SOCIALLY-AWARE RECOMMENDER SYSTEMS MITIGATE USER-CREATOR FEATURES CLUSTERIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Recommender systems shape online interactions by matching users with creators' content to maximize engagement. Creators, in turn, adapt their content to align with users' preferences and enhance their popularity. At the same time, users' preferences evolve under the influence of both suggested content from the recommender system and content shared within their social circles. This feedback loop generates a complex interplay between users, creators, and recommender algorithms, which is the key cause of filter bubbles and polarization. We develop a social network-aware recommender system that explicitly accounts for this user-creators feedback interaction and strategically exploits the topology of the user's own social network to promote diversification. Our approach highlights how accounting for and exploiting users' social network in the recommender system design is crucial to mediate filter bubble effects while balancing content diversity with personalization. We show that a recommender system that greedily optimizes for user satisfaction leads to users' feature clusters formation. Ultimately, the proposed approach shows the power of socially-aware recommender systems in combating polarization and clusterization phenomena.

1 INTRODUCTION

The rise of streaming and e-commerce platforms has increased the need for effective Recommender Systems (RS) (Linden et al., 2003; Covington et al., 2016; Gomez-Uribe and Hunt, 2016), which personalize content to enhance user experience and maximize engagement (Li et al., 2024; Raza et al., 2025). However, by reinforcing user preferences, RS can reduce information diversity, fostering echo chambers and polarization (Cinus et al., 2022). Designing RS that maintain high user satisfaction while mitigating such global effects remains challenging (Slokom et al., 2025; Su et al., 2013), raising the key question of how to balance individual satisfaction with societal outcomes (Lanzetti et al., 2023).

The core principle of RS is to leverage past user interactions and profile information to generate personalized recommendations (Yi et al., 2019). Collaborative filtering, content-based filtering, and their hybrids are the dominant approaches (Li et al., 2024). However, while personalization improves individual satisfaction, it can amplify negative macroscopic effects such as opinion radicalization (Rossi et al., 2022; Lanzetti et al., 2023; Lin et al., 2024). While conventional countermeasures may be effective in isolated experimental settings, they fail to account for the dynamic response of content creators who strategically adapt their content to the potential target audiences (Lin et al., 2024; Dean et al., 2024b). Furthermore, most approaches neglect interactions among users, limiting their ability to operate at a macroscopic level.

Individuals are embedded in social contexts where interactions shape preferences and behaviors (Proskurnikov and Tempo, 2017). This motivates *Social RS*, which leverage the social network alongside user–item data for personalized recommendations (Ma et al., 2008; Yang et al., 2014). In contrast, we do not use neighbors' opinions to predict preferences; instead, we exploit network structure as a control mechanism to shape long-term features dynamics and mitigate clusterization.

A critical yet underexplored aspect of social RS is how to leverage the social network to mitigate harmful effects without reducing user satisfaction (Hassan, 2019). We argue that, since polarization is a collective phenomenon shaped by social interactions, promoting diversity at the individual level

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

alone is insufficient. Accounting for users' embedding in the social network is essential for designing trustworthy RS (Hassan, 2019; Chandrasekaran et al., 2024).

This paper addresses a key question: how can RS leverage social interactions to mitigate global clusterization while maintaining high user satisfaction? We show that explicitly modeling the dynamic interplay between users, creators, and RS enables a better balance between satisfaction and diversity.

Contributions. We propose a novel theoretical framework that models the RS landscape with dynamic interaction between users, creators, and the RS, where users are embedded in social networks. We leverage the user's network structure to develop a RS that mitigates opinion clusterization while maintaining high user satisfaction. Unlike previous approaches, we explicitly model the interaction between users, creators, and the RS, providing a comprehensive solution to the clusterization problem in modern social media platforms. Our main contributions are as follows:

- We propose a framework capturing the dynamic interplay between socially connected users, strategic content creators, and the RS, linking content personalization to features polarization.
- We show that a RS that greedily optimizes for user satisfaction leads to users clusters formation.
- We propose a social-network–aware recommender, $RS(d)$, where the parameter d controls the tradeoff between user satisfaction and the extent of creator clustering.
- We test our algorithm on both the *Facebook dataset* (McAuley and Leskovec, 2012) and sinthetic data and showcase that when only accounting for engagement maximization, RS increase clusterization effects over the users population.

Scope and Limitations. We analyze the relationship between engagement-based recommendation with polarization, under a theoretical framework. The aim is to bring theoretical insights that find applications in industrial practice. However, we emphasize that our model is a stylized abstraction. While we believe it captures the core dynamics of real-world recommender systems, actual platforms may exhibit additional features other than confirmation bias due to human irrationality, heterogeneous engagement incentives, or mechanisms beyond confirmation bias, that are not captured here.

2 Related Work

Negative impacts of RS. RS algorithms have been linked to societal issues such as polarization, filter bubbles, and echo chambers. Santos et al. (2021) shows that recommending links between highly similar nodes can exacerbate opinion polarization. Ziegler et al. (2005) proposes balancing diversity and personalization to improve user satisfaction, while Cheng et al. (2017); Zhang et al. (2023); Zhang and Hurley (2008) introduce optimization frameworks and metrics that promote diversity beyond accuracy-based objectives. Collectively, these works show that diversification can be improved without significantly sacrificing accuracy, and sometimes even enhancing it. However, this body of work takes a static view of RS, neglecting the dynamic interaction between users and the system.

Opinion dynamics. Opinion dynamics studies how opinions evolve and spread among interacting agents within a social network (Proskurnikov and Tempo, 2017; Parsegov et al., 2017; Friedkin and Johnsen, 1990). We explicitly account for network influence on user's features and assume that both users' and creators' feature vectors evolve dynamically according to the extended Friedkin–Johnsen (FJ) model (Friedkin and Johnsen, 1990).

Link recommendations. A line of work studies the impact of link recommendation on users under FJ dynamics (Wang and Kleinberg, 2023; Zhu et al., 2021; Kühne et al., 2025; Chitra and Musco, 2020). In particular, Wang and Kleinberg (2023) analyze how adding links in a social network affects conflict levels. Rácz and Rigobon (2023) consider how a centralized planner can modify network structure to reduce polarization, linking the problem to maximizing the Laplacian spectral gap. Zhu et al. (2021); Kühne et al. (2025); Chitra and Musco (2020) instead study link recommendation design to minimize polarization and disagreement.

Performative prediction. Performative predictions support decisions that influence the outcomes they aim to predict (Perdomo et al., 2020). This is particularly relevant for RS, which predict relevant content for users. While RS are often designed and evaluated using a supervised learning perspective with static user preferences and content catalogs, in practice they interact with and shape both user and creator behavior, creating a feedback loop (Li et al., 2024; He et al., 2025). A recent line of work

makes this explicit by modeling RS–user interactions and studying how preferences evolve under recommended content (Dean and Morgenstern, 2022; Yao et al., 2024; Sprenger et al., 2024; Lin et al., 2024; Dean et al., 2024a; Chandrasekaran et al., 2024).

On the other hand, Hron et al. (2023); Jagadeesan et al. (2023); Eilat and Rosenfeld (2023); Yao et al. (2023) focus on creators' dynamic adaptation while treating users as static. The position paper Dean et al. (2024b) instead proposes a unified dynamical system for user–creator–recommender interactions. We build on Lin et al. (2024), where users and creators co-evolve in a feedback loop. That work shows that such dynamics can induce feature polarization and that standard diversity-promoting strategies are insufficient to mitigate it.

In contrast to Lin et al. (2024), we consider users embedded in a social network, influenced both by recommended creator content and by content shared by other users. Incorporating social influence proved to be key in balancing content personalization and diversity and mitigating users' preferences clusterization. Appendix G provides a schematic overview of related work and our positioning.

3 Problem Setup

We consider a setup where users interact bidirectionally with content creators, mediated by a RS. Formally, we consider two sets of agents with n-dmensional preferences/feature vectors:

- **Users:** $\mathcal{U}^t = \{u_0^t, u_1^t, ..., u_{N-1}^t\}$, where $\boldsymbol{u}_i^t \in [-1, 1]^n$ represents the preferences/feautres vector of user i at time t. The k-th entry of $\boldsymbol{u}_i^t$ indicates the k-th feature of user i at time t. We define the global user feature's vector at time t as $\mathbf{u}^t = [(u_0^t)^\top \dots (u_{N-1}^t)^\top]^\top \in \mathbb{R}^{Nn}$.
- **Content creators:** $\mathcal{C}^t = \{c_0^t, c_1^t, ..., c_{M-1}^t\}$, where $\boldsymbol{c}_j^t \in [-1, 1]^n$ represents the content/feature vector of creator j at time t. The k-th entry of $\boldsymbol{c}_j^t$ is the k-th feature of creator j at time t. We define the global creator feature's vector at time t as $\mathbf{c}^t = [(c_0^t)^\top \dots (c_{M-1}^t)^\top]^\top \in \mathbb{R}^{Mn}$.

At each timestep t, creators publish content reflecting their current feature vectors in $\mathcal{C}^t$. We extend the framework introduced by Lin et al. (2024) to account for the social network effects in the features evolution and overall user behavioral trends. We assume closed-loop interactions among all agents, creating a dynamic feedback system where: 1) recommended content influences user interests over time, 2) content creators adapt their content based on audience implicit feedback, and 3) users influence each other on a social network. While there is a mutual influence between users and creators, the difference between these two classes of agents lies in the nature of their interactions:

- **User-user interactions:** mediated via the social network. This is, two users influence one another if they are connected directly through their social network.
- **Creator-user interactions:** mediated via a RS, where the RS presents each user with a subset of creators. In turn, creators only receive feedback from the users they reach through the RS.

The dynamics resulting from these interactions are modeled as:

$$u_i^{t+1} = f_i(\mathbf{u}^t) + h_i^t(\mathbf{c}^t), \tag{1a}$$

$$c_j^{t+1} = p_j^t(\mathbf{u}^t) + q_j(\mathbf{c}^t), \tag{1b}$$

$$c_j^t \sim \mathcal{R}(u_i^t), \qquad \forall i = 1, \dots, N \quad j = 1, \dots, M, \tag{1c}$$

where the creators in (1a) are determined by user i's decision to engage with content from creator j at time t, which is sampled from a probability distribution $\mathcal{R}$ defined by the recommendation assigned to user i by the RS at time t, as per (1c). The functional relationships for f, h, p, q and definitions in (1) are given by the FJ model (Parsegov et al., 2017) extended to include the influence of the RS in the same fashion as in Sprenger et al. (2024). Details are discussed in the following sections. For a complete discussion of the model, we refer to Appendix A.

3.1 User-user interaction

In our framework, users influence each other by means of interactions through their social network.

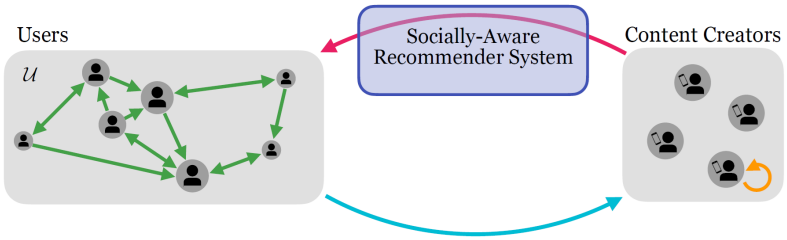


Figure 1: Overview of the dynamic framework described in equation (1). Users $\mathcal{U}$ influence each other via dynamics $f(\cdot)$, and are influenced by creators via $h^t(\cdot)$, mediated by $\mathcal{R}$. Content creators have internal features dynamics $q(\cdot)$, and are influenced by users via $p^t(\cdot)$.

Definition 1 (Social Network). *Let $A \in \mathbb{R}^{N\times N}$ be an adjacency matrix, where $A_{ij} > 0$ indicates that user j influences user i with weight A_{ij}. The* social network *is a directed and weighted graph $\mathcal{G}(\mathcal{U}, \mathcal{E}, A)$, where each node $i \in \mathcal{U}$ corresponds to a user, and $\mathcal{E}$ is the set of edges. An edge $(j, i) \in \mathcal{E}$ exists if and only if user j has a direct influence on user i, i.e. $A_{ij} \neq 0$.*

Given Definition 1 and the model in Friedkin and Johnsen (1990) for one single topic, i.e. $n = 1$,

$$f_i(\mathbf{u}^t) = (1-\lambda_i)\sum_{j=1}^{N} a_{ij} {u_j}^t + \lambda_i u_i^0, \tag{2}$$

where a_{ij} is the (i,j)-th entry of the sub-stochastic adjacency matrix of the social network $A \in [0,1]^{N\times N}$, while $\lambda_i \in (0,1]$ capture the resistance to features change ("stubbornness") of users, and u_i^0 the user's i initial preferences. We will adopt the multi-dimensional version of (2) from Parsegov et al. (2017) that is explained in detail in Section A.

We note that asymmetric social relationships, common in modern platforms with "following" dynamics, are naturally captured by this representation. More specifically, for each user we can define the set of all users that have an influence on the user via its d-hop social network. The parameter d controls the breadth of social context: larger values incorporate more distant connections.

Definition 2 (d-hop influencers). *For a graph $\mathcal{G}(\mathcal{U}, \mathcal{E}, A)$, the d-hop influencers of user $i \in \mathcal{U}$, is $in_i(d) = \{j \in \mathcal{U} \mid dist(j \to i) \leq d\}$, and $i \in in_i(d)$.*

3.2 Creator-user interaction

Creators and users influence each other by means of a RS which functions as a probability distribution over content creators: for each user i, $c_j^t \sim \mathcal{R}(u_i^t)$, where the distribution is dependent on the user's feature vector u_i^t. This is, the RS adjusts the probability distribution of content to the user's preferences; details on the recommendation strategies studied in this work are provided in the following sections.[1].

Definition 3. *Let $\mathcal{F}_1^t, \ldots, \mathcal{F}_M^t$ be the set of disjoint partitions of $\mathcal{U}$ at time t, where $\mathcal{F}_j^t$ contains all users that consume content from creator j at time t, i.e., $i \in \mathcal{F}_j^t$ if c_j^t is sampled from $\mathcal{R}(u_i^t)$ at time t.*

This allows for the definition of how content creators and users influence each other. The influence of content creators towards user i follows as

$$h_i^t(\mathbf{c}^t) = (1-\lambda_i)\sum_{j=1}^{M} b_{ij}^t c_j^t, \tag{3}$$

where b_{ij} is the $(i.j)$-th entry of the matrix $B^t \in [0,1]^{N\times M}$, with B^t such that $[A\, B^t]\vec{1}_{N+M} = \vec{1}_N$, i.e. is row stochastic, similar to Sprenger et al. (2024). b_{ij}^t describes the influence power of the recommendation on user i preferemces', where $b_{ij}^t > 0$ if $i \in \mathcal{F}_j^t$ at time t and $b_{ij}^t = 0$ otherwise.

[1]A common choice of distribution is the softmax of a utility function that maximizes confirmation bias (Nickerson, 1998). See Anas (1983); Chee et al. (2024); Kalimeris et al. (2021); Hazrati and Ricci (2022) for additional details.

This is, creator j can only directly influence user's i preferences at time t if they consume their content.

Similar to users, we assume a degree of critical thinking in the creators. This is captured by:

$$q_j(\mathbf{c}^t) = (1-\gamma_j)e_j c_j^t + \gamma_j c_j^0, \tag{4}$$

where $e_j, \gamma_j \in [0,1]$, govern the temporal consistency of the creator's features, determining the influence of their previous stance, while c_j^0 captures the creator's j initial content.

The influence of users towards content creators follows as

$$p_j^t(\mathbf{u}^t) = (1-\gamma_j)\sum_{k=1}^{N} c_{jk}^t u_k^t, \tag{5}$$

where c_{jk}^t is the (j,k)-th entry of the matrix $C^t \in [0,1]^{M\times N}$, with C^t such that $[E\, C^t]\vec{1}_{M+N} = \vec{1}_M$. c_{jk}^t describes the influence power of user k feedback on creator's j contents', where $C_{jk}^t > 0$ if $k \in \mathcal{F}_j^t$ at time t and $C_{jk} = 0$ otherwise. This is, creator j only receives feedback from the set of users consuming their content.

3.3 Recommender System

The RS mediates directly the creator $\rightarrow$ user interaction, and indirectly the user $\rightarrow$ creator interaction. The goal of a RS is to sort and present content to users to maximize their engagement. According to Nickerson (1998), this goal can be achieved by recommending content that perfectly matches the users' interests. However, such a greedy approach was shown to result in polarization and clusterization (Del Vicario et al., 2017). Hence, we study the design of a RS that maximizes users' satisfaction while reducing clusterization effects on a global scale by having the RS explicitly accounting for the social network. To that end, we provide definitions of satisfaction and clusterization.

Satisfaction. Satisfaction is quantified by measuring the cumulative distance between a user's features vector and the selected content over the entire time sequence $\{0,...,T-1\}$. Motivated by the fact that the engagement of users is driven by alignment with users' features vector (Nickerson, 1998), the RS will recommend content to maximize user satisfaction, defined as follows.

Definition 4 (User Satisfaction)**.** *The satisfaction of user i with creator j at time T is:*

$$sat(u_i^T) = -\frac{1}{T}\sum\nolimits_{t=0}^{T-1} \|u_i^t - c_j^t\|_2, \quad u_i^t \in \mathcal{F}_j^t \tag{6}$$

where $\mathcal{F}_j^t$ represents the users audience of creator j at time t.

Definition 4 formalizes the fact that users' engage more with content that is aligned with their current preferences, i.e. $u_i^t \approx c_j^t$On a global scale, the RS wants to maximize global satisfaction defined as follows.

Definition 5 (Global Satisfaction)**.** *Let $\mathcal{U}^t$ be the set of N users at time t, the global satisfaction is:*

$$sat(\mathcal{U}^t) = \frac{1}{N}\sum\nolimits_{i=0}^{N-1} sat(u_i^t). \tag{7}$$

Clusterization. Features clusterization is quantified by the silhouette coefficient (Lai et al., 2025) that each user has in its assigned cluster (as computed via k-means for the features vectors)[2]. A high silhouette coefficient indicates that a user's features are well-matched to its assigned cluster and poorly matched to neighboring clusters. If users can be clearly assigned to a cluster and thus have high silhouettes, the users features landscape is clusterized.

Definition 6 (User Silhouette)**.** *Given a user features vector $u_i^t \in \mathcal{F}_i$, the silhouette coefficient is defined as:*

$$s(u_i^t) = \frac{b(u_i^t) - a(u_i^t)}{\max\{a(u_i^t), b(u_i^t)\}} \in [-1,1], \tag{8}$$

where $a(u_i^t) = \frac{1}{|\mathcal{F}_i|-1}\sum_{u_j^t\in\mathcal{F}_i, j\neq i} ||u_i^t - u_j^t||_2$ is the average intra-cluster distance, and $b(u_i^t) = \min_{l\neq i}\frac{1}{|\mathcal{F}_l|}\sum_{\boldsymbol{u}_k^t\in\mathcal{F}_l} ||u_i^t - u_k^t||_2$ is the minimum average outer-cluster distance.

[2]The optimal number of clusters is determined by running k-means for various values of k and selecting the value that yields the highest average silhouette.

Definition 7 (Global Clusterization). *The global feature clusterization of the set $\mathcal{U}^t$ is defined as*

$$cl(\mathcal{U}^t) = \frac{1}{N}\sum_{i=0}^{N-1} s(u_i^t). \tag{9}$$

In what follows, we present a RS design that ensures balance between satisfaction and clusterization by taking into account the social network dynamics.

4 Recommender System Design

We present the design of our *socially-aware RS*. As standard in personalized feed mechanisms employed by online platforms (Koren et al., 2009), we consider a two-stage algorithmic curation strategy that determines which subset of available content reaches each user. At each time t, the RS:

(i) Computes a reference recommendation $r_i^t \in \mathbb{R}^n$ for each user i as a function of their features vector u_i^t, i.e. $r_i^t := r(u_i^t)$.

(ii) Provides the `top-k` recommendations by presenting the user with the k content items closest to this reference point through k-nearest-neighbor search:

$$\mathcal{R}(u_i^t) = \left\{ c_j^t \mid \|c_j^t - r_i^t\|_2 \in k\text{-top min}_j \{\|c_j^t - r_i^t\|_2\}_{j=1}^M \right\}. \tag{10}$$

Each user i samples one piece of content from $\mathcal{R}(u_i^t)$ at every time t, i.e., $c_j^t \sim \mathcal{R}(u_i^t)$ according to softmax with temperature parameter β^{-1}. In what follows, we provide a theoretical result that informs how the choice of $r(u_i^t)$ impacts user satisfaction and clusterization, offering a design choice that, by mimicking the structure of the social network, achieves a sweet spot between the two metrics.

4.1 Recommendation Strategies for Satisfaction and Clusterization

By studying the dynamics of the feedback mechanism in (1), it is possible to derive insights on how to design the RS to avoid clusterization while promoting user satisfaction. Consider the dynamics in (1) where no RS is intervening and the user partitions are static. In this case, (1) becomes:

$$u_i^{t+1} = (1-\lambda_i)\sum_{j=1}^{N} u_j^t + \lambda u_i^0 + (1-\lambda_i)\sum_{j=1}^{M} b_{ij} c_j^t, \tag{11a}$$

$$c_j^{t+1} = (1-\gamma_j) e_{jj} c_j^t + \gamma_j c_j^0 + (1-\gamma)\sum_{k=1}^{N} c_{jk} u_k^t, \tag{11b}$$

where b_{ij} and c_{jk} are static, since the user partitions are static. This is, users engage with the same content creator over time. In what follows, we explore the influence of the dynamic components (social network A, creator's influence B, etc.) on the user's features evolution.

We also study how different recommendation strategies influence the overall dynamic behavior and emphasize the key role the RS and the social connections have in this framework. To do that, we consider a recommendation strategy as in (10) with $k = 1$. This implies that the RS becomes a deterministic map, i.e. $c_j^t = R(u_i^t)$. Note that this relaxation allows for the bypassing of the stochastic sampling of content given a recommendation – this assumption is solely used for the theoretical results; simulations consider the full stochastic setup.

4.1.1 Social Interactions Mitigate the Effect of Recommendations

Theorem 1. *Given the system with dynamics in* (11) *and* $[A, B]\vec{1}_{N+1} = \vec{1}_N$*, the users' features reach a steady state, i.e.,* $\exists \mathbf{u}^* \in \mathbb{R}^N$ *such that* $\lim_{t\to\infty} \mathbf{u}^t = \mathbf{u}^*$*. Moreover, for a generic user i, consider a perturbation of* (11) *where B_i is increased by τ and all nonzero entries a_{ik} are decreed proportionally so that the sum is preserved. Let $u_i^*(\tau)$ denote the steady state of i as function of τ. Then, $sgn(du_i^*/d\tau) = sgn(c^* j_i - \bar{u}_i^*)$, where $\bar{u}_i^*$ is the weighted average of i's neighbors' steady states and $c_{j_i}^*$ is the recommended creator's feature. Hence, increasing recommender influence at the expense of social influence ($\tau > 0$) moves u_i^* monotonically toward $c_{j_i}^*$ and away from $\bar{u}_i^*$, and vice versa.*

The proof can be found in Appendix F.1. Intuitively, Theorem 1 states that when users consistently engage with the same creator over time, they gradually approach the creator's vectors. Creators, in

turn, adapt to their fixed audience until an equilibrium is reached. The degree to which this happens depends on the user's and creator's stubbornness (as captured by Λ and Γ, respectively) as well as the strength of the social connections (captured in A). The theorem shows that the social network and the recommended content exert opposing influences on the steady-state users' features, a phenomenon driven by users' limited attention span (Simon, 1996). The social network drives partitioned dynamics by propagating interests influences through connections, yet indirect links across clusters enable cross-partition diffusion and prevent static assignments.

4.1.2 RS THAT MAXIMIZE SATISFACTION INCREASE CLUSTERIZATION

The main goal of a RS is to maximize engagement. A greedy RS is one that solely maximizes satisfaction by providing the creator that most closely aligns with users features (Nickerson, 1998):

$$R(u_i^t) = \min_j(||c_j^t - u_i^t||_2). \tag{12}$$

This strategy achieves the maximum satisfaction as per definition of *user satisfaction* in Definition 5. For this strategy, we establish the following result, whose proof can be found in Appendix F.2.

Theorem 2. *Consider the system with dynamics as in* (1)*. For each user $u_i \in \mathcal{U}$, let us define $j_i := \arg\min_{j\in\mathcal{C}} \|c_j^0 - u_i^0\|^2$, and the separation margin $\Delta_{ik}^t := \|c_k^t - u_i^t\|^2 - \|c_{j_i}^t - u_i^t\|^2$. Let A be a diagonal matrix and $\Delta_{ik}^0 > F(A, \Gamma, \Lambda, E)$, for all $i \in \mathcal{U}$ and $k \in \mathcal{C} \setminus \{j_i\}$, where $F(A, \Gamma, \Lambda, E)$ is a threshold reported in* (20)*, Appendix F.2. Then, a greedy RS as in* (12) *induces a static user partition $\mathcal{F}_1, ..., \mathcal{F}_M$. Moreover, if $\lambda_i = 0$, then $\|u_i^t - c_{j_i}^t\| \to 0$.*

Intuitively, if the users are sufficiently closer to a creator, then they will fall into filter bubbles where they engage with creators who provide users with content that closely aligns with their preferences, yielding a static partition. This effect is aggravated when the users do not have social interactions to counterbalance the effect of the RS, as per Theorem 1, since static user-creator mappings lead to the formation of an equilibrium which fosters clusterization. Moreover, the closer a user's feature vector gets to the creators' contents, the smaller the chances for them to exit the bubble. The critical insight is that global polarization emerges when social network effects are insufficient to counterbalance the RS's homogenizing influence. This tipping point, where clustering effects begin to dominate, depends on the relative strength of social connections versus recommendation influence. We showcase this phenomenon empirically in the experiments section. For the case of stubborn creators, Theorem 2 simplifies as follows, with proof in Appendix F.3.

Lemma 1. *Consider the system with dynamics as in* (1)*. Let A be a diagonal matrix and creators being stubborn, i.e. $\Gamma = I_M$. Then, a greedy RS as in* (12) *induces a static user partition $\mathcal{F}_1, ..., \mathcal{F}_M$ and for any user $u_i \in \mathcal{F}_j$, all features approach the recommended creator's content, namely $||u_i^t - c_j^t||_2$ decreases monotonically in time.*

4.1.3 RS THAT REDUCE CLUSTERIZATION DECREASE SATISFACTION

We have shown that targeting only satisfaction, in the absence of social interactions, promotes users' clusterization around creators. To counter the static assignment of each user to creators, the RS has to depart from engagement maximization and promote content diversity.

Corollary 1. *Let $c_j^t = R(u_i^t)$ in* (11) *$\forall i, j$ result from a non-greedy RS, that is, $R(u_i^t) \neq min_j(||c_j^t - u_i^t||_2)$. Then the user satisfaction in* (7) *is suboptimal.*

Together with Theorem 2 and Lemma 1, this statement (proof in Appendix F.4) promotes the search of a tradeoff between maximizing users satisfaction while countering clusterization.

4.2 A SOCIALLY-AWARE RECOMMENDER SYSTEM MIMICS THE SOCIAL NETWORK

Since social connections mitigate clustering, we propose a socially aware strategy that leverages the network structure. Rather than optimizing only for individual preferences, using the concept of *d-hop influencers* (Definition 2), the RS mimics social influence by basing recommendations on the mean interests of each user's d-hop neighborhood.

Definition 8 (d-hop socially-aware RS)**.** *Let $in_i(d)$ be the d-hop influencers of user i. The d-hop socially aware RS produces a recommendation reference that is given as:*

$$R(u_i^t) = \frac{1}{|in_i(d)|} \sum\nolimits_{j\in in_i(d)} u_j^t. \tag{13}$$

Parameter d controls the tradeoff between personalization and diversification: $d = 0$ recovers the greedy strategy in (12), while larger values incorporate broader social influence.

Intuitively, by the dynamics in (11), if the RS provides content aligned with the d-hop neighborhood average preferences then users' features are pulled towards the average preferences of their neighborhood, including neighbors from other clusters. This mechanisms will reduce the separation between clusters. Increasing d enhances content diversity at the cost of reduced user satisfaction, as recommendations deviate further from individual preferences. In the next section, we empirically analyze this tradeoff and identify parameter values that minimize global clustering while maintaining user satisfaction.

Remark 1 (Scalability of d-hop socially aware RS)**.** *Note that the d-hop socially aware recommendation strategy is fully parallelizable, since each user's recommendation only depends on their local environment.*

5 EXPERIMENTS

We evaluate the performance of our RS using the satisfaction and clusterization metrics as in (5) and (7). First, we study the closed loop interaction of our RS with a synthetic social network. Then, we demonstrate our findings on the Facebook-ego dataset (McAuley and Leskovec, 2012), which comprises a social network of 4039 users. Throughout this section, the temperature parameter governing the users choice is chosen as $\beta^{-1} = 0.5$. We present our main findings below.

Synthetic Dataset. We consider a network of $N = 600$ users and $M = 50$ creators, with average user in-degree 11, and parameters in Appendix B.1. We set the feature vector dimension to $n = 2$, initializing all users and creators uniformly in $[-1, 1]^n$. The social network is randomly generated with edge probabilities decreasing with features distance, so users with similar preferences are more likely to connect, consistent with homophily (McPherson et al., 2001). Details of the connection probability are in Appendix B.2. The RS follows a $\texttt{top}$–k recommendation strategy with $k = 5$. The impact of different k and varying interaction parameters in (1) on all performance metrics for the following settings is in Appendix C.

Results. We investigate three distinct RS strategies: (i) RS only optimizes for user engagement (greedy RS with $d = 0$), (ii) RS only accounts for users diversity (*socially aware RS* with $d = 6$, high diversification), and (iii) hybrid RS that balances satisfaction and diversity (socially aware RS with $d = 3$, intermediate localization). Strategy (i) only accounts for personalized recommendations, strategy (ii) emphasizes social interactions by considering broad network effects, while strategy (iii) seeks trades-off between user satisfaction and content diversity.

Figure 2 shows the evolution of user and creator features under the three RS strategies. User clusters (via k-means) are depicted as ellipsoids with axes equal to twice the standard deviation along principal components, with transparency decreasing as global clusterization increases; clusters below 0.5 are omitted. The last column reports negative global clusterization and user satisfaction over time, both ideally increasing. The greedy RS (i) forms distinct, increasingly polarized clusters around creators; the socially-aware RS (ii) reduces clusterization but lowers satisfaction; the hybrid RS (iii) maintains high satisfaction while mitigating clusterization, balancing both objectives.

Figure 3 displays negative global clusterization and user satisfaction as functions of the localization parameter d at $t = 50$ and $t = 500$, along with their variance (computed over users as in (7) and (9)). As d increases, both clusterization and satisfaction decline. A sweet spot occurs at $d = 3$, where clusterization drops sharply while satisfaction remains high.

Real Dataset. The ego-Facebook dataset comprises $N = 4039$ users on an undirected social network with average degree 45. We consider $M = 120$ creators with random initial features and dimension $n = 3$ (parameters in Appendix D.1). To generate initial user features that reflect the homophilic structure of social networks, we use spectral clustering (Luxburg, 2004) to identify communities and assign similar features to users within the same community (see Appendix D.2).

Results. The RS operates under a $\texttt{top}$–k recommendation strategy with $k = 5$.

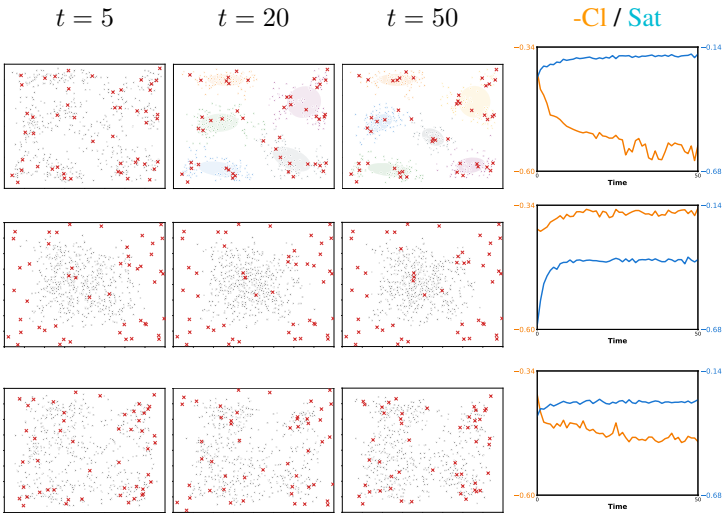
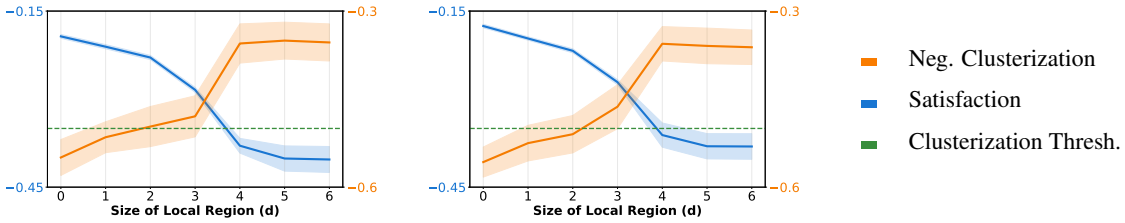


Figure 2: Snapshots of the features environment simulated with a localized region (i) $d = 0$, first row, (ii) $d = 6$, second row (iii) $d = 3$, third row. $\times$ denotes the creator, $\bullet$ the users respectively.



(a) After 50 timesteps (b) After 500 timesteps

Figure 3: Global clusterization and global user satisfaction plotted as d varies after (a) 50 and (b) 500 timesteps. Clusterization thresh is set to -0.5, for which clusters are no longer distinguishable.

Figure 4 displays the user landscapes after $t = 20$ time-steps for (i) a greedy RS with $d = 0$ and (ii) a hybrid RS with $d = 3$. For illustration purposes, we omitted the creators features in the 3-dimensional figure on the left side and only display every fourth feature for the users and the creators features respectively in the figures showing the projections on the xy-, xz- and yz-planes.

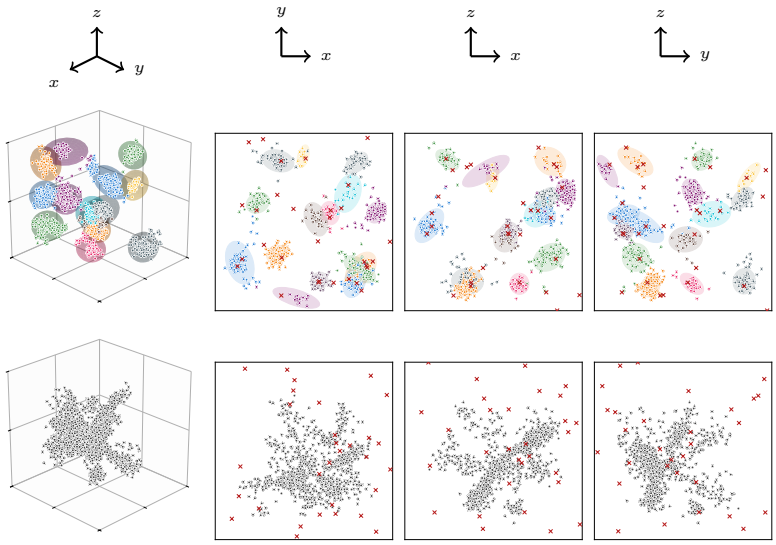


Figure 4: Snapshots after $t = 20$ timesteps of the features environment simulated with a localized region (i) $d = 0$, first row, (ii) $d = 3$, second row. $\times$ denotes the creator, $\bullet$ the users respectively.

REFERENCES

Alex Anas. Discrete choice theory, information-theory and the multinomial logit and gravity models. *Transportation Research Part B: Methodological*, 17:13–23, 02 1983.

Sanjay Chandrasekaran, Giulia De Pasquale, Giuseppe Belgioioso, and Florian Dörfler. Network-aware recommender system via online feedback optimization, 2024.

Jerry Chee, Shankar Kalyanaraman, Sindhu Kiranmai Ernala, Udi Weinsberg, Sarah Dean, and Stratis Ioannidis. Harm mitigation in recommender systems under user preference dynamics. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD '24, page 255–265. ACM, 2024.

Peizhe Cheng, Shuaiqiang Wang, Jun Ma, Jiankai Sun, and Hui Xiong. Learning to recommend accurate and diverse items. In *Proceedings of the 26th International Conference on World Wide Web*, WWW '17, page 183–192, Republic and Canton of Geneva, CHE, 2017. International World Wide Web Conferences Steering Committee.

Uthsav Chitra and Christopher Musco. Analyzing the impact of filter bubbles on social network polarization. In *Proceedings of the 13th International Conference on Web Search and Data Mining*, WSDM '20, page 115–123, New York, NY, USA, 2020. Association for Computing Machinery.

Federico Cinus, Marco Minici, Corrado Monti, and Francesco Bonchi. The effect of people recommenders on echo chambers and polarization. *Sixteenth International AAAI Conference on Web and Social Media (ICWSM 2022)*, pages 90–101, 2022.

Paul Covington, Jay Adams, and Emre Sargin. Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems*, RecSys '16, page 191–198, New York, NY, USA, 2016. ACM.

Sarah Dean and Jamie Morgenstern. Preference dynamics under personalized recommendations. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 795–816, 2022.

Sarah Dean, Evan Dong, Meena Jagadeesan, and Liu Leqi. Accounting for AI and users shaping one another: The role of mathematical models. *Transactions on Machine Learning Research*, pages 1–25, 2024a.

Sarah Dean, Evan Dong, Meena Jagadeesan, and Liu Leqi. Recommender systems as dynamical systems: Interactions with viewers and creators. In *Workshop on Recommendation Ecosystems: Modeling, Optimization and Incentive Design*, 2024b.

Michela Del Vicario, Antonio Scala, Guido Caldarelli, H Eugene Stanley, and Walter Quattrociocchi. Modeling confirmation bias and polarization. *Scientific reports*, 7(1):40391, 2017.

Itay Eilat and Nir Rosenfeld. Performative recommendation: Diversifying content via strategic incentives. In *International Conference on Machine Learning*, 2023. URL https://api.semanticscholar.org/CorpusID:256697126.

Noah E Friedkin and Eugene C Johnsen. Social influence and opinions. *Journal of mathematical sociology*, 15(3-4):193–206, 1990.

Carlos A. Gomez-Uribe and Neil Hunt. The netflix recommender system: Algorithms, business value, and innovation. *ACM Transaction Management Information Systems*, 6(4):1–19, 2016.

Taha Hassan. Trust and trustworthiness in social recommender systems. In *Companion Proceedings of The 2019 World Wide Web Conference*, WWW '19, page 529–532, New York, NY, USA, 2019. ACM.

Naieme Hazrati and Francesco Ricci. Recommender systems effect on the evolution of users' choices distribution. *Information Processing and Management*, 59:102766, 01 2022.

Zhiyu He, Saverio Bolognani, Florian Dörfler, and Michael Muehlebach. Decision-dependent stochastic optimization: The role of distribution dynamics, 2025. URL https://arxiv.org/abs/2503.07324.

Jiri Hron, Karl Krauth, Michael I Jordan, Niki Kilbertus, and Sarah Dean. Modeling content creator incentives on algorithm-curated platforms. In *International Conference on Learning Representations*, 2023.

Meena Jagadeesan, Nikhil Garg, and Jacob Steinhardt. Supply-side equilibria in recommender systems. *Advances in Neural Information Processing Systems*, 36:14597–14608, 2023.

Dimitris Kalimeris, Smriti Bhagat, Shankar Kalyanaraman, and Udi Weinsberg. Preference amplification in recommender systems. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, KDD '21, page 805–815. ACM, 2021.

Yehuda Koren, Robert Bell, and Chris Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009. doi: 10.1109/MC.2009.263.

Marino Kühne, Panagiotis D. Grontas, Giulia De Pasquale, Giuseppe Belgioioso, Florian Dorfler, and John Lygeros. Optimizing social network interventions via hypergradient-based recommender system design. In *Forty-second International Conference on Machine Learning*, 2025.

Huixia Lai, Tao Huang, Binlong Lu, Shi Zhang, and Ruliang Xiao. Silhouette coefficient-based weighting k-means algorithm. *Neural Computing and Applications*, 37(5):3061–3075, 2025. doi: 10.1007/s00521-024-10706-0.

Nicolas Lanzetti, Florian Dörfler, and Nicolò Pagan. The impact of recommendation systems on opinion dynamics: Microscopic versus macroscopic effects. In *2023 62nd IEEE Conference on Decision and Control (CDC)*, pages 4824–4829, 2023.

Yang Li, Kangbo Liu, Ranjan Satapathy, Suhang Wang, and Erik Cambria. Recent developments in recommender systems: A survey [review article]. *IEEE Computational Intelligence Magazine*, 19 (2):78–95, 2024.

Tao Lin, Kun Jin, Andrew Estornell, Xiaoying Zhang, Yiling Chen, and Yang Liu. User-creator feature polarization in recommender systems with dual influence. In *Proceedings of the 38th International Conference on Neural Information Processing Systems*, NeurIPS '24, Red Hook, NY, USA, 2024.

Greg Linden, Brent Smith, and Jeremy York. Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Computing*, 7(1):76–80, 2003.

Ulrike Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17:395–416, 01 2004.

Hao Ma, Haixuan Yang, Michael R Lyu, and Irwin King. Sorec: social recommendation using probabilistic matrix factorization. In *Proceedings of the 17th ACM conference on Information and knowledge management*, pages 931–940, 2008.

Julian McAuley and Jure Leskovec. Learning to discover social circles in ego networks. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'12, page 539–547, Red Hook, NY, USA, 2012. Curran Associates Inc.

Miller McPherson, Lynn Smith-Lovin, and James M. Cook. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27:415–444, 2001.

Raymond Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2:175–220, 06 1998.

Sergey E. Parsegov, Anton Proskurnikov, Roberto Tempo, and Noah Friedkin. Novel multidimensional models of opinion dynamics in social networks. *IEEE Transactions on Automatic Control*, 62:2270–2285, 05 2017.

Juan Perdomo, Tijana Zrnic, Celestine Mendler-Dünner, and Moritz Hardt. Performative prediction. In *International Conference on Machine Learning*, pages 7599–7609. PMLR, 2020.

Anton V. Proskurnikov and Roberto Tempo. A tutorial on modeling and analysis of dynamic social networks. Part I. *Annual Reviews in Control*, 43:65–79, 2017. ISSN 1367-5788.

Shaina Raza, Mizanur Rahman, Safiullah Kamawal, Armin Toroghi, Ananya Raval, Farshad Navah, and Amirmohammad Kazemeini. A comprehensive review of recommender systems: Transitioning from theory to practice. *arXiv:2407.13699*, pages 1–66, 2025.

Wilbert Samuel Rossi, Jan Willem Polderman, and Paolo Frasca. The closed loop between opinion formation and personalized recommendations. *IEEE Transactions on Control of Network Systems*, 9(3):1092–1103, 2022.

Miklos Z. Rácz and Daniel E. Rigobon. Towards consensus: Reducing polarization by perturbing social networks. *IEEE Transactions on Network Science and Engineering*, 10(6):3450–3464, 2023. doi: 10.1109/TNSE.2023.3262970.

Fernando Santos, Yphtach Lelkes, and Simon A. Levin. Link recommendation algorithms and dynamics of polarization in online social networks. *Proceedings of the National Academy of Sciences*, 50(118), 2021.

Herbert A Simon. Designing organizations for an information-rich world. *International Library of Critical Writings in Economics*, 70:187–202, 1996.

Manel Slokom, Savvina Danil, and Laura Hollink. How to diversify any personalized recommender? *Advances in Information Retrieval. ECIR 2025*, 15575:1–17, 2025.

Ben Sprenger, Giulia De Pasquale, Raffaele Soloperto, John Lygeros, and Florian Dörfler. Control strategies for recommendation systems in social networks. *IEEE Control Systems Letters*, 8: 634–639, 2024.

Ruilong Su, Li'Ang Yin, Kailong Chen, and Yong Yu. Set-oriented personalized ranking for diversified top-n recommendation. In *Proceedings of the 7th ACM Conference on Recommender Systems*, RecSys '13, page 415–418, New York, NY, USA, 2013. ACM.

Yanbang Wang and Jon Kleinberg. On the relationship between relevance and conflict in online social link recommendations. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, *Advances in Neural Information Processing Systems*, volume 36, pages 36708–36725. Curran Associates, Inc., 2023.

Xiwang Yang, Yang Guo, Yong Liu, and Harald Steck. A survey of collaborative filtering based social recommender systems. *Computer communications*, 41:1–10, 2014.

Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang, and Haifeng Xu. How bad is top-k recommendation under competing content creators? In *International Conference on Machine Learning*, pages 39674–39701. PMLR, 2023.

Fan Yao, Yiming Liao, Mingzhe Wu, Chuanhao Li, Yan Zhu, James Yang, Jingzhou Liu, Qifan Wang, Haifeng Xu, and Hongning Wang. User welfare optimization in recommender systems with competing content creators. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 3874–3885, 2024.

Xinyang Yi, Ji Yang, Lichan Hong, Derek Zhiyuan Cheng, Lukasz Heldt, Aditee Kumthekar, Zhe Zhao, Li Wei, and Ed Chi. Sampling-bias-corrected neural modeling for large corpus item recommendations. In *Proceedings of the 13th ACM Conference on Recommender Systems*, RecSys '19, page 269–277, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450362436.

Mi Zhang and Neil Hurley. Avoiding monotony: improving the diversity of recommendation lists. In *Proceedings of the 2008 ACM Conference on Recommender Systems*, RecSys '08, page 123–130, New York, NY, USA, 2008. ACM.

Xiaoying Zhang, Hongning Wang, and Hang Li. Disentangled representation for diversified recommendations. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, WSDM '23, page 490–498, New York, NY, USA, 2023. ACM.

Liwang Zhu, Qi Bao, and Zhongzhi Zhang. Minimizing polarization and disagreement in social networks via link recommendation. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan, editors, *Advances in Neural Information Processing Systems*, volume 34, pages 2072–2084. Curran Associates, Inc., 2021.

Cai-Nicolas Ziegler, Sean M. McNee, Joseph A. Konstan, and Georg Lausen. Improving recommendation lists through topic diversification. In *Proceedings of the 14th International Conference on World Wide Web*, WWW '05, page 22–32, New York, NY, USA, 2005. ACM.

A Multi-Dimensional Friedkin-Johnsen Model

The multi-dimensional FJ model (Parsegov et al., 2017) extends the classical scalar model in (2) to the case where each user is characterized by a vectorial representation. Let $u_i^t \in \mathbb{R}^n$ denote the feature vector of user i at time t, where each entry corresponds to a distinct feature. Stacking all N users' features into a single vector $\mathbf{u} \in \mathbb{R}^{Nn}$, the features update rule can be written as

$$f(\mathbf{u}^t) = ((I_N - \Lambda)\hat{A} \otimes C)\mathbf{u}^t + (\Lambda \otimes I_n)\mathbf{u}^0, \tag{14}$$

where $\hat{A} \in \mathbb{R}^{N\times N}$ is the row-stochastic influence matrix describing interpersonal influence in the network, $\Lambda \in \mathbb{R}^{N\times N}$ is the diagonal susceptibility matrix capturing how attached each influence is to their own prejudice $u^0 \in \mathbb{R}^{Nn}$ versus the social influence, and $\otimes$ denotes the Kronecker product. The matrix C is a correlation matrix among different features. We consider the case of uncorrelated features, and hence set $C = I_n$. In this special case, the model reduces to

$$f(\mathbf{u}^t) = ((I_N - \Lambda)\hat{A} \otimes I_n)\mathbf{u}^t + (\Lambda \otimes I_n)\mathbf{u}^0. \tag{15}$$

B Environment Variables for Synthetic Dataset

B.1 Parameters for user-creator Dynamics

The parameters governing the dynamics in (1) are sampled independently from uniform distributions with bounds given in Table 1.

Table 1: Simulation Parameters for Uniform Distribution Sampling

Parameter	**Lower Bound**	**Upper Bound**
User Parameters		
User Stubbornness λ_i	0.0	0.5
User Self-Influence a_{ii}	0.5	0.8
Recommender Influence b_{ij}	0.2	0.8
Neighbor Influence a_{ij}	0.025	0.05
Creator Parameters		
Creator Stubbornness γ_j	0.0	0.5
Creator Self-Influence e_j	0.5	0.8
User-Creator Influence c_j	0.2	0.8

The user-creator influence is evenly distributed among the audience set of creator j; specifically, for creator j with audience set $\mathcal{F}_j$, each user $i \in \mathcal{F}_j$ exerts influence $c_{ji} = c_j/|\mathcal{F}_j|$. The overall social influence on user i is determined by summing the influences from all neighbors. Each user is influenced by exactly one creator. Thus, referring to the FJ model in section A, we obtain the stochastic constraints: $a_{ii} + b_{ij} + \sum_{j=0}^{N-1} a_{ij} = 1$ for users and $c_j + e_j = 1$ for creators.

B.2 User-User interaction Probability

For each user j, we assume the connection to any other user with a probability that is given as:

$$\text{Prob(user } j \text{ influences user } i) = \exp(-\delta||u_i^0 - u_j^0||_2)$$

where $u_i^0, u_j^0 \in [-1, 1]^n$ are the features vectors of users i and j at time 0, and $\delta > 0$ is a parameter controlling the connectivity of the network. Different parameters of δ lead to different number of connections. Different choices of parameter δ and the resulting average node degrees are displayed in Table 2. We choose $\delta = 9$, to recover 11 neighbors.

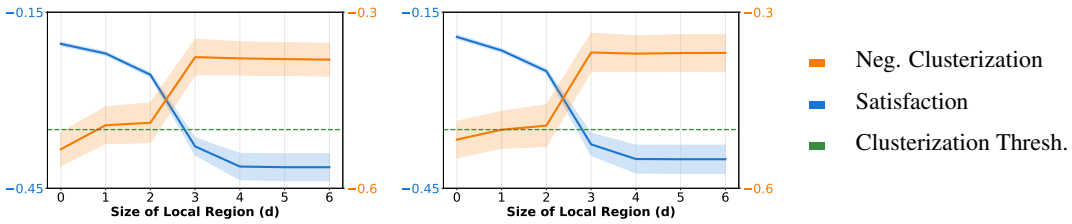
Table 2: Network connectivity for different parameters δ with $N = 600$ users initialized randomly

Parameter δ	**Average Connections**
6	21
7	17
8	14
9	11

C Results for variational environment on synthetic data

C.1 Varying the Simulation Parameters for the dynamics

We increase the number of social interactions by setting $\delta = 6$, yielding an average of 21 connections per user. The expanded social network exposes users to a broader spectrum of features through peer interactions. Figure 5a and fig. 5b present the global clusterization and satisfaction metrics after 50 and 500 timesteps, respectively. The results demonstrate that increased social connectivity mitigates clusterization, even under the greedy recommender system ($d = 0$).



(a) After 50 timesteps (b) After 500 timesteps

Figure 5: Global clusterization and global user satisfaction plotted as d varies after (a) 50 and (b) 500 timesteps with more social interactions as opposed to fig. 3. Clusterization thresh is set to -0.5, for which clusters are no longer distinguishable.

C.2 Varying k for top-k

The feature dynamics under socially-aware recommender systems are examined for localization parameters $d \in \{0, 3\}$ and recommendation set sizes $k \in \{1, 2, 3, 4\}$. Figures 6–9 display the corresponding features landscapes. For $k = 1$, the dynamics becomes stationary after $t = 20$ (fig. 6). Cluster visualization is omitted given the low global clusterization score. The number of clusters increases as k decreases, reflecting reduced creator-user interaction diversity. The recommender system with $d = 3$ produces fewer clusters than $d = 0$ across all values of k.

D Environment Variables for Real Dataset

D.1 Parameters for User-Creator Dynamics

The parameters governing the dynamics in (1) are sampled independently from uniform distributions with bounds given in Table 3.

The user-creator influence is evenly distributed among the audience set of creator j; specifically, for creator j with audience set $\mathcal{F}_j$, each user $i \in \mathcal{F}_j$ exerts influence $c_{ji} = c_j/|\mathcal{F}_j|$. The overall social influence on user i is determined by summing the influences from all neighbors. Each user

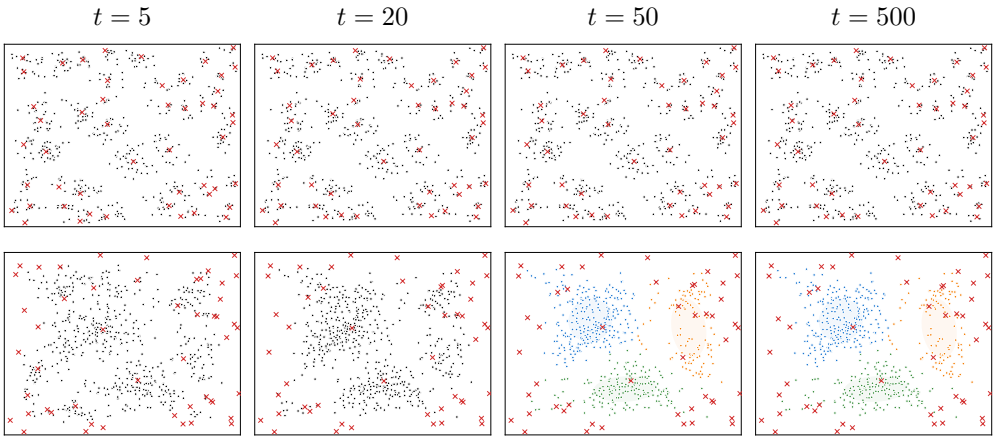


Figure 6: Snapshots of the features environment with $k = 1$ simulated with a localized region $d = 0$, first row, $d = 3$, second row. $\times$ denotes the creator, $\bullet$ the users respectively.

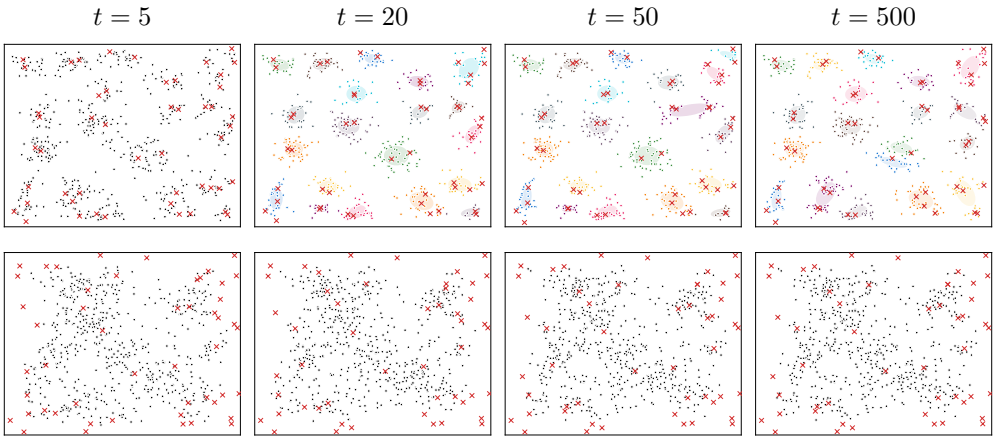


Figure 7: Snapshots of the features environment with $k = 2$ simulated with a localized region $d = 0$, first row, $d = 3$, second row. $\times$ denotes the creator, $\bullet$ the users respectively.

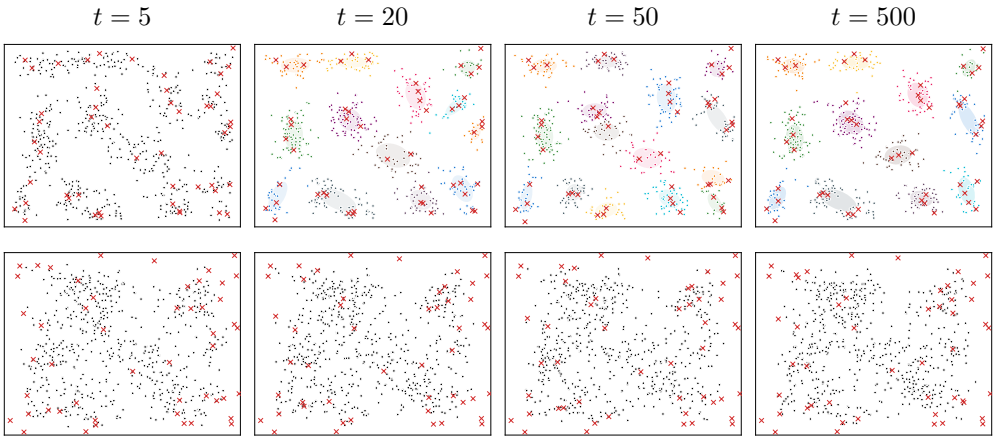


Figure 8: Snapshots of the features environment with $k = 3$ simulated with a localized region $d = 0$, first row, $d = 3$, second row. $\times$ denotes the creator, $\bullet$ the users respectively.

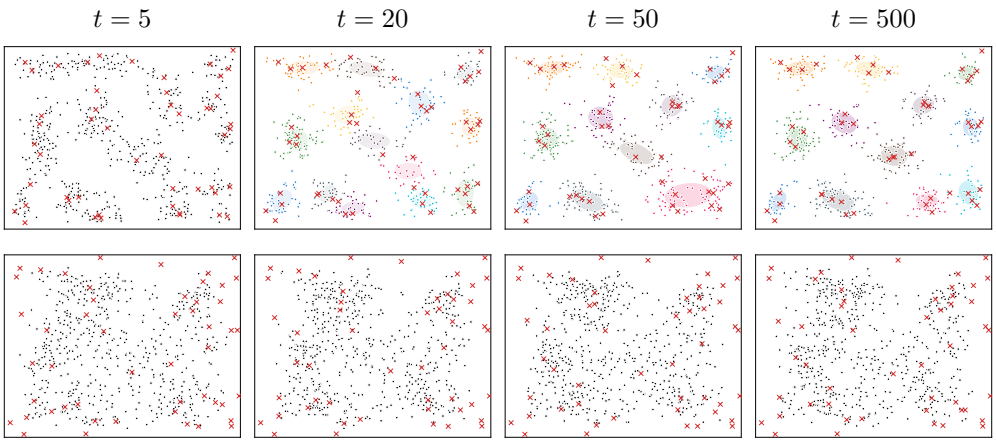


Figure 9: Snapshots of the features environment with $k = 4$ simulated with a localized region $d = 0$, first row, $d = 3$, second row. $\times$ denotes the creator, $\bullet$ the users respectively.

Table 3: Simulation Parameters for Uniform Distribution Sampling

Parameter	**Lower Bound**	**Upper Bound**
User Parameters		
User Stubbornness λ_i	0.0	0.5
User Self-Influence a_{ii}	0.5	0.8
Recommender Influence b_{ij}	0.2	0.8
Neighbor Influence $\sum_{j=0}^{N-1} a_{ij}$	0.25	0.5
Creator Parameters		
Creator Stubbornness γ_j	0.0	0.5
Creator Self-Influence e_j	0.5	0.8
User-Creator Influence c_j	0.2	0.8

is influenced by exactly one creator. Thus, referring to the FJ model in section A, we obtain the stochastic constraints: $a_{ii} + b_{ij} + \sum_{j=0}^{N-1} a_{ij} = 1$ for users and $c_j + e_j = 1$ for creators.

D.2 EGO-FACEBOOK DATASET

The network comprises 4039 anonymous users and their social connections. The resulting graph has an average degree of 45 with a number of neighbors reaching from 2 to 1046. We identify 34 community centers $\{C_1, C_2, ..., C_{34}\}$ that are randomly dispersed in $[0, 1]^3$ and apply spectral clustering to assign each user to one of the specified communities. After assignment, any user i, assigned to community center j, is initialized with $u_i^0 = C_j + \epsilon_i$, with $\epsilon_i \sim \mathcal{N}(0, 0.15)$.

E Compute Resources

All simulations and experiments were conducted on a MacBook Air equipped with an Apple M2 chip and 8 GB of unified memory, running macOS 15.6.1.

F Theorem and Lemmas Proofs

F.1 Proof of Theorem 1

Proof. The proof follows by the extended FJ dynamics with

$$\begin{bmatrix} u^{t+1} \\ c^{t+1} \end{bmatrix} = \begin{bmatrix} I-\Lambda & 0 \\ 0 & I-\Gamma \end{bmatrix} \underbrace{\begin{bmatrix} A & B \\ C & E \end{bmatrix}}_{\Pi} \begin{bmatrix} u^t \\ c^t \end{bmatrix} + \begin{bmatrix} \Lambda & 0 \\ 0 & \Gamma \end{bmatrix} \begin{bmatrix} u^0 \\ c^0 \end{bmatrix} \tag{16}$$

with $\Pi \in \mathbb{R}^{(N+M)\times(N+M)}$ row stochastic. From (Sprenger et al., 2024, Theorem 2) one gets that the limit for $t \to \infty$ of (16) exists and that in the limit we have

$$\begin{bmatrix} u^{\infty} \\ c^{\infty} \end{bmatrix} = \left(\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} - \underbrace{\begin{bmatrix} I-\Lambda & 0 \\ 0 & I-\Gamma \end{bmatrix} \begin{bmatrix} A & B \\ C & E \end{bmatrix}}_{J} \right)^{-1} \begin{bmatrix} \Lambda u^0 \\ \Gamma c^0 \end{bmatrix}. \tag{17}$$

Observe that the rows of J associated with the users sum to $1-\lambda_i$, while those associated with creators sum to $1-\gamma_j$. Hence, by Gershgorin's circle theorem, it holds that the largest eigenvalue un modulus of J satisfies $\mu_{\max} \leq 1 - \min(\min_i \lambda_i, \min_j \gamma_j) < 1$, since $\lambda_i, \gamma_j \in (0,1]$. Hence $(I-J)^{-1} = \sum_{k\geq 0} J^k \geq 0$ exists, proving convergence to a steady state.

Consider now the scenario in which we perturb the ith row of J, increasing B_i by τ and decreasing the entries A_{ij} proportionally to preserve the sum. So $x^*(\tau) = (I - J(\tau))^{-1} y$ is well defined and continuously differentiable. Hence, we can differentiate it.

Differentiating $(I - J(\tau))x^*(\tau) = y$ gives $\dot{x}^*(\tau) = (I - J(\tau))^{-1} J'(\tau) x^*(\tau)$. Since only row i of $J(\tau)$ varies, with $\partial_\tau[(1-\lambda_i)a_{ik}] = -(1-\lambda_i)a^0_{ik}/S^0_i$ and $\partial_\tau[(1-\lambda_i)B_i] = 1-\lambda_i$, we get

$$J'(\tau)x^*(\tau) = (1-\lambda_i)\big(c^*_{j_i}(\tau) - \bar{u}^*_i(\tau)\big)\, e_i, \tag{18}$$

yielding

$$\frac{d}{d\tau} u^*_i(\tau) = (1-\lambda_i)\big[(I - J(\tau))^{-1}\big]_{ii} \big(c^*_{j_i}(\tau) - \bar{u}^*_i(\tau)\big), \tag{19}$$

and since $(I-J(\tau))^{-1} = I + J(\tau) + J(\tau)^2 + \cdots \geq I$ entrywise, $[(I-J(\tau))^{-1}]_{ii} \geq 1 > 0$. Thus $\operatorname{sgn}(du^*_i(\tau)/d\tau) = \operatorname{sgn}(c^*_{j_i} - \bar{u}^*_i)$, independent of higher-order terms. Monotonicity of $u^*_i(\tau)$ on any interval where $c^*_{j_i}(\tau) - \bar{u}^*_i(\tau)$ keeps constant sign follows by the mean value theorem. □

F.2 Proof of Theorem 2

Before detailing the proof, we explicitly report here the value of the threshold:

$$F(A,\Gamma,\Lambda,E) := 8n \frac{5 + \beta_{\max} - \theta_{\max} - 2\alpha_{\max} - 2\gamma_{\max}}{(1-\alpha_{\max})(1-\theta_{\max}) - \beta_{\max}(1-\gamma_{\min})}, \tag{20}$$

where $\alpha_{\max} := \max_i (1-\lambda_i)a_{ii}$, $\beta_{\max} := \max_i (1-\lambda_i)(1-a_{ii})$, $\theta_{\max} := \max_j (1-\gamma_j)e_{jj}$, and $\gamma_{\min} := \min_j \gamma_j$.

Interestingly, this critical threshold is lowered when users and creators are highly inertial (large a_{ii} and e_{jj}) and stubborn (large Λ_i and γ_j), making clusterization more likely.

Proof. The proof consists of three steps. First, we derive bounds on the maximal total variation of the features of the users and of the creator while partitions remains constant. Then, we use these bounds to prove that under the assumptions of Lemma 2, the initial user partition $\mathcal{F}_1, ..., \mathcal{F}_M$ remains

constant. Finally, we study the behaviour of $||u_i^t - c_{j_i}^t||$, with c_{j_i} the closest creator to user u_i at time 0, i.e. $j_i := \arg\min_{j \in \mathcal{C}} \|c_j^0 - u_i^0\|^2$, showing that it converges to zero as time goes to infinity.

Step I. Under the assumption that A is diagonal and partitions are constant up to time t, we write

$$u_i^{t+1} = (1-\lambda_i)a_{ii}u_i^t + \lambda_i u_i^0 + (1-\lambda_i)b_i c_{j_i}^t, \tag{21}$$

where $b_i = 1 - a_{ii}$, due to the fact that rows of $[A\,B]$ sum to 1. For the sake of simplicity, define $\alpha_i := (1-\lambda_i)a_{ii}$ and $\beta_i := (1-\lambda_i)(1-a_{ii})$. Hence, bound

$$u_i^{t+1} - u_i^t = \alpha_i(u_i^t - u_i^{t-1}) + \beta_i(c_{j_i}^t - c_{j_i}^{t-1}), \tag{22}$$

and, using the triangular inequality

$$\|u_i^{t+1} - u_i^t\| \leq \alpha_i \|u_i^t - u_i^{t-1}\| + \beta_i \|c_{j_i}^t - c_{j_i}^{t-1}\|. \tag{23}$$

Defining $x_i^t := \||u_i^{t+1} - u_i^t\|$ and $y_j^t := \|c_j^t - c_j^{t-1}\|$, (23) can be re-written as the recursive inequality:

$$x_i^{t+1} \leq \alpha_i x_i^t + \beta_i y_{j_i}^t \leq \alpha_i^{t+1} x_i^0 + \beta_i \sum_{s=0}^{t} \alpha_i^{t-s} y_j^s \tag{24}$$

Defining $R_i = \sum_{s=1}^t x_i^s$ and $S_j = \sum_{s=1}^t y_j^s$, and summing the terms of the recursion in (24) with the initial bound that $x_i^0 \leq 2\sqrt{n}$ due to the bounded domain of $u_i^0 \in [-1,1]^n$, we obtain

$$R_i \leq \frac{2\sqrt{n}}{1-\alpha_i} + \frac{\beta_i}{1-\alpha_i} S_{j_i}. \tag{25}$$

From the creators' dynamics, assuming constant partitioning up to time t, we write

$$c_j^{t+1} = (1-\gamma_j)e_{jj}c_j^t + \gamma_j c_j^0 + (1-\gamma_j)\sum_{i \in \mathcal{F}_j} c_{ji} u_i^t. \tag{26}$$

We define $\theta_j := (1-\gamma_j)e_{jj}$, from which we derive

$$c_j^{t+1} - c_j^t = \theta_j(c_j^t - c_j^{t-1}) + (1-\gamma_j)\sum_{i \in \mathcal{F}_j} c_{ji}(u_i^t - u_i^{t-1}). \tag{27}$$

Hence, following a similar argument as above, we obtain the recursive inequality:

$$y_j^{t+1} \leq \theta_j y_j^t + (1-\gamma_j)\sum_{i \in \mathcal{F}_j} c_{ji} x_i^t, \tag{28}$$

which ultimately yields

$$S_j \leq \frac{2\sqrt{n}}{1-\theta_j} + \frac{1-\gamma_j}{1-\theta_j}\sum c_{ji} R_i. \tag{29}$$

While the two sets of inequalities can be solved, this requires to invert matrix C, which hinders the derivation of a closed-form expression. To this aim, we derive a conservative bound uniform in the model parameters, by introducing $R_{\max} := \max_i R_i$ and $S_{\max} := \max_j S_j$, as well as $\alpha_{\max} := \max_i \alpha_i$, $\beta_{\max} := \max_i \beta_i$, $\theta_{\max} := \max_j \theta_j$, and $\gamma_{\min} := \min_j \gamma_j$. Taking the worst-case over all users and creators, (25) and (29) read

$$R_{\max} \leq \frac{2\sqrt{n}}{1-\alpha_{\max}} + \frac{\beta_{\max}}{1-\alpha_{\max}} S_{\max} \tag{30a}$$

and

$$S_{\max} \leq \frac{2\sqrt{n}}{1-\theta_{\max}} + \frac{1-\gamma_{\min}}{1-\theta_{\max}} R_{\max}, \tag{30b}$$

where we have used that C is (sub-)stochastic, hence $\sum_i C_{ji} S_i \leq S_{\max}$.

From the two coupled linear inequalities in (30), we finally bound

$$R_{\max} \leq \frac{2\sqrt{n}(1 + \beta_{\max} - \theta_{\max})}{(1-\alpha_{\max})(1-\theta_{\max}) - \beta_{\max}(1-\gamma_{\min})} \tag{31a}$$

and

$$S_{\max} \leq \frac{2\sqrt{n}(2-\alpha_{\max}-\gamma_{\max})}{(1-\alpha_{\max})(1-\theta_{\max})-\beta_{\max}(1-\gamma_{\min})}. \tag{31b}$$

Step II. Now, consider the separation margin as a function of the model variables:

$$\Delta_{ik}^t(u_i, c_k, c_{j_i}) := \|c_k^t - u_i^t\|^2 - \|c_{j_i}^t - u_i^t\|^2. \tag{32}$$

Preliminarily, we observe that the initial user partition $\mathcal{F}_1, ..., \mathcal{F}_M$ remains constant if and only if, for all $i \in \mathcal{U}$ and $k \in \mathcal{C} \setminus \{j_i\}$, it holds $\Delta_{ik}^t > 0$ for all $t \geq 0$. Hence, in order to prove that the user partition is constant, we prove that the latter holds true, under the assumptions of our statement. We proceed by induction. We prove that the partition is constant up to time $t-1$, then it remains constant at time t.

First, we observe that

$$\nabla_{u_i}\Delta_{ik}^t = 2(c_{j_i} - c_k), \quad \nabla_{c_k}\Delta_{ik}^t = 2(c_k - u), \quad \nabla_{c_{j_i}}\Delta_{ik}^t = 2(u - c_{j_i}). \tag{33}$$

Being $u_i, c_k, c_{j_i} \in [-1,1]^n$, then all gradients in (33) have norm at most $4\sqrt{n}$. Thus, the mean value inequality implies that $\Delta_{ik}^t(u_i, c_k, c_{j_i})$ in (32) is a Lipschitz-continuous function with Lipschitz constant $4\sqrt{n}$. Hence, we bound

$$|\Delta_{ik}^{t+1} - \Delta_{ik}^t| \leq 4\sqrt{n}(\|u_i^{t+1} - u_i^t\| + \|c_k^{t+1} - c_k^t\| + \|c_{j_i}^{t+1} - c_{j_i}^t\|). \tag{34}$$

Summing (34) over time, we bound

$$|\Delta_{ik}^t - \Delta_{ik}^0| \leq 4\sqrt{n}\left(\sum_{s=0}^{t-1}\|u_i^{s+1} - u_i^s\| + \sum_{s=0}^{t-1}\|c_k^{s+1} - c_k^s\| + \sum_{s=0}^{t-1}\|c_{j_i}^{s+1} - c_{j_i}^s\|\right), \tag{35}$$

from which we conclude that

$$\begin{aligned}\Delta_{ik}^t &\geq \Delta_{ik}^0 - 4\sqrt{n}\left(\textstyle\sum_{s=0}^{t-1}\|u_i^{s+1} - u_i^s\| + \sum_{s=0}^{t-1}\|c_k^{s+1} - c_k^s\| + \sum_{s=0}^{t-1}\|c_{j_i}^{s+1} - c_{j_i}^s\|\right) \\ &\geq \Delta_{ik}^0 - 4\sqrt{n}\Big(R_i + S_k + S_{j_i}\Big) \geq \Delta_{ik}^0 - 4\sqrt{n}\Big(R_{\max} + 2S_{\max}\Big).\end{aligned} \tag{36}$$

Finally, since the partition is constant up to time $t-1$, the bounds in (31) hold true, implying that

$$\Delta_{ik}^t \geq \Delta_{ik}^0 - 4\sqrt{n}\Big(R_{\max} + 2S_{\max}\Big) \geq \Delta_{ik}^0 - 8n\frac{5 + \beta_{\max} - \theta_{\max} - 2\alpha_{\max} - 2\gamma_{\max}}{(1-\alpha_{\max})(1-\theta_{\max}) - \beta_{\max}(1-\gamma_{\min})} > 0, \tag{37}$$

where the last inequality holds true, due to (20). Hence, the partition remains constant at time t. Hence, by induction, the partition remains constant for all $t \geq 0$.

Step III. Now, we focus on studying the dynamics of $d_i^t := \|u_i^t - c_{j_i}^t\|$. Using the notation developed in the previous steps of the proof, and using $\lambda_i = 0$, subtracting c_j^{t+1} from both sides of (21), we obtain

$$u_i^{t+1} - c_{j_i}^{t+1} = \alpha_i u_i^t + \beta_i c_{j_i}^t - c_{j_i}^{t+1}. \tag{38}$$

Then, adding and subtracting $c_{j_i}^t$ to the right-hand side, we get

$$u_i^{t+1} - c_{j_i}^{t+1} = \alpha_i(u_i^t - c_{j_i}^t) + (c_{j_i}^t - c_{j_i}^{t+1}), \tag{39}$$

which yields the recursive inequality

$$d_i^{t+1} \leq \alpha_i d_i^t + \|c_{j_i}^{t+1} - c_{j_i}^t\|. \tag{40}$$

Since the last term is vanishing (being S_{j_i} convergent) and $\alpha_i < 1$, we get $d_i^t \to 0$, from which the claim follows.

□

F.3 Proof of Lemma 1

Proof. We note that under a greedy RS for each user i, we have $r(u_i^t) = \min_j(||c_j^t - u_i^t||_2)$, and thus $u_i^t \in \mathcal{F}_j^t$. Furthermore, given $\Gamma = I_M$, we have $c_j^0 = c_j^t, \forall t$ and we simply write c_j.

Induction Hypothesis Assume for any user i, with feature vector u_i^t, governed by the system dynamics in (1), that $u_i^t, u_i^{t-1} \in \mathcal{F}_j^t \times \mathcal{F}_j^{t-1}$ with $(u_i^t - c_j) = \alpha^t(u_i^{t-1} - c_j)$, $\alpha^t \in [\eta, 1]$, where $\eta = (1-\lambda_i)a_{ii} \in [0,1]$. That is, $(u_i^t - c_j)$ and $(u_i^{t-1} - c_j)$ are parallel, point in the same direction, and $||u_i^t - c_j|| \leq ||u_i^{t-1} - c_j||$.

Induction Step Using (1) we can write

$$\begin{aligned} u_i^t &= (1-\lambda_i)(a_{ii}u_i^{t-1} + b_i c_j) + \lambda_i u_i^0 \\ u_i^{t+1} &= (1-\lambda_i)(a_{ii}u_i^{t} + b_i c_j) + \lambda_i u_i^0. \end{aligned} \tag{41}$$

Where we used the fact that matrix A is diagonal and that $u_i^t, u_i^{t-1} \in \mathcal{F}_j^t \times \mathcal{F}_j^{t-1}$. For notational simplicity, we omit the Kronecker product in (41). Using $\lambda_i u_i^0 = u_i^t - (1-\lambda_i)(a_{ii}u_i^{t-1} + b_i c_j)$, we get

$$\begin{aligned} &u_i^{t+1} = -(1-\lambda_i)a_{ii}(u_i^{t-1} - u_i^t) + u_i^t \Rightarrow \\ &(u_i^{t+1} - c_j) = (u_i^t - c_j) - (1-\lambda_i)a_{ii}(u_i^{t-1} - u_i^t). \end{aligned}$$

Intuitively, this states that the users $u_i^{t-1}, u_i^t, u_i^{t+1}$, and the creators c_j lie on straight line. Now let $\eta = (1-\lambda_i)a_{ii} \in [0,1]$, we can write

$$\begin{aligned} u_i^{t+1} - c_j &= (u_i^t - c_j) - \eta(u_i^{t-1} - u_i^t) \\ &= (u_i^t - c_j) - \eta((u_i^{t-1} - c_j) - (u_i^t - c_j)) \\ &= (1+\eta)(u_i^t - c_j) - \eta(u_i^{t-1} - c_j) \\ &= (1+\eta)(u_i^t - c_j) - \eta/\alpha^t(u_i^t - c_j) \\ &= (1+\eta-\eta/\alpha^t)(u_i^t - c_j). \end{aligned}$$

The fourth equality follows from the induction hypothesis, namely: $(u_i^t - c_j) = \alpha^t(u_i^{t-1} - c_j)$. Now we observe that

$$\alpha^t \in [\eta, 1] \Rightarrow (1+\eta-\eta/\alpha^t) \in [\eta, 1].$$

Now let $\alpha^{t+1} = (1+\eta-\eta/\alpha^t)$, which leads to the desired property: $u_i^{t+1} - c_j = \alpha^{t+1}(u_i^t - c_j)$, with $\alpha^{t+1} \in [\eta, 1]$, where $\eta = (1-\lambda_i)a_{ii}$. We can use this property to further deduce that $||u_i^{t+1} - c_j||_2 \leq ||u_i^t - c_j||_2$. Because all other creators are stubborn as well, this directly implies that under the greedy RS: $u_i^{t+1} \in F_j^{t+1}$.

Base Case Induction now follows by:

$$\begin{aligned} &u_i^0 \in \mathcal{F}_j^0 \Rightarrow \\ &u_i^1 = (1-\lambda_i)(a_{ii}u_i^0 + B_i c_j) + \lambda_i u_i^0 \\ &u_i^1 = (1-\lambda_i)(a_{ii}u_i^0 + (1-a_{ii})c_j) + \lambda_i u_i^0 \\ &u_i^1 = ((1-\lambda_i)a_{ii} + \lambda_i)u_i^0 + (1-\lambda_i)(1-a_{ii})c_j \Rightarrow \\ &u_i^1 - c_j = ((1-\lambda_i)a_{ii} + \lambda_i)(u_i^0 - c_j). \end{aligned}$$

The proof now follows by noting that: $\eta = (1-\lambda_i)a_{ii} \leq ((1-\lambda_i)a_{ii} + \lambda_i) = \alpha^1 \leq 1$. We conclude $u_i^1 - c_j = \alpha^1(u_i^0 - c_j)$, with $\alpha^1 \in [\eta, 1]$, which given stubborn users, implies $u_i^0, u_i^1 \in \mathcal{F}_j^0 \times \mathcal{F}_j^1$. Thus the user partitions $\mathcal{F}_0, .., \mathcal{F}_M$ are in fact static under the greedy RS and for any user $i \in \mathcal{F}_j$, the distance to the creator $||u_i^t - c_j||_2$, decreases monotonically with α^t. □

F.4 Proof of Corollary 1

Proof. The proof follows directly from (7). That is $c_k^t = R(u_i^t)$ with $||u_i^t - c_k^t||_2 > \min_j ||u_i^t - c_j^t||_2$, for some $j \in 0, ..., M-1$, will lead to reduced satisfaction of user i. □

G Related Work

The better position our paper, the following table provides a schematic summary of the related work, by classifying user and creators as Static (S) or Dynamic (D) and wheather they are seen as embedded in a Network (N) or seen as Isolated (I). For the Recommender System we distinguish if it is Fixed (F), namely taken from the literature, or Explicitly Designed (ED).

	Users	**Creators**	**Recommender System**
Us	D,N	D, I	ED
Lin et al. (2024)	D,I	D,I	F
Rossi et al. (2022)	D,I	N/A	ED
Lanzetti et al. (2023)	D,I	N/A	ED
Dean et al. (2024b)	D,I	D,I	ED
Ma et al. (2008)	S,N	N/A	F
Chandrasekaran et al. (2024)	D,N	N/A	ED
Ziegler et al. (2005)	S,I	N/A	ED
Cheng et al. (2017)	S, I	N/A	ED
Zhang et al. (2023)	S,I	N/A	ED
Zhang and Hurley (2008)	S,I	N/A	ED