DYNAMICS IN TWO-SIDED ATTENTION MARKETS: A NON-CONVEX OPTIMIZATION VIEW

Anonymous authors
Paper under double-blind review

ABSTRACT

How do attention markets — where creators compete for engagement, users choose what to consume, and platforms mediate via recommender systems — behave in the long run? Existing theory addresses each side in isolation, leaving open whether the coupled dynamics of user choice, platform ranking and creator strategy converge to a coherent outcome at all.

For a dynamic model of attention markets encompassing heterogeneous creator costs, multinomial logit user choice and a spectrum of ranking schemes, we show that this three-way feedback loop is equivalent to mirror descent or its delayed-gradient variant on a global potential function. This potential decomposes into four interpretable socioeconomic forces: content quality, fairness, production efficiency and diversity, whose balance is directly shaped by the platform's ranking policy. This characterization provides a theoretical foundation for designing healthier creator ecosystems by identifying key policy parameters that govern these trade-offs. Moreover, non-convexity of the potential, intensified by social influence and popularity-based ranking, explains the path-dependent and winner-take-all outcomes observed on real-world platforms.

1 INTRODUCTION

Every minute, millions of users scroll through algorithmically curated feeds, while thousands of creators decide what and how much effort to put into their next posts. Platforms like YouTube, Spotify and TikTok sit at the center of this activity, using recommender systems to match content to audiences. As illustrated in Figure 1, this creates a three-way feedback loop: user engagement generates popularity signals, platform's recommender system translates these into visibility rankings, and creators in turn adapt their content quality in response to those rankings. Despite the enormous cultural and economic stakes of this cycle, we have a surprisingly poor theoretical understanding of where these markets end up: what equilibria, if any, do the self-interested actions of creators, users and platforms collectively converge to?

This question connects to one of the oldest ideas in economics: Adam Smith's (1776) *invisible hand*, the notion that decentralized, self-interested behavior can inadvertently lead to collectively desirable outcomes. But digital attention markets are fundamentally different from the conventional markets — they are two-sided, and digital content has no explicit price as a coordinating signal. Whether an invisible hand operates in this modern setting has remained an open question.

We answer it in the affirmative. Our central insight is that the feedback loop of Figure 1, despite appearing as an opaque process, is secretly performing a global optimization. Specifically, we show that these market dynamics are equivalent to mirror descent on a single potential function, which encodes a balance between content quality, creator fairness, production efficiency and diversity. Rather than an inscrutable black box, the attention economy has a mathematical structure that can be read, interpreted, and crucially, designed. To appreciate why this is non-trivial, consider what prior work has and has not established. Zhu et al. (2023) showed that popularity dynamics under fixed content quality admit a mirror descent interpretation on a convex potential. Ben-Porat et al. (2020) showed convergence to Nash equilibria in creator competition, but without user dynamics. What has been missing is an analysis of the fully coupled system in Figure 1. This coupling introduces non-convexity into the potential function, which turns out to be not a technical nuisance but the

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

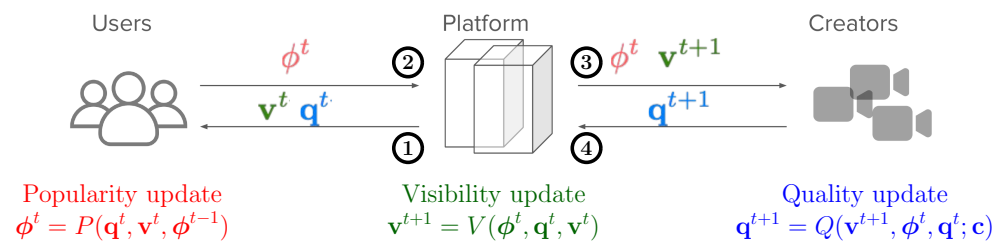

Figure 1: Feedback loop of a two-sided attention market. ① Users engage with content items based on their latent *quality* and platform-assigned *visibility*. ② These engagements generate *popularity* signals. ③ The platform processes the signals to recalibrate its recommender system rankings. ④ Popularity signals and updated rankings provide the strategic incentives for content creators to adapt their content quality. Notations correspond to those in Section 2.

right mathematical language for real-world phenomena: path-dependence, market volatility, and the sudden dominance of a single creator going viral.

Our main contributions are as follows. First, we identify a global potential function (Equation (10), Sections 3 and 4) which combines interpretable socio-economic criteria, and show that platform ranking policies directly shape this landscape — giving designers a concrete lever for steering markets toward desirable outcomes. Second, we establish a formal equivalence between market dynamics and mirror descent or the proximal incremental aggregated gradient (PIAG) method of Zhang et al. (2021) (Theorems 2–3; Remark 1). Specifically, when creators best-respond and platforms rank content by popularity and quality, the resulting market updates are identical to the steps of these optimization methods. Third, we prove global convergence when the potential is convex (Theorem 4), and local convergence results for the non-convex case (Theorem 5). We hope this optimization lens serves as a unifying foundation for future work: the potential function offers a natural objective for empirical calibration of platform parameters, while the mirror descent equivalence opens the door to importing algorithmic tools — acceleration, noise robustness, online learning — into the study of attention markets.

1.1 Related Work

Online content as attention markets. The idea that social influence drives cultural market outcomes long predates modern recommender systems. The seminal Musiclab experiments (Salganik et al., 2006) demonstrated that social influence produces unpredictable and unequal market shares, while Krumme et al. (2012) were the first to apply a multinomial choice model to explain this quantitatively. A well-documented consequence of popularity-driven ranking is algorithmic popularity bias (Ciampaglia et al., 2018): a reinforcing feedback loop that amplifies already-popular content and suppresses higher-quality alternatives. Van Hentenryck et al. (2016) and Ciampaglia et al. (2018) proposed quality-aware ranking strategies to counteract this, while Ghosh & Hummel (2014) modeled the strategic side — how creators trade off the visibility gains of high-quality content against its production costs. Our model synthesizes these threads, combining quality-aware ranking with strategic creator behavior in a unified dynamic framework.

The algorithmic invisible hand. Eisenberg & Gale (1959) showed that competitive equilibria of linear Fisher markets correspond to Nash social welfare maximizers. This inspired a line of work in algorithmic game theory connecting market dynamics such as tâtonnement and proportional response to first-order optimization algorithms (Birnbaum et al., 2011; Brânzei et al., 2021; Cheung et al., 2018; 2020; Gao & Kroer, 2020; Brânzei et al., 2018; Cheung et al., 2025; Goktas et al., 2023). For digital attention markets, Zhu et al. (2023) studied user-platform interactions, and showed equivalence to mirror descent on a convex potential. Our work offers several generalisations (Table 1). By incorporating strategic creator behavior and dynamic platform ranking, we arrive at a non-convex potential — a generalization that is both technically more challenging and empirically more realistic. Chaudhury et al. (2024; 2025) explored non-convex potentials in chore markets; their setting uses Euclidean gradient descent in price space, while ours uses mirror descent on exposure probabilities.

Attention markets as dynamical systems. Maldonado et al. (2018) and Zhu et al. (2023) analyzed user-driven dynamics under multinomial logit choice and established convergence to unique equilibria. Ben-Porat et al. (2020) proved convergence of creator best-response dynamics to pure Nash equilibria, albeit potentially in exponential time. Dean et al. (2024) emphasized the importance of analyzing the coupled creator-user-platform system, and Kiyohara et al. (2025) modeled participation dynamics for both users and creators, though relying on aggregated update rules rather than utility-grounded behavior. Our optimization lens provides a tractable analytical framework that these coupled systems have lacked, yielding convergence results without sacrificing the realism of the three-way interaction.

Game-theoretic approaches. Creator competition has been studied extensively through a game-theoretic lens, with work examining content quality choices, clickbait incentives, and the alignment of Nash equilibria with social welfare via Price of Anarchy and fairness metrics (Hron et al., 2023; Immorlica et al., 2024; Jagadeesan et al., 2023; Yao et al., 2023a; 2024). Several of these settings are instances of potential games, where best-response or no-regret dynamics provably converge (Ben Basat et al., 2017; Ben-Porat & Tennenholtz, 2018; Yao et al., 2023b). Our work complements this literature by making the potential function explicit and interpretable, and by coupling the creator game to a responsive user population and a dynamically updating platform. The potential functions in these potential games can be viewed as an aggregation of creators' utilities as a function of individual actions. In contrast, the potential function in our setting (Equation (10)) combines interests from creators, users and the platform, and is driven by content exposure.

2 Two-sided Attention Markets

A two-sided attention market comprises three groups of agents: content creators, users, and the platform's recommender system (RecSys) acting as mediator. As depicted in Figure 1, the market evolves in discrete epochs via a three-way feedback loop. Each epoch, creators publish content items to compete for RecSys recommendations and user engagement; the RecSys allocates visibility based on creators' historical performance and content quality; users are exposed to recommended items and stochastically decide whether to engage. These patterns then drive strategic quality adjustments by creators and ranking recalibrations by the RecSys for the following epoch.

Now we provide a formal model of the agents' interactions. Let $\mathcal{C}$ denote the set of content creators. We assume that each creator produces *one* piece of content. In the resultant market dynamics, the RecSys controls *visibility parameters* $v_j \in [0,1]$ for every $j \in \mathcal{C}$, ranking items in descending order of v_j. Each creator j controls a *quality parameter* $q_j \in [0,1]$, representing the probability that their item will be engaged by a user upon recommendation. While a scalar quality for each item may sound simple, it encapsulates the degree of match between user preference and content type as implemented by modern recommender systems (Yao et al., 2024). Quality is also a crucial creator decision variable, accounting for the higher cost in time and material needed to produce better content (Ghosh & Hummel, 2014; Jagadeesan et al., 2023). Specifically, this is encoded in cost function $c_j(q) : [0,1] \to \mathbb{R}^+$ for each creator j, which is increasing from $c_j(0) = 0$. The collective response by users determines the *popularity* ϕ_j for every creator $j \in \mathcal{C}$, which measures the fraction of total engagements captured by that creator. One could also call them *market shares*, which reside on the probability simplex, such that $\sum_{j\in\mathcal{C}} \phi_j = 1$.

The system starts from strictly positive $\boldsymbol{\phi}^0$, $\mathbf{v}^0$, and $\mathbf{q}^0$; the visibility and quality updates then produce $\mathbf{v}^1$ and $\mathbf{q}^1$ in that order. In each epoch $t \geq 1$, these parameters update via operators $\mathcal{P}$, $\mathcal{V}$, and $\mathcal{Q}$ (summarized in Section A.1):

- **Quality update** ($\mathbf{q}^{t+1} = \mathcal{Q}(\mathbf{v}^{t+1}, \boldsymbol{\phi}^t, \mathbf{q}^t; \mathbf{c})$): Creators act strategically. They best-respond by choosing quality levels for their next items, balancing the expected visibility/popularity rewards against their fixed production cost functions $\mathbf{c} := \{c_j : j \in \mathcal{C}\}$.
- **Popularity update** ($\boldsymbol{\phi}^t = \mathcal{P}(\mathbf{q}^t, \mathbf{v}^t, \boldsymbol{\phi}^{t-1})$): Users engage with items via platform recommendation $\mathbf{v}^t$, determining the attention market share (popularity) of each item for the current epoch.
- **Visibility update** ($\mathbf{v}^{t+1} = \mathcal{V}(\boldsymbol{\phi}^t, \mathbf{q}^t, \mathbf{v}^t)$): The RecSys observes user engagements and explicitly adjusts the items' visibilities for the subsequent epoch.

We use calendar-time indexing: $\mathbf{s}^t$ denotes exposure at the start of epoch t, computed from $\mathbf{v}^t$ and $\boldsymbol{\phi}^{t-1}$. After users produce $\boldsymbol{\phi}^t$, the platform chooses $\mathbf{v}^{t+1}$; creators then choose $\mathbf{q}^{t+1}$ before the next popularity update.

We start with the creators' strategic behavior. At the end of epoch t, creators update the qualities of their next items. The platform choice $\mathbf{v}^{t+1}$ and current popularity $\boldsymbol{\phi}^t$ determine s_j^{t+1}. Taking this exposure as fixed, creator j chooses q_j to maximize $s_j^{t+1} q_j - c_j(q_j)$. When the cost function is strictly convex, there is a unique maximizer depicted by the lemma below.[1]

Lemma 1. *Suppose that c_j is strictly convex and continuously differentiable, with $c_j'(0) = 0$ and $c_j'(1) \geq 1$. Let $\zeta_j := (c_j')^{-1}$. The unique best-response quality is*

$$q_j^{t+1} = \left[\mathcal{Q}_{\mathrm{BR}}(\mathbf{v}^{t+1}, \boldsymbol{\phi}^t, \mathbf{c})\right]_j := \zeta_j(s_j^{t+1}).$$

In this paper, we also consider the case that the creators keep their qualities constant across epochs, which is denoted by the operator $\mathcal{Q}_0$.

We now specify how user behavior gives rise to the exposure probabilities introduced above and the resulting popularity profile $\boldsymbol{\phi}^t$. We model popularity evolution using generalized multinomial logit choice and consider two models governing intra-epoch evolution.

Proportional Response (PR) In a short epoch, users preview and engage nearly simultaneously. This combines multinomial-logit choice from microeconomic discrete-choice models with feedback resembling proportional response in Fisher markets (Zhu et al., 2023); related probabilistic user-choice rules are common in strategic recommendation (Yao et al., 2023a; 2024). The exposure probability for item j is:

$$s_j^t := \frac{v_j^t (\phi_j^{t-1})^r}{\sum_i v_i^t (\phi_i^{t-1})^r}, \tag{1}$$

where $r \geq 0$ captures the degree of user conformity to social influence. Upon previewing, a user engages with probability q_j^t. In the limit of user population being very large, popularity becomes proportional to $s_j^t q_j^t$.

Definition 1 (Proportional Response (PR))**.** The market dynamic with the popularity operator

$$\left[\mathcal{P}_{\mathrm{PR}}(\mathbf{v}^t, \mathbf{q}^t, \boldsymbol{\phi}^{t-1})\right]_j := \frac{v_j^t q_j^t (\phi_j^{t-1})^r}{\sum_{i \in \mathcal{C}} v_i^t q_i^t (\phi_i^{t-1})^r} \quad \forall j \in \mathcal{C}, \tag{2}$$

is called a Proportional Response dynamic.

Equilibrium Response (ER) ER describes a longer interval between RecSys and creator updates: visibility and quality remain fixed while users repeatedly respond to current popularity until it settles. At each intra-epoch time-step, exposure probabilities update based on current popularity $\phi_j^{\tau,t}$. For $0 \leq r < 1$ and strictly positive visibility and quality, the market admits a unique interior trial-offer market equilibrium (TOME) (Maldonado et al., 2018):

$$\phi_j^{t,*} := \frac{(v_j^t q_j^t)^{\frac{1}{1-r}}}{\sum_i (v_i^t q_i^t)^{\frac{1}{1-r}}}. \tag{3}$$

Theorem 1 (Maldonado et al. (2018); Van Hentenryck et al. (2016))**.** *Suppose $r \in [0, 1)$, $\mathbf{v}^t > \mathbf{0}$, and $\mathbf{q}^t > \mathbf{0}$. From strictly positive initial popularity, the intra-epoch stochastic process converges almost surely to the interior TOME (3).*

The theorem motivates using this settled TOME as each epoch's popularity; the ER rule below takes that long-horizon outcome as exact in the fluid limit.

Definition 2 (Equilibrium Response (ER))**.** The market dynamic with the popularity operator

$$\left[\mathcal{P}_{\mathrm{ER}}(\mathbf{v}^t, \mathbf{q}^t, \boldsymbol{\phi}^{t-1})\right]_j := \frac{(v_j^t q_j^t)^{\frac{1}{1-r}}}{\sum_i (v_i^t q_i^t)^{\frac{1}{1-r}}} \quad \forall j \in \mathcal{C}, \tag{4}$$

is called an Equilibrium Response dynamic.

[1] Proofs and derivations for this and all subsequent results are found in the appendix.

We finally specify the platform's visibility update. Despite the complexity of modern RecSys, their macroscopic behavior effectively interpolates between historical popularity (engagement bias) and inherent content quality (relevance bias). Thus, most policies can be statistically approximated as a *quality-popularity mixed ranking*.

Definition 3 (Mixed Ranking)**.** For $\alpha, \beta \geq 0$, an (α, β)-*mixed ranking* assigns visibilities following a generalized logit distribution:

$$\left[\mathcal{V}_{\text{mix}}(\mathbf{q}^t, \boldsymbol{\phi}^t, \mathbf{v}^t)\right]_j := \frac{\mu_j (q_j^t)^\alpha (\phi_j^t)^\beta}{\sum_i \mu_i (q_i^t)^\alpha (\phi_i^t)^\beta} \quad \forall j \in \mathcal{C}\,. \tag{5}$$

μ_j is a creator-specific weight estimated separately and assumed to be known (or estimated from other data) in this work. It can encode factors such as longevity or reputation on the platform. Since a common rescaling of the weights leaves the ranking unchanged, we assume $\mu_j > 0$ and $\sum_j \mu_j = 1$ throughout. This encompasses α-*quality ranking* ($\beta = 0$), β-*popularity ranking* ($\alpha = 0$), and *constant ranking* ($\alpha = \beta = 0$).

With all three update operators specified, we are ready to define equilibrium of a two-sided attention market.

Definition 4 (Two-sided Attention Market Equilibrium (2SAME))**.** A 2SAME is a fixed point $(\boldsymbol{\phi}^*, \mathbf{v}^*, \mathbf{q}^*)$ of the dynamical system governed by $(\mathcal{P}, \mathcal{V}, \mathcal{Q})$. The equilibrium is *interior* if $\phi_j^* > 0$ for all j.

3 Potential Functions and Mirror Descent in Attention Markets

Table 1 summarizes our main results on interpreting market updates as descent steps. The table is organized along three axes that together pin down a market dynamic: the popularity update model ($\mathcal{P}_{\text{PR}}$ or $\mathcal{P}_{\text{ER}}$), the quality operator ($\mathcal{Q}_0$ if creators keep qualities constant, or $\mathcal{Q}_{\text{BR}}$ if they best-respond), and the ranking scheme used by the platform's RecSys (constant, β-popularity, α-quality, or (α, β)-mixed; see Definition 3). Prior work (Zhu et al., 2023) established a mirror descent equivalence only in the upper-left cell of this grid — constant quality, PR popularity, and constant ranking — thus leaving open the realistic regime in which both creators and the RecSys react. We close this gap by proving the equivalence (Theorems 2 and 3) for the entire grid with two exceptions: for constant quality under any of the four ranking schemes and either popularity model, and for best-responding creators under PR with any ranking scheme as well as under ER with the constant or β-popularity ranking. The two remaining cells, ER paired with $\mathcal{Q}_{\text{BR}}$ and either α-quality or (α, β)-mixed ranking, do not match standard mirror descent but have a lag correction, discussed in the remark at the end of this section. Interpretations of the corresponding potential function as an interpolation between measures on quality, fairness and production efficiency follow in Section 4.

Table 1: Overview of results. Gray text denotes existing results; black text denotes our new results.

Quality operators	Popularity operators	Constant ranking	β-popularity ranking	α-quality ranking	(α, β)-mixed ranking
Constant, $\mathcal{Q}_0$	PR, $\mathcal{P}_{\text{PR}}$	Zhu et al. (2023)	Equivalent to mirror descent on Φ_0 (Theorem 2)		
	ER, $\mathcal{P}_{\text{ER}}$	One epoch [2]			
Best-response, $\mathcal{Q}_{\text{BR}}$	PR, $\mathcal{P}_{\text{PR}}$	Equivalent to mirror descent on Φ_{BR} (Theorem 3)			
	ER, $\mathcal{P}_{\text{ER}}$	Equivalent to mirror descent on Φ_{BR} (Theorem 3)		Equivalent to PIAG (Remark 1)	

To provide context, recall that mirror descent is a first-order optimization algorithm for minimizing a function f over a convex domain X via

$$\mathbf{x}^{t+1} = \arg\min_{\mathbf{x} \in X} \left\{ \left\langle \nabla f(\mathbf{x}^t), \mathbf{x} - \mathbf{x}^t \right\rangle + \frac{1}{\eta} \cdot d_h(\mathbf{x}, \mathbf{x}^t) \right\}, \tag{6}$$

where d_h is the Bregman divergence. When d_h is the KL divergence and X is the probability simplex Δ_n, the update simplifies to $x_j^{t+1} \propto x_j^t \exp(-\eta \nabla_j f(\mathbf{x}^t))$.

[2] The dynamical system reaches equilibrium after one epoch by definition in Section 2.

Returning to market dynamics, arithmetic manipulations yield the following update rules for exposure probabilities $\mathbf{s}^{t+1}$ in Equation (1) under PR and ER with (α, β)-mixed ranking, respectively:

$$s_j^{t+1} = \frac{\mu_j (q_j^t)^{r+\alpha+\beta} (s_j^t)^{r+\beta}}{\sum_i \mu_i (q_i^t)^{r+\alpha+\beta} (s_i^t)^{r+\beta}} . \tag{7}$$

$$s_j^{t+1} = \frac{(\mu_j)^{\frac{1}{1-r}} (q_j^t)^{\alpha+\frac{r+\beta}{1-r}} (q_j^{t-1})^{\frac{\alpha r}{1-r}} (s_j^t)^{\frac{\beta}{1-r}}}{\sum_i (\mu_i)^{\frac{1}{1-r}} (q_i^t)^{\alpha+\frac{r+\beta}{1-r}} (q_i^{t-1})^{\frac{\alpha r}{1-r}} (s_i^t)^{\frac{\beta}{1-r}}} . \tag{8}$$

Under best-response quality, $q_j^t = \zeta_j(s_j^t)$ for $t \geq 1$. Hence the PR reduced recurrence begins at $t = 1$, while the lagged ER recurrence begins at $t = 2$ from $(\mathbf{s}^1, \mathbf{s}^2)$. The convergence theorems re-index these post-startup states as their initial exposure vectors.

Lemma 2. *(Template Potential Function) Let*

$$\Phi(\mathbf{s}) = -\sum_i \left(s_i \log C_i + A \int_0^{s_i} \log \zeta_i(z)\, \mathrm{d}z + B s_i \log s_i \right) ,$$

where A, B, C_i are constants. Assume $C_i > 0$ for every i and, if $A \neq 0$, $\int_0^1 |\log \zeta_i(z)|\, \mathrm{d}z < \infty$ for every i. When $A = 0$, the integral term is omitted. Mirror descent w.r.t. KL divergence with step-size η yields the update: $s_j^{t+1} \propto (C_j)^\eta (s_j^t)^{\eta B+1} (\zeta_j(s_j^t))^{\eta A}$.

The above lemma is the technical backbone of our equivalence results. Recalling $q_j^t = \zeta_j(s_j^t)$ from Lemma 1, we can instantiate (A, B, C_i) to match market update rules Equation (7) and Equation (8), which yields the following theorems. Throughout this section, ER assumes $0 \leq r < 1$, as in Definition 2.

Theorem 2. *Under constant quality $\mathcal{Q}_0$ ($q_j^t \equiv q_j > 0$) and $\mu_j > 0$ for every j, let Φ_0 be the function in Lemma 2 with $A = 0$, $B = r + \beta - 1$, and $C_i = \mu_i (q_i)^{r+\alpha+\beta}$. The PR update (7) is equivalent to mirror descent on Φ_0 with $\eta = 1$. The ER update (8) is equivalent to mirror descent on Φ_0 with $\eta = \frac{1}{1-r}$.*

Theorem 3. *Under best-response quality $\mathcal{Q}_{\mathrm{BR}}$, let Φ_{BR} be the function in Lemma 2 with $A = r + \alpha + \beta$, $B = r + \beta - 1$, and $C_i = \mu_i$. The PR update (7) is equivalent to mirror descent on Φ_{BR} with $\eta = 1$. The ER update (8) for β-popularity ranking ($\alpha = 0$) is equivalent to mirror descent on Φ_{BR} with $\eta = \frac{1}{1-r}$.*

Remark 1. *Under best-response quality and mixed ranking, the ER update is equivalent to the non-Euclidean proximal incremental aggregated gradient (PIAG) update studied by Zhang et al. (2021):*

$$\mathbf{s}^{t+1} = \arg\min_{\mathbf{s} \in \Delta} \left\{ \langle \mathbf{g}^t, \mathbf{s} - \mathbf{s}^t \rangle + \frac{1}{\eta} \mathrm{KL}(\mathbf{s} \| \mathbf{s}^t) \right\}, \qquad \eta = \frac{1}{1-r}. \tag{9}$$

Here $\mathbf{g}^t$ is the aggregated gradient with a one-step delay defined in Section F.1. Global and local convergence are established in Theorems 4–5.

Section 4 interprets the potential socio-economically; Section 5 establishes convergence guarantees.

4 A Socio-economic Interpretation of the Potential Function

The socio-economic intepretation of Φ_{BR} starts with a rewrite of $-\Phi_{\mathrm{BR}}(\mathbf{s}^t)$ into four parts, applying integration by parts with $z = c_i'(y)$ (Appendix C). The result is the sum of the four terms is being maximized and scaled by $\tau_1 = r + \alpha + \beta$ and $\tau_2 = r + \beta$. Recall that $q_j^t = \zeta_j(s_j^t)$ from Lemma 1.

$$-\Phi_{\mathrm{BR}}(\mathbf{s}^t) = \tau_1 \cdot \underbrace{\sum_j s_j^t \log q_j^t}_{\text{expected log-quality}} + \underbrace{\left(-\mathrm{KL}(\mathbf{s}^t \| \boldsymbol{\mu})\right)}_{\text{fairness}} - \tau_2 \cdot \underbrace{H(\mathbf{s}^t)}_{\text{diversity}} + \tau_1 \cdot \underbrace{\left(-\sum_j \int_0^{q_j^t} \frac{c_j'(y)}{y}\, \mathrm{d}y \right)}_{\text{skewed production cost}} , \tag{10}$$

This decomposition reveals four distinct forces driving the market, representing the interests across all three stakeholders: log-quality and diversity align with user welfare; production cost, fairness and diversity govern creator incentives; the platform directly mediates the first three via exposure probabilities. We now interpret each in turn.

Expected log-quality ($\tau_1 \cdot \sum_j s_j^t \log q_j^t$) This term is the expected logarithm of quality over the exposure distribution over creators. It is the most direct proxy for user welfare in the decomposition. Its coefficient $\tau_1 = r + \alpha + \beta$ reveals that quality incentives are amplified jointly by social influence (r), quality bias (α), and popularity bias (β), meaning that any of these forces, not just explicit quality ranking, can drive the market toward higher content quality, which arguably is in the public interest.

Fairness ($-\mathrm{KL}(\mathbf{s}^t \| \boldsymbol{\mu})$). This term pulls the exposure distribution $\mathbf{s}^t$ toward the reference vector $\boldsymbol{\mu}$. Setting $\boldsymbol{\mu}$ to the uniform distribution promotes equal exposure across creators; alternatively, the platform can tune$\boldsymbol{\mu}$ as a targeted intervention to increase visibility to reputable creators who are known to produce high quality content, or under-represented creators. Notably, this is the only term whose normative meaning is directly controllable by the platform through explicit design.

Diversity ($-\tau_2 H(s^t)$). The negative Shannon entropy discourages uniform attention and encourages concentration, with strength governed by $\tau_2 = r + \beta$. This stands in direct tension with the fairness term: while fairness pulls $\mathbf{s}^t$ toward $\boldsymbol{\mu}$, (negative) diversity pulls it toward a single dominant creator. The balance between these two forces — shaped by the platform's choice of β and $\boldsymbol{\mu}$ — determines the balance between a pluralistic or oligopolistic equilibrium. High social influence (r) or heavy popularity bias (β) tips the balance toward oligopoly, as captured by the τ_2 coefficient.

Skewed production cost ($-\tau_1 \sum_j \int_0^{q_j^t} \frac{c_j'(y)}{y} dy$). Unlike the standard production cost $c_j(q) = \int_0^q c_j'(y)\, dy$, the $\frac{1}{y}$ factor gives greater weight to marginal costs at lower quality levels. For a fixed target quality, the reweighted cost is therefore more sensitive to production inefficiencies at those levels. Economically, this amplification functions as a structural barrier to entry (McAfee et al., 2004; Lutz et al., 2010) — a force that is privately efficient but whose social desirability depends on whether it concentrates quality in the hands of a few or raises the floor for all producers. This tension is not resolved by the potential alone and depends on the equilibrium selected.

The decomposition provides an invisible-hand interpretation of the dynamics: creator responses, user engagement, and ranking can be described through a single potential over exposure. We note that it does not by itself establish that the resulting equilibrium is socially optimal.

5 Global and Local Convergence

We now analyze the convergence of the market dynamics by leveraging their equivalence to mirror descent on the potential function Φ_{BR}. Because Φ_{BR} can be non-convex, we establish both global convergence under convexity conditions and local convergence for non-convex landscapes. Missing proofs are deferred to the Appendix.

5.1 Global Convergence under Convexity

We first give a sufficient condition for convexity of Φ_{BR}. Since the cost c_j is strictly convex, $\zeta_j = (c_j')^{-1}$ is strictly increasing.

Lemma 3. *Φ_{BR} is $(1 - r - \beta)$-Bregman-smooth.*

Suppose, in addition, that each ζ_j is differentiable on $(0, 1)$. If, for every $j \in \mathcal{C}$ and every $x \in (0, 1)$,

$$(r + \alpha + \beta) x \zeta_j'(x) + (r + \beta - 1)\zeta_j(x) \leq 0. \tag{11}$$

then Φ_{BR} is convex on Δ.

This condition requires the multiplicative growth rate of the response function ζ_i to be moderate; creators must not drastically alter quality in response to slight income increases. Notably, increasing social influence (r), popularity bias (β), or quality bias (α) makes this condition harder to satisfy. Thus, strong social influence or aggressive ranking drives the potential landscape toward non-convexity, contributing to market volatility and unpredictability (Salganik et al., 2006; Krumme et al., 2012).

For the global and local convergence theorems below, assume $n \geq 2$, $0 \leq r < 1$, $\alpha, \beta \geq 0$, and $\mu_j > 0$. Each response ζ_j is continuous and strictly increasing on $[0,1]$, with $\zeta_j(0) = 0$ and $\int_0^1 |\log \zeta_j(z)|\, \mathrm{d}z < \infty$.

Theorem 4 (Global Convergence under convexity). *Suppose Φ_{BR} is convex on Δ, and let $\mathbf{s}^\star$ be any global minimizer, with value $\Phi^\star_{\mathrm{BR}}$. For every $T \geq 1$:*

- ***Mirror-descent cases.*** *From any $\mathbf{s}^0 \in \Delta^\circ$,*

$$\Phi_{\mathrm{BR}}(\mathbf{s}^T) - \Phi^\star_{\mathrm{BR}} \leq \frac{\mathrm{KL}(\mathbf{s}^\star \| \mathbf{s}^0)}{\eta T}, \tag{12}$$

  where $\eta = 1$ for PR and $\eta = (1-r)^{-1}$ for ER with $\alpha = 0$.

- ***General ER.*** *From any $\mathbf{s}^0, \mathbf{s}^1 \in \Delta^\circ$, the reduced ER sequence converges to a global minimizer, and its average $\widehat{\mathbf{s}}^T := T^{-1} \sum_{t=1}^T \mathbf{s}^{t+1}$ satisfies*

$$\Phi_{\mathrm{BR}}(\widehat{\mathbf{s}}^T) - \Phi^\star_{\mathrm{BR}} \leq \frac{1-r}{T} \left[\mathrm{KL}(\mathbf{s}^\star \| \mathbf{s}^1) + \mathrm{KL}(\mathbf{s}^\star \| \mathbf{s}^0) \right]. \tag{13}$$

If Φ_{BR} is strictly convex, all these dynamics converge to its unique minimizer. Any interior limit induces a 2SAME.

5.2 Local Convergence on Non-Convex Landscapes

For local convergence, call a neighborhood $\Omega(\mathbf{s}^\sharp)$ (γ, κ)-smooth, with $\gamma, \kappa > 0$, if for every $\mathbf{s} \in \Omega(\mathbf{s}^\sharp) \cap \Delta$,

$$\Phi_{\mathrm{BR}}(\mathbf{s}^\sharp) \geq \Phi_{\mathrm{BR}}(\mathbf{s}) + \left\langle \nabla \Phi_{\mathrm{BR}}(\mathbf{s}), \mathbf{s}^\sharp - \mathbf{s} \right\rangle, \quad \left\langle \nabla \Phi_{\mathrm{BR}}(\mathbf{s}), \mathbf{s} - \mathbf{s}^\sharp \right\rangle \geq \gamma \|\mathbf{s} - \mathbf{s}^\sharp\|_2^2,$$

and

$$\|\nabla \Phi_{\mathrm{BR}}(\mathbf{s}) - \nabla \Phi_{\mathrm{BR}}(\mathbf{s}^\sharp)\|_2 \leq \kappa \|\mathbf{s} - \mathbf{s}^\sharp\|_2.$$

Theorem 5 (Local Convergence). *Let $\mathbf{s}^\sharp \in \Delta^\circ$ be a strict local minimizer of Φ_{BR}. From sufficiently close interior initializations ($\mathbf{s}^0$ for the first claim; both $\mathbf{s}^0, \mathbf{s}^1$ for the second):*

- ***Mirror-descent cases.*** *Suppose $\mathbf{s}^\sharp$ has a (γ, κ)-smooth neighborhood and $\eta \leq 2\gamma/\kappa^2$, where $\eta = 1$ for PR and $\eta = (1-r)^{-1}$ for ER with $\alpha = 0$. Then $\mathbf{s}^T \to \mathbf{s}^\sharp$ and, for every $T \geq 1$,*

$$\Phi_{\mathrm{BR}}(\mathbf{s}^T) - \Phi_{\mathrm{BR}}(\mathbf{s}^\sharp) \leq \frac{\mathrm{KL}(\mathbf{s}^\sharp \| \mathbf{s}^0)}{\eta T}.$$

- ***General ER.*** *If Φ_{BR} is convex near $\mathbf{s}^\sharp$, then $\mathbf{s}^t \to \mathbf{s}^\sharp$. If additionally Φ_{BR} is C^2 there and its Hessian at $\mathbf{s}^\sharp$ is positive definite on $\{\mathbf{v} : \mathbf{1}^\top \mathbf{v} = 0\}$, convergence is geometric: $\|\mathbf{s}^t - \mathbf{s}^\sharp\|_2 \leq C q^t$ for some $C > 0$ and $q \in (0,1)$.*

In the $\alpha = 0$ standard-mirror-descent regime, ER uses the larger learning rate $\eta = \frac{1}{1-r} > 1$ and therefore satisfies the sufficient local-stability condition less readily than PR, which uses $\eta = 1$. This is consistent with a greater propensity to leave a shallow local basin.

Finally, strong feedback can amplify a creator's ranking or initial exposure advantage into extinction of competitors or monopoly under mixed ranking.

Proposition 1 (Dominance under mixed ranking). *Consider the deterministic PR or ER dynamics with best-response quality, $0 \leq r < 1$, $\alpha, \beta \geq 0$, positive ranking weights, and interior initial exposure states. Fix $j \neq i$ with $\mu_j \geq \mu_i$ and $\zeta_j \in C^1((0,1))$. Suppose that, for some $\varepsilon \geq 0$ and every $x \in (0,1)$, $\zeta_j(x) \geq \zeta_i(x)$ and*

$$(r + \alpha + \beta) x \zeta_j'(x) + (r + \beta - 1) \zeta_j(x) \geq \varepsilon \zeta_j(x). \tag{14}$$

Assume $s_j^1 \geq s_i^1$; for ER with $\alpha r > 0$, also assume $s_j^2 \geq s_i^2$. Then $s_i^t \to 0$ if either $\mu_j > \mu_i$, or $\varepsilon > 0$ and all required initial exposure inequalities are strict. If these conditions hold against every $i \neq j$, then $s_j^t \to 1$.

Condition (14) requires the feedback between exposure, creator investment, and ranking to be strong enough to sustain or amplify an exposure advantage throughout the market's evolution.

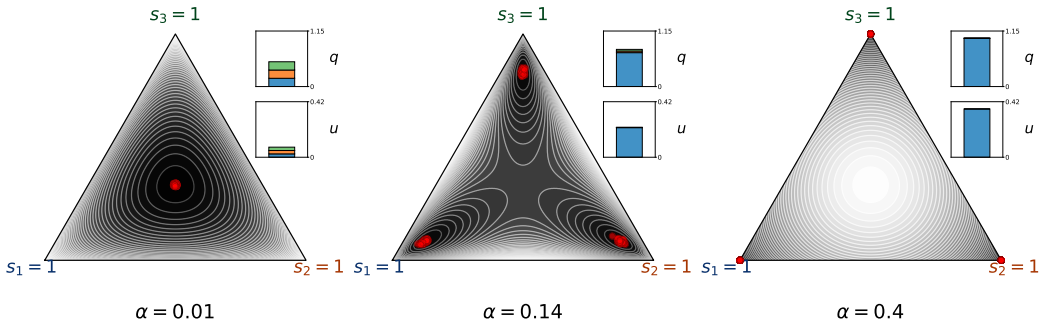

Figure 2: Stochastic PR outcomes for three creators with $\zeta(s) = s\exp(s-1)$, $r = 0.06$, $\beta = 0.29$, and $\mu_j = 1/3$, for $\alpha \in \{0.01, 0.14, 0.40\}$ (left to right). Red dots show 200 terminal exposures over Φ_{BR} contours (darker is lower). Insets show quality q_j (top) and profit u_j (bottom) at the largest K-means cluster center, sorted by exposure and plotted on shared scales; corner labels identify creators.

6 Empirical Simulations

We compare finite-population PR outcomes with the potential landscape while varying the quality-bias exponent α.

Setup. We simulate three identical creators with best-response map $\zeta(s) = s\exp(s-1)$ and matching cost $c'(q) = \zeta^{-1}(q)$, $c(0) = 0$. We set $\mu_j = 1/3$, $r = 0.06$, $\beta = 0.29$, and $\alpha \in \{0.01, 0.14, 0.40\}$. Each epoch, 50,000 users sample previews by current exposure and engage independently with probability q_j^t; their empirical popularities feed the mixed ranking and creator best responses. For each α, we run 200 trajectories of 1,000 epochs from $s_j^0 = 1/3$ and $q_j^0 = 0.5$. Appendix E.1 reports a complementary β sweep.

Recall from Definition 3 that α is the quality-bias exponent of the mixed ranking $\propto \mu_j q_j^\alpha \phi_j^\beta$: small α leaves the platform close to a pure popularity ranking that largely ignores quality, while large α corresponds to an aggressive quality-based ranking that allocates exposure heavily toward whichever creators currently produce the best content. At $\alpha = 0.01$, the potential in Figure 2 is globally convex and terminal exposures cluster near its symmetric minimum. At $\alpha = 0.14$, the potential has three symmetric local minima and the center is a local maximum; stochastic fluctuations determine which creator becomes favored, echoing the empirically observed unpredictability of cultural-market outcomes reported by Salganik et al. (2006). At $\alpha = 0.40$, the potential is globally concave and terminal exposures concentrate near the simplex vertices.

The insets show $q_j = \zeta(s_j)$ and $u_j = s_j q_j - c(q_j)$. At low α, nearly uniform exposure leaves creators at similar, low best-response quality and limits profit through $s_j q_j$. At high α, one creator approaches maximal quality and captures almost all profit, while the others approach zero on both measures; aggregate gains can mask this disparity. In (10), α strengthens log-quality and skewed-cost terms while the coefficient of $-H(\mathbf{s}^t)$ remains $r + \beta$. None of the displayed settings combines high diversity with high quality.

7 Conclusion

We showed that the coupled updates of user choice, creator best responses, and platform ranking admit mirror-descent or PIAG representations through a potential over exposure. These representations give global convergence under convexity and local convergence near suitable minima under the stated conditions. In finite-population simulations, increasing quality bias reshapes the potential landscape and, in the studied three-creator setting, raises total quality and profit while concentrating exposure. Future work could extend the model to personalized recommendations, multidimensional quality, and richer creator rewards; calibrate its parameters empirically; and study correlated attention and barriers to creator entry.

AI USE STATEMENT

The authors used Gemini, Claude and Codex to assist in developing and preparing this paper. These tools helped explore the optimization interpretation of the market dynamics, formulate and check mathematical claims and proof steps, and design and implement the simulations, including suggestions of response functions and parameter choices. They also assisted with editing prose and LaTeX formatting, and locating relevant literature.

The authors independently reviewed and revised all AI-assisted mathematical arguments, text and code. The authors also inspected the resulting figures, and checked bibliographic entries against their original sources. The authors take responsibility for the final submission, including all AI-assisted statements, proofs, code and figures.

REFERENCES

Ran Ben Basat, Moshe Tennenholtz, and Oren Kurland. A game theoretic analysis of the adversarial retrieval setting. *Journal of Artificial Intelligence Research*, 60:1127–1164, 2017. doi: 10.1613/jair.5547. URL https://www.jair.org/index.php/jair/article/view/11104.

Omer Ben-Porat and Moshe Tennenholtz. A game-theoretic approach to recommendation systems with strategic content providers. In *Advances in Neural Information Processing Systems*, volume 31, pp. 1118–1128, 2018. URL https://proceedings.neurips.cc/paper/2018/file/a9a1d5317a33ae8cef33961c34144f84-Paper.pdf.

Omer Ben-Porat, Itay Rosenberg, and Moshe Tennenholtz. Content provider dynamics and coordination in recommendation ecosystems. In Hugo Larochelle, Marc'Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan, and Hsuan-Tien Lin (eds.), *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/dabd8d2ce74e782c65a973ef76fd540b-Abstract.html.

Benjamin Birnbaum, Nikhil R Devanur, and Lin Xiao. Distributed algorithms via gradient descent for fisher markets. In *Proceedings of the 12th ACM conference on Electronic commerce*, pp. 127–136, 2011.

Simina Brânzei, Ruta Mehta, and Noam Nisan. Universal growth in production economies. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems (NeurIPS)*, pp. 1975, 2018. URL https://proceedings.neurips.cc/paper/2018/hash/692f93be8c7a41525c0baf2076aecfb4-Abstract.html.

Simina Brânzei, Nikhil R. Devanur, and Yuval Rabani. Proportional dynamics in exchange economies. In Péter Biró, Shuchi Chawla, and Federico Echenique (eds.), *Proceedings of the 22nd ACM Conference on Economics and Computation (EC)*, pp. 180–201. ACM, 2021. doi: 10.1145/3465456.3467644. URL https://doi.org/10.1145/3465456.3467644.

Bhaskar Ray Chaudhury, Christian Kroer, Ruta Mehta, and Tianlong Nan. Competitive equilibrium for chores: from dual eisenberg-gale to a fast, greedy, lp-based algorithm. In Dirk Bergemann, Robert Kleinberg, and Daniela Sabán (eds.), *Proceedings of the 25th ACM Conference on Economics and Computation, EC 2024, New Haven, CT, USA, July 8-11, 2024*, pp. 40. ACM, 2024. doi: 10.1145/3670865.3673516. URL https://doi.org/10.1145/3670865.3673516.

Bhaskar Ray Chaudhury, Christian Kroer, Ruta Mehta, and Tianlong Nan. Tâtonnement dynamics for fisher markets with chores. *CoRR*, abs/2511.21162, 2025. doi: 10.48550/ARXIV.2511.21162. URL https://doi.org/10.48550/arXiv.2511.21162.

Yun Kuen Cheung, Richard Cole, and Yixin Tao. Dynamics of distributed updating in fisher markets. In *Proceedings of the 19th ACM Conference on Economics and Computation (EC)*, pp. 351–368, Ithaca, NY, USA, 2018. ACM.

Yun Kuen Cheung, Richard Cole, and Nikhil R. Devanur. Tatonnement beyond gross substitutes? Gradient descent to the rescue. *Games and Economic Behavior*, 123:295–326, 2020. ISSN 0899-8256. doi: 10.1016/j.geb.2019.03.014.

Yun Kuen Cheung, Richard Cole, and Yixin Tao. Proportional response dynamics in gross substitutes markets. In Itai Ashlagi and Aaron Roth (eds.), *Proceedings of the 26th ACM Conference on Economics and Computation, EC 2025, Stanford University, Stanford, CA, USA, July 7-10, 2025*, pp. 389. ACM, 2025. doi: 10.1145/3736252.3742559. URL https://doi.org/10.1145/3736252.3742559.

Giovanni Luca Ciampaglia, Azadeh Nematzadeh, Filippo Menczer, and Alessandro Flammini. How algorithmic popularity bias hinders or promotes quality. *Scientific Reports*, 8(1):15951, 2018.

Sarah Dean, Evan Dong, Meena Jagadeesan, and Liu Leqi. Recommender systems as dynamical systems: Interactions with viewers and creators. In *Workshop on Recommendation Ecosystems: Modeling, Optimization and Incentive Design*, 2024. URL https://openreview.net/forum?id=LAhDyjC3HF.

Edmund Eisenberg and David Gale. Consensus of subjective probabilities: The pari-mutuel method. *The Annals of Mathematical Statistics*, 30(1):165–168, 1959.

Yuan Gao and Christian Kroer. First-order methods for large-scale market equilibrium computation. In *Proceedings of the 34th International Conference on Neural Information Processing Systems (Neurips)*, volume 33, pp. 21738–21750, 2020.

Arpita Ghosh and Patrick Hummel. A game-theoretic analysis of rank-order mechanisms for user-generated content. *Journal of Economic Theory*, 154:349–374, 2014. ISSN 0022-0531. doi: https://doi.org/10.1016/j.jet.2014.09.009. URL https://www.sciencedirect.com/science/article/pii/S0022053114001306.

Denizalp Goktas, Jiayi Zhao, and Amy Greenwald. Tâtonnement in homothetic fisher markets. In Kevin Leyton-Brown, Jason D. Hartline, and Larry Samuelson (eds.), *Proceedings of the 24th ACM Conference on Economics and Computation, EC 2023, London, United Kingdom, July 9-12, 2023*, pp. 760–781. ACM, 2023. doi: 10.1145/3580507.3597746. URL https://doi.org/10.1145/3580507.3597746.

Jiri Hron, Karl Krauth, Michael I. Jordan, Niki Kilbertus, and Sarah Dean. Modeling content creator incentives on algorithm-curated platforms. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023. URL https://openreview.net/forum?id=l6CpxixmUg.

Nicole Immorlica, Meena Jagadeesan, and Brendan Lucier. Clickbait vs. quality: How engagement-based optimization shapes the content landscape in online platforms. In *Proceedings of the ACM Web Conference 2024*, pp. 36–45. ACM, 2024. doi: 10.1145/3589334.3645353. URL https://dl.acm.org/doi/10.1145/3589334.3645353.

Meena Jagadeesan, Nikhil Garg, and Jacob Steinhardt. Supply-side equilibria in recommender systems. In Alice Oh, Tristan Naumann, Amir Globerson, Kate Saenko, Moritz Hardt, and Sergey Levine (eds.), *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*, 2023. URL http://papers.nips.cc/paper_files/paper/2023/hash/2f1486343c2c942a617e4f5bb0cc64c8-Abstract-Conference.html.

Haruka Kiyohara, Fan Yao, and Sarah Dean. Policy design for two-sided platforms with participation dynamics. In *International Conference on Machine Learning*, Proceedings of Machine Learning Research, pp. 30966–30985. PMLR / OpenReview.net, 2025. URL https://proceedings.mlr.press/v267/kiyohara25a.html.

Coco Krumme, Manuel Cebrian, Galen Pickard, and Sandy Pentland. Quantifying social influence in an online cultural market. *PloS one*, 7(5):e33785, 2012.

Clemens HM Lutz, Ron GM Kemp, and S Gerhard Dijkstra. Perceptions regarding strategic and structural entry barriers. *Small Business Economics*, 35:19–33, 2010.

Felipe Maldonado, Pascal Van Hentenryck, Gerardo Berbeglia, and Franco Berbeglia. Popularity signals in trial-offer markets with social influence and position bias. *European Journal of Operational Research*, 266(2):775–793, 2018.

Preston R. McAfee, Hugo M. Mialon, and Michael A. Williams. What is a barrier to entry? *American Economic Review*, 94(2):461–465, May 2004. doi: 10.1257/0002828041302235. URL https://www.aeaweb.org/articles?id=10.1257/0002828041302235.

Matthew J Salganik, Peter Sheridan Dodds, and Duncan J Watts. Experimental study of inequality and unpredictability in an artificial cultural market. *science*, 311(5762):854–856, 2006.

Adam Smith. *An Inquiry into the Nature and Causes of the Wealth of Nations*. W. Strahan and T. Cadell, London, 1776. URL https://www.econlib.org/library/Smith/smWN.html.

Pascal Van Hentenryck, Andrés Abeliuk, Franco Berbeglia, Felipe Maldonado, and Gerardo Berbeglia. Aligning popularity and quality in online cultural markets. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 10, pp. 398–407, 2016.

Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang, and Haifeng Xu. How bad is top-k recommendation under competing content creators? In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (eds.), *International Conference on Machine Learning, ICML 2023, 23-29 July 2023, Honolulu, Hawaii, USA*, volume 202 of *Proceedings of Machine Learning Research*, pp. 39674–39701. PMLR, 2023a. URL https://proceedings.mlr.press/v202/yao23b.html.

Fan Yao, Chuanhao Li, Karthik Abinav Sankararaman, Yiming Liao, Yan Zhu, Qifan Wang, Hongning Wang, and Haifeng Xu. Rethinking incentives in recommender systems: Are monotone rewards always beneficial? In Alice Oh, Tristan Naumann, Amir Globerson, Kate Saenko, Moritz Hardt, and Sergey Levine (eds.), *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*, 2023b. URL http://papers.nips.cc/paper_files/paper/2023/hash/ebfabf372037aaa4a8d92c9b457ece3e-Abstract-Conference.html.

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 Ricardo Baeza-Yates and Francesco Bonchi (eds.), *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD 2024, Barcelona, Spain, August 25-29, 2024*, pp. 3874–3885. ACM, 2024. doi: 10.1145/3637528.3672021. URL https://doi.org/10.1145/3637528.3672021.

Hui Zhang, Yu-Hong Dai, Lei Guo, and Wei Peng. Proximal-like incremental aggregated gradient method with linear convergence under bregman distance growth conditions. *Mathematics of Operations Research*, 46(1):61–81, 2021. doi: 10.1287/moor.2019.1047. URL https://doi.org/10.1287/moor.2019.1047.

Haiqing Zhu, Yun Kuen Cheung, and Lexing Xie. Stability and efficiency of personalised cultural markets. In *Proceedings of the ACM Web Conference 2023*, pp. 3447–3455, 2023.

PROOFS AND ADDITIONAL EXPERIMENTS FOR THIS PAPER

A ADDITIONAL INFORMATION FOR SECTION 2

A.1 RECAP: UPDATE RULES OF USERS, RECOMMENDER SYSTEMS AND CONTENT CREATORS

In Section 2, we gave elaborate discussions of different update rules of users, recommender systems (RS) and content creators. We summarize the key notations and formulae here. Recall that $\boldsymbol{\phi}$ denotes popularity vector, $\mathbf{v}$ denotes visibility vector, $\mathbf{q}$ denotes quality vector, and $\mathbf{c}$ denotes the vector of cost functions of different content creators.

The general form of dynamical systems of two-sided attention markets is

$$\begin{aligned} \boldsymbol{\phi}^t &= \mathcal{P}(\mathbf{q}^t, \mathbf{v}^t, \boldsymbol{\phi}^{t-1}) && \text{(users' popularity update)} \\ \mathbf{v}^{t+1} &= \mathcal{V}(\boldsymbol{\phi}^t, \mathbf{q}^t, \mathbf{v}^t) && \text{(RS's visibility update)} \\ \mathbf{q}^{t+1} &= \mathcal{Q}(\mathbf{v}^{t+1}, \boldsymbol{\phi}^t, \mathbf{q}^t; \mathbf{c}) && \text{(content creators' quality update)} \end{aligned}$$

An important variable is preview probabilities

$$s_j^t \;:=\; \mathbb{P}\left[\text{item } j \text{ is chosen for preview in epoch } t\right] = \frac{v_j^t (\phi_j^{t-1})^r}{\sum_i v_i^t (\phi_i^{t-1})^r} \;,$$

POPULARITY UPDATE

We examine the following two rules:

- Proportional Response (PR) Dynamic (Definition 1), where

$$\left[\mathcal{P}_{\mathrm{PR}}(\mathbf{v}^t, \mathbf{q}^t, \boldsymbol{\phi}^{t-1})\right]_j := \frac{q_j^t v_j^t (\phi_j^{t-1})^r}{\sum_i q_i^t v_i^t (\phi_i^{t-1})^r} \;.$$

- Equilibrium Response (ER) Dynamic (Definition 2), where

$$\left[\mathcal{P}_{\mathrm{ER}}(\mathbf{v}^t, \mathbf{q}^t, \boldsymbol{\phi}^{t-1})\right]_j := \frac{(q_j^t v_j^t)^{\frac{1}{1-r}}}{\sum_i (q_i^t v_i^t)^{\frac{1}{1-r}}} \;.$$

VISIBILITY UPDATE

We examine the following two groups of rules for visibility update:

- Constant: $\mathcal{V}_0(\mathbf{q}^t, \boldsymbol{\phi}^t, \mathbf{v}^t) := \mathbf{v}$, where each $v_j > 0$, and $\sum_j v_j = 1$.
- Popularity-Quality mixed ranking: This includes a range of ranking mechanisms, specified by the parameters α, β and $(\mu_1, \mu_2, \ldots, \mu_{|C|})$ as below:

$$\left[\mathcal{V}_{\mathrm{mix}}(\mathbf{q}^t, \boldsymbol{\phi}^t, \mathbf{v}^t)\right]_j := \frac{\mu_j (q_j^t)^\alpha (\phi_j^t)^\beta}{\sum_i \mu_i (q_i^t)^\alpha (\phi_i^t)^\beta} \;.$$

  In particular,
  - When $\beta = 1$ and $\alpha = 0$, this is called popularity ranking.
  - When $\beta = 0$ and $\alpha = 1$, this is called quality ranking.

QUALITY UPDATE

We examine best-response update. We need the assumptions that for every $j \in \mathcal{C}$, c_j is strictly convex and continuously differentiable, $c_j'(0) = 0$ and $c_j'(1) \geq 1$. Then the update rule can be explicitly written as

$$q_j^{t+1} = \left[\mathcal{Q}_{\mathrm{BR}}(\mathbf{v}^{t+1}, \boldsymbol{\phi}^t, \mathbf{c})\right]_j := \zeta_j(s_j^{t+1}).$$

where $\zeta_j := (c_j')^{-1}$.

A.2 Proof of Lemma 1

Proof. Fix $s = s_j^{t+1} \in [0,1]$. Creator j maximizes

$$u_j(q; s) = sq - c_j(q) \qquad \text{over } q \in [0,1].$$

Because c_j is continuously differentiable and strictly convex, $u_j(\cdot; s)$ is strictly concave. Moreover,

$$c_j'(0) = 0 \le s \le 1 \le c_j'(1).$$

Thus there is a unique optimizer $q^\star \in [0,1]$ satisfying

$$u_j'(q^\star; s) = s - c_j'(q^\star) = 0.$$

Therefore,

$$q_j^{t+1} = q^\star = (c_j')^{-1}(s_j^{t+1}) = \zeta_j(s_j^{t+1}).$$

□

B Proofs and Derivations for Section 3

B.1 The Update Rules for Exposure Probabilities s in PR and ER Dynamics

We derive the update rules Equations (7) and (8) here.

For PR, recall from Definition 1 that $\phi_j^t \propto v_j^t q_j^t (\phi_j^{t-1})^r$. By Definition 3 of mixed ranking, $v_j^{t+1} \propto \mu_j (q_j^t)^\alpha (\phi_j^t)^\beta$. Hence,

$$\begin{aligned} s_j^{t+1} &\propto v_j^{t+1} (\phi_j^t)^r \\ &\propto \mu_j (q_j^t)^\alpha (\phi_j^t)^{r+\beta} \\ &\propto \mu_j (q_j^t)^{r+\alpha+\beta} \left(v_j^t (\phi_j^{t-1})^r \right)^{r+\beta} \\ &\propto \mu_j (q_j^t)^{r+\alpha+\beta} (s_j^t)^{r+\beta}. \end{aligned}$$

which completes the derivation of (7) since $\sum_j s_j^{t+1} = 1$.

For ER, by Definition 2, $\phi_j^t \propto (q_j^t v_j^t)^{\frac{1}{1-r}}$. Hence,

$$\begin{aligned} s_j^{t+1} &\propto v_j^{t+1} (\phi_j^t)^r \\ &\propto \mu_j (q_j^t)^\alpha (\phi_j^t)^{r+\beta} \\ &\propto \mu_j (q_j^t)^{\alpha + \frac{r+\beta}{1-r}} (v_j^t)^{\frac{r+\beta}{1-r}}. \end{aligned}$$

Splitting the exponent on v_j^t and expanding $(v_j^t)^{\frac{r}{1-r}}$ using the mixed-ranking update gives

$$s_j^{t+1} \propto \mu_j (q_j^t)^{\alpha + \frac{r+\beta}{1-r}} (v_j^t)^{\frac{\beta}{1-r}} \mu_j^{\frac{r}{1-r}} (q_j^{t-1})^{\frac{\alpha r}{1-r}} (\phi_j^{t-1})^{\frac{\beta r}{1-r}}.$$

In the RHS, $(v_j^t)^{\frac{\beta}{1-r}} (\phi_j^{t-1})^{\frac{\beta r}{1-r}} \propto (s_j^t)^{\frac{\beta}{1-r}}$ by the definition of s_j^t. Hence,

$$s_j^{t+1} \propto (\mu_j)^{\frac{1}{1-r}} (q_j^t)^{\alpha + \frac{r+\beta}{1-r}} (q_j^{t-1})^{\frac{\alpha r}{1-r}} (s_j^t)^{\frac{\beta}{1-r}},$$

which completes the derivation of (8).

B.2 Proofs for Lemma 2, Theorem 2 and Theorem 3

We provide more details on the calculations of the equivalence between mirror descents of the potential functions and market dynamic updates here. We start with the formal definition of Bregman divergence.

Definition 5. Let X be a compact and convex set, and let h be a differentiable convex function on X. The *Bregman divergence* w.r.t. h, denoted by d_h, is

$$d_h(\mathbf{x}, \mathbf{y}) = h(\mathbf{x}) - h(\mathbf{y}) - \langle \nabla h(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle ,$$

for any $\mathbf{x} \in X$ and $\mathbf{y} \in \mathrm{rint}(X)$, where $\mathrm{rint}(X)$ denotes the relative interior of the set X. The Kullback–Leibler (KL) divergence, denoted by $KL(\mathbf{x}\|\mathbf{y})$, is the Bregman divergence generated by the function $h(\mathbf{x}) = \sum_j x_j \log x_j - x_j$.

Given a Bregman divergence d_h, the corresponding mirror descent update rule is

$$\mathbf{x}^{t+1} = \arg\min_{\mathbf{x} \in X} \left\{ \left\langle \nabla f(\mathbf{x}^t), \mathbf{x} - \mathbf{x}^t \right\rangle + \frac{1}{\eta} \cdot d_h(\mathbf{x}, \mathbf{x}^t) \right\}, \tag{15}$$

where η is considered as the step-size. If d_h is the KL divergence and X is the probability simplex $\Delta_n = \{\mathbf{x} \geq 0 \mid \sum_{j=1}^n x_j = 1\}$, then the mirror descent update can be written as

$$x_j^{t+1} = \frac{x_j^t \exp(-\eta \nabla_j f(\mathbf{x}^t))}{\sum_i x_i^t \exp(-\eta \nabla_i f(\mathbf{x}^t))} \qquad \forall j = 1, 2, \ldots, n \,. \tag{16}$$

Then, the results in Lemma 2, Theorem 2 and Theorem 3 can be obtained by direct calculation of the gradients of the objectives.

C Derivation for Section 4

We establish the integral identity used in Equation (10) without assuming that c_j'' exists. Assume the log-integrability condition in Lemma 2. Fix $s_j \in (0, 1]$ and set $q_j = \zeta_j(s_j)$. Since $c_j'(1) \geq 1$, we have $0 < q_j \leq 1$. For $z > 0$,

$$-\log \zeta_j(z) = \int_{\zeta_j(z)}^1 \frac{\mathrm{d}y}{y}.$$

Tonelli's theorem and the inverse relation $\zeta_j^{-1} = c_j'$ give

$$\begin{aligned} -\int_0^{s_j} \log \zeta_j(z)\, \mathrm{d}z &= \int_0^{q_j} \frac{c_j'(y)}{y}\, \mathrm{d}y + s_j \int_{q_j}^1 \frac{\mathrm{d}y}{y} \\ &= \int_0^{q_j} \frac{c_j'(y)}{y}\, \mathrm{d}y - s_j \log q_j. \end{aligned}$$

Thus,

$$\int_0^{s_j} \log \zeta_j(z)\, \mathrm{d}z = s_j \log \zeta_j(s_j) - \int_0^{\zeta_j(s_j)} \frac{c_j'(y)}{y}\, \mathrm{d}y.$$

The identity extends to $s_j = 0$ by continuity under the log-integrability condition in Lemma 2. In particular, with $q_j^t = \zeta_j(s_j^t)$,

$$\int_0^{s_j^t} \log \zeta_j(z)\, \mathrm{d}z = s_j^t \log q_j^t - \int_0^{q_j^t} \frac{c_j'(y)}{y}\, \mathrm{d}y.$$

D Proofs for Section 5

D.1 Proof of Lemma 3

Proof of Lemma 3: To prove it is Bregman smooth, recall that

$$\Phi_{\mathrm{BR}}(\mathbf{s}) \;=\; -\sum_i \left(s_i \log \mu_i \;+\; A \int_0^{s_i} \log \zeta_i(z)\, \mathrm{d}z \;+\; B s_i \log s_i \right),$$

where $A = r + \alpha + \beta$ and $B = r + \beta - 1$, and hence

$$\nabla_i \Phi_{\mathrm{BR}}(\mathbf{s}') \;=\; -\left(\log \mu_i + A \log \zeta_i(s_i') + B(\log s_i' + 1)\right).$$

Thus,

$$\begin{aligned}
&\Phi_{\mathrm{BR}}(\mathbf{s}) - \Phi_{\mathrm{BR}}(\mathbf{s}') - \langle \nabla \Phi_{\mathrm{BR}}(\mathbf{s}'), \mathbf{s} - \mathbf{s}' \rangle \\
&= -\sum_i \left(A \int_{s_i'}^{s_i} \log \zeta_i(z)\, \mathrm{d}z + B(s_i \log s_i - s_i' \log s_i') - A(s_i - s_i') \log \zeta_i(s_i') - B(\log s_i' + 1)(s_i - s_i') \right) \\
&= -\sum_i \left[A \left(\int_{s_i'}^{s_i} \log \zeta_i(z)\, \mathrm{d}z - (s_i - s_i') \log \zeta_i(s_i') \right) + B s_i \log \frac{s_i}{s_i'} \right] \qquad (\text{since } \mathbf{s}, \mathbf{s}' \in \Delta, \textstyle\sum_i (s_i - s_i') = 0) \\
&= (1 - r - \beta) \cdot \mathrm{KL}(\mathbf{s} \| \mathbf{s}') \; - \; (r + \alpha + \beta) \sum_i \int_{s_i'}^{s_i} \log \frac{\zeta_i(z)}{\zeta_i(s_i')}\, \mathrm{d}z \,.
\end{aligned}$$

Since ζ_i is an increasing function, the last integral is non-negative, and we complete the part of the proof that it is Bregman smooth.

For the convexity claim, differentiability of ζ_j on $(0, 1)$ gives

$$\frac{\partial^2 \Phi_{\mathrm{BR}}}{\partial s_j^2}(\mathbf{s}) = -\frac{(r + \alpha + \beta)\zeta_j'(s_j)}{\zeta_j(s_j)} - \frac{r + \beta - 1}{s_j}, \qquad \mathbf{s} \in \Delta^\circ.$$

Since $s_j \zeta_j(s_j) > 0$, condition (11) at every $x \in (0, 1)$ makes each diagonal entry nonnegative throughout Δ°. Separability therefore makes Φ_{BR} convex on Δ°. The stated log-integrability condition makes Φ_{BR} continuous up to the boundary, so it is convex on Δ.

□

D.2 Proof of Theorem 4

Definition 6. A function f is L-Bregman-smooth with respect to Bregman divergence d_h, if for any $\mathbf{y} \in \mathrm{rint} X$ and $\mathbf{x} \in X$,

$$f(\mathbf{x}) \leq f(\mathbf{y}) + \langle \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle + L \cdot d_h(\mathbf{x}, \mathbf{y}).$$

Moreover, it is L-Bregman-convex if the function is both L-Bregman-smooth and convex.

Theorem 6 ((Birnbaum et al., 2011, Theorem 3))**.** *Suppose f is an L-Bregman-convex function with respect to d_h, let $\mathbf{x}^T$ be the point reached by running mirror descent Equation (6). Then with the learning rate $\eta \leq \frac{1}{L}$, the following inequality holds for any $T \geq 1$:*

$$f(\mathbf{x}^T) - f(\mathbf{x}^\star) \leq \frac{d_h(\mathbf{x}^\star, \mathbf{x}^0)}{\eta \cdot T}.$$

Proof of Theorem 4. By Lemma 3 and $\beta \geq 0$, the convex potential Φ_{BR} is $(1 - r)$-Bregman-convex. The mirror-descent equivalence in Theorem 3 and Theorem 6 give (12), since both prescribed step sizes satisfy $\eta \leq (1 - r)^{-1}$. The general-ER claim follows from Theorem 7 and (34). Under strict convexity, the mirror-descent sequences also converge to the unique minimizer: their objective gaps tend to zero, and the potential is continuous on the compact simplex.

□

D.3 Proof of Theorem 5

We first prove the mirror-descent claim. Set $f := \Phi_{\mathrm{BR}}$ and $\mathbf{x}^t := \mathbf{s}^t$. By Lemma 3, f is L-Bregman-smooth, so

$$f(\mathbf{x}) - f(\mathbf{y}) - \langle \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \leq L \cdot \mathrm{KL}(\mathbf{x}, \mathbf{y}). \tag{17}$$

Hence, it is true that, for $\frac{1}{\eta} \geq L$, we have

$$\begin{aligned}
f(\mathbf{x}^{t+1}) &\leq f(\mathbf{x}^t) + \left\langle \nabla f(\mathbf{x}^t), \mathbf{x}^{t+1} - \mathbf{x}^t \right\rangle + \frac{1}{\eta} \mathrm{KL}(\mathbf{x}^{t+1}, \mathbf{x}^t) \\
&\leq f(\mathbf{x}^t) + \left\langle \nabla f(\mathbf{x}^t), \mathbf{x}^t - \mathbf{x}^t \right\rangle + \frac{1}{\eta} \mathrm{KL}(\mathbf{x}^t, \mathbf{x}^t) = f(\mathbf{x}^t),
\end{aligned}$$

where the first inequality follows from Equation (17) and positivity of KL divergence, the second inequality follows from the optimality of $\mathbf{x}^{t+1}$ given by the mirror descent update rule. Now, consider the following lemma.

Lemma 4. *For some open ball $\mathcal{B}_k(x^\star)$ with $k > 0$, if $\mathbf{x}^\star$ is a local minimizer on $\mathcal{B}_k(x^\star) \cap \Delta$. Then, there exists $k' > 0$ such that $\mathbf{x}^\star$ is a local minimizer of the set $\Omega := \{\mathbf{z} \in \Delta : \mathrm{KL}(\mathbf{x}^\star, \mathbf{z}) < k'\}$.*

Proof. For any point $\mathbf{z} \in \Omega$, set $k' = \frac{k^2}{2}$, we have

$$\|\mathbf{x}^\star - \mathbf{z}\|_2 \le \|\mathbf{x}^\star - \mathbf{z}\|_1 \le \sqrt{2\mathrm{KL}(\mathbf{x}^\star, \mathbf{z})} \le \sqrt{2 \cdot \frac{k^2}{2}} = k,$$

where the second inequality follows from Pinsker inequality. Hence, we have shown that $\Omega \subset \mathcal{B}_k(x^\star) \cap \Delta$. Since $k' > 0$, by continuity of KL divergence, we could also observe that Ω is open with respect to the subspace topology on Δ. □

For the mirror-descent claim, let $\mathbf{x}^\star := \mathbf{s}^\sharp$ and let $\Omega(\mathbf{x}^\star)$ be a (γ, κ)-smooth neighborhood as defined in Section 5.2: for every $\mathbf{y} \in \Omega(\mathbf{x}^\star) \cap \Delta$,

$$f(\mathbf{x}^\star) \ge f(\mathbf{y}) + \langle \nabla f(\mathbf{y}), \mathbf{x}^\star - \mathbf{y} \rangle, \tag{18}$$

$$\langle \nabla f(\mathbf{y}), \mathbf{y} - \mathbf{x}^\star \rangle \ge \gamma \|\mathbf{y} - \mathbf{x}^\star\|_2^2, \tag{19}$$

$$\|\nabla f(\mathbf{y}) - \nabla f(\mathbf{x}^\star)\|_2 \le \kappa \|\mathbf{y} - \mathbf{x}^\star\|_2. \tag{20}$$

These are explicit assumptions of the mirror-descent claim; they are not implied by an arbitrary strict local minimum. A positive-definite Hessian on the tangent space at $\mathbf{x}^\star$, together with twice continuous differentiability, is a sufficient condition after shrinking the neighborhood.

Lemma 5. *Let $\mathbf{x}^\star \in \operatorname{rint} \Delta$ and suppose (19)–(20) hold at $\mathbf{x}^t$. For the KL mirror-descent update with $\eta \le 2\gamma/\kappa^2$,*

$$\mathrm{KL}(\mathbf{x}^\star, \mathbf{x}^{t+1}) \le \mathrm{KL}(\mathbf{x}^\star, \mathbf{x}^t).$$

Proof. Write $\mathbf{g}^t = \nabla f(\mathbf{x}^t)$ and $\mathbf{g}^\star = \nabla f(\mathbf{x}^\star)$. Since $\mathbf{x}^\star$ is an interior constrained minimizer, the KKT conditions give $\mathbf{g}^\star = \lambda \mathbf{1}$ for some λ. Set $\mathbf{y} = \mathbf{g}^\star - \mathbf{g}^t$ and $D_t = \mathrm{KL}(\mathbf{x}^\star, \mathbf{x}^t)$. The multiplicative form of KL mirror descent gives

$$D_{t+1} - D_t = \eta \langle \mathbf{x}^\star, \mathbf{g}^t \rangle + \log\left(\sum_i x_i^t e^{-\eta g_i^t} \right).$$

Using (19), (20), and $\langle \mathbf{g}^\star, \mathbf{x}^t - \mathbf{x}^\star \rangle = 0$, we obtain

$$D_{t+1} - D_t \le -\eta \langle \mathbf{x}^t, \mathbf{y} \rangle - \frac{\eta\gamma}{\kappa^2} \|\mathbf{y}\|_2^2 + \log\left(\sum_i x_i^t e^{\eta y_i} \right). \tag{21}$$

To bound the last term, define $\psi(\mathbf{y}) = \log\left(\sum_i x_i^t e^{\eta y_i}\right)$. If $w_i(\mathbf{y}) = x_i^t e^{\eta y_i} / \sum_k x_k^t e^{\eta y_k}$, then

$$\nabla^2 \psi(\mathbf{y}) = \eta^2 \left(\mathrm{Diag}(\mathbf{w}) - \mathbf{w}\mathbf{w}^\top \right) \preceq \eta^2 I.$$

Since $\psi(\mathbf{0}) = 0$ and $\nabla\psi(\mathbf{0}) = \eta \mathbf{x}^t$, smoothness yields

$$\psi(\mathbf{y}) \le \eta \langle \mathbf{x}^t, \mathbf{y} \rangle + \frac{\eta^2}{2} \|\mathbf{y}\|_2^2.$$

Substituting this inequality into (21) gives

$$D_{t+1} - D_t \le \eta \left(\frac{\eta}{2} - \frac{\gamma}{\kappa^2} \right) \|\mathbf{y}\|_2^2 \le 0.$$

This proves the claim, including the equality case $\eta = 2\gamma/\kappa^2$. □

Choose $\rho > 0$ sufficiently small that

$$\mathcal{K}_\rho := \{\mathbf{x} \in \Delta : \mathrm{KL}(\mathbf{x}^\star, \mathbf{x}) \le \rho\} \subset \Omega(\mathbf{x}^\star) \cap \operatorname{rint} \Delta.$$

Pinsker's inequality guarantees this containment for sufficiently small ρ, and continuity of KL at $\mathbf{x}^\star$ puts every sufficiently close interior initialization in $\mathcal{K}_\rho$. Initialize $\mathbf{x}^0 \in \mathcal{K}_\rho$. The preceding lemma

implies inductively that $\mathbf{x}^t \in \mathcal{K}_\rho$ for every t, so all three local inequalities remain valid. Moreover, Lemma 3 gives $L = 1 - r - \beta$. Because $\beta \geq 0$, the PR step $\eta = 1$ satisfies $1/\eta \geq L$, and, when $\alpha = 0$, the ER step $\eta = 1/(1-r)$ satisfies $1/\eta = 1 - r \geq L$. Thus the descent calculation at the start of this proof shows that $f(\mathbf{x}^t)$ is nonincreasing.

The mirror-descent three-point inequality and (18) now give

$$\begin{aligned} f(\mathbf{x}^{t+1}) &\leq f(\mathbf{x}^t) + \langle \nabla f(\mathbf{x}^t), \mathbf{x}^{t+1} - \mathbf{x}^t \rangle + \frac{1}{\eta}\mathrm{KL}(\mathbf{x}^{t+1}, \mathbf{x}^t) \\ &\leq f(\mathbf{x}^t) + \langle \nabla f(\mathbf{x}^t), \mathbf{x}^\star - \mathbf{x}^t \rangle + \frac{1}{\eta}\big(\mathrm{KL}(\mathbf{x}^\star, \mathbf{x}^t) - \mathrm{KL}(\mathbf{x}^\star, \mathbf{x}^{t+1})\big) \\ &\leq f(\mathbf{x}^\star) + \frac{1}{\eta}\big(\mathrm{KL}(\mathbf{x}^\star, \mathbf{x}^t) - \mathrm{KL}(\mathbf{x}^\star, \mathbf{x}^{t+1})\big)\,. \end{aligned}$$

Summing from $t = 0$ to $T - 1$ yields

$$\sum_{t=0}^{T-1} \big(f(\mathbf{x}^{t+1}) - f(\mathbf{x}^\star)\big) \leq \frac{\mathrm{KL}(\mathbf{x}^\star, \mathbf{x}^0)}{\eta}.$$

Since the objective values are nonincreasing,

$$f(\mathbf{x}^T) - f(\mathbf{x}^\star) \leq \frac{\mathrm{KL}(\mathbf{x}^\star, \mathbf{x}^0)}{\eta T}.$$

Finally, $\mathcal{K}_\rho$ is convex, and integrating (19) along the segment from $\mathbf{x}^\star$ to any $\mathbf{x} \in \mathcal{K}_\rho$ gives

$$f(\mathbf{x}) - f(\mathbf{x}^\star) \geq \frac{\gamma}{2}\|\mathbf{x} - \mathbf{x}^\star\|_2^2.$$

Consequently,

$$\|\mathbf{x}^T - \mathbf{x}^\star\|_2 \leq \sqrt{\frac{2\,\mathrm{KL}(\mathbf{x}^\star, \mathbf{x}^0)}{\gamma\eta T}} = \mathcal{O}(T^{-1/2}),$$

which proves the mirror-descent convergence and rates in Theorem 5.

For general ER, strict local minimality and local convexity allow a convex neighborhood in which $\mathbf{s}^\sharp$ is the unique minimizer. Theorem 8 then gives convergence to $\mathbf{s}^\sharp$ from sufficiently close initial pairs. For the geometric-rate claim, when $A := r + \alpha + \beta > 0$, (24) shows that C^2 regularity of Φ_{BR} implies the required local smoothness of G; the proof of Corollary 1 therefore applies and gives $\|\mathbf{s}^t - \mathbf{s}^\sharp\|_2 \leq Cq^t$ after adjusting the constants. When $A = 0$, ER reaches the unique minimizer in one update, so the claim also holds.

D.4 Proof of Proposition 1

Proof of Proposition 1: Fix $j \neq i$ and write

$$A := r + \alpha + \beta, \qquad \lambda := \alpha r, \qquad \eta := \frac{1}{1-r}, \qquad R_t := \frac{s_j^t}{s_i^t}, \qquad m := \frac{\mu_j}{\mu_i} \geq 1.$$

Positive ranking weights and best responses preserve interior exposures. Condition (14) forces $A > 0$. Set

$$e := \max\left\{0, \frac{1 - r - \beta + \varepsilon}{A}\right\}.$$

The feedback condition and monotonicity of ζ_j imply $x\zeta_j'(x)/\zeta_j(x) \geq e$. Integrating this inequality and using $\zeta_j(y) \geq \zeta_i(y)$ gives

$$\frac{\zeta_j(x)}{\zeta_i(y)} \geq \frac{\zeta_j(x)}{\zeta_j(y)} \geq \left(\frac{x}{y}\right)^e, \qquad 0 < y \leq x < 1. \tag{22}$$

For PR, (7) and $q_k^t = \zeta_k(s_k^t)$ give, whenever $t \geq 1$ and $R_t \geq 1$,

$$R_{t+1} = mR_t^{r+\beta}\left(\frac{\zeta_j(s_j^t)}{\zeta_i(s_i^t)}\right)^A \geq mR_t^{r+\beta+Ae} \geq mR_t^{1+\varepsilon}.$$

Starting from $R_1 \geq 1$, this preserves $R_t \geq 1$. If $m > 1$, then $R_t \geq m^{t-1} R_1 \to \infty$. If instead $\varepsilon > 0$ and $R_1 > 1$, then $R_t \geq R_1^{(1+\varepsilon)^{t-1}} \to \infty$.

For ER and $t \geq 2$, (8) gives

$$R_{t+1} = m^{\eta} R_t^{\eta\beta} \left(\frac{\zeta_j(s_j^t)}{\zeta_i(s_i^t)} \right)^{\eta(A-\lambda)} \left(\frac{\zeta_j(s_j^{t-1})}{\zeta_i(s_i^{t-1})} \right)^{\eta\lambda}.$$

Here $A - \lambda = r + \beta + \alpha(1-r) > 0$ and $\lambda \geq 0$. For ordered current and previous exposures, (22) therefore yields

$$R_{t+1} \geq m^{\eta} R_t^{\eta[\beta+(A-\lambda)e]} R_{t-1}^{\eta\lambda e}.$$

Both exponents are nonnegative, and their sum satisfies

$$\eta(\beta + Ae) \geq 1 + \eta\varepsilon.$$

When $\lambda = 0$, the lag factor disappears, so for $t \geq 1$, $R_{t+1} \geq m^{\eta} R_t^{1+\eta\varepsilon}$. The PR argument then applies from R_1.

When $\lambda > 0$, put $M_t := \min\{R_t, R_{t-1}\}$ for $t \geq 2$. The post-startup ordering gives $M_2 \geq 1$, and induction yields

$$R_{t+1} \geq m^{\eta} M_t^{1+\eta\varepsilon}, \qquad M_{t+1} \geq M_t \geq 1.$$

Applying the first inequality at two consecutive updates gives

$$M_{t+2} \geq m^{\eta} M_t^{1+\eta\varepsilon}.$$

If $m > 1$, the minimum grows by at least the factor $m^{\eta} > 1$ every two updates. If instead $\varepsilon > 0$ and $R_1, R_2 > 1$, then $M_2 > 1$ and the power $1 + \eta\varepsilon > 1$ likewise forces $M_t \to \infty$. Thus $R_t \to \infty$ in both cases.

For either dynamic, $s_i^t = s_j^t / R_t \leq 1/R_t \to 0$. If the assumptions hold against every other creator, summing their vanishing exposures gives $s_j^t = 1 - \sum_{i \neq j} s_i^t \to 1$. □

E Additional Empirical Results

E.1 Objective landscape with changing β

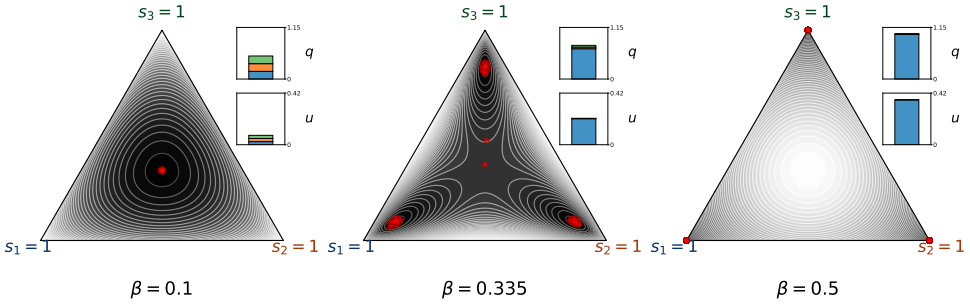

Figure 3: Stochastic PR simulations with $\zeta(s) = s\exp(s-1)$, $r = \alpha = 0.06$, and $\mu_j \equiv 1/3$, for $\beta \in \{0.10, 0.335, 0.50\}$ (left to right). Red dots show terminal exposure profiles from 200 runs over potential contours; darker shading marks lower potential. Insets show representative qualities and profits as in Figure 2.

Using the response function and simulation protocol of Section 6, Figure 3 varies popularity bias while fixing $\alpha = 0.06$. The three settings yield globally convex, nonconvex, and globally concave potentials, respectively. At $\beta = 0.335$, most terminal exposures cluster near the three symmetric local minima; two trajectories remain near the center after 1,000 updates.

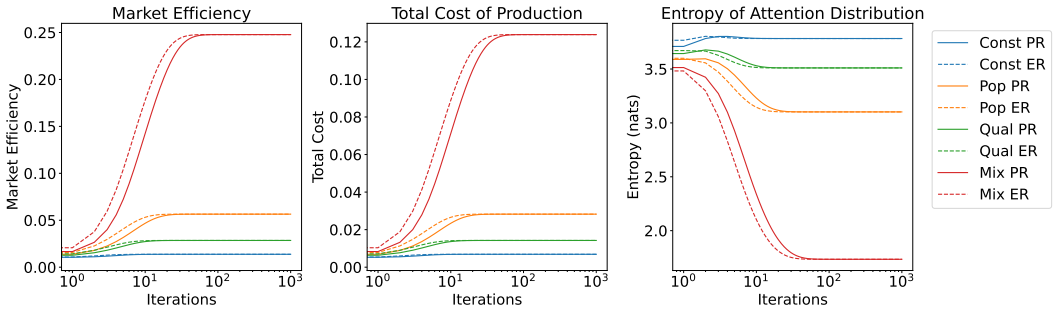

Figure 4: Demonstration of simulation results with $r = 0.3, \alpha = \beta = 0.1$. Each simulation is run for 1000 epochs. Each line is the average value of 100 runs. The shaded area indicates the standard deviations of the metrics across 100 runs. *Left*: Market efficiency – higher value indicates higher expected quality against the attention distribution $\mathbf{s}^t$, *Middle*: Total production cost across all creators, *Right*: Shannon Entropy of the attention distribution $\mathbf{s}^t$ – higher values indicate a more uniform distribution. An entropy of zero means $s_j^t = 1$ for some j and $s_i^t = 0$ for all $i \neq j$.

E.2 Market efficiency, cost of production and diversity

We simulate an instance of two-sided market with 50 producers under all four ranking strategies presented in Definition 3 w.r.t. PR/ER dynamics. The cost function of each producer is set to be $c_j(q_j) = p_j \cdot q_j^2$ with p_j sampled uniformly from $[0.5, 5]$. We set $v_j = \mu_j = \frac{1}{50}$ for all $j \in \mathcal{C}$. We select 100 initialisations $\mathbf{s}^0$ uniformly at random and report the average values of these 100 runs of the following three metrics:

- **Market efficiency** ($\sum_{j \in \mathcal{C}} s_j^t q_j^t$): The expected quality under the attention distribution, equivalently the probability that a user at epoch successfully purchases an item.
- **Total cost of production** ($\sum_{j \in \mathcal{C}} c_j(q_j^t)$): The aggregate input from all producers, indicating how strongly the market dynamics incentivize creator effort.
- **Shannon entropy** ($-\sum_{j \in \mathcal{C}} s_j^t \log s_j^t$): A measure of polarization in the attention distribution.

The results for $r = 0.3$, $\alpha = \beta = 0.1$ are shown in Figure 4. In all cases, the three metrics converge after about 10^2 iterations. Relative to their initial values, both the market efficiency and total production cost increase significantly, whereas the Shannon entropy decreases. This indicates that the dynamics enhance market efficiency and incentivize production but also polarize the attention distribution, consistent with the theoretical insights from the market's potential functions.

From the left and middle panels of Figure 4, the mixed ranking strategy achieves the highest market efficiency, followed by the popularity and quality ranking strategies, while the constant ranking strategy performs the worst. This contrasts with generic trial-offer markets studied in Van Hentenryck et al. (2016); Zhu et al. (2023), where the quality ranking strategy outperforms all alternatives. The findings support our theoretical prediction that all ranking strategies enhance market efficiency yet polarize attention distribution. Since the mixed ranking strategy combines both quality and popularity factors, it achieves the greatest efficiency and lowest entropy.

As shown in the right panel of Figure 4, the Shannon entropy remains positive across all ranking strategies, indicating convergence to equilibria in which multiple producers hold positive market shares. Moreover, the standard deviations of these metrics converge to zero, demonstrating that different initializations lead to the same equilibrium under each ranking strategy—a reflection of dynamic stability. Finally, all ER dynamics converge faster than PR dynamics, affirming our theoretical result that ER corresponds to larger learning rates.

F A Bregman Proximal Proof of Convergence for General Mixed-Ranking ER

General mixed-ranking ER has the algebraic form of the non-Euclidean PIAG update of Zhang et al. (2021, Eq. (3.2)). Their convergence theorems do not apply directly: the objective split reproducing ER includes a concave component, whereas their Assumption B1 requires convex components. We instead adapt the Bregman three-point and Lyapunov approach used in their Section 4. An exact proximal reformulation specific to this market yields the convergence results below, with a self-contained proof.

F.1 Objective splitting and an exact proximal reformulation

Throughout this section, assume $n \geq 2$, $0 \leq r < 1$, $\alpha, \beta \geq 0$, and $\mu_j > 0$. Each response ζ_j is continuous and strictly increasing on $[0,1]$, with $\zeta_j(0) = 0$, and satisfies $\int_0^1 |\log \zeta_j(z)|\,\mathrm{d}z < \infty$. The last condition makes the potential finite and continuous on the closed simplex. Define

$$w(\mathbf{s}) := \sum_j s_j \log s_j, \qquad G(\mathbf{s}) := \sum_j \int_0^{s_j} \log \zeta_j(z)\,\mathrm{d}z,$$
$$A := r + \alpha + \beta, \qquad \lambda := \alpha r, \qquad \eta := \frac{1}{1-r}. \tag{23}$$

We use $0 \log 0 = 0$ and write $\Phi := \Phi_{\mathrm{BR}}$, so that

$$\Phi(\mathbf{s}) = -\langle \log \boldsymbol{\mu}, \mathbf{s}\rangle + (1 - r - \beta) w(\mathbf{s}) - A G(\mathbf{s}). \tag{24}$$

Both G and w are strictly convex. For $K \in \{G, w\}$, let

$$D_K(\mathbf{u}, \mathbf{v}) := K(\mathbf{u}) - K(\mathbf{v}) - \langle \nabla K(\mathbf{v}), \mathbf{u} - \mathbf{v}\rangle, \qquad \mathbf{u} \in \Delta, \quad \mathbf{v} \in \Delta^\circ.$$

In particular, $D_w(\mathbf{u}, \mathbf{v}) = \mathrm{KL}(\mathbf{u} \| \mathbf{v})$.

With the simplex constraint explicit, the PIAG regularizer is $h \equiv 0$. Take two components $f_0 := \Phi + \lambda G$ and $f_1 := -\lambda G$, with delays zero and one, respectively. The resulting aggregate gradient is

$$\mathbf{g}^t = \nabla f_0(\mathbf{s}^t) + \nabla f_1(\mathbf{s}^{t-1}) = \nabla \Phi(\mathbf{s}^t) + \lambda\big(\nabla G(\mathbf{s}^t) - \nabla G(\mathbf{s}^{t-1})\big).$$

The step minimizing $\langle \mathbf{g}^t, \mathbf{s} - \mathbf{s}^t\rangle + \eta^{-1} D_w(\mathbf{s}, \mathbf{s}^t)$ over Δ is exactly Remark 1. The following second representation is what permits convergence analysis despite the concavity of f_1.

Proposition 2 (Exact Bregman proximal representation)**.** *For every interior pair $(\mathbf{s}^t, \mathbf{s}^{t-1})$, the reduced ER update is the unique solution of*

$$\mathbf{s}^{t+1} = \operatorname*{arg\,min}_{\mathbf{s} \in \Delta} \left\{ \Phi(\mathbf{s}) + (A - \lambda) D_G(\mathbf{s}, \mathbf{s}^t) + \lambda D_G(\mathbf{s}, \mathbf{s}^{t-1}) + \beta D_w(\mathbf{s}, \mathbf{s}^t) \right\}. \tag{25}$$

This solution belongs to Δ°, even when Φ is nonconvex.

Proof. Expand the divergences and use (24). Up to a constant independent of $\mathbf{s}$, the minimized expression becomes

$$(1 - r) w(\mathbf{s}) - \left\langle \log \boldsymbol{\mu} + (A - \lambda)\nabla G(\mathbf{s}^t) + \lambda \nabla G(\mathbf{s}^{t-1}) + \beta \nabla w(\mathbf{s}^t), \mathbf{s} \right\rangle.$$

The coefficient $1 - r$ is positive, and all coefficients of the linear term are finite. Entropy minimization on the simplex therefore has a unique interior solution, given by

$$s_j^{t+1} \propto \mu_j^\eta \zeta_j(s_j^t)^{\eta(A-\lambda)} \zeta_j(s_j^{t-1})^{\eta\lambda} (s_j^t)^{\eta\beta}.$$

This is (8) after best response. All proximal weights are nonnegative; moreover, $A - \lambda = r + \beta + \alpha(1 - r) > 0$ whenever $A > 0$. □

Proposition 3 (Potential characterization of equilibrium)**.** *For $\mathbf{s}^\star \in \Delta^\circ$, the following are equivalent:*

(i) $\mathbf{s}^\star$ is a constrained stationary point of Φ on Δ;

(ii) $(\mathbf{s}^\star, \mathbf{s}^\star)$ is a fixed state of the reduced ER recurrence;

(iii) $\mathbf{s}^\star$ *is the exposure vector of an interior 2SAME.*

Consequently, every interior local or global minimizer of Φ *induces an equilibrium.*

Proof. By (24), constrained stationarity means that $\mu_j \zeta_j(s_j^\star)^A (s_j^\star)^{r+\beta-1}$ is constant across j. Equivalently,

$$s_j^\star \propto \mu_j \zeta_j(s_j^\star)^A (s_j^\star)^{r+\beta}. \tag{26}$$

This is the fixed-point equation of (8) after best response, proving (i)⇔(ii).

Set $q_j^\star := \zeta_j(s_j^\star)$ and define

$$\phi_j^\star := \frac{s_j^\star q_j^\star}{\sum_i s_i^\star q_i^\star}, \qquad v_j^\star := \frac{\mu_j (q_j^\star)^\alpha (\phi_j^\star)^\beta}{\sum_i \mu_i (q_i^\star)^\alpha (\phi_i^\star)^\beta}.$$

Equation (26) gives $s_j^\star \propto v_j^\star (\phi_j^\star)^r$, and hence $(v_j^\star q_j^\star)^\eta \propto \phi_j^\star$. Together with $q_j^\star = \zeta_j(s_j^\star)$, these are the three fixed-point conditions for an interior 2SAME. Reversing the calculation proves the converse. □

F.2 A Lyapunov inequality with one-step memory

For a comparison point $\mathbf{u} \in \Delta$, define

$$\mathcal{E}_t(\mathbf{u}) := A D_G(\mathbf{u}, \mathbf{s}^t) + \beta D_w(\mathbf{u}, \mathbf{s}^t) + \lambda D_G(\mathbf{u}, \mathbf{s}^{t-1}), \tag{27}$$

$$\mathcal{R}_t := (A-\lambda) D_G(\mathbf{s}^{t+1}, \mathbf{s}^t) + \lambda D_G(\mathbf{s}^{t+1}, \mathbf{s}^{t-1}) + \beta D_w(\mathbf{s}^{t+1}, \mathbf{s}^t) \geq 0. \tag{28}$$

Lemma 6 (Bregman Lyapunov inequality)**.** *If* Φ *is convex on a convex set containing* $\mathbf{u}$ *and* $\mathbf{s}^{t+1}$*, then*

$$\Phi(\mathbf{s}^{t+1}) - \Phi(\mathbf{u}) + \mathcal{R}_t \leq \mathcal{E}_t(\mathbf{u}) - \mathcal{E}_{t+1}(\mathbf{u}). \tag{29}$$

Proof. Write $\mathbf{z} = \mathbf{s}^{t+1}$. Interior optimality in (25), paired with $\mathbf{z} - \mathbf{u}$, gives

$$\begin{aligned}\langle \nabla\Phi(\mathbf{z}), \mathbf{z}-\mathbf{u}\rangle &= (A-\lambda)\langle \nabla G(\mathbf{s}^t) - \nabla G(\mathbf{z}), \mathbf{z}-\mathbf{u}\rangle \\ &\quad + \lambda\langle \nabla G(\mathbf{s}^{t-1}) - \nabla G(\mathbf{z}), \mathbf{z}-\mathbf{u}\rangle + \beta\langle \nabla w(\mathbf{s}^t) - \nabla w(\mathbf{z}), \mathbf{z}-\mathbf{u}\rangle.\end{aligned}$$

The simplex multiplier vanishes because $\sum_j (z_j - u_j) = 0$. Apply the three-point identity

$$\langle \nabla K(\mathbf{v}) - \nabla K(\mathbf{z}), \mathbf{z}-\mathbf{u}\rangle = D_K(\mathbf{u}, \mathbf{v}) - D_K(\mathbf{u}, \mathbf{z}) - D_K(\mathbf{z}, \mathbf{v})$$

to each term. Their sum is $\mathcal{E}_t(\mathbf{u}) - \mathcal{E}_{t+1}(\mathbf{u}) - \mathcal{R}_t$. Finally, convexity gives $\Phi(\mathbf{z}) - \Phi(\mathbf{u}) \leq \langle \nabla\Phi(\mathbf{z}), \mathbf{z}-\mathbf{u}\rangle$. □

The memory term in (27) absorbs the delayed contribution exactly. Thus no bound on a sum of successive KL divergences, and no additional step-size reduction, is needed in this argument. The quantity that decreases is $\mathcal{E}_t$ relative to a minimizer; the potential itself need not decrease at each iteration.

F.3 Global convergence under convexity

We first record two elementary boundary properties used to prove convergence of the full exposure sequence. Strict convexity and continuity of G imply

$$D_G(\mathbf{x}, \mathbf{y}) \geq G(\mathbf{x}) + G(\mathbf{y}) - 2G\big((\mathbf{x}+\mathbf{y})/2\big). \tag{30}$$

The right side is continuous on the compact set $\Delta \times \Delta$ and strictly positive off the diagonal. Consequently, $D_G(\mathbf{x}^k, \mathbf{y}^k) \to 0$ implies $\|\mathbf{x}^k - \mathbf{y}^k\|_2 \to 0$. Also, for fixed $\mathbf{u} \in \Delta$ and interior $\mathbf{v}^k \to \mathbf{u}$,

$$D_G(\mathbf{u}, \mathbf{v}^k) \to 0, \qquad D_w(\mathbf{u}, \mathbf{v}^k) \to 0. \tag{31}$$

For a coordinate with $u_j = 0$, put $g_j(v) = \int_0^v \log \zeta_j(z)\,\mathrm{d}z$. Its contribution to D_G is $-g_j(v) + v \log \zeta_j(v)$, which tends to zero: integrability gives $g_j(v) \to 0$, and, for small v, $g_j(v) \leq v \log \zeta_j(v) \leq 0$. The other coordinates follow by continuity, and the entropy case follows from $v \log v \to 0$. Inequality (30) follows by applying the supporting-hyperplane inequality at $\mathbf{y}$ to the midpoint.

Theorem 7 (Global convergence of convex mixed-ranking ER)**.** *Under the assumptions of this section, suppose Φ is convex on Δ. From every pair $\mathbf{s}^0, \mathbf{s}^1 \in \Delta^\circ$, the reduced ER sequence converges to a global minimizer $\bar{\mathbf{s}} \in \Delta$ of Φ. For every $\mathbf{s}^\star \in \arg\min_\Delta \Phi$ and $T \geq 1$,*

$$\sum_{t=1}^{T} \left[\Phi(\mathbf{s}^{t+1}) - \Phi(\mathbf{s}^\star)\right] \leq \mathcal{E}_1(\mathbf{s}^\star). \tag{32}$$

In particular, the average $\widehat{\mathbf{s}}^T := T^{-1}\sum_{t=1}^{T} \mathbf{s}^{t+1}$ satisfies

$$0 \leq \Phi(\widehat{\mathbf{s}}^T) - \min_\Delta \Phi \leq \frac{\mathcal{E}_1(\mathbf{s}^\star)}{T}. \tag{33}$$

Strict convexity makes the limiting minimizer unique. If the limit is interior, it induces an interior 2SAME by Proposition 3.

Proof. If $A = 0$, nonnegativity forces $r = \alpha = \beta = 0$; the next exposure is $s_j = \mu_j / \sum_i \mu_i$, the unique minimizer of Φ, and the claims follow. Suppose $A > 0$. Continuity on compact Δ guarantees a minimizer $\mathbf{s}^\star$. Apply (29) with this comparison point. The objective gaps, $\mathcal{E}_t$, and $\mathcal{R}_t$ are nonnegative, so telescoping gives (32) and $\sum_{t\geq 1} \mathcal{R}_t < \infty$. Thus $\Phi(\mathbf{s}^t) \to \min_\Delta \Phi$. Jensen's inequality gives (33).

Because $A - \lambda > 0$, the summability of $\mathcal{R}_t$ gives $D_G(\mathbf{s}^{t+1}, \mathbf{s}^t) \to 0$. By (30), consecutive exposures approach each other. Choose a subsequence $\mathbf{s}^{t_k} \to \bar{\mathbf{s}}$ in compact Δ. Then $\mathbf{s}^{t_k - 1} \to \bar{\mathbf{s}}$ as well, and continuity of Φ makes $\bar{\mathbf{s}}$ a global minimizer. Equation (31) implies $\mathcal{E}_{t_k}(\bar{\mathbf{s}}) \to 0$. Applying (29) with $\mathbf{u} = \bar{\mathbf{s}}$ shows that $\mathcal{E}_t(\bar{\mathbf{s}})$ is nonincreasing, hence tends to zero along the entire sequence. Since it bounds $AD_G(\bar{\mathbf{s}}, \mathbf{s}^t)$ from above, (30) gives $\mathbf{s}^t \to \bar{\mathbf{s}}$. □

The initial energy also admits an explicit KL bound. For $\mathbf{u} \in \Delta$ and $\mathbf{v} \in \Delta^\circ$, convexity and (24) give

$$0 \leq \Phi(\mathbf{u}) - \Phi(\mathbf{v}) - \langle \nabla\Phi(\mathbf{v}), \mathbf{u} - \mathbf{v}\rangle = (1 - r - \beta) D_w(\mathbf{u}, \mathbf{v}) - AD_G(\mathbf{u}, \mathbf{v}).$$

Using $0 \leq \lambda \leq A$ and $\beta \geq 0$ in (27), it follows that

$$\begin{aligned} \mathcal{E}_1(\mathbf{u}) &\leq (1 - r)\,\mathrm{KL}(\mathbf{u}\|\mathbf{s}^1) + (1 - r - \beta)\,\mathrm{KL}(\mathbf{u}\|\mathbf{s}^0) \\ &\leq (1 - r)\left[\mathrm{KL}(\mathbf{u}\|\mathbf{s}^1) + \mathrm{KL}(\mathbf{u}\|\mathbf{s}^0)\right]. \end{aligned} \tag{34}$$

Combining this bound with (33) proves (13).

Convexity here is convexity restricted to the simplex, not a requirement that every ambient coordinate Hessian entry be nonnegative. Boundary minimizers are permitted, and the rate in (33) is for averaged iterates; it is not a claim of monotone or $O(1/T)$ last-iterate decay.

F.4 Local convergence and a geometric rate under positive curvature

Theorem 8 (Local convergence under local convexity)**.** *Under the assumptions of this section, let $\mathbf{s}^\dagger \in \Delta^\circ$ be a constrained stationary point of Φ. Suppose Φ is convex on a relatively open convex neighborhood $U \subset \Delta^\circ$ of $\mathbf{s}^\dagger$. There is a neighborhood $\mathcal{N}$ of $(\mathbf{s}^\dagger, \mathbf{s}^\dagger)$ such that every initialization $(\mathbf{s}^1, \mathbf{s}^0) \in \mathcal{N}$ stays in U and converges to a minimizer of Φ on U, hence to an interior local minimum and an interior 2SAME. If $\mathbf{s}^\dagger$ is the only minimizer in U, the limit is $\mathbf{s}^\dagger$.*

Proof. The case $A = 0$ is immediate as above. First take $A > 0$ and $\lambda > 0$. Stationarity and local convexity make $\mathbf{s}^\dagger$ a minimizer on U, and Proposition 3 makes $(\mathbf{s}^\dagger, \mathbf{s}^\dagger)$ a fixed state of the continuous ER map. Choose a closed relative ball K centered at $\mathbf{s}^\dagger$ with $K \subset U$. For an interior pair $(\mathbf{x}, \mathbf{y})$, set

$$\mathcal{E}(\mathbf{x}, \mathbf{y}; \mathbf{s}^\dagger) := AD_G(\mathbf{s}^\dagger, \mathbf{x}) + \beta D_w(\mathbf{s}^\dagger, \mathbf{x}) + \lambda D_G(\mathbf{s}^\dagger, \mathbf{y}).$$

By (30), sufficiently small energy forces both coordinates of the state arbitrarily close to $\mathbf{s}^\dagger$. Continuity of the update at its fixed state therefore lets us choose $\varepsilon > 0$ so that $\mathcal{E}(\mathbf{x}, \mathbf{y}; \mathbf{s}^\dagger) \leq \varepsilon$ implies both $(\mathbf{x}, \mathbf{y}) \in K \times K$ and that the next exposure lies in U.

Initialize with $\mathcal{E}_1(\mathbf{s}^\dagger) < \varepsilon$. Whenever this energy bound holds, the next exposure lies in U, so (29) applies and its objective gap is nonnegative. Hence $\mathcal{E}_{t+1}(\mathbf{s}^\dagger) \leq \mathcal{E}_t(\mathbf{s}^\dagger)$. Induction proves that the

energy sublevel is invariant and all exposures remain in compact K. It also proves summability of the objective gaps and of $\mathcal{R}_t$.

Every cluster point in K consequently minimizes Φ on U, and consecutive exposures approach each other. For any such cluster point $\bar{\mathbf{s}}$, the argument in the global proof gives $\mathcal{E}_{t_k}(\bar{\mathbf{s}}) \to 0$ along its subsequence. Local convexity permits (29) with this comparison point at every step, so $\mathcal{E}_t(\bar{\mathbf{s}}) \to 0$ and the whole sequence converges to $\bar{\mathbf{s}}$. When $\lambda = 0$, the recurrence and energy depend only on the current exposure; the same invariant-sublevel argument applies to this single coordinate, without restricting the unused previous exposure. □

Corollary 1 (Local geometric convergence)**.** *Suppose the assumptions of Theorem 8 hold with $A > 0$, each ζ_j is continuously differentiable near $s_j^\dagger$, and the Hessian of Φ at $\mathbf{s}^\dagger$ is positive definite on $\{\mathbf{v} : \sum_j v_j = 0\}$. Then sufficiently close initial states converge geometrically to $\mathbf{s}^\dagger$. More precisely, choose the invariant neighborhood small enough that, on its containing convex ball K, Φ is ν-strongly convex relative to the simplex, and $\nabla^2 G \preceq L_G I$, $\nabla^2 w \preceq L_w I$. With $L := AL_G + \beta L_w > 0$ and $q := L/(L+\nu) < 1$,*

$$\mathcal{E}_{t+2}(\mathbf{s}^\dagger) \le q\mathcal{E}_t(\mathbf{s}^\dagger), \qquad \|\mathbf{s}^{t+1} - \mathbf{s}^\dagger\|_2^2 \le \frac{2\mathcal{E}_1(\mathbf{s}^\dagger)}{\nu} q^{\lfloor (t-1)/2 \rfloor}, \quad t \ge 1. \tag{35}$$

Proof. Continuity of the Hessian and positive definiteness on the tangent space give such a ball and $\nu > 0$. Smoothness and compactness inside Δ° give the finite upper bounds. Write $d_t := \|\mathbf{s}^t - \mathbf{s}^\dagger\|_2$ and $\mathcal{E}_t := \mathcal{E}_t(\mathbf{s}^\dagger)$. Strong convexity and stationarity give $\Phi(\mathbf{s}^t) - \Phi(\mathbf{s}^\dagger) \ge \nu d_t^2/2$. Applying (29) at two successive times yields

$$\mathcal{E}_t - \mathcal{E}_{t+2} \ge \frac{\nu}{2}(d_{t+1}^2 + d_{t+2}^2).$$

The Hessian upper bounds give $D_G(\mathbf{s}^\dagger, \mathbf{s}^t) \le L_G d_t^2/2$ and the analogous bound for D_w. Since $\lambda \le A$,

$$\mathcal{E}_{t+2} \le \frac{L}{2}(d_{t+2}^2 + d_{t+1}^2) \le \frac{L}{\nu}(\mathcal{E}_t - \mathcal{E}_{t+2}).$$

Rearranging proves the first inequality of (35). Monotonicity of the energy gives $\mathcal{E}_t \le q^{\lfloor (t-1)/2 \rfloor}\mathcal{E}_1$. Finally, the one-step inequality gives $\nu d_{t+1}^2/2 \le \mathcal{E}_t$. □

These results concern the original ER step size and nonnegative ranking parameters. Local convexity supplies an invariant neighborhood through the energy argument, while positive tangent curvature supplies the local rate. No convergence claim for arbitrary nonconvex potentials or signed ranking parameters is made here.