Supplementary Material - Bandit Online Learning on Graphs via Adaptive Optimization

Peng Yang[†], Peilin Zhao[‡], Xin Gao[†]
[†]KAUST, Saudi Arabia;[‡]SCUT, China
yangpeng1985521@gmail.com; xin.gao@kaust.edu.sa;
peilinzhao@hotmail.com

April 27, 2018

1 Proof of Theorems

1.1 Proof of Theorem 1 - Adaptive Online Learning on Graphs

Theorem 1. *Assume that $1 + a_T \mathbf{x}_T^\top \mathbf{A}_{T-1}^{-1} \mathbf{x}_T - a_T \leq 0$, then the optimal prediction of the min-max function $\hat{y}_T = \arg\min_{\hat{y}_T} \max_{y_T} F(\mathbf{y}_T, \mathbf{f}_T)$ is*

$$\hat{y}_T = \arg\max_{i \in [K]} (\mathbf{B}_{T-1}^\top \mathbf{A}_{T-1}^{-1} \mathbf{x}_T)_i. \tag{1}$$

Proof. Since we have

$$\begin{aligned}
&\mathrm{tr}(\mathbf{B}_T^\top \mathbf{A}_T^{-1} \mathbf{B}_T) \\
=&\mathrm{tr}(\mathbf{B}_{T-1}^\top \mathbf{A}_T^{-1} \mathbf{B}_{T-1}) + 2a_T \mathbf{B}_{T-1}^\top \mathbf{A}_T^{-1} \mathbf{x}_T \cdot \mathbf{y}_T + a_T^2 \mathbf{x}_T^\top \mathbf{A}_T \mathbf{x}_T \|\mathbf{y}_T\|^2,
\end{aligned}$$

we obtain the following form by omitting unrelated terms,

$$\begin{aligned}
&\min_{\hat{y}_T} \max_{y_T} F(\mathbf{y}_T, \mathbf{f}_T) \\
=&\min_{\hat{y}_T} \max_{y_T} \sum_{t=1}^{T} (\|\mathbf{f}_t - \mathbf{y}_t\|^2 - a_t \|\mathbf{y}_t\|^2) + \mathrm{tr}(\mathbf{B}_T^\top \mathbf{A}_T^{-1} \mathbf{B}_T) \\
=&\min_{\hat{y}_T} \max_{y_T} \alpha(a_T) \|\mathbf{y}_T\|^2 + 2\beta(a_T, \mathbf{f}_T) \cdot \mathbf{y}_T + \|\mathbf{f}_T\|^2,
\end{aligned}$$

where

$$\alpha(a_T) = \frac{1 + a_T \mathbf{x}_T^\top \mathbf{A}_{T-1}^{-1} \mathbf{x}_T - a_T}{1 + a_T \mathbf{x}_T^\top \mathbf{A}_{T-1}^{-1} \mathbf{x}_T}, \ \beta(a_T, \mathbf{f}_T) = a_T \mathbf{B}_{T-1}^\top \mathbf{A}_T^{-1} \mathbf{x}_T - \mathbf{f}_T.$$

We consider two cases: (1) $1+a_T\mathbf{x}_T^\top\mathbf{A}_{T-1}^\top\mathbf{x}_T - a_T < 0$ and (2) $1+a_T\mathbf{x}_T^\top\mathbf{A}_{T-1}^\top\mathbf{x}_T - a_T = 0$. Starting with the first case, the function $F(\mathbf{y}_T, \mathbf{f}_T)$ is strictly-concave with respect to $\mathbf{y}_T$ since $\alpha(a_T) < 0$. Thus, $F(\mathbf{y}_T, \mathbf{f}_T)$ attains a unique maximal value when $\mathbf{y}_T = -\frac{\beta(a_T,\mathbf{f}_T)}{\alpha(a_T)}$, i.e.,

$$
\begin{aligned}
&\min_{\hat{y}_T} F(-\frac{\beta(a_T,\mathbf{f}_T)}{\alpha(a_T)}, \mathbf{f}_T) \\
=&\frac{-a_T}{1+a_T\mathbf{x}_T^\top\mathbf{A}_{T-1}^{-1}\mathbf{x}_T - a_T}\|\mathbf{f}_T\|^2 + \frac{2a_T\mathbf{B}_{T-1}^\top\mathbf{A}_T^{-1}\mathbf{x}_T(1+a_T\mathbf{x}_T^\top\mathbf{A}_{T-1}^{-1}\mathbf{x}_T)}{1+a_T\mathbf{x}_T^\top\mathbf{A}_{T-1}^{-1}\mathbf{x}_T - a_T}\cdot\mathbf{f}_T - \gamma(a_T).
\end{aligned}
$$

This objective is strictly-convex with respect to $\mathbf{f}_T$, and it is easy to obtain the optimal predictor,

$$
\begin{aligned}
\mathbf{f}_T =&\mathbf{B}_{T-1}^\top\mathbf{A}_T^{-1}\mathbf{x}_T(1+a_T\mathbf{x}_T^\top\mathbf{A}_{T-1}^{-1}\mathbf{x}_T) \\
=&\mathbf{B}_{T-1}^\top\mathbf{A}_T^{-1}\mathbf{x}_T + \mathbf{B}_{T-1}^\top\mathbf{A}_T^{-1}(\mathbf{A}_T - \mathbf{A}_{T-1})\mathbf{A}_{T-1}^{-1}\mathbf{x}_T \\
=&\mathbf{B}_{T-1}^\top\mathbf{A}_{T-1}^{-1}\mathbf{x}_T.
\end{aligned}
$$

Considering the second case for which, $1+a_T\mathbf{x}_T^\top\mathbf{A}_{T-1}^{-1}\mathbf{x}_T - a_T = 0$, the problem can be formulated as $\min_{\hat{y}_T}\max_{y_T} 2(a_T\mathbf{B}_{T-1}^\top\mathbf{A}_T^{-1}\mathbf{x}_T - \mathbf{f}_T)\cdot\mathbf{y}_T + \|\mathbf{f}_T\|^2$. If $\mathbf{f}_T \neq a_T\mathbf{B}_{T-1}^\top\mathbf{A}_T^{-1}\mathbf{x}_T$, the objective value is not-bounded. Thus, the optimal min-max prediction is $\mathbf{f}_T = a_T\mathbf{B}_{T-1}^\top\mathbf{A}_T^{-1}\mathbf{x}_T = \mathbf{B}_{T-1}^\top\mathbf{A}_{T-1}^{-1}\mathbf{x}_T$, given $a_T = \frac{1}{1-\mathbf{x}_T^\top\mathbf{A}_{T-1}^{-1}\mathbf{x}_T}$, and $\mathbf{A}_t^{-1} = \mathbf{A}_{t-1}^{-1} - \mathbf{A}_{t-1}^{-1}\mathbf{x}_t\mathbf{x}_t^\top\mathbf{A}_{t-1}^{-1}$. $\square$

1.2 Proof of Theorem 2 - Error-driven Algorithm

Theorem 2. *Assume* $\{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_T, y_T)\}$ *is a node sequence, an online algorithm predicts with* $\hat{y}_T = \arg\max_{i\in[K]}(\mathbf{B}_{T-1}^\top\mathbf{A}_{T-1}^{-1}\mathbf{x}_T)_i$*, and updates* $\mathbf{A}_T$ *and* $\mathbf{B}_T$ *with Eq. (6). Let* $\mathcal{N} = \{t|\hat{\Delta}_t = \mathbf{f}_t\cdot\mathbf{y}_t \leq 0\}$ *be updated trials, then the following inequality holds for any* $\mathbf{U}\in\mathbb{R}^{d\times K}$*,*

$$
M \leq \sum_{t\in\mathcal{N}} a_t\tilde{\mathcal{L}}(\mathbf{y}_t\cdot\mathbf{U}^\top\mathbf{x}_t) + \frac{1}{2}tr(\mathbf{U}^\top\mathbf{A}_\mathcal{N}\mathbf{U}) + \frac{b}{2(b-1)}\log|\frac{1}{b}\mathbf{A}_\mathcal{N}|.
$$

Proof: Since a update is issued when an error occurs, the update trials are defined as $\mathcal{N} = \{t : \hat{\Delta}_t \leq 0, \hat{y}_t \neq y_t\}$ with $Z = |\mathcal{N}|$ includes the indices on which an error occurs. Given $\ell_t(alg) = \|\mathbf{y}_t - \mathbf{f}_t\|^2$, we derive when $t\in\mathcal{N}$,

$$
\begin{aligned}
&\ell_t(alg) + \inf_{\mathbf{U}}(b\|\mathbf{U}\|^2 + L_{t-1}^{\mathbf{a}}(\mathbf{U})) - \inf_{\mathbf{U}}(b\|\mathbf{U}\|^2 + L_t^{\mathbf{a}}(\mathbf{U})) \\
=&\|\mathbf{f}_t - \mathbf{y}_t\|^2 - a_t\|\mathbf{y}_t\|^2 - \mathrm{tr}(\mathbf{B}_{t-1}^\top\mathbf{A}_{t-1}^{-1}\mathbf{B}_{t-1}) + \mathrm{tr}(\mathbf{B}_t^\top\mathbf{A}_t^{-1}\mathbf{B}_t) \\
=&(1-a_t)\|\mathbf{y}_t\|^2 - 2\mathbf{y}_t\cdot\mathbf{f}_t + \|\mathbf{f}_t\|^2 - \mathrm{tr}(\mathbf{B}_{t-1}^\top\mathbf{A}_{t-1}^{-1}\mathbf{B}_{t-1}) + \mathrm{tr}(\mathbf{B}_t^\top\mathbf{A}_t^{-1}\mathbf{B}_t) \\
=&(1-a_t)\|\mathbf{y}_t\|^2 - 2\mathbf{y}_t\cdot(a_t\mathbf{B}_{t-1}^\top\mathbf{A}_t^{-1}\mathbf{x}_t) + \mathrm{tr}((\mathbf{B}_{t-1} + a_t\mathbf{x}_t\mathbf{y}_t^\top)^\top\mathbf{A}_t^{-1}(\mathbf{B}_{t-1} + a_t\mathbf{x}_t\mathbf{y}_t^\top)) \\
&+ \mathrm{tr}(\mathbf{B}_{t-1}^\top\mathbf{A}_{t-1}^{-1}\mathbf{x}_t\mathbf{x}_t^\top\mathbf{A}_{t-1}^{-1}\mathbf{B}_{t-1}) - \mathrm{tr}(\mathbf{B}_{t-1}^\top\mathbf{A}_{t-1}^{-1}\mathbf{B}_{t-1}) \\
=&\mathrm{tr}(\mathbf{B}_{t-1}^\top(\mathbf{A}_{t-1}^{-1}\mathbf{x}_t\mathbf{x}_t^\top\mathbf{A}_{t-1}^{-1} - \mathbf{A}_{t-1}^{-1} + \mathbf{A}_t^{-1})\mathbf{B}_{t-1}) + \mathrm{tr}(a_t^2\mathbf{y}_t\mathbf{x}_t^\top\mathbf{A}_t^{-1}\mathbf{x}_t\mathbf{y}_t^\top) + (1-a_t)\|\mathbf{y}_t\|^2 \\
=&(a_t^2\mathbf{x}_t^\top\mathbf{A}_t^{-1}\mathbf{x}_t + 1 - a_t)\mathrm{tr}(\mathbf{y}_t\mathbf{y}_t^\top) = a_t^2\mathbf{x}_t^\top\mathbf{A}_t^{-1}\mathbf{x}_t - a_t + 1.
\end{aligned}
$$

When no error occurs, $\mathbf{U}_t = \mathbf{U}_{t-1}$ yields $\inf_{\mathbf{U}} G_t(\mathbf{U}) = \inf_{\mathbf{U}} G_{t-1}(\mathbf{U})$. When an error occurs, there is a parameter update:

$$\inf_{\mathbf{U}} G_t(\mathbf{U}) - \inf_{\mathbf{U}} G_{t-1}(\mathbf{U}) = \ell_t(alg) - a_t^2 \mathbf{x}_t^\top \mathbf{A}_t^{-1} \mathbf{x}_t + a_t - 1,$$

holds for all trial $t \in \mathcal{N}$, which is similar to the proof of [2]. Summing over $t = 1, \dots, T$ with $\|\mathbf{y}_t\|^2 = 1$, we obtain with expanding the square,

$$\begin{aligned}
&\sum_{t \in \mathcal{N}} (a_t \|\mathbf{y}_t\|^2 - 2\mathbf{y}_t \cdot \mathbf{f}_t - a_t^2 \mathbf{x}_t^\top \mathbf{A}_t^{-1} \mathbf{x}_t + \|\mathbf{f}_t\|^2) \\
&= \inf_{\mathbf{U}} (b\|\mathbf{U}\|^2 + \sum_t a_t \|\mathbf{y}_t - \mathbf{U}^\top \mathbf{x}_t\|^2) - (\inf_{\mathbf{U}} (b\|\mathbf{U}\|^2 + L_0^{\mathbf{a}}(\mathbf{U}))) \\
&\leq \sum_{t \in \mathcal{N}} a_t (\|\mathbf{y}_t\|^2 - 2\mathbf{y}_t \cdot \mathbf{U}^\top \mathbf{x}_t) + \mathrm{tr}(\mathbf{U}^\top (b\mathbf{I} + \sum_{t \in \mathcal{N}} a_t \mathbf{x}_t \mathbf{x}_t^\top) \mathbf{U}).
\end{aligned}$$

Assume that $\mathbf{A}_{\mathcal{N}} = b\mathbf{I} + \sum_{t \in \mathcal{N}} a_t \mathbf{x}_t \mathbf{x}_t^\top$, and $\sigma_t = \frac{1}{2} a_t^2 \mathbf{x}_t^\top \mathbf{A}_t^{-1} \mathbf{x}_t$ we obtain,

$$\sum_{t \in \mathcal{N}} (-\mathbf{f}_t \mathbf{y}_t - \sigma_t) \leq -\sum_{t \in \mathcal{N}} a_t \mathbf{y}_t \cdot \mathbf{U}^\top \mathbf{x}_t + \frac{1}{2} \mathrm{tr}(\mathbf{U}^\top \mathbf{A}_{\mathcal{N}} \mathbf{U}),$$

where we omit $\|\mathbf{f}_t\|^2$ since it does not affect the upper bound. We add $\sum_t a_t$ on the both sides with $a_t = \frac{1}{1 - \mathbf{x}_t^\top \mathbf{A}_{t-1}^{-1} \mathbf{x}_t} \geq 1$,

$$\begin{aligned}
\sum_{t \in \mathcal{N}} (1 - \mathbf{f}_t \mathbf{y}_t - \sigma_t) &\leq \sum_{t \in \mathcal{N}} (a_t - \mathbf{f}_t \mathbf{y}_t - \sigma_t) \\
&\leq \sum_{t \in \mathcal{N}} a_t (1 - \mathbf{y}_t \cdot \mathbf{U}^\top \mathbf{x}_t) + \frac{1}{2} \mathrm{tr}(\mathbf{U}^\top \mathbf{A}_{\mathcal{N}} \mathbf{U}) \\
&\leq \sum_{t \in \mathcal{N}} a_t \tilde{\mathcal{L}}(\mathbf{y}_t \cdot \mathbf{U}^\top \mathbf{x}_t) + \frac{1}{2} \mathrm{tr}(\mathbf{U}^\top \mathbf{A}_{\mathcal{N}} \mathbf{U}),
\end{aligned} \tag{2}$$

where the last inequality holds due to hinge loss $\tilde{\mathcal{L}}(x) = \max(0, 1 - x) \geq 1 - x$. Here, update trials are the ones when an error occurs, i.e., $t \in \mathcal{N}$ with $M = |\mathcal{N}|$ and $-\mathbf{f}_t \mathbf{y}_t \geq 0$,

$$\sum_{t \in \mathcal{N}} (1 - \mathbf{f}_t \mathbf{y}_t - \sigma_t) \geq M - \sum_{t \in \mathcal{N}} \sigma_t;$$

Combining this bound with the upper bound (2), and substituting the inequality $\sum_{t \in \mathcal{N}} \sigma_t \leq \frac{b}{2(b-1)} \log(\frac{1}{b} \mathbf{A}_{\mathcal{N}})$ inspired by [3], we finish the proof. □

1.3 Proof of Theorem 3 - MOLG-F

Theorem 3. *Algorithm 1 runs on an arbitrary node sequence* $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_T, y_T)\}$ *and update model when* $\Theta_t = \hat{\Delta}_t - \sigma_t \leq 0$. *Let* $\tilde{\mathcal{L}}(x) = \max(0, 1 - x)$ *be hinge loss, for any* $\mathbf{U} \in \mathbb{R}^{d \times K}$, *the following inequality holds,*

$$M \leq \sum_{t \in \mathcal{N}} a_t \tilde{\mathcal{L}}(\mathbf{y}_t \cdot \mathbf{U}^\top \mathbf{x}_t) + \frac{1}{2} tr(\mathbf{U}^\top \mathbf{A}_{\mathcal{N}} \mathbf{U}) + \frac{b}{b-1} \log |\frac{1}{b} \mathbf{A}_{\mathcal{N}}| - D.$$

Proof: In Algorithm 1, the update trials are partitioned into two disjoint sets, $\mathcal{M} = \{t : \hat{\Delta}_t \leq 0, \hat{y}_t \neq y_t\}$ with $M = |\mathcal{M}|$ includes the indices on which an update is issued when an error occurs, and $\mathcal{D} = \{t : 0 < \hat{\Delta}_t < \sigma_t, \hat{y}_t = y_t\}$ with $D = |\mathcal{D}|$ includes the indices on which an aggressive update is issued for low-confident prediction, even if the prediction is correct. Let $\mathcal{N} = \{t : N_t = 1\}$ with $N = |\mathcal{N}|$ be the update trials containing $N = M + D$. Similar with Eq. (2) in lemma 1, we derive for $t \in \mathcal{N}$,

$$\sum_{t\in\mathcal{N}}(1 - \mathbf{f}_t\mathbf{y}_t - \sigma_t) \leq \sum_{t\in\mathcal{N}} a_t\tilde{\mathcal{L}}(\mathbf{y}_t \cdot \mathbf{U}^\top\mathbf{x}_t) + \frac{1}{2}\mathrm{tr}(\mathbf{U}^\top\mathbf{A}_\mathcal{N}\mathbf{U}), \tag{3}$$

There are two types of update trials: (I) when an error occurs, i.e., $t \in \mathcal{M}$ and $-\mathbf{f}_t\mathbf{y}_t \geq 0$,

$$\sum_{t\in\mathcal{M}}(1 - \mathbf{f}_t\mathbf{y}_t - \sigma_t) \geq M - \sum_{t\in\mathcal{M}}\sigma_t;$$

and (II) when no error occurs, i.e., $t \in \mathcal{D}$ and $0 \leq \mathbf{f}_t\mathbf{y}_t \leq \sigma_t \Rightarrow -\mathbf{f}_t\mathbf{y}_t + \sigma_t \geq 0$,

$$\sum_{t\in\mathcal{D}}(1 - \mathbf{f}_t\mathbf{y}_t + \sigma_t - 2\sigma_t) \geq D - 2\sum_{t\in\mathcal{D}}\sigma_t.$$

Combining two cases with the upper bound (3), and substituting the inequality $\sum_{t\in\mathcal{M}\cup\mathcal{D}} 2\sigma_t \leq \frac{b}{b-1}\log(\frac{1}{b}\mathbf{A}_\mathcal{N})$, we finish the proof. $\square$

Conclusion: Empirically, the update number of adaptive-margin method can be comparable with or smaller than that of error-driven algorithm, due to a fast convergence of adaptive-margin learning. Due to the deduction of the low-confident update trials $|\mathcal{D}|$, the error bound of Algorithm 1 can be lower than that of the weighted min-max algorithm using error-driven update rules.

1.4 Proof of Theory 4 - MOLG-B

Theorem 4. *Algorithm 2 runs on a node-label sequence* $\{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_T, y_T)\}$. *If we set* $\varphi_t^{i2} = (\frac{b}{b-1})^2(4(b-1)\|\mathbf{U}\|_F^2 + 4\gamma\log|\frac{1}{b}\mathbf{A}_{t-1}^i| + 144\log\frac{t+4}{\delta})$ *where* $\ell_t(\mathbf{U}) \leq \gamma$, *then for any* $\mathbf{U} \in \mathbb{R}^{d\times K}$, *such that* $|\mathbf{u}^{i\top}\mathbf{x}_t| \leq 1$, *the inequality holds,*

$$R_T \leq \sqrt{(\frac{b}{b-1})^3 T}(\sqrt{H_1H_2} + H_2),$$

with probability at least $1-\delta$ *over* T *trials, where* $H_1 = 2(b-1)\|\mathbf{U}\|_F^2 + 72\log\frac{t+4}{\delta}$ *and* $H_2 = 2Kd\gamma\log(1 + \frac{T}{Kdb})$.

Proof: Note that the update rule is

$$\mathbf{B}_t = \mathbf{B}_{t-1} + a_t\mathbf{x}_t\mathbf{y}_t^\top \quad \mathbf{A}_t = \mathbf{A}_{t-1} + a_t\mathbf{x}_t\mathbf{x}_t^\top,$$

or $\mathbf{A}_t^{-1} = \mathbf{A}_{t-1}^{-1} - \mathbf{A}_{t-1}^{-1}\mathbf{x}_t\mathbf{x}_t^\top\mathbf{A}_{t-1}^{-1}$ according to Woodbury identity. Given an annotation $\mathcal{D}_t(\mathbf{U}, \mathbf{V}) = \|\mathbf{U} - \mathbf{V}\|_{\mathbf{A}_t}^2$, the following equations can be derived,

$$\begin{aligned} &a_t\left(\|\mathbf{y}_t - \mathbf{W}_{t-1}^\top\mathbf{x}_t\|^2 - \|\mathbf{y}_t - \mathbf{U}^\top\mathbf{x}_t\|^2\right) \\ =&\mathcal{D}_{t-1}(\mathbf{U}, \mathbf{W}_{t-1}) - \mathcal{D}_t(\mathbf{U}, \mathbf{W}_t) + \mathcal{D}_t(\mathbf{W}_{t-1}, \mathbf{W}_t), \end{aligned} \tag{4}$$

$$a_t^2\|\mathbf{y}_t - \mathbf{W}_{t-1}^\top \mathbf{x}_t\|^2 \mathbf{x}_t^\top \mathbf{A}_t^{-1}\mathbf{x}_t = \mathcal{D}_t(\mathbf{W}_{t-1}, \mathbf{W}_t). \tag{5}$$

Assume that $\ell_t(\mathbf{U}) = \|\mathbf{y}_t - \mathbf{U}^\top \mathbf{x}_t\|^2 \le r, (r > 1)$ for any $\mathbf{U} \in \mathbb{R}^{n\times K}$ and $a_t = \frac{1}{1-\mathbf{x}_t^\top \mathbf{A}_{t-1}^{-1}\mathbf{x}_t} \le \frac{b}{b-1}$ for any $t \in [T]$, the cumulative sum of Eq. (5) can be bounded,

$$\begin{aligned}\sum_{t=1}^{T} \mathcal{D}_t(\mathbf{W}_{t-1}, \mathbf{W}_t) &= \sum_{t=1}^{T} a_t^2\|\mathbf{y}_t - \mathbf{W}_{t-1}^\top \mathbf{x}_t\|^2 \mathbf{x}_t^\top \mathbf{A}_t^{-1}\mathbf{x}_t \\ &\le \frac{rb}{b-1}\sum_{t=1}^{T} a_t \mathbf{x}_t^\top \mathbf{A}_t^{-1}\mathbf{x}_t \le \frac{rb}{b-1}\sum_{t=1}^{T}\log\frac{|\mathbf{A}_t|}{|\mathbf{A}_t - a_t\mathbf{x}_t\mathbf{x}_t^\top|} \\ &= \frac{rb}{b-1}\log|\frac{1}{b}\mathbf{A}_T|,\end{aligned} \tag{6}$$

where the last inequality is similar to the proof of Theorem 5 in [3]. Equipped with the bound (6), the cumulative sum of Eq. (4) can be bounded

$$\begin{aligned}&\sum_{s=1}^{t-1} a_s\left(\|\mathbf{y}_s - \mathbf{W}_{s-1}^\top \mathbf{x}_s\|^2 - \|\mathbf{y}_s - \mathbf{U}^\top \mathbf{x}_s\|^2\right) \\ \le&\mathcal{D}_0(\mathbf{U}, \mathbf{0}) - \mathcal{D}_{t-1}(\mathbf{U}, \mathbf{W}_{t-1}) + \frac{rb}{b-1}\log|\frac{1}{b}\mathbf{A}_{t-1}|.\end{aligned} \tag{7}$$

According to the Cauchy-Schwarz inequality (dual norms), we have

$$\|\hat{\Delta}_t - \Delta_t\|^2 = \|(\mathbf{W}_{t-1}^\top - \mathbf{U}^\top)\mathbf{x}_t\|^2 \le 2\mathbf{x}_t^\top \mathbf{A}_{t-1}^{-1}\mathbf{x}_t \mathcal{D}_{t-1}(\mathbf{U}, \mathbf{W}_{t-1}). \tag{8}$$

Due to $a_t = \frac{1}{1-\mathbf{x}_t^\top \mathbf{A}_{t-1}^{-1}\mathbf{x}_t} > 1$ for all t, we can infer according to the proof of Lemma 2 to Lemma 5 in [1],

$$\sum_{s=1}^{t-1} a_s\left(\|\mathbf{y}_s - \mathbf{W}_{s-1}^\top \mathbf{x}_s\|^2 - \|\mathbf{y}_s - \mathbf{U}^\top \mathbf{x}_s\|^2\right) \ge -36\log\frac{t+4}{\delta} \tag{9}$$

holds with probability at least $1-\delta$ over the t rounds. Substituting Eq. (7)-Eq. (9), we obtain for any $\mathbf{U} \in \mathbb{R}^{n\times K}$,

$$\|\hat{\Delta}_t - \Delta_t\|^2 \le 2\mathbf{x}^\top \mathbf{A}_{t-1}^{-1}\mathbf{x}_t\left(b\|\mathbf{U}\|_F^2 + \frac{rb}{b-1}\log|\frac{1}{b}\mathbf{A}_{t-1}| + 36\log\frac{t+4}{\delta}\right) \tag{10}$$

hold with probability at least $1-\delta$ over the $t-1$ rounds.

Since $\mathbf{A}_t^{-1} \preceq \mathbf{A}_{t-1}^{-1}$, we have $\mathbf{x}_t^\top \mathbf{A}_{t-1}^{-1}\mathbf{x}_t \le \ldots \le \mathbf{x}_t^\top \mathbf{A}_0^{-1}\mathbf{x}_t = \frac{1}{b}\|\mathbf{x}_t\|^2$. Assume that $\|\mathbf{x}_t\| \le 1$, we infer that $0 \le \mathbf{x}_t^\top \mathbf{A}_{t-1}^{-1}\mathbf{x}_t \le \frac{1}{b}$ where we let $b > 1$. Thus,

$$1 - \mathbf{x}_t^\top \mathbf{A}_{t-1}^{-1}\mathbf{x}_t \ge 1 - \frac{1}{b} \;\Rightarrow\; \frac{b}{b-1}(1 - \mathbf{x}_t^\top \mathbf{A}_{t-1}^{-1}\mathbf{x}_t) \ge 1.$$

Multiplying $\mathbf{x}_t^\top \mathbf{A}_{t-1}^{-1}\mathbf{x}_t$ on both sides, we obtain

$$\mathbf{x}_t^\top \mathbf{A}_{t-1}^{-1}\mathbf{x}_t \le \frac{b}{b-1}\mathbf{x}_t^\top(\mathbf{A}_{t-1}^{-1} - \mathbf{A}_{t-1}^{-1}\mathbf{x}_t\mathbf{x}_t^\top \mathbf{A}_{t-1}^{-1})\mathbf{x}_t = \frac{b}{b-1}\mathbf{x}_t^\top \mathbf{A}_t^{-1}\mathbf{x}_t. \tag{11}$$

Substituting Eq. (11) into Eq. (10), we obtain,

$$
\begin{aligned}
\|\hat{\Delta}_t - \Delta_t\|^2 &\leq \frac{1}{2}(\frac{b}{b-1})^2 \mathbf{x}_t^\top \mathbf{A}_t^{-1}\mathbf{x}_t \left(4(b-1)\|\mathbf{U}\|_F^2 + 4r\log|\frac{1}{b}\mathbf{A}_{t-1}| + \frac{b-1}{b}144\log\frac{t+4}{\delta}\right) \\
&\leq \sigma_t(\frac{b}{b-1})^2 \left(4(b-1)\|\mathbf{U}\|_F^2 + 4r\log|\frac{1}{b}\mathbf{A}_{t-1}| + 144\log\frac{t+4}{\delta}\right),
\end{aligned}
$$

where $\sigma_t = \frac{1}{2}a_t^2\mathbf{x}_t^\top \mathbf{A}_t^{-1}\mathbf{x}_t$ and the last inequality holds due to $a_t \geq 1$. We assume that

$$
\varphi_t^2 = (\frac{b}{b-1})^2 \left(4(b-1)\|\mathbf{U}\|_F^2 + 4r\log|\frac{1}{b}\mathbf{A}_{t-1}| + 144\log\frac{t+4}{\delta}\right), \tag{12}
$$

and bound the cumulative sum of σ_t for $t \in [T]$,

$$
\sum_{s=1}^{T}\sigma_s = \frac{1}{2}\sum_{s=1}^{T} a_s^2\mathbf{x}_s^\top \mathbf{A}_s^{-1}\mathbf{x}_s \leq \frac{b}{2(b-1)}\log|\frac{1}{b}\mathbf{A}_T| \leq \frac{b}{2(b-1)}Kn\log(1+\frac{T}{Knb}).
$$

Assume that

$$
H_1 = 2(b-1)\|\mathbf{U}\|_F^2 + 72\log\frac{t+4}{\delta}, \quad H_2 = 2Knr\log(1+\frac{T}{Knb})
$$

we have that

$$
\sum_{t=1}^{T}\|\hat{\Delta}_t - \Delta_t\|^2 \leq \sum_{t=1}^{T}\varphi_t^2\sigma_t \leq 2(H_1+H_2)(\frac{b}{b-1})^2\sum_{t=1}^{T}\sigma_t \leq (\frac{b}{b-1})^3(H_1+H_2)H_2. \tag{13}
$$

with probability at least $1-\delta$ over T rounds. Since $\sum_{t=1}^{T}A_t^2 \leq M$ implies $\sum_{t=1}^{T}A_t \leq \sqrt{TM}$, we obtain

$$
\begin{aligned}
&\sum_{t=1}^{T}(\mathbb{P}_t(y_t \neq \hat{y}_t) - \mathbb{P}_t(y_t \neq y_t^*)) \leq \sum_{t=1}^{T}|\Delta_t - \hat{\Delta}_t| \\
\leq &\sqrt{(\frac{b}{b-1})^3T}\sqrt{H_1H_2 + H_2^2} \leq \sqrt{(\frac{b}{b-1})^3T}(\sqrt{H_1H_2} + H_2),
\end{aligned} \tag{14}
$$

where the last inequality holds due to $\sqrt{A+B} \leq \sqrt{A} + \sqrt{B}$. □

References

[1] Koby Crammer and Claudio Gentile. Multiclass classification with bandit feedback using adaptive regularization. *Machine learning*, 90(3):347–383, 2013.

[2] Jürgen Forster. On relative loss bounds in generalized linear regression. In *Fundamentals of Computation Theory*, pages 269–280, 1999.

[3] Edward Moroshko and Koby Crammer. Weighted last-step min–max algorithm with improved sub-logarithmic regret. *Theoretical Computer Science*, 558:107–124, 2014.