Theorems and Complete Proof

1 Joint Optimization Objective

In this section, we draw inspiration from redistribution theory in economics, and propose an objective function to optimize diversity and fairness while maintaining accuracy. We explain the diversity and fairness objectives separately, and then discuss some nice properties of our objective function.

Overall objective. As illustrated in the Introduction, diversity corresponds to the Sufficientarian principle while fairness corresponds to the Rawlsian principle. Drawing inspiration from this correspondence, we propose an objective function to optimize both diversity and fairness in recommender systems. We design the function to be piecewise, where diversity is optimized in the first stage and fairness becomes the target in the latter stage.

Define $f\colon [0,1]^{\mathcal{I}} \to \mathbb{R}$ for utililty vector $\boldsymbol{v} \in [0,1]^{\mathcal{I}}$ to be the full objective function:

$$f(v) = \sum_{i=1}^{\mathcal{I}} \mathcal{F}(v_i), \tag{1}$$

where $\mathcal{F}\colon [0,1] \to \mathbb{R}$ is the function for the utility $v_i \in [0,1]$ of a single item i:

$$\mathcal{F}(v_i) = \underbrace{\lambda\, C^*_\delta(v_i)}_{\text{diversity objective}} + \underbrace{U^*_\alpha(v_i)}_{\text{fairness objective}}, \tag{2}$$

where $C^*_\delta(\cdot)$ and $U^*_\alpha(\cdot)$ are diversity and fairness terms, respectively. $\mathcal{F}$ is a piecewise function divided into diversity stage and fairness stage, as detailed in Sections 1.1 and 1.2. $\delta \in (0,1]$ is the coverage threshold to measure diversity. It is also the dividing point between two stages. $\alpha \in [0,\infty)$ represents the tax rate, where a higher value results in fairer results. Coefficient λ is to ensure the smoothness of the function at the dividing point, which we will discuss in Section 1.3.

1.1 Diversity Objective

In this section, we introduce how we design the diversity function. First, we define the diversity objective in the following form:

$$C^*_\delta(v) = \begin{cases} 1 - e^{-v/\delta}, & 0 \le v < \delta, \\ 1 - e^{-1}, & \delta \le v \le 1. \end{cases}$$

which takes item utility $\boldsymbol{v}$ as input, and outputs how much it contributes to diversity with coverage threshold set to δ.

Our diversity function aligns with the Bass diffusion model with external influence [2], which models the growth of production adoptions. This function form incentivizes the surge of item utilities to reach the coverage threshold at the early stage of online recommendation, analogous to the Sufficientarian principle in redistribution economics. It "saturates" and falls into a constant when the item is already covered, which aligns with the definition of aggregate diversity. Concavity exhibits a diminishing return growth pattern and encourages breadth (exposing many items) over depth (heavily exposing a few items).

Diversity axioms in RS. To satisfy a set of axioms which commonly hold for diversity, Theorem 1 shows that such an objective is unique in RS.

Based on the concept of aggregate diversity and a rich line of related work [12, 14, 15], the diversity function $C_\delta(\cdot)$ should satisfy the following axioms.

A1 (Monotonicity). $C'_\delta(v) \geq 0, \forall v \geq 0$. More utility never *reduces* diversity.

A2 (Law of diminishing returns). $C''_\delta(v) \leq 0, \forall v \geq 0$. The marginal gain in diversity decreases as accumulated utility grows. Diminishing return is a well-understood fundamental principle of economics [3].

A3 (Capped potential). $C_\delta(0) = 0, \lim_{v\to\infty} C_\delta(v) = 1$. Diversity is zero before any exposure happens, and has an upper ceiling of 1.

A4 (Memoryless gains). For all $v, w \geq 0$, $C_\delta(v+w) = C_\delta(v) + \big(1 - C_\delta(v)\big)\,\psi(w)$, where $\psi(w)$ does not depend on v. A fixed extra utility w converts the *remaining* potential into diversity by the *same factor*, regardless of history [12, 14]. Mathematically, there exists a function $\psi : [0,\infty) \to [0,1]$ independent of v such that the gain from adding w to accumulated utility v satisfies a consistent multiplicative structure.

A5 (Threshold-invariant). $\exists C_0 \in (0,1)$ such that $C_\delta(\delta) = C_0, \forall \delta > 0$. C_0 is the same constant for all δ, independent of the threshold value.

Based on these axioms, we can show the uniqueness of the objective.

Theorem 1. *Let $C : [0,\infty) \to [0,1)$ be a C^1 function satisfying:(i) $C(0) = 0$; (ii) $C'(v) \geq 0$; (iii) for all $v, w \geq 0$, the* memoryless gains *identity $C(v+w) = C(v) + (1 - C(v))\,C(w)$ holds. Then there exists a constant $k \geq 0$ such that for all $v \geq 0$, $C(v) = 1 - e^{-kv}$.*

In particular, $C(\delta) = 1 - e^{-k\delta}$ is a constant for any threshold $\delta > 0$. For simplicity and w.l.o.g., take $k = 1/\delta$ and then $C(v) = 1 - e^{-v/\delta}$.

It is easy to verify that function $C(\cdot)$ satisfies Axioms A1 to A5.

Proof. Let the coverage level after accumulated utility v be the scalar $X(v) = C_\delta(v) \in [0,1)$. Observables are bounded Borel functions $f : [0,1) \to \mathbb{R}$ in $\mathcal{B} = \{f : \|f\|_\infty < \infty\}$, the Banach space under the sup norm $\|\cdot\|_\infty$.

For every budget chunk $w \geq 0$, define an operator

$$(T_w f)(x) = f\big(x + (1-x)C_\delta(w)\big), \quad x \in [0,1). \tag{1.1}$$

Claim. $\{T_w\}_{w\geq 0} \subset \mathcal{L}(\mathcal{B})$ is a contractive C_0-semigroup.

- **Semigroup property.**

  Compose T_{w_1} and T_{w_2},

$$\begin{aligned}(T_{w_1}T_{w_2}f)(x) &= T_{w_1}\big[z \mapsto f\big(z + (1-z)C_\delta(w_2)\big)\big](x)\\ &= f\big(x + (1-x)C_\delta(w_1) + (1-y)C_\delta(w_2)\big),\\ y &:= x + (1-x)C_\delta(w_1).\end{aligned} \tag{$\star$}$$

  Expand $1-y$,

$$1 - y = 1 - \big[x + (1-x)C_\delta(w_1)\big] = (1-x)\big[1 - C_\delta(w_1)\big].$$

  Substituting into ($\star$) yields

$$(T_{w_1}T_{w_2}f)(x) = f\Big(x + (1-x)\big[C_\delta(w_1) + (1 - C_\delta(w_1))\,C_\delta(w_2)\big]\Big).$$

  By Axiom A4,

$$C_\delta(w_1 + w_2) = C_\delta(w_1) + \big(1 - C_\delta(w_1)\big)\,C_\delta(w_2).$$

  Hence the bracket equals $C_\delta(w_1 + w_2)$, and

$$(T_{w_1}T_{w_2}f)(x) = f\big(x + (1-x)C_\delta(w_1 + w_2)\big) = (T_{w_1+w_2}f)(x),$$

  as required.

- **Strong continuity.** $T_w \to I$ pointwise as $w \downarrow 0$, and the operators are uniformly bounded, $\|T_w\|_{\mathrm{op}} \le 1$; hence $\{T_w\}$ is strongly continuous on $\mathcal{B}$.

Thus $\{T_w\}_{w \ge 0} \subset \mathcal{L}(\mathcal{B})$ is a contractive C_0-semigroup.

The infinitesimal generator of any C_0-semigroup is

$$Af = \lim_{w \downarrow 0} \frac{T_w f - f}{w}, \tag{3}$$

$$\mathcal{D}(A) = \{f \in \mathcal{B} : \text{limit exists in } \mathcal{B}\}. \tag{4}$$

Fix $f \in C^1([0,1)) \subset \mathcal{D}(A)$. Using (1.1) and $C_\delta(w) = o(1)$ as $w \to 0$, we have

$$\frac{T_w f(x) - f(x)}{w} = \frac{f\big(x + (1-x)C_\delta(w)\big) - f(x)}{w} \xrightarrow{w \to 0} f'(x)\,(1-x)\,C'_\delta(0).$$

Denote

$$\kappa := C'_\delta(0) = \frac{dC_\delta}{dv}\bigg|_{v=0} > 0 \quad \text{(A1 ensures } \kappa \text{ exists)}.$$

Hence, for all $f \in C^1$:

$$(Af)(x) = -\kappa\,(1-x)\,f'(x). \tag{1.2}$$

Equation (1.2) is the Kolmogorov backward generator of a deterministic flow with drift $-(1-x)\,\kappa$.

The semigroup satisfies

$$\frac{d}{dv} T_v f = A T_v f, \quad T_0 = I.$$

Evaluating on the identity observable $f(x) = x$ gives an ODE for $m(v) := T_v f(0) = C_\delta(v)$

$$\frac{dm}{dv} = -\kappa(1 - m(v)), \quad m(0) = 0,$$

whose solution is

$$m(v) = 1 - e^{-\kappa v}. \tag{1.3}$$

Thus every contractive semigroup compatible with A4 must produce the exponential coverage curve (1.3) for some scalar $\kappa > 0$.

Axiom A5 imposes $C_\delta(\delta) = C_0 \implies \kappa\delta = -\ln(1 - C_0)$, choose $\kappa\delta = 1$ and then $C_0 = 1 - e^{-1}$, hence $C_\delta(v) = 1 - e^{-v/\delta}$.

□

Theorem 1 shows that our objective satisfies different assumptions in RS, and can be well generalized to different situations.

1.2 Fairness Objective

In this section, we introduce how we design the fairness function based on the iso-elastic utility function and the Rawlsian principle in redistribution economics.

We design a function $U^*_\alpha(v)$ to pursue fairness in the second stage of recommendation when item utility v exceeds coverage threshold δ. From the perspective of Rawlsian redistribution, fairness in re-ranking objectives imposes taxes on richer items and redistributes to poorer items via taxation [**?**]. Drawing inspiration from iso-elastic utility function [5], we design the fairness function to be:

$$U^*_\alpha(v) = \begin{cases} U_\alpha(\delta), & 0 \le v < \delta, \\ U_\alpha(v), & \delta \le v \le 1. \end{cases}, \; U_\alpha(v) = \begin{cases} \dfrac{\gamma}{1-\alpha}\, v^{\,1-\alpha}, & \alpha \ne 1, \\ \gamma \ln v, & \alpha = 1. \end{cases},$$

where $\alpha \in [0, \infty)$ denote the tax parameter, where larger values correspond to stronger emphasis on fairness. The coefficient γ represents the importance weight of each item, and is commonly set to 1 across all items for simplicity.

The fairness component $U_\alpha(v)$ penalizes items with higher utility and redistributes that utility to those with lower utility. To maintain a constant level of overall social welfare, measured as $\sum_i U^*_\alpha(v_i)$, the utility gained by a low-performing item must be offset by a proportional reduction from a high-performing one. More precisely, for each unit of utility transferred to item j, item i must give up a tax amount of $\frac{\gamma_j}{\gamma_i}(\frac{\boldsymbol{v}_i}{\boldsymbol{v}_j})^\alpha$, ensuring the fairness-preserving transformation.

When $\alpha = 0$, the objective simplifies to a utility-weighted sum, aligning with Utilitarian principles that prioritize efficiency. Setting $\alpha = 1$ leads to the Nash solution, which equalizes the ratio of item utilities to their weights, $\boldsymbol{v}_i : \boldsymbol{v}_j = \gamma_i : \gamma_j, \forall i, j \in \mathcal{I}$. As α approaches infinity, the objective function prioritizes the least advantaged item, converging to a max-min fairness criterion of the form $\min_{i \in \mathcal{I}} \boldsymbol{v}_i$, thus promoting uniform utility distribution.

1.3 Properties of Objective Function

After introducing the fairness and diversity objectives, we analyze the properties of the overall objective function (shown in Equation 1).

In redistribution economics, smoothness of economic transition is crucial when it comes to the turning point where the leading principle switches from Sufficientarian to Rawlsian. Likewise, to make our piecewise objective function smooth, it is essential to ensure a continuous function value and derivative value between two stages.

To ensure the smoothness of our objective function, we have the following theorem:

Theorem 2 (Properties of Function f)**.** *The objective function f is L-smooth and μ-strongly concave with*

$$L = \gamma \delta^{-(\alpha+1)} \max\{e, \alpha\}, \;\; \mu = \gamma \min\{\alpha, \delta^{-(\alpha+1)}\},$$

where $\lambda \left(C_\delta^*\right)'(\delta^-) = \left(U_\alpha^*\right)'(\delta^+)$, *which leads to* $\lambda = \gamma e \delta^{1-\alpha}$.

Proof. From definitions in Section 1, we derive the first and second derivatives of $C_\delta^*(\cdot)$ and $U_\alpha^*(\cdot)$.

$$C_\delta^{*\,\prime}(v) = \begin{cases} \dfrac{e^{-v/\delta}}{\delta}, & 0 \le v < \delta, \\ 0, & \delta < v \le 1, \end{cases}$$

$$C_\delta^{*\,\prime\prime}(v) = \begin{cases} -\dfrac{e^{-v/\delta}}{\delta^2}, & 0 \le v < \delta, \\ 0, & \delta < v \le 1. \end{cases}$$

$$U'_\alpha(v) = \begin{cases} \gamma\, v^{-\alpha}, & \alpha \ne 1, \\ \dfrac{\gamma}{v}, & \alpha = 1, \end{cases} \qquad U_\alpha^{*\,\prime}(v) = \begin{cases} 0, & 0 \le v < \delta, \\ U'_\alpha(v), & \delta < v \le 1, \end{cases}$$

$$U''_\alpha(v) = \begin{cases} -\alpha\gamma\, v^{-\alpha-1}, & \alpha \ne 1, \\ -\dfrac{\gamma}{v^2}, & \alpha = 1. \end{cases} \qquad U_\alpha^{*\,\prime\prime}(v) = \begin{cases} 0, & 0 \le v < \delta, \\ U''_\alpha(v), & \delta < v \le 1, \end{cases}$$

Therefore, we have the derivatives of $\mathcal{F}(\cdot)$.

$$\mathcal{F}'(v) = \begin{cases} \dfrac{1}{\lambda}\,\dfrac{e^{-v/\delta}}{\delta}, & 0 \le v < \delta, \\ \gamma\, v^{-\alpha}, & \delta < v \le 1, \end{cases} \qquad \mathcal{F}'(\delta^-) = \mathcal{F}'(\delta^+) = \frac{e^{-1}}{\lambda\,\delta}.$$

and $\mathcal{F}'_1(v)$ is continuous at the single "kink" point $v = \delta$.

$$\mathcal{F}''(v) = \begin{cases} -\dfrac{1}{\lambda}\,\dfrac{e^{-v/\delta}}{\delta^2}, & 0 \le v < \delta, \\ -\alpha\gamma v^{-\alpha-1}, & \delta < v \le 1, \\ \text{undefined (kink)}, & v = \delta. \end{cases}$$

By deriving the maximum and minimum value of $|\mathcal{F}''(v)|$, one can quickly find that

$$L \triangleq \max_v |\mathcal{F}''(v)| = \gamma\, \delta^{-(\alpha+1)} \max\{e, \alpha\},$$

$$\mu \triangleq \min_v |\mathcal{F}''(v)| = \gamma \, \min\{\alpha, \delta^{-(\alpha+1)}\}.$$

Then we first derive Hessian bounds of the second derivative of function f, and then show the smoothness and strong concavity.

Hessian-bounds. Because f is separable, $\nabla^2 f(v) = \operatorname{diag}\big(\mathcal{F}''(v_1), \ldots, \mathcal{F}''(v_{\mathcal{I}})\big)$. For $v_i \neq \delta$ the explicit formulas above give

$$|\mathcal{F}''(v_i)| \;\leq\; L,$$

and

$$|\mathcal{F}''(v_i)| \;\geq\; \mu.$$

Hence for every $v \in [0,1]^{\mathcal{I}} \setminus \{v_i = \delta\}$,

$$-\,L\,\mathrm{I} \;\preceq\; \nabla^2 f(v) \;\preceq\; -\,\mu\,\mathrm{I}, \tag{5}$$

At $v_i = \delta$ the second derivative is undefined, but this set has Lebesgue measure 0 and will not affect the integral arguments below.

***L*-smoothness.** For any $v, w \in [0,1]^{\mathcal{I}}$ let $\phi(\tau) = \nabla f\big(w + \tau(v-w)\big)$. By the fundamental theorem of calculus,

$$\nabla f(v) - \nabla f(w) = \int_0^1 \nabla^2 f\big(w + \tau(v-w)\big)(v-w)\, d\tau.$$

Using $\|\nabla^2 f(\cdot)\|_2 \leq L$ from (5),

$$\|\nabla f(v) - \nabla f(w)\|_2 \;\leq\; \int_0^1 L\, \|v-w\|_2 \, d\tau = L\|v-w\|_2.$$

***μ*-strong concavity.** By integrating the Hessian bound a second time,

$$f(w) - f(v) - \nabla f(v)^{\top}(w-v) = \int_0^1 (1-\tau)(v-w)^{\top} \nabla^2 f\big(w + \tau(v-w)\big)(v-w)\, d\tau.$$

Because $\nabla^2 f \preceq -\mu \mathrm{I}$, the integrand is $\leq -\mu\|v-w\|_2^2$, giving $f(w) \leq f(v) + \nabla f(v)^{\top}(w-v) - \frac{\mu}{2}\|w-v\|_2^2$. The zero-measure non-differentiable slice $v_i = \delta$ does not affect the integrals, so the bounds hold on all of $[0,1]^{\mathcal{I}}$.

Therefore, we showed that the objective function f defined in Section 1 is L-smooth and μ-strongly concave.

□

As a result, function $\mathcal{F}$ belongs to class C^1 and owns the properties of strong concavity and smoothness, which are essential to the theoretical analysis of our algorithm in Section 2.3.

2 Our Approach: DivFair

Building on the design of the objective function, we propose a re-ranking algorithm, DivFair, which aims to efficiently optimize the objective proposed in Section 1. To achieve this, we formulate our multi-objective optimization problem as a mathematical program, and then propose an algorithm called DivFair for online recommendation.

2.1 Optimization Objective

First, the multi-objective re-ranking can be formulated as an integer programming problem. The optimization objective is

$$\begin{aligned}
\mathcal{W}_{\mathrm{OPT}} &= \max_{\mathbf{x}_t \in \mathcal{X}} \sum_{t=1}^{T} f\big(\mathbf{x}_t\big), \\
\text{s.t.} \quad \mathbf{e} &= \frac{1}{T}\sum_{t=1}^{T} \mathbf{x}_t \leq \boldsymbol{\tau}, \\
\mathcal{X} &= \Big\{ x \;\Big|\; x_{t,i} \in \{0,1\},\; \sum_{i \in \mathcal{I}} x_{t,i} = K,\; \forall\, t \in [1, 2, \ldots, T] \Big\}.
\end{aligned} \tag{6}$$

where f is the objective function based on the ranking list $L_K(u_t)$ for user u_t. $\mathbf{e} \in \mathbb{R}^{\mathcal{I}}$ is the average exposure vector for items across all T users, and $\tau \in \mathbb{R}^{\mathcal{I}}$ denotes the exposure upper limits.

More specifically, based on Equation 1, the diverse and fair re-ranking problem is expressed as

$$\begin{aligned}
\mathcal{W}_{\mathrm{OPT}} &= \max_{\mathbf{x}_t \in \mathcal{X}} \sum_{t=1}^{T} f\big(\mathbf{x}_t\big), \\
\text{s.t.} \quad \mathbf{e} &= \frac{1}{T}\sum_{t=1}^{T} \mathbf{x}_t \le \boldsymbol{\tau}, \\
f\big(\mathbf{x}_t\big) &= \sum_{i \in \mathcal{I}} \frac{1}{\lambda} C_{\delta}^{*}(v_{t,i}) \; + \; U_{\alpha}^{*}(v_{t,i}), \\
v_{t,i} &= \frac{1}{T}\sum_{t'=1}^{t} w_{t',i} x_{t',i}, \forall i \in \mathcal{I}, \\
\mathcal{X} &= \big\{ x \;\big|\; x_{t,i} \in \{0,1\}, \sum_{i \in \mathcal{I}} x_{t,i} = K, \; \forall\, t \in [1,2,\ldots,T] \big\}.
\end{aligned} \tag{7}$$

where $\mathbf{x}_t \in \{0,1\}^{|\mathcal{I}|}$ is the decision vector for user u_t. For each item i, $\mathbf{x}_{ti} = 1$ if it is added to the ranking list $L_K(u_t)$, otherwise, $\mathbf{x}_{ti} = 0$. $\mathcal{X}$ is the feasible region of variable $\boldsymbol{x}$. It ensures that each recommendation list is of size K. The first constraint in Equation 7 suggests that the exposures of each item i are the accumulated exposures over all users. $v_{t,i}$ is utility of item i and here we normalize it to the range $[0, 1]$.

We observe that directly optimizing the Equation 7 is NP-hard since it is a non-linear, large-scale, and integer programming [4]. The time-separable form of the optimization objective inspires us to design an online algorithm to make recommendations to user u_t without access to information beyond time step t. Next we will discuss our proposed model, DivFair, for online recommendation.

2.2 DivFair for Online Recommendation

In this section, we first derive the dual of the original problem, which conveniently turns the primal problem into a convex optimization problem. Then we propose a mirror gradient descent algorithm for efficient online learning with provable error bounds.

Dual problem. From Equation 7 we know that directly solving this integer program offline is NP-hard. However, by deriving its dual problem, we convert the primal problem into a convex optimization problem which is linear with respect to the non-integral dual variable.

Theorem 3 (Dual Problem). *The dual problem of Equation 6 can be written as:*

$$\mathcal{W}_{OPT} \le \mathcal{W}_{Dual} = \min_{\boldsymbol{\mu} \in \mathcal{U}} \Big[\max_{\mathbf{x_t} \in \mathcal{X}} \Big[\sum_{t=1}^{T} f\big(\mathbf{x}_t\big) \; - \; \boldsymbol{\mu}^{\top} \frac{1}{T} \sum_{t=1}^{T} \mathbf{x}_t \Big] \; + \; \max_{\mathbf{e} \le \boldsymbol{\tau}} \boldsymbol{\mu}^{\top} \mathbf{e} \Big], \tag{8}$$

where $\mathcal{U} = \{\boldsymbol{\mu} | \boldsymbol{\mu}_i \ge 0, \; \forall i \in [1,2,\ldots,\mathcal{I}]\} = \mathbb{R}_{+}^{\mathcal{I}}$ *is the feasible region of dual variable* $\boldsymbol{\mu}$.

Proof. Since the exposure vector $\mathbf{e} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{x}_t \le \boldsymbol{\tau}$, we can treat the $\mathbf{e}$ as an auxiliary variable and move the exposure constraints to the objective using a vector of Lagrange multipliers $\boldsymbol{\mu} \in \mathbb{R}^{|\mathcal{I}|}$. From minimax inequality, we have

$$\begin{aligned}
\mathcal{W}_{\mathrm{OPT}} &= \max_{\mathbf{x}_t \in \mathcal{X}, \mathbf{e} \le \boldsymbol{\tau}} \min_{\boldsymbol{\mu} \in \mathcal{U}} \left[\sum_{t=1}^{T} f(\mathbf{x}_t) - \boldsymbol{\mu}^{\top} \left(-\mathbf{e} + \frac{1}{T} \sum_{t=1}^{T} \mathbf{x}_t \right) \right] \\
&\le \min_{\boldsymbol{\mu} \in \mathcal{U}} \Big[\max_{\mathbf{x_t} \in \mathcal{X}} \Big[\sum_{t=1}^{T} f\big(\mathbf{x}_t\big) \; - \; \boldsymbol{\mu}^{\top} \frac{1}{T} \sum_{t=1}^{T} \mathbf{x}_t \Big] \; + \; \max_{\mathbf{e} \le \boldsymbol{\tau}} \boldsymbol{\mu}^{\top} \mathbf{e} \Big] \\
&= \mathcal{W}_{Dual}
\end{aligned}$$

Algorithm 1: DivFair

Input: User arriving order $\{u_i\}_{i=1}^{T}$, user size T, ranking size K, user-item preference score $w_{t,i}, \forall t, i$, and exposure constraints $\boldsymbol{\tau}$.

Output: The decision variables $\{\mathbf{x}_i, i = 1, 2, \ldots, N\}$

1: Initialize dual solution $\boldsymbol{\mu}_1 = 0$, phased budget $\boldsymbol{\theta}_1 = 0$, item utility $\mathbf{v}_0 = 0$, and gradient $\mathbf{g}_0 = 0$.

2: **for** $t = 1, \cdots, T$ **do**

3: User u_t arrives.

4: $\mathbf{m}_i = \begin{cases} 0, & \boldsymbol{\theta}_{ti} \le \boldsymbol{\tau_i} \\ \infty, & \text{otherwise} \end{cases}$

5: // Online recommendation

6: $\mathbf{x}_t = \arg\max_{\mathbf{x}_t \in \mathcal{X}} \left[f(\mathbf{v}_t) - (\boldsymbol{\mu}_t + \mathbf{m})^\top \mathbf{x}_t / T \right]$

7: $\approx \arg\max_{\mathbf{x}_t \in \mathcal{X}} \left[f(\mathbf{v}_{t-1}) + \nabla f(\mathbf{v}_{t-1})^\top (\mathbf{v}_t - \mathbf{v}_{t-1}) - (\boldsymbol{\mu}_t + \mathbf{m})^\top \mathbf{x}_t / T \right]$

8: $= \arg\max_{\mathbf{x}_t \in \mathcal{X}} \left[\nabla f(\mathbf{v}_{t-1})^\top (\mathbf{w}_t \odot \mathbf{x}_t)/T - (\boldsymbol{\mu}_t + \mathbf{m})^\top \mathbf{x}_t / T \right]$

9: $= \arg\max_{\mathbf{x}_t \in \mathcal{X}} \left[(\nabla f(\mathbf{v}_{t-1}) \odot \mathbf{w}_t - \boldsymbol{\mu}_t - \mathbf{m})^\top \mathbf{x}_t / T \right]$

10: $\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t + \mathbf{x}_t / T$

11: $\mathbf{e}_t = \arg\max_{\mathbf{e_t} \le \boldsymbol{\theta}_t} \boldsymbol{\mu_t}^\top \mathbf{e}_t$

12: $\mathbf{v}_t = \mathbf{v}_{t-1} + \mathbf{w}_t \odot \mathbf{x}_t / T$

13: $\mathbf{g}_t = \mathbf{e}_t - \mathbf{x}_t / T$

14: $\boldsymbol{\mu}_{t+1} = \arg\min_{\boldsymbol{\mu} \in \mathcal{U}} \left[\langle \mathbf{g}_t, \boldsymbol{\mu} \rangle + \|\boldsymbol{\mu} - \boldsymbol{\mu}_t\|^2 / (2\eta) \right]$

15: **end for**

where $\mathcal{U}$ is the feasible region of dual variable $\boldsymbol{\mu}$. To make the conjugate function $r^*(\boldsymbol{\mu}) = \max_{\mathbf{e} \le \boldsymbol{\tau}} \boldsymbol{\mu}^\top \mathbf{e}$ finite, it suffices to have non-negative vector $\boldsymbol{\mu}$. So $\mathcal{U} = \{\boldsymbol{\mu} | \boldsymbol{\mu}_i \ge 0, \ \forall i \in [1, 2, \ldots, \mathcal{I}]\} = \mathbb{R}_+^{\mathcal{I}}$. $\mathcal{U}$ is a closed and convex set with positive orthant inside its recession cone.

□

Since item utility $v_t = v_{t-1} + \frac{1}{T} w_t \odot x_t$ and $f(\cdot)$ depends on v_t, we use notation $f(v_t)$ in subsequent sections for brevity.

DivFair algorithm. Algorithm 1 illustrates DivFair, our proposed method for online optimization. DivFair is designed to operate in the dual space. It leverages linear approximation for real-time decision-making and employs a mirror-descent approach to ensure stable convergence.

DivFair operates on sequences of users of length T. For each user sequence, the algorithm maintains three key components: a dual variable $\boldsymbol{\mu}_t$, a phased budget $\boldsymbol{\theta}_t$ tracking the limits on average exposure, and the accumulated item utility vector $\mathbf{v}_t$. The procedure begins with initialization of $\boldsymbol{\mu}_1 = 0$, $\boldsymbol{\theta}_1 = 0$, and gradient $\mathbf{g}_0 = 0$. For each arriving user u_t, the algorithm first computes a penalty term $\mathbf{m}$ to ensure only items whose exposure does not exceed the phased budget are recommended. From line 6 to line 9, we provide a linear approximation of the dual function by Taylor expansion, and make a recommendation decision: $\mathbf{x}_t = \arg\max_{\mathbf{x}_t \in \mathcal{X}} \left[\nabla f(\mathbf{v}_{t-1})^\top (\mathbf{w}_t \odot \mathbf{x}_t)/T - (\boldsymbol{\mu}_t + \mathbf{m})^\top \mathbf{x}_t / T \right]$.

Through this approximation, the scoring function becomes linear with respect to $\mathbf{x_t}$. Therefore, $\mathbf{x_t}$ can be computed using a top-K sorting algorithm with $\mathcal{O}(\mathcal{I} log \mathcal{K})$ complexity, making it suitable for large-scale online recommendation scenarios.

Following the recommendation, we update the phased budget $\boldsymbol{\theta}_t$, accumulated exposure $\mathbf{e}_t$, and the utility vector $\mathbf{v}_t$. Then the subgradient of the dual function is obtained through $\widetilde{\mathbf{g}}_t = \mathbf{e}_t - \mathbf{x}_t / T$. Finally, the dual variable is updated through mirror descent. When the gradient $g_{t,i}$ is positive, the mirror descent step decreases the dual variable, so the algorithm naturally recommends more of this item. With a negative gradient, the value of the dual variable is increased, and the algorithm recommends fewer of this item. Intuitively, in classical welfare economics, the dual variable $\boldsymbol{\mu}$ can be viewed as the shadow price [10], reflecting the cost of allocating additional exposure to an item.

2.3 Theoretical Analysis

In this section, we first provide an upper bound on the linear approximation error. Then we measure the regret of our online algorithm and show that DivFair achieves a sub-linear regret bound. Finally, we compare DivFair with the representative re-ranking model P-MMF, and prove that DivFair ensures a tighter regret bound and more stable convergence.

Linear approximation. Benefiting from the properties of function f shown in Theorem 2, we have the following theorems to bound the error of linear approximation in Algorithm 1.

Lemma 1 (Taylor Expansion)**.** *The error of using first-order Taylor expansion to approximate the objective function $f(\cdot)$ is of order $\mathcal{O}(\frac{1}{T^2})$.*

Proof. Let $\Delta_f(v_t) := \widetilde{f}(v_t) - f(v_t)$ where $\widetilde{f}(v_t) = f(v_{t-1}) + \nabla f(v_{t-1})^\top (v_t - v_{t-1})$.

From the proof of Theorem 2 we know that function $f : [0,1]^{\mathcal{I}} \to \mathbb{R}$ is L-smooth with a bounded Hessian ($\|H_f\| \leq L$). f is strongly concave so $\Delta_f(v) \geq 0$. The first derivative of f is continuous, and f has an Lebesgue integrable second derivative almost everywhere, so the first derivative of f is absolutely continuous on the closed interval between 0 and 1.

Let $z_t = v_t - v_{t-1} = \frac{1}{T} w_t \odot x_t$. According to Taylor's theorem, we have the integral form of the remainder that

$$
\begin{aligned}
&\Delta_f(v_t) \\
=& \mid \widetilde{f}(v_t) - f(v_t) \mid \\
=& \mid f(v_{t-1}) + \nabla f(v_{t-1})^\top z_t - f(v_{t-1} + z_t) \mid \\
=& \mid \int_{v_{t-1}}^{v_t} f''(t)(v_t - t)\mathrm{d}t \mid \\
\leq& L \int_{v_{t-1}}^{v_t} (v_t - t)\mathrm{d}t \\
=& \frac{L}{2} \mid z_t^2 \mid \\
=& \frac{L}{2T^2} \mid w_t \odot x_t \mid^2 \\
\leq& \frac{L w_{\max}^2 k}{2T^2} \\
\in& \mathcal{O}(\frac{1}{T^2}).
\end{aligned}
$$

□

Using Lemma 1, we have Theorem 4 to bound the error of linear approximation of per-round Lagrangian $\mathcal{L}(\cdot)$.

Theorem 4 (Linear Approximation)**.** *Let $\mathcal{L}_t(x_t) := f(\mathbf{v}_t) - (\boldsymbol{\mu}_t + \mathbf{m})^\top \mathbf{x}_t / T$ be the Lagrangian function in line 6 of Algorithm 1. Then in line 7, the error introduced by the linear approximation of $\mathcal{L}_t(\cdot)$ is of order $\mathcal{O}(\frac{1}{T^2})$.*

Proof. Let

$$\mathcal{L}_t(x_t) := f(\mathbf{v}_t) - (\boldsymbol{\mu}_t + \mathbf{m})^\top \mathbf{x}_t / T,$$

$$\widetilde{\mathcal{L}}_t(x_t) := f(\mathbf{v}_{t-1}) + \nabla f(\mathbf{v}_{t-1})^\top (\mathbf{v}_t - \mathbf{v}_{t-1}) - (\boldsymbol{\mu}_t + \mathbf{m})^\top \mathbf{x}_t / T,$$

and

$$\Delta_{\mathcal{L}}(x_t) := \mathcal{L}_t(x_t) - \widetilde{\mathcal{L}}_t(x_t),$$

$$\epsilon_t := \mathcal{L}_t(x_t^*) - \mathcal{L}_t(x_t),$$

where $x_t^* = \arg\max_x \mathcal{L}_t(x), x_t = \arg\max_x \widetilde{\mathcal{L}}_t(x)$,

Let $v_t^* = v_{t-1} + \frac{1}{T} w_t \odot x_t^*$, $v_t = v_{t-1} + \frac{1}{T} w_t \odot x_t$. Using Lemma 1 it is easy to see that $\Delta_{\mathcal{L}}(x_t^*) = \Delta_f(v_t^*) \in \mathcal{O}(\frac{1}{T^2})$, and $\Delta_{\mathcal{L}}(x_t) = \Delta_f(v_t) \in \mathcal{O}(\frac{1}{T^2})$. Since $x_t^* = \arg\max_x \mathcal{L}_t(x), x_t = \arg\max_x \widetilde{\mathcal{L}}_t(x)$, we have

$$
\begin{aligned}
&\mathcal{L}_t(x_t^*) \\
&= \widetilde{\mathcal{L}}_t(x_t^*) - \Delta_{\mathcal{L}}(x_t^*) \\
&\le \widetilde{\mathcal{L}}_t(x_t) - \Delta_{\mathcal{L}}(x_t^*) \\
&= \mathcal{L}_t(x_t) + \Delta_{\mathcal{L}}(x_t) - \Delta_{\mathcal{L}}(x_t^*).
\end{aligned}
$$

Therefore, the approximation error $\epsilon_t = \mathcal{L}_t(x_t^*) - \mathcal{L}_t(x_t) \le \Delta_{\mathcal{L}}(x_t) - \Delta_{\mathcal{L}}(x_t^*) \in \mathcal{O}(\frac{1}{T^2})$, which completes the proof. □

From Theorem 4 we see that linear approximation introduces an error of $\mathcal{O}(\frac{1}{T^2})$ per round which will not affect the regret bound shown in the next section.

Regret bound. We measure the regret of an algorithm as the difference between the optimal performance $\mathcal{W}_{OPT}$ of the offline problem and the performance of the online algorithm.

$$
\text{Regret(f)} = W_{OPT} - W_{online}. \tag{9}
$$

In Algorithm 1, $\mathbf{g}_t = \mathbf{e}_t - \mathbf{x}_t/T$. Since $\mathbf{e}_t \le \boldsymbol{\tau}$, we know that there exists a constant $G \in \mathbb{R}^+$ such that $\|\boldsymbol{g}_t\| \le G$. We can see that $\|\boldsymbol{\mu}_{t+1} - \boldsymbol{\mu}_t\| \le \eta G$, so there exists a constant $D \in \mathbb{R}^+$ such that $\|\boldsymbol{\mu}_1\| \le D$. With these conditions satisfied, we provide a regret bound in Theorem 5.

Theorem 5 (Regret Bound)**.** *Assume there exists a constant $G \in \mathbb{R}^+$ and $D \in \mathbb{R}^+$ such that $\|\boldsymbol{g}_t\| \le G$, $\|\boldsymbol{\mu}_1\| \le D$. Then, the regret can be bounded as follows:*

$$
\mathit{Regret}(f) \le \frac{D^2}{2\eta} + \frac{\eta G^2 T}{2} + \sum_{t=1}^{T} \epsilon_t, \tag{10}
$$

where $\sum_{t=1}^T \epsilon_t \in \mathcal{O}(\frac{1}{T})$.

Setting the learning rate as $\eta = \mathcal{O}(\frac{1}{\sqrt{T}})$, we obtain a sublinear regret bound of order $\mathcal{O}(\sqrt{T})$.

Proof. Regret(f) $= W_{OPT} - W_{online} = \sum_{t=1}^T \big(W_{\text{opt}}^t - f(v_t)\big)$, where $v_t = v_{t-1} + \frac{1}{T} w_t \odot x_t$ is calculated using the solution $\boldsymbol{x}_t$ we get from each round in Algorithm 1. Using the same notations with Theorem 4, we have

$$
\begin{aligned}
&W_{\text{opt}}^t - f(v_t) \\
&= W_{\text{opt}}^t - \mathcal{L}_t(x_t) - (\mu_t + m)^\top x_t/T \\
&= W_{\text{opt}}^t - \mathcal{L}_t(x_t^*) + \epsilon_t - (\mu_t + m)^\top x_t/T,
\end{aligned} \tag{11}
$$

From weak duality, we know that $W_{\text{opt}}^t \le \mathcal{L}_t(x_t^*) + \max_{e_t \le \theta_t} \mu_t^T e_t$. And from Algorithm 1 we know $m^T x_t = 0$. Therefore,

$$
\begin{aligned}
&W_{\text{opt}}^t - f(v_t) \\
&\le \max_{e_t \le \theta_t} \mu_t^T (e_t - \frac{1}{T} x_t) + \epsilon_t \\
&= \mu_t^T g_t + \epsilon_t.
\end{aligned} \tag{12}
$$

From Theorem 4 we know $\sum_{t=1}^T \epsilon_t \in \mathcal{O}(\frac{1}{T})$. Next we show the bound for $\sum_{t=1}^T \mu_t^T g_t$.

For the mirror decent step in line 14 of Algorithm 1, if we take the gradient of the objective we get $g_t + \frac{1}{\eta}(\mu_{t+1} - \mu_t) = 0$, so $\mu_{t+1} = \arg\min_{\mu \in \mathcal{U}} [\mu_t - \eta g_t] = \Pi_{\mathcal{U}}\big(\mu_t - \eta\, g_t\big)$. It follows that $\|\mu_{t+1} - \mu_t\| \le \eta \|g_t\| \le \eta G$.

Let

$$
x \;=\; \mu_t - \eta\, g_t, \qquad z \;=\; \mu_{t+1} \;=\; \Pi_{\mathcal{U}}(x).
$$

Since the feasible region $\mathcal{U}$ is non-negative orthant which is convex and closed, according to first-order necessary optimality condition, we have

$$
(\mu - z)^\top (x - z) \;\le\; 0, \tag{13}
$$

for any $\mu \in \mathcal{U}$. Substituting $x = \mu_t - \eta g_t$ and $z = \mu_{t+1}$ gives

$$(\mu - \mu_{t+1})^\top \left[(\mu_t - \eta g_t) - \mu_{t+1}\right] \ \leq \ 0. \tag{14}$$

Rearrange the bracket as $(\mu_t - \mu_{t+1}) - \eta g_t$, move $-\eta g_t$ to the right, and divide by η:

$$\frac{1}{\eta}(\mu - \mu_{t+1})^\top (\mu_t - \mu_{t+1}) \ \leq \ (\mu - \mu_{t+1})^\top g_t. \tag{15}$$

So we have

$$\begin{aligned} &(\mu_t - \mu)^\top g_t \\ &= (\mu_t - \mu_{t+1})^\top g_t \ - \ (\mu - \mu_{t+1})^\top g_t \\ &\leq (\mu_t - \mu_{t+1})^\top g_t \ - \ \tfrac{1}{\eta}(\mu - \mu_{t+1})^\top (\mu_t - \mu_{t+1}). \end{aligned} \tag{16}$$

Let

$$a := \mu_t - \mu_{t+1}, \qquad b := \mu - \mu_{t+1},$$

so that (16) reads

$$(\mu_t - \mu)^\top g_t \ \leq \ a^\top g_t \ - \ \frac{1}{\eta}\, a^\top b. \tag{17}$$

We now bound each inner product. By Young's inequality,

$$a^\top g_t \ \leq \ \frac{\|a\|_2^2}{2\eta} \ + \ \frac{\eta}{2}\,\|g_t\|_2^2. \tag{18}$$

Using $\|b - a\|_2^2 = \|b\|_2^2 + \|a\|_2^2 - 2\, a^\top b$, we get

$$a^\top b = \frac{\|b\|_2^2 - \|b - a\|_2^2 + \|a\|_2^2}{2}. \tag{19}$$

Notice $b - a = \mu - \mu_t$. Hence

$$-\frac{1}{\eta}\, a^\top b = -\frac{1}{2\eta}\Big(\|\mu - \mu_t\|_2^2 - \|\mu - \mu_{t+1}\|_2^2 + \|a\|_2^2\Big). \tag{20}$$

Substituting (18) and (20) into (17), the $\frac{\|a\|_2^2}{2\eta}$ terms cancel, yielding the bound

$$(\mu_t - \mu)^\top g_t \ \leq \ \frac{\|\mu - \mu_t\|_2^2 - \|\mu - \mu_{t+1}\|_2^2}{2\eta} \ + \ \frac{\eta}{2}\,\|g_t\|_2^2, \tag{21}$$

for any $\mu \in \mathcal{U}$.

Summing up (21) from t = 1 to T yields

$$\begin{aligned} &\sum_{t=1}^{T} (\mu_t - \mu)^\top g_t \\ &\leq \sum_{t=1}^{T} \frac{\|\mu - \mu_t\|_2^2 - \|\mu - \mu_{t+1}\|_2^2}{2\eta} + \frac{\eta}{2} \sum_{t=1}^{T} \|g_t\|_2^2 \\ &= \frac{\|\mu - \mu_1\|_2^2 \ - \ \|\mu - \mu_{T+1}\|_2^2}{2\eta} + \frac{\eta}{2} \sum_{t=1}^{T} \|g_t\|_2^2. \end{aligned} \tag{22}$$

Drop the non-negative term $\|\mu - \mu_{T+1}\|_2^2$ and take $\mu = 0$, we get

$$\begin{aligned} &\sum_{t=1}^{T} \mu_t^\top g_t \\ &\leq \ \frac{\|\mu_1\|_2^2}{2\eta} \ + \ \frac{\eta}{2} \sum_{t=1}^{T} \|g_t\|_2^2 \\ &\leq \frac{D^2}{2\eta} \ + \ \frac{\eta G^2 T}{2}. \end{aligned} \tag{23}$$

Combining (23) with (12), if we set learning rate $\eta = \mathcal{O}(\frac{1}{\sqrt{T}})$, we obtain $\mathcal{O}(\sqrt{T})$ bound on the regret. $\square$

We compare DivFair with the representative re-ranking model P-MMF, which is adapted to individual-level item fairness in online recommendation.

Corollary 1. *DivFair has a strictly tighter regret bound than P-MMF, although they belong to the same order of magnitude.*

Proof. We use superscript D and P to represent DivFair algorithm and P-MMF algorithm, respectively. Since both gradients $\|g_t^D\|$ and $\|g_t^P\| \in \mathcal{O}(1)$, there exists a common constant $G \in \mathbb{R}^+$ such that $\|g_t\| \leq G$ for both algorithms.

For DivFair we obtain

$$\text{Regret}_D \leq \frac{D^2}{2}\sqrt{T} + \frac{G^2}{2}\sqrt{T} + O(\frac{1}{T}),$$

so the coefficient of $\sqrt{T}$ term is $C_D = \frac{D^2+G^2}{2}$.

For P-MMF we have

$$\text{Regret}_P \leq H\sqrt{T} + \frac{G^2}{c}\sqrt{T} + \frac{G^2\,\sigma}{2c^2}\sqrt{T} + O(1),$$

hence $C_P = H + \frac{G^2}{c}\left(1 + \frac{\sigma}{2c}\right)$, where $c = (1-\alpha)\,\sigma$. Therefore

$$C_P = H + \frac{G^2}{c}\left(1 + \tfrac{1}{2(1-\alpha)}\right) \geq H + \frac{G^2}{c}\left(1 + \tfrac{1}{2}\right) = H + \frac{3G^2}{2c}.$$

Moreover in practice $0 < c \leq 1$ (since $\alpha \lesssim 1$ and $\sigma \leq 1$), so $C_P \geq H + \frac{3G^2}{2} \geq \frac{3G^2}{2}$.

In DivFair, the gradient descent $\|\mu_{t+1} - \mu_t\| \leq \eta G$ gives $C_D \leq G^2$. Therefore, $C_D < C_P$ for all admissible (α, σ), which gives a strictly tighter regret bound. □

Corollary 1 shows that our regret is strictly smaller than P-MMF although they have the same order of magnitude.

Variance of dual variables. In this section, we compare DivFair with a representative re-ranking model P-MMF, adapted to individual-level item fairness. We provide Theorem 6 and Theorem 7 to compare the variance of dual variables in two models, and show that DivFair leads to more stable and faster convergence.

Notations and assumptions. We use superscript D and P to represent DivFair and P-MMF, respectively. Let S_t be the size-K set recommended at round t. Let $\pi^{\mathsf{D}}_{t,i} := \Pr\big(i \in S_t \mid S^{\mathsf{D}}_{t-1}\big)$, $\pi^{\mathsf{P}}_{t,i} := \Pr\big(i \in S_t \mid S^{\mathsf{P}}_{t-1}\big)$. $x_{t,i}$ is a Bernoulli variable which equals 1 with probability $\pi_{t,i}$ and 0 otherwise.

To facilitate online exploration and stable tie-breaking, we introduce an infinitesimal Gumbel-top-K perturbation to the logits, whose marginals remain Boltzmann-distributed. This is widely adopted in large-scale recommender systems for online exploration and smooth gradient estimates [7, 9]. As β increases the noise becomes negligible and the sampler converges to the deterministic top-K implementation. Lemma 2 in [6] shows that for *every* round and item, a softmax sampling with temperature β gives $\frac{K}{I}\,e^{-2\beta S_{\max}} \leq \pi_{t,i} \leq \frac{K}{I}\,e^{+2\beta S_{\max}}$.

We denote the maximum and minimum variance of $x_{t,i}$ by $v_{\max}$ and $v_{\min}$. In P-MMF algorithm, $\gamma_{\max}$ and α denote the maximum resource and momentum coeffcient in P-MMF, respectively. $Var\big[\mu^{\mathsf{D}}_{T,i}\big]$ and $Var\big[\mu^{\mathsf{P},\perp}_{T,i}\big]$ denote the variance of dual variables in DivFAir and P-MMF, respectively. Then we calculate the variance of dual variable in P-MMF algorithm in Lemma 2.

Lemma 2 (P-MMF variance)**.** *Under the P-MMF updates one has*

$$Var\big[\mu^{\mathsf{P},\perp}_{T,i}\big] = \eta^2 \kappa_T(\alpha)\,\Phi^{\mathsf{P}}_i,$$

where

$$\kappa_T(\alpha) := \frac{\alpha}{2-\alpha}T \;-\; \mathcal{E}(\alpha, T), \quad 0 < \mathcal{E}(\alpha, T) < \alpha e^{-2\alpha T},$$

$$\Phi^{\mathsf{P}}_i := \left(1 - \tfrac{1}{I}\right)^2 \frac{\bar{v}^{\mathsf{P}}_i}{\gamma_i^2} \;+\; \frac{1}{I^2}\sum_{j \neq i} \frac{\bar{v}^{\mathsf{P}}_j}{\gamma_j^2},$$

$$\bar{v}^{\mathsf{P}}_i := \frac{1}{T}\sum_{t=1}^{T} v^{\mathsf{P}}_{t,i}.$$

Proof. In P-MMF algorithm, from the update $g^{\mathsf{P}}_{t,i} = \alpha\,\tilde g_{t,i} + (1-\alpha)\,g^{\mathsf{P}}_{t-1,i}$ one obtains by induction

$$g^{\mathsf{P}}_{t,i} = \alpha \sum_{s=0}^{t-1} (1-\alpha)^s\,\tilde g_{t-s,i}, \qquad \tilde g_{t,i} = -x_{t,i} + e_{t,i}.$$

From $\mu^{\mathsf{P}}_{t+1,i} = \mu^{\mathsf{P}}_{t,i} - \eta\frac{1}{\gamma_i} g^{\mathsf{P}}_{t,i}$, we know before projection, the scaled dual variable after T rounds is

$$\mu^{\mathsf{P}}_{T,i} = -\eta\frac{1}{\gamma_i}\sum_{t=1}^{T} g^{\mathsf{P}}_{t,i} = -\eta\frac{1}{\gamma_i}\alpha\sum_{t=1}^{T}\sum_{s=0}^{t-1}(1-\alpha)^s\,\tilde g_{t-s,i}.$$

Due to independence across t,

$$Var[\mu^{\mathsf{P}}_{T,i}] = \eta^2\frac{1}{\gamma_i^2}\sum_{s=0}^{T-1}\alpha^2(1-\alpha)^{2s}\sum_{t=s+1}^{T} v^{\mathsf{P}}_{t-s,i}.$$

Re-index $t-s\mapsto t$ and let $\bar v^{\mathsf{P}}_i := \frac{1}{T}\sum_{t=1}^{T} v^{\mathsf{P}}_{t,i}$.

$$Var[\mu^{\mathsf{P}}_{T,i}] = \eta^2\frac{1}{\gamma_i^2}\sum_{s=0}^{T-1}\alpha^2(1-\alpha)^{2s}\big(T\,\bar v^{\mathsf{P}}_i\big).$$

By Geometric series,

$$\sum_{s=0}^{T-1}(1-\alpha)^{2s} \;=\; \frac{1-(1-\alpha)^{2T}}{1-(1-\alpha)^2}.$$

Thus,

$$Var[\mu^{\mathsf{P}}_{T,i}] = \eta^2\,\bar v^{\mathsf{P}}_i\,\frac{1}{\gamma_i^2}\Big(\frac{\alpha}{2-\alpha}T - T\,\mathcal{E}(\alpha,T)\Big).$$

where $\mathcal{E}(\alpha,T) = \frac{\alpha}{2-\alpha}(1-\alpha)^{2T}$. $0<\mathcal{E}(\alpha,T)<\alpha e^{-2\alpha T}$, and $\mathcal{E}(\alpha,T)\ll 10^{-3}$ once $T \gtrsim \frac{3}{2\alpha}$.

By simplex-projection lemma [11, 8], the projection calculates the common average

$$\bar h_t \;:=\; \frac{1}{I}\sum_{j=1}^{I}\frac{1}{\gamma_j}\,g^{\mathsf{P}}_{t,j}.$$

According to [1] and [13], for a single coordinate i, the post-projection update can be written as

$$\mu^{\mathsf{P},\perp}_{T,i} := -\,\eta\sum_{t=1}^{T}\Big(\frac{1}{\gamma_i}\,g^{\mathsf{P}}_{t,i} \;-\; \bar h_t\Big).$$

$$Var\big[\mu^{\mathsf{P},\perp}_{T,i}\big] = Var\big[\mu^{\mathsf{P}}_{T,i}\big] + \eta^2\,Var\Big(\sum_{t=1}^{T}\bar h_t\Big) - 2\,\eta^2\,Cov\Big(\sum_{t=1}^{T}\frac{g^{\mathsf{P}}_{t,i}}{\gamma_i}\;,\;\sum_{t=1}^{T}\bar h_t\Big).$$

To simplify notation introduce

$$\beta_t \;:=\; \sum_{j=1}^{I}\frac{1}{\gamma_j}\,g^{\mathsf{P}}_{t,j}, \qquad w_{t,i} \;:=\; \frac{1}{\gamma_i}\,g^{\mathsf{P}}_{t,i}.$$

Because the $g^{\mathsf{P}}_{t,j}$ are independent across j given $\mathcal{H}_{t-1}$, $Cov(w_{t,i},\beta_t) = Var(w_{t,i})$. Summing over t yields

$$Var[\sum_{t=1}^{T}\bar h_t] = \frac{1}{I^2}\sum_{j=1}^{I}Var[\sum_{t=1}^{T}w_{t,j}], \quad Cov\Big(\sum_{t=1}^{T}w_{t,i},\sum_{t=1}^{T}\bar h_t\Big) = \frac{1}{I}\,Var[\sum_{t=1}^{T}w_{t,i}].$$

The exact closed-form variance after projection is

$$Var\big[\mu^{\mathsf{P},\perp}_{T,i}\big] = \Big(1-\tfrac{1}{I}\Big)^2 Var\big[\mu^{\mathsf{P}}_{T,i}\big] + \frac{1}{I^2}\sum_{j\neq i}Var\big[\mu^{\mathsf{P}}_{T,j}\big] = \eta^2\kappa_T(\alpha)\,\Phi^{\mathsf{P}}_i.$$

with $\kappa_T(\alpha) = \frac{\alpha}{2-\alpha}\,T - \mathcal{E}(\alpha,T)$, and $\Phi^{\mathsf{P}}_i = \Big(1-\frac{1}{I}\Big)^2\frac{\bar v^{\mathsf{P}}_i}{\gamma_i^2} + \frac{1}{I^2}\sum_{j\neq i}\frac{\bar v^{\mathsf{P}}_j}{\gamma_j^2}$. This completes the proof. $\square$

Then we compare the variance of dual variables in DivFair and P-MMF, and show that DivFair has smaller variance than P-MMF, leading to more stable convergence.

Theorem 6 (Lower Variance)**.** *Let*

$$A := \frac{\alpha}{2-\alpha}\Big(1-\frac{1}{I}\Big)\frac{v^{\mathsf{P}}_{\min}}{\gamma^2_{\max}},$$

$$B := \Big(1-\frac{1}{I}\Big)\frac{v^{\mathsf{P}}_{\min}}{\gamma^2_{\max}\alpha},$$

$$C := v^{\mathsf{D}}_{\max}, \quad \textit{and}$$

$$T_* := \frac{B+\sqrt{B^2+4AC}}{2A}.$$

Then for every item i, when $T > T_*$, we have:

$$Var[\mu^{\mathsf{D}}_{T,i}] < Var[\mu^{\mathsf{P},\perp}_{T,i}]. \tag{24}$$

Proof. Using notations above we have

$$\bar{v}^{\mathsf{P}}_i \geq v^{\mathsf{P}}_{\min}, \qquad \frac{1}{T}\sum_{t=1}^{T} v^{\mathsf{D}}_{t,i} \leq v^{\mathsf{D}}_{\max}, \qquad \gamma_i \geq \gamma_{\min}.$$

From the online algorithm, in each round e_t is determined by exposure history $\mathcal{H}_{t-1}$, which is fixed when sampling the recommendation set x_t, $Var[g^{\mathsf{D}}_{t,i}] = Var[-\frac{1}{T}\,x_{t,i}] = \frac{1}{T^2}v^{\mathsf{D}}_{t,i}$. For each coordinate i, $\mu^{\mathsf{D}}_{T,i} = \sum_{t=1}^{T}\Pi_{R^I_+}[-\eta\, g^{\mathsf{D}}_{t,i}]$. Projection onto R^I_+ is 1-Lipschitz, hence $Var[\mu^{\mathsf{D}}_{T,i}] \leq \eta^2\sum_{t=1}^{T} Var[g^{\mathsf{D}}_{t,i}] = \eta^2\,\frac{1}{T^2}\sum_{t=1}^{T} v^{\mathsf{D}}_{t,i} \leq \eta^2\,\frac{v^D_{\max}}{T}$.

Then from the exact variance–gap formula $\Delta_{T,i} = Var[y^{\mathsf{P},\perp}_{T,i}] - Var[\mu^{\mathsf{D}}_{T,i}]$ we obtain from Lemma 2

$$\Delta_{T,i} \;\geq\; \eta^2\Big[\Big(\tfrac{\alpha}{2-\alpha}\,T - \tfrac{1}{\alpha}\Big)\Big(\big(1-\tfrac{1}{I}\big)^2\frac{v^{\mathsf{P}}_{\min}}{\gamma^2_{\max}} + \frac{I-1}{I^2}\,\frac{v^{\mathsf{P}}_{\min}}{\gamma^2_{\max}}\Big) \;-\; \frac{v^{\mathsf{D}}_{\max}}{T}\Big].$$

Noting that $\big(1-\frac{1}{I}\big)^2 + \frac{I-1}{I^2} = 1-\frac{1}{I}$, Hence $\Delta_{T,i} \geq \eta^2\big(AT - B - \frac{C}{T}\big)$, where

$$A := \frac{\alpha}{2-\alpha}\,\frac{v^{\mathsf{P}}_{\min}}{\gamma^2_{\max}}\Big(1-\tfrac{1}{I}\Big) > 0, \quad B := \Big(1-\frac{1}{I}\Big)\frac{v^{\mathsf{P}}_{\min}}{\gamma^2_{\max}\,\alpha}, \quad C := v^{\mathsf{D}}_{\max}.$$

To ensure $\Delta_{T,i} > 0$, we require

$$AT - B - \frac{C}{T} > 0 \iff AT^2 - B\,T - C > 0.$$

The quadratic inequality $AT^2 - BT - C > 0$ holds exactly for

$$T \;>\; \frac{B+\sqrt{B^2+4AC}}{2A} \;=:\; T_*.$$

Because $A, B, C > 0$, the threshold T_* is finite and yields the desired variance gap $\Delta_{T,i} > 0$ for all $T > T_*$.

□

Remark. All quantities $K, I, M_f, \sigma, \gamma_{\max}, \alpha$ are observable *before* training; thus T_* is a closed-form numerical constant. Typically, based on Steam dataset, we have $I = 8000, K = 20, \alpha = 0.60, \gamma_{\max} = 1, \sigma = 0.30, v^D_{\max} = 0.25, v^P_{\min} = 0.18$, then $T_* \approx 9.2$ rounds before dual variable in DivFair gains smaller variance than P-MMF.

Faster convergence of dual variables. In online optimization, sub-gradient $g_{t,i} > 0$ means item i is overexposed in this round and underexposed otherwise. Theorem 7 shows DivFair reduces the exposure imbalance more rapidly than P-MMF, converging faster by order of magnitude in empirical evaluations.

Lemma 3 (Freedman's inequality). *Let $\{X_t\}$ be a martingale-difference sequence satisfying $|X_t| \leq b$ almost surely. Define*

$$V_T \;:=\; \sum_{t=1}^{T} Var\big(X_t \mid \mathcal{X}_{t-1}\big).$$

Then Freedman's inequality gives, for any $\lambda > 0$,

$$\Pr\Big[\sum_{t=1}^{T} X_t \geq \lambda\Big] \;\leq\; \exp\Big(-\frac{\lambda^2}{2\big(V_T + \frac{b\lambda}{3}\big)}\Big).$$

Theorem 7 (Accumulative error band). *$\forall \delta > 0$, we have*

$$\Pr\Big[\forall i,\ \big|\tfrac{1}{T}\sum_t g^{\mathrm{D}}_{t,i}\big| \leq c_O\,\frac{\sqrt{\ln(I/\delta)}}{T^{3/2}}\Big] \;\geq 1-\delta,$$

$$\Pr\Big[\forall i,\ \big|\tfrac{1}{T}\sum_t g^{\mathrm{P}}_{t,i}\big| \leq c_P\,\frac{\sqrt{\ln(I/\delta)}}{T^{1/2}}\Big] \;\geq 1-\delta.$$

where $c_{\mathrm{D}}, c_{\mathrm{P}}$ are explicit functions of $\alpha, v^{\mathrm{D}}_{\max}, v^{\mathrm{P}}_{\max}$.

Thus DivFair's per-item exposure error shrinks orders of magnitude faster $\big(T^{-3/2}$ vs. $T^{-1/2}\big)$ in high probability.

Proof. From Lemma 2, $V^{\mathrm{D}}_T(i) := \sum_{t=1}^{T} Var\big(g^{\mathrm{D}}_{t,i} \mid \mathcal{H}_{t-1}\big) \leq \frac{v^{\mathrm{D}}_{\max}}{T}$, $\quad V^{\mathrm{P}}_T(i) := \sum_{t=1}^{T} Var\big(g^{\mathrm{P}}_{t,i} \mid \mathcal{H}_{t-1}\big) \leq \frac{\alpha}{2-\alpha}\, v^{\mathrm{P}}_{\max}\, T$.

Set the right-hand side of Lemma 3 to δ/I and solve for λ. ForDivFair,

$$\lambda_{\mathrm{D}} = O\Big(\sqrt{\tfrac{v^{\mathrm{D}}_{\max}}{T}\ \ln(I/\delta)}\Big).$$

For P-MMF,

$$\lambda_{\mathrm{P}} = O\Big(\sqrt{\tfrac{\alpha}{2-\alpha}\, v^{\mathrm{P}}_{\max}\, T \ln(I/\delta)}\Big).$$

Divide by T, $\frac{\lambda_{\mathrm{D}}}{T} \in D\big(T^{-3/2}\big)$, $\frac{\lambda_{\mathrm{P}}}{T} \in D\big(T^{-1/2}\big)$. By union bound over all items, we get $\forall$ item i,

$$\big|\tfrac{1}{T}\sum_{t=1}^{T} g^{\mathrm{D}}_{t,i}\big| \;\leq\; c_{\mathrm{D}}\,\frac{\sqrt{\ln(I/\delta)}}{T^{3/2}}, \qquad \big|\tfrac{1}{T}\sum_{t=1}^{T} g^{\mathrm{P}}_{t,i}\big| \;\leq\; c_{\mathrm{P}}\,\frac{\sqrt{\ln(I/\delta)}}{T^{1/2}},$$

with probability at least $1-\delta$. This completes the proof. □

Theorem 7, together with Theorem 6, shows that stochastic noise in the update of dual variable is much smaller in DivFair. In practice, after only a few dozen rounds, DivFair algorithm is already nearly budget-balanced for items, whereas P-MMF algorithm still shows appreciable systematic over- or under-exposure.

References

[1] Lukás Adam and Václav Mácha. Projections onto the canonical simplex with additional linear inequalities. *Optimization Methods and Software*, 37(2):451–479, 2022.

[2] Frank M. Bass. A new product growth for model consumer durables. *Management Science*, 15(5):215–227, 1969.

[3] W.J. Baumol and A.S. Blinder. *Microeconomics: Principles and Policy*. Cengage Learning, 2011.

[4] Dimitri P Bertsekas. Nonlinear programming. *Journal of the Operational Research Society*, 48(3):334–334, 1997.

[5] Dimitris Bertsimas, Vivek F. Farias, and Nikolaos Trichakis. On the efficiency-fairness trade-off. *Management Science*, 58(12):2234–2250, 2012.

[6] Nicolò Cesa-Bianchi, Claudio Gentile, Gábor Lugosi, and Gergely Neu. Boltzmann exploration done right, 2017.

[7] Minmin Chen, Alex Beutel, Paul Covington, Sagar Jain, Francois Belletti, and Ed Chi. Top-k off-policy correction for a reinforce recommender system. *arXiv preprint arXiv:1812.02353*, 2018.

[8] John C. Duchi, Shai Shalev-Shwartz, Yoram Singer, and Tushar Chandra. Efficient projections onto the l1-ball for learning in high dimensions. In *International Conference on Machine Learning*, 2008.

[9] Wouter Kool, Herke van Hoof, and Max Welling. Stochastic beams and where to find them: The gumbel-top-k trick for sampling sequences without replacement. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, pages 3499–3508, 2019.

[10] Peter J. Lambert. *The Distribution and Redistribution of Income*. Springer, 1992.

[11] M. Minoux. *Mathematical Programming: Theory and Algorithms*. A Wiley-Interscience publication. Wiley, 1986.

[12] Shameem Puthiya Parambath, Siwei Liu, Christos Anagnostopoulos, Roderick Murray-Smith, and Iadh Ounis. Parameter tuning of reranking-based diversification algorithms using total curvature analysis. In *Proceedings of the 2022 ACM SIGIR International Conference on Theory of Information Retrieval*, pages 108–114, 2022.

[13] Sinong Wang and Ness Shroff. Towards fast-convergence, low-delay and low-complexity network optimization. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 1(2):1–32, 2017.

[14] Le Wu, Qi Liu, Enhong Chen, Nicholas Jing Yuan, Guangming Guo, and Xing Xie. Relevance meets coverage: A unified framework to generate diversified recommendations. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 7(3):1–30, 2016.

[15] 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*, pages 22–32, 2005.