# Supplementary Material for LTP-MMF

Anonymous Author(s)

## ABSTRACT

Due to the limitation of the space, we only prove the skeleton of our proof of theorems. The detailed proof can be found in this materiel.

**ACM Reference Format:**

Anonymous Author(s). 2023. Supplementary Material for LTP-MMF. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

## 1 APPENDIX

### 1.1 Proof of Theorem 1

Proof. For max-min fairness, we have the regularizer as $r(\mathbf{e}) = \min_{p\in\mathcal{P}}\left(\mathbf{e}_p/\boldsymbol{\gamma}_p\right)$, we can easily proof that the exposure vector $\mathbf{e}$ can be represented as the dot-product between decision varible $\mathbf{x}_t$ and the item-provider adjacent matrix $\mathbf{A}$: $\mathbf{e} = \sum_{t=1}^{T}\left(\mathbf{A}^\top\mathbf{x}_t\right)$. Then we treat the $\mathbf{e}$ as the auxiliary variable, and the Equation of ideal objective can be written as:

$$
\begin{aligned}
W_{OPT} = &\max_{\mathbf{x}_t\in\mathcal{X},\mathbf{e}\leq\boldsymbol{\gamma}}\left[\sum_{t=1}^{T} g(\mathbf{x}_t)/T + \lambda r(\mathbf{e})\right]\\
&s.t. \mathbf{e} = \sum_{t=1}^{T}\left(\mathbf{A}^\top\mathbf{x}_t\right),
\end{aligned}
$$

where $\mathcal{X} = \{\mathbf{x}_t | \mathbf{x}_t \in 0,1 \wedge \sum_{i\in\mathcal{I}}\mathbf{x}_{ti} = K\}$. Then we move the constraints to the objective using a vector of Lagrange multipliers $\boldsymbol{\mu}\in\mathbb{R}^{|\mathcal{P}|}$:

$$
\begin{aligned}
W_{OPT} &= \max_{\mathbf{x}_t\in\mathcal{X},\mathbf{e}\leq\boldsymbol{\gamma}}\min_{\boldsymbol{\mu}\in\mathcal{D}}\left[\sum_{t=1}^{T} g(\mathbf{x}_t)/T + \lambda r(\mathbf{e}) - \boldsymbol{\mu}^\top\left(-\mathbf{e}+\sum_{t=1}^{T}\mathbf{A}^\top\mathbf{x}_t\right)\right]\\
&\leq \min_{\boldsymbol{\mu}\in\mathcal{D}}\left[\max_{\mathbf{x}_t\in\mathcal{X}}\left[\sum_{t=1}^{T} g(\mathbf{x}_t)/T - \boldsymbol{\mu}^\top\sum_{t=1}^{T}\mathbf{A}^\top\mathbf{x}_t\right] + \max_{\mathbf{e}\leq\boldsymbol{\gamma}}\left(\lambda r(\mathbf{e}) - \boldsymbol{\mu}^\top\mathbf{e}\right)\right]\\
&= \min_{\boldsymbol{\mu}\in\mathcal{D}}\left[f^*(\mathbf{A}\boldsymbol{\mu}) + \lambda r^*(-\boldsymbol{\mu})\right] = W_{Dual},
\end{aligned}
$$

where $\mathcal{D} = \boldsymbol{\mu}|r^*(-\boldsymbol{\mu}) < \infty\}$ is the feasible region of dual variable $\boldsymbol{\mu}$. According to the Lemma 1 in the Balseiro et al. [2], we have $\mathcal{D}$ is convex and positive orthant is inside the recession cone of $\mathcal{D}$.

We let the variable $\mathbf{z}_p = (\mathbf{e}_p/\boldsymbol{\gamma}_p - 1)$, we have:

$$
\begin{aligned}
r^*(\boldsymbol{\mu}) &= \max_{\mathbf{e}\leq\boldsymbol{\gamma}}\left[\min\left(\mathbf{e}_p/\boldsymbol{\gamma}_p\right) + \boldsymbol{\mu}^\top\mathbf{e}/\lambda\right]\\
&= \boldsymbol{\mu}^\top\boldsymbol{\gamma}/\lambda + 1 + \max_{\mathbf{z}_p\leq 0}\left[\min(\mathbf{z}_p) + 1/\lambda\sum_{p\in\mathcal{P}}\boldsymbol{\mu}_p\boldsymbol{\gamma}_p\mathbf{z}_p\right]
\end{aligned}
$$

Let $s(\mathbf{z}) = \min_p \mathbf{z}_p$ and $\mathbf{v} = (\boldsymbol{\mu}\odot\boldsymbol{\gamma})/\lambda$, $\odot$ is the hadamard product. Then we define $s^*(\mathbf{v}) = \max_{\mathbf{z}\leq 0}\left(s(\mathbf{z}) + \mathbf{z}^T\mathbf{v}\right)$. We firstly show that if $\sum_{p\in\mathcal{S}}\mathbf{v}_p \geq -1, \forall\mathcal{S}\in\mathcal{P}_s$, then $s^*(\mathbf{v}) = 0$ and $\mathbf{z} = 0$ is the optimal solution, otherwise $s^*(\mathbf{v}) = \infty$.

We can equivalently write $\mathcal{D} = \{\mathbf{v}|\sum_{p\in\mathcal{S}}\mathbf{v}_p \geq -1, \forall\mathcal{S}\in\mathcal{P}_s\}$. We firstly show that $s^*(\mathbf{v}) = \infty$ for $\mathbf{v}\notin\mathcal{D}$. Suppose that there exists a subset $\mathcal{S}\in\mathcal{P}_s$ such that $\sum_{p\in\mathcal{S}}\mathbf{v}_p < -1$. For any $b > 1$, we can get a feasible solution:

$$
\mathbf{v}_p = \begin{cases} -b, & p\in\mathcal{S}\\ 0, & otherwise. \end{cases}
$$

Then, because such solution is feasible and $s(\mathbf{z}) = -b$, we obtain that

$$
s^*(\mathbf{v}) \geq s(\mathbf{z}) - b(\sum_{p\in\mathcal{S}}\mathbf{v}_p) = -b(\sum_{p\in\mathcal{S}}\mathbf{v}_p + 1).
$$

Let $b\to\infty$, we have $s^*(\mathbf{v})\to\infty$.

Then we show that $s^*(\boldsymbol{\mu}) = 0$ for $\mathbf{v}\in\mathcal{D}$. Note that $\mathbf{z} = 0$ is feasible. Therefore, we have

$$
s^*(\mathbf{v}) \geq s^*(0) = 0.
$$

Then we have $\mathbf{z}\leq 0$ and without loss of generality, that the vector $\mathbf{z}$ is sorted in increasing order, i.e., $\mathbf{z}_1\leq\mathbf{z}_2,\cdots,\leq\mathbf{z}_{|\mathcal{P}|}$. The objective value is

$$
\begin{aligned}
s^*(\mathbf{v}) &= \mathbf{z}_1 + \sum_{j\in|\mathcal{P}|}\mathbf{z}_j\mathbf{v}_j\\
&= \sum_{j=1}^{|\mathcal{P}|}\left(\mathbf{z}_j - \mathbf{z}_{j+1}\right)\left(1 + \sum_{i=1}^{j}\mathbf{v}_j\right) \leq 0.
\end{aligned}
$$

Thus we can have $s^*(\boldsymbol{\mu}) = 0$ for $\mathbf{v}\in\mathcal{D}$ and we have

$$
r^*(-\boldsymbol{\mu}) = \boldsymbol{\mu}^\top\boldsymbol{\gamma}/\lambda + 1.
$$

□

### 1.2 Proof of Theorem 2

Proof. Our proof will take the following four steps:

**Bias Term Bound** Firstly, we will bound the bias term $\|\hat{v}_{u,t} - v_u^*\|_{A_{u,t}}, \|\hat{v}_{i,t} - v_i^*\|_{C_{i,t}}$:

we define the bias term upper bound as $\alpha_t, \beta_t$ respectively. We will bound the bias term as follows: by taking the gradient of the objective function respect to $v_u, v_i$, we have,

$$
A_{u,t}(\hat{v}_{u,t} - v_u^*) = \sum_{j=1}^{t}\hat{v}_{i,j}(v_i^* - \hat{v}_{i,j})^\top v_u^* + \sum_{j=1}^{t}(\hat{v}_{i,j})\epsilon_j - \lambda_u v_u^*,
$$

where $\epsilon_j$ is the Gaussian noise at time $j$. Without loss of generality, in ranking task, we can always scale the user-item score $s_{u,i}$, so that the l2-norm of $v_u, v_i$ can be bounded by a constant factor:

$$
\|v_u\|_2 \leq 1, \quad \|v_i\|_2 \leq 1
$$



*Conference'17, July 2017, Washington, DC, USA*



ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

Therefore, we can bound the function norm of the above equation as

$$\begin{aligned}\|\hat{\boldsymbol{v}}_{u,t}-v_u^*\|_{A_{u,t}} &= \|\sum_{j=1}^{t}\hat{\boldsymbol{v}}_{i,j}(v_i^*-\hat{\boldsymbol{v}}_{i,j})^\top \boldsymbol{v}_u^* + \sum_{j=1}^{t}(\hat{\boldsymbol{v}}_{i,j})\epsilon_j - \lambda_u \boldsymbol{v}_u^*\|_{A_{u,t}} \\ &\le \|\sum_{j=1}^{t}\hat{\boldsymbol{v}}_{i,j}\epsilon_j\|_{A_{u,t}} + \frac{1}{\sqrt{\lambda_u}}\sum_{j=1}^{t}\|\boldsymbol{v}_i^*-\hat{\boldsymbol{v}}_{i,j}\|_2 + \sqrt{\lambda_u}\end{aligned}$$

Since the $q$-linearly convergent to the optimizer of parameter $\boldsymbol{v}_u, \boldsymbol{v}_i$ in [4], we have for every $\epsilon_q > 0$ and $0 < q < 1$, we have

$$\|\boldsymbol{v}_i^*-\hat{\boldsymbol{v}}_{i,t+1}\|_2 \le (q+\epsilon_q)\|\boldsymbol{v}_i^*-\hat{\boldsymbol{v}}_{i,t}\|_2 \tag{1}$$

Therefore, by applying the self-normalized vector-valued martingales [1], we have for any $\sigma > 0$, with probability at least $1-\sigma$,

$$\|\hat{\boldsymbol{v}}_{u,t}-v_u^*\|_{A_{u,t}} \le \alpha_t,$$

$$\alpha_t = \sqrt{\lambda_u} + \frac{2(q+\epsilon_q)(1-(q+\epsilon_q)^t)}{1-q-\epsilon_q} + \sqrt{d\ln\frac{\lambda_u d+t}{\lambda_u d\sigma}}.$$

Similarly, we have

$$\|\hat{\boldsymbol{v}}_{i,t}-v_i^*\|_{C_{i,t}} \le \beta_t,$$

$$\beta_t = \sqrt{\lambda_i} + \frac{2(q+\epsilon_q)(1-(q+\epsilon_q)^t)}{1-q-\epsilon_q} + \sqrt{d\ln\frac{\lambda_i d+t}{\lambda_i d\sigma}}$$

. Therefore, the bias term $\alpha_t, \beta_t$ are comparable with $O((1-(q+\epsilon_r)^t)\sqrt{\ln t})$ for every $\epsilon_r > 0$.

**Variance Term Bound** Next, we will bound the variance term Then we prove the converge rate of variance term: $\|\hat{\boldsymbol{v}}_{u,t}\|^2_{A_{u,t}}, \|\hat{\boldsymbol{v}}_{i,t}\|^2_{C_{i,t}}$ as follows:

$$\begin{aligned}\|\hat{\boldsymbol{v}}_{u,t}\|^2_{A_{u,t}} &= \hat{\boldsymbol{v}}_{u,t}^\top A_{u,t}\hat{\boldsymbol{v}}_{u,t} \\ &= \hat{\boldsymbol{v}}_{u,t}^\top(\lambda_u \boldsymbol{I} + \sum_{j=1}^{t}\hat{\boldsymbol{v}}_{i,t}\hat{\boldsymbol{v}}_{i,t}^\top)\hat{\boldsymbol{v}}_{u,t} \\ &\le \|\hat{\boldsymbol{u}}_{i,t}\|_2^2 + \sum_{j=1}^{T}\|\hat{\boldsymbol{v}}_{i,j}^\top\hat{\boldsymbol{v}}_{u,t}\|_2^2 \le (t+1) \sim O(t),\end{aligned}$$

since we can always scale the l2-norm of $v_u, v_i$ by any rate in ranking tasks (we only care the relative score). Thus we can easily obtain that

$$\|\hat{\boldsymbol{v}}_{u,t}\|_{A^{-1}_{u,t}} \sim (1/\sqrt{t}).$$

**Collaborative Variance Term Bound** Next, we will bound the variance term of collaborative term $\|\hat{\boldsymbol{v}}_{u,t}-\boldsymbol{v}_u^*\|^2_{A^{-1}_{u,t}}, \|\hat{\boldsymbol{v}}_{i,t}-\boldsymbol{v}_i^*\|^2_{C^{-1}_{i,t}}$ as follows:

$$\|\hat{\boldsymbol{v}}_{u,t}-\boldsymbol{v}_u^*\|_{A^{-1}_{u,t}} \le \|\hat{\boldsymbol{v}}_{u,t}-\boldsymbol{v}_u^*\|_2 \le (q+\epsilon_q)^t,$$

where $q, \epsilon_q$ follows the Eq. (1). Similarly, we have

$$\|\hat{\boldsymbol{v}}_{i,t}-\boldsymbol{v}_i^*\|_{C^{-1}_{i,t}} \le \|\hat{\boldsymbol{v}}_{i,t}-\boldsymbol{v}_i^*\|_2 \le (q+\epsilon_q)^t.$$

Let's abbreviate the upper bound of collaborative error term $(q+\epsilon_q)^t$ as $C$

**Upper Confidence Bound** Finally, the upper confidence bound of user-item score can have easily by putting the aforementioned term together:

$$\begin{aligned}s^*_{u,i} - \hat{s}_{u,t,i} &= (\boldsymbol{v}_u^*)^\top \boldsymbol{v}_i^* - \hat{v}_{u,t}^\top\hat{v}_{i,t} \\ &= 1/2\left[(\boldsymbol{v}_u^*+\hat{v}_{u,t})^\top(\boldsymbol{v}_i^*-\hat{v}_{i,t}) + (\boldsymbol{v}_u^*-\hat{v}_{u,t})^\top(\boldsymbol{v}_i^*+\hat{v}_{i,t})\right] \\ &\le \alpha_t(C/2+\|\hat{v}_{i,t}\|_{A^{-1}_{u,t}}) + \beta_t(C/2+\|\hat{v}_{u,t}\|_{C^{-1}_{i,t}}).\end{aligned}$$

We can easily have the upper bound of user-item score have the converge term

$$O(\frac{(1-(q+\epsilon_q)^t)\sqrt{\ln t}}{\sqrt{t}}),$$

which is a decrease function of $t$ when the $t$ becomes large.

Q.E.D □

### 1.3 Proof of Theorem 3

Proof. Firstly, in practice, we normalize the user-item preference score $s_{u,i}$ to $[0, 1]$. Therefore, $\sum_{t=1}^{T} g(\mathbf{x}_t)/T \le K$. In max-min regularizer $r(\mathbf{e})$. Let's abbreviate its upper bound to $\bar{r}$. In practice, $\bar{r} \le 1$ We have

$$W_{OPT} \le K + \lambda\bar{r}. \tag{2}$$

We consider the stopping time $\tau$ of Algorithm **??** as the first time the provider will have the maximum exposures, i.e.

$$\sum_{t=1}^{\tau}\mathbf{M}^\top\mathbf{x}_t \ge \boldsymbol{\gamma}.$$

Note that is $\tau$ a random variable.

Similarly, followed the prove idea of Balseiro et al. [2], First, we analysis the primal performance of the objective function. Second, we bound the complementary slackness term by the momentum gradient descent. Finally, We conclude by putting it to achieve the final regret bound.

**Primal performance proof**: Consider a time $t < \tau$, the recommender action will not violate the resource constraint. Therefore, we have:

$$g(\mathbf{x}_t)/T = g^*(\boldsymbol{M}\boldsymbol{\mu}_t) + \lambda\boldsymbol{\mu_t}^T\boldsymbol{M}^\top\mathbf{x}_t,$$

and we have $\mathbf{e}_t = \arg\max_{\mathbf{e}\le\boldsymbol{\gamma}}\{r(\mathbf{e}) + \boldsymbol{\mu}^\top\mathbf{e}/\lambda\}$

$$r(\mathbf{e}_t) = r^*(-\boldsymbol{\mu}) - \boldsymbol{\mu_t}^T\mathbf{e}_t/\lambda.$$

We make the expectations for the current time step $t$ for the primal functions:

$$\begin{aligned}\mathbb{E}\left[g(\mathbf{x}_t)/T + \lambda r(\mathbf{e}_t)\right] &= \mathbb{E}\left[g^*(\boldsymbol{M}\boldsymbol{\mu}_t) + \boldsymbol{\mu_t}^T\boldsymbol{M}^\top\mathbf{x}_t + \lambda r^*(-\boldsymbol{\mu}) - \boldsymbol{\mu_t}^T\mathbf{e}_t\right] \\ &= W_{Dual}(\boldsymbol{\mu}_t) - \mathbf{E}\left[\boldsymbol{\mu}_t^T(-\boldsymbol{M}^\top\mathbf{x}_t+\mathbf{e}_t)\right].\end{aligned}$$

Consider the process $Z_t = \sum_{j=1}^{T}\boldsymbol{\mu}_j^t(-\boldsymbol{M}^\top\mathbf{x}_t+\mathbf{e}_t) - \mathbf{E}\left[\boldsymbol{\mu}_t^T(-\boldsymbol{M}^\top\mathbf{x}_t+\mathbf{e}_t)\right]$ is a martingale process. The Optional Stopping Theorem in martingale process [5] implies that $\mathbb{E}[Z_\tau] = 0$. Consider the variable $w_t(\boldsymbol{\mu}_t) = \boldsymbol{\mu}_t^T(-\mathbf{A}^\top\mathbf{x}_t+\mathbf{e}_t)$, we have

$$\mathbb{E}\left[\sum_{t=1}^{\tau}w_t(\boldsymbol{\mu}_t)\right] = \mathbb{E}\left[\sum_{t=1}^{\tau}\mathbb{E}\left[w_t(\boldsymbol{\mu}_t)\right]\right]$$

Moreover, in MMF, the dual function $W_{Dual}$ is convex proofed in Theorem ??, we have

$$\begin{aligned}\mathbb{E}\left[\sum_{t=1}^{\tau} g(\mathbf{x}_t)/T + \lambda r(\mathbf{e}_t)\right] &= \mathbb{E}\left[\sum_{t=1}^{\tau} W_{Dual}(\boldsymbol{\mu}_t)\right] - \mathbb{E}\left[\sum_{t=1}^{\tau} w_t(\boldsymbol{\mu}_t)\right] \\ &\le \mathbb{E}\left[\tau W_{Dual}(\widetilde{\boldsymbol{\mu}_\tau})\right] - \mathbb{E}\left[\sum_{t=1}^{\tau} w_t(\boldsymbol{\mu}_t)\right],\end{aligned} \tag{3}$$

where $\widetilde{\boldsymbol{\mu}_\tau} = \sum_{t=1}^{\tau} \boldsymbol{\mu}_t/\tau$.

Next, we will bound the the bias of the primal performance due to the estimation error of the ranking model.

From the proof of theorem 1, we can bound the $W_{Dual}(\boldsymbol{\mu}_t)$ as follows: at iteration $n$, for any $\sigma > 0$, with probability at least $1-\sigma$,

$$\begin{aligned}W_{Dual}(\boldsymbol{\mu}_t) - \hat{W}_{Dual}(\boldsymbol{\mu}_t) &= g^*(\boldsymbol{M}\boldsymbol{\mu}_t) + \lambda r^*(-\mu) - \hat{g}^*(\boldsymbol{M}\boldsymbol{\mu}_t) - \lambda r^*(-\mu) \\ &= s_u^\top \boldsymbol{x}_t - (\boldsymbol{M}\boldsymbol{\mu}_t)^\top \boldsymbol{x}_t - \mathbf{s}_u^\top \boldsymbol{x}_t - (\boldsymbol{M}\boldsymbol{\mu}_t)^\top \boldsymbol{x}_t \\ &\le \sum_{i,\boldsymbol{x}_{ti}=1} \triangle f_{u_t,i}^n,\end{aligned}$$

**Complementary slackness proof** Then we aim to proof the complementary slackness $\sum_{t=1}^{T} w_t(\boldsymbol{\mu}_t) - w_t(\boldsymbol{\mu})$ is bounded. Suppose there exists $G, s.t.$ the gradient norm is bounded $\|\widetilde{\mathrm{g}}_t\| \le G$. Then we have:

$$\sum_{t=1}^{\tau} w_t(\boldsymbol{\mu}_t) - w_t(\boldsymbol{\mu}) \le \frac{L^2}{\eta} + \frac{G^2}{(1-\alpha)\sigma}\eta(\tau-1) + \frac{G^2}{2(1-\alpha)^2\sigma\eta}, \tag{4}$$

where the project function $\|\boldsymbol{\mu} - \boldsymbol{\mu}_t\|_{\boldsymbol{\gamma}}^2$ is $\sigma-$strongly convex.

Next we prove the inequality in Equation . According to the Theorem 1 in [3], we have

$$\|\mathrm{g}_t\|_2^2 = \|(1-\alpha)\sum_{s=1}^{t} \alpha^{t-s}(\widetilde{\mathrm{g}}_s)\|_2^2 \le G^2,$$

and

$$\sum_{t=1}^{\tau} w_t(\boldsymbol{\mu}_t) - w_t(\boldsymbol{\mu}) \le \frac{\|\boldsymbol{\mu}_t - \boldsymbol{\mu}_0\|_{\boldsymbol{\gamma}^2}^2}{\eta} + \frac{G^2}{(1-\alpha)\sigma}\eta(\tau-1) + \frac{G^2}{2(1-\alpha)^2\sigma\eta}, \forall \boldsymbol{\mu}.$$

We have $\|\boldsymbol{\mu}_t - \boldsymbol{\mu}_0\|_{\boldsymbol{\gamma}^2}^2 \le L^2$ according to the Cauchy-Schwarz' inequality. The results follows. Let $M = \frac{L^2}{\eta} + \frac{G^2}{(1-\alpha)\sigma}\eta(T-1) + \frac{G^2}{2(1-\alpha)^2\sigma\eta}$. We now choose a proper $\boldsymbol{\mu}$, s.t. the complementary stackness can be further bounded.

For $\boldsymbol{\mu} = \hat{\boldsymbol{\mu}} + \theta$, where $\theta \in \mathbb{R}^{|P|}$ is non-negative to be determined later and $\hat{\boldsymbol{\mu}} = \arg\max_{\boldsymbol{\mu}} -\boldsymbol{\mu}^\top(\sum_{i=1}^{T} \mathbf{A}^\top \mathbf{x}_t)/\lambda$. According to the constraint $\mathbf{e} = \sum_{i=1}^{T} \mathbf{A}^\top \mathbf{x}_t$, we have that

$$\sum_{t=1}^{T}(r(\mathbf{e}_t) + \boldsymbol{\mu}^\top \mathbf{e}_t\lambda) \le r^*(-\hat{\boldsymbol{\mu}}) = r(\sum_{i=1}^{T} \mathbf{A}^\top \mathbf{x}_t) + \hat{\boldsymbol{\mu}}^T(\sum_{i=1}^{T} \mathbf{A}^\top \mathbf{x}_t)/\lambda.$$

Note that in proof of Theorem 1, the feasible region $\mathcal{D}$ is recession cone, therefore, $\boldsymbol{\mu} \in \mathcal{D}$.

Therefore, we have

$$\sum_{t=1}^{\tau} w_t(\boldsymbol{\mu}_t) = \sum_{t=1}^{T} w_t(\hat{\boldsymbol{\mu}}) - \sum_{t=\tau+1}^{T} w_t(\hat{\boldsymbol{\mu}}) + \sum_{t=1}^{\tau} w_t(\theta) + M. \tag{5}$$

For each iteration $n$, we have

$$\begin{aligned}w_t\boldsymbol{\mu}_t - \hat{w}_t\boldsymbol{\mu}_t &= \boldsymbol{\mu}_t((\boldsymbol{M} - \boldsymbol{M})^\top \boldsymbol{x}_t) \\ &\le \sum_{i,\boldsymbol{x}_{ti}=1} \triangle f_{u_t,i}^n L\end{aligned}$$

**Put them together:** For each batch $T$ at iteration $n$, we obtain that

$$\begin{aligned}W_{OPT} &= \frac{\tau}{T} W_{OPT} + \frac{T-\tau}{T} W_{OPT} \\ &\le \tau W_{Dual}(\widetilde{\boldsymbol{\mu}_\tau}) + (T-\tau)(K + \lambda\bar{r}) \\ &\le \tau \hat{W}_{Dual}(\widetilde{\boldsymbol{\mu}_\tau}) + \sum_{t=1}^{\tau}\sum_{i,\boldsymbol{x}_{ti}} \triangle f_{u_t,i}^n + (T-\tau)(K + \lambda\bar{r})\end{aligned} \tag{6}$$

Let's abbreviate $\triangle r(n) = \sum_{t=1}^{\tau}\sum_{i,\boldsymbol{x}_{ti}} \triangle f_{u_t,i}^n L$. Therefore, combining Eq. (10,12,13) the regret $\mathrm{Regret}(n)$ of iteration $n$, can be bounded as:

$$\begin{aligned}\mathrm{Regret}(n) &= \mathbb{E}\left[W_{OPT} - \hat{W}\right] \\ &\le \mathbb{E}\left[W_{OPT} - \sum_{t=1}^{\tau}\left(\hat{g}(\mathbf{x}_t)/T - \lambda r(\boldsymbol{M}^\top \mathbf{x}_t/\boldsymbol{\gamma})\right)\right] \\ &\le \mathbb{E}\left[W_{OPT} - \tau\hat{W}_{Dual}(\widetilde{\boldsymbol{\mu}}_t) + \sum_{t=1}^{\tau}\hat{w}_t(\boldsymbol{\mu}_t) + \sum_{t=1}^{T}(\mathbf{e}_t - \boldsymbol{M}^\top \mathbf{x}_t)\right] \\ &\le \mathbb{E}\left[(T-\tau)(K+\lambda\bar{r}) + \sum_{t=1}^{T} w_t(\hat{\boldsymbol{\mu}}) + \sum_{t=1}^{\tau} w_t(\theta)\right] + M + \triangle r(n) \\ &\le (T-\tau)(K + \lambda\bar{r} + \lambda K) + \sum_{t=1}^{\tau} w_t(\theta) + M + \triangle r(n) \\ &= T\mathrm{Regret}(\pi^F)/N + \triangle r(n)\end{aligned} \tag{7}$$

Let $C = K + \lambda\bar{r} + \lambda K$, then setting the $\theta = C\min_p \boldsymbol{\gamma}_p \mathbf{u}_p$, where $\mathbf{u}_p$ is the p-th unit vector. We have

$$\sum_{t=1}^{\tau} w_t(\theta) = C/(\min_p \boldsymbol{\gamma}_p) - C(T-\tau).$$

Then the $\mathrm{Regret}(n) \le M + C/(\min_p \boldsymbol{\gamma}_p)$, when we set $\eta = O(T^{-1/2})$, the $\mathrm{Regret}(\pi^F)$ is comparable with $O(T^{-1/2})$.

According to our algorithm in Algorithm of LTP-MMF, the user will interact with the system in $N/T$ iterations. Following the Lemma 11 in Abbasi-Yadkori et al. [1], the error $\mathrm{Regret}(\pi)$ raised of accuracy module is bounded as

$$\mathrm{Regret}(\pi) = \sum_{n=1}^{N/T} \triangle r(n) \le KT[\alpha_{N/T}(\rho_u + \kappa) + \beta_{N/T}(\rho_i + \kappa)],$$

where

$$\kappa = \frac{(q+\epsilon_q)(1-(q+\epsilon_q)^{N/T})}{1-q-\epsilon_q},$$

and

$$\rho_u = \sqrt{2d\frac{N}{T}\ln(1+\frac{N}{T\lambda_u d})}, \rho_i = \sqrt{2d\frac{N}{T}\ln(1+\frac{N}{T\lambda_i d})}.$$

From the conclusion, we can see that the regret of accuracy module is comparable with $O(\sqrt{NT\ln\frac{N}{T}})$.

Finally, the total regret can be bounded as

$$\text{Regret}(\pi, \pi^F) = \sum_{n=1}^{N/T} \text{Regret}(n) = \text{Regret}(\pi^F) + \text{Regret}(\pi)L \quad (8)$$

Setting the learning rate as $\eta = O(T^{-1/2})$, we can obtain a fairness regret Regret$(\pi^F)$ upper bound of order $O(\frac{N}{\sqrt{T}})$. Overall, the long term regret of LTP-MMF can be obtained of order $O(N \ln N)$.

Q.E.D

□

# REFERENCES


[1] Yasin Abbasi-Yadkori, Dávid Pál, and Csaba Szepesvári. 2011. Improved algorithms for linear stochastic bandits. *Advances in neural information processing systems* 24 (2011).

[2] Santiago Balseiro, Haihao Lu, and Vahab Mirrokni. 2021. Regularized online allocation problems: Fairness and beyond. In *International Conference on Machine Learning*. PMLR, 630–639.

[3] Santiago R. Balseiro, Haihao Lu, Vahab Mirrokni, and Balasubramanian Sivan. 2022. From Online Optimization to PID Controllers: Mirror Descent with Momentum. https://doi.org/10.48550/ARXIV.2202.06152

[4] Huazheng Wang, Qingyun Wu, and Hongning Wang. 2016. Learning hidden features for contextual bandits. In *Proceedings of the 25th ACM international on conference on information and knowledge management*. 1633–1642.

[5] David Williams. 1991. *Probability with martingales*. Cambridge university press.