User-Creator Feature Polarization in Recommender Systems with Dual Influence

Tao Lin[*]
Harvard University
tlin@g.harvard.edu

Kun Jin[*]
Google
kunjin@google.com

Andrew Estornell
ByteDance
andrew.estornell@bytedance.com

Xiaoying Zhang
ByteDance
zhangxiaoying.xy@bytedance.com

Yiling Chen
Harvard University
yiling@seas.harvard.edu

Yang Liu
University of California, Santa Cruz
yangliu@ucsc.edu

Abstract

Recommender systems serve the dual purpose of presenting relevant content to users and helping content creators reach their target audience. The dual nature of these systems naturally influences both users and creators: users' preferences are affected by the items they are recommended, while creators may be incentivized to alter their content to attract more users. We define a model, called user-creator feature dynamics, to capture the dual influence of recommender systems. We prove that a recommender system with dual influence is guaranteed to polarize, causing diversity loss in the system. We then investigate, both theoretically and empirically, approaches for mitigating polarization and promoting diversity in recommender systems. Unexpectedly, we find that common diversity-promoting approaches do not work in the presence of dual influence, while relevancy-optimizing methods like top-k truncation can prevent polarization and improve diversity of the system.

1 Introduction

From restaurant selection, video watching, to apartment renting, recommender systems play a pivotal role across a plethora of real-world domains. These systems match users with content they like, and help creators (those producing the content) identify their target audiences. Nevertheless, behind such success, concerns have emerged regarding possible harmful outcomes of recommender systems, in particular, *filter bubbles* [32, 5] and *polarization* [36] – outcomes with insufficient *recommendation diversity* and *creation diversity*. Recommendation diversity, meaning the diversity of the contents recommended to a user, is key to users' engagement and retention on the platform. Meanwhile, creation diversity, meaning the variety of content created on the platform, is a determinant of the platform's long-term health. In extreme cases, insufficient creation diversity can lead to consensus or polarization, where the latter can cause conflict and hatred, diminish people's mutual understanding, and cause societal crises. Therefore, from both business and social responsibility perspectives, championing and improving diversity in recommender systems is equally important as optimizing recommendation relevancy.

---

[*]The first two authors contributed equally to this work. And this work was done when Kun was at TikTok.

There is increasing emphasis in academia and industry on investigating and improving the diversity of recommender systems, combating filter bubbles and polarization. Popular diversity-boosting approaches include applying post-processing procedures such as re-ranking [11, 47] and setting diversity-aware objectives in addition to relevance maximization [38, 44, 22, 39, 12]. These methods aim to increase the recommendation diversity for users. Assuming that the contents on the platform are static, these methods have been shown to bring diversity gain to the system.

However, an important aspect that is overlooked in the aforementioned approaches is that: users and contents on a recommendation platform are not static entities – they can be *influenced* by the recommendation made by the system. In content creation platforms like YouTube, TikTok, and Twitter, recommendations naturally affect both content users and content creators. It is well known that the exposure to recommended items can shift a user's preference [24, 26, 14]. On the other hand, the creators have the incentive to change their creation styles constantly to attract their audience better (and to make more profits from the platform) [15, 20, 23]. While the effects of recommendation on either users or creators have been investigated separately, to our knowledge no previous work considers both effects. The dual influence of recommendation on users and creators causes complicated dynamics where users and creators interact and their preferences evolve together. Such evolution might exacerbate filter bubble and polarization effects. Whether the aforementioned diversity-boosting approaches still work in a dynamic environment with dual influence is questionable.

Our contributions The first contribution of our work is to define a novel, natural dynamics model that captures the dual influence of a recommender system on users and creators, which we call user-creator feature dynamics (Section 2). We leverage the users' and items'/creators' embedding vectors to represent their preferences and creation styles, and use cosine similarity to characterize the relevance of creations and users' interests (which is common in the recommender system literature and practice). This model allows us to formally reason about the impact of various design choices on the long-term diversity of a recommender system with dual influence.

Our second contribution is to demonstrate that, under realistic conditions, the user-creator feature dynamics of any recommender system with dual influence must unavoidably converge to polarization (Section 3), i.e., the preferences of users and the contents of creators will become tightly clustered into two opposite groups, significantly reducing the diversity of the system. We demonstrate that this phenomenon still occurs even after applying diversity-boosting interventions to the system.

Then, (in Section 4) we investigate some real-world designs of recommendation algorithms in order to look for techniques that mitigate polarization. Interestingly, we find that some common efficiency-improving methods, such as top-k truncation, can both prevent the system from polarization and improve the creation diversity. We also provide empirical results (Section 5) on both synthetic and real-world (MovieLens) data. As predicted by our theory, we find that systems with dual influence more easily converge to polarization under diversity-boosting designs, while efficiency-oriented and relevance-optimizing designs can in fact improve the long-term diversity of the system. This could explain why polarization does not always happen in reality. Section 6 concludes.

1.1 Related Work

Diversity in recommendations Diversity, filter bubbles, and polarization in recommendations have been important research topics in recent years, and they are closely related concepts with different focuses. On the one hand, filter bubbles are frequently defined as decreasing recommendation diversity over time [5], which describes both the process and the outcome of insufficiently diverse recommendations. On the other hand, polarization describes the negative outcome of insufficient mutual understanding between people [36]. In content platforms, an example of polarization is people creating content with strong agreement or disagreement with other content under the same topic, e.g., political opinions. To combat these negative outcomes, previous works propose diversity-boosting approaches including re-ranking [11, 47] and diversity-aware objective optimization [38, 44, 22, 39, 12, 45]. Despite having positive effects in situations where user preferences and creation styles are fixed, these approaches overlooked the dynamic nature of recommender systems and our work shows that certain approaches will make long-term outcomes worse under the dual influence.

Opinion dynamics Opinion dynamics study the effect of people exchanging opinions with others on social networks [37, 17, 29, 4]. Our model of a recommender system with dual influence on users

and creators resembles a bipartite social network, and our conclusion that the system converges to polarization is conceptually similar to people reaching consensus on social networks [1, 10, 31, 46]. However, the technique we use to prove our conclusion (absorbing Markov chain) significantly differs from the main technique (stability of ODE) in the mentioned works.

Performative effects of recommender systems The phenomenon that predictive systems like recommender systems can impact the individuals interacting with those systems (e.g., users and creators) is related to the literature of performative prediction [34, 18]. These impacts can be direct, such as individuals ostensibly modifying their features in order to obtain more desirable outcomes [27]. Prior works on the performative effects of recommender systems (e.g., [7, 24, 14, 41, 15, 42, 35, 20, 3, 43, 2, 23]) only consider one-sided impact, either on users or on creators. Differing from them, our work studies two-sided impacts, i.e., on both users and creators. We provide a table to compare our work with previous works in Appendix A.

2 Model: User-Creator Feature Dynamics

We define a *dynamics* model for user preferences and content/creator features in a recommender system. Let $\boldsymbol{U}^t = [\boldsymbol{u}_j^t]_{j=1}^m = [\boldsymbol{u}_1^t, \ldots, \boldsymbol{u}_m^t] \in \mathbb{R}^{d \times m}$ be a population of m users and $\boldsymbol{V}^t = [\boldsymbol{v}_i^t]_{i=1}^n = [\boldsymbol{v}_1^t, \ldots, \boldsymbol{v}_n^t] \in \mathbb{R}^{d \times n}$ be a population of n creators at time t, where each vector $\boldsymbol{u}_j^t, \boldsymbol{v}_i^t \in \mathbb{S}^{d-1}$ represent the preference/feature vector of each user and creator respectively, assumed to be on the unit sphere $\mathbb{S}^{d-1}$ with ℓ_2-norm. Then $(\boldsymbol{U}^t, \boldsymbol{V}^t)$ denotes the state of the dynamics at time t. The dynamics evolve as follows at each time step $t \geq 0$:

1) Recommendation: Each user $j \in [m]$ is recommended a creator, where creator $i \in [n]$ is chosen with a probability

$$p_{ij}^t = p_{ij}^t(\boldsymbol{U}^t, \boldsymbol{V}^t). \tag{1}$$

While we allow a wide array of different functions $p_{ij}^t(\cdot)$, a common example of such functions is the so-called *softmax function*:

$$p_{ij}^t = \text{softmax}(\boldsymbol{u}_j^t, \boldsymbol{V}^t; \beta) = \frac{\exp(\beta \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle)}{\sum_{i=1}^n \exp(\beta \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle)}. \tag{2}$$

A larger β means that the recommendation is more sensitive to the *relevance* of a creator to a user, measured by $\langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle$.

2) User update: After recommendation, each user $j \in [m]$ updates their feature vector $\boldsymbol{u}_j^t$, based on which creator, say i_j^t, was recommended to them:

$$\boldsymbol{u}_j^{t+1} = \mathcal{P}\big(\boldsymbol{u}_j^t + \eta_u f(\boldsymbol{v}_{i_j^t}^t, \boldsymbol{u}_j^t) \boldsymbol{v}_{i_j^t}^t\big). \tag{3}$$

Here, $\eta_u \in [0, 1]$ is a parameter controlling the rate of update, $f(\boldsymbol{v}_i, \boldsymbol{u}_j)$ is a function that quantifies the impact of creator i's content on user j (discussed in detail later), and $\mathcal{P}(\boldsymbol{x}) = \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}$ is the projection back onto the unit sphere. Our user update model generalizes [14], which considers $\boldsymbol{u}_j^{t+1} = \mathcal{P}(\boldsymbol{u}_j^t + \eta_u \langle \boldsymbol{v}_{i_j^t}^t, \boldsymbol{u}_j^t \rangle \boldsymbol{v}_{i_j^t}^t)$, by replacing the inner product with a general function f.

3) Creator update: Creators also update their feature vectors based on which users are recommended their content. For each creator $i \in [n]$, let $J_i^t = \{j : i_j^t = i\}$ be the set of users being recommended creator i, then $\boldsymbol{v}_i^t$ is updated by:

$$\boldsymbol{v}_i^{t+1} = \mathcal{P}\Big(\boldsymbol{v}_i^t + \frac{\eta_c}{|J_i^t|} \sum_{j \in J_i^t} g(\boldsymbol{u}_j^t, \boldsymbol{v}_i^t) \boldsymbol{u}_j^t\Big), \tag{4}$$

where $\eta_c \in [0, 1]$ is a parameter controlling the rate of update, and $g(\boldsymbol{u}_j, \boldsymbol{v}_i)$ is a function that quantifies the impact of user j on creator i.

Impact functions f and g Our results apply to any impact functions f and g that satisfy the following natural assumptions. First, $f(\boldsymbol{v}_i, \boldsymbol{u}_j)$ and the inner product $\langle \boldsymbol{v}_i, \boldsymbol{u}_j \rangle$ have the same sign:

$f(\boldsymbol{v}_i, \boldsymbol{u}_j)$ is $\begin{cases} > 0 & \text{if } \langle \boldsymbol{v}_i, \boldsymbol{u}_j \rangle > 0 \\ < 0 & \text{if } \langle \boldsymbol{v}_i, \boldsymbol{u}_j \rangle < 0 \\ = 0 & \text{if } \langle \boldsymbol{v}_i, \boldsymbol{u}_j \rangle = 0. \end{cases}$ This means that if a user *likes* the content ($\langle \boldsymbol{v}_i^t, \boldsymbol{u}_j^t \rangle > 0$), then the

user vector $\boldsymbol{u}_j^t$ will be updated *towards* the direction of the creator vector $\boldsymbol{v}_j^t$. If the user *dislikes* the content ($\langle \boldsymbol{v}_i^t, \boldsymbol{u}_j^t \rangle < 0$), then the user vector $\boldsymbol{u}_j^t$ will move *away from* $\boldsymbol{v}_j^t$. Such "biased assimilation" user behavior is well documented in the literature [14]. Further, we assume upper and lower bounds on $|f|$:

$$|f(\boldsymbol{v}_i, \boldsymbol{u}_j)| \leq 1, \qquad |f(\boldsymbol{v}_i, \boldsymbol{u}_j)| \geq L_f > 0 \text{ whenever } \langle \boldsymbol{v}_i, \boldsymbol{u}_j \rangle \neq 0.$$

The lower bound $|f(\boldsymbol{v}_i, \boldsymbol{u}_j)| \geq L_f$ means that the exposure to an item that a user likes or dislikes always has some non-negligible impact on the user's preference. For example, $f(\boldsymbol{v}_i, \boldsymbol{u}_j) = \text{sign}(\langle \boldsymbol{v}_i, \boldsymbol{u}_j \rangle)a + b\langle \boldsymbol{v}_i, \boldsymbol{u}_j \rangle$ satisfies both assumptions when $L_f = a > 0$ and $b \geq 0$.

For g, likewise assume that its sign is the same as $\langle \boldsymbol{u}_j, \boldsymbol{v}_i \rangle$: $g(\boldsymbol{u}_j, \boldsymbol{v}_i)$ is $\begin{cases} > 0 & \text{if } \langle \boldsymbol{u}_j, \boldsymbol{v}_i \rangle > 0 \\ < 0 & \text{if } \langle \boldsymbol{u}_j, \boldsymbol{v}_i \rangle < 0 \\ = 0 & \text{if } \langle \boldsymbol{u}_j, \boldsymbol{v}_i \rangle = 0. \end{cases}$ Intuitively, this captures the incentive of a creator who aims to maximize the average ratings from users who are recommended their items. On video platforms for example, if the creators are rewarded based on the average rating of their videos, they will try to reinforce their creation styles based on the users who give positive feedback ($\langle \boldsymbol{u}_j, \boldsymbol{v}_i \rangle > 0$) so that their creations are more likely to be recommended to those users. Meanwhile, the creators will also change their creation styles based on negative feedback ($\langle \boldsymbol{u}_j, \boldsymbol{v}_i \rangle < 0$), but in the opposite direction of the negative-feedback users' interests, so that their creations are less likely to be recommended to those users. Taking both scenarios into account, the creator moves towards the weighted average of user preferences $\sum_{j \in J_i^t} g(\boldsymbol{u}_j^t, \boldsymbol{v}_i^t)\boldsymbol{u}_j^t$, which is captured by our update rule (4). A particular example of g is the sign function $g(\boldsymbol{u}_j, \boldsymbol{v}_i) = \text{sign}(\langle \boldsymbol{u}_j, \boldsymbol{v}_i \rangle) \in \{-1, 0, 1\}$. We will only consider the sign function g in order to simplify the theoretical presentation. We believe that all our results can be generalized to other g functions satisfying similar conditions as f; the details are left as future work.

3 Unavoidable Polarization

Having defined the user-creator feature dynamics in a recommender system with dual influence, we now theoretically study how such dynamics evolve. Our main result is: if every creator can be recommended to every user with some non-zero probability, then the dynamics must eventually *polarize*.

Definition 3.1 (consensus and bi-polarization)**.** *Let $R > 0$. The dynamics $(\boldsymbol{U}^t, \boldsymbol{V}^t)$ is said to reach:*

- *R-consensus if there exists a vector $\boldsymbol{c} \in \mathbb{R}^d$ such that every feature vector is R-close to $\boldsymbol{c}$: $\forall \boldsymbol{u}_j^t, \|\boldsymbol{u}_j^t - \boldsymbol{c}\|_2 \leq R$ and $\forall \boldsymbol{v}_i^t, \|\boldsymbol{v}_i^t - \boldsymbol{c}\|_2 \leq R$.*

- *R-bi-polarization if there exists a vector $\boldsymbol{c} \in \mathbb{R}^d$ such that every feature vector is R-close to $+\boldsymbol{c}$ or $-\boldsymbol{c}$: $\forall \boldsymbol{u}_j^t, \|\boldsymbol{u}_j^t - \boldsymbol{c}\|_2 \leq R$ or $\|\boldsymbol{u}_j^t + \boldsymbol{c}\|_2 \leq R$, and $\forall \boldsymbol{v}_i^t, \|\boldsymbol{v}_i^t - \boldsymbol{c}\|_2 \leq R$ or $\|\boldsymbol{v}_i^t + \boldsymbol{c}\|_2 \leq R$.*

The dynamics is said to reach $(R, \boldsymbol{c})$-consensus (or $(R, \boldsymbol{c})$-bi-polarization) if the dynamics reaches R-consensus (or R-bi-polarization) with the vector $\boldsymbol{c}$.

Consensus is any state where all users and creators have similar feature vectors (with maximum difference R), implying that they have similar interests or preferences. Bi-polarization is any state where all users and creators are clustered into two groups with exactly opposite features (e.g., Republicans vs Democrats). Mathematically, consensus is a special case of bi-polarization.

Proposition 3.2. *Bi-polarization states are absorbing: once the dynamics reaches $(R, \boldsymbol{c})$-bi-polarization with some $R \in [0, 1]$ and $\boldsymbol{c} \in \mathbb{S}^{d-1}$, it will satisfy $(R, \boldsymbol{c})$-bi-polarization forever. The same holds for consensus.*

A natural property of a recommender system is that every creator can be recommended to every user with some non-zero probability: $p_{ij}^t \geq p_0 > 0$ with some constant p_0. This is satisfied by the softmax function, which is a rough model of real-world recommendation algorithms [13, 26]: $p_{ij}^t = \frac{\exp(\beta\langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle)}{\sum_{i=1}^n \exp(\beta\langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle)} \geq \frac{\exp(-\beta)}{n \exp(\beta)} = p_0 > 0$. Moreover, many large-scale real-world recomendation systems (e.g., Yahoo! [28] and Kuaishou [16]) intentionally insert small random traffic attempting to improve recommendation diversity or explore users' interests [25, 40], which will cause all recommendation probabilities to be non-zero. We show in Theorem 3.3 that, however, a recommender system satisfying $p_{ij}^t \geq p_0 > 0$ must converge to polarization, under some additional conditions on the users' and creators' update rates:

Theorem 3.3. *Suppose $g(\boldsymbol{u}_j, \boldsymbol{v}_i) = \mathrm{sign}(\langle \boldsymbol{u}_j, \boldsymbol{v}_i \rangle)$, the update rates $\eta_c \leq \frac{\eta_u L_f}{2}$ and $\eta_u < \frac{1}{2}$, and the recommendation probability $p_{ij}^t \geq p_0 > 0, \forall i, j, t$. Then, from almost all initial states, the dynamics $(\boldsymbol{U}^t, \boldsymbol{V}^t)$ will eventually reach R-consensus or R-bi-polarization for any $R > 0$.*

In other words, if the users' and creators' updates are not too fast and all recommendation probabilities are non-zero, then all users and creators will eventually converge to at most two clusters (regardless of the feature dimension d). Since creators in one cluster produce similar contents, users in such a polarized system can never receive diverse recommendations. This means that the naïve attempt of imposing $p_{ij}^t \geq p_0 > 0$ cannot improve the diversity of a recommender system with dual influence. The conditions on the update rates η_u, η_c are only assumed to simplify the proof of Theorem 3.3. Our experiments (in Section 5) will show that polarization still occurs even without those conditions.

Theorem 3.3 does not characterize the rate of convergence of the user-creator feature dynamics to polarization, which we leave as an open question.

The proof of Theorem 3.3 is an absorbing Markov chain argument. It uses the following lemma:

Lemma 3.4. *Suppose $\eta_c \leq \frac{\eta_u L_f}{2}$ and $\eta_u < \frac{1}{2}$. For any $R > 0$, for almost every state $(\boldsymbol{U}^t, \boldsymbol{V}^t)$ in the state space, there exists a path $(\boldsymbol{U}^t, \boldsymbol{V}^t) \to (\boldsymbol{U}^{t+1}, \boldsymbol{V}^{t+1}) \to \cdots \to (\boldsymbol{U}^{t+T}, \boldsymbol{V}^{t+T})$ of finite length that leads to an R-bi-polarization state $(\boldsymbol{U}^{t+T}, \boldsymbol{V}^{t+T})$.*

The proof of this lemma (in Appendix F) is involved. It uses induction on the number of creators n. The base case of $n = 1$ is proved by a potential function argument. For $n \geq 2$, we first construct a path that leads the *subsystem* of $n - 1$ creators and all users to R-bi-polarization. Then, depending on where the remaining creator is, we construct a sequence of recommendations that leads the remaining creator to one of the two clusters formed by the $n - 1$ creators and all users. Such recommendations will move some users out of the formed clusters, which requires extra care in the proof.

Proof of Theorem 3.3. For any state $(\boldsymbol{U}^t, \boldsymbol{V}^t)$ in the state space, by Lemma 3.4 there exists a path $(\boldsymbol{U}^t, \boldsymbol{V}^t) \to \cdots \to (\boldsymbol{U}^{t+T}, \boldsymbol{V}^{t+T})$ of length T that leads to R-bi-polarization. Because every creator can be recommended to a user with probability at least p_0, each transition $(\boldsymbol{U}^{t'}, \boldsymbol{V}^{t'}) \to (\boldsymbol{U}^{t'+1}, \boldsymbol{V}^{t'+1})$ happens with probability at least p_0^m. So, the path of length T has probability at least $p_0^{mT} > 0$, and the probability that the dynamics *does not* reach R-bi-polarization after KT steps is at most $(1 - p_0^{mT})^K$, which $\to 0$ as $K \to \infty$. Therefore, with probability 1 the dynamics will reach R-bi-polarization eventually. $\qquad\square$

4 Discussions on Real-World Designs

Next, we discuss how 4 types of real-world recommender system designs affect the user-creator feature dynamics: top-k truncation, threshold truncation, diversity-boosting, and uniform traffic.

(1) Top-k Truncation A prevalent practice in modern two-stage recommendation algorithms on large-scale platforms, such as YouTube [13], is to first filter out items that are unlikely to be relevant to a user, then make recommendations from the remaining items. In particular, we consider the top-k truncation policy: for every user j, find the k most relevant creators, namely, the k creators whose inner products with the user $\langle \boldsymbol{v}_i^t, \boldsymbol{u}_j^t \rangle$ are largest (equivalently, the k creators whose probabilities p_{ij}^t of being recommended to user j are highest), then recommend one of those k creators to user j with probability proportional to p_{ij}^t. The other creators will not be recommended. This practice significantly reduces the computation cost and improves the relevancy of recommendations. Interestingly, we show that such a practice also has the potential to improve the long-term diversity of a recommender system with dual influence.

Definition 4.1 (clusters). *We say a state $(\boldsymbol{U}^t, \boldsymbol{V}^t)$ forms q clusters if there exist $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_q \in \mathbb{R}^d$ and a small number $R > 0$ such that every feature vector is in the ℓ_2 ball of some $\boldsymbol{c}_i$ with radius R (denoted by $B(\boldsymbol{c}_\ell, R) = \{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{c}_\ell\|_2 \leq R\}$), and $B(\boldsymbol{c}_\ell, 2R) \cap B(\boldsymbol{c}_{\ell'}, 2R) = \emptyset$ for $\ell \neq \ell'$.*

It is clear that consensus has a single cluster, and bi-polarization has two.

Proposition 4.2. *With top-k truncation, there exist states $(\boldsymbol{U}^t, \boldsymbol{V}^t)$ that form $\lfloor n/k \rfloor$ clusters and are absorbing (i.e., once the system forms $\lfloor n/k \rfloor$ clusters, it forms $\lfloor n/k \rfloor$ clusters forever).*

This result is in contrast with Theorem 3.3 which shows that a recommender system where every creator can be recommended to every user ($p_{ij}^t > 0$) is doomed to polarize. With top-k truncation where some $p_{ij}^t = 0$, polarization can be avoided. Experiments in Section 5.3 support our prediction that top-k truncation can reduce polarization and improve diversity.

(2) Threshold Truncation   Besides top-k truncation, threshold truncation is another way to filter out irrelevant creators: set a threshold $\tau \in [-1, 1]$ such that any user-creator pair with inner product $\langle \boldsymbol{u}_i, \boldsymbol{v}_j \rangle < \tau$ is not recommended. A natural choice is $\tau = 0$, meaning that users will not receive recommendations predicted to be "disliked" by them. Increasing τ is similar to increasing the β in the softmax function, which improves recommendation relevance.

Proposition 4.3. *In d-dimensional feature space, if user-creator pairs with $\langle \boldsymbol{u}_i, \boldsymbol{v}_j \rangle < 0$ are not recommended, then there exist stable states with $d + 1$ clusters.*

Although truncation at $\tau = 0$ allows stable states with $d + 1$ clusters to exist, the dynamics does not necessarily converge to such states; it can still end up with stable states with fewer clusters. In fact, experiments (in Appendix B.2) show that truncation at $\tau = 0$ is *not good* for diversity and causes severe polarization, while truncation at a large threshold like $\tau = 0.707$ is better at reducing polarization.

(3) Diversity Boosting   Diversity boosting aims to explore users' interests and improve users' experience by diversifying recommendation. For example, when making recommendations, the model optimizes the objective:

$$h_{rel}(\langle \boldsymbol{u}_i, \boldsymbol{v}_j \rangle) + \rho h_{div}(list_i, \boldsymbol{v}_j), \tag{5}$$

where h_{rel}, h_{div} rewards the recommendation relevance and diversity respectively and $list_i$ records the recent list of recommended items to user i. h_{div} can take a simple form of $\sum_{j' \in list_i} 1 - \langle \boldsymbol{v}_{j'}, \boldsymbol{v}_j \rangle$, and $\rho > 0$ controls the strength of diversity-boosting. Despite being successful when users' preferences and items are fixed, this design alone cannot prevent bi-polarization in our dual-influence dynamics, since the conditions in Theorem 3.3 are still satisfied and the users' and creators' update rules remain the same. Experiments in Section 5.2 support our claim.

(4) Uniform Traffic   Adding a small fraction of uniform traffic to the personalized recommendations is another method proposed in previous works to improve recommendation diversity or to explore user preferences [25, 16, 9, 8, 30]. This method gives a non-zero lower bound on the probability of every creator being recommended to every user. So, as a corollary of our Theorem 3.3, it causes a recommender system with dual influence to polarize. Such an observation is striking as it demonstrates that optimizing for recommendation diversity in a static setting can ultimately lead to a huge loss of the system diversity in the long run.

5   Experiments

We present experimental results on the behavior of user-creator feature dynamics on synthetic data and real-world (MovieLens 20M) data and the effect of top-k truncation and threshold truncation on the dynamics.

5.1   Synthetic Data Experiments

Setup   The dynamics is initialized by randomly generating user and creator features on the unit sphere in $\mathbb{R}^d$. We pick $d = 10$, number of creators $n = 50$, number of users $m = 100$. We use the softmax recommendation probability function (2). We simulate the dynamics for $T = 1000$ steps, repeated 100 times each with a new initialization. We choose the sign impact function $g(\boldsymbol{u}_j, \boldsymbol{v}_i) = \text{sign}(\langle \boldsymbol{u}_j, \boldsymbol{v}_i \rangle)$ for creator updates. For user updates, we choose inner product $f(\boldsymbol{v}_i, \boldsymbol{u}_j) = \langle \boldsymbol{v}_i, \boldsymbol{u}_j \rangle$. The inner product function is studied in previous works on users' preference dynamics (but not creators') [14]. Note that the inner product does not satisfy the condition $|f(\boldsymbol{v}_i, \boldsymbol{u}_j)| \geq L_f$ needed in Theorem 3.3. However, we still observe convergence to polarization in nearly all experiments. Thus, even when this condition does not hold, users and creators still tend towards polarization in practice.

Three key parameters in our model are β (sensitivity of the softmax function), η_c (creator update rate), and η_u (user update rate). We set them to $\beta = 1, \eta_c = \eta_u = 0.1$, and change one parameter at
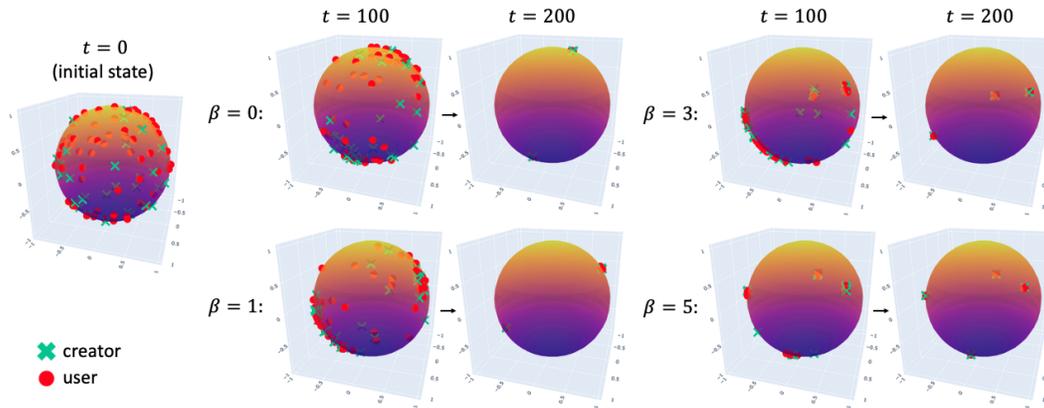
Figure 1: Snapshots of the dynamics simulated with the same initialization but different recommendation sensitivity β. A larger β resulted in more clusters at time step $t = 200$.

a time to see its effect on the dynamics. We also test what happens when some dimensions of the user features are *fixed* features that are not updated.

Measures To quantify the behavior of the dynamics, given user and creator feature vectors $(\boldsymbol{U}, \boldsymbol{V})$ we compute the following measures, which cover diversity, relevancy, and polarization of the system:

- *Creator Diversity* (CD): diversity of the creator features, measured by their average pairwise distance [47, 33]: $\mathrm{CD}(\boldsymbol{V}) = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j \neq i} \|\boldsymbol{v}_i - \boldsymbol{v}_j\|$.

- *Recommendation Diversity* (RD): diversity of the contents recommended to a user, measured by the weighted variance of the contents: $\mathrm{RD}(\boldsymbol{U}, \boldsymbol{V}; \beta) = \frac{1}{m} \sum_{j=1}^{m} \sum_{i=1}^{n} p_{ij} \|\boldsymbol{v}_i - \overline{\boldsymbol{v}}_j\|^2$, where $\overline{\boldsymbol{v}}_j = \sum_{i=1}^{n} p_{ij} \boldsymbol{v}_i$ and $p_{ij} = \frac{\exp(\beta \langle \boldsymbol{u}_j, \boldsymbol{v}_i \rangle)}{\sum_{i=1}^{n} \exp(\beta \langle \boldsymbol{u}_j, \boldsymbol{v}_i \rangle)}$.

- *Recommendation Relevance* (RR): relevance of the contents recommended to a user, measured by the weighted average of inner products: $\mathrm{RR}(\boldsymbol{U}, \boldsymbol{V}; \beta) = \frac{1}{m} \sum_{j=1}^{m} \sum_{i=1}^{n} p_{ij} \langle \boldsymbol{u}_j, \boldsymbol{v}_i \rangle$.

- *Tendency to Polarization* (TP): This is a novel measure we propose to quantify how close the system is to consensus or bi-polarization, measured by the average absolute inner products between the creators: $\mathrm{TP}(\boldsymbol{V}) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{k=1}^{n} |\langle \boldsymbol{v}_i, \boldsymbol{v}_k \rangle|$. $\mathrm{TP}(\boldsymbol{V})$ being closer to 1 means that the system is more polarized, because the term $|\langle \boldsymbol{v}_i, \boldsymbol{v}_k \rangle|$ is 1 iff the two vectors $\boldsymbol{v}_i, \boldsymbol{v}_k$ are equal or opposite to each other.

It is worth noting that a high creator diversity is necessary for simultaneously achieving high recommendation relevance and high recommendation diversity. For example, they cannot be simultaneously achieved in a polarized state.

Sensitivity Parameter β A larger β means that a user will be recommended more relevant content/creator with a higher probability. $\beta = 0$, on the other hand, means that the user receives uniform recommendations across all creators. **Our main observation** from the experiments is: *a larger β leads to higher creator diversity and alleviated polarization in the long run.*

Figure 1 shows snapshots of the dynamics at different time steps under different β values. Here, we choose dimension $d = 3$ instead of 10 so the feature vectors can be visualized on a 3d sphere. We see that the system tends to form more clusters at time $t = 200$ as β increases.

Figure 2 shows the changes of the 4 measures CD, RD, RR, TP over time under different β values. We see that a more diverse recommendation policy (a smaller β) leads to lower creator diversity and a higher level of polarization in the long run. In particular, while Creator Diversity reaches a similar level under different β in the end, it *drops at a slower rate* with a *larger β* (see $\beta = 5, 6$). Moreover, from the plot of Tendency to Polarization, we see that a larger β *alleviates* polarization, which means improvement in the diversity of the whole system.

An explanation for our observation is the following: When β is smaller, each user receives more uniform recommendations across all creators. So, for different creators, the sets of users recommended
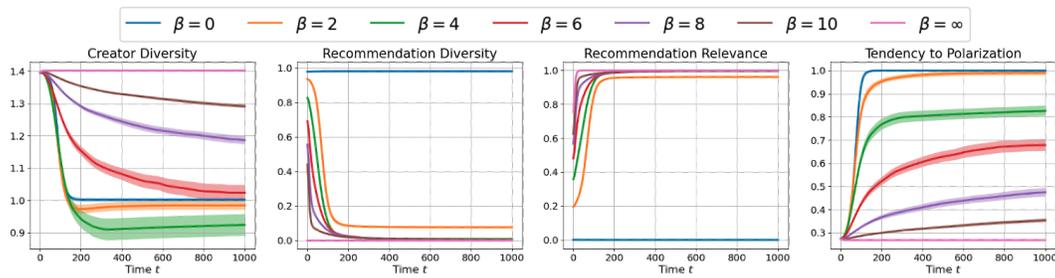
Figure 2: Changes of measures over time under different sensitivity parameter β, on synthetic data. $\beta = 0$ means uniform (non-personalized) recommendation. $\beta = \infty$ means hard-max recommendation: only recommend the single most relevant creator to a user. Larger β reduces the tendency to polarization.

to those creators have larger intersections. Since the creator updates are based on the sets of recommended users, different creators will be moving towards more similar directions. This leads to faster polarization. One can also predict this observation from Theorem 3.3: when β is large, the minimum recommendation probability p_0 of the softmax function tends to 0, so it might take a long time for the system to converge to polarization, while with a small β the system polarizes quickly.

Update Rates η_c and η_u   A larger η_c means that creator features are updated faster, and intuitively should lead to faster polarization. This is validated in experiments: Figure 3 shows that a larger η_c indeed causes more extreme polarization and lower diversity (both CD and RD). A larger η_u means that user features are updated faster. It has a similar effect of exacerbating polarization as η_c does, as shown in Figure 4.
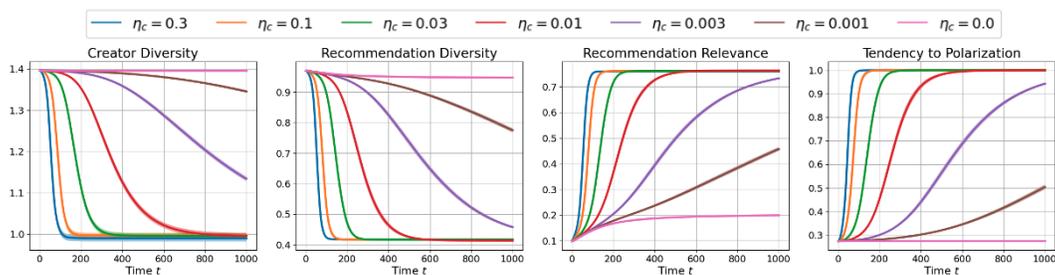

Figure 3: Changes of measures over time under different creator update rate η_c, on synthetic data
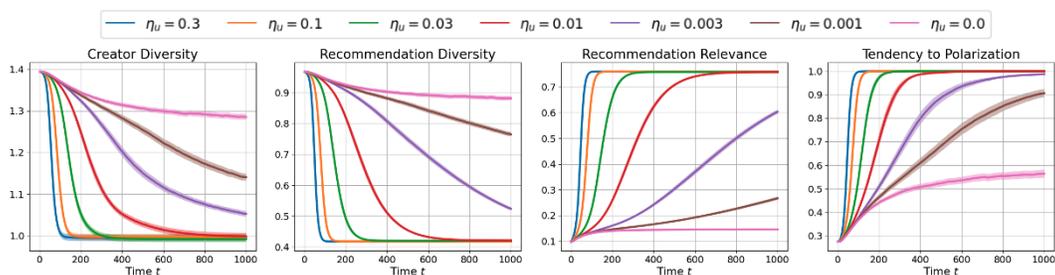

Figure 4: Changes of measures over time under different user update rate η_u, on synthetic data

Number of Fixed Dimensions   We also consider the scenario where some dimensions of the user feature vectors are fixed features and thus not updated from round to round (e.g., age, gender), which is a realistic scenario. Detailed results are in Appendix B.1. The **main observation** is: *as the number of fixed dimensions increases, the diversity of the system improves and the degree of polarization is reduced.* This is similar to the effect of decreasing user update rate η_u. The observation that

fixed dimensions of user features help to improve diversity might be a reason why the recommender systems in practice are not as polarized as our theoretical prediction.

5.2 Real-World Data Experiments

In this part, we conduct experiments on the MovieLens 20M dataset [19]. We use a real-world two-tower recommendation model with 16-dimensional tower tops as the user and creator embeddings (Figure 9). The model is initialized by fitting a two-tower model [21] on the existing MovieLens rating data and using the tower tops as the initial user and creator embeddings. Then we follow Algorithm 1 to simulate the dynamics.

Figure 5 shows the effect of the recommendation sensitivity parameter β on the system. Similar to the synthetic data experiments, a smaller β (more diverse recommendation for the users in the short term) results in faster polarization. We note that the joint results on CD and TP are more informative than each one alone: despite $\beta = 0$ has a higher creator diversity than $\beta = 2$ at $T = 500$, the system reaches polarization more quickly under $\beta = 0$. The higher creator diversity under $\beta = 0$ is because the two clusters in the bi-polarized state are more balanced so the average pairwise distance between the creators is higher under $\beta = 0$ than under $\beta = 2$.

Figure 6 shows the effect of using diversity-aware objective (Eq. 5) for diversity boosting. We see that myopically promoting the short-term recommendation diversity (using a larger ρ) results a higher creation diversity but also a higher tendency to polarization. A possible explanation for this phenomenon is, similar to the case with β, the system polarizes into two balanced clusters which actually have a large average pairwise distance. In this case, Tendency to Polarization is a better measure for diversity loss than Creator Diversity (average pairwise distance).
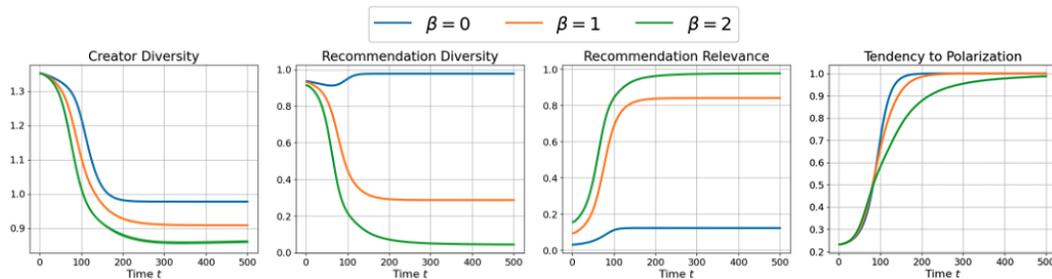

Figure 5: Experiment on MovieLens 20M dataset under different recommendation sensitivity β
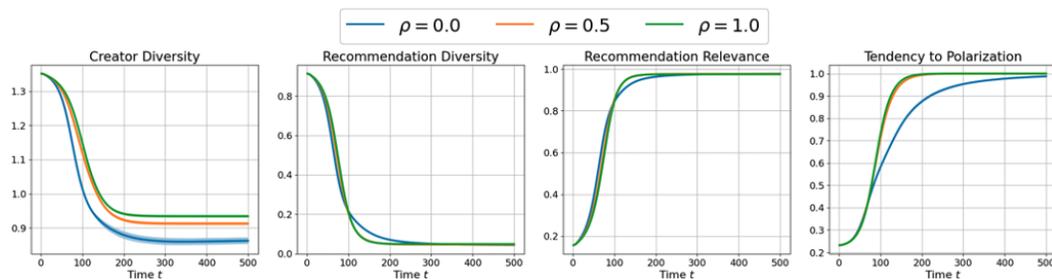
Figure 6: Experiment on MovieLens 20M dataset with diversity-aware objective under different ρ

5.3 Top-k Truncation and Threshold Truncation

We experimented with top-k truncation on the synthetic data (Table 1) and the MovieLens dataset (Appendix C). Our **main observation** is: *a small k improves the diversity of the recommender system and reduces polarization*. This is consistent with our theoretical prediction (Proposition 4.2). However, there is a tradeoff between the diversity of recommendations to users (RD) and the diversity of creations in the system (CD and TP). A top-k truncation policy with small k is "not diverse" for users because it exposes a user only to a small set of contents. However, such a policy can lead to a more diverse outcome in the whole system. This tradeoff is worth further studying.

Table 1: Diversity improvement by top-k truncation on synthetic data

β	k	Creator Diversity	Recommendation Diversity	Recommendation Relevance	Tendency to Polarization
1	50	$1.00_{\pm.03}$	$\mathbf{0.42_{\pm 0.01}}$	$0.76_{\pm 0.01}$	$1.00_{\pm 10^{-3}}$
	25	$0.52_{\pm.32}$	$0.03_{\pm 0.03}$	$0.97_{\pm 0.02}$	$0.91_{\pm 0.13}$
	20	$0.91_{\pm.15}$	$0.00_{\pm 0.01}$	$1.00_{\pm 0.01}$	$0.68_{\pm 0.12}$
	10	$1.17_{\pm.06}$	$0.00_{\pm 10^{-3}}$	$1.00_{\pm 10^{-3}}$	$0.50_{\pm 0.07}$
	5	$1.31_{\pm.02}$	$0.00_{\pm 10^{-3}}$	$1.00_{\pm 10^{-3}}$	$0.35_{\pm 0.03}$
	1	$\mathbf{1.40_{\pm 10^{-3}}}$	$0.00_{\pm 10^{-3}}$	$\mathbf{1.00_{\pm 10^{-3}}}$	$\mathbf{0.27_{\pm 10^{-3}}}$
3	50	$0.95_{\pm.14}$	$\mathbf{0.02_{\pm 0.02}}$	$0.99_{\pm 0.01}$	$0.91_{\pm 0.10}$
	25	$0.80_{\pm.24}$	$0.00_{\pm 0.01}$	$1.00_{\pm 10^{-3}}$	$0.77_{\pm 0.13}$
	20	$0.89_{\pm.13}$	$0.00_{\pm 10^{-3}}$	$1.00_{\pm 10^{-3}}$	$0.74_{\pm 0.11}$
	10	$1.18_{\pm.05}$	$0.00_{\pm 10^{-3}}$	$1.00_{\pm 10^{-3}}$	$0.49_{\pm 0.07}$
	5	$1.31_{\pm.02}$	$0.00_{\pm 10^{-3}}$	$1.00_{\pm 10^{-3}}$	$0.34_{\pm 0.03}$
	1	$\mathbf{1.40_{\pm 10^{-3}}}$	$0.00_{\pm 10^{-3}}$	$\mathbf{1.00_{\pm 10^{-3}}}$	$\mathbf{0.27_{\pm 10^{-3}}}$

We also experimented with threshold truncation on synthetic data (Appendix B.2) and MovieLens data (Appendix C). The effect of a large truncation threshold τ is similar to the effect of a small k in top-k truncation.

6 Conclusion

Our work defines a dynamics model to capture the dual influence of recommender systems on user preferences and content creation. Although our model is a theoretical abstraction, we believe that it captures the essence of a real-world recommender system, and our effort is an important initial endeavor to study diversity in recommender systems with dual influence. (See Appendix H for some additional discussions on real-world recommender systems.) We specifically point out different concepts of diversity in recommender systems (creation diversity, recommendation diversity, and tendency to polarization) and provide theoretical and empirical evidences to show that, due to dual influence, myopically optimizing recommendation diversity might hurt the long-term creation diversity and result in polarization of the system. We also explore popular design choices in recommender systems and show an interesting and somewhat counter-intuitive result that designs purely targeting efficiency improvement (e.g., top-k truncation) can alleviate polarization. We believe that the insights from our work are valuable to building healthy and sustainable recommender systems, and our results can inspire more sophisticated solutions for improving the long-term diversity of recommender systems to be developed.

References

[1] Daron Acemoğlu, Giacomo Como, Fabio Fagnani, and Asuman Ozdaglar. Opinion fluctuations and disagreement in social networks. *Mathematics of Operations Research*, 38(1):1–27, 2013. ISSN 0364765X, 15265471. URL http://www.jstor.org/stable/23358646.

[2] Krishna Acharya, Varun Vangala, Jingyan Wang, and Juba Ziani. Producers equilibria and dynamics in engagement-driven recommender systems. *arXiv preprint arXiv:2401.16641*, 2024.

[3] Arpit Agarwal and William Brown. Online recommendations for agents with discounted adaptive preferences. *arXiv preprint arXiv:2302.06014*, 2023.

[4] Claudio Altafini and Gabriele Lini. Predictable dynamics of opinion forming for networks with antagonistic interactions. *IEEE Transactions on Automatic Control*, 60(2):342–357, 2015. doi: 10.1109/TAC.2014.2343371.

[5] Guy Aridor, Duarte Goncalves, and Shan Sikdar. Deconstructing the filter bubble: User decision-making recommender systems. In *Proceedings of the 14th ACM Conference on Recommender Systems*, RecSys '20, page 82–91, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450375832. doi: 10.1145/3383313.3412246. URL https://doi.org/10.1145/3383313.3412246.

[6] Nikhil Bansal and Anupam Gupta. Potential-Function Proofs for Gradient Methods. *Theory of Computing*, 15(4):1–32, 2019. doi: 10.4086/toc.2019.v015a004. URL `https://theoryofcomputing.org/articles/v015a004`. Publisher: Theory of Computing.

[7] Omer Ben-Porat and Moshe Tennenholtz. A game-theoretic approach to recommendation systems with strategic content providers. *Advances in Neural Information Processing Systems*, 31, 2018.

[8] Stephen Bonner and Flavian Vasile. Causal embeddings for recommendation. In *Proceedings of the 12th ACM Conference on Recommender Systems*, pages 104–112, Vancouver British Columbia Canada, September 2018. ACM. ISBN 978-1-4503-5901-6. doi: 10.1145/3240323.3240360. URL `https://dl.acm.org/doi/10.1145/3240323.3240360`.

[9] Christian Borgs, Jennifer Chayes, Christian Ikeokwu, and Ellen Vitercik. Bursting the Filter Bubble: Disincentivizing Echo Chambers in Social Networks. In *Proceedings of EAAMO'23: ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, 2023.

[10] Marco Caponigro, Anna Chiara Lai, and Benedetto Piccoli. A nonlinear model of opinion formation on the sphere. *Discrete & Continuous Dynamical Systems - A*, 35(9):4241–4268, 2015. ISSN 1553-5231. doi: 10.3934/dcds.2015.35.4241. URL `http://aimsciences.org//article/doi/10.3934/dcds.2015.35.4241`.

[11] Jaime Carbonell and Jade Goldstein. The use of mmr, diversity-based reranking for reordering documents and producing summaries. In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '98, page 335–336, New York, NY, USA, 1998. Association for Computing Machinery. ISBN 1581130155. doi: 10.1145/290941.291025. URL `https://doi.org/10.1145/290941.291025`.

[12] 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. ISBN 9781450349130. doi: 10.1145/3038912.3052585. URL `https://doi.org/10.1145/3038912.3052585`.

[13] Paul Covington, Jay Adams, and Emre Sargin. Deep Neural Networks for YouTube Recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems*, pages 191–198, Boston Massachusetts USA, September 2016. ACM. ISBN 978-1-4503-4035-9. doi: 10.1145/2959100.2959190. URL `https://dl.acm.org/doi/10.1145/2959100.2959190`.

[14] Sarah Dean and Jamie Morgenstern. Preference Dynamics Under Personalized Recommendations. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 795–816, Boulder CO USA, July 2022. ACM. ISBN 978-1-4503-9150-4. doi: 10.1145/3490486.3538346. URL `https://dl.acm.org/doi/10.1145/3490486.3538346`.

[15] Itay Eilat and Nir Rosenfeld. Performative Recommendation: Diversifying Content via Strategic Incentives, June 2023. URL `http://arxiv.org/abs/2302.04336`. arXiv:2302.04336 [cs].

[16] Chongming Gao, Shijun Li, Yuan Zhang, Jiawei Chen, Biao Li, Wenqiang Lei, Peng Jiang, and Xiangnan He. Kuairand: An unbiased sequential recommendation dataset with randomly exposed videos. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, CIKM '22, page 3953–3957, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450392365. doi: 10.1145/3511808.3557624. URL `https://doi.org/10.1145/3511808.3557624`.

[17] Benjamin Golub and Matthew O. Jackson. Naïve learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1):112–49, February 2010. doi: 10.1257/mic.2.1.112. URL `https://www.aeaweb.org/articles?id=10.1257/mic.2.1.112`.

[18] Moritz Hardt, Meena Jagadeesan, and Celestine Mendler-Dünner. Performative power. *Advances in Neural Information Processing Systems*, 35:22969–22981, 2022.

[19] F Maxwell Harper and Joseph A Konstan. The movielens datasets: History and context. *Acm transactions on interactive intelligent systems (tiis)*, 5(4):1–19, 2015.

[20] Jiri Hron, Karl Krauth, Michael I. Jordan, Niki Kilbertus, and Sarah Dean. Modeling Content Creator Incentives on Algorithm-Curated Platforms, July 2023. URL `http://arxiv.org/abs/2206.13102`. arXiv:2206.13102 [cs, stat].

[21] Po-Sen Huang, Xiaodong He, Jianfeng Gao, Li Deng, Alex Acero, and Larry Heck. Learning deep structured semantic models for web search using clickthrough data. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*, pages 2333–2338, San Francisco California USA, October 2013. ACM. ISBN 978-1-4503-2263-8. doi: 10.1145/2505515.2505665. URL `https://dl.acm.org/doi/10.1145/2505515.2505665`.

[22] Neil J. Hurley. Personalised ranking with diversity. In *Proceedings of the 7th ACM Conference on Recommender Systems*, RecSys '13, page 379–382, New York, NY, USA, 2013. Association for Computing Machinery. ISBN 9781450324090. doi: 10.1145/2507157.2507226. URL `https://doi.org/10.1145/2507157.2507226`.

[23] Meena Jagadeesan, Nikhil Garg, and Jacob Steinhardt. Supply-side equilibria in recommender systems. *Advances in Neural Information Processing Systems*, 36, 2024.

[24] Ray Jiang, Silvia Chiappa, Tor Lattimore, András György, and Pushmeet Kohli. Degenerate Feedback Loops in Recommender Systems. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pages 383–390, Honolulu HI USA, January 2019. ACM. ISBN 978-1-4503-6324-2. doi: 10.1145/3306618.3314288. URL `https://dl.acm.org/doi/10.1145/3306618.3314288`.

[25] Natali Helberger Judith Möller, Damian Trilling and Bram van Es. Do not blame it on the algorithm: an empirical assessment of multiple recommender systems and their impact on content diversity. *Information, Communication & Society*, 21(7):959–977, 2018. doi: 10.1080/1369118X.2018.1444076. URL `https://doi.org/10.1080/1369118X.2018.1444076`.

[26] 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, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383325. doi: 10.1145/3447548.3467298. URL `https://doi.org/10.1145/3447548.3467298`.

[27] Sagi Levanon and Nir Rosenfeld. Generalized strategic classification and the case of aligned incentives. In *International Conference on Machine Learning*, pages 12593–12618. PMLR, 2022.

[28] Lihong Li, Wei Chu, John Langford, and Robert E. Schapire. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, page 661–670, New York, NY, USA, 2010. Association for Computing Machinery. ISBN 9781605587998. doi: 10.1145/1772690.1772758. URL `https://doi.org/10.1145/1772690.1772758`.

[29] Wei Li and Mark W. Spong. Unified cooperative control of multiple agents on a sphere for different spherical patterns. *IEEE Transactions on Automatic Control*, 59(5):1283–1289, 2014. doi: 10.1109/TAC.2013.2286897.

[30] Dugang Liu, Pengxiang Cheng, Zinan Lin, Xiaolian Zhang, Zhenhua Dong, Rui Zhang, Xiuqiang He, Weike Pan, and Zhong Ming. Bounding System-Induced Biases in Recommender Systems with a Randomized Dataset. *ACM Transactions on Information Systems*, 41(4):1–26, October 2023. ISSN 1046-8188, 1558-2868. doi: 10.1145/3582002. URL `https://dl.acm.org/doi/10.1145/3582002`.

[31] Johan Markdahl, Johan Thunberg, and Jorge Goncalves. Almost Global Consensus on the n-Sphere. *IEEE Transactions on Automatic Control*, 63(6):1664–1675, June 2018. ISSN 0018-9286, 1558-2523. doi: 10.1109/TAC.2017.2752799. URL `https://ieeexplore.ieee.org/document/8038829/`.

[32] Farzan Masrour, Tyler Wilson, Heng Yan, Pang-Ning Tan, and Abdol Esfahanian. Bursting the filter bubble: Fairness-aware network link prediction. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(01):841–848, Apr. 2020. doi: 10.1609/aaai.v34i01.5429. URL `https://ojs.aaai.org/index.php/AAAI/article/view/5429`.

[33] Tien T. Nguyen, Pik-Mai Hui, F. Maxwell Harper, Loren Terveen, and Joseph A. Konstan. Exploring the filter bubble: the effect of using recommender systems on content diversity. In *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14, page 677–686, New York, NY, USA, 2014. Association for Computing Machinery. ISBN 9781450327442. doi: 10.1145/2566486.2568012. URL `https://doi.org/10.1145/2566486.2568012`.

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

[35] Siddharth Prasad, Martin Mladenov, and Craig Boutilier. Content prompting: Modeling content provider dynamics to improve user welfare in recommender ecosystems. *arXiv preprint arXiv:2309.00940*, 2023.

[36] Fernando P. 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*, 118(50):e2102141118, 2021. doi: 10.1073/pnas.2102141118. URL `https://www.pnas.org/doi/abs/10.1073/pnas.2102141118`.

[37] Alain Sarlette, Rodolphe Sepulchre, and Naomi Ehrich Leonard. Autonomous rigid body attitude synchronization. In *2007 46th IEEE Conference on Decision and Control*, pages 2566–2571, 2007. doi: 10.1109/CDC.2007.4434153.

[38] 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. Association for Computing Machinery. ISBN 9781450324090. doi: 10.1145/2507157.2507207. URL `https://doi.org/10.1145/2507157.2507207`.

[39] Mark Wilhelm, Ajith Ramanathan, Alexander Bonomo, Sagar Jain, Ed H. Chi, and Jennifer Gillenwater. Practical diversified recommendations on youtube with determinantal point processes. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, CIKM '18, page 2165–2173, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450360142. doi: 10.1145/3269206.3272018. URL `https://doi.org/10.1145/3269206.3272018`.

[40] Longqi Yang, Yin Cui, Yuan Xuan, Chenyang Wang, Serge Belongie, and Deborah Estrin. Unbiased offline recommender evaluation for missing-not-at-random implicit feedback. In *Proceedings of the 12th ACM Conference on Recommender Systems*, RecSys '18, page 279–287, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450359016. doi: 10.1145/3240323.3240355. URL `https://doi.org/10.1145/3240323.3240355`.

[41] Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang, and Haifeng Xu. Learning from a learning user for optimal recommendations. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 25382–25406. PMLR, 17–23 Jul 2022. URL `https://proceedings.mlr.press/v162/yao22a.html`.

[42] Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang, and Haifeng Xu. How Bad is Top-K Recommendation under Competing Content Creators? In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 39674–39701. PMLR, July 2023. URL `https://proceedings.mlr.press/v202/yao23b.html`.

[43] 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*, KDD '24, page 3874–3885, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400704901. doi: 10.1145/3637528.3672021. URL `https://doi.org/10.1145/3637528.3672021`.

[44] 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. Association for Computing Machinery. ISBN 9781605580937. doi: 10.1145/1454008.1454030. URL https://doi.org/10.1145/1454008.1454030.

[45] 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*, pages 490–498, 2023.

[46] Ziqiao Zhang, Said Al-Abri, and Fumin Zhang. Opinion Dynamics on the Sphere for Stable Consensus and Stable Bipartite Dissensus. *9th IFAC Conference on Networked Systems NECSYS 2022*, 55(13):288–293, January 2022. ISSN 2405-8963. doi: 10.1016/j.ifacol.2022.07.274.

[47] 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. Association for Computing Machinery. ISBN 1595930469. doi: 10.1145/1060745.1060754.

A Additional Discussions on Related Works

Table 2: Comparison between our work and some previous works on performative effects of recommender systems

Works	Adaptive Users?	Adaptive Creators?	Creator Reward	Dynamics or Equilibrium?	Content Adjustment Model
Ours	Yes	Yes	User engagement	Dynamics	Conditioned on previous time step; implicit cost of content adjustment
[15]	No	Yes	Exposure	Dynamics	Conditioned on previous time step; explicit cost of content adjustment
[42]	No	Yes	User engagement	Dynamics	Freely choose without cost
[35]	No	Yes	User engagement	Dynamics	Freely choose without cost
[23]	No	Yes	Exposure	Equilibrium	Freely choose with cost
[20]	No	Yes	Exposure	Equilibrium	Freely choose without cost
[7]	No	Yes	Exposure	Equilibrium	Freely choose without cost
[2]	No	Yes	User engagement	Equilibrium	Freely choose without cost
[43]	No	Yes	Designed by a welfare-maximizing platform	Dynamics	Freely choose without cost
[14]	Yes	No[1]	N/A	Dynamics	N/A
[41]	Yes	No[1]	N/A	Dynamics	N/A
[3]	Adaptive and adversarial	No[1]	N/A	Dynamics	N/A

[1]: These works study the design of recommendation algorithms for the platform with a fixed set of content, without explicitly modeling the content creators.

B Additional Experiments on Synthetic Data

B.1 Number of Fixed Dimensions

In this part, we consider the case where some dimensions of the user feature vectors are fixed features and thus not updated from round to round (e.g., ages, genders). Formally, we fix the first $k \leq d$ dimensions. The remaining $d - k$ dimensions $\boldsymbol{u}_j^t[k + 1 : d] = (u_j^t[k + 1], \ldots, u_j^t[d])$ are updated according to the following rule: $\boldsymbol{u}_j^{t+1}[k + 1 : d] = \|\boldsymbol{u}_j^t[k + 1 : d]\| \cdot \mathcal{P}(\boldsymbol{u}_j^t[k + 1 : d] + \eta_u f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t)\boldsymbol{v}_i^t[k+1 : d])$. The multiplication by $\|\boldsymbol{u}_j^t[k + 1 : d]\|$ ensures unit norm $\|\boldsymbol{u}_j^{t+1}\| = 1$. The effect of the number of fixed dimensions on the dynamics is shown in Figure 7. We see that the diversity of the system *improves* as the number of fixed dimensions increases, and the degree of polarization is reduced. This is similar to the effect of decreasing user update rate η_u in Figure 4. The observation that fixed user features encourage diversity might be a reason why the recommender systems in practice are not as polarized as our theoretical prediction.
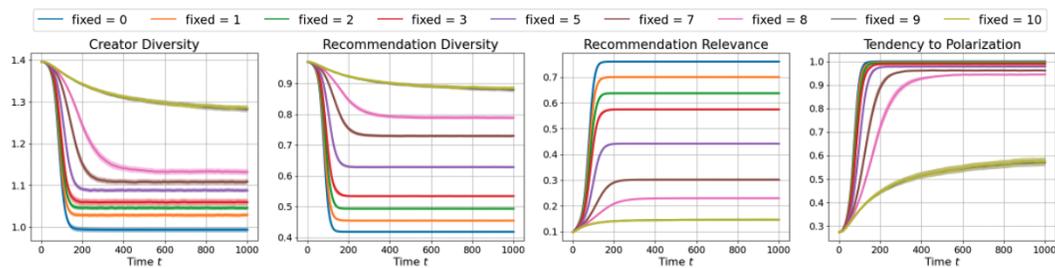

Figure 7: Changes of measures over time under different numbers of fixed dimensions, on synthetic data

B.2 Threshold Truncation

Table 3 shows that the effect of different thresholds in threshold truncation on the long-term diversity of the system. We see that truncating at $\tau = 0$, which corresponds to $90°$ angle between $\boldsymbol{u}_j$ and $\boldsymbol{v}_i$, is *not good* for diversity, resulting in the lowest creator diversity measure (CD) and highest tendency to polarization (TP). Truncating at a large threshold like 0.707 is good for diversity, instead.

Figure 8 shows how the diversity measures change over time, under different truncation thresholds.

Table 3: Diversity improvement by threshold truncation on synthetic data

| β	threshold τ	CD	RD	RR	TP
0	$-\cos(60°) = -0.5$	1.00 ± 0.03	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.99 ± 10^{-3}
	$-\cos(72°) = -0.309$	0.96 ± 0.06	$\mathbf{0.01 \pm 0.02}$	1.00 ± 0.02	0.92 ± 0.10
	$\cos(90°) = 0$	0.03 ± 0.16	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.99 ± 0.04
	$\cos(72°) = 0.309$	0.72 ± 0.30	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.81 ± 0.12
	$\cos(60°) = 0.5$	1.16 ± 0.11	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.47 ± 0.10
	$\cos(45°) = 0.707$	$\mathbf{1.37 \pm 0.02}$	0.00 ± 10^{-3}	$\mathbf{1.00 \pm 10^{-3}}$	$\mathbf{0.33 \pm 0.02}$
	$\cos(30°) = 0.866$	1.30 ± 0.03	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.55 ± 0.05
1	$-\cos(60°) = -0.5$	0.98 ± 0.04	0.00 ± 0.02	1.00 ± 0.01	0.96 ± 0.04
	$-\cos(72°) = -0.309$	0.92 ± 0.08	0.00 ± 0.02	0.99 ± 0.02	0.87 ± 0.10
	$\cos(90°) = 0$	0.13 ± 0.31	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.97 ± 0.08
	$\cos(72°) = 0.309$	0.85 ± 0.16	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.76 ± 0.11
	$\cos(60°) = 0.5$	1.21 ± 0.07	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.43 ± 0.08
	$\cos(45°) = 0.707$	$\mathbf{1.38 \pm 0.01}$	0.00 ± 10^{-3}	$\mathbf{1.00 \pm 10^{-3}}$	$\mathbf{0.30 \pm 0.01}$
	$\cos(30°) = 0.866$	1.33 ± 0.02	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.47 ± 0.04
3	$-\cos(60°) = -0.5$	0.91 ± 0.18	$\mathbf{0.01 \pm 0.02}$	1.00 ± 0.01	0.83 ± 0.10
	$-\cos(72°) = -0.309$	0.85 ± 0.23	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.78 ± 0.11
	$\cos(90°) = 0$	0.64 ± 0.33	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.81 ± 0.12
	$\cos(72°) = 0.309$	1.01 ± 0.14	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.64 ± 0.14
	$\cos(60°) = 0.5$	1.26 ± 0.05	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.38 ± 0.06
	$\cos(45°) = 0.707$	$\mathbf{1.39 \pm 0.01}$	0.00 ± 10^{-3}	$\mathbf{1.00 \pm 10^{-3}}$	$\mathbf{0.28 \pm 0.01}$
	$\cos(30°) = 0.866$	1.37 ± 0.01	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.34 ± 0.01
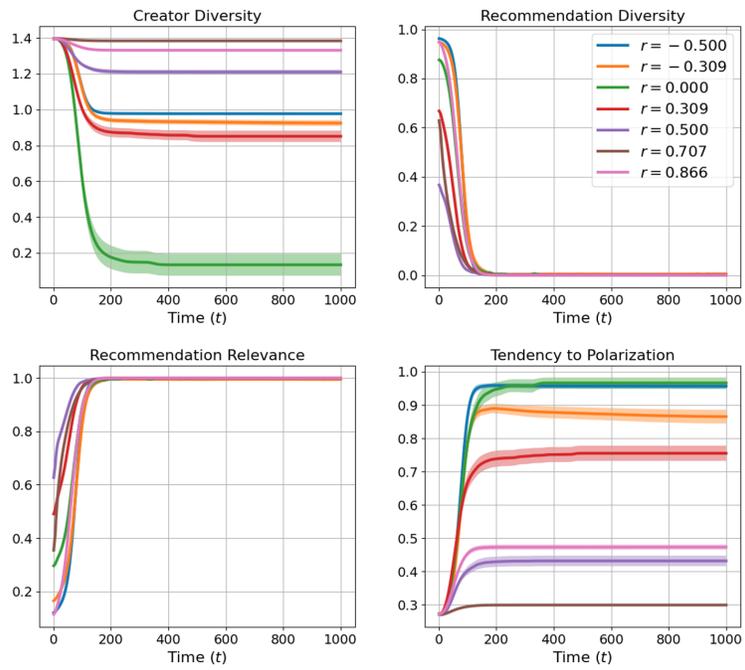
Figure 8: Changes of measures over time under different truncation threshold τ, with $\beta = 1$, on synthetic data

C Additional Experiments on Real-World Dataset

C.1 Details of the Recommendation Algorithm
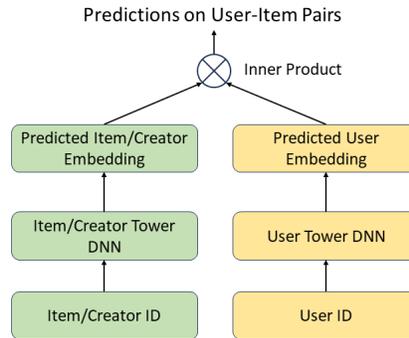
Figure 9: Two tower model for the MovieLens experiment, where the two towers both have size 16×16 with linear layers and ReLu activations.

Algorithm 1 Real-world Recommendation with Dual Influence

Input: $t = 0$, actual embedding $U^{(0)}, V^{(0)}$, true labels $Y_{ij}^{(0)} := y(u_i^{(0)}, v_j^{(0)})$, initial parameter $\boldsymbol{\theta}^{(0)}$ (which includes the predicted embedding $\hat{U}^{(0)}, \hat{V}^{(0)}$)
repeat
    Let temporary parameter $\boldsymbol{w}^{(0)} \leftarrow \boldsymbol{\theta}^{(t)}$
    Compute loss $\mathcal{L}(\boldsymbol{\theta}^{(t)}, Y^{(t)})$
    for $s = 1$ **to** $m - 1$ **do**
        $\boldsymbol{w}^{(s+1)} \leftarrow \boldsymbol{\theta}^{(s)} - \eta \nabla_{\boldsymbol{w}} \mathcal{L}(\boldsymbol{w}^{(s)}, Y^{(t)})$
    end for
    $\boldsymbol{\theta}^{(t+1)} \leftarrow \boldsymbol{w}^{(m)}$
    Deliver recommendations based on $\hat{U}^{(t+1)}, \hat{V}^{(t+1)}$
    Update $U^{(t+1)}, V^{(t+1)}$, and $Y^{(t+1)}$
    $t \leftarrow t + 1$
until $\|\boldsymbol{\theta}^{(t)} - \boldsymbol{\theta}^{(t-1)}\|_2 \leq \delta$

C.2 Top-k and Threshold Truncations

We also try top-k and threshold truncations (Section 4) on the MovieLens 20M dataset. Here, we have $n = 2000$ creators and $m = 2000$ users, with feature dimension $d = 16$. The results for top-k truncation are in Table 4 and Figure 10. Similar to the experiments with synthetic data, we see that a smaller k improves Creator Diversity (CD) and Recommendation Relevance (RR), reduces Tendency to Polarization (TP), yet worsens Recommendation Diversity (RD).

Table 4: Diversity improvement by top-k truncation on MovieLens 20M dataset

| β	k	CD	RD	RR	TP
0	2000	1.00 ± 10^{-3}	$\mathbf{1.00 \pm 10^{-3}}$	0.00 ± 10^{-3}	1.00 ± 10^{-3}
	1000	0.30 ± 0.04	0.03 ± 0.01	0.88 ± 0.01	1.00 ± 10^{-3}
	500	1.10 ± 0.06	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.43 ± 0.03
	100	1.36 ± 10^{-3}	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.28 ± 0.01
	10	1.40 ± 10^{-3}	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.20 ± 10^{-3}
	1	$\mathbf{1.40 \pm 10^{-3}}$	0.00 ± 10^{-3}	$\mathbf{1.00 \pm 10^{-3}}$	$\mathbf{0.20 \pm 10^{-3}}$
1	2000	1.00 ± 10^{-3}	$\mathbf{0.42 \pm 10^{-3}}$	0.92 ± 0.01	1.00 ± 10^{-3}
	1000	0.61 ± 0.16	0.03 ± 0.01	0.97 ± 0.01	0.90 ± 0.06
	500	1.14 ± 0.04	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.41 ± 0.04
	100	1.35 ± 0.01	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.27 ± 10^{-3}
	10	1.40 ± 10^{-3}	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.20 ± 10^{-3}
	1	$\mathbf{1.40 \pm 10^{-3}}$	0.00 ± 10^{-3}	$\mathbf{1.00 \pm 10^{-3}}$	$\mathbf{0.20 \pm 10^{-3}}$
3	2000	0.92 ± 0.07	$\mathbf{0.02 \pm 0.01}$	0.99 ± 10^{-3}	0.91 ± 0.05
	1000	0.65 ± 0.18	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.69 ± 0.14
	500	1.07 ± 0.07	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.48 ± 0.11
	100	1.36 ± 0.01	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.27 ± 0.01
	10	1.40 ± 10^{-3}	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.20 ± 10^{-3}
	1	$\mathbf{1.40 \pm 10^{-3}}$	0.00 ± 10^{-3}	$\mathbf{1.00 \pm 10^{-3}}$	$\mathbf{0.20 \pm 10^{-3}}$
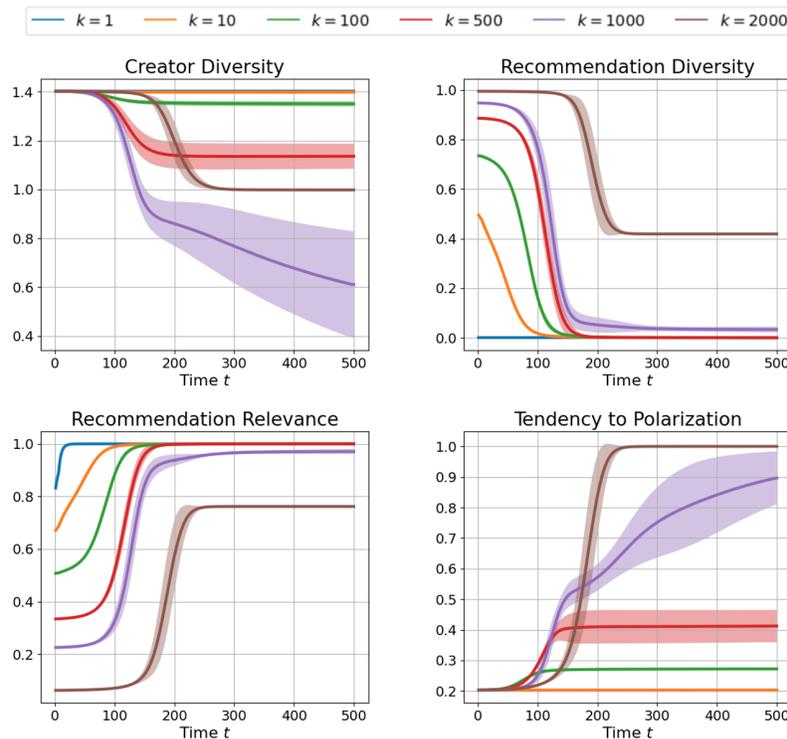
Figure 10: Changes of measures over time under different k, with $\beta = 1$, on MovieLens 20M dataset

Results for threshold truncation are in Table 5 and Figure 11. Similar to synthetic data, we see that a large (but not too large) threshold like 0.707 is good for improving CD and TP.

Table 5: Threshold truncation with different thresholds on MovieLens 20M dataset

| β	threshold τ	CD	RD	RR	TP
0	$-\cos(60°) = -0.5$	1.00 ± 10^{-3}	0.00 ± 10^{-3}	1.00 ± 10^{-3}	1.00 ± 10^{-3}
	$-\cos(72°) = -0.309$	1.00 ± 10^{-3}	0.00 ± 10^{-3}	1.00 ± 10^{-3}	1.00 ± 10^{-3}
	$\cos(90°) = 0$	0.01 ± 0.01	0.00 ± 10^{-3}	1.00 ± 10^{-3}	1.00 ± 10^{-3}
	$\cos(72°) = 0.309$	0.83 ± 0.08	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.72 ± 0.09
	$\cos(60°) = 0.5$	1.20 ± 0.05	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.46 ± 0.07
	$\cos(45°) = 0.707$	$\mathbf{1.39 \pm 10^{-3}}$	0.00 ± 10^{-3}	1.00 ± 10^{-3}	$\mathbf{0.20 \pm 10^{-3}}$
	$\cos(30°) = 0.866$	1.36 ± 10^{-3}	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.40 ± 0.01
1	$-\cos(60°) = -0.5$	1.00 ± 10^{-3}	0.00 ± 10^{-3}	1.00 ± 10^{-3}	1.00 ± 10^{-3}
	$-\cos(72°) = -0.309$	0.96 ± 0.03	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.95 ± 0.03
	$\cos(90°) = 0$	0.02 ± 0.02	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.99 ± 10^{-3}
	$\cos(72°) = 0.309$	0.83 ± 0.07	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.66 ± 0.10
	$\cos(60°) = 0.5$	1.18 ± 0.06	0.00 ± 10^{-3}	1.00 ± 0.01	0.44 ± 0.07
	$\cos(45°) = 0.707$	$\mathbf{1.40 \pm 10^{-3}}$	0.00 ± 10^{-3}	1.00 ± 10^{-3}	$\mathbf{0.20 \pm 10^{-3}}$
	$\cos(30°) = 0.866$	1.35 ± 0.01	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.40 ± 0.02
3	$-\cos(60°) = -0.5$	0.77 ± 0.27	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.86 ± 0.09
	$-\cos(72°) = -0.309$	0.80 ± 0.24	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.79 ± 0.13
	$\cos(90°) = 0$	0.04 ± 0.02	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.98 ± 0.01
	$\cos(72°) = 0.309$	0.99 ± 0.11	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.55 ± 0.13
	$\cos(60°) = 0.5$	1.26 ± 0.05	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.36 ± 0.06
	$\cos(45°) = 0.707$	$\mathbf{1.40 \pm 10^{-3}}$	0.00 ± 10^{-3}	1.00 ± 10^{-3}	$\mathbf{0.20 \pm 10^{-3}}$
	$\cos(30°) = 0.866$	1.36 ± 10^{-3}	0.00 ± 10^{-3}	1.00 ± 10^{-3}	0.39 ± 0.01
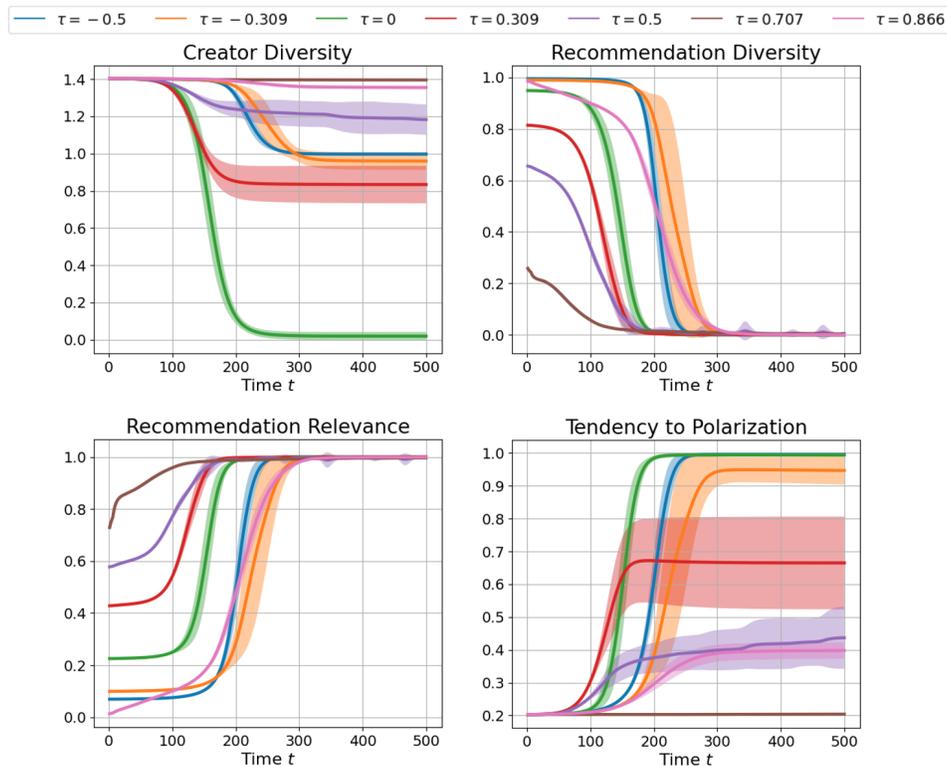
Figure 11: Changes of measures over time under truncation with different threshold τ, with $\beta = 1$, on MovieLens 20M dataset

D   Useful Lemmas

This section provides some lemmas that will be used in the proofs. They are mainly about some properties of the dynamics update rule.

Claim D.1. *For vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ with unit norm $\|\boldsymbol{x}\|_2 = \|\boldsymbol{y}\|_2 = 1$, we have:*

- $\|\boldsymbol{x} - \boldsymbol{y}\|_2^2 = 2(1 - \langle \boldsymbol{x}, \boldsymbol{y} \rangle)$.

- $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 1 - \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{y}\|_2^2$.

Lemma D.2 (Convex Cone Property). *Let $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k \in \mathbb{R}^d$ be vectors with norm $\|\boldsymbol{z}_i^t\|_2 = 1$. Suppose $\langle \boldsymbol{z}_i, \boldsymbol{y} \rangle > 0$ for every $i = 1, \ldots, k$ for some $\boldsymbol{y} \in \mathbb{R}^d$. Let $\boldsymbol{x} = \mathcal{P}(\sum_{i=1}^k a_i \boldsymbol{z}_i)$ for some $a_1, \ldots, a_k \geq 0$ (namely, $\boldsymbol{x}$ is the normalization of some vector in the convex cone formed by $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k$). Then, we have*

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle \geq \min_{i=1}^k \langle \boldsymbol{z}_i, \boldsymbol{y} \rangle > 0 \quad \text{and} \quad \|\boldsymbol{x} - \boldsymbol{y}\|_2 \leq \max_{i=1}^k \|\boldsymbol{z}_i - \boldsymbol{y}\|_2 > 0.$$

Proof.

$$
\begin{aligned}
\langle \boldsymbol{x}, \boldsymbol{y} \rangle &= \Big\langle \frac{\sum_{i=1}^k a_i \boldsymbol{z}_i}{\|\sum_{i=1}^k a_i \boldsymbol{z}_i\|_2}, \boldsymbol{y} \Big\rangle = \frac{1}{\|\sum_{i=1}^k a_i \boldsymbol{z}_i\|_2} \sum_{i=1}^k a_i \langle \boldsymbol{z}_i, \boldsymbol{y} \rangle \\
&\geq \frac{1}{\|\sum_{i=1}^k a_i \boldsymbol{z}_i\|_2} \sum_{i=1}^k a_i \min_{i=1}^k \langle \boldsymbol{z}_i, \boldsymbol{y} \rangle = \min_{i=1}^k \langle \boldsymbol{z}_i, \boldsymbol{y} \rangle \frac{\sum_{i=1}^k a_i}{\|\sum_{i=1}^k a_i \boldsymbol{z}_i\|_2} \\
&\geq \min_{i=1}^k \langle \boldsymbol{z}_i, \boldsymbol{y} \rangle \frac{\sum_{i=1}^k a_i}{\sum_{i=1}^k a_i} = \min_{i=1}^k \langle \boldsymbol{z}_i, \boldsymbol{y} \rangle.
\end{aligned}
$$

This proves the first inequality. To prove the second inequality, we use Claim D.1 and the first inequality:

$$\|\boldsymbol{x} - \boldsymbol{y}\|_2 = \sqrt{2(1 - \langle \boldsymbol{x}, \boldsymbol{y} \rangle)} \leq \sqrt{2(1 - \min_i \langle \boldsymbol{z}_i, \boldsymbol{y} \rangle)} = \max_i \sqrt{2(1 - \langle \boldsymbol{z}_i, \boldsymbol{y} \rangle)} = \max_{i=1}^k \|\boldsymbol{z}_i - \boldsymbol{y}\|_2.$$

$\square$

Lemma D.3. *Let $\boldsymbol{x}^t, \boldsymbol{y}, \boldsymbol{z}^t \in \mathbb{R}^d$ be vectors with norm $\|\boldsymbol{x}^t\|_2 = 1$, $\|\boldsymbol{y}\|_2 \geq 0$, $\|\boldsymbol{z}^t\|_2 \leq 1$. Suppose $\langle \boldsymbol{x}^t, \boldsymbol{y} \rangle \geq 0$, $\langle \boldsymbol{z}^t, \boldsymbol{y} \rangle \geq 0$. After the update $\boldsymbol{x}^{t+1} = \mathcal{P}(\boldsymbol{x}^t + \eta \boldsymbol{z}^t)$, we have*

$$\langle \boldsymbol{x}^{t+1} - \boldsymbol{x}^t, \boldsymbol{y} \rangle \geq \frac{\eta}{1 + \eta \|\boldsymbol{z}^t\|_2} \Big(\langle \boldsymbol{z}^t, \boldsymbol{y} \rangle - \|\boldsymbol{z}^t\|_2 \langle \boldsymbol{x}^t, \boldsymbol{y} \rangle \Big).$$

As a corollary, if $\boldsymbol{y} = \boldsymbol{z}^t$ and $\|\boldsymbol{z}^t\|_2 = 1$, then

$$\langle \boldsymbol{x}^{t+1} - \boldsymbol{x}^t, \boldsymbol{z}^t \rangle \geq \frac{\eta}{1 + \eta} \Big(1 - \langle \boldsymbol{x}^t, \boldsymbol{z}^t \rangle \Big).$$

Proof. By definition,

$$
\begin{aligned}
\langle \boldsymbol{x}^{t+1} - \boldsymbol{x}^t, \boldsymbol{y} \rangle &= \Big\langle \frac{\boldsymbol{x}^t + \eta \boldsymbol{z}^t}{\|\boldsymbol{x}^t + \eta \boldsymbol{z}^t\|_2} - \boldsymbol{x}^t, \boldsymbol{y} \Big\rangle \\
&= \Big(\frac{1}{\|\boldsymbol{x}^t + \eta \boldsymbol{z}^t\|_2} - 1 \Big) \cdot \langle \boldsymbol{x}^t, \boldsymbol{y} \rangle + \frac{\eta}{\|\boldsymbol{x}^t + \eta \boldsymbol{z}^t\|_2} \cdot \langle \boldsymbol{z}^t, \boldsymbol{y} \rangle \\
\text{(because } \|\boldsymbol{x}^t + \eta \boldsymbol{z}^t\|_2 \leq 1 + \eta \|\boldsymbol{z}^t\|_2 \text{)} \quad &\geq \Big(\frac{1}{1 + \eta \|\boldsymbol{z}^t\|_2} - 1 \Big) \cdot \langle \boldsymbol{x}^t, \boldsymbol{y} \rangle + \frac{\eta}{1 + \eta \|\boldsymbol{z}^t\|_2} \cdot \langle \boldsymbol{z}^t, \boldsymbol{y} \rangle \\
&= \frac{\eta}{1 + \eta \|\boldsymbol{z}^t\|_2} \Big(\langle \boldsymbol{z}^t, \boldsymbol{y} \rangle - \|\boldsymbol{z}^t\|_2 \langle \boldsymbol{x}^t, \boldsymbol{y} \rangle \Big).
\end{aligned}
$$

$\square$

Lemma D.4. *Let $\boldsymbol{x}^t, \boldsymbol{z}^t \in \mathbb{R}^d$ be vectors with norm $\|\boldsymbol{x}^t\|_2 = 1$, $\|\boldsymbol{z}^t\|_2 \leq 1$. Suppose $\langle \boldsymbol{x}^t, \boldsymbol{z}^t \rangle \geq 0$ and $\eta > 0$. Then the update $\boldsymbol{x}^{t+1} = \mathcal{P}(\boldsymbol{x}^t + \eta \boldsymbol{z}^t)$ satisfies*

- $\langle x^{t+1} - x^t, z^t \rangle \geq \frac{1}{\eta} \|x^{t+1} - x^t\|_2^2$.

- $\|x^{t+1} - x^t\|_2 \leq \eta \|z^t\|_2$.

Proof. Let $\tilde{x}^{t+1} = x^t + \eta z^t$, so $x^t = \mathcal{P}(\tilde{x}^{t+1})$ and $z^t = \frac{1}{\eta}(\tilde{x}^{t+1} - x^t)$. Then we have

$$\langle x^{t+1} - x^t, z^t \rangle = \frac{1}{\eta} \langle x^{t+1} - x^t, \tilde{x}^{t+1} - x^t \rangle.$$

Because $\langle x^t, z^t \rangle \geq 0$, the vector $\tilde{x}^{t+1} = x^t + \eta z^t$ has length ≥ 1 and hence is outside (or on the surface) of the d-dimensional unit ball. Since $x^t = \mathcal{P}(\tilde{x}^{t+1})$ is the projection of $\tilde{x}^{t+1}$ onto the unit ball, and z^t is another vector inside the unit ball, by the "Pythagorean property" (Proposition 2.2 in [6]), we must have $\langle x^t - x^{t+1}, \tilde{x}^{t+1} - x^{t+1} \rangle \leq 0$. This implies

$$\langle x^{t+1} - x^t, z^t \rangle \geq \frac{1}{\eta} \Big(\langle x^{t+1} - x^t, \tilde{x}^{t+1} - x^t \rangle + \langle x^t - x^{t+1}, \tilde{x}^{t+1} - x^{t+1} \rangle \Big)$$
$$= \frac{1}{\eta} \langle x^{t+1} - x^t, x^{t+1} - x^t \rangle = \frac{1}{\eta} \|x^{t+1} - x^t\|_2^2,$$

which proves the first claim. To prove the second claim, we use Cauchy-Schwarz inequality:

$$\frac{1}{\eta} \|x^{t+1} - x^t\|_2^2 \leq \langle x^{t+1} - x^t, z^t \rangle \leq \|x^{t+1} - x^t\|_2 \|z^t\|_2.$$

This implies $\|x^{t+1} - x^t\|_2 \leq \eta \|z^t\|_2$. $\qquad \square$

Lemma D.5. *Consider a creator v_i^t and a user u_j^t. Suppose the user is always recommended creator i (so the user is updated by $u_j^{t+1} = \mathcal{P}(u_j^t + \eta_u f(v_i^t, u_j^t) v_i^t)$), and creator i is updated by $v_i^{t+1} = \mathcal{P}(v_i^t + \eta_c \alpha_i^t)$ with $\|\alpha_i^t\|_2 \leq 1$ and $\langle v_i^t, \alpha_i^t \rangle \geq 0$ at each time step. Assume:*

- *The inner product $\langle u_j^0, v_i^0 \rangle > 0$ initially. (Note that $\langle u_j^0, u_{j'}^0 \rangle$ needs not hold.)*

- *There exists some constant $L_f > 0$ such that $f(v_i, u_j) \geq L_f > 0$ whenever $\langle u_j, v_i \rangle > 0$.*

- $\eta_c \leq \frac{\eta_u L_f}{2}$ *and* $0 \leq \eta_u < \frac{1}{2}$.

Then, we have $\langle u_j^t, v_i^t \rangle > 0$ in all time steps.

Proof. We prove by induction. Suppose $\langle u_j^t, v_i^t \rangle > 0$ already holds. We prove that $\langle u_j^{t+1}, v_i^{t+1} \rangle > 0$ will also hold. Take the difference between $\langle u_j^{t+1}, v_i^{t+1} \rangle$ and $\langle u_j^t, v_i^t \rangle$:

$$\langle u_j^{t+1}, v_i^{t+1} \rangle - \langle u_j^t, v_i^t \rangle = \langle u_j^{t+1}, v_i^{t+1} - v_i^t \rangle + \langle u_j^{t+1} - u_j^t, v_i^t \rangle.$$

For $\langle u_j^{t+1} - u_j^t, v_i^t \rangle$, using Lemma D.3 with $x^t = u_j^t$, $z^t = v_i^t$, and $\eta = \eta_u f(v_i^t, u_j^t)$, we get

$$\langle u_j^{t+1} - u_j^t, v_i^t \rangle \geq \frac{\eta_u f(v_i^t, u_j^t)}{1 + \eta_u f(v_i^t, u_j^t)} \big(1 - \langle u_j^t, v_i^t \rangle \big) \geq \frac{\eta_u L_f}{1 + \eta_u L_f} \big(1 - \langle u_j^t, v_i^t \rangle \big).$$

For $\langle u_j^{t+1}, v_i^{t+1} - v_i^t \rangle$, by Cauchy-Schwarz inequality and Lemma D.4,

$$\langle u_j^{t+1}, v_i^{t+1} - v_i^t \rangle \geq - \|u_j^{t+1}\|_2 \cdot \|v_i^{t+1} - v_i^t\|_2 \geq -1 \cdot \eta_c \|\alpha_i^t\|_2 \geq -\eta_c.$$

- If $1 - \langle u_j^t, v_i^t \rangle > \frac{1}{2}(1 + \eta_u L_f)$, then we have

$$\langle u_j^{t+1}, v_i^{t+1} \rangle - \langle u_j^t, v_i^t \rangle > \eta_u L_f \frac{1}{2} - \eta_c \geq 0$$

  by the assumption of $\eta_c \leq \frac{\eta_u L_f}{2}$.

- If $1 - \langle u_j^t, v_i^t \rangle \leq \frac{1}{2}(1 + \eta_u L_f)$, then we have

$$\langle u_j^{t+1}, v_i^{t+1} \rangle - \langle u_j^t, v_i^t \rangle \geq 0 - \eta_c$$
$$\implies \langle u_j^{t+1}, v_i^{t+1} \rangle \geq \langle u_j^t, v_i^t \rangle - \eta_c \geq \frac{1}{2} - \frac{1}{2}\eta_u L_f - \eta_c > 0$$

  under the assumption of $\eta_c \leq \frac{\eta_u L_f}{2}$ and $\eta_u < \frac{1}{2}$.

The above two cases together ensure $\langle \boldsymbol{u}_j^{t+1}, \boldsymbol{v}_i^{t+1} \rangle > 0$. $\hfill \square$

Lemma D.6. *Consider a system of one user and one creator that satisfies $\langle \boldsymbol{u}_j^0, \boldsymbol{v}_i^0 \rangle > 0$ and $\langle \boldsymbol{u}_j^0, \boldsymbol{y} \rangle > \langle \boldsymbol{v}_i^0, \boldsymbol{y} \rangle > 0$ for some $\boldsymbol{y} \in \mathbb{R}^d$ with $\|\boldsymbol{y}\| \leq 1$ initially. The creator is always recommended to the user (so the updates are $\boldsymbol{u}_j^{t+1} = \mathcal{P}(\boldsymbol{u}_j^t + \eta_u f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t) \boldsymbol{v}_i^t)$ and $\boldsymbol{v}_i^{t+1} = \mathcal{P}(\boldsymbol{v}_i^t + \eta_c \boldsymbol{u}_j^t))$. Suppose $\eta_c \leq \frac{\eta_u L_f}{2}$ and $0 \leq \eta_u < \frac{1}{2}$. Then, we have:*

- *$\langle \boldsymbol{u}_j^t, \boldsymbol{y} \rangle > \langle \boldsymbol{v}_i^t, \boldsymbol{y} \rangle > 0$ for all $t \geq 1$.*

- *Suppose $\langle \boldsymbol{u}_j^0, \boldsymbol{y} \rangle - \langle \boldsymbol{v}_i^0, \boldsymbol{y} \rangle = D > 0$. For any $R < D$, after $T = \frac{8}{3 \eta_u L_f} \ln \frac{2}{R^2}$ steps, we have $\langle \boldsymbol{v}_i^T, \boldsymbol{y} \rangle - \langle \boldsymbol{v}_i^0, \boldsymbol{y} \rangle \geq \frac{\eta_c}{\eta_u + \eta_c} (D - R)$.*

Proof. We prove the first item by induction. Suppose $\langle \boldsymbol{u}_j^t, \boldsymbol{y} \rangle > \langle \boldsymbol{v}_i^t, \boldsymbol{y} \rangle > 0$ holds. Consider $t + 1$. First, by Lemma D.2, $\langle \boldsymbol{v}_i^{t+1}, \boldsymbol{y} \rangle > 0$ holds. Then, we prove $\langle \boldsymbol{u}_j^{t+1}, \boldsymbol{y} \rangle > \langle \boldsymbol{v}_i^{t+1}, \boldsymbol{y} \rangle$. Let $f = f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t)$.

$$
\begin{aligned}
\langle \boldsymbol{u}_j^{t+1}, \boldsymbol{y} \rangle - \langle \boldsymbol{v}_i^{t+1}, \boldsymbol{y} \rangle &= \left\langle \frac{\boldsymbol{u}_j^t + \eta_u f \boldsymbol{v}_i^t}{\|\boldsymbol{u}_j^t + \eta_u f \boldsymbol{v}_i^t\|_2}, \boldsymbol{y} \right\rangle - \left\langle \frac{\boldsymbol{v}_i^t + \eta_c \boldsymbol{u}_j^t}{\|\boldsymbol{v}_i^t + \eta_c \boldsymbol{u}_j^t\|_2}, \boldsymbol{y} \right\rangle \\
&= \left(\frac{1}{\|\boldsymbol{u}_j^t + \eta_u f \boldsymbol{v}_i^t\|_2} - \frac{\eta_c}{\|\boldsymbol{v}_i^t + \eta_c \boldsymbol{u}_j^t\|_2} \right) \langle \boldsymbol{u}_j^t, \boldsymbol{y} \rangle - \left(\frac{1}{\|\boldsymbol{v}_i^t + \eta_c \boldsymbol{u}_j^t\|_2} - \frac{\eta_u f}{\|\boldsymbol{u}_j^t + \eta_u f \boldsymbol{v}_i^t\|_2} \right) \langle \boldsymbol{v}_i^t, \boldsymbol{y} \rangle \\
&> \left(\frac{1}{\|\boldsymbol{u}_j^t + \eta_u f \boldsymbol{v}_i^t\|_2} - \frac{\eta_c}{\|\boldsymbol{v}_i^t + \eta_c \boldsymbol{u}_j^t\|_2} \right) \langle \boldsymbol{v}_i^t, \boldsymbol{y} \rangle - \left(\frac{1}{\|\boldsymbol{v}_i^t + \eta_c \boldsymbol{u}_j^t\|_2} - \frac{\eta_u f}{\|\boldsymbol{u}_j^t + \eta_u f \boldsymbol{v}_i^t\|_2} \right) \langle \boldsymbol{v}_i^t, \boldsymbol{y} \rangle \\
&= \left(\frac{1 + \eta_u f}{\|\boldsymbol{u}_j^t + \eta_u f \boldsymbol{v}_i^t\|_2} - \frac{1 + \eta_c}{\|\boldsymbol{v}_i^t + \eta_c \boldsymbol{u}_j^t\|_2} \right) \langle \boldsymbol{v}_i^t, \boldsymbol{y} \rangle \\
&= \left(\frac{1 + \eta_u f}{\sqrt{1 + 2\eta_u f \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle + (\eta_u f)^2}} - \frac{1 + \eta_c}{\sqrt{1 + 2\eta_c \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle + (\eta_c)^2}} \right) \langle \boldsymbol{v}_i^t, \boldsymbol{y} \rangle.
\end{aligned}
$$

Let $a = \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle \leq 1$. We note that the function

$$
h(\eta) = \frac{1 + \eta}{\sqrt{1 + 2\eta a + \eta^2}} = \sqrt{\frac{1 + 2\eta + \eta^2}{1 + 2\eta a + \eta^2}} = \sqrt{1 + \frac{(2 - 2a)\eta}{1 + 2\eta a + \eta^2}} = \sqrt{1 + \frac{2(1 - a)}{\frac{1}{\eta} + 2a + \eta}}
$$

is increasing in $\eta \in [0, 1]$. Under the assumption of $\eta_c \leq \frac{\eta_u L_f}{2} \leq \frac{\eta_u f}{2} < \eta_u f$, we have $h(\eta_c) \leq h(\eta_u f)$ and hence

$$
\langle \boldsymbol{u}_j^{t+1}, \boldsymbol{y} \rangle - \langle \boldsymbol{v}_i^{t+1}, \boldsymbol{y} \rangle > \left(h(\eta_u f) - h(\eta_c) \right) \langle \boldsymbol{v}_i^t, \boldsymbol{y} \rangle \geq 0.
$$

We then prove the second item. Using Lemma D.3 for $\boldsymbol{v}_i^{t+1} = \mathcal{P}(\boldsymbol{v}_i^t + \eta_c \boldsymbol{u}_j^t)$, we get

$$
\langle \boldsymbol{v}_i^{t+1} - \boldsymbol{v}_i^t, \boldsymbol{y} \rangle \geq \frac{\eta_c}{1 + \eta_c} \left(\langle \boldsymbol{u}_j^t, \boldsymbol{y} \rangle - \langle \boldsymbol{v}_i^t, \boldsymbol{y} \rangle \right).
$$

Using Lemma D.3 for $\boldsymbol{u}_j^{t+1} = \mathcal{P}(\boldsymbol{u}_j^t + \eta_u f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t) \boldsymbol{v}_i^t)$ and using the fact $\langle \boldsymbol{v}_i^t, \boldsymbol{y} \rangle - \langle \boldsymbol{u}_j^t, \boldsymbol{y} \rangle < 0$ proved in item 1,

$$
\langle \boldsymbol{u}_j^{t+1} - \boldsymbol{u}_j^t, \boldsymbol{y} \rangle \geq \frac{\eta_u f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t)}{1 + \eta_u f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t)} \left(\langle \boldsymbol{v}_i^t, \boldsymbol{y} \rangle - \langle \boldsymbol{u}_j^t, \boldsymbol{y} \rangle \right) \geq \frac{\eta_u}{1 + \eta_u} \left(\langle \boldsymbol{v}_i^t, \boldsymbol{y} \rangle - \langle \boldsymbol{u}_j^t, \boldsymbol{y} \rangle \right).
$$

Rearranging the above two inequalities:

$$
\frac{1 + \eta_c}{\eta_c} \left(\langle \boldsymbol{v}_i^{t+1}, \boldsymbol{y} \rangle - \langle \boldsymbol{v}_i^t, \boldsymbol{y} \rangle \right) \geq \langle \boldsymbol{u}_j^t, \boldsymbol{y} \rangle - \langle \boldsymbol{v}_i^t, \boldsymbol{y} \rangle;
$$

$$
\frac{1 + \eta_u}{\eta_u} \left(\langle \boldsymbol{u}_j^{t+1}, \boldsymbol{y} \rangle - \langle \boldsymbol{u}_j^t, \boldsymbol{y} \rangle \right) \geq \langle \boldsymbol{v}_i^t, \boldsymbol{y} \rangle - \langle \boldsymbol{u}_j^t, \boldsymbol{y} \rangle.
$$

Summing the above two inequalities over $t = 0, 1, \ldots, T-1$:

$$\frac{1+\eta_c}{\eta_c}\Big(\langle \boldsymbol{v}_i^T, \boldsymbol{y}\rangle - \langle \boldsymbol{v}_i^0, \boldsymbol{y}\rangle\Big) + \frac{1+\eta_u}{\eta_u}\Big(\langle \boldsymbol{u}_j^T, \boldsymbol{y}\rangle - \langle \boldsymbol{u}_j^0, \boldsymbol{y}\rangle\Big) \geq 0. \tag{6}$$

According to Lemma F.1, after at most $T = \frac{8}{3\eta_u L_f} \ln \frac{2}{R^2}$ steps, we have $\|\boldsymbol{u}_j^T - \boldsymbol{v}_i^T\|_2 \leq R$. This implies $\langle \boldsymbol{u}_j^T, \boldsymbol{y}\rangle - \langle \boldsymbol{v}_i^T, \boldsymbol{y}\rangle = \langle \boldsymbol{u}_j^T - \boldsymbol{v}_i^T, \boldsymbol{y}\rangle \leq \|\boldsymbol{u}_j^T - \boldsymbol{v}_i^T\| \leq R$ and hence

$$\Big(\langle \boldsymbol{v}_i^T, \boldsymbol{y}\rangle - \langle \boldsymbol{v}_i^0, \boldsymbol{y}\rangle\Big) - \Big(\langle \boldsymbol{u}_j^T, \boldsymbol{y}\rangle - \langle \boldsymbol{u}_j^0, \boldsymbol{y}\rangle\Big) = \Big(\langle \boldsymbol{u}_j^0, \boldsymbol{y}\rangle - \langle \boldsymbol{v}_i^0, \boldsymbol{y}\rangle\Big) - \Big(\langle \boldsymbol{u}_j^T, \boldsymbol{y}\rangle - \langle \boldsymbol{v}_i^T, \boldsymbol{y}\rangle\Big) \geq D - R. \tag{7}$$

Multiplying (7) by $\frac{1+\eta_u}{\eta_u}$ and adding to (6):

$$\Big(\frac{1+\eta_c}{\eta_c} + \frac{1+\eta_u}{\eta_u}\Big)\Big(\langle \boldsymbol{v}_i^T, \boldsymbol{y}\rangle - \langle \boldsymbol{v}_i^0, \boldsymbol{y}\rangle\Big) \geq \frac{1+\eta_u}{\eta_u}(D - R).$$

This implies

$$\langle \boldsymbol{v}_i^T, \boldsymbol{y}\rangle - \langle \boldsymbol{v}_i^0, \boldsymbol{y}\rangle \geq \frac{\frac{1+\eta_u}{\eta_u}}{\frac{1+\eta_c}{\eta_c} + \frac{1+\eta_u}{\eta_u}}(D - R) = \frac{\eta_c(1+\eta_u)}{\eta_u(1+\eta_c) + \eta_c(1+\eta_u)}(D - R) \geq \frac{\eta_c}{\eta_u + \eta_c}(D - R).$$

given $\eta_c \leq \eta_u$. $\qquad\square$

The following lemma shows that, when we *reflect* some of the feature vectors in a system $(U^t, V^t) = (\{\boldsymbol{u}_j^t\}_{j\in[m]}, \{\boldsymbol{v}_i^t\}_{i\in[n]})$, there is a correspondence between the behaviors of the system with the reflected vectors and the original system.

Lemma D.7 (Reflection). *Let $(U^t, V^t) = (\{\boldsymbol{u}_j^t\}_{j\in[m]}, \{\boldsymbol{v}_i^t\}_{i\in[n]})$ be a system of m users and n creators with impact functions f, g. Let $a_i, b_j \in \{+1, -1\}$, $\forall i \in [n]$, $\forall i \in [m]$ be some binary constants. Define:*

$$\tilde{\boldsymbol{u}}_j^t = b_j \boldsymbol{u}_j^t = \pm \tilde{\boldsymbol{u}}_j^t, \qquad \tilde{\boldsymbol{v}}_i^t = a_i \boldsymbol{v}_i^t = \pm \tilde{\boldsymbol{v}}_i^t.$$

and impact functions

$$\tilde{f}(\tilde{\boldsymbol{v}}_i, \tilde{\boldsymbol{u}}_j) = a_i b_j f(\boldsymbol{v}_i, \boldsymbol{u}_j), \qquad \tilde{g}(\tilde{\boldsymbol{u}}_j, \tilde{\boldsymbol{v}}_i) = a_i b_j g(\boldsymbol{u}_j, \boldsymbol{v}_i).$$

Then:

- *There is a "correspondence" between the evolution of the system (U^t, V^t) with impact functions f, g and the evolution of the system $(\tilde{U}^t, \tilde{V}^t) = (\{\tilde{\boldsymbol{u}}_j^t\}_{j\in[m]}, \{\tilde{\boldsymbol{v}}_i^t\}_{i\in[n]})$ with impact functions $\tilde{f}, \tilde{g}$. Formally, suppose every user is recommended the same creator in the two systems, then the updated vectors in the two systems still satisfy the relations: $\tilde{\boldsymbol{u}}_j^{t+1} = b_j \boldsymbol{u}_j^{t+1}$, $\tilde{\boldsymbol{v}}_i^{t+1} = a_i \boldsymbol{v}_i^{t+1}$.*

- *If the system $(\tilde{U}^t, \tilde{V}^t)$ is in R-bi-polarization, then the original system (U^t, V^t) is also in R-bi-polarization.*

Proof. Consider the first item. Suppose user i is recommended creator j at time step t in the two systems. Then by definition, the updated user vectors in the two systems satisfy

$$\begin{aligned}
\tilde{\boldsymbol{u}}_j^{t+1} &= \mathcal{P}\big(\tilde{\boldsymbol{u}}_j^t + \eta_u \tilde{f}(\tilde{\boldsymbol{v}}_i^t, \tilde{\boldsymbol{u}}_j^t)\tilde{\boldsymbol{v}}_i^t\big) = \mathcal{P}\big(b_j \boldsymbol{u}_j^t + \eta_u a_i b_j f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t) a_i \boldsymbol{v}_i^t\big) \\
&= \mathcal{P}\big(b_j \boldsymbol{u}_j^t + \eta_u b_j f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t)\boldsymbol{v}_i^t\big) = b_j \mathcal{P}\big(\boldsymbol{u}_j^t + \eta_u f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t)\boldsymbol{v}_i^t\big) = b_j \boldsymbol{u}_j^{t+1}
\end{aligned}$$

Suppose creator i is recommended to the set of users J at time step t in the two systems. Then,

$$\begin{aligned}
\tilde{\boldsymbol{v}}_i^{t+1} &= \mathcal{P}\big(\tilde{\boldsymbol{v}}_i^t + \tfrac{\eta_c}{|J|}\sum_{j\in J} g(\tilde{\boldsymbol{u}}_j^t, \tilde{\boldsymbol{v}}_i^t)\tilde{\boldsymbol{u}}_j^t\big) \\
&= \mathcal{P}\big(a_i \boldsymbol{v}_i^t + \tfrac{\eta_c}{|J|}\sum_{j\in J} a_i b_j g(\boldsymbol{u}_j^t, \boldsymbol{v}_i^t) b_j \boldsymbol{u}_j^t\big) \\
&= \mathcal{P}\big(a_i \boldsymbol{v}_i^t + \tfrac{\eta_c}{|J|}\sum_{j\in J} a_i g(\boldsymbol{u}_j^t, \boldsymbol{v}_i^t)\boldsymbol{u}_j^t\big) \\
&= a_i \mathcal{P}\big(\boldsymbol{v}_i^t + \tfrac{\eta_c}{|J|}\sum_{j\in J} g(\boldsymbol{u}_j^t, \boldsymbol{v}_i^t)\boldsymbol{u}_j^t\big) = a_i \boldsymbol{v}_i^{t+1}.
\end{aligned}$$

This means that the evolution of the system $(\tilde{U}^t, \tilde{V}^t)$ has a correspondence to the evolution of the original system (U^t, V^t).

Consider the second item. Suppose $(\tilde{U}^t, \tilde{V}^t)$ is in R-bi-polarization, so $\tilde{v}_i^t = \pm v_i^t$ is R-close to $\pm c$ and $\tilde{u}_j^t = \pm u_j^t$ is R-close to $\pm c$ with some vector $c \in \mathbb{S}^{d-1}$. This implies that v_i^t is R-close to $\pm c$ and u_j^t is R-close to $\pm c$. So, the system (U^t, V^t) satisfies R-bi-polarization. $\qquad\square$

E Proof of Proposition 3.2

Proof. Let $(\boldsymbol{U}^t, \boldsymbol{V}^t)$ be an $(R, \boldsymbol{c})$-bi-polarization state with $R \in [0, 1]$ and $\boldsymbol{c} \in \mathbb{S}^{d-1}$, where all $\boldsymbol{u}_j^t$ and $\boldsymbol{v}_i^t$ are within distance R to $+\boldsymbol{c}$ or $-\boldsymbol{c}$. We show that, after one step of update, $\boldsymbol{u}_j^{t+1}$ and $\boldsymbol{v}_i^{t+1}$ are still within distance R to $+\boldsymbol{c}$ or $-\boldsymbol{c}$, so $(\boldsymbol{U}^{t+1}, \boldsymbol{V}^{t+1})$ still satisfies $(R, \boldsymbol{c})$-bi-polarization.

Consider $\boldsymbol{u}_j^t$. Without loss of generality, suppose $\boldsymbol{u}_j^t$ is close to $+\boldsymbol{c}$, so $\|\boldsymbol{u}_j^t - \boldsymbol{c}\|_2 \leq R$. Suppose user j is recommended creator i at step t. Let $\tilde{\boldsymbol{v}}_i^t = \boldsymbol{v}_i^t$ if $\langle \boldsymbol{v}_i^t, \boldsymbol{u}_j^t \rangle \geq 0$ and $\tilde{\boldsymbol{v}}_i^t = -\boldsymbol{v}_i^t$ if $\langle \boldsymbol{v}_i^t, \boldsymbol{u}_j^t \rangle < 0$. Then, the user update is

$$\boldsymbol{u}_j^{t+1} = \mathcal{P}\Big(\boldsymbol{u}_j^t + \eta_u f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t)\boldsymbol{v}_i^t\Big) = \mathcal{P}\Big(\boldsymbol{u}_j^t + \eta_u |f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t)|\tilde{\boldsymbol{v}}_i^t\Big).$$

Since $\tilde{\boldsymbol{v}}_i^t$ is close to $+\boldsymbol{c}$ or $-\boldsymbol{c}$, $\langle \tilde{\boldsymbol{v}}_i^t, \boldsymbol{u}_j^t \rangle > 0$, and $\boldsymbol{u}_j^t$ is close to $+\boldsymbol{c}$, it must be that $\tilde{\boldsymbol{v}}_i^t$ is close to $+\boldsymbol{c}$, so $\|\tilde{\boldsymbol{v}}_i^t - \boldsymbol{c}\|_2 \leq R$. Then, since $\boldsymbol{u}_j^{t+1}$ is the normalization of a vector in the convex cone formed by $\boldsymbol{u}_j^t$ and $\tilde{\boldsymbol{v}}_i^t$, by Lemma D.2, we have

$$\|\boldsymbol{u}_j^{t+1} - \boldsymbol{c}\|_2 \ \leq \ \max\big\{\|\boldsymbol{u}_j^t - \boldsymbol{c}\|_2, \|\tilde{\boldsymbol{v}}_i^t - \boldsymbol{c}\|_2\big\} \ \leq \ R.$$

Consider $\boldsymbol{v}_i^t$. Suppose $\|\boldsymbol{v}_i^t - \boldsymbol{c}\|_2 \leq R$. Let J be the set of users that are recommended creator i at step t. For each $j \in J$, let $\tilde{\boldsymbol{u}}_j^t = \boldsymbol{u}_j^t$ if $\langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle \geq 0$ and $\tilde{\boldsymbol{u}}_j^t = -\boldsymbol{u}_j^t$ if $\langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle < 0$. Then, the creator update is

$$\boldsymbol{v}_i^{t+1} = \mathcal{P}\Big(\boldsymbol{v}_i^t + \frac{\eta_c}{|J|}\sum_{j \in J} g(\boldsymbol{u}_j^t, \boldsymbol{v}_i^t)\boldsymbol{u}_j^t\Big) = \mathcal{P}\Big(\boldsymbol{v}_i^t + \frac{\eta_c}{|J|}\sum_{j \in J} |g(\boldsymbol{u}_j^t, \boldsymbol{v}_i^t)|\tilde{\boldsymbol{u}}_j^t\Big).$$

We note that every $\tilde{\boldsymbol{u}}_j^t$ satisfies $\|\tilde{\boldsymbol{u}}_j^t - \boldsymbol{c}\|_2 \leq R$ (by the same reasoning as above). Then, since $\boldsymbol{v}_i^{t+1}$ is the normalization of a vector in the convex cone formed by $\boldsymbol{v}_i^t$ and $\{\tilde{\boldsymbol{u}}_j^t\}_{j \in J}$, by Lemma D.2, we have

$$\|\boldsymbol{v}_i^{t+1} - \boldsymbol{c}\|_2 \ \leq \ \min\Big\{\|\boldsymbol{v}_i^t - \boldsymbol{c}\|_2, \ \min_{j \in J}\|\tilde{\boldsymbol{u}}_j^t - \boldsymbol{c}\|_2\Big\} \ \leq \ R. \qquad\square$$

F Proof of Lemma 3.4

Lemma 3.4 is proved by induction on the number n of creators. We first show that any system with 1 creator and multiple users must converge to R-bi-polarization in finite steps for any $R > 0$. Using the result for 1 creator, we then construct a finite length path that leads to R-bi-polarization for any system with $n \geq 2$ creators.

F.1 Base Case: Convergence Results for $n = 1$ Creator

We prove some convergence results for the special case of only one creator. This will serve as the basis for the proof for $n \geq 2$ creators. Recall that we have the following dynamics update rule:

- User: $\boldsymbol{u}_j^{t+1} = \mathcal{P}(\boldsymbol{u}_j^t + \eta_u f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t)\boldsymbol{v}_i^t)$ where $\boldsymbol{v}_i^t$ is the creator recommended to user j; $f(\boldsymbol{v}_i, \boldsymbol{u}_j)$ satisfies:

$$f(\boldsymbol{v}_i, \boldsymbol{u}_j) \text{ is } \begin{cases} > 0 & \text{if } \langle \boldsymbol{v}_i, \boldsymbol{u}_j \rangle > 0 \\ < 0 & \text{if } \langle \boldsymbol{v}_i, \boldsymbol{u}_j \rangle < 0 \\ = 0 & \text{if } \langle \boldsymbol{v}_i, \boldsymbol{u}_j \rangle = 0. \end{cases} \tag{8}$$

- Creator: $\boldsymbol{v}_i^{t+1} = \mathcal{P}(\boldsymbol{v}_j^t + \frac{\eta_c}{|J|}\sum_{j \in J} g(\boldsymbol{u}_j^t, \boldsymbol{v}_i^t)\boldsymbol{u}_j^t)$ where J is the set of users being recommended creator i.

Lemma F.1. *Consider a system of 1 creator $\boldsymbol{v}_i^t$ and $|J|$ users $\{\boldsymbol{u}_j^t\}_{j\in J}$, where the creator is recommended to all users at every time step. Assume:*

- *Initially, $\forall j \in J, \langle \boldsymbol{u}_j^0, \boldsymbol{v}_i^0 \rangle > 0$.*

- *There exists some constant $L_f > 0$ such that $f(\boldsymbol{v}_i, \boldsymbol{u}_j) \geq L_f > 0$ whenever $\langle \boldsymbol{v}_i, \boldsymbol{u}_j \rangle > 0$.*

- *$g(\boldsymbol{u}_j, \boldsymbol{v}_i) = 1$ when $\langle \boldsymbol{u}_j, \boldsymbol{v}_i \rangle > 0$.*

- *$\eta_c \leq \frac{\eta_u L_f}{2}$ and $0 \leq \eta_u < \frac{1}{2}$.*

Then, for any $R > 0$, after at most $\frac{8}{3\eta_u L_f} \ln \frac{2|J|}{R^2}$ steps, $\sum_{j\in J} \|\boldsymbol{u}_j^t - \boldsymbol{v}_i^t\|_2^2 \leq R^2$ will hold forever. In particular, each user vector will satisfy $\|\boldsymbol{u}_j^t - \boldsymbol{v}_i^t\|_2 \leq R$.

Proof. We first note that, by Lemma D.5, all user vectors satisfy $\langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle > 0$ in all time steps $t > 0$. Hence, the creator update is always $\boldsymbol{v}_i^{t+1} = \mathcal{P}(\boldsymbol{v}_i^t + \frac{\eta_c}{|J|} \sum_{j\in J} g(\boldsymbol{u}_j^t, \boldsymbol{v}_i^t)\boldsymbol{u}_j^t) = \mathcal{P}(\boldsymbol{v}_i^t + \eta_c \frac{1}{|J|} \sum_{j\in J} \boldsymbol{u}_j^t)$.

Let $a_t = 1/(1 - \frac{3\eta_u L_f}{8})^t$. Define the following potential function:

$$\Phi^t \;=\; a_t \sum_{j\in J} \frac{1}{2}\|\boldsymbol{u}_j^t - \boldsymbol{v}_i^t\|_2^2 \;=\; a_t \sum_{j\in J} \left(1 - \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle\right). \tag{9}$$

We will show that Φ^t is monotonically decreasing. Take the difference between Φ^{t+1} and Φ^t:

$$\Phi^{t+1} - \Phi^t \;=\; a_{t+1} \sum_{j\in J} \left(\langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle - \langle \boldsymbol{u}_j^{t+1}, \boldsymbol{v}_i^{t+1} \rangle\right) \;+\; (a_{t+1} - a_t) \sum_{j\in J} \left(1 - \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle\right)$$

$$=\; a_{t+1}\left(\sum_{j\in J}\langle \boldsymbol{v}_i^t, \boldsymbol{u}_j^t - \boldsymbol{u}_j^{t+1} \rangle + \sum_{j\in J}\langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t - \boldsymbol{v}_i^{t+1} \rangle + \sum_{j\in J}\langle \boldsymbol{u}_j^{t+1} - \boldsymbol{u}_j^t, \boldsymbol{v}_i^t - \boldsymbol{v}_i^{t+1} \rangle\right)$$

$$+\; (a_{t+1} - a_t) \sum_{j\in J} \left(1 - \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle\right).$$

Using Lemma D.3 with $\boldsymbol{x}^t = \boldsymbol{u}_j^t$, $\boldsymbol{z}^t = \boldsymbol{v}_i^t$, and $\eta = \eta_u f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t)$, we get

$$\langle \boldsymbol{v}_i^t, \boldsymbol{u}_j^t - \boldsymbol{u}_j^{t+1} \rangle \;\leq\; -\frac{\eta_u f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t)}{1 + \eta_u f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t)}\left(1 - \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle\right) \;\leq\; -\frac{\eta_u L_f}{2}\left(1 - \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle\right).$$

Using Lemma D.4 with $\boldsymbol{x}^t = \boldsymbol{u}_j^t$, $\boldsymbol{z}^t = \boldsymbol{v}_i^t$, and $\eta = \eta_u f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t)$, we get

$$\langle \boldsymbol{v}_i^t, \boldsymbol{u}_j^t - \boldsymbol{u}_j^{t+1} \rangle \;\leq\; -\frac{1}{\eta_u f(\boldsymbol{v}_i^t, \boldsymbol{u}_j^t)}\|\boldsymbol{u}_j^{t+1} - \boldsymbol{u}_j^t\|_2^2 \;\leq\; -\frac{1}{\eta_u}\|\boldsymbol{u}_j^{t+1} - \boldsymbol{u}_j^t\|_2^2.$$

Using Lemma D.4 with $\boldsymbol{x}^t = \boldsymbol{v}_i^t$, $\boldsymbol{z}^t = \frac{1}{|J|} \sum_{j\in J} \boldsymbol{u}_j^t$, and $\eta = \eta_c$, we get

$$\sum_{j\in J}\langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t - \boldsymbol{v}_i^{t+1} \rangle \;=\; |J|\langle \frac{1}{|J|}\sum_{j\in J}\boldsymbol{u}_j^t, \boldsymbol{v}_i^t - \boldsymbol{v}_i^{t+1} \rangle \;\leq\; -\frac{|J|}{\eta_c}\|\boldsymbol{v}_i^{t+1} - \boldsymbol{v}_i^t\|_2^2.$$

125688

Using the above three inequalities, we can upper bound $\Phi^{t+1} - \Phi^t$:

$$\Phi^{t+1} - \Phi^t$$

$$= a_{t+1}\left(\frac{3}{4}\sum_{j\in J}\langle \boldsymbol{v}_i^t, \boldsymbol{u}_j^t - \boldsymbol{u}_j^{t+1}\rangle \;+\; \frac{1}{4}\sum_{j\in J}\langle \boldsymbol{v}_i^t, \boldsymbol{u}_j^t - \boldsymbol{u}_j^{t+1}\rangle \right.$$

$$\left. + \sum_{j\in J}\langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t - \boldsymbol{v}_i^{t+1}\rangle \;+\; \sum_{j\in J}\langle \boldsymbol{u}_j^{t+1} - \boldsymbol{u}_j^t, \boldsymbol{v}_i^t - \boldsymbol{v}_i^{t+1}\rangle \right) \;+\; (a_{t+1}-a_t)\sum_{j\in J}\left(1 - \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t\rangle\right)$$

$$\le a_{t+1}\left(-\frac{3}{4}\sum_{j\in J}\frac{\eta_u L_f}{2}\left(1 - \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t\rangle\right) \;-\; \frac{1}{4}\sum_{j\in J}\frac{1}{\eta_u}\|\boldsymbol{u}_j^{t+1} - \boldsymbol{u}_j^t\|_2^2 \right.$$

$$\left. -\frac{|J|}{\eta_c}\|\boldsymbol{v}_i^{t+1} - \boldsymbol{v}_i^t\|_2^2 \;+\; \sum_{j\in J}\|\boldsymbol{u}_j^{t+1} - \boldsymbol{u}_j^t\|_2 \cdot \|\boldsymbol{v}_i^{t+1} - \boldsymbol{v}_i^t\|_2 \right) \;+\; (a_{t+1}-a_t)\sum_{j\in J}\left(1 - \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t\rangle\right)$$

$$= a_{t+1}\left(-\frac{3\eta_u L_f}{8}\sum_{j\in J}\left(1 - \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t\rangle\right) \right.$$

$$\left. - \sum_{j\in J}\Big(\underbrace{\frac{1}{4\eta_u}\|\boldsymbol{u}_j^{t+1} - \boldsymbol{u}_j^t\|_2^2 + \frac{1}{\eta_c}\|\boldsymbol{v}_i^{t+1} - \boldsymbol{v}_i^t\|_2^2 \;-\; \|\boldsymbol{u}_j^{t+1} - \boldsymbol{u}_j^t\|_2 \cdot \|\boldsymbol{v}_i^{t+1} - \boldsymbol{v}_i^t\|_2}_{\geq 2\sqrt{\frac{1}{4\eta_u \eta_c}\|\boldsymbol{u}_j^{t+1}-\boldsymbol{u}_j^t\|_2^2\|\boldsymbol{v}_i^{t+1}-\boldsymbol{v}_i^t\|_2^2}}\Big)\right)$$

$$+ (a_{t+1}-a_t)\sum_{j\in J}\left(1 - \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t\rangle\right)$$

$$\le a_{t+1}\left(-\frac{3\eta_u L_f}{8}\sum_{j\in J}\left(1 - \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t\rangle\right) \;-\; \sum_{j\in J}\Big(\underbrace{\sqrt{\tfrac{1}{\eta_u\eta_c}}-1}_{\geq 0}\Big)\|\boldsymbol{u}_j^{t+1} - \boldsymbol{u}_j^t\|_2 \cdot \|\boldsymbol{v}_i^{t+1} - \boldsymbol{v}_i^t\|_2 \right)$$

$$+ (a_{t+1}-a_t)\sum_{j\in J}\left(1 - \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t\rangle\right)$$

$$\le a_{t+1}\left(-\frac{3\eta_u L_f}{8}\sum_{j\in J}\left(1 - \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t\rangle\right) \;+\; 0 \right) \;+\; (a_{t+1}-a_t)\sum_{j\in J}\left(1 - \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t\rangle\right)$$

$$= \left(\left(1 - \frac{3\eta_u L_f}{8}\right)a_{t+1} - a_t\right)\sum_{j\in J}\left(1 - \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t\rangle\right)$$

$$= 0$$

where the last step is because $(1 - \frac{3\eta_u L_f}{8})a_{t+1} = a_t$.

We have shown that Φ^t is monotonically decreasing. Thus,

$$\frac{1}{2}\sum_{j\in J}\|\boldsymbol{u}_j^T - \boldsymbol{v}_i^T\|^2 \;=\; \frac{\Phi^T}{a_T} \;\le\; \frac{\Phi^0}{a_T} \;\le\; \frac{\sum_{j\in J}1}{a_T} \;=\; \left(1 - \frac{3\eta_u L_f}{8}\right)^T|J| \;\le\; e^{-\frac{3\eta_u L_f}{8}T}|J| \;\le\; \frac{1}{2}R^2$$

whenever $T \ge \frac{8}{3\eta_u L_f}\ln\frac{2|J|}{R^2}$. $\qquad\qquad\square$

Corollary F.2 (of Lemma F.1). *Consider a system of 1 creator $\boldsymbol{v}_i^t$ and $|J|$ users $\{\boldsymbol{u}_j^t\}_{j\in J}$, where the creator is recommended to all users at every time step. Assume:*

- *Initially, $\langle \boldsymbol{u}_j^0, \boldsymbol{v}_i^0\rangle \ne 0$ for every $j \in J$.*

- *There exists some constant $L_f > 0$ such that $|f(\boldsymbol{v}_i, \boldsymbol{u}_j)| \ge L_f > 0$ whenever $\langle \boldsymbol{v}_i, \boldsymbol{u}_j\rangle \ne 0$.*

- *$g(\boldsymbol{u}_j, \boldsymbol{v}_i) = \mathrm{sign}(\langle \boldsymbol{u}_j, \boldsymbol{v}_i\rangle)$.*

- *$\eta_c \le \frac{\eta_u L_f}{2}$ and $0 \le \eta_u < \frac{1}{2}$.*

Then, for any $R > 0$, after at most $\frac{8}{3\eta_u L_f}\ln\frac{2|J|}{R^2}$ steps, the system will reach R-bi-polarization.

Proof. Let $J^+ = \{j \in J : \langle \boldsymbol{u}_j^0, \boldsymbol{v}_i^0 \rangle > 0\}$ be the set of users with positive inner products with creator i initially; let $J^- = \{j \in J : \langle \boldsymbol{u}_j^0, \boldsymbol{v}_i^0 \rangle < 0\}$. Let $\tilde{\boldsymbol{u}}_j^t = -\boldsymbol{u}_j^t$ for $j \in J^-$ and $\tilde{\boldsymbol{u}}_j^t = \boldsymbol{u}_j^t$ for $j \in J^+$. Then, the system consisting of $\{\tilde{\boldsymbol{u}}_j^t\}_{j \in J}$ and $\boldsymbol{v}_i^t$ satisfies the initial condition $\langle \tilde{\boldsymbol{u}}_j^0, \boldsymbol{v}_i^0 \rangle > 0$ in Lemma F.1. So, by Lemma F.1, it reaches R-consensus after at most $\frac{8}{3\eta_u L_f} \ln \frac{2|J|}{R^2}$ steps. Then by the reflection lemma (Lemma D.7), the original system, consisting of $\{\boldsymbol{u}_j^t\}_{j \in J}$ and $\boldsymbol{v}_i^t$, must reach R-bi-polarization. $\qquad\square$

F.2 Inductive Step: Proof of Lemma 3.4

Lemma F.3. *Consider a system of $n \geq 1$ creators $\{\boldsymbol{v}_1^t, \ldots, \boldsymbol{v}_n^t\}$ and $|J|$ users $\{\boldsymbol{u}_j^t\}_{j \in J}$. Assume:*

- *Initially, $\langle \boldsymbol{v}_i^0, \boldsymbol{v}_{i'}^0 \rangle > 0$ for every i, i', and $\langle \boldsymbol{v}_i^0, \boldsymbol{u}_j^0 \rangle > 0$ for every i, j.*

- *Assumptions of Lemma F.1.*

Then, for any $R \in (0, 1)$, there exists a path of finite length that leads the initial state $(\boldsymbol{U}^0, \boldsymbol{V}^0)$ to R-consensus.

Proof. Fix any $R \in (0, 1)$. Choose R_1 such that $\sqrt{(\frac{\eta_u}{\eta_c} + 2) 4 R_1} = R$. Clearly, $R_1 < R$. We construct a path that leads the state $(\boldsymbol{U}^0, \boldsymbol{V}^0)$ to R-consensus as follows.

Step (1): Consider the subsystem of the first $n - 1$ creators and all users J. By induction, there exists a path of length $T_1 = L_{n-1, R_1} < +\infty$ that leads the subsystem to $(R_1, \boldsymbol{c}^{T_1})$-consensus with some $\boldsymbol{c}^{T_1} \in \mathbb{S}^{d-1}$. So, after these T_1 steps, all creators $i \in \{1, \ldots, n - 1\}$ and all users $j \in J$ satisfy $\|\boldsymbol{v}_i^{T_1} - \boldsymbol{c}^{T_1}\| \leq R_1$ and $\|\boldsymbol{u}_j^{T_1} - \boldsymbol{c}^{T_1}\| \leq R_1$. Creator n does not update during these T_1 steps, so $\boldsymbol{v}_n^{T_1} = \boldsymbol{v}_n^0$, and it still has positive inner products with the first $n - 1$ creators and all users by the convex cone property (Lemma D.2). Let's then consider the distance between creators n and the consensus center $\boldsymbol{c}^{T_1}$: $\|\boldsymbol{v}_n^{T_1} - \boldsymbol{c}^{T_1}\|$. If $\|\boldsymbol{v}_n^{T_1} - \boldsymbol{c}^{T_1}\| \leq R$, then the system has satisfied $(R, \boldsymbol{c}^{T_1})$-consensus, so our construction is finished. Otherwise, $\|\boldsymbol{v}_n^{T_1} - \boldsymbol{c}^{T_1}\| > R$. We continue the construction as follows:

Step (2): Pick any user $j_0 \in J$, recommend creator n to user j_0 for $T_2 = \frac{8}{3\eta_u L_f} \ln \frac{2}{R_1^2}$ steps, while recommending creator 1 to all other users. From the $(R_1, \boldsymbol{c}^{T_1})$-consensus in step (1) we know $\|\boldsymbol{u}_{j_0}^{T_1} - \boldsymbol{c}^{T_1}\| \leq R_1$, so

$$\langle \boldsymbol{u}_{j_0}^{T_1}, \boldsymbol{c}^{T_1} \rangle = 1 - \tfrac{1}{2}\|\boldsymbol{u}_{j_0}^{T_1} - \boldsymbol{c}^{T_1}\|^2 \geq 1 - \tfrac{R_1^2}{2} > 1 - \tfrac{R^2}{2} \geq 1 - \tfrac{1}{2}\|\boldsymbol{v}_n^{T_1} - \boldsymbol{c}^{T_1}\|^2 = \langle \boldsymbol{v}_2^{T_1}, \boldsymbol{c}^{T_1} \rangle.$$

Thus, we can apply Lemma D.6 with $\boldsymbol{y} = \boldsymbol{c}^{T_1}$ to derive that, after these T_2 steps,

$$
\begin{aligned}
\langle \boldsymbol{v}_n^{T_1 + T_2}, \boldsymbol{c}^{T_1} \rangle - \langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle &\geq \tfrac{\eta_c}{\eta_u + \eta_c}\left(\langle \boldsymbol{u}_{j_0}^{T_1}, \boldsymbol{c}^{T_1} \rangle - \langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle - R_1 \right) \\
&\geq \tfrac{\eta_c}{\eta_u + \eta_c}\left(1 - \tfrac{R_1^2}{2} - \langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle - R_1 \right).
\end{aligned}
$$
$$\implies \quad \langle \boldsymbol{v}_n^{T_1 + T_2}, \boldsymbol{c}^{T_1} \rangle \geq \langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle + \tfrac{\eta_c}{\eta_u + \eta_c}\left(1 - \tfrac{R_1^2}{2} - \langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle - R_1 \right). \tag{10}$$

For the inner product between creator n and user j_0, by Lemma F.1 $\|\boldsymbol{v}_n^{T_1 + T_2} - \boldsymbol{u}_{j_0}^{T_1 + T_2}\| \leq R_1$, so

$$\langle \boldsymbol{v}_n^{T_1 + T_2}, \boldsymbol{u}_{j_0}^{T_1 + T_2} \rangle = 1 - \tfrac{1}{2}\|\boldsymbol{v}_n^{T_1 + T_2} - \boldsymbol{u}_{j_0}^{T_1 + T_2}\|^2 \geq 1 - \tfrac{R_1^2}{2}. \tag{11}$$

Consider the inner products between creator n and the first $n - 1$ creators and the users in $J \setminus \{j_0\}$. Because the first $n - 1$ creators and the users in $J \setminus \{j_0\}$ form $(R_1, \boldsymbol{c}^{T_1})$-consensus at time step T_1, by Observation 3.2, they still form $(R_1, \boldsymbol{c}^{T_1})$-consensus at time step $T_1 + T_2$, so $\|\boldsymbol{v}_i^{T_1 + T_2} - \boldsymbol{c}^{T_1}\| \leq R_1$ and $\|\boldsymbol{u}_j^{T_1 + T_2} - \boldsymbol{c}^{T_1}\| \leq R_1$. This implies, for $i \neq n$,

$$
\begin{aligned}
\langle \boldsymbol{v}_n^{T_1 + T_2}, \boldsymbol{v}_i^{T_1 + T_2} \rangle &\geq \langle \boldsymbol{v}_n^{T_1 + T_2}, \boldsymbol{c}^{T_1} \rangle - \|\boldsymbol{v}_i^{T_1 + T_2} - \boldsymbol{c}^{T_1}\| \\
&\geq \langle \boldsymbol{v}_2^{T_1 + T_2}, \boldsymbol{c}^{T_1} \rangle - R_1 \\
(10) &\geq \langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle + \tfrac{\eta_c}{\eta_u + \eta_c}\left(1 - \tfrac{R_1^2}{2} - \langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle - R_1 \right) - R_1, \tag{12}
\end{aligned}
$$

and for $j \in J \setminus \{j_0\}$,

$$\langle \boldsymbol{v}_n^{T_1+T_2}, \boldsymbol{u}_j^{T_1+T_2} \rangle \geq \langle \boldsymbol{v}_n^{T_1+T_2}, \boldsymbol{c}^{T_1} \rangle - \|\boldsymbol{u}_j^{T_1+T_2} - \boldsymbol{c}^{T_1}\|$$

$$\geq \langle \boldsymbol{v}_2^{T_1+T_2}, \boldsymbol{c}^{T_1} \rangle - R_1$$

$$(10) \geq \langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle + \tfrac{\eta_c}{\eta_u+\eta_c}\left(1 - \tfrac{R_1^2}{2} - \langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle - R_1\right) - R_1. \qquad (13)$$

Step (3): Consider the subsystem of the first $n-1$ creators and all users J. By induction, there exists a path of length $T_3 = L_{n-1,R_1} < +\infty$ that leads the subsystem to $(R_1, \boldsymbol{c}^{T_1+T_2+T_3})$-consensus with some $\boldsymbol{c}^{T_1+T_2+T_3} \in \mathbb{S}^{d-1}$. So, we have $\|\boldsymbol{v}_i^{T_1+T_2+T_3} - \boldsymbol{c}^{T_1+T_2+T_3}\| \leq R_1$ for every $i \in \{1, \ldots, n-1\}$ and $\|\boldsymbol{u}_j^{T_1+T_2+T_3} - \boldsymbol{c}^{T_1+T_2+T_3}\| \leq R_1$ for every $j \in J$, and $\boldsymbol{v}_n^{T_1+T_2+T_3} = \boldsymbol{v}_n^{T_1+T_2}$. Consider the inner product between creator n and any of the first $n-1$ creators $i \in \{1, \ldots, n-1\}$. By the convex cone property (Lemma D.2),

$$\langle \boldsymbol{v}_n^{T_1+T_2+T_3}, \boldsymbol{v}_i^{T_1+T_2+T_3} \rangle = \langle \boldsymbol{v}_n^{T_1+T_2}, \boldsymbol{v}_i^{T_1+T_2+T_3} \rangle$$

$$\text{Lemma D.2} \geq \min\left\{\langle \boldsymbol{v}_n^{T_1+T_2}, \boldsymbol{v}_i^{T_1+T_2} \rangle, \min_{j \in J}\langle \boldsymbol{v}_n^{T_1+T_2}, \boldsymbol{u}_j^{T_1+T_2} \rangle\right\}$$

$$\text{by (11), (12), (13)} \geq \min\left\{\langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle + \tfrac{\eta_c}{\eta_u+\eta_c}\left(1 - \tfrac{R_1^2}{2} - \langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle - R_1\right) - R_1, \ 1 - \tfrac{R_1^2}{2}\right\}$$

$$= \langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle + \tfrac{\eta_c}{\eta_u+\eta_c}\left(1 - \tfrac{R_1^2}{2} - \langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle - R_1\right) - R_1 \qquad (14)$$

where the last equality is because, under the assumption of $\|\boldsymbol{v}_n^{T_1} - \boldsymbol{c}^{T_1}\| > R = \sqrt{(\tfrac{\eta_u}{\eta_c}+2)4R_1}$,

$$\langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle + \tfrac{\eta_c}{\eta_u+\eta_c}\left(1 - \tfrac{R_1^2}{2} - \langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle - R_1\right) - R_1$$

$$= \tfrac{\eta_u}{\eta_u+\eta_c}\langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle + \tfrac{\eta_c}{\eta_u+\eta_c}\left(1 - \tfrac{R_1^2}{2} - R_1\right) - R_1$$

$$= \tfrac{\eta_u}{\eta_u+\eta_c}\left(1 - \tfrac{1}{2}\|\boldsymbol{v}_2^{T_1} - \boldsymbol{c}^{T_1}\|^2\right) + \tfrac{\eta_c}{\eta_u+\eta_c}\left(1 - \tfrac{R_1^2}{2} - R_1\right) - R_1$$

$$\leq \tfrac{\eta_u}{\eta_u+\eta_c}\left(1 - \tfrac{1}{2}(\tfrac{\eta_u}{\eta_c}+2)4R_1\right) + \tfrac{\eta_c}{\eta_u+\eta_c}\left(1 - \tfrac{R_1^2}{2} - R_1\right) - R_1$$

$$\leq \max\{1 - \tfrac{1}{2}(\tfrac{\eta_u}{\eta_c}+2)4R_1, \ 1 - \tfrac{R_1^2}{2} - R_1\} - R_1$$

$$= 1 - \tfrac{R_1^2}{2} - R_1 - R_1 \leq 1 - \tfrac{R_1^2}{2}.$$

From (14) and $\|\boldsymbol{v}_i^{T_1+T_2+T_3} - \boldsymbol{c}^{T_1+T_2+T_3}\| \leq R_1$,

$$\langle \boldsymbol{v}_n^{T_1+T_2+T_3}, \boldsymbol{c}^{T_1+T_2+T_3} \rangle \geq \langle \boldsymbol{v}_n^{T_1+T_2}, \boldsymbol{v}_i^{T_1+T_2+T_3} \rangle - \|\boldsymbol{v}_i^{T_1+T_2+T_3} - \boldsymbol{c}^{T_1+T_2+T_3}\|$$

$$\geq \langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle + \tfrac{\eta_c}{\eta_u+\eta_c}\left(1 - \tfrac{R_1^2}{2} - \langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle - R_1\right) - 2R_1.$$

Using 1 to minus the above inequality, we obtain

$$1 - \langle \boldsymbol{v}_n^{T_1+T_2+T_3}, \boldsymbol{c}^{T_1+T_2+T_3} \rangle \leq \tfrac{\eta_u}{\eta_u+\eta_c}\left(1 - \langle \boldsymbol{v}_n^{T_1}, \boldsymbol{c}^{T_1} \rangle\right) + \tfrac{\eta_c}{\eta_u+\eta_c}\left(\tfrac{R_1^2}{2} + R_1\right) + 2R_1.$$

Let $F^t = 1 - \langle \boldsymbol{v}_n^t, \boldsymbol{c}^t \rangle$, then

$$F^{T_1+T_2+T_3} \leq \tfrac{\eta_u}{\eta_u+\eta_c}F^{T_1} + \tfrac{\eta_c}{\eta_u+\eta_c}\left(\tfrac{R_1^2}{2} + R_1\right) + 2R_1. \qquad (15)$$

Repeat steps (2) and (3) for K times. Then, using (15) for K times,

$$
\begin{aligned}
F^{T_1+K(T_2+T_3)} \;\leq\;& \tfrac{\eta_u}{\eta_u+\eta_c} F^{T_1+(K-1)(T_2+T_3)} + \tfrac{\eta_c}{\eta_u+\eta_c}\left(\tfrac{R_1^2}{2}+R_1\right)+2R_1 \\
\leq\;& \tfrac{\eta_u}{\eta_u+\eta_c}\left(\tfrac{\eta_u}{\eta_u+\eta_c}F^{T_1+(K-2)(T_2+T_3)}+\tfrac{\eta_c}{\eta_u+\eta_c}\left(\tfrac{R_1^2}{2}+R_1\right)+2R_1\right)+\tfrac{\eta_c}{\eta_u+\eta_c}\left(\tfrac{R_1^2}{2}+R_1\right)+2R_1 \\
&\vdots \\
\leq\;& \left(\tfrac{\eta_u}{\eta_u+\eta_c}\right)^K F^{T_1} + \left(1+\tfrac{\eta_u}{\eta_t+\eta_c}+\cdots+\left(\tfrac{\eta_u}{\eta_t+\eta_c}\right)^{K-1}\right)\left(\tfrac{\eta_c}{\eta_u+\eta_c}\left(\tfrac{R_1^2}{2}+R_1\right)+2R_1\right) \\
\leq\;& \left(\tfrac{\eta_u}{\eta_u+\eta_c}\right)^K \cdot 1 + \tfrac{1}{1-\frac{\eta_u}{\eta_u+\eta_c}}\left(\tfrac{\eta_c}{\eta_u+\eta_c}\left(\tfrac{R_1^2}{2}+R_1\right)+2R_1\right) \\
=\;& \left(\tfrac{\eta_u}{\eta_u+\eta_c}\right)^K + \tfrac{R_1^2}{2}+R_1+\tfrac{\eta_u+\eta_c}{\eta_c}2R_1 \\
\leq\;& \tfrac{R_1^2}{2}+\tfrac{R_1^2}{2}+R_1+\tfrac{\eta_u+\eta_c}{\eta_c}2R_1 \;\leq\; \left(\tfrac{\eta_u}{\eta_c}+2\right)2R_1,
\end{aligned}
$$

by choosing $K=\dfrac{\ln\frac{2}{R_1^2}}{\ln\frac{\eta_u+\eta_c}{\eta_u}}\leq\dfrac{\eta_u+\eta_c}{\eta_c}\ln\frac{2}{R_1^2}$. This means that, after repeating steps (2) and (3) for K times, we must have

$$
\begin{aligned}
\|\boldsymbol{v}_n^{T_1+K(T_2+T_3)}-\boldsymbol{c}^{T_1+K(T_2+T_3)}\| \;&=\; \sqrt{2\big(1-\langle\boldsymbol{v}_n^{T_1+K(T_2+T_3)},\boldsymbol{c}^{T_1+K(T_2+T_3)}\rangle\big)} \\
&=\; \sqrt{2F^{T_1+K(T_2+T_3)}} \;\leq\; \sqrt{2\big(\tfrac{\eta_u}{\eta_c}+2\big)2R_1} \;=\; R.
\end{aligned}
$$

The above inequality, together with the fact that other creators $i\neq n$ and all users in J already satisfy $(R_1\leq R, \boldsymbol{c}^{T_1+K(T_2+T_3)})$-consensus after step (3), implies that the whole system has reached $(R,\boldsymbol{c}^{T_1+K(T_2+T_3)})$-consensus.

The length of the path constructed above is at most:

$$
T_1+K(T_2+T_3) \;\leq\; L_{n-1,R_1}+\tfrac{\eta_u+\eta_c}{\eta_c}\ln\tfrac{2}{R_1^2}\left(\tfrac{8}{3\eta_u L_f}\ln\tfrac{2|J|}{R_1^2}+L_{n-1,R_1}\right) \;=\; L_{n,R} \;<\; +\infty,
$$

which is finite. $\qquad\square$

Lemma F.4. *Consider a subsystem of n creators $\{\boldsymbol{v}_1^t,\ldots,\boldsymbol{v}_n^t\}$ and $|J|$ users $\{\boldsymbol{u}_j^t\}_{j\in J}$. Assume:*

- *Initially, the first $n-1$ creators and all users are in R_0-consensus: $\|\boldsymbol{v}_i^0-\boldsymbol{c}\|\leq R_0$, $\|\boldsymbol{u}_j^0-\boldsymbol{c}\|\leq R_0$, with $0<R_0<\frac{\eta_c}{5(\eta_c+\eta_u)}$.*

- $\langle\boldsymbol{v}_n^0,\boldsymbol{u}_{j_0}^0\rangle>0$ *for some $j_0\in J$.*

- $g(\boldsymbol{u}_j,\boldsymbol{v}_i)=\mathrm{sign}(\langle\boldsymbol{u}_j,\boldsymbol{v}_i\rangle)$.

- *Assumption of Lemma F.1.*

Then, for any $R\in(0,1)$, there exists a path of finite length that leads the initial state $(\boldsymbol{U}^0,\boldsymbol{V}^0)$ to R-consensus.

Proof. First, we recommend creator n to user j_0 for $T=\frac{8}{3\eta_u L_f}\ln\frac{2}{R_0^2}$ steps, while recommending other creators to other users arbitrarily. Applying Lemma D.6 with $\boldsymbol{y}=\boldsymbol{u}_{j_0}^0$, we get

$$
\langle\boldsymbol{v}_n^T,\boldsymbol{u}_{j_0}^0\rangle-\langle\boldsymbol{v}_n^0,\boldsymbol{u}_{j_0}^0\rangle \;\geq\; \tfrac{\eta_c}{\eta_u+\eta_c}\Big(\langle\boldsymbol{u}_{j_0}^0,\boldsymbol{u}_{j_0}^0\rangle-\langle\boldsymbol{v}_n^0,\boldsymbol{u}_{j_0}^0\rangle-R_0\Big) \;=\; \tfrac{\eta_c}{\eta_u+\eta_c}\Big(1-\langle\boldsymbol{v}_n^0,\boldsymbol{u}_{j_0}^0\rangle-R_0\Big). \tag{16}
$$

On the other hand, because the first $n-1$ creators and all users in $J\setminus\{j_0\}$ form an $(R_0,\boldsymbol{c})$-consensus at time step 0, according to Observation 3.2, they still form an $(R_0,\boldsymbol{c})$-consensus at time step T, so $\|\boldsymbol{v}_i^T-\boldsymbol{c}\|\leq R_0$ for every $i\in\{1,\ldots,n-1\}$. This implies, for every $i\in\{1,\ldots,n-1\}$,

$$
\langle\boldsymbol{v}_n^T,\boldsymbol{v}_i^T\rangle-\langle\boldsymbol{v}_n^T,\boldsymbol{u}_{j_0}^0\rangle \;\geq\; -\|\boldsymbol{v}_i^T-\boldsymbol{u}_{j_0}^0\| \;\geq\; -\|\boldsymbol{v}_i^T-\boldsymbol{c}\|-\|\boldsymbol{c}-\boldsymbol{u}_{j_0}^0\| \;\geq\; -2R_0. \tag{17}
$$

Adding (16) and (17) and moving $\langle \boldsymbol{v}_n^0, \boldsymbol{u}_{j_0}^0 \rangle$ to the right side, we get

$$\begin{aligned}
\langle \boldsymbol{v}_n^T, \boldsymbol{v}_i^T \rangle &\geq \langle \boldsymbol{v}_n^0, \boldsymbol{u}_{j_0}^0 \rangle + \tfrac{\eta_c}{\eta_u+\eta_c}\left(1 - \langle \boldsymbol{v}_n^0, \boldsymbol{u}_{j_0}^0 \rangle - R_0\right) - 2R_0 \\
&= \tfrac{\eta_u}{\eta_u+\eta_c}\langle \boldsymbol{v}_n^0, \boldsymbol{u}_{j_0}^0 \rangle + \tfrac{\eta_c}{\eta_u+\eta_c}\left(1 - R_0\right) - 2R_0 \\
&> 0 + \tfrac{\eta_c}{\eta_u+\eta_c}\left(1 - R_0\right) - 2R_0 > 0,
\end{aligned}$$

under the condition of $R_0 < \frac{\eta_c}{5(\eta_u+\eta_c)}$. Moreover, for every $j \in J \setminus \{j_0\}$, because $\|\boldsymbol{u}_j^T - \boldsymbol{v}_i^T\| \leq \|\boldsymbol{u}_j^T - \boldsymbol{c}\| + \|\boldsymbol{c} - \boldsymbol{v}_i^T\| \leq 2R_0$,

$$\langle \boldsymbol{v}_n^T, \boldsymbol{u}_j^T \rangle \geq \langle \boldsymbol{v}_n^T, \boldsymbol{v}_i^T \rangle - \|\boldsymbol{u}_j^T - \boldsymbol{v}_i^T\| \geq \tfrac{\eta_c}{\eta_u+\eta_c}\left(1 - R_0\right) - 4R_0 > 0.$$

For j_0, by Lemma F.1, $\|\boldsymbol{v}_n^T - \boldsymbol{u}_{j_0}^T\| \leq R_0$, so

$$\langle \boldsymbol{v}_n^T, \boldsymbol{u}_{j_0}^T \rangle = 1 - \tfrac{1}{2}\|\boldsymbol{v}_n^T - \boldsymbol{u}_{j_0}^T\|^2 \geq 1 - \tfrac{R_0^2}{2} > 0.$$

For the inner product between any creator $i \in \{1, \ldots, n-1\}$ and the users:

$$\langle \boldsymbol{v}_i^T, \boldsymbol{u}_{j_0}^T \rangle \geq \langle \boldsymbol{v}_i^T, \boldsymbol{v}_n^T \rangle - \|\boldsymbol{v}_n^T - \boldsymbol{u}_{j_0}^T\| \geq \tfrac{\eta_c}{\eta_u+\eta_c}\left(1 - R_0\right) - 2R_0 - R_0 = \tfrac{\eta_c}{\eta_u+\eta_c}\left(1 - R_0\right) - 3R_0 > 0;$$

$$\forall j \in J \setminus \{j_0\}, \quad \langle \boldsymbol{v}_i^T, \boldsymbol{u}_j^T \rangle = 1 - \tfrac{1}{2}\|\boldsymbol{v}_i^T - \boldsymbol{u}_j^T\|^2 \geq 1 - \tfrac{1}{2}\left(\|\boldsymbol{v}_i^T - \boldsymbol{c}\| + \|\boldsymbol{c} - \boldsymbol{u}_j^T\|\right)^2 > 1 - \tfrac{1}{2}(2R_0)^2 > 0.$$

All of the "> 0" inequalities above show that the system of $\{\boldsymbol{v}_i^T\}_{i\in[n]}$ and $\{\boldsymbol{u}_j^T\}_{j\in J}$ satisfies the condition of Lemma F.3. So, there exists a path of finite length $T_2 < +\infty$ that leads the system to R-consensus by Lemma F.3. The total length of path $T + T_2 = \frac{8}{3\eta_u L_f}\ln\frac{2}{R_0^2} + T_2 < +\infty$ is finite. $\qquad\square$

Lemma 3.4. *Suppose $\eta_c \leq \frac{\eta_u L_f}{2}$ and $\eta_u < \frac{1}{2}$. For any $R > 0$, for almost every state $(\boldsymbol{U}^t, \boldsymbol{V}^t)$ in the state space, there exists a path $(\boldsymbol{U}^t, \boldsymbol{V}^t) \to (\boldsymbol{U}^{t+1}, \boldsymbol{V}^{t+1}) \to \cdots \to (\boldsymbol{U}^{t+T}, \boldsymbol{V}^{t+T})$ of finite length that leads to an R-bi-polarization state $(\boldsymbol{U}^{t+T}, \boldsymbol{V}^{t+T})$.*

Proof. We prove this lemma by induction on the number of creators n. The case for $n = 1$ directly follows from Corollary F.2 which shows that, for any system of $n = 1$ creator and $|J|$ users with no $\langle \boldsymbol{v}_i^0, \boldsymbol{u}_j^0 \rangle = 0$, there exists a path of length at most $L_1^R = \frac{8}{3\eta_u L_F}\ln\frac{2|J|}{R^2} < +\infty$ that leads to R-bi-polarization.

Consider $n \geq 2$. Consider the subsystem consisting of the first $n - 1$ creators $\{\boldsymbol{v}_1^t, \ldots, \boldsymbol{v}_{n-1}^t\}$ and all users. Let $R_0 = \frac{\eta_c}{6(\eta_c+\eta_u)}$. By induction, there exists a path of finite length $T_1 = L_{n-1}^{R_0} < +\infty$ that leads the subsystem to R_0-bi-polarization, with some vector $\boldsymbol{c}_0 \in \mathbb{S}^{d-1}$, so every $\boldsymbol{v}_i^{T_1}$ is R_0-close to $+\boldsymbol{c}_0$ or $-\boldsymbol{c}_0$, for $i \neq n$, and every $\boldsymbol{u}_j^{T_1}$ is R_0-close to $+\boldsymbol{c}_0$ or $-\boldsymbol{c}_0$. Define:

$$\tilde{\boldsymbol{v}}_i^t = \begin{cases} \boldsymbol{v}_i^t & \text{if } \boldsymbol{v}_i^{T_1} \text{ is } R_0\text{-close to } +\boldsymbol{c} \\ -\boldsymbol{v}_i^t & \text{if } \boldsymbol{v}_i^{T_1} \text{ is } R_0\text{-close to } -\boldsymbol{c} \end{cases} \forall i \neq n, \qquad \tilde{\boldsymbol{u}}_j^t = \begin{cases} \boldsymbol{u}_j^t & \text{if } \boldsymbol{u}_j^{T_1} \text{ is } R_0\text{-close to } +\boldsymbol{c} \\ -\boldsymbol{u}_j^t & \text{if } \boldsymbol{u}_j^{T_1} \text{ is } R_0\text{-close to } -\boldsymbol{c} \end{cases} \forall j \in J.$$

By definition, we have

$$\|\tilde{\boldsymbol{v}}_i^{T_1} - \boldsymbol{c}_0\| \leq R_0, \quad \forall i \neq n, \qquad \|\tilde{\boldsymbol{u}}_j^{T_1} - \boldsymbol{c}_0\| \leq R_0, \quad \forall j \in J.$$

This means that $\{\tilde{\boldsymbol{v}}_i^{T_1}\}_{i\neq n}$ and $\{\tilde{\boldsymbol{u}}_j^{T_1}\}_{j\in J}$ form an $(R_0, \boldsymbol{c}_0)$-consensus. Consider creator n. Let

$$\tilde{\boldsymbol{v}}_n^t = \begin{cases} \boldsymbol{v}_n^t & \text{if } \langle \boldsymbol{v}_n^{T_1}, \tilde{\boldsymbol{u}}_{j_0}^{T_1} \rangle > 0 \text{ for some } j_0 \in J \\ -\boldsymbol{v}_n^t & \text{if } \langle \boldsymbol{v}_n^{T_1}, \tilde{\boldsymbol{u}}_j^{T_1} \rangle < 0 \text{ for all } j \in J. \end{cases}$$

(The case where $\langle \boldsymbol{v}_n^{T_1}, \tilde{\boldsymbol{u}}_j^{T_1} \rangle = 0$ for some $j \in J$ is ignored because the initial states that can lead to such states have measure 0.) By definition, we have

$$\langle \tilde{\boldsymbol{v}}_n^{T_1}, \tilde{\boldsymbol{u}}_{j_0}^{T_1} \rangle > 0 \text{ for some } j_0 \in J.$$

Note that, at time step T_1, the system consisting of $\{\tilde{\boldsymbol{v}}_i^{T_1}\}_{i\in[n]}$ and $\{\tilde{\boldsymbol{u}}_j^{T_1}\}_{j\in J}$ satisfies the condition of Lemma F.4, so there exists a path of length $T_2 = \tilde{L}_n^R < +\infty$ that leads the system to R-consensus. Then by the reflection lemma (Lemma D.7), the original system $\{\boldsymbol{v}_i^t\}_{i\in[n]}$, $\{\boldsymbol{u}_j^t\}_{j\in J}$ must reach R-bi-polarization. The total length of path that leads to this R-bi-polarization is $L_n^R = T_1 + T_2 = L_{n-1}^{R_0} + \tilde{L}_n^R < +\infty$. $\qquad\square$

G   Missing Proofs from Section 4

G.1   Proof of Proposition 4.2

Let $R > 0$ be any small number. Let $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_{\lfloor n/k \rfloor} \in \mathbb{R}^d$ be $\lfloor n/k \rfloor$ vectors that satisfy $B(\boldsymbol{c}_\ell, 2R) \cap B(\boldsymbol{c}_{\ell'}, 2R) = \emptyset$ for $\ell \neq \ell'$, where $B(\boldsymbol{c}, R)$ is the ball centered at $\boldsymbol{c}$ with radius R: $\{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x} - \boldsymbol{c}\|_2 \leq R\}$. Consider user and creator features $(\boldsymbol{U}^t, \boldsymbol{V}^t)$ that satisfy: every ball $B(\boldsymbol{c}_\ell, R)$ ($\ell = 1, \ldots, \lfloor n/k \rfloor$) contains k creator vectors, and every user vector $\boldsymbol{u}_j^t$ is in one of the balls $B(\boldsymbol{c}_\ell, R)$. By definition, $(\boldsymbol{U}^t, \boldsymbol{V}^t)$ form $\lfloor n/k \rfloor$ clusters. We show that, after one step of update, the new state $(\boldsymbol{U}^{t+1}, \boldsymbol{V}^{t+1})$ must still form $\lfloor n/k \rfloor$ clusters. Consider any user j. Suppose $\boldsymbol{u}_j^t \in B(\boldsymbol{c}_\ell, R)$, then the distance from $\boldsymbol{u}_j^t$ to any creator $\boldsymbol{v}_i^t \in B(\boldsymbol{c}_\ell, R)$ is at most $2R$:

$$\|\boldsymbol{u}_j^t - \boldsymbol{v}_i^t\| \leq 2R.$$

The distance from $\boldsymbol{u}_j^t$ to any creator $\boldsymbol{v}_{i'}^t$ not in $B(\boldsymbol{c}_\ell, R)$ is greater than $2R$:

$$\|\boldsymbol{u}_j^t - \boldsymbol{v}_{i'}^t\| > 2R$$

because $\boldsymbol{v}_{i'}^t$ is in some other ball $B(\boldsymbol{c}_{\ell'}, R)$ that satisfies $B(\boldsymbol{c}_{\ell'}, 2R) \cap B(\boldsymbol{c}_\ell, 2R) = \emptyset$. This implies that the inner products between user j and the creators in ball $B(\boldsymbol{c}_\ell, R)$ are greater than that with the creators in other ball:

$$\forall \boldsymbol{v}_i^t \in B(\boldsymbol{c}_\ell, R), \ \ \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle = 1 - \frac{1}{2}\|\boldsymbol{u}_j^t - \boldsymbol{v}_i^t\|_2^2 \geq 1 - \frac{1}{2}(2R)^2 > 1 - \frac{1}{2}\|\boldsymbol{u}_j^t - \boldsymbol{v}_{i'}^t\| = \langle \boldsymbol{u}_j^t, \boldsymbol{v}_i^t \rangle, \ \ \forall \boldsymbol{v}_{i'}^t \in B(\boldsymbol{c}_{\ell'}, R).$$

Since $B(\boldsymbol{c}_\ell, R)$ contains k creators, these k creators are the k-most relevant ones to user j, so user j will only be recommended these creators. Then, by applying Observation 3.2 to each of the $\lfloor n/k \rfloor$ balls separately, we see that each ball is a R-consensus and hence absorbing. So, the new state $(\boldsymbol{U}^{t+1}, \boldsymbol{V}^{t+1})$ still forms $\lfloor n/k \rfloor$ clusters with these $\lfloor n/k \rfloor$ balls.

G.2   Proof of Proposition 4.3

The d-dimensional simplex centered at the original has $d+1$ vectors with negative inner products with each other. They form $d+1$ clusters. Since user-creator pairs with negative inner product $\langle \boldsymbol{u}_i, \boldsymbol{v}_j \rangle < 0$ are not recommended, recommendations only happen within each cluster. By Observation 3.2, each cluster is absorbing, so the whole system is stable, keep forming $d+1$ clusters forever.

H   Additional Discussion on Real-World Recommender Systems

Here we further discuss real-world recommender systems' properties and designs that are currently not covered in our main paper. We plan to generalize our model in the future to further capture these features and discuss insightful findings, but having them in the current paper may be a distraction to our main findings.

H.1   User and Creator Retention and Activeness

In our current model, the users and creators will stay in the system from the start to the end. However, in real-world recommender systems, users and creators may leave the platform either permanently or for a certain period. Meanwhile, new users and creators will join the platform. Such join and leave dynamics are also influenced by the recommendations' relevance and diversity, which further complicate the problem. Moreover, users and creators have different activeness levels on the platform, e.g., some users may watch a lot more videos than others, and some creators may post a lot more creations, these effects will also be strongly correlated with the dual influence of the recommender system.

H.2   Creation Quality

Creation quality is a major factor influencing users' feedback in addition to the creation style, e.g., well-made cuisine videos could also be fun and liked by gamers and pet lovers, which we need more than a collaborative filtering type of modeling like our current model to capture such features. A

potential solution to boost both long-term system diversity and single-shot recommendation diversity is to design mechanisms that can incentivize creators to create higher-quality videos instead of changing their creation styles.

H.3 Cold Start

Cold Start is widely used in real-world recommender systems for newly published items. Due to the lack of user-item interactions on new items, the systems randomly recommend these new items to users and collect data for collaborative filtering. In our current model, if we consider the creators creating new items in each time step under their current time creation style, then cold start guarantees the conditions in Theorem 3.3. But if we consider the system to have good enough content understanding ability and can accurately predict the new creations' embeddings, the cold start is not necessary and our model and results in the top-k truncation and threshold truncation parts are valid. We also highlight a subtle difference between cold start and random traffic, if cold start is used on creators instead of items, then after the creator is exposed to users a certain number of times, the system will not guarantee to provide a non-zero probability of recommending this creator, and thus the conditions in Theorem 3.3 may not hold.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification:

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Limitations are discussed in Section 6 and Section H.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory Assumptions and Proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification: Proof sketches are given in the body and the full proofs are in the appendix.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental Result Reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: Experiment details are provided in Section 5 and Section C.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
    (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
    (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
    (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
    (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

Justification: Provided in the supplemental file.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental Setting/Details**

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [Yes]

Justification: See Section 5.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment Statistical Significance**

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification:

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments Compute Resources**

Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer: [NA]

Justification: Computer resources are not a limitation in our experiments.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code Of Ethics**

Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

Answer: [Yes]

Justification:

Guidelines:

- The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader Impacts**

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [Yes]

Justification: Discussed in Section 6.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to

generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.

- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification:

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Justification:

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New Assets**

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

125700

Answer: [NA]

Justification:

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and Research with Human Subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [NA]

Justification:

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Justification:

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.