respect to structural network metrics.

Structural metrics

Clustering coefficient Clustering coefficient of a node is the ratio of triangles it forms Δ_i and the maximum number of triangles it could have formed given it's current degree d_i; i.e. $c_i = \frac{2\Delta_i}{d_i(d_i-1)}$. The clustering coefficient captures the micro-level cohesion in the network [Fer+22]. The average clustering coefficient is the metric averaged over all the nodes. Averaging the metric over communities measure clustering at the community level.

Gini coefficient We measure the inequality in the degree distribution via Gini coefficient. An increase of the Gini coefficient resulting from the use of recommendations is often used to demonstrate biases in link recommendation [Fer+22; Fab+21; Fab+22]. The Gini coefficient can be computed for an ordered list of node degrees as: $G = \frac{2\sum_{i=1}^{n} i d_i}{n\sum_{i=1}^{n} d_i} - \frac{n+1}{n}$. Similarly, this metric can be computed for the entire graph or restricted to communities.

Homophily We define monocromatic and bichromatic edges to be edges that link two nodes from the same community and nodes from different communities, respectively. The homophily of a community measures the propensity of nodes to favor within-group connections compared to a non-preferential baseline. We compute the homophily of a community g as follows: $H_g = \frac{|E_{gg}|}{|E_g|} - \frac{n_g}{n}$, where $|E_{gg}|$ denotes the number of monochromatic edges within g, $|E_g|$ denotes the number of total edges that have at least one node in community g. Finally n_g denotes the size of g, and n denotes the total size of the network, the ratio $\frac{n_g}{n}$ is the baseline fraction of within-group links when nodes have no group-based preferences. The vast majority of existing simulation-based evaluations study recommendation-induced changes in homophily [Fab+20; Abe+22; DGM10].

Temporal evaluation

We denote by $\mathcal{G}$ the evolution of the network under natural dynamics and by $\mathcal{G}(\texttt{rec}, [\underline{t}, \overline{t}])$ the evolution when the network receives link recommendation according to `rec` algorithm over the $[\underline{t}, \overline{t}]$ intervention interval. The subscript T in $\mathcal{G}_T(\cdot)$ is used to denote the snapshot of the network at time T. For a structural metric m, we denote by $m(\mathcal{G}_t(\texttt{rec}, [\underline{t}, \overline{t}]))$ the value of this metric evaluated at time t. Given a network evolution model, we simulate the counter-factual trajectories of the metrics for intervention intervals $[\underline{t}, \overline{t}]$ of different lengths.

Total effects. The total effect of intervening in a network evolution $\mathcal{G}$ with recommender `rec` for intervention interval $[\underline{t}, \overline{t}]$ on the metric m at time T is defined as the difference

between two counterfactual measurements:

$$\text{Effect}_T(\mathcal{G}, \texttt{rec}, [\underline{t}, \bar{t}]) := m(\mathcal{G}_T(\texttt{rec}, [\underline{t}, \bar{t}])) - m(\mathcal{G}_T).$$

This definition compares between the "treatment" universe where the whole network received an algorithmic intervention and a counterfactual "control" universe where the network evolved without algorithmic recommendations.

Delayed effects. When the measurement occurs during the intervention interval, i.e. $T \in [\underline{t}, \bar{t}]$ we call the effect *immediate*. Conversely, when the measurement occurs after the intervention stops, i.e. $T > \bar{t}$, we can define the notion of *delayed* impact as:

$$\text{DelayedEffect}_T(\mathcal{G}, \texttt{rec}, [\underline{t}, \bar{t}]) = \text{Effect}_T(\mathcal{G}, \texttt{rec}, [\underline{t}, \bar{t}]) - \text{Effect}_{\bar{t}}(\mathcal{G}, \texttt{rec}, [\underline{t}, \bar{t}])$$

The notion of delayed impacts allows us to characterize the impact of link recommendations into three broad categories: diminishing, amplifying, and persistent based on the sign of the delayed effect ($\text{DelayedEffect}_T(\mathcal{G}, \texttt{rec}, [\underline{t}, \bar{t}])$). Note that the delayed impacts measure the difference in effect sizes between time $\bar{t}$ and time T, rather than the difference in the metric $m(\mathcal{G}_T(\texttt{rec}, [\underline{t}, \bar{t}])) - m(\mathcal{G}_{\bar{t}}(\texttt{rec}, [\underline{t}, \bar{t}]))$. These two notions are equivalent only when the metric remains constant from time $\bar{t}$ to time T under natural network evolution dynamics.

Indirect effects. The temporal evolution of the network in the presence of link recommendations is affected *directly* by the addition of algorithmic edges; but also *indirectly*, as the addition of algorithmic edges biases the natural growth dynamics. In our model, recommendations have an indirect impact on natural dynamics in the "Meeting Friends" phase of the network evolution. Upon arrival, a new node forms the initial neighborhood by "Meeting Strangers". In the next phase, the arriving node connects with nodes at distance 2 from itself (neighbors of neighbors). In the presence of algorithmic edges, there can be nodes at distance 2 that require algorithmic edges in order to be reachable. For instance, in Figure 4.2, at time $t + 2$ node F arrives and connects with B in the first phase of natural growth. In the second phase, nodes $\{A, C, D, E\}$ are reachable for the purposes of "Meeting Friends", however node A is only reachable because of algorithmic edge $B - A$. If the incoming node F connects to a node such as A, the resulting edge $F - A$ is said to be *mediated* by recommendation.

To measure direct impacts, we design a counterfactual experimental procedure to remove the indirect influence of algorithmic edges on natural growth by discounting mediated edges. Constructing a post-hoc counterfactual by simply removing mediated edges from the network would inadvertently decrease the edge density in the network and thus bias the analysis. To address this, the counterfactual procedure is defined as follows: upon arrival of a new node i, we modify the "Meeting Friends" phase to only consider node candidates that are at distance 2 from i via non-algorithmic edges. For this counterfactual, the network contains no mediated edges. Since organic link are created only when a new node arrives, this counterfactual removes all *indirect* effects of recommenders on natural growth. Formally, we refer to this unmediated evolution as $\widetilde{\mathcal{G}}(\texttt{rec}, [\underline{t}, \bar{t}])$.

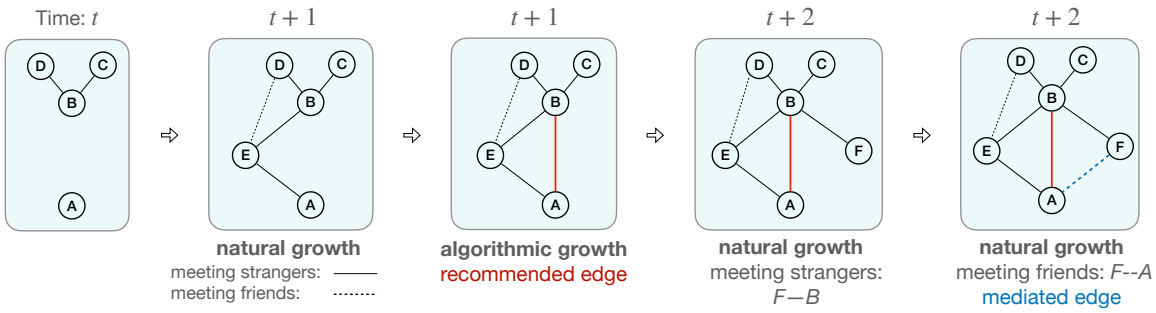


Figure 4.2: Mediated edges. Edge $A - B$ is a recommended edge and a direct effect of the recommendation. Edge $F - A$ is a mediated edge as node A would not be reachable from F in the absence of $A - B$ and thus an indirect effect of recommendation.

Finally we can define direct effects as the difference between the metric evaluated for the unmediated counterfactual and the natural growth trajectory. The indirect effects are the difference between total effect and direct effect:

$$\begin{aligned}\text{DirectEffect}_T(\mathcal{G}, \texttt{rec}, [\underline{t}, \overline{t}]) &= m(\widetilde{\mathcal{G}}_T(\texttt{rec}, [\underline{t}, \overline{t}])) - m(\mathcal{G}_T),\\ \text{IndirectEffect}_T(\mathcal{G}, \texttt{rec}, [\underline{t}, \overline{t}]) &= \text{Effect}_T(\mathcal{G}, \texttt{rec}, [\underline{t}, \overline{t}]) - \text{DirectEffect}_T(\mathcal{G}, \texttt{rec}, [\underline{t}, \overline{t}]).\end{aligned}$$

Longitudinal evaluation. In a simulated setting, one has access to both counterfactuals: the evolution of the network in the presence of recommendations and in their absence. In longitudinal evaluations (i.e. observational studies), effects are measured as the difference in a metric before and after the intervention:

$$\widehat{\text{Effect}}_T^{obs}(\mathcal{G}, \texttt{rec}, [\underline{t}, \overline{t}]) := m(\mathcal{G}_T, \texttt{rec}, [\underline{t}, \overline{t}]) - m(\mathcal{G}_{\underline{t}})$$

This measure is biased whenever $m(\mathcal{G}_T) \neq m(\mathcal{G}_{\underline{t}})$.

A/B evaluation. The lack of valid counterfactuals motivates the use of A/B tests in settings with interventional access. In an A/B test, nodes are divided into two groups: treatment and control; the treatment group receives recommendations, while the control group does not. However, in networks, the Stable Unit Treatment Value Assumption (SUTVA) [IR15] does not hold due to network interference. There are various methods for choosing treatment nodes, as well as methods to correct for network interference in estimation procedures.

Let a scheme for choosing treatment nodes exist and let $\mathcal{G}^{AB}(\texttt{rec}, [\underline{t}], \overline{t})$ be the network evolution where a group of nodes are assigned to the treatment group and receive recommendations. To estimate the impact of recommendations on a metric of interest, we compare the values of the metric on the treatment and control groups:

$$\widehat{\text{Effect}}_T^{AB}(\mathcal{G}, \texttt{rec}, [\underline{t}, \overline{t}]) := m^{treatment}(\mathcal{G}_T^{AB}, \texttt{rec}, [\underline{t}, \overline{t}]) - m^{control}(\mathcal{G}_T^{AB}, \texttt{rec}, [\underline{t}, \overline{t}]).$$

In evaluating the impact of recommendations, we consider both naive estimators that do not correct for network interference, and more sophisticated ones that take network externalities into account. We simulate counterfactual evolutions of the network by applying different recommendation interventions for various time periods. Although accessing counterfactuals in reality is infeasible, these simulations offer insights into dynamic and temporal phenomena.

Experimental Setup

Our results consider a simple experimental setup to illustrate delayed and indirect effects of `FoF` and `Latent`[3] recommenders. In the baseline setup, we consider two equally sized communities. We sample latent embeddings for the nodes independently from $\mathcal{N}(\mu_c, \sigma I)$, with $\mu_1 = [0, 1]$ and $\mu_2 = [1, 0]$; for both communities, the variance of the embeddings is set to $\sigma = 0.05$. We sample 50 nodes in each group and initialize both within-group and across-group edges by connecting pairs of nodes $i - j$ with probability proportional to the inner product of their embeddings. Then, for each node in the network, we consider their neighbors at distance 2 and connect to them with probability $p_1 = 0.05$. This results in a slightly homophilic initial network with homophily $h_1 = h_2 \approx 0.1$. Upon initialization, at each time step, 5 new nodes arrive. For natural growth, we consider $N_s = N_f = 100$ and the connection probability of connecting to a candidate node in the "Meeting Friends" phase: $p_2 = 0.05$ (which adds an average of 5 edges per node). Further, we assume a behavioral model where nodes accept recommended edges with a constant probability, 0.5. At each time step we measure structural metrics of the network such as clustering coefficient, Gini coefficient and homophily. We repeat each trajectory for 5 random seeds and report average results along with confidence bands. [4]

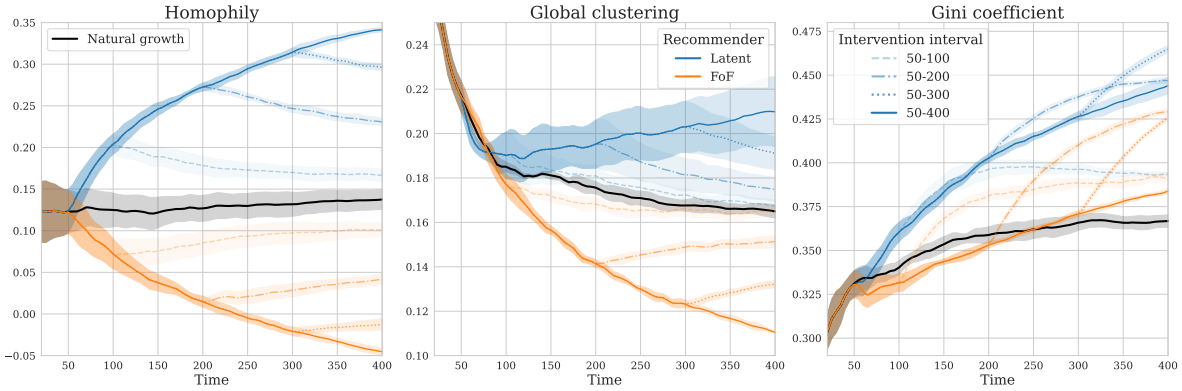
In the section below we vary these assumptions by considering majority-minority structure, differentiated homophily, and within-group heterogeneity. Finally we illustrate evaluation phenomena by considering additional behavioral assumptions, different variants of the underlying dynamics and modifications to the recommendation algorithms.

4.4 Delayed Effects

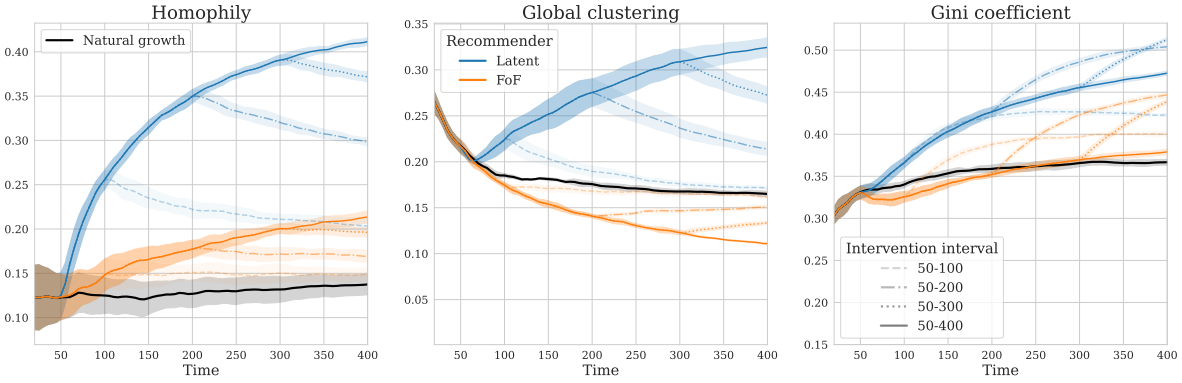
We find that affinity and neighborhood-based recommenders have different long-term effects on homophily and clustering coefficients. `Latent` leads to increases in homophily as well as global clustering whereas `FoF` recommendations reduces clustering. Both recommenders have diminishing delayed impacts with respect to homophily and clustering, as upon the end of the intervention, the trajectory of the metric regresses to the counterfactual natural growth trajectory. Meanwhile, effects with respect to the Gini coefficient are qualitatively different. `Latent` recommendations lead to increases in degree inequality both in the short and in the

[3]For `Latent` recommendations we use softmax temperature $\beta = 10$, which in our simulations corresponds to $\approx 80\%$ chance of recommending 'closest' node.

[4]Code for reproducing experimental results can be found here.



(a) Constant recommendation acceptance probability



(b) Choice homophily in recommendation acceptance probability

Figure 4.3: Counterfactual trajectories of homophily, global clustering and Gini coefficient for `Latent` and `FoF` recommendations applied over varying intervention intervals. The solid black lines represent the evolution of the metric under natural growth dynamics. The blue lines represent trajectories for the `Latent` recommender whereas the orange lines correspond to `FoF`. The shaded area corresponds to the 95% confidence intervals calculated over 5 independent trajectories.

long term. Conversely, `FoF` recommendations decrease the Gini coefficient in the short term but increase it in the long term. Figure 4.3 illustrates these findings.

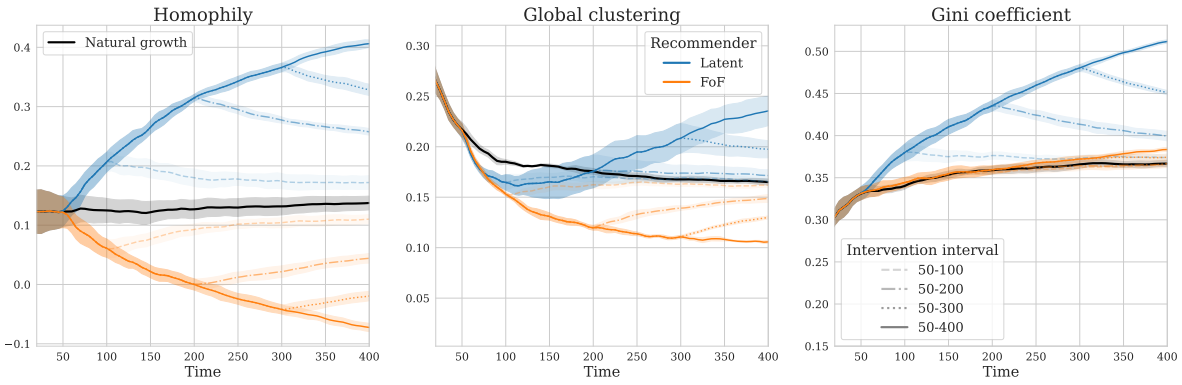
Clustering and homophily. In the case of `Latent` recommendations, if nodes i, j, and k are similar in embedding space, it is likely they are from the same community and all three pairs of edges have been proposed as recommendations, leading to a strong bias for closing triangles within the community. This results in increased homophily and clustering, especially with the behavioral assumption that nodes accept links based on embedding similarity (see Figure 4.3b and Figure 4.4b). The bias is further exacerbated for high-temperature β. Conversely, with a low β value, which corresponds to a more stochastic recommender, it is less likely that all three edges of a given triplet will close, thus slightly lowering the homophily value compared to the high-β case.

The `FoF` recommender intervention mimics the "Meeting Friends" phase of natural growth, which has a bias towards forming cross-community links, resulting in decreased homophily. Decreased clustering occurs because random `FoF` recommenders lack the bias of connecting nodes with a large common neighborhood, unlike neighborhood-based models such as Adamic-Adar [AA03], which favor links with nodes with a large number of common neighborhoods and thus increase clustering (see Figure 4.5).

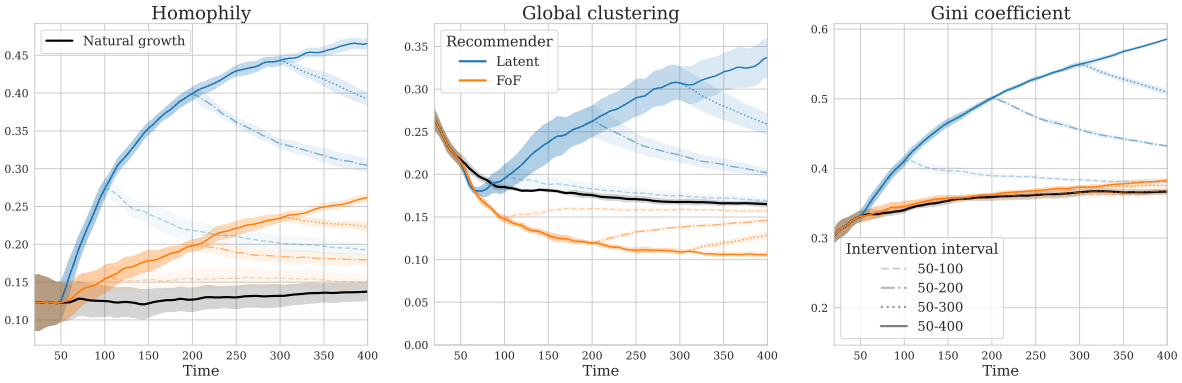
Gini coefficient. The Gini coefficient for recommendations shows amplifying delayed effects. The natural growth trend exhibits a slight increase in inequality over time. `Latent` recommendations exacerbate the wealth gap between popular and unpopular nodes through a bias towards nodes with high embedding norms, creating the "rich-get-richer" effect. `FoF` recommendations initially reduce degree inequality as they are less biased towards popular nodes. Surprisingly, once recommendations stop, the Gini coefficient increases dramatically for both recommenders, particularly for `FoF`. This is due to a "relative-rich-minority" effect caused by differences in edge density between "older" and "newer" nodes. Through recommendations, the existing nodes have accumulated a large number of connections. After recommendations stop, as new nodes are arriving and natural growth is continuing to take place, the set of "rich" nodes becomes relatively smaller, while the majority of nodes, who come after recommendations stop, have much fewer connections. This results in higher inequality as measured by the Gini coefficient. If a rewiring behavioral model is assumed (see Figure 4.4), where nodes receiving recommendations do not change their degree, the delayed impacts diminish, similar to the case of clustering and homophily.

Intervention Variants

In Figure 4.5a, we consider different choices of temperature parameter β for the `Latent` recommender. All recorded metrics increase as the recommendations become more deterministic (increasing β). Note that $\beta = 4$ and $\beta = 10$ are nearly indistinguishable in the homophily graph whereas $\beta = 4$ and $\beta = 2$ are close in the graph of the clustering coefficient.



(a) Edge rewiring + Constant recommendation acceptance probability



(b) Edge rewiring + Choice homophily

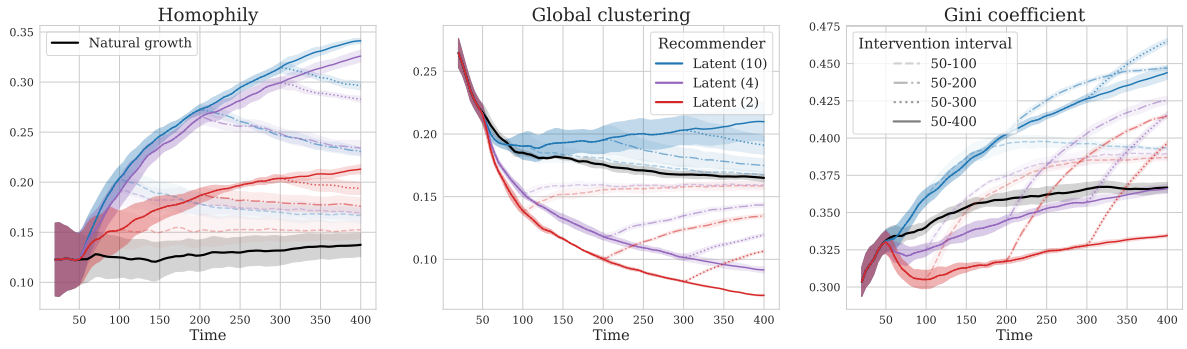
Figure 4.4: Counterfactual trajectories of homophily, global clustering and Gini coefficient for `Latent` and `FoF` under edge rewring assumptions

This suggests that different metrics have different sensitivities to changes in recommendation intervention.

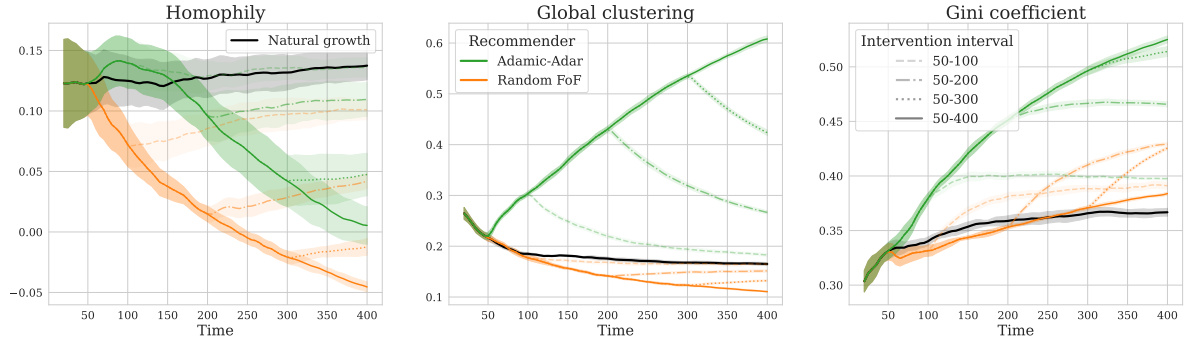
In Figure 4.5b we compare `FoF` with a neighborhood-based recommendation algorithm that recommends node pairs with higher Adamic-Adar index. The Adamic-Adar index assigns a higher similarity score to pairs of low-degree nodes that share many neighbors in common. This change leads to an extreme increase in the clustering coefficient.

4.5 Indirect Effects

The "Meeting Friends" phase occurs in feedback loop with algorithmic recommendations, leading to the mediation phenomenon illustrated in Figure 4.2. We find that mediated edges are common, long-lasting and biased. For instance, in the case of `Latent` recommendations, the mediated edges are more likely to connect nodes from the same community, resulting in significant effects on homophily.



(a) `Latent` recommender



(b) `FoF` recommender

Figure 4.5: Counterfactual trajectories for variants of `Latent` and `FoF` recommenders

Prevalence, persistence and bias of mediated edges. Mediated edges play a substantial role in the "Meeting Friends" stage of natural growth dynamics. Figure 4.6 shows that the proportion of mediated edges increases with the duration of the intervention period and remains constant even after recommendations have stopped for both `Latent` and `FoF` recommenders. The bias of mediated edges in terms of bichromaticism is distinct for each recommender, relative to the proportion of bichromatic edges under natural growth. For instance, about a third of unmediated "Meeting Friends" edges are bichromatic for both recommenders. However, the proportion of bichromatic edges among mediated edges is lower for `Latent` and higher for `FoF`. Our findings about homophily in Section 4.4 suggest that algorithmic edges are biased towards monochromatic for `Latent` and bichromatic for `FoF`. This experiment supports the hypothesis that recommendations have a compounding effect by inducing biases in the natural network evolution.

Counterfactual measurements of indirect effects. To measure the indirect effects on structural metrics, we apply the counterfactual procedure from Section 4.3 and isolate the direct effects. We find that without mediation, structural metrics trend faster to their natural evolution. Figure 4.6 shows the comparison of homophily under natural growth, algorithmic

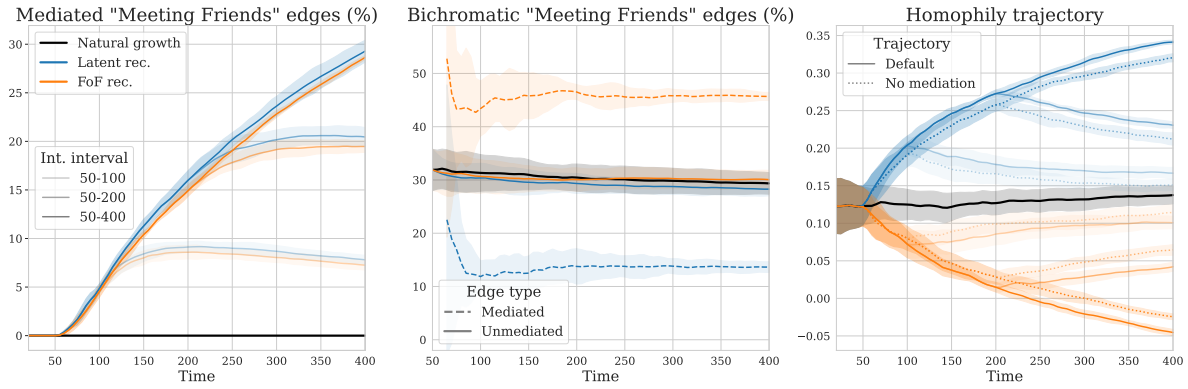


Figure 4.6: Indirect effects. Solid black lines correspond to natural growth dynamics. The blue and orange lines represent trajectories for `Latent` and `FoF` recommenders, respectively. Left plot illustrates the prevalence of algorithmically mediated edges in the "Meeting Friends" phase of the natural growth dynamics. Lighter and darker lines correspond to various intervention intervals. Center plot illustrates the bias in the mediated edges relative to unmediated ones. Solid lines correspond to fraction of bicromatic edges among unmediated edges created in the "Meeting Friends" phase; the dashed lines record the fraction of bicromatic mediated edges. Right plot shows the trajectories of homophily metrics. Dotted lines correspond to the trajectories of the metrics for the unmediated counterfactual evolution of the network.

growth, and unmediated algorithmic growth. The difference between the latter two quantifies the magnitude of indirect effects. Additionally, indirect effects grow relatively stronger over time, even after recommendations end, highlighting the persistence of indirect impacts.

4.6 Impacts of Group Structure

Previous studies have highlighted the unequal impact of recommendations on different communities, especially when they are divided into majority and minority groups and display varying levels of homophily [SRC18; Fer+22; Fab+20]. Here, we examine the effects of differential homophily between majority and minority groups as well as within-community heterogeneity.

Community heterogeneity. We examine the impacts of homogeneity and heterogeneity in the latent representation of nodes on the results of link recommendation. We model within-group heterogeneity by varying the variance of the latent embeddings. In the heterogeneous setting, we set the variance of the embedding distribution to $\sigma^2 = 0.1$, and to $\sigma^2 = 0.01$ in the homogeneous setting; in the extreme case when the variance is 0 this recovers the JR variant of [Abe+22].

Figure 4.7 shows that when there is high within-group heterogeneity, `Latent` recommendations greatly increase the global clustering coefficient, while for high within-group homogeneity,

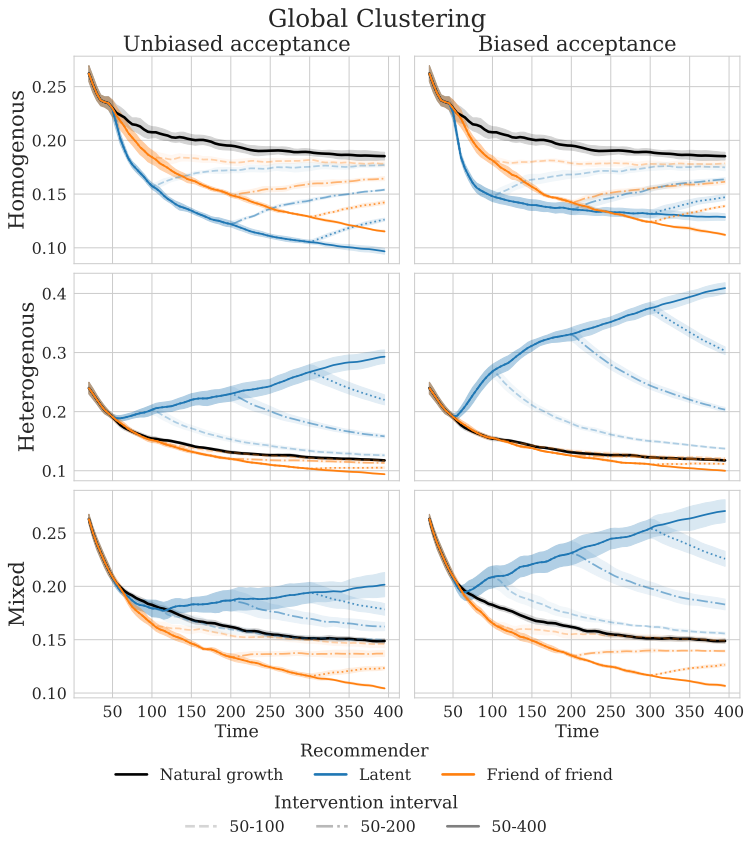


Figure 4.7: Effects of within-group heterogeneity. Black trajectories correspond to the natural evolution of the global clustering metric. Blue and orange lines indicate the trajectories for `Latent` and `FoF` recommenders, respectively.

`Latent` recommenders have the opposite effect, reducing global clustering. This unexpected phenomenon is due to `Latent` recommendations favoring nodes with high embedding norms. For any existing edge $i - j$, there is a disproportionately high chance that both i and j will receive recommendations concentrated on a small subset of large-normed nodes, creating closed triangles. In the homogeneous setting, such "collisions" are less likely to occur, as the probability of two nodes independently being recommended with the same node is lower. These findings highlight the importance of investigating not only between-group differences but also within-group differences.

Homophilic minority and heterophilic majority. In previous studies, it has been observed that homophilic minority groups receive an excessive amount of exposure from recommendations, leading to increased disparities in homophily between groups. We simulate

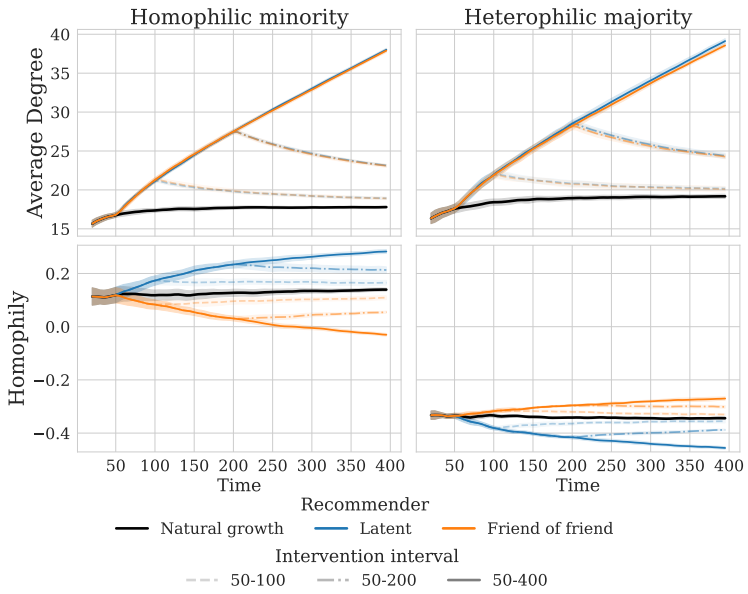


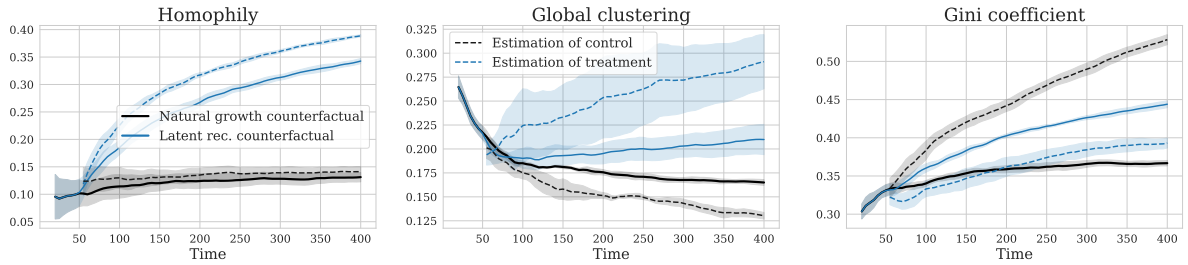
Figure 4.8: Homophily impacts. Trajectories of average degree and community homophily for `Latent` and `FoF` for the heterophilic majority and homophilic minority.

the network evolution for a majority fraction of 60%, $\mu_1 = [0, 1]$ and $\mu_2 = [1.1, 1.1]$[5] and find that our results support this conclusion for `Latent` recommendations. In contrast, Figure 4.8 shows that while `FoF` increases homophily for the majority, it decreases it significantly for the minority. The `FoF` recommender amplifies the visibility of the majority more than that of the minority. This phenomenon can be explained by the heterophilic nature of the majority group, where most of its direct connections are with minority nodes. For a majority node i, its neighbors' neighbors that are in the minority group are likely to also be direct neighbors of i. Therefore, there is a larger probability that i is recommended with other majority nodes at distance 2, increasing the homophily for the majority group.

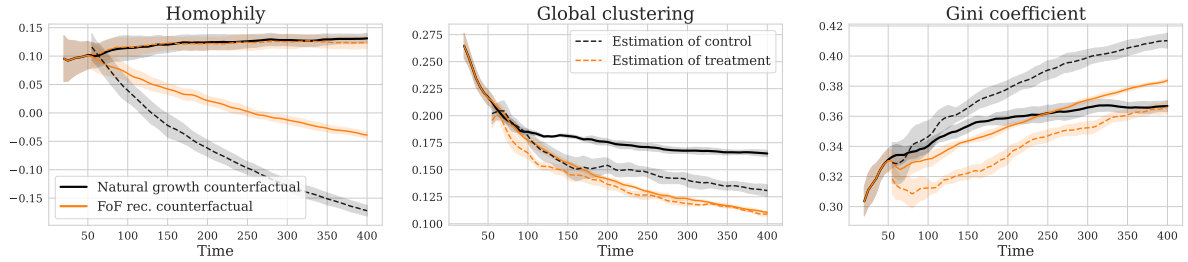
4.7 Evaluation Biases

As we do not have access to the full range of counterfactual measurements, we often resort to either longitudinal or A/B evaluations. However, these methods have their limitations and may fail to accurately capture the impacts of recommendations. In this section, we explore the potential biases and limitations of these evaluation procedures.

[5]by increasing the norm of latent representations in group 2 we increase the within-group connection probabilities.



(a) `Latent` recommender



(b) `FoF` recommender

Figure 4.9: A/B evaluations (naively): Solid lines correspond to ground truth counterfactual evaluations of homophily, clustering and Gini coefficient. Dashed lines correspond to trajectories estimated *naively* from running an A/B test.

Longitudinal evaluation. Simulating a longitudinal evaluation mimics an observational study. In this setting, a single trajectory is observed. The estimated total effect of link recommendations is the temporal difference between the value of the metric before and after the intervention. If the underlying dynamics for a metric are stationary, as is the case for homophily in Figure 4.3, then a longitudinal evaluation is unbiased. Conversely, when under natural dynamics a metric is non-stationary, a naive comparison between the initial network at time $\underline{t}$ and the network at time T could yield qualitatively misleading estimates. For instance, in Figure 4.3, when measuring the impact of `FoF` on clustering coefficient, a longitudinal measurement would compare the solid orange trajectory between time $\underline{t} = 50$ and $T = 400$, thus over-estimating about double the size of the true effects. Furthermore, the naive observational measurement for `Latent` would incorrectly suggest a negative effect on network clustering, when in reality it has a positive impact compared to the natural evolution.

A/B evaluation. A/B evaluations in dynamic networks become complex due to changes of network structure over time caused by natural dynamics and algorithmic interventions. We simulate A/B tests aimed at estimating causal effects of `Latent` and `FoF` link recommendations. We perform random treatment ($p = 0.5$) on node assignment for clustering and Gini coefficient and community-based treatment assignment for measuring homophily.

The estimates in Figure 4.9 are not adjusted for network interference. The quality of

A/B estimates varies based on the metric and intervention. Without correction, the A/B estimated trajectories not only produce quantitatively wrong results (magnitude of effect), but often produce quantitatively wrong conclusions (sign of effect). Across most settings, the quality of the metric deteriorates over time, further supporting the existence of dynamic effects that compound network interference effects.

4.8 Discussion

In this chapter, we explored the dynamic effects of link recommendations on network evolution through simulations. Our proposed extension of the Jackson-Rogers model provides insight into the impact of link recommendations on network structure. By emphasizing the importance of temporal dynamics and measurement timing, our simulations revealed surprising and persistent effects of link recommendations.

Using synthetic data, simple network evolution models, and recommendation algorithms, we addressed "what-if" scenarios in a controlled setting. This approach serves as a valuable first step in understanding these effects in real networks. Our results demonstrated that link recommendations can have delayed and indirect impacts on network structure, with long-lasting effects even after the recommendations have ceased. These findings underscore the significant cascading indirect effects over time and highlight the need for further research in evaluating link recommendation algorithms in dynamic networks.

Moreover, our study sheds light on potential biases that can arise when evaluating link recommendation algorithms in dynamic networks. We found that longitudinal metrics or A/B tests can produce biased estimates if the underlying network dynamics are not stationary. This underscores the need for advanced estimation procedures that account for both network interference and underlying dynamic effects to accurately assess the impacts of link recommendation algorithms.

We identify several opportunities for future research. To improve the validity of our conclusions, one avenue is to validate modeling assumptions against real-world networks. Refining the modeling assumptions, such as allowing for non-recommender-driven edge creation between existing nodes, could better reflect reality and potentially result in greater indirect effects. Another important area of interest is to further develop methods for measuring direct and indirect effects in different network formation models.

A second direction of future work is to study downstream impacts of link recommendation. Modeling node embeddings as dynamic properties, influenced by the local neighborhood through biased assimilation [DGL13] or mere exposure effects [Cur+22], could provide insight into how opinion formation on networks is influenced by recommendations. Finally, examining recommendation scenarios where network edges in the social network are formed by content recommendations, rather than user recommendations, which is typical of social media sites like Instagram, is a promising area for investigation.

Chapter 5

Emergent specialization from participation dynamics and multi-learner retraining

Transitioning from the feedback dynamics in social network ecosystems discussed in Chapter 4, we now broaden our scope to examine markets where, in addition to a multitude of users, there is also a multitude of algorithmic decision-makers. In this chapter, we analyze a class of dynamics where users allocate their participation among services to minimize their individual risk, while services update their model parameters to reduce their own risk based on their current user population. We refer to these dynamics as risk-reducing, encompassing a broad range of common model updates, including gradient descent and multiplicative weights. For this general class of dynamics, we show that asymptotically stable equilibria are always segmented, with sub-populations committed to a single learner. Contrary to previous findings that repeated risk minimization with a single learner can lead to representation disparity and high overall loss, we find that repeated myopic updates with multiple learners yield better outcomes. We illustrate this phenomenon through a simulated example initialized from real data, demonstrating the benefits of multiple algorithmic decision-makers in dynamic market settings.

This chapter is based on the paper "Emergent Specialization from Participation Dynamics and Multi-Learner Retraining" written in collaboration with Sarah Dean, Lillian Ratliff, Jamie Morgenstern and Maryam Fazel.

5.1 Background

Many online platforms, including social media networks, personalized recommendation engines, and advertising auction systems, collect user data and make incremental adjustments to the models they use to personalize content. These continuous updates are motivated by many

factors, though large amongst them is the fact that the systems operate in non-stationary environments, where the preferences of their users change as the system operates. Changes in user preferences might occur exogeneously of service settings (e.g., global events might spur interest in new topics) or endogeneously (e.g., increasing the ranking of certain content on a platform might lead the content to "go viral"). The fact that user behavior might depend on service settings can take on many forms: people may learn to ignore or avoid clicking on advertisements; they may choose to use the service only for tasks at which it already works well; or they may choose to switch to a different service if they have a better experience with the second service. The latter two examples of adaptation, where users might opt for services that already suit their needs, affect the system's capacity to learn about its user base and improve its overall performance.

In this chapter, we study a particular form of endogenously shifting distributions over multiple rounds, in contexts where individuals prefer to use services whose predictions are more accurate for them. Much of the existing work on endogeneous distribution shift focuses on users who modify their features to achieve desired outcomes, as in strategic classification [Har+16] and related problems [Per+20; MPZ21]. While important, this model of data manipulation does not capture the most straightforward way that individuals express their preferences in a market: by choosing amongst alternative providers. In fact, recent work has shown that in the presence of a choice of participation between competing providers, individuals do not have an incentive to perform costly data manipulations [HJM22].

Consider, as an example, a social media platform. If the platform recommends content that does not appeal to the tastes of younger generations, these users will spend a smaller fraction of their time on that platform. This results in a positive (i.e. self-reinforcing) feedback loop, where a services's poor performance on young customers dissuades them from using the service, leading to less data and diminishing weight placed on making better predictions for young customers in the future. Within a single service, these effects may lead to representation disparity [Has+18].

However, in a broader ecosystem, individuals can choose *amongst* services. If a new social media platform can predict the tastes of younger users more accurately, the younger users may spend more of their time on the new service, and correspondingly less on an existing platform. The new platform will then receive more data and improve its performance on young customers, while the old platform's predictions may deteriorate, reinforcing their exit. Similarly, in the context of Large Language Models (LLM), if one LLM performs particularly well on creative tasks and another on answering homework questions, the distribution of prompts each receives may shift towards their existing expertise. Such feedback loops can also arise in settings such as music recommendation or healthcare, where demographic and socio-economic factors explain some of the emerging specialization.

Contributions

- **Formulation:** In Section 5.2 we introduce *risk-reducing* populations and services who choose their actions myopically, incrementally improving their utility based on current

conditions.

- **Analysis:** We present a complete characterization of stable fixed points for this general class of dynamics in Section 5.4 and in Section 5.7 we illustrate our theory with simulated experiments.

- **Social implications:** By drawing a connection between the dynamics and the *total risk* we characterize the implications of this dynamic in terms of a utilitarian notion of *social welfare*, and argue that increasing the number of available services leads to better outcomes in terms of accurate predictions and user experience.

Related Work

The study of equilibria in the presence of utility optimizing agents has classical roots in game theory, and optimization over decision-dependent probabilities is classically studied by stochastic optimization and control (e.g., the review article by [HBT18]); we narrow our focus to the most relevant literature on this as it arises in machine learning systems.

Endogenous Distribution Shifts. In the study of machine learning systems, a large body of literature studies exogenous distribution shifts such as covariate, label, or concept drift [Qui+08]. A more recent trend is to study shifts in the underlying data distribution due to endogenous reactions, for example due to strategic behavior exhibited by a user population.

The work of [Per+20] introduces *performative prediction* as a model capturing user reaction via endogenous distribution shifts. This work models a single decision-maker facing a risk minimization problem subject to an underlying decision-dependent data distribution. Following its introduction, several relevant solution concepts have been explored and algorithms for achieving them proposed [IYZ21; DX20; Men+20; MPZ21]. A variant of the single decision-maker performative prediction problem studies time-dependent dynamics of the data distribution, with both exogenous [WBD21; CDH21] and endogenous [Ray+22; BHK22] sources. These works primarily consider strategic covariate shifts in a single distribution. In contrast, we consider a mixture of distributions: sub-populations of users whose participation choices result in attrition and retention dynamics which are not studied in the aforementioned distribution shift literature.

Multiple Decision-Makers. Endogenous distribution shift has also been studied in settings with multiple decision-makers as a continuous game. For instance, the multi-player performative prediction problem extends the original problem by allowing for multiple competing decision-makers [Nar+22; PY22; WD22]. This line of work differs from ours in that the population is modeled as homogeneous and stateless. These works focus on characterizing the existence and uniqueness of different types of competitive equilibria for the game, and analyze learning dynamics that lead to different equilibrium concepts. In contrast, we the focus is on asymptotically stable points (equilibrium) for the combined dynamical system resulting from myopic optimization by non-anticipating decision-makers and stateful user participation updates.

Retention. User retention in machine learning systems is closely related to the population participation dynamics we consider [Has+18; Zha+19]. In settings with multiple subpopulations of users of different types, the question of retention has been explored in parallel with the issue of fairness. [Has+18] coined the term *representation disparity* for the phenomenon in which the traditional approach of minimizing average performance leads to high overall accuracy coupled with low accuracy on minority groups, causing an exodus of said groups. For single learners, systems which instead perform robust risk minimization avoid such disparity. Our work generalizes the single-learner retention setting and analyzes the fixed points of dynamics between multiple systems and populations without modifying risk functions to be robust. [Gin+21] also consider user choice between multiple learning systems, with an empirical investigation and theoretical results in restricted settings focused on finite sample effects. In contrast, we propose a general class of *risk-reducing* dynamics and develop a comprehensive theoretical understanding.

Motivating Examples

We discuss several real-world examples that exhibit degrees of market segmentation across characteristics such as nationality, age, and race. In these examples, market conditions are certainly affected by more complex phenomena, from network effects to explicit competition between firms. While we do not claim that the dynamics we study are necessarily the main contributing factor, our simple model isolates the potential contribution of learning dynamics: namely, to reinforce such segmentation. This perspective highlights the potential effects of efforts to incorporate data or improve personalization.

Social Media

Usage of various social media sites in the US varies across genders[1] and age groups[2]. For example, the users of Facebook and LinkedIn skew older while Snapchat, Tiktok, Tumblr, and Twitch are more heavily used by the younger population. Similarly users of Pinterest strongly skew female while users of Twitch are more likely to be male. Figure 5.1 shows the disparities along gender and age for leading social media platforms. These disparities across platforms are reinforced by user behaviors: imaging the experience of a 45 year old logging onto Twitch for the first time compared with a 14 year old; or instead imagine a 14 year old logging into Facebook. Because the usage patterns determine the data available to the platforms, the disparities are also reinforced by the behavior of the platforms themselves. Similarly, Pinterest algorithms are more likely to be tailored to the tastes of an female demographic, while Twitch's to a demographic more representative of males.

[1] Gender distribution
[2] Age distribution

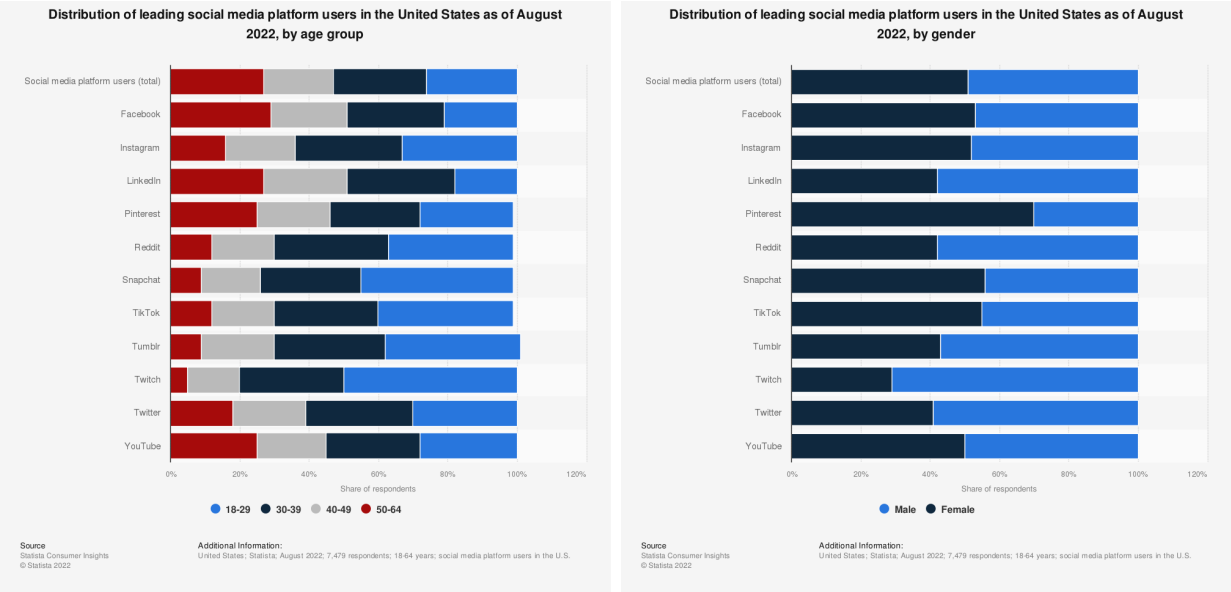
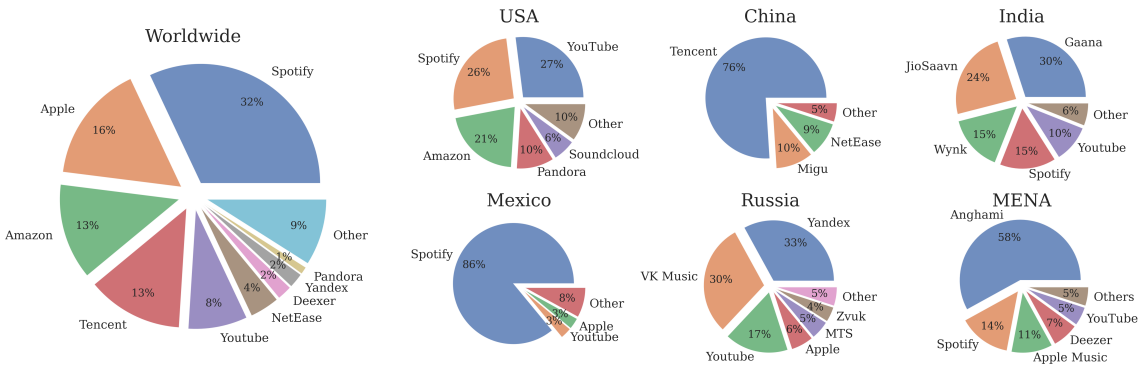


Figure 5.1: Social media usage across leading social media platforms. Left: Age distribution. Right: Gender distribution

Music Streaming

Worldwide market share of music streaming services is split between several companies (see Figure 5.2). However, the distribution of music streaming by country shows clear patterns: most users in China use Tencent, most users in Mexico use Spotify, and most users in the Middle East and Northern Africa (MENA) use Anghami. On the other hand, the markets United States, Russia, and India are not dominated by a single service. However, the handful of most used services in these regions have a small market outside of their main market. Due to this segmented market, only certain platforms collect large scale data about music preferences in certain regions. If many users from western cultures make playlists containing both Arabic and Indian music, Spotify may learn to associate those genres in a way that is undesirable or even offensive to users from those cultures. This leads to a self-reinforcing effect: services who make bad predictions for users from certain cultures are unlikely to correct this bias as those users choose instead to use services that more accurately reflect their tastes.

[3]All statistics recorded from Statista: Worldwide, United States, China, India,Mexico, Russia, Middle East and North Africa.



(a) Worldwide usage

(b) Usage by country

Figure 5.2: Usage of music streaming services in different markets[3]. Left: Worldwide marked share. Right: Market share in USA, China, India, Mexico, Russia, Middle East and Northern Africa (MENA).

Personalized health

The growing popularity of direct-to-consumer genetic testing is driven by the growth of two market leaders: AncestryDNA and 23andMe[4]. These tests are used both for determining ancestry as well as receiving polygenic risk scores for various medical conditions. The accuracy of the tests varies across ethnic groups; with Latino, Middle Eastern and, African ancestry being most under-represented. This issue is self re-inforcing; for instance people of African descent are less likely to use a large service like 23andMe and more likely to use a specialized service such as AfricanAncestry[5].

5.2 Framework and Setting

We consider a setting where the population of individuals is composed of n subpopulations spreading their participation amongst m learners (service providers or decision-makers). Figure 5.3 illustrates a simple example. Each subpopulation $i \in \{1, \dots, n\} =: [n]$ has features and labels distributed according to a fixed distribution $(x, y) =: z \sim \mathcal{D}_i$ and makes up β_i proportion of the total population, so that $\sum_{i=1}^{n} \beta_i = 1$. An α_{ij} proportion of subpopulation i is associated to each learner $j \in [m]$, normalized so that $\sum_{j=1}^{m} \alpha_{ij} = 1$. The subpopulations therefore redistribute their participation among the various learners. Further, to model the ability of subpopulations to opt-out, one can include a static "null learner".

A subpopulation can be as broad as a demographic or affinity group and as granular as a single individual. The allocation of a subpopulation can represent several things: the

[4] https://www.statista.com/chart/17023/commercial-genetic-testing/

[5] https://africanancestry.com/

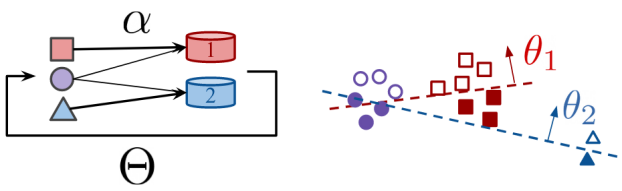


Figure 5.3: $n = 3$ subpopulations (□, ○, △) select among $m = 2$ learners (red, blue) based on classification accuracy with respect to label (solid, hollow). Parameters $\Theta = (\theta_1, \theta_2)$ (decision lines) update in response to subpopulations participation $\alpha_{i,j}$. At the current state, the circle subpopulation will shift participation towards blue learner.

fraction of a subpopulation which uses a given service, or the *fraction of time* users from that subpopulation choose to spend using learners' systems. Accordingly, the relative size β_i of the population can represent the proportion of individuals or the *total time* individuals spend. This framework also allows for a subpopulation to represent *types of tasks or activities* a user wishes to accomplish, allocating these tasks to learners based upon which systems perform best on which tasks. *The only assumption we make about any subpopulation is that individual samples comprising it are i.i.d.*

Throughout, we assume that there are fewer learners than subpopulations, $m \leq n$. Each learner j observes data from the subpopulations who participate in it. Formally it observes features and labels drawn from the mixture distribution determined by the participation and subpopulation sizes:

$$(x, y)_j = z_j \sim \frac{\sum_{i=1}^{n} \alpha_{ij}\beta_i \mathcal{D}_i}{\sum_{i=1}^{n} \alpha_{ij}\beta_i}$$

Learners make predictions or decisions according to a parameter $\theta_j \in \mathbb{R}^d$. Beyond the information encoded in the features and labels, the learners are unaware of which subpopulation individual data points are.

The quality of predictions made by parameter $\theta_j \in \mathbb{R}^d$ for an individual instance z_j is quantified by the loss $\ell(\theta_j; z_j)$. The quality of θ_j for a subpopulation is quantified by the average loss, i.e. the *risk* $\mathcal{R}_i(\theta_j) = \mathbf{E}_{z\sim\mathcal{D}_i}[\ell(\theta_j; z)]$. Throughout, we will make the additional assumption that the risk function for each subpopulation $\mathcal{R}_i(\theta)$ is convex and differentiable. Figure 5.4 illustrates an example of the risk functions arising in linear regression.

Decision dynamics of learners and subpopulations

Subpopulations and learners react to each other; Updates in subpopulation allocations lead to updates in learners parameters $\Theta^t = (\theta_1^t, \ldots, \theta_m^t)$, and vice versa. We introduce a broad class of update dynamics by way of a canonical example. Suppose that each subpopulation i updates its allocation by increasing the participation proportional to the quality of various models; for example, by spending more time on recommendation platforms that suggest more

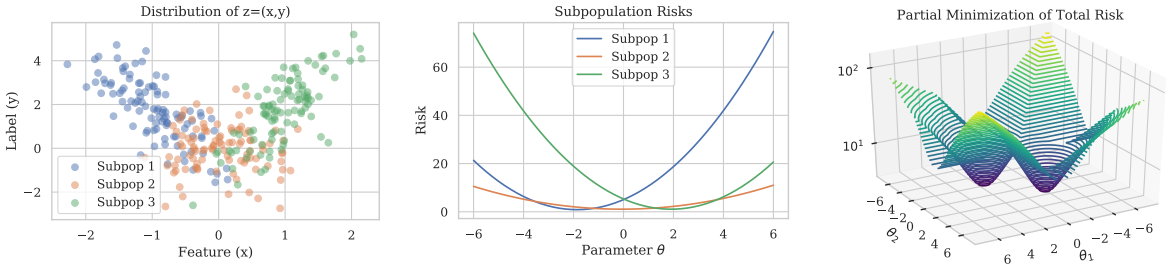


Figure 5.4: An example arising from least-squares linear regression with $n = 3$ subpopulations and $m = 2$ learners. Left: The distribution of $z = (x, y)$, colored by subpopulation. Middle: The subpopulation risks $\mathcal{R}_i(\theta)$ arising from least-squares linear regression $\ell(\theta; z) = (y - \theta x)^2$. Right: A visualization of the non-convex total risk as a function of learner parameters, via the partial minimization over subpopulation allocation: $\min_\alpha \sum_{i=1}^3 \sum_{j=1}^2 \alpha_{ij} \mathcal{R}_i(\theta_j) = \sum_{i=1}^3 \min\{\mathcal{R}_i(\theta_1), \mathcal{R}_i(\theta_2)\}$.

engaging content. Recalling that the risk (i.e. average loss) quantifies quality, this manifests as a *multiplicative weights update*: $\alpha_{ij}^{t+1} \propto \alpha_{ij}^t \cdot \exp(-\gamma \mathcal{R}_i(\theta_j))$ for $j \in [m]$ and some parameter $\gamma > 0$. This is similar to the retention function studied by [Has+18] and has connections to replicator dynamics, a foundational evolutionary dynamic that can be interpreted as a process of information diffusion and imitation [San20].

Recall that each learner j observes data from the mixture distribution:

$$\left(\sum_{i=1}^n \alpha_{ij} \beta_i \right)^{-1} \sum_{i=1}^n \alpha_{ij} \beta_i \mathcal{D}_i,$$

for which we use the shorthand $\mathcal{D}(\alpha_{:,j})$, where $\alpha_{:,j} \in \mathbb{R}^n$ denotes the vector of allocations from all subpopulations to learner j. Suppose the learners update their parameters using gradient descent to reduce the average loss over this data (e.g. to improve the prediction of user engagement). Setting aside finite sample issues, for a step size γ_t the gradient update takes the form $\theta_j^{t+1} = \theta_j^t - \gamma_t \nabla_\theta \mathbb{E}_{z \sim \mathcal{D}(\alpha_{:,j})} \left[\ell(\theta_j^t; z) \right]$. This is an incremental version of the *repeated retraining dynamics* which have been studied in the single learner setting by [Has+18; Per+20].

Despite the apparent simplicity of independent update rules, the evolution of subpopulations and learners is highly coupled. The sequential interaction between subpopulations and learners leads to complex nonlinear dynamics: i.e. multiplicative weights over non-stationary risks (due to learner updates) and gradient descent over non-stationary data distributions (due to subpopulation updates).

To study this complex behavior, we now formalize key properties. The first observation is that updates are *stateful*, with subpopulation allocations and learner parameters depending on previous values. This motivates a description of the dynamics arising from interactions between n subpopulations and m learners in terms of the overall state $\alpha \in \Delta_m \times \cdots \times \Delta_m =: \Delta_m^n$

and $\Theta \in \mathbb{R}^{m\times d}$. We thus define for each subpopulation i a general Markovian allocation function $\nu_i^t : \Delta_m \times \mathbb{R}^{m\times d} \to \Delta_m$ which describes the participation update $\alpha_{i,:}^{t+1} = \nu_i^t\left(\alpha_{i,:}^t, \Theta^t\right)$ at time t, where $\alpha_{i,:} \in \Delta_m$ denotes the vector of allocations from the subpopulation i to all learners. Similarly, define $\mu_j^t : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^d$ so learner j updates their parameter according to $\theta_j^{t+1} = \mu_j^t(\theta_j^t, \alpha_{:,j}^t)$.

The second observation is that the basis for the updates is the average loss, i.e. risk. This motivates the following definition: given participation α and parameters Θ, the average risk experienced by each subpopulation i and each learner j is:

$$\bar{\mathcal{R}}_i^{\mathsf{subpop}}(\alpha_{i,:}, \Theta) := \mathop{\mathbb{E}}_{j\sim\alpha_{i,:}}\left[\mathop{\mathbb{E}}_{z\sim\mathcal{D}_i}\left[\ell(\theta_j; z)\right]\right],$$
$$\bar{\mathcal{R}}_j^{\mathsf{learner}}(\alpha_{:,j}, \theta_j) := \mathop{\mathbb{E}}_{z\sim\mathcal{D}(\alpha_{:,j})}\left[\ell(\theta_j; z)\right].$$

In the recommendation example, $\bar{\mathcal{R}}^{\mathsf{subpop}}$ captures the dissatisfaction with content for a subpopulation and $\bar{\mathcal{R}}^{\mathsf{learner}}$ corresponds to the average prediction error of the platform. Intuitively, multiplicative weights reduces the average subpopulation risk while gradient descent reduces the average learner risk.

Definition 5.2.1 (Reducing and Minimizing Dynamics)**.** A u update rule is *P-reducing* w.r.t. v if $P(u^{t+1}, v^t) \le P(u^t, v^t)$ for all t and any sequence of v^t. It is further *P-minimizing in the limit* if the inequality is strict when u^t is not a minimizer and $\lim_{t\to\infty} P(u^t, v) = \min_u P(u, v)$.

We call a subpopulation i *risk-reducing* (resp. minimizing) when the allocation update on $\alpha_{i,:}$ is $\bar{\mathcal{R}}_i^{\mathsf{subpop}}$-reducing (resp. minimizing in the limit) with respect to Θ. Similarly, we call a learner j *risk-reducing* (resp. minimizing) when the parameter update on θ_j is $\bar{\mathcal{R}}_j^{\mathsf{learner}}$-reducing (resp. minimizing in the limit) with respect to $\alpha_{:,j}$.

We remark that the notion of risk minimizing in the limit is reasonable for subpopulations because their average risk is linear in $\alpha_{i,:}$. It is also reasonable for learners because their average risk is convex in θ_j (due to the assumption that risks $\mathcal{R}_i(\theta_j)$ are convex). However, risk-reducing/minimizing is only a property defined with respect to the participation α or parameter Θ observed at a previous time step. Thus it does not necessarily hold that $\bar{\mathcal{R}}_j^{\mathsf{learner}}$ or $\bar{\mathcal{R}}_i^{\mathsf{subpop}}$ decrease when the state evolves $(\alpha^t, \Theta^t) \to (\alpha^{t+1}, \Theta^{t+1})$ by sequential updates of ν^t and μ^t. Our experiments (Figure 5.6a) illustrate the non-monotonicity of the coupled updates.

Example 5.2.2 (Semi-static participation)**.** Suppose a population has a constant allocation of 20% to one learner, while the remaining 80% is allocated to the remaining learners inversely proportional to the learner's risk on that population. This is risk-reducing but not risk minimizing in the limit.

Example 5.2.3 (Full risk minimization)**.** Suppose that a learner updates its parameter to minimize the average risk function $\bar{\mathcal{R}}_j^{\mathsf{learner}}(\alpha_{:,j}^t, \cdot)$ at each timestep. This has been studied as *repeated retraining dynamics* in the single learner case by [Has+18; Per+20].

Example 5.2.4 (Non-continuity of allocation updates)**.** Suppose a population prefers one learner over others, and only shifts participation away from the preferred learner if there is another with risk smaller by at least $R_0 > 0$. This is risk-reducing but not minimizing in the limit.

Example 5.2.5 (Shifting to lower-risk models)**.** If a subpopulation's allocation updates always shift allocation from learners with high subpopulation risk to learners with lower subpopulation risk, then the allocation is risk-reducing. It may or may not be risk minimizing in the limit.

Example 5.2.6 (Allocations determined by gradient descent)**.** Consider an allocation determined by (projected) gradient descent with respect to a subpopulation's average risk. This is risk-reducing, and may be risk minimizing in the limit depending on the step-size.

Proposition 5.2.7. *A subpopulation i updating their participation with multiplicative weights is risk minimizing in the limit if $\gamma > 0$ and $\alpha_{ij}^0 > 0\ \forall j$. A learner updating its parameter with gradient descent is risk minimizing in the limit when the risk functions $\mathcal{R}_i(\theta)$ are L smooth and the step size satisfies $\gamma^t < \frac{2}{L}$, $\sum_{t=0}^{\infty} \gamma^t = \infty$, and $\sum_{t=1}^{\infty} (\gamma^t)^2 < \infty$.*

Proof. To see that the subpopulation is risk minimizing, first see that

$$\begin{aligned}
\bar{\mathcal{R}}_i^{\mathsf{subpop}}\left(\alpha_{i,:}^{t+1}, \Theta\right) &= \sum_{j=1}^{m} \alpha_{ij}^{t+1} \mathcal{R}_i(\theta_j) \\
&= \sum_{j=1}^{m} \frac{\alpha_{ij}^t \cdot \exp(-\gamma \mathcal{R}_i(\theta_j))}{\sum_{j=1}^{m} \alpha_{ij}^t \cdot \exp(-\gamma \mathcal{R}_i(\theta_j))} \mathcal{R}_i(\theta_j) \\
&< \sum_{j=1}^{m} \alpha_{ij}^t \mathcal{R}_i(\theta_j) = \bar{\mathcal{R}}_i^{\mathsf{subpop}}\left(\alpha_{i,:}^t, \Theta\right)
\end{aligned}$$

where the strict inequality holds as long as α_{ij}^t is not on the boundary of the simplex. Second, observe that for a fixed Θ, $\alpha_{ij}^t \to 1$ if and only if $\mathcal{R}_i(\theta_j)$ is minimal over all learners for which $\alpha_{ij}^0 > 0$.

To see that the learner is risk minimizing, notice that the gradient update is equivalently

$$\theta_j^{t+1} = \theta_j^t - \gamma_t \nabla_\theta \bar{\mathcal{R}}_j^{\mathsf{learner}}\left(\alpha_{:,j}, \theta_j\right) .$$

Gradient descent on an L-smooth and convex function leads to strictly decreasing objective values when θ_j^t is not at a minimum and the step size satisfies $\gamma^t < \frac{2}{L}$. It further converges to a minimum in the limit as long as the step size satisfies the Robbins-Munroe condition (see, e.g. [LY22; Ora20]). □

5.3 Equilibria and Stability

We focus on the equilibrium states resulting from risk-reducing subpopulations and learners.

Definition 5.3.1 (Equilibrium)**.** The state $(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}})$ is an equilibrium state if it is stationary under the dynamics update $\{\nu_i^t\}$, $\{\mu_j^t\}$; i.e. that for all $i \in [n]$ and $j \in [m]$:

$$\alpha_{i,:}^{\mathsf{eq}} = \nu_i^t(\alpha_{i,:}^{\mathsf{eq}}, \Theta^{\mathsf{eq}}) \quad \text{and} \quad \theta_j^{\mathsf{eq}} = \mu_j^t(\alpha_{:,j}^{\mathsf{eq}}, \theta_j^{\mathsf{eq}}) \, .$$

If learners and subpopulations are in an equilibrium state, they will remain that way indefinitely. However, some equilibrium configurations may be unstable to perturbations.

Definition 5.3.2 (Stable Equilibrium)**.** The state $(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}})$ is a *stable* equilibrium state if it is an equilibrium and for each $\varepsilon_\alpha, \varepsilon_\theta > 0$, there exist $\delta_\alpha, \delta_\theta > 0$ such that

$$\begin{array}{l} \|\alpha^0 - \alpha^{\mathsf{eq}}\| < \delta_\alpha, \\ \|\Theta^0 - \Theta^{\mathsf{eq}}\| < \delta_\theta \end{array} \implies \begin{array}{l} \|\alpha^t - \alpha^{\mathsf{eq}}\| \leq \varepsilon_\alpha, \, \forall \; t \geq 0. \\ \|\Theta^t - \Theta^{\mathsf{eq}}\| \leq \varepsilon_\theta \end{array}$$

It is further *asymptotically stable* if $\lim_{t\to\infty} \|\alpha^t - \alpha^{\mathsf{eq}}\| = 0$ and $\lim_{t\to\infty} \|\Theta^t - \Theta^{\mathsf{eq}}\| = 0$.

Stability analysis identifies qualitatively different equilibrium states. For the class of risk-reducing dynamics that we study, equilibria may be unstable, stable, or asymptotically stable;

Examples

To illustrate the subtleties of determining stability when the total risk function has non-isolated local minima, we consider a setting with $n = m = 2$ subpopulations and learners where $\mathcal{R}_1(\theta) = \mathcal{R}_2(\theta) = \theta^2$. Then the total risk function is minimized for any $\alpha \in \Delta_m^n$ and $\Theta = (0, 0)$. This continuum of minima can contain equilibria of risk minimizing dynamics, and those equilibria may be stable, asymptotically stable, or unstable, which we illustrate with the following examples.

Example 5.3.3 (Continuum of stable balanced markets)**.** Suppose that subpopulations update their allocation via any Lipschitz continuous risk minimizing update rule which is stationary whenever learners are risk equivalent (i.e. $\mathcal{R}_i(\theta_1) = \mathcal{R}_i(\theta_2)$). Suppose that learners update via full risk minimization. Then equilibria will have the form $(\alpha^{\mathsf{eq}}, (0, 0))$ for any $\alpha^{\mathsf{eq}} \in \Delta_2^2$.

Then starting from any (α^0, Θ^0) with a $\delta_\alpha, \delta_\theta$ ball of any equilibrium $(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}})$,

$$\alpha^1 = \nu(\alpha^0, \Theta^0), \quad \Theta^1 = (0, 0)$$

at which point the system is in a new equilibrium, since any allocation α is a fixed point when $\Theta = (0, 0)$ so $\alpha^t = \alpha^1$ and $\Theta^t = \Theta^1$ for all t. We have that $\|\Theta^{\mathsf{eq}} - \Theta_0\| = 0$ and

$$\|\alpha^{\mathsf{eq}} - \alpha^1\| = \|\nu(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}}) - \nu(\alpha^0, \Theta^0)\|$$

By the assumption of Lipschitzness, this distance will scale linearly in $\delta_\alpha, \delta_\theta$ so the definition of stability is satisfied for δ chosen proportionally to ε depending on the Lipschitz constant of ν.

In this example, any perturbation converges to a new fixed point within one time step. The continuity of the update functions ensures that the new fixed point is within a bounded distance of the original, satisfying the definition of stability. This example is not asymptotically stable: the allocation does not convergence back to the original point.

Example 5.3.4 (Asymptotically stable segmentation)**.** Consider the subpopulation and learner update rules as in the prior example, with one amendment. When $\mathcal{R}_i(\theta_1) = \mathcal{R}_i(\theta_2)$, subpopulation 1 re-allocates half of its mass from learner 2 to learner 1, and while subpopulation 2 re-allocates half its mass from learner 1 to learner 2. Thus the subpopulation update can be written as

$$\alpha_{1,:}^{t+1} = \begin{cases} \nu_1(\alpha_{1,:}^t, \Theta^t) & \mathcal{R}_1(\theta_1) \neq \mathcal{R}_1(\theta_2) \\ \begin{bmatrix} 1 & 1/2 \\ 0 & 1/2 \end{bmatrix} \alpha_{1,:}^t & \mathcal{R}_1(\theta_1) = \mathcal{R}_1(\theta_2) \end{cases}, \quad \alpha_{2,:}^{t+1} = \begin{cases} \nu_2(\alpha_{2,:}^t, \Theta^t) & \mathcal{R}_2(\theta_1) \neq \mathcal{R}_2(\theta_2) \\ \begin{bmatrix} 0 & 1/2 \\ 1 & 1/2 \end{bmatrix} \alpha_{2,:}^{t+1} & \mathcal{R}_2(\theta_1) = \mathcal{R}_2(\theta_2) \end{cases}$$

The only equilibrium has α^{eq} segmented with subpopulation i associated to learner i for $i = 1, 2$ and $\Theta^{\mathsf{eq}} = (0, 0)$. It is straightforward to see that this is an asymptotically stable equilibrium, since for any $a \in \Delta_2$,

$$\lim_{t\to\infty} \begin{bmatrix} 1 & 1/2 \\ 0 & 1/2 \end{bmatrix}^t a = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} a = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \lim_{t\to\infty} \begin{bmatrix} 0 & 1/2 \\ 1 & 1/2 \end{bmatrix}^t a = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} a = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Example 5.3.5 (Asymptotically stable balanced market)**.** Consider a setting similar to the previous example except that when $\mathcal{R}_i(\theta_1) = \mathcal{R}_i(\theta_2)$, subpopulation i moves half the mass from group 1 to group 2 and half the mass from group 2 to group 1 for all i. Then the subpopulation update can be written as

$$\alpha_{1,:}^{t+1} = \begin{cases} \nu_1(\alpha_{1,:}^t, \Theta^t) & \mathcal{R}_1(\theta_1) \neq \mathcal{R}_1(\theta_2) \\ \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \alpha_{1,:}^t & \mathcal{R}_1(\theta_1) = \mathcal{R}_1(\theta_2) \end{cases}, \alpha_{2,:}^{t+1} = \begin{cases} \nu_2(\alpha_{2,:}^t, \Theta^t) & \mathcal{R}_2(\theta_1) \neq \mathcal{R}_2(\theta_2) \\ \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \alpha_{2,:}^{t+1} & \mathcal{R}_2(\theta_1) = \mathcal{R}_2(\theta_2) \end{cases}$$

The only equilibrium has $\alpha_{i,:}^{\mathsf{eq}} = [1/2, 1/2]$ for $i = 1, 2$ and $\Theta^{\mathsf{eq}} = (0, 0)$. It is straightforward to see that this is an asymptotically stable equilibrium, since for any $a \in \Delta_2$,

$$\lim_{t\to\infty} \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}^t a = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} a = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}.$$

Example 5.3.6 (Unstable balanced market)**.** Suppose that subpopulation allocations follow a projected gradient descent update for all i:

$$\alpha_{i1}^{t+1} = \mathrm{Proj}_{[0,1]}\left(\alpha_{i1}^t - \gamma(\mathcal{R}_i(\theta_1) - \mathcal{R}_i(\theta_2))\right)$$

and $\alpha_{i2} = 1 - \alpha_{i1}$. Further suppose learners update with gradient descent:

$$\theta_j^{t+1} = \theta_j^t - \frac{1}{2\sqrt{t}} \nabla \bar{\mathcal{R}}_j^{\mathsf{learner}}(\alpha_{:,j}^t, \theta_j^t) = \sqrt{\frac{t}{t+1}} \theta_j^t$$

Both rules are risk minimizing in the limit (note that $\theta_j^t = \frac{1}{\sqrt{t}} \theta_j^0$) and have a continuum of equilibria at any $\alpha^{\mathsf{eq}} \in \Delta_m^n$ and $\Theta^{\mathsf{eq}} = (0, 0)$. However, we show that the equilibria are not stable. Consider the initial condition $(\alpha^{\mathsf{eq}}, (\delta_\theta, 0))$. We have that

$$\alpha_{i1}^{t+1} = \mathrm{Proj}_{[0,1]} \left(\alpha_{i1}^0 - \gamma \delta_\theta^2 \sum_{k=1}^{t} \frac{1}{k} \right) \to 0 \quad \text{as} \quad t \to \infty.$$

No matter how small the perturbation δ_θ is, the summation increases with t and participation will converge all weight to learner 2. A similar argument shows that perturbations exist that will send all participation to learner 1.

In this example, the learners update slowly. Despite eventual convergence to the minimizing parameter, the accumulating error causes the participation allocation to shift completely to the unperturbed learner, precluding stability.

While a quantitative understanding of convergence may also be of interest, it would require stronger assumptions on the behavior of subpopulations and learners; here we favor generality and leave this to future work. Furthermore, characterizing stable equilibria sets the foundation for understanding high probability behavior of systems under noisy updates which are risk-reducing only in expectation [Kus67]. This sets the stage for finite sample risk minimization or multi-agent user models, a challenge which we leave to future work.

5.4 Main Results

We study a large class of feedback dynamics between risk-reducing learners and subpopulations described by the sequential updates: $\alpha^{t+1} = \nu^t(\alpha^t, \Theta^t)$ and $\Theta^{t+1} = \mu^t(\alpha^{t+1}, \Theta^t)$.

Our analysis allows for learners and subpopulations who exhibit a diverse range of behaviors. We do not require that every learner or every subpopulation update their parameter or allocation in the same manner or even at every timestep, allowing for any number of round-robin schemes. Our only assumption on learner and subpopulation updates is that they are risk-reducing or minimizing.

Figure 5.5 presents a summary of the equilibria characterization that we present in this section. All omitted proofs can be found in Section 5.9.

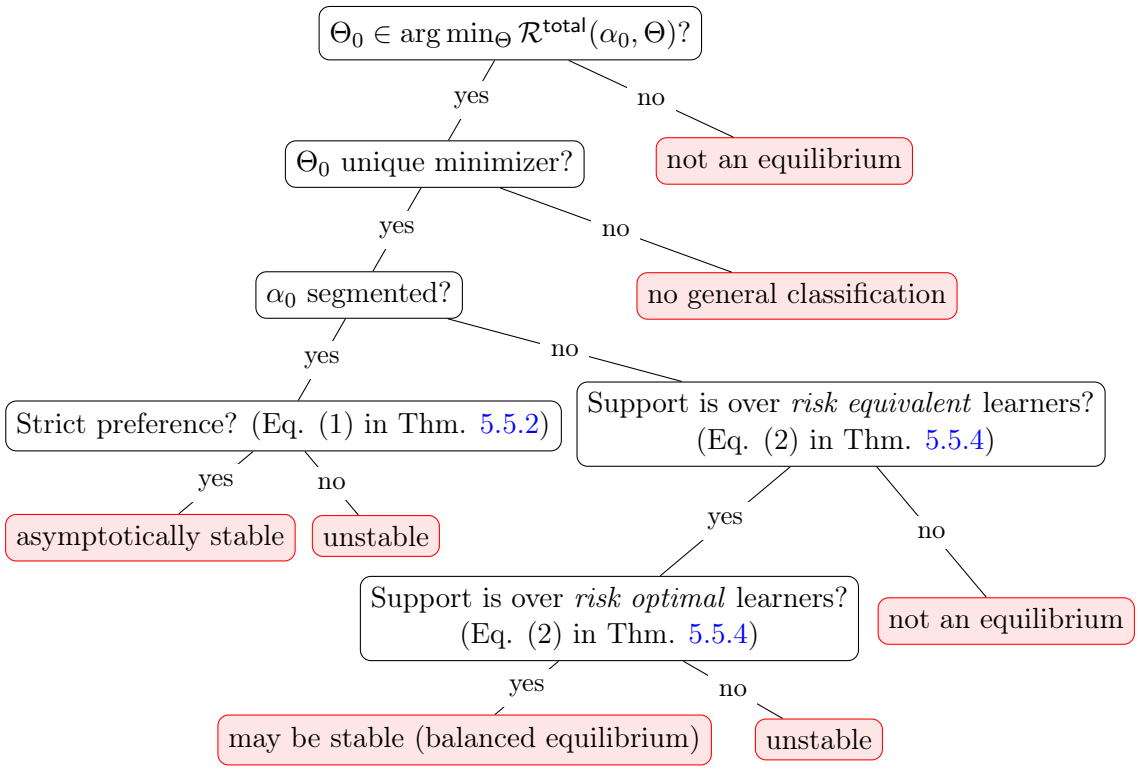


Figure 5.5: A summary of our main results on equilibria classification for a given participation α_0 and model parameters Θ_0. These results hold for dynamics which are risk minimizing in the limit and loss functions that are convex.

Total Risk Reduction

Definition 5.4.1 (Total Risk)**.** The *total risk* of all subpopulations over all learners is the weighted sum

$$\mathcal{R}^{\text{total}}(\alpha, \Theta) := \sum_{i=1}^{n} \sum_{j=1}^{m} \beta_i \alpha_{ij} \mathcal{R}_i(\theta_j).$$

The total risk maps $\Delta_m^n \times \mathbb{R}^{m \times d} \to \mathbb{R}$. While our assumption that the loss is convex implies that the total risk is convex in Θ, it is not jointly convex in (α, Θ), illustrated in the right panel of Figure 5.4.

Our first result shows that the total risk $\mathcal{R}^{\text{total}}(\alpha^t, \Theta^t)$ is non-increasing over time.

Proposition 5.4.2. *For any risk-reducing subpopulation and learner dynamics, the total risk is non-increasing:* $\mathcal{R}^{\text{total}}(\alpha^{t+1}, \Theta^{t+1}) \leq \mathcal{R}^{\text{total}}(\alpha^t, \Theta^t), \forall t$. *If subpopulations and learners are risk minimizing in the limit, then the total risk is strictly decreasing unless* (α^t, Θ^t) *is a local minimizer of* $\mathcal{R}^{\text{total}}$.

Proof Sketch. First note that the total risk can be decomposed into either a weighted sum of average subpopulation risk or average learner risk. Thus the fact that learner and subpopulation dynamics are risk-reducing ensures that the total risk is decreasing after the sequential updates. □

Thus, the total risk acts like a potential function for the feedback dynamics of learners and subpopulations. When the subpopulation and learner dynamics are risk minimizing in the limit, there is a strong connection between properties of the total risk function and equilibria of the dynamics.

Theorem 5.4.3. *For any learners and subpopulations who are risk minimizing in the limit, an equilibrium* $(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}})$ *is asymptotically stable if it is an isolated local minimizer of the total risk* $\mathcal{R}^{\mathsf{total}}$*. If it is not a local minimizer of the total risk, then it is not stable.*

Proof Sketch. By Proposition 5.4.2 the function $V(\alpha, \Theta) := \mathcal{R}^{\mathsf{total}}(\alpha, \Theta) - \mathcal{R}^{\mathsf{total}}(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}})$ is potential function for the autonomous dynamical system $(\alpha^t, \Theta^t) \to (\alpha^{t+1}, \Theta^{t+1})$. The stability result follow from Lyapunov arguments. □

The connection between stability and the total risk function is significant in at least two ways: first, it means that under general classes of myopic and self-interested behaviors on the part of subpopulations and learners, the total risk is driven to at least a local minimum. Second, it is a technically useful connection that will enable us to characterize and classify the stable equilibria for dynamics which are risk minimizing in the limit. We remark that Theorem 5.4.3 leaves open the question of stability for equilibria which are non-isolated minima of the total risk function. In Section 5.3, we provide examples which show that such points may be asymptotically stable, stable, or unstable depending on the particular instantiation of dynamics. The following existence result further motivates our focus dynamics which are risk minimizing, rather than just reducing.

Corollary 5.4.4. *Equilibria exist when learners and subpopulations are risk minimizing in the limit and the total risk function has isolated local minima. They may not exist otherwise.*

Example of dynamics without equilibria. Consider subpopulations with risk functions minimized at the same value θ^*. If learners use full risk minimization, the setting lacks isolated minima because the total risk is uniform across all allocations α. Assuming that risk-minimizing subpopulations randomly choose among equivalent learners, no equilibrium exists as allocations randomly switch between learners once the learners converge to the optimum θ^*. □

5.5 Segmented and Balanced Equilibria

Definition 5.5.1 (Segmented allocation)**.** An allocation is *segmented* if $\alpha_{ij} \in \{0, 1\}$ for all i, j.

In a segmented allocation, each subpopulation is associated with a single learner, and thus the population is partitioned across learners.

For allocation dynamics like multiplicative weights, such configurations are clearly equilibria for any parameter choice Θ on the part of the learners. We thus consider the set of possible segmented equilibria and characterize which are asymptotically stable.

Theorem 5.5.2. *Suppose learners and subpopulations are risk minimizing in the limit, α^{eq} is segmented, and $\mathcal{R}^{\mathsf{total}}(\alpha^{\mathsf{eq}}, \Theta)$ has a unique minimizer Θ^{eq}. Define a mapping $\gamma : [n] \to [m]$ such that $\gamma(i) = j$ is the learner with nonzero mass in $\alpha^{\mathsf{eq}}_{i,:}$.*

If every subpopulation strictly prefers their current learner:

$$\mathcal{R}_i(\theta^{\mathsf{eq}}_{\gamma(i)}) < \mathcal{R}_i(\theta^{\mathsf{eq}}_j)\,, \tag{5.1}$$

for all i and learners $j \neq \gamma(i)$, then $(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}})$ is an asymptotically stable equilibrium. If there is a subpopulation who would strictly prefer to switch learners, then $(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}})$ is not stable.

When risks are strongly convex, there is always such a unique minimizer Θ^{eq}. In particular, in a segmented allocation, each θ^{eq}_j minimizes the average loss over the group of subpopulations assigned to them.

Corollary 5.5.3. *Suppose that risk functions satisfy $\mathcal{R}_i(\theta) < \mathcal{R}_i(\theta') \iff \|\theta - \phi_i\| < \|\theta' - \phi_i\|$ for ϕ_i the subpopulation optimal parameter. Then in an asymptotically stable segmented equilibrium, the convex hulls of the grouped subpopulations optimal parameters $\{\phi_i\}$ are non-intersecting.*

Proof Sketch. Consider a partition where the convex hulls intersect for some pair of learners.

Then there exists at least one subpopulation who would be better off switching to the other learner, and thus the risk condition in Theorem 5.5.2 cannot hold. □

Applying the Corollary to the example in Figure 5.4, we see that a segmented equilibrium with subpopulation 1 and 3 participating in the same learner cannot be stable.

Theorem 5.5.2 leaves open the question of stability in the case that the risks in Equation (5.1) are equal. Under such *risk equivalence*, is it natural to consider equilibria where a subpopulation has support over multiple learners.

Theorem 5.5.4. *Consider dynamics which are risk minimizing in the limit and an α^{eq} with any subpopulation i having nonzero support on set of two or more learners $j \in \mathcal{J}$. Assume risks are strongly convex and define $\Theta^{\mathsf{eq}} = \arg\min \mathcal{R}^{\mathsf{total}}(\alpha^{\mathsf{eq}}, \Theta)$. Then $(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}})$ cannot be stable unless it is "balanced" in the sense that learners in $\mathcal{J}$ are* risk equivalent *and* optimal *for i, i.e. for all $j, j' \in \mathcal{J}$,*

$$\mathcal{R}_i(\theta^{\mathsf{eq}}_j) = \mathcal{R}_i(\theta^{\mathsf{eq}}_{j'}) \quad \textit{and} \quad \nabla\mathcal{R}_i(\theta^{\mathsf{eq}}_j) = 0\,. \tag{5.2}$$

If it is balanced, so are all allocations for subpopulation i with support over $\mathcal{J}$.

Finally, all stable equilibria must be either balanced or segmented.

This result characterizes a set of possibly stable equilibria. It demonstrates that risk *optimality*, in addition to *equivalence*, is necessary. Guaranteeing the stability of such balanced equilibria requires further information about the dynamics, and it is not possible to make a general statement.

Examples in Section 5.3 demonstrate that such balanced equilibria may be asymptotically stable, stable, or unstable.

Furthermore, the balance condition is fragile in the sense that it would not hold under small perturbations to the underlying risk functions. While the number of possible balanced equilibria is combinatorial in the number of learners and subpopulations, risk functions are continuous, so it is possible to find arbitrarily small perturbations to any the risk functions that would destabilize all balanced equilibria.

Proof Sketch of Theorem 5.5.2 and Theorem 5.5.4. By Theorem 5.4.3, characterizing the stable equilibria is equivalent to characterizing isolated and non-isolated local minima of the total risk. We show that it suffices to characterize local minima of the partial minimization $F(\alpha) = \min_\Theta \mathcal{R}^{\mathsf{total}}(\alpha, \Theta)$ over the simplex product $\Delta_m^n = \Delta_m \times \cdots \times \Delta_m$. Since F is concave, all minima occur on the boundary, i.e. a face or a vertex. Since F is still concave when restricted to a face of the simplex, the same argument shows the minima are on the boundary, hence vertices, except for the degenerate case where F takes a constant value over the face.

Thus, the isolated local minima occur at vertices of the simplex product, which correspond to segmented allocation. Further analysis of F yields the conditions presented in Theorem 5.5.2. The local minima in the degenerate case are characterized by the balanced equilibria conditions in Theorem 5.5.4. □

5.6 Social Welfare for Segmented Populations

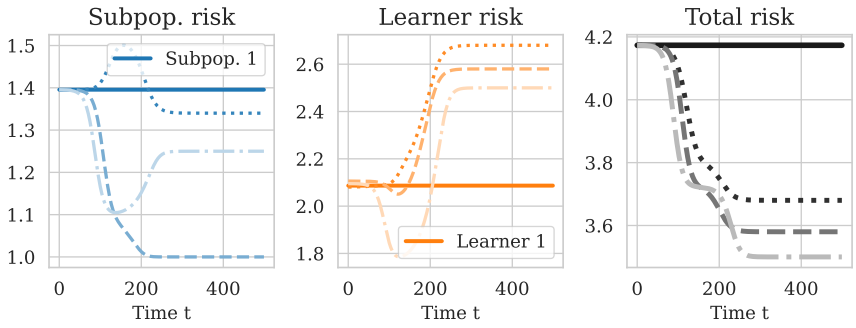
Definition 5.6.1. The *social welfare* of a state (α, Θ) is strictly decreasing in the total risk $\mathcal{R}^{\mathsf{total}}(\alpha, \Theta)$.

This definition of social welfare is utilitarian in the sense that it depends on the cumulative quality of individuals' experiences. Maximizing the social welfare corresponds to minimizing the total risk, which can be posed as the following optimization problem

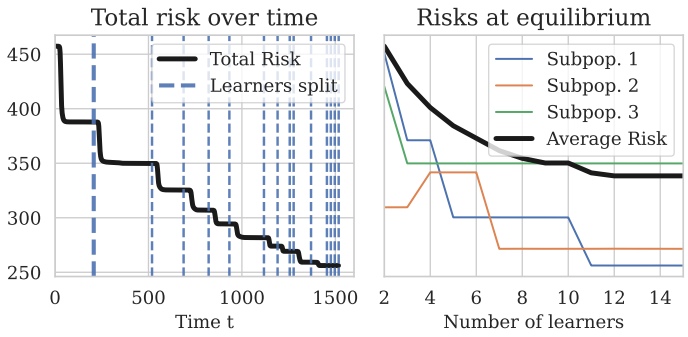
$$\begin{aligned} (\alpha^\star, \Theta^\star) \in \arg\min_{\alpha,\Theta} \quad & \mathcal{R}^{\mathsf{total}}(\alpha, \Theta) \\ \text{s.t.} \quad & \alpha_{i,:} \in \Delta_m \ \ \forall \ \ i = \{1, \dots n\} \, . \end{aligned} \tag{5.3}$$

Here, $(\alpha^\star, \Theta^\star)$ is the social welfare maximizer.

Our discussion of stable equilibria has so far focused on only local minimizers of the total risk. In fact, global minimization of this objective (and therefore maximization of social welfare) is a hard problem. The total risk objective can be viewed as an instance of the k-means clustering problem with $k = m$. In the language of this literature (e.g., [SI84]), each



(a) Risk dynamics



(b) Impact of competition on social welfare

Figure 5.6: **Synthetic settings:** Figure (a) illustrates a setting with 3 subpopulations and 2 learners. The dsolid lines correspond to the risk trajectory for the unstable balanced equilibrium at initialization. Dotted and dashed lines illustrate risk trajectories under three different slight perturbations from the initialization. In Figure (b), the left plot illustrates the reduction in total risk over time. The dashed blue lines indicate when a new learner joins. The right plot shows the equilibrium-risk for a subset of the subpopulations as the number of learners increases.

subpopulation is a data point and the parameter selected by each learner is a cluster center. The allocations described by α correspond to (fuzzy) cluster assignment and each risk function $\mathcal{R}_i(\theta_j)$ corresponds to a measure of "dissimilarity" between data points (subpopulations) and cluster centers (learners).

The connection to k-means clustering elucidates the difficulty of minimizing the total risk. The "minimum sum-of-squares clustering" problem (i.e., squared Euclidean norm dissimilarity) is NP hard with general dimension even when $k = 2$ [Alo+09]. When the number of clusters and dimension are fixed, [IKI94] present an algorithm for solving the minimum sum-of-squares clustering problem which is polynomial in the number of datapoints. Translated to our setting, its complexity is $O(n^{md})$. It is therefore unrealistic to hope that a myopic dynamic might generally lead to social welfare maximization. However, due to the connections with

total risk, risk-reducing dynamics are at least well-behaved with regards to social welfare.

Proposition 5.6.2. *For risk-reducing subpopulations and learners, social welfare is non-decreasing over time. If the dynamics are furthermore risk minimizing in the limit, social welfare is strictly increasing and stable equilibria correspond to local social welfare maxima.*

Proof. Social welfare is non-decreasing (or increasing) if and only if total risk is non-increasing (or decreasing), as guaranteed by Proposition 5.4.2. Maxima of the social welfare are equivalent to minima of the total risk and therefore the connections to stable equilibria follow by Theorem 5.4.3. □

Local maximization is not a panacea: Example 5.6.3 shows a local maximum of the social welfare can be much worse than the global one.

Example 5.6.3 (Arbitrarily high total risk at local optimum)**.** Consider three subpopulations with

$$\mathcal{R}_1(\theta) = \theta^2, \quad \mathcal{R}_2(\theta) = (\theta - 1)^2, \quad \mathcal{R}_2(\theta) = (\theta - \phi)^2$$

for some $\phi > 2$. Suppose that subpopulation sizes are $\beta_1 = \beta_2 = \beta$ and $\beta_3 = 1 - 2\beta$ for some $0 < \beta < 1/2$. In addition, suppose that there are two learners. Up to permutation, the social welfare optimum is $\theta_1 = 1/2$ and $\theta_2 = \phi$, with total risk $\beta/2$. However, as long as $\phi < \frac{1-\beta}{1-2\beta}$, there is another stable equilibrium. Let $\phi = \frac{1-\beta}{1-2\beta} - \varepsilon$. Then the following is a stable equilibrium: $\theta_1 = 0$ and $\theta_2 = 1 - \varepsilon$. The total risk is $\beta + \frac{(\beta-\varepsilon)^2}{1-2\beta}$. For β close to $1/2$, this risk can be arbitrarily larger than the social optimum.

In this example, a large gap between a stable local optimum and the global optimum arises in part due to a large difference in subpopulations' sizes. We further remark that minority groups can be under-served particularly when considering worst-case risk over subpopulations [Has+18]. Even at a social welfare maximizer $(\alpha^\star, \Theta^\star)$, the worst-case subpopulation risk can be arbitrarily bad. It is straightforward to construct such examples even in the single learner case: consider a minority group with vanishingly small population proportion and arbitrarily high risk at the optimal parameter for the majority group (Example 5.6.7).

Despite these inherent difficulties, we find that the situation improves as the number of learners increases. It is straightforward to see that the maximal social welfare will increase: any point which is optimal for m learners can be trivially transformed into a feasible point for $m + 1$ learners which achieves the same social welfare, by allocating no subpopulations to the new learner. There is more nuance involved when considering any possible stable equilibria. Instead, we make a statement about a particular learner growth process which corresponds to existing learner m "splitting in half".

Proposition 5.6.4. *Suppose that risks are strongly convex, there are m learners, $(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}})$ is an equilibrium, and at least one subpopulation i allocated to learner m does not have optimal subpopulation risk, so $\nabla\mathcal{R}_i(\theta_m^{\mathsf{eq}}) \neq 0$. The state is amended to add an additional learner: $\widetilde{\Theta}^{\mathsf{eq}} = [\Theta^{\mathsf{eq}}, \theta_m^{\mathsf{eq}}]$ and*

$$\widetilde{\alpha}^{\mathsf{eq}}_{:,j} = \begin{cases} \alpha^{\mathsf{eq}}_{:,j} & j \leq m \\ \frac{1}{2}\alpha^{\mathsf{eq}}_{:,m} & j \in \{m, m+1\} \end{cases}$$

Under dynamics which are risk minimizing in the limit, the equilibrium $(\widetilde{\alpha}^{\mathsf{eq}}, \widetilde{\Theta}^{\mathsf{eq}})$ *is not stable, so a small perturbation will send the system to a state with strictly lower total risk (higher social welfare).*

Proof. By construction $(\widetilde{\alpha}^{\mathsf{eq}}, \widetilde{\Theta}^{\mathsf{eq}})$ is not segmented, and neither is it a stable balanced equilibrium (by the non-optimality assumption). Therefore, it is not stable (Theorem 5.5.4), and thus not a local minimum of the total risk (Theorem 5.4.3). A perturbation will thus send the system along a risk-reducing trajectory. □

re optima.

Example 5.6.5 (Stability vs. optimality)**.** Consider three subpopulations $i \in \{1, 2, 3\}$ with risks $\|\theta - \phi_i\|_2^2$, sizes β_i, and two learners $j \in \{1, 2\}$. Suppose that the α^{eq} is such that the subpopulations are partitioned into $\{1\}$ and $\{2, 3\}$. Then we have that

$$\theta_1^{\mathsf{eq}} = \phi_1, \quad \theta_2^{\mathsf{eq}} = \frac{\beta_2}{\beta_2 + \beta_3}\phi_2 + \frac{\beta_3}{\beta_2 + \beta_3}\phi_3$$

By Theorem 5.5.2, this is stable if and only if

$$\|\phi_2 - \phi_3\|_2 \leq (\beta_2 + \beta_3) \min\left\{ \frac{\|\phi_2 - \phi_1\|_2}{\beta_3}, \frac{\|\phi_3 - \phi_1\|_2}{\beta_2} \right\} .$$

However, it is only social optimal if and only if ϕ_2 and ϕ_3 are relatively close to each other than to ϕ_1, i.e.

$$\|\phi_2 - \phi_3\|_2 \leq \min\left\{\|\phi_2 - \phi_1\|_2, \|\phi_3 - \phi_1\|_2\right\} .$$

The set of subpopulation optima $\{\phi_1, \phi_2, \phi_3\}$ satisfying the optimality condition are a subset of those satisfying the stability condition. As the difference between β_2 and β_3 becomes more extreme, the number of settings satisfying the stability but not optimality condition increases.

We use this generic example to illustrate a scenario in which the total risk can be arbitrarily high at a stable equilibria.

Example 5.6.6. Suppose there are two learners and three subpopulations with sizes $\beta_1 = \beta_2 = \beta$ and $\beta_3 = 1 - 2\beta$ for some $0 < \beta < 1/2$. Consider the following: $\mathcal{R}_1(\theta) = \theta^2$, $\mathcal{R}_2(\theta) = (\theta - 1)^2$, $\mathcal{R}_2(\theta) = (\theta - \frac{1-\beta}{1-2\beta} + \varepsilon)^2$. The social welfare optimizing decision $\Theta^\star = (1/2, \frac{1-\beta}{1-2\beta} - \varepsilon)$ corresponds to total risk $\beta/2$. However, there is a stable equilibrium at $\Theta^{\mathsf{eq}} = (0, 1 + \varepsilon)$ with total risk $\beta + \frac{(\beta-\varepsilon)^2}{1-2\beta}$. For $\beta \to 1/2$, the gap becomes arbitrarily large.

Finally, we present an example which illustrates that even in the single learner setting, the risk of a subpopulation can be arbitrarily worse than the total risk.

Example 5.6.7 (Arbitrarily high risk for minority subpopulation)**.** Consider two subpopulations with $\mathcal{R}_1(\theta) = \theta^2$ and $\mathcal{R}_2(\theta) = (\theta - \phi)^2$ with $\beta_1 = \beta$ and $\beta_2 = 1 - \beta$ and a single learner. The single equilibrium and total risk minimizer is $\theta_1 = (1 - \beta)\phi$ with total risk $\beta(1 - \beta)\phi^2$ and $\mathcal{R}_2(\theta^\star) = \beta^2\phi^2$. The difference between the two quantities can be arbitrarily high as β gets close to 1.

5.7 Experiments

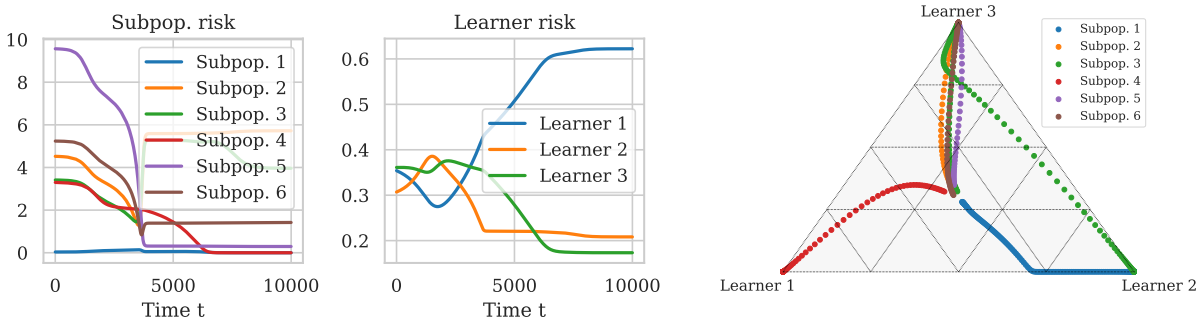
We illustrate the salient properties of the decision dynamics in simulation[6]. We consider both a synthetic setting as well as one instantiated from a prediction task on census data.

Synthetic In Figure 5.6a we consider a simple scenario with $n = 3$ subpopulations of equal sizes $\beta_i = 1/3$, quadratic risk functions $\mathcal{R}_i = \|\phi_i - \theta\|^2 + 1$ with distinct risk minimizing decisions ϕ_i and $m = 2$ learners. The learners minimize their risk according to *full risk minimization* (Example 5.2.3) and the subpopulations update their participation via multiplicative weights update (Section 5.2). When $\alpha^0_{i,j} = 1/2$ for all i, j the risk equality condition from Theorem 5.5.4 is satisfied with $\theta^{\mathsf{eq}}_j = (\phi_1 + \phi_2 + \phi_3)/3$, however the optimality condition is not. We therefore observe that this equilibrium is not stable, and slightly perturbing the initial conditions leads to split-market equilibria. Figure 5.6a illustrates trajectories from three different perturbations. It demonstrates that the total risk is non-increasing whereas the average risks for both learners and subpopulations are not monotonic. Each of the perturbations has different risk trajectories and equilibrates at a different split-market equilibrium. We repeat these experiments with noisy dynamics, we consider both exogenous noise that independently perturbs the decisions of the learners and/or populations as well as intrinsic noise due to making updates with finite sample estimates rather than at population level. We find that the key properties of the dynamics hold when the updates are noisy, detailed experiments are presented in Section 5.9.

Another set of experiments in Figure 5.6b illustrates how a larger number of learners lead to better outcomes in terms of total risk. We consider a set of $m = 2$ learners and $n = 50$ subpopulations. We simulate the dynamics until the market has reached equilibrium, at which point a randomly chosen learner breaks up into two identical learners with half the user base. From this unstable equilibrium (Proposition 5.6.4) we slightly perturb the parameters of the two learners and allow the system to reach a new equilibrium state. The procedure repeats until the number of learners reaches number of subpopulations. These simulations illustrate that more competition improves social welfare, however the improvements are not uniform for all subpopulations with some groups seeing their risk at equilibrium increase with the addition of new learners.

Census data We consider a semi-synthetic setting where subpopulations and their risk functions are instantiated by a prediction task on real data. Using `folktables` [Din+21] we consider a modified version of *ACSTravelTime* prediction problem derived from the

[6]Implementation details and reproduction instructions at:
https://github.com/mcurmei627/MultiLearnerRiskReduction



(a) Risk dynamics

(b) Allocation trajectories

Figure 5.7: **Empirical subpopulations from Census data:** Figure (a) displays the relative risk with respect to the best achievable risk for the subpopulation over time. Figure (b) illustrates how allocations initialized near $(1/3, 1/3, 1/3)$ converge to a split market equilibrium.

2018 California Census data. We consider 6 subpopulations corresponding to racial groups with relative size ranging from 1.2% to 61%. We define the least-squares risk functions as $\mathcal{R}_i(\theta) = \frac{1}{N_i}\|X_i\theta - y_i\|^2$ where $X_i \in \mathbb{R}^{N_i \times d}$ are the features (containing demographics, educational attainment, income levels, and modes of transportation) and $y_i \in \mathbb{R}^{N_i}$ are the labels (log transform of the daily commute time in minutes) for individuals within subpopulation i. We simulate risk-reducing dynamics from a perturbed balanced equilibrium over 3 learners. As in the synthetic example, the risks of learners and subpopulations are not all monotone (Figure 5.7a), but the total risk function is. Finally Figure 5.7b illustrates the convergence of allocation dynamics to a segmented equilibrium.

5.8 Discussion

In this chapter, we study the feedback dynamics of user retention for loss minimizing learners, where subpopulations choose between providers. We introduce a formal notion of *risk-reducing* and *minimzing* to capture this feedback, and show that there is a close connection between such dynamics and the *total risk* summed over subpopulations and learners. We provide a comprehensive characterization of stable equilibria and investigate the implications in terms of a utilitarian social welfare. This work relates to questions of fairness and minority representation in several ways. First, our results imply that risk-minimizing dynamics in multi-learner settings can result in higher welfare for small subpopulations compared with single-learner settings, as studied by [Has+18; Zha+19]. This resonates with recent work showing that monopolies have higher *performative power* and lead to lower individual utility [HJM22].

The dynamics that we study often lead to *segmentation* of subpopulations across learners

as an emergent phenomenon[7]. This segmentation can lead to pointwise lower risks for subpopulations, especially when subpopulations have considerably different risk profiles. In some contexts, the benefits of the reduced risk among subpopulations may outweigh possible harms from segregation. In others, where proportional representation of groups across learners, models, or clusters [KAM19; Kle+19] is important, our work implies that independent risk minimization can lead to undesirable outcomes. In short, this work analyzes natural dynamics with consequences for the distribution of subpopulations amongst independent learners; whether or not the consequences are desirable depend on the specific application considered.

We highlight several directions for future work. Our results lay the groundwork for an investigation of the stochastic dynamics that occur for finite sample approximations to the risk or participation driven by decisions of individuals. Such behaviors are risk-reducing in *expectation*, so we expect the noisy trajectories to converge with high probability to sets around the asymptotically stable equilbria we characterize. There are many interesting and relevant questions in the finite sample setting: What is the effect of sample size on the ability of new learners to enter a market and minority subpopulations to be adequately represented? Can we model heterogeneous learners who differ in which features they measure and with how much noise? Are there trade-offs between the expressivity of models and the practical difficulty of minimizing risk from finite samples in high dimensions?

It would also be interesting to consider extensions or alternative dynamics models for the learner and subpopulation decisions. One could investigate competitive learners who explicitly strategize to capture subpopulations [BT19; Ari+20]; this setting is related to facility location and Hotelling games [OD98; Hot29]. One might imagine that subpopulations do not act uniformly and may not even be entirely independent of each other—the participation update may depend on some underlying social network.

The connections between total risk reduction and k-means clustering algorithms suggest interventions such as subpopulationaware initialization [Bos+23] that could improve social welfare. Results on "ground truth recovery" may yield insight into particular population structures that lead to simpler dynamics or restricted sets of equilibria.

5.9 Ommitted Proofs

Preliminaries

We introduce a compact notation. The simplex product is defined as

$$\Delta_m^n = \left\{ A \in \mathbb{R}^{n \times m} \mid \sum_{j=1}^{m} A_{ij} = 1 \right\}$$

so that the *rows* sum to 1. Then the state space of subpopulation allocations and learner parameters is $\mathcal{X} = \Delta_m^n \times \mathbb{R}^{m \times d}$. For a square matrix A, we use the notation $\mathrm{diag}(A)$ to

[7]This connects to economic literature on "rational" discrimination, where competitors have no inherent preference to discriminate and yet equilibria are segregated, e.g. [FV92]

represent the vector containing the diagonal entries of A. For a vector a, $\mathrm{Diag}(a)$ is a diagonal matrix with a along the diagonal. Furthermore we will say $a \leq b$ for vectors a, b if the inequality holds elementwise.

Define a matrix valued risk function $R : \mathbb{R}^{m\times d} \to \mathbb{R}^{n\times m}$ so that $R_{ij}(\Theta) = \mathcal{R}_i(\theta_j)$. Recall that in Section 5.2, the subpopulation and learner risks played a key role. We therefore define vector valued functions $\bar{\mathcal{R}}^{\mathsf{subpop}} : \mathcal{X} \to \mathbb{R}^n$ and $\bar{\mathcal{R}}^{\mathsf{learner}} : \mathcal{X} \to \mathbb{R}^m$ as follows:

$$\bar{\mathcal{R}}^{\mathsf{subpop}}(\alpha, \Theta) = \mathrm{diag}(\alpha R(\Theta)^\top), \quad \bar{\mathcal{R}}^{\mathsf{learner}}(\alpha, \Theta) = \mathrm{diag}\left(\mathrm{Diag}(\alpha^\top\beta)^{-1}\alpha^\top \mathrm{Diag}(\beta) R(\Theta)\right) .$$

Then the definition of risk-reducing dynamics for subpopulations and learners can be written as

$$\bar{\mathcal{R}}^{\mathsf{subpop}}(\alpha^{t+1}, \Theta) \leq \bar{\mathcal{R}}^{\mathsf{subpop}}(\alpha^t, \Theta) \quad \text{and} \quad \bar{\mathcal{R}}^{\mathsf{learner}}(\alpha, \Theta^{t+1}) \leq \bar{\mathcal{R}}^{\mathsf{learner}}(\alpha, \Theta^t) .$$

Risk minimizing in the limit is defined similarly, where the inequality is strict for at least one entry of the vectors unless the state is at a local minimum.

The total risk can be written as

$$\mathcal{R}^{\mathsf{total}}(\alpha, \Theta) := \mathrm{tr}(\mathrm{diag}(\beta)\alpha R(\Theta)^\top) .$$

Lemma 5.9.1. *Under the assumption that all loss functions are continuous, the risk function R is continuous w.r.t. to Θ, and thus $\mathcal{R}^{\mathsf{total}}$ is continuous w.r.t. α and Θ.*

The sequential dynamics updates described in Section 5.2 can be written as

$$\begin{bmatrix}\alpha^{t+1}\\ \Theta^{t+1}\end{bmatrix} = \begin{bmatrix}\nu(\alpha^t, \Theta^t)\\ \mu(\alpha^{t+1}, \Theta^t)\end{bmatrix} = \begin{bmatrix}\nu(\alpha^t, \Theta^t)\\ \mu(\nu(\alpha^t, \Theta^t), \Theta^t)\end{bmatrix} =: f(\alpha^t,\ \Theta^t) . \tag{5.4}$$

Lemma 5.9.2. *As long as the subpopulation and learner updates described in Section 5.2 are locally Lipschitz, so is the dynamics function f defined in (5.4).*

Background

For completeness, we include important results and definitions that our proofs will make use of. First, we state two theorems about Lyapunov theory for stability.

Theorem 5.9.3 (Theorem 1.2 in [BCS18])**.** *Let $x_{\mathsf{eq}} \in \mathcal{D}$ be an equilibrium point for the autonomous systems $x_{t+1} = f(x_t)$ where $f : \mathcal{D} \to \mathcal{X}$ is locally Lipschitz in $\mathcal{D} \subseteq \mathcal{X}$. Suppose there exists a function $V : \mathcal{D} \to \mathbb{R}$ which is continuous and such that*

$$V(x_{\mathsf{eq}}) = 0 \quad \text{and} \quad V(x) > 0 \ \ \forall \ \ x \in \mathcal{D} - \{x_{\mathsf{eq}}\}$$
$$V(f(x)) - V(x) \leq 0 \quad (\text{resp. } < 0) \ \ \forall \ \ x \in \mathcal{D}$$

Then x_{eq} is stable (resp. asymptotically stable).

Theorem 5.9.4 (Theorem 1.5 in [BCS18]). *Let $x_{\mathsf{eq}} \in \mathcal{D}$ be an equilibrium point for the autonomous systems $x_{t+1} = f(x_t)$ where $f : \mathcal{D} \to \mathcal{X}$ is locally Lipschitz in $\mathcal{D} \subseteq \mathcal{X}$. Let $V : \mathcal{D} \to \mathbb{R}$ be a continuous function with $V(x_{\mathsf{eq}}) = 0$ and $V(x_0) > 0$ for some x_0 arbitrarily close to x_{eq}. Let $r > 0$ be such that $B_r(x_{\mathsf{eq}}) \subseteq \mathcal{D}$ and $\mathcal{U} = \{x \in B_r(x_{\mathsf{eq}}) \mid V(x) > 0\}$, and suppose that $V(f(x)) - V(x) > 0$ for all $x \in \mathcal{U}$. Then x_{eq} is not stable.*

Next, we state the definition of a (isolated) local minimum.

Definition 5.9.5. The point $u_\star$ is a local minimum (resp. isolated local minimum) of a function h over a domain $\mathcal{U}$ if there is a $\delta > 0$ such that for any $u \in \mathcal{U}$ with $\|u - u_\star\| \leq \delta$, $h(u_\star) \leq h(u)$ (resp. $h(u_\star) < h(u)$).

Next, we state the implicit function theorem.

Theorem 5.9.6 (Implicit Function Theorem). *Let $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$ be open sets and $f : U \times V \to \mathbb{R}$ is C^r for some $r \geq 1$. For some $x_0 \in U$, $y_0 \in V$ assume the partial derivative in the second argument $D_2 f(x_0, y_0) : \mathbb{R}^m \to \mathbb{R}$ is an isomorphism. Then there are neighborhoods U_0 of x_0 and W_0 of $f(x_0, y_0)$ and a unique C^r map $g : U_0 \times W_0 \to V$ such that for all $(x, w) \in U_0 \times W_0$, $f(x, g(x, w)) = w$.*

Finally we prove a property of intersecting convex hulls.

Lemma 5.9.7. *Let $x_1, x_2, \cdots x_n$ and $y_1, y_2, \ldots, y_m$ be some points in $\mathbb{R}^d$. Define by $\mathcal{C}_x$ and $\mathcal{C}_y$ the convex hulls of $\{x_i\}_{i=1}^n$ and $\{y_i\}_{i=1}^m$ respectively. Then there do not exist points $\bar{x} \in \mathbb{R}^d$ and $\bar{y} \in R^d$ such that the following inequalities are satisfied:*

$$\|x_i - \bar{x}\| < \|x_i - \bar{y}\| \ \forall i = 1, 2, \ldots, n$$
$$\|y_j - \bar{y}\| < \|y_j - \bar{x}\| \ \forall j = 1, 2, \ldots, m$$

Proof. Assume by contradiction that the inequalities above hold. Define $\mathcal{H}_x := \{z \in R^d \mid \|z - \bar{x}\| < \|z - \bar{y}\|\}$ and $\mathcal{H}_y := \{z \in R^d \mid \|z - \bar{y}\| < \|z - \bar{x}\|\}$. The sets $\mathcal{H}_x$ and $\mathcal{H}_y$ are disjoint half-spaces (without boundary) then defined by the hyperplane bisecting the segment connecting $\bar{x}$ and $\bar{y}$. By assumption then we have that $x_i \in \mathcal{H}_x$ for all i and $y_j \in \mathcal{H}_y$ for all j; since $\mathcal{H}_x$ and $\mathcal{H}_y$ are convex, it follows that $\mathcal{C}_x \subset \mathcal{H}_x$ and $\mathcal{C}_y \subset \mathcal{H}_y$. Therefore $\mathcal{C}_x \cap \mathcal{C}_y = \emptyset$, which leads to a contradiction. □

Properties of partial minimization

In this section, we state a handful of important results about the partial minimization of the total risk. This is somewhat similar to the analysis presented by [SI84] in the context of clustering algorithms.

Lemma 5.9.8. *Define the function $F : \mathbb{R}^{m \times n} \to \mathbb{R}$ as $F(\alpha) = \min_\Theta \mathcal{R}^{\mathsf{total}}(\alpha, \Theta)$. This function is concave and a point (α^0, Θ^0) is a local minimum of $\mathcal{R}^{\mathsf{total}}$ over the domain $\mathcal{X} =$*

$\mathcal{X}_\alpha \times \mathbb{R}^{m\times d}$ *if* α^0 *is a local minimum of* F *over the domain* $\mathcal{X}_\alpha$ *and* $\Theta^0 \in \arg\min_\Theta \mathcal{R}^{\mathsf{total}}(\alpha^0, \Theta)$. *Furthermore, in the case that* Θ^0 *is the unique minimizer of* $\mathcal{R}^{\mathsf{total}}(\alpha^0, \Theta)$, *then* (α^0, Θ^0) *is a local minimum (resp. isolated local minimum) if and only* α^0 *is a local minimum (resp. isolated local minimum).*

Proof. $F(\alpha)$ is well defined due to the convexity of the risk functions. Concavity follows from the observation that F is the point-wise minimum of a family of functions which are linear in α (since for every fixed Θ, the total risk is linear in α).

We break the proof of equivalence into two implications.

1. F minimized $\implies$ $\mathcal{R}^{\mathsf{total}}$ minimized

There is a $\delta > 0$ such that for any $\alpha \in \mathcal{X}_\alpha$ with $\|\alpha^0 - \alpha\| \le \delta$, $F(\alpha^0) \le F(\alpha)$, i.e.

$$\mathcal{R}^{\mathsf{total}}(\alpha^0, \Theta^0) \le \mathcal{R}^{\mathsf{total}}(\alpha, \Theta^*(\alpha))$$

for any minimizing $\Theta^*(\alpha)$. For fixed allocation α define $\mathcal{R}^{\mathsf{total}}_\alpha(\Theta) = \mathcal{R}^{\mathsf{total}}(\alpha, \Theta)$ which is is convex and minimized at $\Theta^*(\alpha)$ and hence:

$$\mathcal{R}^{\mathsf{total}}(\alpha, \Theta^*(\alpha)) \le \mathcal{R}^{\mathsf{total}}(\alpha, \Theta), \quad \forall \ \ \Theta\,.$$

Combining the inequalities yields: $\mathcal{R}^{\mathsf{total}}(\alpha^0, \Theta^0) \le \mathcal{R}^{\mathsf{total}}(\alpha, \Theta)$, and thus (α^0, Θ^0) is a local minimum of $\mathcal{R}^{\mathsf{total}}$. The implication for the isolated local minimum case follows by the same arguments with strict inequalities on the total risk, noting that if Θ^0 is a unique minimizer, it must also be isolated.

2. $\mathcal{R}^{\mathsf{total}}$ minimized $\implies$ F minimized

Recall that $\mathcal{R}^{\mathsf{total}}(\alpha, \Theta)$ can be written as $\mathrm{tr}(\mathrm{diag}(\beta)\alpha R(\Theta)^\top)$. Then

$$\mathcal{R}^{\mathsf{total}}(\alpha^0 + D, \Theta^0) - \mathcal{R}^{\mathsf{total}}(\alpha^0, \Theta^0) = \mathrm{tr}(\mathrm{diag}(\beta) D R(\Theta^0)^\top) \ge 0$$

where inequality holds for all $D \in \mathbb{R}^{n\times m}$ such that $\alpha^0 + D \in \mathcal{X}_\alpha$ by the fact that α^0 is a minimum. Recognizing the gradient from Lemma 5.9.9 and using the uniqueness of Θ^0, the expression is equivalently $\langle \nabla_\alpha F(\alpha^0), D\rangle \ge 0$. In other words, the directional derivative in any feasible direction D is non-negative. Hence, α^0 is a local minimum of F. The implication for the isolated local minimum case follows by the same arguments with strict inequalities on the total risk.

□

Lemma 5.9.9. *For* $F : \mathbb{R}^{n\times m} \to \mathbb{R}$ *defined as in Lemma 5.9.8, suppose the minimizier* $\Theta^*(\alpha) = \arg\min_\Theta \mathcal{R}^{\mathsf{total}}(\alpha, \Theta)$ *is unique. The gradient is*

$$\nabla_\alpha F(\alpha) = \mathrm{diag}(\beta) R\left(\Theta^*(\alpha)\right), \ \ \textit{i.e.} \quad \frac{\partial F(\alpha)}{\partial \alpha_{ij}} = \beta_i \mathcal{R}_i(\theta_j^*(\alpha))\,.$$

Further suppose that the risks are strongly convex. Then second partial derivatives are given by

$$\frac{\partial^2 F(\alpha)}{\partial \alpha_{k\ell}\partial \alpha_{ij}} = \begin{cases} 0 & k \ne j \\ -\beta_i \nabla \mathcal{R}_i(\theta_j^\star)^\top \left(\sum_{\ell'} \beta_{\ell'} \alpha_{\ell' j} \nabla^2 \mathcal{R}_{\ell'}(\theta_j^\star)\right) \nabla \mathcal{R}_\ell(\theta_j^\star) & k = j \end{cases}.$$

Proof. Computing the gradient:

$$\nabla_\alpha F(\alpha) = \nabla_\alpha \mathcal{R}^{\mathsf{total}}(\alpha, \Theta^\star(\alpha)) + \nabla\Theta^*(\alpha)\nabla_\theta \mathcal{R}^{\mathsf{total}}(\alpha, \Theta^\star(\alpha)) = \mathrm{diag}(\beta)R(\Theta)\,.$$

The first equality follows by product rule. The second equality follows because 1) the total risk is linear in α and 2) the second term is zero due to the optimality of $\Theta^*(\alpha)$.

Now notice that

$$\frac{\partial}{\partial \alpha_{k\ell}} \mathcal{R}_i(\theta_j^\star(\alpha)) = \left\langle \frac{\partial \theta_j^\star(\alpha)}{\partial \alpha_{k\ell}}, \nabla_\theta \mathcal{R}_i(\theta_j^\star(\alpha)) \right\rangle$$

To compute the derivatives of $\theta_j^\star(\alpha)$ we use the implicit function theorem and the assumption that the risks are strongly convex. We apply the implicit function theorem to the first order optimality condition

$$\theta_j^\star(\alpha) \in \arg\min_{\theta_j} \bar{\mathcal{R}}_j^{learner}(\alpha_{:,j}, \theta_j)$$

The Hessian $\nabla_\theta^2 \bar{\mathcal{R}}_j^{learner}(\alpha, \Theta))$ is non-degenerate due to strong convexity of the subpopulation risks. There exists a neighborhood U_0 of α and a unique (sufficiently smooth) map $\theta_j^*(\cdot)$ such that for all $\alpha \in U_0$, we have that $\nabla_\theta \bar{\mathcal{R}}_j^{learner}(\alpha, \theta^*(\alpha)) = 0$. Then by implicit function theorem we obtain

$$\nabla\theta_j^\star(\alpha) = -\nabla_\theta^2 \bar{\mathcal{R}}_j^{learner} \circ \nabla_{\alpha\theta} \bar{\mathcal{R}}_j^{learner}(\alpha_{:,j}, \theta_j^\star(\alpha))$$

by taking the derivative of the first order condition differentiating through $\theta_j^\star(\cdot)$ and setting it to zero. We have that

$$\nabla_\theta^2 \bar{\mathcal{R}}_j^{learner} = \sum_{\ell'} \beta_{\ell'}\alpha_{\ell' j}\nabla^2 \mathcal{R}_{\ell'}(\theta_j^\star), \quad \frac{\partial}{\partial \alpha_{k\ell}} \nabla_\theta \bar{\mathcal{R}}_j^{\mathsf{learner}} = \begin{cases} 0 & k \neq j \\ \nabla \mathcal{R}_\ell(\theta_j^{\mathsf{eq}}) & k = j \end{cases}.$$

The result follows by combining the expressions. □

Connections between dynamics and total risk

Proposition 5.4.2. *For any risk-reducing subpopulation and learner dynamics, the total risk is non-increasing: $\mathcal{R}^{\mathsf{total}}(\alpha^{t+1}, \Theta^{t+1}) \leq \mathcal{R}^{\mathsf{total}}(\alpha^t, \Theta^t), \forall t$. If subpopulations and learners are risk minimizing in the limit, then the total risk is strictly decreasing unless (α^t, Θ^t) is a local minimizer of $\mathcal{R}^{\mathsf{total}}$.*

Proof. The key to seeing that the total risk acts like a potential for the market dynamics is to note two equivalent decompositions of the total risk:

$$\mathcal{R}^{\mathsf{total}}(\alpha, \Theta) = \beta^\top \bar{\mathcal{R}}^{\mathsf{subpop}}(\alpha, \Theta) = \beta^\top \alpha \bar{\mathcal{R}}^{\mathsf{learner}}(\alpha, \Theta)\,.$$

Being risk-reducing learners' updates satisfy:

$$\bar{\mathcal{R}}^{\mathsf{learner}}(\alpha^t, \Theta^{t+1}) \leq \bar{\mathcal{R}}^{\mathsf{learner}}(\alpha^t, \Theta^t) \implies \mathcal{R}^{\mathsf{total}}(\alpha^t, \Theta^{t+1}) \leq \mathcal{R}^{\mathsf{total}}(\alpha^t, \Theta^t)\,.$$

Similarly risk-reducing subpopulations satisfy:

$$\bar{\mathcal{R}}^{\mathsf{subpop}}(\alpha^{t+1}, \Theta^{t+1}) \leq \bar{\mathcal{R}}^{\mathsf{subpop}}(\alpha^{t}, \Theta^{t+1}) \implies \mathcal{R}^{\mathsf{total}}(\alpha^{t+1}, \Theta^{t+1}) \leq \mathcal{R}^{\mathsf{total}}(\alpha^{t}, \Theta^{t+1}) \,.$$

Finally, combining the two updates yields the desired inequality.

In the case that learners and subpopulations are risk minimizing in the limit, the same argument holds with strict inequality, unless $(\alpha^{t+1}, \Theta^{t+1})$ is a local minimum. □

Theorem 5.4.3. *For any learners and subpopulations who are risk minimizing in the limit, an equilibrium $(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}})$ is asymptotically stable if it is an isolated local minimizer of the total risk $\mathcal{R}^{\mathsf{total}}$. If it is not a local minimizer of the total risk, then it is not stable.*

Proof of Theorem 5.4.3. We break this proof into two implications.

1. Isolated local min $\implies$ Asymptotic stability

Define $V(\alpha, \Theta) = \mathcal{R}^{\mathsf{total}}(\alpha, \Theta) - \mathcal{R}^{\mathsf{total}}(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}})$. The dynamics f are Lipschitz by Lemma 5.9.2 and this V satisfies the conditions of Theorem 5.9.3 with strict inequality, thus we conclude that $(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}})$ is an asymptotically stable equilibrium.

2. Not local min $\implies$ Not stable

Define $V(\alpha, \Theta) = \mathcal{R}^{\mathsf{total}}(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}}) - \mathcal{R}^{\mathsf{total}}(\alpha, \Theta)$ which will increase along trajectories. Since we are not at a local min, there must be some arbitrarily close α^0, Θ^0 such that $V(\alpha, \Theta) > 0$. Then we apply Theorem 5.9.4 which guarantees that the equilibrium is not stable. □

Corollary 5.4.4. *Equilibria exist when learners and subpopulations are risk minimizing in the limit and the total risk function has isolated local minima. They may not exist otherwise.*

Proof. We first argue that if the dynamics are risk minimizing, then an isolated local minimum of the total risk must be an equilibria. Let (α^0, Θ^0) denote the isolated local minima of the total risk. It must be that α^0 is an isolated, and thus unique, minimizer of $\mathcal{R}^{\mathsf{total}}(\alpha, \Theta^0)$ since the function is linear in α. We can thus conclude that $\nu(\alpha^0, \Theta^0) = \alpha^0$. It also must be that Θ^0 is a unique minimizer of $\mathcal{R}^{\mathsf{total}}(\alpha^0, \Theta)$ since the function is convex in Θ. We can thus conclude that $\mu(\alpha^0, \Theta^0) = \Theta^0$. Therefore (α^0, Θ^0) is equilibrium of the dynamics.

We next show that equilibria may not exist when the dynamics are not risk minimizing in the limit. To show that they may not exist otherwise, consider the following example. Let all learners be static and identical so $\Theta^{t+1} = \Theta^t$ and $\Theta = (\theta, \theta, \ldots, \theta)$. Let the subpopulation update break ties among equivalent learners randomly. Then the subpopulations will randomly switch between learners. Though these dynamics satisfy the definition of risk-reducing (at equality), they will not converge to an equilibrium.

We lastly show that equilibria may not exist when the total risk function does not have an isolated local minima. Suppose that learners update with full risk minimization and all subpopulations have risk uniquely minimized at the same value θ. Finally suppose that subpopulations will break ties among equivalent learners randomly (and are otherwise risk minimizing). As in the previous example, the subpopulations will randomly switch between learners and no equilibrium exists. □

Stability

Theorem 5.5.2. *Suppose learners and subpopulations are risk minimizing in the limit, α^{eq} is segmented, and $\mathcal{R}^{\mathsf{total}}(\alpha^{\mathsf{eq}}, \Theta)$ has a unique minimizer Θ^{eq}. Define a mapping $\gamma : [n] \to [m]$ such that $\gamma(i) = j$ is the learner with nonzero mass in $\alpha^{\mathsf{eq}}_{i,:}$.*

If every subpopulation strictly prefers their current learner:

$$\mathcal{R}_i(\theta^{\mathsf{eq}}_{\gamma(i)}) < \mathcal{R}_i(\theta^{\mathsf{eq}}_j)\,, \tag{5.1}$$

for all i and learners $j \neq \gamma(i)$, then $(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}})$ is an asymptotically stable equilibrium. If there is a subpopulation who would strictly prefer to switch learners, then $(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}})$ is not stable.

Proof. First note that it must be that every learner is associated to at least one subpopulation. Otherwise, the total risk would not have a unique minimizer over Θ.

We start with the first statement, and show that the stated conditions imply that $(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}})$ is isolated local minimum of the total risk. By Theorem 5.4.3, this implies asymptotic stability.

We specifically argue the conditions are sufficient for guaranteeing an isolated local minimum with respect to $F(\alpha)$, appealing to Lemma 5.9.8. First notice that we have the unique $\Theta^{\mathsf{eq}} = \arg\min_\Theta \mathcal{R}^{\mathsf{total}}(\alpha^{\mathsf{eq}}, \Theta)$ as required. Suppose by contradiction that there is some perturbation to α that causes $F(\alpha)$ to decrease or remain the same. Equivalently, the projection of the negative gradient onto the simplex points towards some other vertex, i.e. the component of the gradient in the direction of learner j is less than or equal to that in the direction of $\gamma(i)$ for some $j \neq \gamma(i)$. We can write this condition as

$$\frac{\partial F(\alpha)}{\partial \alpha_{i\gamma(i)}} \geq \frac{\partial F(\alpha)}{\partial \alpha_{ij}} \iff \mathcal{R}_i(\theta^{\mathsf{eq}}_{\gamma(i)}) \geq \mathcal{R}_i(\theta^{\mathsf{eq}}_j)$$

where we use Lemma 5.9.9. This violates the risk comparison condition (5.1), and therefore there must be no such perturbation, and thus α^{eq} is an isolated local minimum.

We turn our attention to the second statement. Theorem 5.4.3, it is equivalent to argue about minima of the total risk function. Suppose that for some subpopulation, there is some learner for which $\mathcal{R}_i(\theta^{\mathsf{eq}}_{\gamma(i)}) > \mathcal{R}_i(\theta^{\mathsf{eq}}_j)$. Then any small perturbation of that subpopulations's allocation towards that learner will decrease the total risk, and thus the point is not a minimum. □

In a segmented allocation, each θ^{eq}_j will minimize the average loss over the group of subpopulations assigned to them. Denote the parameter which minimizes risk of subpopulation i as $\phi_i := \arg\min_{\theta \in \mathbb{R}^d} \mathcal{R}_i(\theta)$. Then each θ^{eq}_j is a convex combination of ϕ_i for i in jth partition. Using this perspective, we provide an intuitive necessary (but not sufficient) condition for a class of symmetric risk functions.

Corollary 5.9.10. *Suppose that risk functions satisfy $\mathcal{R}_i(\theta) < \mathcal{R}_i(\theta') \iff \|\theta - \phi_i\| < \|\theta' - \phi_i\|$ for ϕ_i the subpopulation optimal parameter. Then in an asymptotically stable segmented equilibrium, the convex hulls of the grouped subpopulations optimal parameters $\{\phi_i\}$ are non-intersecting.*

Applying this Corollary to the example in Figure 5.4, we see that a segmented equilibrium with subpopulation 1 and 3 participating in the same learner cannot be stable.

Proof. Let $\phi_1, \phi_2, \cdots, \phi_k \in \mathbb{R}^d$ be the optimal decisions for the subpopulations allocated to the first learner and $\psi_1, \psi_2, \cdots, \psi_l \in \mathbb{R}^d$ be the optimal decisions for the subpopulations allocated to the second learner. Let θ_1 and θ_2 be the decisions of each learner. Assume that the convex hulls of $\{\phi_i\}_{i=1}^k$ and $\{\psi\}_{i=1}^l$ intersect. By Lemma 5.9.7, there exists i such that $\|\phi_i - \theta_2\| \leq \|\phi_i - \theta_1\|$. By the assumption about the risk runctions, this implies $\mathcal{R}_i(\theta_2) < \mathcal{R}_i(\theta_1)$. In other words, there exist a subpopulation that would prefer to switch learners. Thus by Theorem 5.5.2 these allocation of subpopulations to learner is not stable and so the convex hulls must not intersect. □

Theorem 5.5.4. *Consider dynamics which are risk minimizing in the limit and an α^{eq} with any subpopulation i having nonzero support on set of two or more learners $j \in \mathcal{J}$. Assume risks are strongly convex and define $\Theta^{\mathsf{eq}} = \arg\min \mathcal{R}^{\mathsf{total}}(\alpha^{\mathsf{eq}}, \Theta)$. Then $(\alpha^{\mathsf{eq}}, \Theta^{\mathsf{eq}})$ cannot be stable unless it is "balanced" in the sense that learners in $\mathcal{J}$ are* risk equivalent *and* optimal *for i, i.e. for all $j, j' \in \mathcal{J}$,*

$$\mathcal{R}_i(\theta_j^{\mathsf{eq}}) = \mathcal{R}_i(\theta_{j'}^{\mathsf{eq}}) \quad \textit{and} \quad \nabla \mathcal{R}_i(\theta_j^{\mathsf{eq}}) = 0\,. \tag{5.2}$$

If it is balanced, so are all allocations for subpopulation i with support over $\mathcal{J}$.

Finally, all stable equilibria must be either balanced or segmented.

Proof. Theorem 5.4.3 shows that an equilibrium cannot be stable if it is not a local minimum of the total risk. We therefore develop conditions under which an equilibrium point will be a local minimum. By Lemma 5.9.8, it is equivalent to argue about the local minima of the concave function $F(\alpha)$ over the simplex product Δ_m^n. All minima of the total risk will occur on the boundary of the simplex product, i.e. a face or a vertex. Since F is still concave when restricted to a face of the simplex, the same argument shows the minima are on the boundary, hence vertices, except for the degenerate case where F takes a constant value over the face.

We now characterize this degenerate case. F takes a constant value over the face if and only if 1) the gradient of F is perpendicular to the face at α and 2) remains perpendicular along the face. The face is described by a set of indices $\mathcal{J} \subseteq [m]$. Mathematically, we write the two conditions as: for all pairs $j, j' \in \mathcal{J}$, $\ell \in [n]$, and $k \in [m]$,

$$\frac{\partial F(\alpha)}{\partial \alpha_{ij}} = \frac{\partial F(\alpha)}{\partial \alpha_{ij'}} \quad \text{and} \quad \frac{\partial}{\partial \alpha_{\ell k}} \left(\frac{\partial F(\alpha)}{\partial \alpha_{ij}} - \frac{\partial F(\alpha)}{\partial \alpha_{ij'}} \right) = 0 \tag{5.5}$$

Using Lemma 5.9.8, the first expression simplifies to the *risk equivalent* condition that $\mathcal{R}_i(\theta_j^{\mathsf{eq}}) = \mathcal{R}_i(\theta_{j'}^{\mathsf{eq}})$. Turning to the second expression in (5.5), we first use Lemma 5.9.9 to compute

$$\frac{\partial}{\partial \alpha_{\ell k}} \frac{\partial F(\alpha)}{\partial \alpha_{ij}} = \begin{cases} 0 & k \neq j \\ -\beta_i \nabla \mathcal{R}_i(\theta_j^{\mathsf{eq}})^\top \left(\sum_{\ell'} \beta_{\ell'} \alpha_{\ell' j} \nabla^2 \mathcal{R}_{\ell'}(\theta_j^{\mathsf{eq}}) \right) \nabla \mathcal{R}_\ell(\theta_j^{\mathsf{eq}}) & k = j \end{cases}$$

Thus, the condition trivially holds for $k \notin \{j, j'\}$. Otherwise, when $\ell = i$, the condition in (5.5) requires that

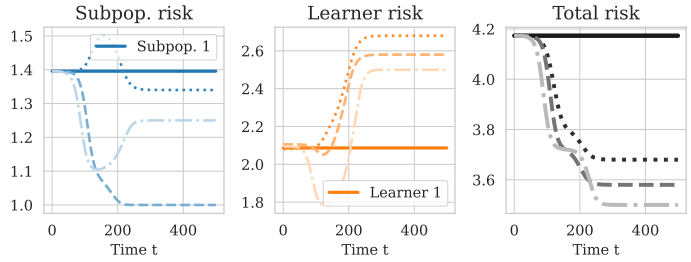
$$\nabla \mathcal{R}_i(\theta_k^{\mathsf{eq}})^\top \left(\sum_{\ell'} \beta_{\ell'} \alpha_{\ell' k} \nabla^2 \mathcal{R}_{\ell'}(\theta_k^{\mathsf{eq}}) \right) \nabla \mathcal{R}_i(\theta_k^{\mathsf{eq}}) = 0, \quad k \in \{j, j'\}$$

Due to the strong convexity of the risks, the Hessian matrix is positive definite. Thus it must be that $\nabla \mathcal{R}_i(\theta_j^{\mathsf{eq}}) = 0$ for all $j \in \mathcal{J}$, i.e. the *risk optimal* condition. Risk optimality implies that the condition holds also when $\ell \neq i$ and thus the characterization is complete. □

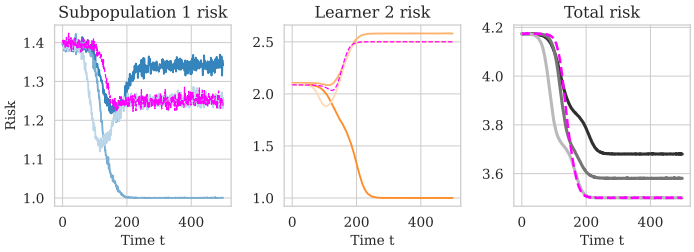
Additional Experiments with Noisy Dynamics

Figure 5.8a replicates Figure 5.6a. The magenta-highlighted trajectory starts precisely at the unstable equilibrium, while the other three, initiated near this point, converge to the three possible split market equilibria, ordered by hue intensity: {(1,2), (3)}, {(2,3), (1)}, and {(1,3), (2)}. In Figure 5.8b, while sub-population dynamics remain as in (a), learner updates experience uncorrelated external perturbations, causing trajectories to be different from (a). Nevertheless, the long term dynamics gravitate near stable split equilibria. Figure 5.8c depicts learners updating decisions based on sampled empirical losses, with sub-populations adjusting participation based on aggregate empirical performance. The fact that each learner uses different samples from each sub-population adds sufficient un-correlated noise to create trajectories similar to when exogenous noise is added.

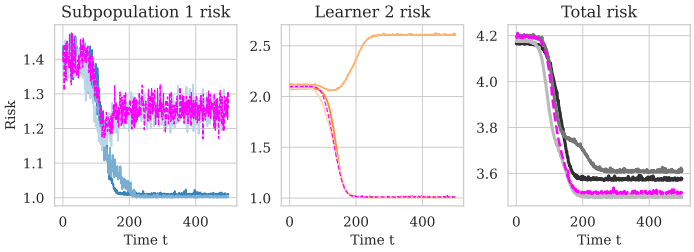
Experimental Details Full experimental details along with instructions for reproducing them can be found at `https://github.com/mcurmei627/MultiLearnerRiskReduction`. The experiments used Python 3.10 on a MacBook Pro 2019.



(a) Learner updates: **noiseless** one-step minimization of **population** loss. Sub-population updates: MWU w.r.t **population** loss



(b) Learner updates: **noisy** one-step minimization of **population** loss. Sub-population updates: MWU w.r.t **population** loss



(c) Learner updates: **noiseless** one-step minimization of **empirical** loss. Sub-population updates: MWU w.r.t **empirical** loss

Figure 5.8: Noisy dynamics

Bibliography

[AA03] Lada A Adamic and Eytan Adar. "Friends and neighbors on the web". In: *Social networks* 25.3 (2003), pp. 211–230.

[AB17] Luca Maria Aiello and Nicola Barbieri. "Evolution of Ego-networks in Social Media with Link Recommendations". In: *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*. arXiv:1702.01398 [physics]. Feb. 2017, pp. 111–120. DOI: 10.1145/3018661.3018733. URL: http://arxiv.org/abs/1702.01398 (visited on 09/27/2022).

[Abd+19] Himan Abdollahpouri et al. "The impact of popularity bias on fairness and calibration in recommendation". In: *arXiv preprint arXiv:1910.05755* (2019).

[Abe+22] Rediet Abebe et al. "On the Effect of Triadic Closure on Network Segregation". In: *Proceedings of the 23rd ACM Conference on Economics and Computation*. arXiv:2205.13658 [cs, econ]. July 2022, pp. 249–284. DOI: 10.1145/3490486.3538322. URL: http://arxiv.org/abs/2205.13658 (visited on 09/30/2022).

[ACF22] Vincenzo Auletta, Antonio Coppola, and Diodato Ferraioli. *On the Impact of Social Media Recommendations on Opinion Consensus*. arXiv:2204.03299 [cs]. Apr. 2022. URL: http://arxiv.org/abs/2204.03299 (visited on 09/27/2022).

[Ado+13] Gediminas Adomavicius et al. "Do recommender systems manipulate consumer preferences? A study of anchoring effects". In: *Information Systems Research* 24.4 (2013), pp. 956–975.

[Akp+22] Nil-Jana Akpinar et al. "Long-term Dynamics of Fairness Intervention in Connection Recommender Systems". In: *arXiv preprint arXiv:2203.16432* (2022).

[Alo+09] Daniel Aloise et al. "NP-hardness of Euclidean sum-of-squares clustering". In: *Machine learning* 75.2 (2009), pp. 245–248.

[ApS19] MOSEK ApS. *MOSEK Optimizer API for Python Release 9.0.88*. 2019. URL: https://docs.mosek.com/9.0/pythonapi.pdf.

[Ari+20] Guy Aridor et al. "Competing bandits: The perils of exploration under competition". In: *arXiv preprint arXiv:2007.10144* (2020).

[Asi+20] Aili Asikainen et al. "Cumulative effects of triadic closure and homophily in social networks". In: *Science Advances* 6.19 (2020), eaax7310.

[BCS18] Nicoletta Bof, Ruggero Carli, and Luca Schenato. "Lyapunov theory for discrete time systems". In: *arXiv preprint arXiv:1809.05289* (2018).

[Bee+13] Joeran Beel et al. "A comparative analysis of offline and online evaluations and discussion of research paper recommender system evaluation". In: *Proceedings of the international workshop on reproducibility and replication in recommender systems evaluation*. 2013, pp. 7–14.

[Ber+21] B Douglas Bernheim et al. "A theory of chosen preferences". In: *American Economic Review* 111.2 (2021), pp. 720–54.

[BHK22] Gavin Brown, Shlomi Hod, and Iden Kalemaj. "Performative prediction in a stateful world". In: *International Conference on Artificial Intelligence and Statistics*. PMLR. 2022, pp. 6045–6061.

[Bor89] Robert F Bornstein. "Exposure and affect: overview and meta-analysis of research, 1968–1987." In: *Psychological bulletin* 106.2 (1989), p. 265.

[Bos+23] Avinandan Bose et al. "Initializing Services in Interactive ML Systems for Diverse Users". In: *arXiv preprint arXiv:2312.11846* (2023).

[Bri71] Philip Brickman. "Hedonic relativism and planning the good society". In: *Adaptation level theory* (1971), pp. 287–301.

[BT19] Omer Ben-Porat and Moshe Tennenholtz. "Regression equilibrium". In: *Proceedings of the 2019 ACM Conference on Economics and Computation*. 2019, pp. 173–191.

[Car+21] Micah Carroll et al. "Estimating and Penalizing Preference Shift in Recommender Systems". In: *Fifteenth ACM Conference on Recommender Systems*. 2021, pp. 661–667.

[Car+23] Micah Carroll et al. "Characterizing manipulation from AI systems". In: *Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*. 2023, pp. 1–13.

[CC02] Dena Cox and Anthony D Cox. "Beyond first impressions: The effects of repeated exposure on consumer liking of visually complex and simple product designs". In: *Journal of the Academy of Marketing Science* 30.2 (2002), pp. 119–130.

[CDH21] Joshua Cutler, Dmitriy Drusvyatskiy, and Zaid Harchaoui. "Stochastic optimization under time drift: iterate averaging, step-decay schedules, and high probability guarantees". In: *Advances in Neural Information Processing Systems* 34 (2021).

[CDR21] Mihaela Curmei, Sarah Dean, and Benjamin Recht. "Quantifying Availability and Discovery in Recommender Systems via Stochastic Reachability". In: *Proceedings of the 38th International Conference on Machine Learning*. Ed. by Marina Meila and Tong Zhang. Vol. 139. Proceedings of Machine Learning Research. PMLR, 2021, pp. 2265–2275. URL: https://proceedings.mlr.press/v139/curmei21a.html.

[Cel10] Oscar Celma. "Music recommendation". In: *Music recommendation and discovery*. Springer, 2010, pp. 43–85.

[Ces+17] Nicolò Cesa-Bianchi et al. "Boltzmann exploration done right". In: *arXiv preprint arXiv:1705.10257* (2017).

[Cha+15] Laurent Charlin et al. "Dynamic poisson factorization". In: *Proceedings of the 9th ACM Conference on Recommender Systems*. 2015, pp. 155–162.

[Che+19] Minmin Chen et al. "Top-k off-policy correction for a REINFORCE recommender system". In: *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*. 2019, pp. 456–464.

[CHH+07] John O Cooper, Timothy E Heron, William L Heward, et al. "Applied behavior analysis". In: (2007).

[Chr+15] Fabian Christoffel et al. "Blockbusters and wallflowers: Accurate, diverse, and scalable recommendations with random walks". In: *Proceedings of the 9th ACM Conference on Recommender Systems*. 2015, pp. 163–170.

[CI20] Sunaina Chugani and Julie R Irwin. "All eyes on you: The social audience and hedonic adaptation". In: *Psychology & Marketing* 37.11 (2020), pp. 1554–1570.

[Cin+22] Federico Cinus et al. "The effect of people recommenders on echo chambers and polarization". In: *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 16. 2022, pp. 90–101.

[CIR15] Sunaina K Chugani, Julie R Irwin, and Joseph P Redden. "Happily ever after: The effect of identity-consistency on product satiation". In: *Journal of Consumer Research* 42.4 (2015), pp. 564–577.

[CSE18] Allison JB Chaney, Brandon M Stewart, and Barbara E Engelhardt. "How algorithmic confounding in recommendation systems increases homogeneity and decreases utility". In: *Proceedings of the 12th ACM Conference on Recommender Systems*. 2018, pp. 224–232.

[Cur+22] Mihaela Curmei et al. "Towards Psychologically-Grounded Dynamic Preference Models". In: *Proceedings of the 16th ACM Conference on Recommender Systems*. 2022, pp. 35–48.

[CVW11] Pablo Castells, Saúl Vargas, and Jun Wang. "Novelty and diversity metrics for recommender systems: choice, discovery and relevance". In: (2011).

[Dac+21] Maurizio Ferrari Dacrema et al. "A troubling analysis of reproducibility and progress in recommender systems research". In: *ACM Transactions on Information Systems (TOIS)* 39.2 (2021), pp. 1–49.

[Dec+09] Alice Dechêne et al. "Mix me a list: Context moderates the truth effect and the mere-exposure effect". In: *Journal of Experimental Social Psychology* 45.5 (2009), pp. 1117–1122.

[DGL13] Pranav Dandekar, Ashish Goel, and David T. Lee. "Biased assimilation, homophily, and the dynamics of polarization". en. In: *Proceedings of the National Academy of Sciences* 110.15 (Apr. 2013). Publisher: National Academy of Sciences Section: Social Sciences, pp. 5791–5796. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1217220110. URL: https://www.pnas.org/content/110/15/5791 (visited on 02/16/2022).

[DGM10] Elizabeth M Daly, Werner Geyer, and David R Millen. "The network effects of recommending social connections". In: *Proceedings of the fourth ACM conference on Recommender systems*. 2010, pp. 301–304.

[Din+21] Frances Ding et al. "Retiring adult: New datasets for fair machine learning". In: *Advances in Neural Information Processing Systems* 34 (2021), pp. 6478–6490.

[DK11] Christian Desrosiers and George Karypis. "A comprehensive survey of neighborhood based recommendation methods". In: *Recommender systems handbook* (2011), pp. 107–144.

[DRR20] Sarah Dean, Sarah Rich, and Benjamin Recht. "Recommendations and user agency: the reachability of collaboratively-filtered information". In: *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. 2020, pp. 436–445.

[DSV19] Gianlorenzo D'Angelo, Lorenzo Severini, and Yllka Velaj. "Recommending links through influence maximization". en. In: *Theoretical Computer Science*. Selected papers of ICTCS 2016 (The Italian Conference on Theoretical Computer Science (ICTCS) 764 (Apr. 2019), pp. 30–41. ISSN: 0304-3975. DOI: 10.1016/j.tcs.2018.01.017. URL: https://www.sciencedirect.com/science/article/pii/S0304397518300641 (visited on 03/16/2022).

[DX20] Dmitriy Drusvyatskiy and Lin Xiao. "Stochastic optimization with decision-dependent distributions". In: *arXiv preprint arXiv:2011.11173* (2020).

[DZ21] Tim Donkers and Jürgen Ziegler. "The Dual Echo Chamber: Modeling Social Media Polarization for Interventional Recommending". In: *Fifteenth ACM Conference on Recommender Systems*. 2021, pp. 12–22.

[Eks+18a] Michael D Ekstrand et al. "All the cool kids, how do they fit in?: Popularity and demographic biases in recommender evaluation and effectiveness". In: *Conference on Fairness, Accountability and Transparency*. PMLR. 2018, pp. 172–186.

[Eks+18b] Michael D Ekstrand et al. "Exploring author gender in book rating and recommendation". In: *Proceedings of the 12th ACM conference on recommender systems*. 2018, pp. 242–250.

[Esp+22] Lisette Espín-Noboa et al. "Inequality and inequity in network-based ranking and recommendation algorithms". In: *Scientific reports* 12.1 (2022), pp. 1–14.

[Fab+20] Francesco Fabbri et al. "The Effect of Homophily on Disparate Visibility of Minorities in People Recommender Systems". en. In: *Proceedings of the International AAAI Conference on Web and Social Media* 14 (May 2020), pp. 165–175. ISSN: 2334-0770. URL: https://ojs.aaai.org/index.php/ICWSM/article/view/7288 (visited on 03/16/2022).

[Fab+21] Francesco Fabbri et al. "Exposure Inequality in People Recommender Systems: The Long-Term Effects". In: *arXiv:2112.08237 [cs]* (Dec. 2021). arXiv: 2112.08237. URL: http://arxiv.org/abs/2112.08237 (visited on 02/17/2022).

[Fab+22] Francesco Fabbri et al. "Exposure Inequality in People Recommender Systems: The Long-Term Effects". In: *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 16. 2022, pp. 194–204.

[FAF11] Asle Fagerstrøm, Erik Arntzen, and Gordon R Foxall. "A study of preferences in a simulated online shopping experiment". In: *The Service Industries Journal* 31.15 (2011), pp. 2603–2615.

[FCF20] Marc Faddoul, Guillaume Chaslot, and Hany Farid. "A Longitudinal Analysis of YouTube's Promotion of Conspiracy Videos". In: *arXiv preprint arXiv:2003.03318* (2020).

[Fer+22] Antonio Ferrara et al. "Link recommendations: Their impact on network structure and minorities". In: *arXiv preprint arXiv:2205.06048* (2022).

[FGR16] Seth Flaxman, Sharad Goel, and Justin M Rao. "Filter bubbles, echo chambers, and online news consumption". In: *Public opinion quarterly* 80.S1 (2016), pp. 298–320.

[Fox04] Gordon Foxall. *Consumer psychology in behavioral perspective*. Beard Books, 2004.

[Fox10] Gordon R Foxall. "Invitation to consumer behavior analysis". In: *Journal of Organizational Behavior Management* 30.2 (2010), pp. 92–109.

[Fox17] Gordon R Foxall. *Behavioral economics in consumer behavior analysis*. 2017.

[FS57] Charles B Ferster and Burrhus Frederic Skinner. "Schedules of reinforcement." In: (1957).

[FSA07] Xiang Fang, Surendra Singh, and Rohini Ahluwalia. "An examination of different explanations for the mere exposure effect". In: *Journal of consumer research* 34.1 (2007), pp. 97–103.

[Ful49] Paul R Fuller. "Operant conditioning of a vegetative human organism". In: *The American journal of psychology* 62.4 (1949), pp. 587–590.

[FV92] Dean P Foster and Rakesh V Vohra. "An economic argument for affirmative action". In: *Rationality and Society* 4.2 (1992), pp. 176–188.

[Gar+14] Florent Garcin et al. "Offline and online evaluation of news recommender systems at swissinfo. ch". In: *Proceedings of the 8th ACM Conference on Recommender systems*. 2014, pp. 169–176.

[Gin+21] Tony Ginart et al. "Competing AI: How does competition feedback affect machine learning?" In: *International Conference on Artificial Intelligence and Statistics*. PMLR. 2021, pp. 1693–1701.

[GLH19] Daniel Geschke, Jan Lorenz, and Peter Holtz. "The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers". In: *British Journal of Social Psychology* 58.1 (2019), pp. 129–149.

[Gui+15] Huan Gui et al. "Network a/b testing: From sampling to estimation". In: *Proceedings of the 24th International Conference on World Wide Web*. 2015, pp. 399–409.

[Har+15] F Maxwell Harper et al. "Putting users in control of their recommendations". In: *Proceedings of the 9th ACM Conference on Recommender Systems*. 2015, pp. 3–10.

[Har+16] Moritz Hardt et al. "Strategic classification". In: *Proceedings of the 2016 ACM conference on innovations in theoretical computer science*. 2016, pp. 111–122.

[Has+18] Tatsunori Hashimoto et al. "Fairness without demographics in repeated loss minimization". In: *International Conference on Machine Learning*. PMLR. 2018, pp. 1929–1938.

[Haz+16] Elad Hazan et al. "Introduction to online convex optimization". In: *Foundations and Trends® in Optimization* 2.3-4 (2016), pp. 157–325.

[HBT18] Lars Hellemo, Paul I Barton, and Asgeir Tomasgard. "Decision-dependent probabilities in stochastic programs with recourse". In: *Computational Management Science* 15.3 (2018), pp. 369–395.

[Hel47] Harry Helson. "Adaptation-level as frame of reference for prediction of psychophysical data". In: *The American journal of psychology* 60.1 (1947), pp. 1–29.

[Her+04] Jonathan Herlocker et al. "Evaluating collaborative filtering recommender systems". In: *ACM transactions on information systems* 22.1 (2004), pp. 5–53.

[HJM22] Moritz Hardt, Meena Jagadeesan, and Celestine Mendler-Dünner. "Performative Power". In: *arXiv preprint arXiv:2203.17232* (2022).

[HK15] F Maxwell Harper and Joseph A Konstan. "The movielens datasets: History and context". In: *Acm transactions on interactive intelligent systems (tiis)* 5.4 (2015), pp. 1–19.

[Hos+18] Seyed Abbas Hosseini et al. "Recurrent poisson factorization for temporal recommendation". In: *IEEE Transactions on Knowledge and Data Engineering* 32.1 (2018), pp. 121–134.

[Hos+20] Homa Hosseinmardi et al. "Evaluating the scale, growth, and origins of right-wing echo chambers on YouTube". In: *arXiv preprint arXiv:2011.12843* (2020).

[Hot29] Harold Hotelling. "Stability in competition". In: *The Economic Journal* (1929).

[HTW13] Paul Hekkert, Clementine Thurgood, and TW Allan Whitfield. "The mere exposure effect for consumer products as a consequence of existing familiarity and controlled exposure". In: *Acta psychologica* 144.2 (2013), pp. 411–417.

[Ie+19] Eugene Ie et al. "SlateQ: A tractable decomposition for reinforcement learning with recommendation sets". In: (2019).

[IKI94] Mary Inaba, Naoki Katoh, and Hiroshi Imai. "Applications of weighted Voronoi diagrams and randomization to variance-based k-clustering". In: *Proceedings of the tenth annual symposium on Computational geometry*. 1994, pp. 332–339.

[IR15] Guido W. Imbens and Donald B. Rubin. *Causal Inference in Statistics, Social, and Biomedical Sciences*. en. Google-Books-ID: Bf1tBwAAQBAJ. Cambridge University Press, Apr. 2015. ISBN: 978-0-521-88588-1.

[IYZ21] Zachary Izzo, Lexing Ying, and James Zou. "How to learn when data reacts to your model: performative gradient descent". In: *International Conference on Machine Learning*. PMLR. 2021, pp. 4641–4650.

[Jan+15] Dietmar Jannach et al. "What recommenders recommend: an analysis of recommendation biases and possible countermeasures". In: *User Modeling and User-Adapted Interaction* 25.5 (2015), pp. 427–491.

[Jay57] Edwin T Jaynes. "Information theory and statistical mechanics". In: *Physical review* 106.4 (1957), p. 620.

[Jeu19] Olivier Jeunen. "Revisiting offline evaluation for implicit-feedback recommender systems". In: *Proceedings of the 13th ACM Conference on Recommender Systems*. 2019, pp. 596–600.

[Jia+19] Ray Jiang et al. "Degenerate feedback loops in recommender systems". In: *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. 2019, pp. 383–390.

[Jia+21] Bohan Jiang et al. "Mechanisms and Attributes of Echo Chambers in Social Media". In: *arXiv preprint arXiv:2106.05401* (2021).

[JJ21] Mathias Jesse and Dietmar Jannach. "Digital nudging with recommender systems: Survey and future directions". In: *Computers in Human Behavior Reports* 3 (2021), p. 100052.

[JR07] Matthew O Jackson and Brian W Rogers. "Meeting strangers and friends of friends: How random are social networks?" In: *American Economic Review* 97.3 (2007), pp. 890–915.

[Kah+99] Daniel Kahneman et al. "Objective happiness". In: *Well-being: The foundations of hedonic psychology* 3.25 (1999), pp. 1–23.

[Kal+21] Dimitris Kalimeris et al. "Preference Amplification in Recommender Systems". In: *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. 2021, pp. 805–815.

[KAM19] Matthäus Kleindessner, Pranjal Awasthi, and Jamie Morgenstern. "Fair k-center clustering for data summarization". In: *International Conference on Machine Learning*. PMLR. 2019, pp. 3448–3457.

[Kar+18] Fariba Karimi et al. "Homophily influences ranking of minorities in social networks". In: *Scientific reports* 8.1 (2018), pp. 1–12.

[Kar+20] Amir-Hossein Karimi et al. "A survey of algorithmic recourse: definitions, formulations, solutions, and prospects". In: *arXiv preprint arXiv:2010.04050* (2020).

[Kaw+15] Jaya Kawale et al. "Efficient thompson sampling for online matrix-factorization recommendation". In: *Advances in neural information processing systems*. 2015, pp. 1297–1305.

[KB15] Yehuda Koren and Robert Bell. "Advances in collaborative filtering". In: *Recommender systems handbook* (2015), pp. 77–118.

[Kih87] John F Kihlstrom. "The cognitive unconscious". In: *Science* 237.4821 (1987), pp. 1445–1452.

[Kle+19] Matthäus Kleindessner et al. "Guarantees for spectral clustering with fairness constraints". In: *International Conference on Machine Learning*. PMLR. 2019, pp. 3458–3467.

[Kor08] Yehuda Koren. "Factorization meets the neighborhood: a multifaceted collaborative filtering model". In: *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2008, pp. 426–434.

[Kor09] Yehuda Koren. "Collaborative filtering with temporal dynamics". In: *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2009, pp. 447–456.

[Kra+20] Karl Krauth et al. "Do Offline Metrics Predict Online Performance in Recommender Systems?" In: *arXiv preprint arXiv:2011.07931* (2020).

[Kus67] Harold J Kushner. *Stochastic stability and control*. Tech. rep. Brown Univ Providence RI, 1967.

[Lev21] Ro'ee Levy. "Social media, news consumption, and polarization: Evidence from a field experiment". In: *American Economic Review* 111.3 (2021), pp. 831–70.

[Lex+21] Elisabeth Lex et al. *Psychology-informed recommender systems*. Now Publishers, 2021.

[Liu+12] Ji Liu et al. "Tensor completion for estimating missing values in visual data". In: *IEEE transactions on pattern analysis and machine intelligence* 35.1 (2012), pp. 208–220.

[Liu+18] Lydia T Liu et al. "Delayed impact of fair machine learning". In: *International Conference on Machine Learning*. PMLR. 2018, pp. 3150–3158.

[Luk+21] Kai Lukoff et al. "How the Design of YouTube Influences User Sense of Agency". In: *arXiv preprint arXiv:2101.11778* (2021).

[LY22] Jun Liu and Ye Yuan. "On almost sure convergence rates of stochastic gradient methods". In: *Conference on Learning Theory*. PMLR. 2022, pp. 2963–2983.

[LZ19] Mark Ledwich and Anna Zaitsev. "Algorithmic extremism: Examining YouTube's rabbit hole of radicalization". In: *arXiv preprint arXiv:1912.11211* (2019).

[Mar09] Stephen Marsland. "Using habituation in machine learning". In: *Neurobiology of learning and memory* 92.2 (2009), pp. 260–266.

[Men+20] Celestine Mendler-Dünner et al. "Stochastic optimization for performative prediction". In: *Advances in Neural Information Processing Systems* 33 (2020), pp. 4929–4939.

[MG21] Simon Markmann and Christian Grimme. "Is YouTube Still a Radicalizer? An Exploratory Study on Autoplay and Recommendation". In: *Multidisciplinary International Symposium on Disinformation in Open Online Media*. Springer. 2021, pp. 50–65.

[MGP15] Jérémie Mary, Romaric Gaudel, and Philippe Preux. "Bandits and recommender systems". In: *International Workshop on Machine Learning, Optimization and Big Data*. Springer. 2015, pp. 325–336.

[Mil+19] Smitha Milli et al. "The social cost of strategic classification". In: *Proceedings of the Conference on Fairness, Accountability, and Transparency*. 2019, pp. 230–239.

[MPZ21] John P Miller, Juan C Perdomo, and Tijana Zrnic. "Outside the echo chamber: Optimizing the performative risk". In: *International Conference on Machine Learning*. PMLR. 2021, pp. 7710–7720.

[Mun19] Luke Munn. "Alt-right pipeline: Individual journeys to extremism online". In: *First Monday* (2019).

[Nar+22] Adhyyan Narang et al. "Multiplayer Performative Prediction: Learning in Decision-Dependent Games". In: *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics (arXiv:2201.03398)* (2022).

[Ngu+14] Tien T Nguyen et al. "Exploring the filter bubble: the effect of using recommender systems on content diversity". In: *Proceedings of the 23rd international conference on World wide web*. 2014, pp. 677–686.

[NK11] Xia Ning and George Karypis. "Slim: Sparse linear methods for top-n recommender systems". In: *2011 IEEE 11th International Conference on Data Mining*. IEEE. 2011, pp. 497–506.

[NM08] Leif D Nelson and Tom Meyvis. "Interrupted consumption: Disrupting adaptation to hedonic experiences". In: *Journal of Marketing Research* 45.6 (2008), pp. 654–664.

[Noo+20] Emil Noordeh et al. "Echo Chambers in Collaborative Filtering Based Recommendation Systems". In: *arXiv preprint arXiv:2011.03890* (2020).

[Nor02] Christie L Nordhielm. "The influence of level of processing on advertising repetition effects". In: *Journal of consumer research* 29.3 (2002), pp. 371–382.

[OD98] Susan Hesse Owen and Mark S Daskin. "Strategic facility location: A review". In: *European journal of operational research* 111.3 (1998), pp. 423–447.

[Ora20] Francesco Orabona. *Almost Sure Convergence of SGD on Smooth Non-Convex Functions on https://parameterfree.com/*. 2020.

[Par11] Eli Pariser. *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin, 2011.

[Pas+21] Francesco Sanna Passino et al. "Where To Next? A Dynamic Model of User Preferences". In: *The 2021 World Wide Web Conference, WWW*. 2021, pp. 19–23.

[Per+20] Juan Perdomo et al. "Performative prediction". In: *International Conference on Machine Learning*. PMLR. 2020, pp. 7599–7609.

[PY22] Georgios Piliouras and Fang-Yi Yu. "Multi-agent Performative Prediction: From Global Stability and Optimality to Chaos". In: *arXiv preprint arXiv:2201.10483* (2022).

[Qui+08] Joaquin Quiñonero-Candela et al., eds. *Dataset shift in machine learning*. Mit Press, 2008.

[Raj+22] Karthik Rajkumar et al. "A causal test of the strength of weak ties". In: *Science* 377.6612 (Sept. 2022). Publisher: American Association for the Advancement of Science, pp. 1304–1310. DOI: 10.1126/science.abl4476. URL: https://www.science.org/doi/10.1126/science.abl4476 (visited on 09/30/2022).

[Ray+22] Mitas Ray et al. "Decision-Dependent Risk Minimization in Geometrically Decaying Dynamic Environments". In: *Proceedings of the Association for the Advancement of Artificial Intelligence Conference on AI (AAAI)*. 2022.

[Ren12] Steffen Rendle. "Factorization Machines with libFM". In: *ACM Trans. Intell. Syst. Technol.* 3.3 (May 2012), 57:1–57:22. ISSN: 2157-6904.

[RF17] Diego Ruiz and Jorge Finke. "Stability of the Jackson-Rogers model". In: 2017, pp. 1803–1808. DOI: 10.1109/CDC.2017.8263909.

[Rib+20] Manoel Horta Ribeiro et al. "Auditing radicalization pathways on YouTube". In: *Proceedings of the 2020 conference on fairness, accountability, and transparency*. 2020, pp. 131–141.

[RPF18] Wilbert Samuel Rossi, Jan Willem Polderman, and Paolo Frasca. "The closed loop between opinion formation and personalised recommendations". In: *arXiv preprint arXiv:1809.04644* (2018).

[RSZ16] Marco Rossetti, Fabio Stella, and Markus Zanker. "Contrasting offline and online results when evaluating recommendation algorithms". In: *Proceedings of the 10th ACM conference on recommender systems*. 2016, pp. 31–34.

[RZK19] Steffen Rendle, Li Zhang, and Yehuda Koren. "On the difficulty of evaluating baselines: A study on recommender systems". In: *arXiv preprint arXiv:1905.01395* (2019).

[San20] William H Sandholm. "Evolutionary game theory". In: *Complex Social and Behavioral Systems: Game Theory and Agent-Based Models* (2020), pp. 573–608.

[Sch+16] Tobias Schnabel et al. "Recommendations as treatments: Debiasing learning and evaluation". In: *international conference on machine learning*. PMLR. 2016, pp. 1670–1679.

[Sha+20] Dougal Shakespeare et al. "Exploring Artist Gender Bias in Music Recommendation". In: *arXiv preprint arXiv:2009.01715* (2020).

[SI84] Shokri Z Selim and Mohamed A Ismail. "K-means-type algorithms: A generalized convergence theorem and characterization of local optimality". In: *IEEE Transactions on pattern analysis and machine intelligence* 1 (1984), pp. 81–87.

[SJ18] Ashudeep Singh and Thorsten Joachims. "Fairness of exposure in rankings". In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2018, pp. 2219–2228.

[Ski38] B. F. Skinner. *The Behavior of Organisms: An Experimental Analysis*. Appleton-Century, 1938.

[SLL21] Fernando P. Santos, Yphtach Lelkes, and Simon A. Levin. "Link recommendation algorithms and dynamics of polarization in online social networks". In: *Proceedings of the National Academy of Sciences* 118.50 (Dec. 2021). Publisher: Proceedings of the National Academy of Sciences, e2102141118. DOI: 10.1073/pnas.2102141118. URL: https://www.pnas.org/doi/10.1073/pnas.2102141118 (visited on 03/08/2022).

[SRC18] Ana-Andreea Stoica, Christopher Riederer, and Augustin Chaintreau. "Algorithmic Glass Ceiling in Social Networks: The effects of social recommendations on network diversity". en. In: *Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18*. Lyon, France: ACM Press, 2018, pp. 923–932. ISBN: 978-1-4503-5639-8. DOI: 10.1145/3178876.3186140. URL: http://dl.acm.org/citation.cfm?doid=3178876.3186140 (visited on 02/23/2022).

[SSG16] Jessica Su, Aneesh Sharma, and Sharad Goel. "The effect of recommendations on network structure". In: *Proceedings of the 25th international conference on World Wide Web*. 2016, pp. 1157–1167.

[Ste18] Harald Steck. "Calibrated recommendations". In: *Proceedings of the 12th ACM conference on recommender systems*. 2018, pp. 154–162.

[Ste19] Harald Steck. "Embarrassingly shallow autoencoders for sparse data". In: *The World Wide Web Conference*. 2019, pp. 3251–3257.

[Sun+20] Zhu Sun et al. "Are we evaluating rigorously? benchmarking recommendation for reproducible evaluation and fair comparison". In: *Fourteenth ACM conference on recommender systems*. 2020, pp. 23–32.

[Tea22] LinkedIn Engineering Team. *People You May Know*. [Online; accessed 31-Jan-2023]. 2022. URL: https://engineering.linkedin.com/teams/data/artificial-intelligence/people-you-may-know.

[Tom+21] Matus Tomlein et al. "An Audit of Misinformation Filter Bubbles on YouTube: Bubble Bursting and Recent Behavior Changes". In: *Fifteenth ACM Conference on Recommender Systems*. 2021, pp. 1–11.

[Tor19] Alexander Torgovitsky. "Nonparametric inference on state dependence in unemployment". In: *Econometrica* 87.5 (2019), pp. 1475–1505.

[USL19] Berk Ustun, Alexander Spangher, and Yang Liu. "Actionable recourse in linear classification". In: *Proceedings of the Conference on Fairness, Accountability, and Transparency*. 2019, pp. 10–19.

[WBD21] Killian Wood, Gianluca Bianchin, and Emiliano Dall'Anese. "Online Projected Gradient Descent for Stochastic Optimization with Decision-Dependent Distributions". In: *IEEE Control Systems Letters* (2021).

[WD22] Killian Wood and Emiliano Dall'Anese. "Stochastic Saddle Point Problems with Decision-Dependent Distributions". In: *arXiv preprint arXiv:2201.02313* (2022).

[Wei+17] Zeng Wei et al. "Reinforcement learning to rank with Markov decision process". In: *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2017, pp. 945–948.

[Wu+20] Fangzhao Wu et al. "Mind: A large-scale dataset for news recommendation". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 2020, pp. 3597–3606.

[Yan+18] Longqi Yang et al. "Unbiased offline recommender evaluation for missing-not-at-random implicit feedback". In: *Proceedings of the 12th ACM Conference on Recommender Systems*. 2018, pp. 279–287.

[Yao+21] Sirui Yao et al. "Measuring Recommender System Effects with Simulated Users". In: *arXiv preprint arXiv:2101.04526* (2021).

[YG15] Yang Yang and Jeff Galak. "Sentimental value and its influence on hedonic adaptation." In: *Journal of personality and social psychology* 109.5 (2015), p. 767.

[Zaj68] Robert B Zajonc. "Attitudinal effects of mere exposure". In: *Journal of personality and social psychology* 9.2 part 2 (1968).

[Zha+19] Xueru Zhang et al. "Group retention when using machine learning in sequential decision making: the interplay between user dynamics and fairness". In: *Advances in Neural Information Processing Systems* 32 (2019).

[Zha+23] Han Zhang et al. "Delayed and Indirect Impacts of Link Recommendations". In: *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 2023, pp. 545–557.

[Zho+08] Yunhong Zhou et al. "Large-scale parallel collaborative filtering for the netflix prize". In: *International conference on algorithmic applications in management*. Springer. 2008, pp. 337–348.

[Zin03] Martin Zinkevich. "Online convex programming and generalized infinitesimal gradient ascent". In: *Proceedings of the 20th international conference on machine learning (icml-03)*. 2003, pp. 928–936.