Supplementary Materials for the TMLR submission "Linear Indexed Minimum Empirical Divergence Algorithms"

A BaseLinUCB Algorithm

Here, we present the BaseLinUCB algorithm used as a subroutine in SubLinIMED (Algorithm 2).

Algorithm 3 BaseLinUCB

1: **Input:** $\gamma = \frac{1}{2t^2}$, $\alpha = \sqrt{\frac{1}{2} \ln \frac{2TK}{\gamma}}$, $\Psi_t \subseteq \{1, 2, ..., t-1\}$
2: $V_t = I_d + \sum_{\tau \in \Psi_t} x_{\tau, A_\tau}^T x_{\tau, A_\tau}$
3: $b_t = \sum_{\tau \in \Psi_t} Y_{\tau, A_\tau} x_{\tau, A_\tau}$
4: $\hat{\theta}_t = V_t^{-1} b_t$
5: Observe K arm features $x_{t,1}, x_{t,2}, \ldots, x_{t,K} \in \mathbb{R}^d$
6: **for** $a \in [K]$ **do**
7: $\quad w_{t,a} = \alpha \sqrt{x_{t,a}^T V_t^{-1} x_{t,a}}$
8: $\quad \hat{Y}_{t,a} = \langle \hat{\theta}_t, x_{t,a} \rangle$
9: **end for**

B Comparison to other related work

Saber et al. (2021) adopts the IMED algorithm to unimodal bandits which achieves asymptotically optimality for one-dimensional exponential family distributions. In their algorithm IMED-UB, they narrow down the search region to the neighboring regions of the empirically best arm and then implement the IMED algorithm for K-armed bandit as in Honda & Takemura (2015). This design is inspired by the lower bound and only involves the neighboring arms of the best arm. The settings in which the algorithm in Saber et al. (2021) is applied to is different from our proposed LinIMED algorithms as we focus on linear bandits, not unimodal bandits.

Liu et al. (2024) proposes an algorithm that achieves $\widetilde{O}(\sqrt{T})$ regret for adversarial linear bandits with stochastic action sets in the absence of a simulator or prior knowledge on the distribution. Although their setting is different from ours, they also use a bonus term $-\alpha_t \hat{\Sigma}_t^{-1}$ in the lifted covariance matrix to encourage exploration. This is similar to our choice of the second term $\log(1/\beta_{t-1}(\gamma)\|x_{t,a}\|_{V_{t-1}^{-1}}^2)$ in LinIMED-1.

C Proof of the regret bound for LinIMED-1 (Complete proof of Theorem 1)

Here and in the following, we abbreviate $\beta_t(\gamma)$ as β_t, i.e., we drop the dependence of β_t on γ, which is taken to be $\frac{1}{t^2}$ per Eqn. (5).

C.1 Statement of Lemmas for LinIMED-1

We first state the following lemmas which respectively show the upper bound of F_1 to F_4:

Lemma 2. *Under Assumption 1, the assumption that $\langle \theta^*, x_{t,a} \rangle \geq 0$ for all $t \geq 1$ and $a \in \mathcal{A}_t$, and the assumption that $\sqrt{\lambda}S \geq 1$, then for the free parameter $0 < \Gamma < 1$, the term F_1 for LinIMED-1 satisfies:*

$$F_1 \leq O(1) + T\Gamma + O\left(\frac{d\beta_T \log(\frac{T}{\Gamma^2})}{\Gamma} \log\left(1 + \frac{L^2\beta_T \log(\frac{T}{\Gamma^2})}{\lambda\Gamma^2}\right)\right). \tag{12}$$

With the choice of Γ as in Eqn. (5),

$$F_1 \leq O\left(d\sqrt{T}\log^{\frac{3}{2}}T\right).$$

Lemma 3. *Under Assumption 1, and the assumption that $\sqrt{\lambda}S \geq 1$, for the free parameter $0 < \Gamma < 1$, the term F_2 for LinIMED-1 satisfies:*

$$F_2 \leq 2T\Gamma + O\left(\frac{d\beta_T \log T}{\Gamma}\right) \log\left(1 + \frac{L^2\beta_T \log T}{\lambda\Gamma^2}\right) . \tag{13}$$

With the choice of Γ as in Eqn. (5),

$$F_2 \leq O\left(d\sqrt{T}\log^{\frac{3}{2}}T\right) .$$

Lemma 4. *Under Assumption 1, and the assumption that $\sqrt{\lambda}S \geq 1$, for the free parameter $0 < \Gamma < 1$, the term F_3 for LinIMED-1 satisfies:*

$$F_3 \leq 2T\Gamma + O\left(\frac{d\beta_T \log(T)}{\Gamma} \log\left(1 + \frac{L^2\beta_T \log(T)}{\lambda\Gamma^2}\right)\right) . \tag{14}$$

With the choice of Γ as in Eqn. (5),

$$F_3 \leq O\left(d\sqrt{T}\log^{\frac{3}{2}}T\right) .$$

Lemma 5. *Under Assumption 1, for the free parameter $0 < \Gamma < 1$, the term F_4 for LinIMED-1 satisfies:*

$$F_4 \leq T\Gamma + O(1) .$$

With the choice of Γ as in Eqn. (5),

$$F_4 \leq O\left(d\sqrt{T}\log^{\frac{3}{2}}T\right) .$$

C.2 Proof of Lemma 2

Proof. From the event C_t and the fact that $\langle\theta^*, x_t^*\rangle = \Delta_t + \langle\theta^*, X_t\rangle \geq \Delta_t$ (here is where we use that $\langle\theta^*, x_{t,a}\rangle \geq 0$ for all t and a), we obtain $\max_{b\in\mathcal{A}_t}\langle\hat{\theta}_{t-1}, x_{t,b}\rangle > (1 - \frac{1}{\sqrt{\log T}})\Delta_t$. For convenience, define $\hat{A}_t := \arg\max_{b\in\mathcal{A}_t}\langle\hat{\theta}_{t-1}, x_{t,b}\rangle$ as the empirically best arm at time step t, where ties are broken arbitrarily, then use $\hat{X}_t$ to denote the corresponding context of the arm $\hat{A}_t$. Therefore from the Cauchy–Schwarz inequality, we have $\|\hat{\theta}_{t-1}\|_{V_{t-1}}\|\hat{X}_t\|_{V_{t-1}^{-1}} \geq \langle\hat{\theta}_{t-1}, \hat{X}_t\rangle > (1 - \frac{1}{\sqrt{\log T}})\Delta_t$. This implies that

$$\|\hat{X}_t\|_{V_{t-1}^{-1}} \geq \frac{(1 - \frac{1}{\sqrt{\log T}})\Delta_t}{\|\hat{\theta}_{t-1}\|_{V_{t-1}}} .$$

On the other hand, we claim that $\|\hat{\theta}_{t-1}\|_{V_{t-1}}$ can be upper bounded as $O(\sqrt{T})$. This can be seen from the fact that $\|\hat{\theta}_{t-1}\|_{V_{t-1}} = \|\hat{\theta}_{t-1} - \theta^* + \theta^*\|_{V_{t-1}} \leq \|\hat{\theta}_{t-1} - \theta^*\|_{V_{t-1}} + \|\theta^*\|_{V_{t-1}}$. Since the event B_t holds, we know the first term is upper bounded by $\sqrt{\beta_{t-1}(\gamma)}$, and since the maximum eigenvalue of the matrix V_{t-1} is upper bounded by $\lambda + TL^2$ and $\|\theta^*\| \leq S$, the second term is upper bounded by $S\sqrt{\lambda + TL^2}$. Hence, $\|\hat{\theta}_{t-1}\|_{V_{t-1}}$ is upper bounded by $O(\sqrt{T})$. Then one can substitute this bound back into Eqn. (2), and this yields

$$\|\hat{X}_t\|_{V_{t-1}^{-1}} \geq \Omega\left(\frac{1}{\sqrt{T}}\left(1 - \frac{1}{\sqrt{\log T}}\right)\Delta_t\right) .$$

Furthermore, by our design of the algorithm, the index of A_t is not larger than the index of the arm with the largest empirical reward at time t. Hence,

$$I_{t,A_t} = \frac{\hat{\Delta}_{t,A_t}^2}{\beta_{t-1}(\gamma)\|X_t\|_{V_{t-1}^{-1}}^2} + \log\frac{1}{\beta_{t-1}(\gamma)\|X_t\|_{V_{t-1}^{-1}}^2} \leq \log\frac{1}{\beta_{t-1}(\gamma)\|\hat{X}_t\|_{V_{t-1}^{-1}}^2} .$$

If $\|X_t\|^2_{V^{-1}_{t-1}} \geq \frac{\Delta^2_t}{\beta_{t-1}}$, by using Corollary 1 and the "peeling device" (Lattimore & Szepesvári, 2020, Chapter 9) on Δ_t such that $2^{-l} < \Delta_t \leq 2^{-l+1}$ for $l = 1, 2, \ldots, \lceil Q \rceil$ where $Q = -\log_2 \Gamma$,

$$\mathbb{E} \sum_{t=1}^T \Delta_t \cdot \mathbb{1}\{B_t, C_t, D_t\} \cdot \mathbb{1}\left\{\|X_t\|^2_{V^{-1}_{t-1}} \geq \frac{\Delta^2_t}{\beta_{t-1}}\right\} \leq \mathbb{E} \sum_{t=1}^T \Delta_t \cdot \mathbb{1}\left\{\|X_t\|^2_{V^{-1}_{t-1}} \geq \frac{\Delta^2_t}{\beta_{t-1}}\right\} \tag{15}$$

$$= \mathbb{E} \sum_{t=1}^T \Delta_t \cdot \mathbb{1}\left\{\|X_t\|^2_{V^{-1}_{t-1}} \geq \frac{\Delta^2_t}{\beta_{t-1}}\right\} \cdot \mathbb{1}\{\Delta_t \leq \Gamma\} + \mathbb{E} \sum_{t=1}^T \Delta_t \cdot \mathbb{1}\left\{\|X_t\|^2_{V^{-1}_{t-1}} \geq \frac{\Delta^2_t}{\beta_{t-1}}\right\} \cdot \mathbb{1}\{\Delta_t > \Gamma\}$$

$$\leq T\Gamma + \mathbb{E} \sum_{t=1}^T \sum_{l=1}^{\lceil Q \rceil} \Delta_t \cdot \mathbb{1}\left\{\|X_t\|^2_{V^{-1}_{t-1}} \geq \frac{\Delta^2_t}{\beta_{t-1}}\right\} \cdot \mathbb{1}\{2^{-l} < \Delta_t \leq 2^{-l+1}\}$$

$$\leq T\Gamma + \mathbb{E} \sum_{t=1}^T \sum_{l=1}^{\lceil Q \rceil} 2^{-l+1} \cdot \mathbb{1}\left\{\|X_t\|^2_{V^{-1}_{t-1}} \geq \frac{2^{-2l}}{\beta_T}\right\}$$

$$\leq T\Gamma + \mathbb{E} \sum_{l=1}^{\lceil Q \rceil} 2^{-l+1} \frac{6d\beta_T}{2^{-2l}} \log\left(1 + \frac{2L^2\beta_T}{\lambda \cdot 2^{-2l}}\right)$$

$$= T\Gamma + \mathbb{E} \sum_{l=1}^{\lceil Q \rceil} 2^l \cdot 12d\beta_T \log\left(1 + \frac{2^{2l+1}L^2\beta_T}{\lambda}\right)$$

$$< T\Gamma + \mathbb{E} \sum_{l=1}^{\lceil Q \rceil} 2^l \cdot 12d\beta_T \log\left(1 + \frac{2^{2Q+3}L^2\beta_T}{\lambda}\right)$$

$$= T\Gamma + (2^{\lceil Q \rceil} - 1) \cdot 24d\beta_T \log\left(1 + \frac{2^{2Q+3}L^2\beta_T}{\lambda}\right)$$

$$< T\Gamma + \frac{48d\beta_T}{\Gamma} \log\left(1 + \frac{8L^2\beta_T}{\lambda\Gamma^2}\right)$$

Then with the choice of Γ as in Eqn. (5),

$$\mathbb{E} \sum_{t=1}^T \Delta_t \cdot \mathbb{1}\{B_t, C_t, D_t\} \cdot \mathbb{1}\left\{\|X_t\|^2_{V^{-1}_{t-1}} \geq \frac{\Delta^2_t}{\beta_{t-1}}\right\}$$

$$< d\sqrt{T} \log^{\frac{3}{2}} T + \frac{48\beta_T \sqrt{T}}{\log^{\frac{3}{2}} T} \log\left(1 + \frac{8L^2\beta_T T}{\lambda d^2 \log^3 T}\right)$$

$$\leq O\left(d\sqrt{T} \log^{\frac{3}{2}} T\right). \tag{16}$$

Otherwise we have $\|X_t\|^2_{V^{-1}_{t-1}} < \frac{\Delta^2_t}{\beta_{t-1}}$, then $\log \frac{1}{\beta_{t-1}\|X_t\|^2_{V^{-1}_{t-1}}} > 0$ since $\Delta_t \leq 1$. Substituting this into Eqn. (4), then using the event D_t and the bound in (3), we deduce that for all T sufficiently large, we have $\|X_t\|^2_{V^{-1}_{t-1}} \geq \Omega\left(\frac{\Delta^2_t}{\beta_{t-1}\log(T/\Delta^2_t)}\right)$. Therefore by using Corollary 1 and the "peeling device" (Lattimore & Szepesvári, 2020, Chapter 9) on Δ_t such that $2^{-l} < \Delta_t \leq 2^{-l+1}$ for $l = 1, 2, \ldots, \lceil Q \rceil$ where $\Gamma := 2^{-Q}$ is a free parameter that we can choose. Consider,

$$\mathbb{E} \sum_{t=1}^T \Delta_t \cdot \mathbb{1}\{B_t, C_t, D_t\} \cdot \mathbb{1}\left\{\|X_t\|^2_{V^{-1}_{t-1}} < \frac{\Delta^2_t}{\beta_{t-1}}\right\}$$

$$\leq \mathbb{E} \sum_{t=1}^T \Delta_t \cdot \mathbb{1}\{B_t, C_t, D_t\} \cdot \mathbb{1}\left\{\Delta_t \leq 2^{-\lceil Q \rceil}\right\} \cdot \mathbb{1}\left\{\|X_t\|^2_{V^{-1}_{t-1}} < \frac{\Delta^2_t}{\beta_{t-1}}\right\}$$

$$+ \mathbb{E} \sum_{t=1}^T \Delta_t \cdot \mathbb{1}\{B_t, C_t, D_t\} \cdot \mathbb{1}\left\{\Delta_t > 2^{-\lceil Q \rceil}\right\} \cdot \mathbb{1}\left\{\|X_t\|^2_{V^{-1}_{t-1}} < \frac{\Delta^2_t}{\beta_{t-1}}\right\}$$

$$\leq O(1) + T\Gamma + \mathbb{E}\sum_{t=1}^{T}\Delta_t \cdot \mathbb{1}\left\{\|X_t\|^2_{V_{t-1}^{-1}} \geq \Omega\left(\frac{\Delta_t^2}{\beta_{t-1}\log(T/\Delta_t^2)}\right)\right\}\mathbb{1}\left\{\Delta_t > 2^{-\lceil Q\rceil}\right\}$$

$$\leq O(1) + T\Gamma + \mathbb{E}\sum_{t=1}^{T}\sum_{l=1}^{\lceil Q\rceil}\Delta_t \cdot \mathbb{1}\left\{\|X_t\|^2_{V_{t-1}^{-1}} \geq \Omega\left(\frac{\Delta_t^2}{\beta_{t-1}\log(T/\Delta_t^2)}\right)\right\}\mathbb{1}\{2^{-l} < \Delta_t \leq 2^{-l+1}\}$$

$$\leq O(1) + T\Gamma + \mathbb{E}\sum_{t=1}^{T}\sum_{l=1}^{\lceil Q\rceil}2^{-l+1} \cdot \mathbb{1}\left\{\|X_t\|^2_{V_{t-1}^{-1}} \geq \Omega\left(\frac{2^{-2l}}{\beta_{t-1}\log(T\cdot 2^{2l})}\right)\right\}$$

$$= O(1) + T\Gamma + \mathbb{E}\sum_{l=1}^{\lceil Q\rceil}2^{-l+1}\sum_{t=1}^{T}\mathbb{1}\left\{\|X_t\|^2_{V_{t-1}^{-1}} \geq \Omega\left(\frac{2^{-2l}}{\beta_{t-1}\log(T\cdot 2^{2l})}\right)\right\}$$

$$\leq O(1) + T\Gamma + \mathbb{E}\sum_{l=1}^{\lceil Q\rceil}2^{-l+1}O\left(2^{2l}d\beta_T\log(T\cdot 2^{2l})\log\left(1 + \frac{2L^2\cdot 2^{2l}\beta_T\log(T\cdot 2^{2l})}{\lambda}\right)\right)$$

$$< O(1) + T\Gamma + \mathbb{E}\sum_{l=1}^{\lceil Q\rceil}2^{l+1}\cdot O\left(d\beta_T\log(\frac{T}{\Gamma^2})\log\left(1 + \frac{L^2\beta_T\log(\frac{T}{\Gamma^2})}{\lambda\Gamma^2}\right)\right)$$

$$\leq O(1) + T\Gamma + O\left(\frac{d\beta_T\log(\frac{T}{\Gamma^2})}{\Gamma}\log\left(1 + \frac{L^2\beta_T\log(\frac{T}{\Gamma^2})}{\lambda\Gamma^2}\right)\right),$$

This proves Eqn. (12). Then with the choice of the parameters as in Eqn. (5),

$$\mathbb{E}\sum_{t=1}^{T}\Delta_t \cdot \mathbb{1}\{B_t, C_t, D_t\} \cdot \mathbb{1}\left\{\|X_t\|^2_{V_{t-1}^{-1}} < \frac{\Delta_t^2}{\beta_{t-1}}\right\}$$

$$< O(1) + d\sqrt{T}\log^{\frac{3}{2}}T + O\left(d\beta_T\log\left(\frac{T^2}{d^2\log^3 T}\right)\frac{\sqrt{T}}{d\log^{\frac{3}{2}}T}\log\left(1 + \frac{L^2\beta_T T}{\lambda d^2\log^3 T}\cdot\log\left(\frac{T^2}{d^2\log^3 T}\right)\right)\right)$$

$$\leq O\left(d\sqrt{T}\log^{\frac{3}{2}}T\right).$$

Hence, we can upper bound F_1 as

$$F_1 = \mathbb{E}\sum_{t=1}^{T}\Delta_t \cdot \mathbb{1}\{B_t, C_t, D_t\} \cdot \mathbb{1}\left\{\|X_t\|^2_{V_{t-1}^{-1}} \geq \frac{\Delta_t^2}{\beta_{t-1}}\right\} + \mathbb{E}\sum_{t=1}^{T}\Delta_t \cdot \mathbb{1}\{B_t, C_t, D_t\} \cdot \mathbb{1}\left\{\|X_t\|^2_{V_{t-1}^{-1}} < \frac{\Delta_t^2}{\beta_{t-1}}\right\}$$

$$\leq O\left(d\sqrt{T}\log^{\frac{3}{2}}T\right) + O\left(d\sqrt{T}\log^{\frac{3}{2}}T\right)$$

$$\leq O\left(d\sqrt{T}\log^{\frac{3}{2}}T\right),$$

which concludes the proof. $\qquad\square$

C.3 Proof of Lemma 3

Proof. Since C_t and $\overline{D}_t$ together imply that $\langle\theta^*, x_t^*\rangle - \delta < \varepsilon + \langle\hat{\theta}_{t-1}, X_t\rangle$, then using the choices of δ and ε, we have $\langle\hat{\theta}_{t-1} - \theta^*, X_t\rangle > \frac{\Delta_t}{\sqrt{\log T}}$. Substituting this into the event B_t and using the Cauchy–Schwarz inequality, we have

$$\|X_t\|^2_{V_{t-1}^{-1}} \geq \frac{\Delta_t^2}{\beta_{t-1}(\gamma)\log T}.$$

Again applying the "peeling device" on Δ_t and Corollary 1, we can upper bound F_2 as follows:

$$F_2 \leq \mathbb{E}\sum_{t=1}^{T}\Delta_t \cdot \mathbb{1}\left\{\|X_t\|^2_{V_{t-1}^{-1}} \geq \frac{\Delta_t^2}{\beta_{t-1}\log T}\right\}$$

$$\leq T\Gamma + \mathbb{E}\sum_{t=1}^{T}\sum_{l=1}^{\lceil Q\rceil}\Delta_t \cdot \mathbb{1}\left\{\|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{\Delta_t^2}{\beta_{t-1}\log T}\right\} \cdot \mathbb{1}\left\{2^{-l} < \Delta_t \leq 2^{-l+1}\right\}$$

$$\leq T\Gamma + \mathbb{E}\sum_{t=1}^{T}\sum_{l=1}^{\lceil Q\rceil} 2^{-l+1} \cdot \mathbb{1}\left\{\|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{2^{-2l}}{\beta_T\log T}\right\}$$

$$\leq T\Gamma + \mathbb{E}\sum_{l=1}^{\lceil Q\rceil} 2^{-l+1} \cdot 2^{2l} \cdot 6d\beta_T(\log T)\log\left(1 + \frac{2^{2l+1}\cdot L^2\beta_T\log T}{\lambda}\right)$$

$$\leq T\Gamma + \mathbb{E}\sum_{l=1}^{\lceil Q\rceil} 2^{l} \cdot 12d\beta_T(\log T)\log\left(1 + \frac{2^{2\lceil Q\rceil+1}\cdot L^2\beta_T\log T}{\lambda}\right)$$

$$= T\Gamma + (2^{\lceil Q\rceil} - 1) \cdot 24d\beta_T(\log T)\log\left(1 + \frac{2^{2\lceil Q\rceil+1}\cdot L^2\beta_T\log T}{\lambda}\right)$$

$$< T\Gamma + \frac{48d\beta_T\log T}{\Gamma}\log\left(1 + \frac{8L^2\beta_T\log T}{\lambda\Gamma^2}\right)$$

$$= T\Gamma + O\left(\frac{d\beta_T\log T}{\Gamma}\log\left(1 + \frac{L^2\beta_T\log T}{\lambda\Gamma^2}\right)\right)$$

This proves Eqn. (13). Hence with the choice of the parameter Γ as in Eqn. (5),

$$F_2 \leq d\sqrt{T}\log^{\frac{3}{2}} T + O\left(d\sqrt{T}\log^{\frac{3}{2}} T\right)$$

$$\leq O\left(d\sqrt{T}\log^{\frac{3}{2}} T\right).$$

$$\square$$

C.4 Proof of Lemma 4

Proof. For F_3, this is the case when the best arm at time t does not perform sufficiently well so that the empirically largest reward at time t is far from the highest expected reward. One observes that minimizing F_3 results in a tradeoff with respect to F_1. On the event $\overline{C}_t$, we can apply the "peeling device" on $\langle\theta^*, x_t^*\rangle - \langle\hat{\theta}_{t-1}, x_t^*\rangle$ such that $\frac{q+1}{2}\delta \leq \langle\theta^*, x_t^*\rangle - \langle\hat{\theta}_{t-1}, x_t^*\rangle < \frac{q+2}{2}\delta$ where $q \in \mathbb{N}$. Then using the fact that $I_{t,A_t} \leq I_{t,a_t^*}$, we have

$$\log\frac{1}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2} < \frac{q^2\delta^2}{4\beta_{t-1}\|x_t^*\|_{V_{t-1}^{-1}}^2} + \log\frac{1}{\beta_{t-1}\|x_t^*\|_{V_{t-1}^{-1}}^2} . \tag{17}$$

On the other hand, using the event B_t and the Cauchy–Schwarz inequality, it holds that

$$\|x_t^*\|_{V_{t-1}^{-1}} \geq \frac{(q+1)\delta}{2\sqrt{\beta_{t-1}}} . \tag{18}$$

If $\|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{\Delta_t^2}{\beta_{t-1}}$, the regret in this case is bounded by $O(d\sqrt{T\log T})$ (similar to the procedure to get from Eqn. (15) to Eqn. (16)). Otherwise $\log\frac{1}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2} > \log\frac{1}{\Delta_t^2} \geq 0$, then combining Eqn. (17) and Eqn. (18) implies that

$$\|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{(q+1)^2\delta^2}{4\beta_{t-1}}\exp\left(-\frac{q^2}{(q+1)^2}\right) .$$

Notice here with $\sqrt{\lambda}S \geq 1$, $\|X_t\|_{V_{t-1}^{-1}}^2 < \frac{\Delta_t^2}{\beta_{t-1}} \leq \frac{1}{\beta_{t-1}} \leq 1$, it holds that for all $q \in \mathbb{N}$,

$$\frac{(q+1)^2\delta^2}{4\beta_{t-1}}\exp\left(-\frac{q^2}{(q+1)^2}\right) < 1 . \tag{19}$$

Using Corollary 1, one can show that:

$$\sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{B_t, \overline{C}_t\right\} \cdot \mathbb{1}\left\{\|X_t\|_{V_{t-1}^{-1}}^2 < \frac{\Delta_t^2}{\beta_{t-1}}\right\}$$

$$\leq T\Gamma + \sum_{t=1}^{T}\sum_{l=1}^{\lceil Q\rceil} \Delta_t \cdot \mathbb{1}\left\{B_t, \overline{C}_t\right\} \cdot \mathbb{1}\left\{\|X_t\|_{V_{t-1}^{-1}}^2 < \frac{\Delta_t^2}{\beta_{t-1}}\right\} \cdot \mathbb{1}\left\{2^{-l} < \Delta_t \leq 2^{-l+1}\right\}$$

$$\leq T\Gamma + \sum_{t=1}^{T}\sum_{l=1}^{\lceil Q\rceil}\sum_{q=1}^{\infty} \Delta_t \cdot \mathbb{1}\left\{B_t\right\} \cdot \mathbb{1}\left\{\frac{q+1}{2}\delta \leq \langle\theta^*, x_t^*\rangle - \langle\hat{\theta}_{t-1}, x_t^*\rangle < \frac{q+2}{2}\delta\right\} \cdot \mathbb{1}\left\{\|X_t\|_{V_{t-1}^{-1}}^2 < \frac{\Delta_t^2}{\beta_{t-1}}\right\}$$
$$\cdot \mathbb{1}\left\{2^{-l} < \Delta_t \leq 2^{-l+1}\right\}$$

$$\leq T\Gamma + \sum_{t=1}^{T}\sum_{l=1}^{\lceil Q\rceil}\sum_{q=1}^{\infty} \Delta_t \cdot \mathbb{1}\left\{1 \geq \|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{(q+1)^2\delta^2}{4\beta_{t-1}} \exp\left(-\frac{q^2}{(q+1)^2}\right)\right\} \cdot \mathbb{1}\left\{2^{-l} < \Delta_t \leq 2^{-l+1}\right\}$$

$$= T\Gamma + \sum_{t=1}^{T}\sum_{l=1}^{\lceil Q\rceil}\sum_{q=1}^{\infty} \Delta_t \cdot \mathbb{1}\left\{1 \geq \|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{(q+1)^2\Delta_t^2}{4\beta_{t-1}\log T} \exp\left(-\frac{q^2}{(q+1)^2}\right)\right\} \cdot \mathbb{1}\left\{2^{-l} < \Delta_t \leq 2^{-l+1}\right\}$$

$$\leq T\Gamma + \sum_{t=1}^{T}\sum_{l=1}^{\lceil Q\rceil}\sum_{q=1}^{\infty} 2^{-l+1} \cdot \mathbb{1}\left\{1 \geq \|X_t\|_{V_{t-1}^{-1}}^2 > \frac{(q+1)^2 \cdot 2^{-2l}}{4\beta_T\log T} \exp\left(-\frac{q^2}{(q+1)^2}\right)\right\}$$

$$\leq T\Gamma + \sum_{l=1}^{\lceil Q\rceil}\sum_{q=1}^{\infty} 2^{-l+1} \cdot 2^{2l} \cdot 24 d\beta_T(\log T) \cdot \frac{\exp\left(\frac{q^2}{(q+1)^2}\right)}{(q+1)^2} \cdot \log\left(1 + \frac{2^{2l} \cdot 8L^2\beta_T\log T}{\lambda} \cdot \frac{\exp\left(\frac{q^2}{(q+1)^2}\right)}{(q+1)^2}\right)$$

$$< T\Gamma + \sum_{l=1}^{\lceil Q\rceil}\sum_{q=1}^{\infty} 2^{l+1} \cdot 24 d\beta_T(\log T) \cdot \frac{\exp\left(\frac{q^2}{(q+1)^2}\right)}{(q+1)^2} \cdot \log\left(1 + \frac{2^{2l+1} \cdot L^2\beta_T\log T}{\lambda}\right)$$

$$= T\Gamma + \sum_{l=1}^{\lceil Q\rceil} 2^{l+1} \cdot 24 d\beta_T(\log T) \cdot \log\left(1 + \frac{2^{2l+1} \cdot L^2\beta_T\log T}{\lambda}\right) \sum_{q=1}^{\infty} \frac{\exp\left(\frac{q^2}{(q+1)^2}\right)}{(q+1)^2}$$

$$\leq T\Gamma + \sum_{l=1}^{\lceil Q\rceil} 2^{l+1} \cdot 24 d\beta_T(\log T) \cdot \log\left(1 + \frac{2^{2l+1} \cdot L^2\beta_T\log T}{\lambda}\right) \cdot (1.09)$$

$$\leq T\Gamma + \sum_{l=1}^{\lceil Q\rceil} 2^{l+1} \cdot 27 d\beta_T(\log T) \cdot \log\left(1 + \frac{2^{2l+1} \cdot L^2\beta_T\log T}{\lambda}\right)$$

$$\leq T\Gamma + \sum_{l=1}^{\lceil Q\rceil} 2^{l+1} \cdot 27 d\beta_T(\log T) \cdot \log\left(1 + \frac{2^{2\lceil Q\rceil+1} \cdot L^2\beta_T\log T}{\lambda}\right)$$

$$< T\Gamma + \sum_{l=1}^{\lceil Q\rceil} \frac{216 d\beta_T\log T}{\Gamma} \cdot \log\left(1 + \frac{8L^2\beta_T\log T}{\lambda\Gamma^2}\right)$$

$$= T\Gamma + O\left(\frac{d\beta_T\log T}{\Gamma}\log\left(1 + \frac{L^2\beta_T\log T}{\lambda\Gamma^2}\right)\right). \tag{20}$$

Hence

$$F_3 = \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{B_t, \overline{C}_t\right\} \cdot \mathbb{1}\left\{\|X_t\|_{V_{t-1}^{-1}}^2 < \frac{\Delta_t^2}{\beta_{t-1}}\right\} + \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{B_t, \overline{C}_t\right\} \cdot \mathbb{1}\left\{\|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{\Delta_t^2}{\beta_{t-1}}\right\}$$

$$< O\left(\frac{d\beta_T}{\Gamma}\log\left(1 + \frac{L^2\beta_T}{\lambda\Gamma^2}\right)\right) + 2T\Gamma + O\left(\frac{d\beta_T\log T}{\Gamma}\log\left(1 + \frac{L^2\beta_T\log T}{\lambda\Gamma^2}\right)\right)$$

$$\leq 2T\Gamma + O\left(\frac{d\beta_T \log(T)}{\Gamma} \log\left(1 + \frac{L^2\beta_T \log(T)}{\lambda\Gamma^2}\right)\right) .$$

This proves Eqn. (14). With the choice of Γ as in Eqn. (5),

$$F_3 \leq 2d\sqrt{T}\log^{\frac{3}{2}} T + O\left(\frac{d\sqrt{T}\beta_T \log T}{d\log^{\frac{3}{2}} T} \log\left(1 + \frac{TL^2\beta_T \log T}{\lambda d^2 \log^3 T}\right)\right)$$

$$< 2d\sqrt{T}\log^{\frac{3}{2}} T + O\left(d\sqrt{T}\log^{\frac{3}{2}} T\right)$$

$$= O\left(d\sqrt{T}\log^{\frac{3}{2}} T\right) .$$

$\square$

C.5 Proof of Lemma 5

Proof. For F_4, the proof is straightforward by using Lemma 1 with the choice of γ. Indeed, one has

$$F_4 = \mathbb{E}\sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{\overline{B}_t\right\} \leq T\Gamma + \mathbb{E}\sum_{t=1}^{T}\sum_{l=1}^{\lceil Q \rceil} \Delta_t \cdot \mathbb{1}\left\{2^{-l} < \Delta_t \leq 2^{-l+1}\right\}\mathbb{1}\left\{\overline{B}_t\right\}$$

$$\leq T\Gamma + \mathbb{E}\sum_{t=1}^{T}\sum_{l=1}^{\lceil Q \rceil} 2^{-l+1}\mathbb{1}\left\{\overline{B}_t\right\} \leq T\Gamma + \sum_{t=1}^{T}\sum_{l=1}^{\lceil Q \rceil} 2^{-l+1}\mathbb{P}\left(\overline{B}_t\right) \leq T\Gamma + \sum_{t=1}^{T}\sum_{l=1}^{\lceil Q \rceil} 2^{-l+1}\gamma$$

$$= T\Gamma + \sum_{t=1}^{T}\frac{1}{t^2}\sum_{l=1}^{\lceil Q \rceil} 2^{-l+1} = T\Gamma + \sum_{t=1}^{T}\frac{2-\Gamma}{t^2} < T\Gamma + \frac{\pi^2}{3} = T\Gamma + O(1) .$$

With the choice of Γ as in Eqn. (5),

$$F_4 < d\sqrt{T}\log^{\frac{3}{2}} T + O(1)$$

$$\leq O\left(d\sqrt{T}\log^{\frac{3}{2}} T\right) .$$

$\square$

C.6 Proof of Theorem 1

Proof. Combining Lemmas 2, 3, 4 and 5,

$$R_T = F_1 + F_2 + F_3 + F_4$$

$$\leq O\left(d\sqrt{T}\log^{\frac{3}{2}} T\right) + O\left(d\sqrt{T}\log^{\frac{3}{2}} T\right) + O\left(d\sqrt{T}\log^{\frac{3}{2}} T\right) + O\left(d\sqrt{T}\log^{\frac{3}{2}} T\right)$$

$$= O\left(d\sqrt{T}\log^{\frac{3}{2}} T\right) .$$

$\square$

D Proof of the regret bound for LinIMED-2 (Proof of Theorem 2)

We choose γ and Γ as follows:

$$\gamma = \frac{1}{t^2} \qquad \Gamma = \frac{\sqrt{d\beta_T}\log T}{\sqrt{T}} . \tag{21}$$

D.1 Statement of Lemmas for LinIMED-2

We first state the following lemmas which respectively show the upper bound of F_1 to F_4:

Lemma 6. *Under Assumption 1, and the assumption that $\sqrt{\lambda}S \geq 1$, for the free parameter $0 < \Gamma < 1$, the term F_1 for LinIMED-3 satisfies:*

$$F_1 \leq T\Gamma + O\left(\frac{d\beta_T \log T}{\Gamma}\right) \log\left(1 + \frac{L^2\beta_T \log T}{\lambda\Gamma^2}\right) .$$

Lemma 7. *Under Assumption 1, and the assumption that $\sqrt{\lambda}S \geq 1$, for the free parameter $0 < \Gamma < 1$, the term F_2 for LinIMED-3 satisfies:*

$$F_2 \leq T\Gamma + O\left(\frac{d\beta_T \log T}{\Gamma}\right) \log\left(1 + \frac{L^2\beta_T \log T}{\lambda\Gamma^2}\right) .$$

Lemma 8. *Under Assumption 1, and the assumption that $\sqrt{\lambda}S \geq 1$, for the free parameter $0 < \Gamma < 1$, the term F_3 for LinIMED-3 satisfies:*

$$F_3 \leq 5T\Gamma + O\left(\frac{d\beta_T \log T}{\Gamma} \log\left(1 + \frac{L^2\beta_T \log T}{\lambda\Gamma^2}\right)\right) + O\left(\sqrt{T \log T} \log\left(\frac{L^2\beta_T \log T}{\lambda\Gamma^2}\right)\right) .$$

Lemma 9. *Under Assumption 1, with the choice of $\gamma = \frac{1}{t^2}$ as in Eqn. (21), for the free parameter $0 < \Gamma < 1$, the term F_4 for LinIMED-3 satisfies:*

$$F_4 \leq T\Gamma + O(1) .$$

D.2 Proof of Lemma 6

Proof. We first partition the analysis into the cases $\hat{A}_t \neq A_t$ and $\hat{A}_t = A_t$ as follows:

$$
\begin{aligned}
F_1 &= \mathbb{E}\sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\{B_t, C_t, D_t\} \\
&= \mathbb{E}\sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\{B_t, C_t, D_t\} \cdot \mathbb{1}\{\hat{A}_t \neq A_t\} + \mathbb{E}\sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\{B_t, C_t, D_t\} \cdot \mathbb{1}\{\hat{A}_t = A_t\}
\end{aligned}
$$

Case 1: If $\hat{A}_t \neq A_t$, this means that the index of A_t is $I_{t,A_t} = \frac{\hat{\Delta}_{t,A_t}^2}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2} + \log\frac{1}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2}$. Using the fact that $I_{t,A_t} \leq I_{t,\hat{A}_t}$ we have:

$$
\begin{aligned}
I_{t,A_t} &= \frac{\hat{\Delta}_{t,A_t}^2}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2} + \log\frac{1}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2} \\
&\leq \log T \wedge \log\frac{1}{\beta_{t-1}\|\hat{X}_t\|_{V_{t-1}^{-1}}^2} \\
&\leq \log T.
\end{aligned}
$$

Therefore

$$\frac{\hat{\Delta}_{t,A_t}^2}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2} + \log\frac{1}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2} \leq \log T . \tag{22}$$

If $\|X_t\|^2_{V_{t-1}^{-1}} \geq \frac{\Delta_t^2}{\beta_{t-1}}$, using the same procedure to get from Eqn. (15) to Eqn. (16), one has:

$$
\mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1} \{B_t, C_t, D_t\} \cdot \mathbb{1} \left\{ \hat{A}_t \neq A_t \right\} \cdot \mathbb{1} \left\{ \|X_t\|^2_{V_{t-1}^{-1}} \geq \frac{\Delta_t^2}{\beta_{t-1}} \right\}
$$
$$
\leq \mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1} \left\{ \|X_t\|^2_{V_{t-1}^{-1}} \geq \frac{\Delta_t^2}{\beta_{t-1}} \right\}
$$
$$
< T\Gamma + \frac{48 d \beta_T}{\Gamma} \log \left(1 + \frac{8 L^2 \beta_T}{\lambda \Gamma^2} \right)
$$
$$
= T\Gamma + O \left(\frac{d \beta_T}{\Gamma} \log \left(1 + \frac{L^2 \beta_T}{\lambda \Gamma^2} \right) \right).
$$

Else if $\|X_t\|^2_{V_{t-1}^{-1}} < \frac{\Delta_t^2}{\beta_{t-1}}$, this implies that $\log \frac{1}{\beta_{t-1} \|X_t\|^2_{V_{t-1}^{-1}}} > \log \frac{1}{\Delta_t^2} \geq 0$. Then substituting the event $D_t := \{\hat{\Delta}_{t,A_t} \geq \varepsilon\}$ into Eqn. (22), we obtain

$$
\frac{\varepsilon^2}{\beta_{t-1} \|X_t\|^2_{V_{t-1}^{-1}}} \leq \log T .
$$

With $\sqrt{\lambda} S \geq 1$ we have $\beta_{t-1} \geq 1$, then one has

$$
\|X_t\|^2_{V_{t-1}^{-1}} \geq \frac{\varepsilon^2}{\beta_{t-1} \log T}.
$$

Hence

$$
\mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1} \left\{ B_t, C_t, D_t, \hat{A}_t \neq A_t, \|X_t\|^2_{V_{t-1}^{-1}} < \frac{\Delta_t^2}{\beta_{t-1}} \right\}
$$
$$
\leq \mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1} \left\{ \|X_t\|^2_{V_{t-1}^{-1}} \geq \frac{\varepsilon^2}{\beta_{t-1} \log T} \right\}.
$$

With the choice of $\varepsilon = (1 - \frac{2}{\sqrt{\log T}})\Delta_t$, when $T \geq 149 > \exp(5)$, $\varepsilon > \frac{\Delta_t}{10}$, then performing the "peeling device" on Δ_t yields

$$
\mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1} \left\{ \|X_t\|^2_{V_{t-1}^{-1}} \geq \frac{\varepsilon^2}{\beta_{t-1} \log T} \right\} \cdot \mathbb{1} \{\Delta_t \geq \Gamma\}
$$
$$
\leq 149 + \mathbb{E} \sum_{t=1}^{T} \sum_{l=1}^{\lceil Q \rceil} \Delta_t \cdot \mathbb{1} \left\{ 2^{-l} < \Delta_t \leq 2^{-l+1}, \|X_t\|^2_{V_{t-1}^{-1}} \geq \frac{\varepsilon^2}{\beta_{t-1} \log T} \right\}
$$
$$
\leq O(1) + \mathbb{E} \sum_{l=1}^{\lceil Q \rceil} 2^{-l+1} \sum_{t=1}^{T} \mathbb{1} \left\{ \|X_t\|^2_{V_{t-1}^{-1}} \geq \frac{\varepsilon^2}{\beta_{t-1} \log T} \right\}
$$
$$
\leq O(1) + \mathbb{E} \sum_{l=1}^{\lceil Q \rceil} 2^{-l+1} \sum_{t=1}^{T} \mathbb{1} \left\{ \|X_t\|^2_{V_{t-1}^{-1}} \geq \frac{2^{-2l}}{100 \beta_T \log T} \right\}
$$
$$
\leq O(1) + \mathbb{E} \sum_{l=1}^{\lceil Q \rceil} 2^{-l+1} \cdot 2^{2l} \cdot 600 d \beta_T (\log T) \log \left(1 + \frac{2^{2l} \cdot 200 L^2 \beta_T \log T}{\lambda} \right)
$$

$$\leq O(1) + \mathbb{E} \sum_{l=1}^{\lceil Q \rceil} 2^{l+1} \cdot 600 d \beta_T (\log T) \log \left(1 + \frac{2^{2 \lceil Q \rceil} \cdot 200 L^2 \beta_T \log T}{\lambda} \right)$$

$$< O(1) + \frac{4800 d \beta_T \log T}{\Gamma} \log \left(1 + \frac{800 L^2 \beta_T \log T}{\lambda \Gamma^2} \right).$$

Considering the event $\{\Delta_t < \Gamma\}$, we can upper bound the corresponding expectation as follows

$$\mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1} \left\{ \|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{\varepsilon^2}{\beta_{t-1} \log T} \right\} \cdot \mathbb{1} \{\Delta_t < \Gamma\} \leq \mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1} \{\Delta_t < \Gamma\} < T\Gamma.$$

Then

$$\mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1} \left\{ B_t, C_t, D_t, \hat{A}_t \neq A_t, \|X_t\|_{V_{t-1}^{-1}}^2 < \frac{\Delta_t^2}{\beta_{t-1}} \right\}$$

$$\leq \mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1} \left\{ \|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{\varepsilon^2}{\beta_{t-1} \log T} \right\}$$

$$= \mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1} \left\{ \|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{\varepsilon^2}{\beta_{t-1} \log T} \right\} \cdot \mathbb{1} \{\Delta_t \geq \Gamma\}$$

$$+ \mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1} \left\{ \|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{\varepsilon^2}{\beta_{t-1} \log T} \right\} \cdot \mathbb{1} \{\Delta_t < \Gamma\}$$

$$\leq O(1) + T\Gamma + \frac{4800 d \beta_T \log T}{\Gamma} \log \left(1 + \frac{800 L^2 \beta_T \log T}{\lambda \Gamma^2} \right).$$

Hence

$$\mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1} \left\{ B_t, C_t, D_t, \hat{A}_t \neq A_t \right\}$$

$$= \mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1} \left\{ B_t, C_t, D_t, \hat{A}_t \neq A_t, \|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{\Delta_t^2}{\beta_{t-1}} \right\}$$

$$+ \mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1} \left\{ B_t, C_t, D_t, \hat{A}_t \neq A_t, \|X_t\|_{V_{t-1}^{-1}}^2 < \frac{\Delta_t^2}{\beta_{t-1}} \right\}$$

$$\leq T\Gamma + O\left(\frac{d \beta_T}{\Gamma} \log \left(1 + \frac{L^2 \beta_T}{\lambda \Gamma^2} \right) \right) + O(1) + T\Gamma + \frac{4800 d \beta_T \log T}{\Gamma} \log \left(1 + \frac{800 L^2 \beta_T \log T}{\lambda \Gamma^2} \right)$$

$$\leq T\Gamma + O\left(\frac{d \beta_T \log T}{\Gamma} \log \left(1 + \frac{L^2 \beta_T \log T}{\lambda \Gamma^2} \right) \right).$$

Case 2: If $\hat{A}_t = A_t$, then from the event C_t and the choice $\delta = \frac{\Delta_t}{\sqrt{\log T}}$ we have

$$\langle \hat{\theta}_{t-1} - \theta^*, X_t \rangle > \left(1 - \frac{1}{\sqrt{\log T}} \right) \Delta_t.$$

Furthermore, using the definition of the event B_t, that implies that

$$\|X_t\|_{V_{t-1}^{-1}}^2 > \frac{(1 - \frac{1}{\sqrt{\log T}})^2 \Delta_t^2}{\beta_{t-1}}.$$

When $T > 8 > \exp(2)$, $(1 - \frac{1}{\sqrt{\log T}})^2 > \frac{1}{16}$, then similarily, we can bound this term by $O(\frac{d \beta_T}{\Gamma}) \log(1 + \frac{L^2 \beta_T}{\lambda \Gamma^2})$

Summarizing the two cases,

$$F_1 \leq O(1) + T\Gamma + O\left(\frac{d\beta_T \log T}{\Gamma}\right) \log\left(1 + \frac{L^2\beta_T \log T}{\lambda\Gamma^2}\right)$$
$$\leq T\Gamma + O\left(\frac{d\beta_T \log T}{\Gamma}\right) \log\left(1 + \frac{L^2\beta_T \log T}{\lambda\Gamma^2}\right).$$

$\square$

D.3 Proof of Lemma 7

Proof. Recall that

$$F_2 = \mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{B_t, C_t, \overline{D}_t\right\}.$$

From C_t and $\overline{D}_t$, we derive that:

$$\langle \theta^*, a_t^* \rangle - \delta < \varepsilon + \langle \hat{\theta}_{t-1}, X_t \rangle.$$

With the choice $\delta = \frac{\Delta_t}{\sqrt{\log T}}, \varepsilon = (1 - \frac{2}{\sqrt{\log T}})\Delta_t$, we have

$$\langle \hat{\theta}_{t-1} - \theta^*, X_t \rangle > \frac{\Delta_t}{\sqrt{\log T}}. \tag{23}$$

Then using the definition of the event B_t in Eqn. (23) yields

$$\|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{\Delta_t^2}{\beta_{t-1} \log T}.$$

Using a similar procedure as in that from Eqn. (15) to Eqn. (16), we can upper bound F_2 by

$$F_2 \leq T\Gamma + O\left(\frac{d\beta_T \log T}{\Gamma}\right) \log\left(1 + \frac{L^2\beta_T \log T}{\lambda\Gamma^2}\right).$$

$\square$

D.4 Proof of Lemma 8

Proof. From the event $\overline{C}_t$, which is $\max_{b \in \mathcal{A}_t} \langle \hat{\theta}_{t-1}, b \rangle \leq \langle \theta^*, x_t^* \rangle - \delta$, the index of the best arm at time t can be upper bounded as:

$$I_{t,a_t^*} \leq \frac{(\langle \theta^*, x_t^* \rangle - \delta - \langle \hat{\theta}_{t-1}, x_t^* \rangle)^2}{\beta_{t-1}\|x_t^*\|_{V_{t-1}^{-1}}^2} + \log \frac{1}{\beta_{t-1}\|x_t^*\|_{V_{t-1}^{-1}}^2}.$$

Case 1: If $\hat{A}_t \neq A_t$, then we have

$$I_{t,a_t^*} \geq I_{t,A_t} \geq \log \frac{1}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2}.$$

Suppose $\frac{q+1}{2}\delta \leq \langle \theta^*, x_t^* \rangle - \langle \hat{\theta}_{t-1}, x_t^* \rangle < \frac{q+2}{2}\delta$ for $q \in \mathbb{N}$, then one has

$$\log \frac{1}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2} \leq \frac{q^2\delta^2}{4\beta_{t-1}\|x_t^*\|_{V_{t-1}^{-1}}^2} + \log \frac{1}{\beta_{t-1}\|x_t^*\|_{V_{t-1}^{-1}}^2}. \tag{24}$$

On the other hand, on the event B_t,

$$\|x_t^*\|_{V_{t-1}^{-1}} \geq \frac{(q+1)\delta}{2\sqrt{\beta_{t-1}}}. \tag{25}$$

If $\|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{\Delta_t^2}{\beta_{t-1}}$, using the same procedure from Eqn. (15) to Eqn. (16), one has:

$$\mathbb{E}\sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{B_t, \overline{C}_t\right\} \cdot \mathbb{1}\left\{\hat{A}_t \neq A_t\right\} \cdot \mathbb{1}\left\{\|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{\Delta_t^2}{\beta_{t-1}}\right\}$$

$$\leq \mathbb{E}\sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{\|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{\Delta_t^2}{\beta_{t-1}}\right\}$$

$$< T\Gamma + \frac{48d\beta_T}{\Gamma} \log\left(1 + \frac{8L^2\beta_T}{\lambda\Gamma^2}\right)$$

$$= T\Gamma + O\left(\frac{d\beta_T}{\Gamma} \log\left(1 + \frac{L^2\beta_T}{\lambda\Gamma^2}\right)\right).$$

Else if $\|X_t\|_{V_{t-1}^{-1}}^2 < \frac{\Delta_t^2}{\beta_{t-1}}$, this implies that $\log\frac{1}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2} > \log\frac{1}{\Delta_t^2} \geq 0$. Then combining Eqn. (24) and Eqn. (25) implies that

$$\|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{(q+1)^2\delta^2}{4\beta_{t-1}} \exp\left(-\frac{q^2}{(q+1)^2}\right).$$

Then using the same procedure to get from Eqn. (19) to Eqn. (20), we have

$$\sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{B_t, \overline{C}_t\right\} \cdot \mathbb{1}\left\{\|X_t\|_{V_{t-1}^{-1}}^2 < \frac{\Delta_t^2}{\beta_{t-1}}, \hat{A}_t \neq A_t\right\}$$

$$< T\Gamma + O\left(\frac{d\beta_T \log T}{\Gamma} \log\left(1 + \frac{L^2\beta_T \log T}{\lambda\Gamma^2}\right)\right). \tag{26}$$

Case 2: $\hat{A}_t = A_t$. If $\|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{\Delta_t^2}{\beta_{t-1}}$, using the same procedure to get from Eqn. (15) to Eqn. (16), one has:

$$\mathbb{E}\sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{B_t, \overline{C}_t\right\} \cdot \mathbb{1}\left\{\hat{A}_t = A_t\right\} \cdot \mathbb{1}\left\{\|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{\Delta_t^2}{\beta_{t-1}}\right\}$$

$$\leq \mathbb{E}\sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{\|X_t\|_{V_{t-1}^{-1}}^2 \geq \frac{\Delta_t^2}{\beta_{t-1}}\right\}$$

$$< T\Gamma + \frac{48d\beta_T}{\Gamma} \log\left(1 + \frac{8L^2\beta_T}{\lambda\Gamma^2}\right)$$

$$= T\Gamma + O\left(\frac{d\beta_T}{\Gamma} \log\left(1 + \frac{L^2\beta_T}{\lambda\Gamma^2}\right)\right).$$

Else $\|X_t\|_{V_{t-1}^{-1}}^2 < \frac{\Delta_t^2}{\beta_{t-1}}$ implies that $\log\frac{1}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2} > \log\frac{1}{\Delta_t^2} \geq 0$.

If $\log \frac{1}{\beta_{t-1}\|X_t\|^2_{V_{t-1}^{-1}}} < \log T$, then using the same procedure to get from Eqn. (24) to Eqn. (26), we have

$$\sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{B_t, \overline{C}_t\right\} \cdot \mathbb{1}\left\{\|X_t\|^2_{V_{t-1}^{-1}} < \frac{\Delta_t^2}{\beta_{t-1}}, \hat{A}_t = A_t, \log \frac{1}{\beta_{t-1}\|X_t\|^2_{V_{t-1}^{-1}}} < \log \frac{T}{\beta_{t-1}}\right\}$$
$$< T\Gamma + O\left(\frac{d\beta_T \log T}{\Gamma} \log\left(1 + \frac{L^2 \beta_T \log T}{\lambda \Gamma^2}\right)\right).$$

If $\log \frac{1}{\beta_{t-1}\|X_t\|^2_{V_{t-1}^{-1}}} \geq \log T$, this means now the index of A_t is $I_{t,A_t} = \log T$, by performing the "peeling device" such that $\frac{q+1}{2}\delta \leq \langle \theta^*, x_t^* \rangle - \langle \hat{\theta}_{t-1}, x_t^* \rangle < \frac{q+2}{2}\delta$ for $q \in \mathbb{N}$, we have

$$\log T \leq \frac{q^2 \delta^2}{4\beta_{t-1}\|x_t^*\|^2_{V_{t-1}^{-1}}} + \log \frac{1}{\beta_{t-1}\|x_t^*\|^2_{V_{t-1}^{-1}}}. \tag{27}$$

On the other hand, using the definition of the event B_t,

$$\|x_t^*\|_{V_{t-1}^{-1}} \geq \frac{(q+1)\delta}{2\sqrt{\beta_{t-1}}}. \tag{28}$$

Combining Eqn. (27) and (28), we have

$$\delta \leq \frac{2\exp(\frac{q^2}{2(q+1)^2})}{(q+1)\sqrt{T}}.$$

Then with $\delta = \frac{\Delta_t}{\sqrt{\log T}}$, this implies that

$$\Delta_t \leq \frac{2\sqrt{\log T}\exp(\frac{q^2}{2(q+1)^2})}{(q+1)\sqrt{T}}.$$

On the other hand, from $\frac{q+1}{2}\delta \leq \sqrt{\beta_{t-1}}\|x_t^*\|_{V_{t-1}^{-1}} \leq \sqrt{\beta_{t-1}} \cdot \frac{L}{\sqrt{\lambda}}$, we have $q + 1 \leq \frac{2L\sqrt{\beta_{t-1}\log T}}{\sqrt{\lambda}\Delta_t}$. Hence,

$$\sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{B_t, \overline{C}_t\right\} \cdot \mathbb{1}\left\{\|X_t\|^2_{V_{t-1}^{-1}} < \frac{\Delta_t^2}{\beta_{t-1}}, \hat{A}_t = A_t, \log \frac{1}{\beta_{t-1}\|X_t\|^2_{V_{t-1}^{-1}}} \geq \log T, \Delta_t \geq \Gamma\right\}$$

$$\leq \mathbb{E}\sum_{q=1}^{\lfloor \frac{2L\sqrt{\beta_T \log T}}{\sqrt{\lambda}\Gamma} - 1\rfloor} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{\Delta_t \leq \frac{2\sqrt{\log T}\exp(\frac{q^2}{2(q+1)^2})}{(q+1)\sqrt{T}}\right\}$$

$$\leq \mathbb{E}\sum_{q=1}^{\lfloor \frac{2L\sqrt{\beta_T \log T}}{\sqrt{\lambda}\Gamma} - 1\rfloor} \sum_{t=1}^{T} \frac{2\sqrt{\log T}\exp(\frac{q^2}{2(q+1)^2})}{(q+1)\sqrt{T}}$$

$$= \mathbb{E}\sum_{q=1}^{\lfloor \frac{2L\sqrt{\beta_T \log T}}{\sqrt{\lambda}\Gamma} - 1\rfloor} \frac{2\sqrt{T\log T}\exp(\frac{q^2}{2(q+1)^2})}{q+1}$$

$$< \mathbb{E}\sum_{q=1}^{\lfloor \frac{2L\sqrt{\beta_T \log T}}{\sqrt{\lambda}\Gamma} - 1\rfloor} \frac{2\sqrt{e}\sqrt{T\log T}}{q+1}$$

$$< 2\sqrt{e}\sqrt{T \log T} \log \left(\frac{2L\sqrt{\log T}}{\sqrt{\lambda}\Gamma} - 1 \right)$$

$$\leq O\left(\sqrt{T \log T} \log \left(\frac{L^2 \beta_T \log T}{\lambda \Gamma^2} \right) \right) .$$

Summarizing the two cases ($\hat{A}_t \neq A_t$ and $\hat{A}_t = A_t$), we see that F_3 is upper bounded by:

$$F_3 < T\Gamma + O\left(\frac{d\beta_T}{\Gamma} \log \left(1 + \frac{L^2 \beta_T}{\lambda \Gamma^2} \right) \right) + T\Gamma + O\left(\frac{d\beta_T \log T}{\Gamma} \log \left(1 + \frac{L^2 \beta_T \log T}{\lambda \Gamma^2} \right) \right)$$

$$+ T\Gamma + O\left(\frac{d\beta_T}{\Gamma} \log \left(1 + \frac{L^2 \beta_T}{\lambda \Gamma^2} \right) \right) + T\Gamma + O\left(\frac{d\beta_T \log T}{\Gamma} \log \left(1 + \frac{L^2 \beta_T \log T}{\lambda \Gamma^2} \right) \right)$$

$$+ T\Gamma + O\left(\sqrt{T\beta_T \log T} \log \left(\frac{L^2 \beta_T \log T}{\lambda \Gamma^2} \right) \right)$$

$$\leq 5T\Gamma + O\left(\frac{d\beta_T \log T}{\Gamma} \log \left(1 + \frac{L^2 \beta_T \log T}{\lambda \Gamma^2} \right) \right) + O\left(\sqrt{T \log T} \log \left(\frac{L^2 \beta_T \log T}{\lambda \Gamma^2} \right) \right) .$$

$\square$

D.5 Proof of Lemma 9

Proof. The proof of this case is straightforward by using Lemma 1 with the choice $\gamma = \frac{1}{t^2}$:

$$F_4 = \mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{ \overline{B}_t \right\}$$

$$= \mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{ \overline{B}_t, \Delta_t < \Gamma \right\} + \mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{ \overline{B}_t, \Delta_t \geq \Gamma \right\}$$

$$< T\Gamma + \mathbb{E} \sum_{t=1}^{T} \sum_{l=1}^{\lceil Q \rceil} \Delta_t \cdot \mathbb{1}\left\{ \overline{B}_t, 2^{-l} < \Delta_t \leq 2^{-l+1} \right\}$$

$$\leq T\Gamma + \mathbb{E} \sum_{t=1}^{T} \sum_{l=1}^{\lceil Q \rceil} 2^{-l+1} \cdot \mathbb{1}\left\{ \overline{B}_t \right\}$$

$$\leq T\Gamma + \sum_{l=1}^{\lceil Q \rceil} 2^{-l+1} \sum_{t=1}^{T} \mathbb{P}\left\{ \overline{B}_t \right\}$$

$$= T\Gamma + \sum_{l=1}^{\lceil Q \rceil} 2^{-l+1} \cdot \frac{\pi^2}{6}$$

$$< T\Gamma + (2 - \Gamma) \cdot \frac{\pi^2}{6}$$

$$< T\Gamma + \frac{\pi^2}{3}$$

$$= T\Gamma + O(1) .$$

$\square$

D.6 Proof of Theorem 2

Proof. Combining Lemmas 6, 7, 8 and 9, with the choices of γ and Γ as in Eqn. (21), the regret of LinIMED-2 is bounded as follows:

$$R_T = F_1 + F_2 + F_3 + F_4$$

$$\leq T\Gamma + O\left(\frac{d\beta_T \log T}{\Gamma}\right) \log\left(1 + \frac{L^2\beta_T \log T}{\lambda\Gamma^2}\right) + T\Gamma + O\left(\frac{d\beta_T \log T}{\Gamma}\right) \log\left(1 + \frac{L^2\beta_T \log T}{\lambda\Gamma^2}\right)$$

$$+ 5T\Gamma + O\left(\frac{d\beta_T \log T}{\Gamma} \log\left(1 + \frac{L^2\beta_T \log T}{\lambda\Gamma^2}\right)\right) + O\left(\sqrt{T\log T}\log\left(\frac{L^2\beta_T \log T}{\lambda\Gamma^2}\right)\right)$$

$$+ T\Gamma + O(1)$$

$$\leq 8T\Gamma + O\left(\frac{d\beta_T \log T}{\Gamma} \log\left(1 + \frac{L^2\beta_T \log T}{\lambda\Gamma^2}\right)\right) + O\left(\sqrt{T\log T}\log\left(\frac{L^2\beta_T \log T}{\lambda\Gamma^2}\right)\right)$$

$$= 8\sqrt{dT\beta_T}\log T + O\left(\sqrt{dT\beta_T}\log\left(1 + \frac{TL^2}{\lambda d \log T}\right)\right) + O\left(\sqrt{T\log T}\log\left(\frac{TL^2}{\lambda d \log T}\right)\right)$$

$$= 8d\sqrt{T}\log^{\frac{3}{2}} T + O\left(d\sqrt{T}\log^{\frac{3}{2}} T\right) + O\left(\sqrt{T}\log^{\frac{3}{2}} T\right)$$

$$\leq O\left(d\sqrt{T}\log^{\frac{3}{2}} T\right).$$

$\square$

E Proof of the regret bound for LinIMED-3 (Proof of Theorem 3)

First we define a_t^* as the best arm in time step t such that $a_t^* = \arg\max_{a\in\mathcal{A}_t}\langle\theta^*, x_{t,a}\rangle$, and use $x_t^* := x_{t,a_t^*}$ denote its corresponding context. Define $\hat{A}_t := \arg\max_{a\in\mathcal{A}_t} \mathrm{UCB}_t(a)$. Let $\Delta_t := \langle\theta^*, x_t^*\rangle - \langle\theta^*, X_t\rangle$ denote the regret in time t. Define the following events:

$$B_t' := \left\{\|\hat{\theta}_{t-1} - \theta^*\|_{V_{t-1}} \leq \sqrt{\beta_{t-1}(\gamma)}\right\}, \quad D_t' := \{\hat{\Delta}_{t,A_t} > \varepsilon\}.$$

where ε is a free parameter set to be $\varepsilon = \frac{\Delta_t}{3}$ in this proof sketch.

Then the expected regret $R_T = \mathbb{E}\sum_{t=1}^T \Delta_t$ can be partitioned by events B_t', D_t' such that:

$$R_T = \underbrace{\mathbb{E}\sum_{t=1}^T \Delta_t \cdot \mathbb{1}\left\{B_t', D_t'\right\}}_{=:F_1} + \underbrace{\mathbb{E}\sum_{t=1}^T \Delta_t \cdot \mathbb{1}\left\{B_t, \overline{D_t'}\right\}}_{=:F_2} + \underbrace{\mathbb{E}\sum_{t=1}^T \Delta_t \cdot \mathbb{1}\left\{\overline{B_t'}\right\}}_{=:F_3}.$$

For the F_1 case:
From D_t' we know $A_t \neq \hat{A}_t$, therefore

$$I_{t,A_t} = \frac{\hat{\Delta}_{t,A_t}^2}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2} + \log\frac{1}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2}. \tag{29}$$

From D_t' and $I_{t,A_t} \leq I_{t,\hat{A}_t} \leq \log\frac{C}{\max_{a\in\mathcal{A}_t}\hat{\Delta}_{t,a}^2}$, we have

$$I_{t,A_t} < \log\frac{C}{\varepsilon^2}. \tag{30}$$

Combining Eqn. (29) and Eqn. (30),

$$\frac{\hat{\Delta}_{t,A_t}^2}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2} + \log\frac{1}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2} < \log\frac{C}{\varepsilon^2}.$$

Then

$$\frac{\hat{\Delta}_{t,A_t}^2}{\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2} < \log\left(\beta_{t-1}\|X_t\|_{V_{t-1}^{-1}}^2 \cdot \frac{C}{\varepsilon^2}\right). \tag{31}$$

If $\|X_t\|^2_{V^{-1}_{t-1}} \geq \frac{\Delta_t^2}{\beta_{t-1}}$, using the same procedure from Eqn. (15) to Eqn. (16), one has:

$$\mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\{B'_t, D'_t\} \cdot \mathbb{1}\left\{\|X_t\|^2_{V^{-1}_{t-1}} \geq \frac{\Delta_t^2}{\beta_{t-1}}\right\}$$

$$\leq \mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{\|X_t\|^2_{V^{-1}_{t-1}} \geq \frac{\Delta_t^2}{\beta_{t-1}}\right\}$$

$$< T\Gamma + \frac{48 d \beta_T}{\Gamma} \log\left(1 + \frac{8 L^2 \beta_T}{\lambda \Gamma^2}\right)$$

$$= T\Gamma + O\left(\frac{d \beta_T}{\Gamma} \log\left(1 + \frac{L^2 \beta_T}{\lambda \Gamma^2}\right)\right).$$

Else $\|X_t\|^2_{V^{-1}_{t-1}} < \frac{\Delta_t^2}{\beta_{t-1}}$,this implies that $\beta_{t-1}\|X_t\|^2_{V^{-1}_{t-1}} < \Delta_t^2$, plug this into Eqn. (31) and with the choice of $\varepsilon = \frac{\Delta_t}{3}$ and D'_t, we have

$$\frac{\Delta_t^2}{9 \beta_{t-1} \|X_t\|^2_{V^{-1}_{t-1}}} < \log(9C) .$$

Since $C \geq 1$ is a constant, then

$$\|X_t\|^2_{V^{-1}_{t-1}} > \frac{\Delta_t^2}{9 \beta_{t-1} \log(9C)} .$$

Using the same procedure from Eqn. (15) to Eqn. (16), one has:

$$\mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\{B'_t, D'_t\} \cdot \mathbb{1}\left\{\|X_t\|^2_{V^{-1}_{t-1}} < \frac{\Delta_t^2}{\beta_{t-1}}\right\}$$

$$\leq \mathbb{E} \sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{\|X_t\|^2_{V^{-1}_{t-1}} > \frac{\Delta_t^2}{9 \beta_{t-1} \log(9C)}\right\}$$

$$< T\Gamma + O\left(\frac{d \beta_T \log C}{\Gamma} \log\left(1 + \frac{L^2 \beta_T \log C}{\lambda \Gamma^2}\right)\right) .$$

Hence

$$F_1 < 2T\Gamma + O\left(\frac{d \beta_T \log C}{\Gamma} \log\left(1 + \frac{L^2 \beta_T \log C}{\lambda \Gamma^2}\right)\right) . \tag{32}$$

For the F_2 case: Since the event B'_t holds,

$$\max_{a \in \mathcal{A}_t} \mathrm{UCB}_t(a) \geq \mathrm{UCB}_t(a^*_t) = \langle \hat{\theta}_{t-1}, x^*_t \rangle + \sqrt{\beta_{t-1}} \|x^*_t\|_{V^{-1}_{t-1}} \geq \langle \theta^*, x^*_t \rangle \tag{33}$$

On the other hand, from $\overline{D'_t}$ we have

$$\max_{a \in \mathcal{A}_t} \mathrm{UCB}_t(a) \leq \mathrm{UCB}_t(A_t) + \varepsilon = \langle \hat{\theta}_{t-1}, X_t \rangle + \sqrt{\beta_{t-1}} \|X_t\|_{V^{-1}_{t-1}} + \varepsilon . \tag{34}$$

Combining Eqn. (33) and Eqn. (34),

$$\langle \theta^*, x^*_t \rangle \leq \langle \hat{\theta}_{t-1}, X_t \rangle + \sqrt{\beta_{t-1}} \|X_t\|_{V^{-1}_{t-1}} + \varepsilon .$$

Hence

$$\Delta_t - \varepsilon \leq \langle \hat{\theta}_{t-1} - \theta^*, X_t \rangle + \sqrt{\beta_{t-1}} \|X_t\|_{V^{-1}_{t-1}} .$$

Then with $\varepsilon = \frac{\Delta_t}{3}$ and B'_t, we have

$$\frac{2}{3}\Delta_t \le 2\sqrt{\beta_{t-1}}\|X_t\|_{V_{t-1}^{-1}} \ ,$$

therefore

$$\|X_t\|_{V_{t-1}^{-1}}^2 > \frac{\Delta_t^2}{9\beta_{t-1}} \ .$$

Using the same procedure from Eqn. (15) to Eqn. (16), one has:

$$F_2 < T\Gamma + O\left(\frac{d\beta_T}{\Gamma}\log\left(1 + \frac{L^2\beta_T}{\lambda\Gamma^2}\right)\right) \ . \tag{35}$$

For the F_3 case:
using Lemma 1 with the choice $\gamma = \frac{1}{t^2}$:

$$
\begin{aligned}
F_3 &= \mathbb{E}\sum_{t=1}^T \Delta_t \cdot \mathbb{1}\left\{\overline{B'_t}\right\} \\
&= \mathbb{E}\sum_{t=1}^T \Delta_t \cdot \mathbb{1}\left\{\overline{B'_t}, \Delta_t < \Gamma\right\} + \mathbb{E}\sum_{t=1}^T \Delta_t \cdot \mathbb{1}\left\{\overline{B'_t}, \Delta_t \ge \Gamma\right\} \\
&< T\Gamma + \mathbb{E}\sum_{t=1}^T \sum_{l=1}^{\lceil Q\rceil} \Delta_t \cdot \mathbb{1}\left\{\overline{B'_t}, 2^{-l} < \Delta_t \le 2^{-l+1}\right\} \\
&\le T\Gamma + \mathbb{E}\sum_{t=1}^T \sum_{l=1}^{\lceil Q\rceil} 2^{-l+1} \cdot \mathbb{1}\left\{\overline{B'_t}\right\} \\
&\le T\Gamma + \sum_{l=1}^{\lceil Q\rceil} 2^{-l+1} \sum_{t=1}^T \mathbb{P}\left\{\overline{B'_t}\right\} \\
&= T\Gamma + \sum_{l=1}^{\lceil Q\rceil} 2^{-l+1} \cdot \frac{\pi^2}{6} \\
&< T\Gamma + (2 - \Gamma) \cdot \frac{\pi^2}{6} \\
&< T\Gamma + \frac{\pi^2}{3} \\
&= T\Gamma + O(1) \ . \tag{36}
\end{aligned}
$$

E.1 Proof of Theorem 3

Proof. Combining Eqn. (32), (35), (36) with the choices of $\gamma = \frac{1}{t^2}$ and $\Gamma = \frac{\beta_T}{\sqrt{T}}$ and $C \ge 1$ is a constant, the regret of LinIMED-3 is bounded as follows:

$$
\begin{aligned}
R_T &= F_1 + F_2 + F_3 + F_4 \\
&< 4T\Gamma + O\left(\frac{d\beta_T\log C}{\Gamma}\log\left(1 + \frac{L^2\beta_T\log C}{\lambda\Gamma^2}\right)\right) + O\left(\frac{d\beta_T}{\Gamma}\log\left(1 + \frac{L^2\beta_T}{\lambda\Gamma^2}\right)\right) + O(1) \\
&< O\left(d\sqrt{T}\log C\log\left(1 + \frac{L^2T\log C}{\lambda}\right)\right) \\
&= O\left(d\sqrt{T}\log(T)\right) \ .
\end{aligned}
$$

This completes the proof. $\qquad\square$

F Proof of the regret bound for SupLinIMED (Proof of Theorem 4)

Define $s_t \in [\lceil \log T \rceil]$ as the index of s when the arm is chosen at time t. For the SupLinIMED, the index of arms except the empirically best arm is defined by $I_{t,a} = \left(\frac{\hat{\Delta}_{t,a}^{s_t}}{w_{t,a}^{s_t}} \right)^2 - 2 \log(w_{t,a}^{s_t})$, whereas the index of the empirically best arm is defined by $I_{t,\hat{A}_t^*} = \log(2T) \wedge (-2\log(w_{t,\hat{A}_t^*}^{s_t}))$ where $\hat{A}_t^* = \arg\max_{a \in \hat{\mathcal{A}}_{s_t}} \langle \hat{\theta}_t^{s_t}, x_{t,a} \rangle$. Define the index of the best arm at time t as $a_t^* := \arg\max_{a \in [K]} \langle \theta^*, x_{t,a} \rangle$.

Remark 1. *Here the upper bound we set for the index of the empirically best arm is $\log(2T)$, which is slightly larger than our previous $\log T$ (Line 10 in the LinIMED algorithm) since in the first step of the of the SupLinIMED algorithm or, more generally, the SupLinUCB-type algorithms, the width of each arm is less than $\frac{1}{\sqrt{T}}$, as a result, the index of each arm is larger than $\log T$.*

Let the set of time indices such that the chosen arm is from Step 1 (Lines 6–9 in Algorithm 2) be Ψ_0. Then the cumulative expected regret of the SupLinIMED algorithm over time horizon T can be defined by the following equation:

$$R_T = \mathbb{E}\left[\sum_{t \in \Psi_0} \langle \theta^*, x_{t,a_t^*} - X_t \rangle \right] + \mathbb{E}\left[\sum_{t \notin \Psi_0} \langle \theta^*, x_{t,a_t^*} - X_t \rangle \right] \tag{37}$$

Since the index set has not changed in Step 1 (see Line 9 in Algorithm 2), the second term of the regret is the same as in the original SupLinUCB algorithm of Chu et al. (2011). For the first term, we partitioned it by the following events:

$$B_t := \bigcap_{t \in [T], s \in [\log T], a \in [K]} \left\{ |\langle \theta^* - \hat{\theta}_t^s, x_{t,a} \rangle| \le \frac{\alpha+1}{\alpha} w_{t,a}^s \right\}, \quad \text{and}$$

$$D_t := \left\{ \hat{\Delta}_{t,A_t}^{s_t} \ge \varepsilon \right\},$$

where $\alpha = \sqrt{\frac{1}{2} \ln \frac{2TK}{\gamma}}$ as in the SupLinUCB (Chu et al., 2011). We choose $\gamma = \frac{1}{2t^2}$ throughout. Furthermore, $\hat{\theta}_t^s$ is the $\hat{\theta}_t$ obtained from Algorithm 3 with Ψ_t^s as the input, i.e.,

$$\hat{\theta}_t^s := \left(I_d + \sum_{\tau \in \Psi_t^s} x_{\tau,A_\tau} x_{\tau,A_\tau}^T \right)^{-1} \sum_{\tau \in \Psi_t^s} Y_{\tau,A_\tau} x_{\tau,A_\tau}.$$

Define $\Delta_t := \langle \theta^*, x_{t,a_t^*} - X_t \rangle$ as the instantaneous regret at each time step t. In addition, choose $\varepsilon = \frac{\Delta_t}{3}$ in the definition of D_t. Then the first term of the expected regret in (37) can be partitioned by the events B_t and D_t as follows:

$$\mathbb{E}\left[\sum_{t \in \Psi_0} \langle \theta^*, x_{t,a_t^*} - X_t \rangle \right] = \underbrace{\mathbb{E}\left[\sum_{t \in \Psi_0} \Delta_t \cdot \mathbb{1}\{B_t, D_t\} \right]}_{=:F_1} + \underbrace{\mathbb{E}\left[\sum_{t \in \Psi_0} \Delta_t \cdot \mathbb{1}\{B_t, \overline{D}_t\} \right]}_{=:F_2} + \underbrace{\mathbb{E}\left[\sum_{t \in \Psi_0} \Delta_t \cdot \mathbb{1}\{\overline{B}_t\} \right]}_{=:F_3}$$

We recall that when $t \in \Psi_0$, $w_{t,a}^{s_t} \le \frac{1}{\sqrt{T}}$ for all $a \in \hat{\mathcal{A}}_{s_t}$.

To bound F_1, we note that since B_t occurs, the actual best arm $a_t^* \in \hat{\mathcal{A}}_{s_t}$ with high probability $(1 - \gamma \log^2 T)$ by Chu et al. (2011, Lemma 5). As such,

$$\max_{a \in \hat{\mathcal{A}}_{s_t}} \langle \hat{\theta}_t^{s_t}, x_{t,a} \rangle \ge \langle \hat{\theta}_t^{s_t}, x_{t,a_t^*} \rangle \ge \langle \theta^*, x_{t,a_t^*} \rangle - \frac{\alpha+1}{\alpha} w_{t,a_t^*}^s \ge \langle \theta^*, x_{t,a_t^*} \rangle - \frac{2}{\sqrt{T}}$$

where the last inequality is from the fact that $\gamma = \frac{1}{2t^2}$ and $\alpha = \sqrt{\frac{1}{2}\ln\frac{2TK}{\gamma}} \geq 1$. Else if, in fact, the best arm $a_t^* \notin \hat{\mathcal{A}}_{s_t}$, the corresponding regret in this case is bounded by:

$$
\mathbb{E}\sum_{t \in \Psi_0} \Delta_t \cdot \mathbb{1}\left\{a_t^* \notin \hat{\mathcal{A}}_{s_t}\right\} \leq \mathbb{E}\sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{a_t^* \notin \hat{\mathcal{A}}_{s_t}\right\}
$$

$$
\leq \mathbb{E}\sum_{t=1}^{T} \mathbb{1}\left\{a_t^* \notin \hat{\mathcal{A}}_{s_t}\right\}
$$

$$
= \sum_{t=1}^{T} \mathbb{P}(a_t^* \notin \hat{\mathcal{A}}_{s_t})
$$

$$
\leq \sum_{t=1}^{T} \gamma \log^2 T
$$

$$
= \sum_{t=1}^{T} \frac{\log^2 T}{2t^2}
$$

$$
< \frac{\pi^2}{12}\log^2 T .
$$

Case 1: If $\hat{A}_t^* \neq A_t$, this means that the index of A_t is $I_{t,A_t} = \frac{(\hat{\Delta}_{t,A_t}^{s_t})^2}{\alpha^2\|X_t\|_{V_t^{-1}}^2} + \log\frac{1}{\alpha^2\|X_t\|_{V_t^{-1}}^2}$. Using the fact that $I_{t,A_t} \leq I_{t,\hat{A}_t^*}$ we have

$$
\frac{(\hat{\Delta}_{t,A_t}^{s_t})^2}{\alpha^2\|X_t\|_{V_t^{-1}}^2} + \log\frac{1}{\alpha^2\|X_t\|_{V_t^{-1}}^2} \leq 2\log T .
$$

Then using the definition of the event D_t and the fact that $(w_{t,a}^{s_t})^2 = \alpha^2\|X_t\|_{V_t^{-1}}^2 \leq \frac{1}{T}$, we have

$$
\Delta_t^2 \leq 9\alpha^2\|X_t\|_{V_{t-1}^{-1}}^2 \log T \leq \frac{9\log T}{T}.
$$

Hence, $\Delta_t \leq \frac{3\sqrt{\log T}}{\sqrt{T}}$. Therefore F_1 in this case is upper bounded as follows:

$$
\mathbb{E}\left[\sum_{t\in\Psi_0} \Delta_t \cdot \mathbb{1}\left\{B_t, D_t\right\} \cdot \mathbb{1}\left\{\hat{A}_t^* \neq A_t\right\} \cdot \mathbb{1}\left\{a_t^* \in \hat{\mathcal{A}}_{s_t}\right\}\right] \leq \mathbb{E}\left[\sum_{t\in\Psi_0} \frac{3\sqrt{\log T}}{\sqrt{T}}\right] \leq 3\sqrt{T\log T} .
$$

Case 2: If $\hat{A}_t^* = A_t$, then using the definition of the event B_t, we have

$$
\langle\hat{\theta}_t^{s_t}, X_t\rangle = \max_{a\in\hat{\mathcal{A}}_{s_t}}\langle\hat{\theta}_t^{s_t}, x_{t,a}\rangle \geq \langle\hat{\theta}_t^{s_t}, x_{t,a_t^*}\rangle \geq \langle\theta^*, x_{t,a_t^*}\rangle - \frac{2}{\sqrt{T}} = \langle\theta^*, X_t\rangle + \Delta_t - \frac{2}{\sqrt{T}}
$$

therefore since event B_t occurs,

$$
\Delta_t \leq \langle\hat{\theta}_t^{s_t} - \theta^*, X_t\rangle + \frac{2}{\sqrt{T}} \leq \frac{3}{\sqrt{T}} .
$$

Hence F_1 in this case is bounded as $2\sqrt{T}$. Combining the above cases,

$$
F_1 \leq 3\sqrt{T\log T} + 3\sqrt{T} + \frac{\pi^2}{12}\log^2 T \leq O(\sqrt{T\log T}) .
$$

To bound F_2, we note from the definition of B_t that

$$\max_{a \in \hat{\mathcal{A}}_{s_t}} \langle \hat{\theta}_t^{s_t}, x_{t,a} \rangle \geq \langle \hat{\theta}_t^{s_t}, x_{t,a_t^*} \rangle \geq \langle \theta^*, x_{t,a_t^*} \rangle - \frac{2}{\sqrt{T}}$$

then on the event $\overline{D}_t$,

$$\langle \theta^*, x_{t,a_t^*} \rangle - \frac{2}{\sqrt{T}} \leq \max_{a \in \hat{\mathcal{A}}_{s_t}} \langle \hat{\theta}_t^{s_t}, x_{t,a} \rangle < \varepsilon + \langle \hat{\theta}_t^{s_t}, X_t \rangle = \frac{\Delta_t}{3} + \langle \hat{\theta}_t^{s_t}, X_t \rangle \ ,$$

therefore

$$\Delta_t < \frac{3}{2} \left(\langle \hat{\theta}_t^{s_t} - \theta^*, X_t \rangle + \frac{2}{\sqrt{T}} \right) \leq \frac{9}{2\sqrt{T}}$$

Hence

$$F_2 = \mathbb{E} \left[\sum_{t \in \Psi_0} \Delta_t \cdot \mathbb{1}\left\{ B_t, \overline{D}_t \right\} \cdot \mathbb{1}\left\{ a_t^* \in \hat{\mathcal{A}}_{s_t} \right\} \right] + \mathbb{E} \left[\sum_{t \in \Psi_0} \Delta_t \cdot \mathbb{1}\left\{ B_t, \overline{D}_t \right\} \cdot \mathbb{1}\left\{ a_t^* \notin \hat{\mathcal{A}}_{s_t} \right\} \right]$$

$$\leq \mathbb{E} \left[\sum_{t=1}^{T} \Delta_t \cdot \mathbb{1}\left\{ B_t, \overline{D}_t \right\} \right] + \frac{\pi^2}{12} \log^2 T$$

$$< \mathbb{E} \left[\sum_{t=1}^{T} \frac{9}{2\sqrt{T}} \cdot \mathbb{1}\left\{ B_t, \overline{D}_t \right\} \right] + \frac{\pi^2}{12} \log^2 T$$

$$< T \cdot \frac{9}{2\sqrt{T}} + \frac{\pi^2}{12} \log^2 T$$

$$= \frac{9}{2}\sqrt{T} + \frac{\pi^2}{12} \log^2 T$$

$$\leq O(\sqrt{T}) \ .$$

To bound F_3, we use the proof as in of Chu et al. (2011, Lemma 1), which is restated as follows.

Lemma 10. *For any $a \in [K]$, $s \in [\lceil \log T \rceil]$, $t \in [T]$,*

$$\mathbb{P} \left[|\langle \theta^* - \hat{\theta}_t^s, x_{t,a} \rangle| > \frac{\alpha+1}{\alpha} w_{t,a}^s \right] \leq \frac{\gamma}{TK}$$

where $\alpha = \sqrt{\frac{1}{2} \ln \frac{2TK}{\gamma}}$.

Then using the union bound, we have for all $t \in [T]$, $s \in [\lceil \log T \rceil]$, for all $a \in [K]$,

$$\mathbb{P} \left[\overline{B}_t \right] = \mathbb{P} \left[\bigcup_{t \in [T], s \in [\lceil \log T \rceil], a \in [K]} \left\{ |\langle \theta^* - \hat{\theta}_t^s, x_{t,a} \rangle| > \frac{\alpha+1}{\alpha} w_{t,a}^s \right\} \right]$$

$$\leq \sum_{t \in [T]} \sum_{s \in [\lceil \log T \rceil]} \sum_{a \in [K]} \mathbb{P} \left[|\langle \theta^* - \hat{\theta}_t^s, x_{t,a} \rangle| > \frac{\alpha+1}{\alpha} w_{t,a}^s \right]$$

$$< \left(TK(1 + \log T) \right) \frac{\gamma}{TK}$$

$$= \gamma(1 + \log T) \ .$$

With the choice $\gamma = \frac{1}{2t^2}$ and the assumption $\Delta_t \leq 1$,

$$
\begin{aligned}
F_3 &= \mathbb{E}\left[\sum_{t \in \Psi_0} \Delta_t \cdot \mathbb{1}\left\{\overline{B}_t\right\}\right] \\
&\leq \sum_{t=1}^{T} \mathbb{P}\left[\overline{B}_t\right] \\
&< \sum_{t=1}^{T} \frac{1 + \log T}{2t^2} \\
&< \frac{\pi^2}{12}(1 + \log T) \\
&\leq O(\log T) .
\end{aligned}
$$

Hence the first term in R_T in (37) is upper bounded by:

$$
\begin{aligned}
\mathbb{E}\left[\sum_{t \in \Psi_0} \langle \theta^*, x_{t,a_t^*} - X_t \rangle\right] &\leq O(\sqrt{T}) + O(\log T) + O(\sqrt{T \log T}) \\
&\leq O(\sqrt{T \log T}) .
\end{aligned}
$$

On the other hand, by Chu et al. (2011, Theorem 1), the second term in R_T in (37) is upper bounded as follows:

$$
\mathbb{E}\left[\sum_{t \notin \Psi_0} \langle \theta^*, x_{t,a_t^*} - X_t \rangle\right] \leq O\left(\sqrt{dT \log^3(KT)}\right).
$$

Hence the regret of our algorithm SupLinIMED is upper bounded as follows:

$$
R_T \leq O\left(\sqrt{dT \log^3(KT)}\right) .
$$

This completes the proof of Theorem 4.

G   Hyperparameter tuning in our empirical study

G.1   Synthetic Dataset

The below tables are the empirical results while tuning the hyperparameter α (scale of the confidence width) for fixed $T = 1000$.

Method	LinUCB			LinTS			LinIMED-1			LinIMED-2			LinIMED-3 ($C = 30$)		
α	0.5	0.55	0.6	0.2	0.25	0.3	0.15	0.2	0.25	0.2	0.25	0.3	0.15	0.2	0.25
Regret	7.780	6.695	6.856	9.769	9.201	12.068	24.086	5.482	6.108	4.999	4.998	7.329	25.588	2.075	2.760

Table 2: Tuning α when $K = 10, d = 2$

Method	LinUCB			LinTS			LinIMED-1			LinIMED-2			LinIMED-3 ($C = 30$)		
α	0.5	0.55	0.6	0.1	0.15	0.2	0.2	0.25	0.3	0.2	0.25	0.3	0.2	0.25	0.3
Regret	7.203	6.832	7.423	54.221	7.042	7.352	6.707	6.053	8.458	6.254	4.918	7.013	4.407	2.562	3.041

Table 3: Tuning α when $K = 100, d = 2$

Method	LinUCB			LinTS			LinIMED-1			LinIMED-2			LinIMED-3 ($C = 30$)		
α	0.5	0.55	0.6	0.1	0.15	0.2	0.15	0.2	0.25	0.2	0.25	0.3	0.15	0.2	0.25
Regret	7.919	5.679	7.063	69.955	6.925	7.037	24.393	5.625	6.335	6.335	4.831	7.040	41.355	1.936	2.250

Table 4: Tuning α when $K = 500, d = 2$

Method	LinUCB			LinTS			LinIMED-1			LinIMED-2			LinIMED-3 ($C = 30$)		
α	0.45	0.5	0.55	0.1	0.15	0.2	0.1	0.15	0.2	0.1	0.15	0.2	0.1	0.15	0.2
Regret	9.164	9.094	14.183	14.252	9.886	14.680	19.663	6.463	10.643	15.685	5.399	8.373	8.024	2.062	3.342

Table 5: Tuning α when $K = 10, d = 20$

Method	LinUCB			LinTS			LinIMED-1			LinIMED-2			LinIMED-3 ($C = 30$)		
α	0.25	0.3	0.35	0.1	0.15	0.2	0.05	0.1	0.15	0.1	0.15	0.2	0.05	0.1	0.15
Regret	7.923	7.085	10.981	14.983	9.565	19.300	58.278	6.165	9.225	8.916	8.575	13.483	142.704	2.816	3.497

Table 6: Tuning α when $K = 10, d = 50$

We run these algorithms on the same dataset with different choices of α, we choose the best α with the corresponding least regret.

G.2 MovieLens Dataset

The below tables are the empirical results while tuning the hyperparameter α (scale of the confidence width) for fixed $T = 1000$.

Method	LinUCB			LinTS			LinIMED-1			LinIMED-2			LinIMED-3 ($C = 30$)			IDS		
α	0.7	0.75	0.8	0.05	0.1	0.15	0.15	0.2	0.25	0.15	0.2	0.25	0.2	0.25	0.3	0.25	0.3	0.35
CTR	0.608	0.675	0.668	0.615	0.705	0.679	0.740	0.823	0.766	0.740	0.823	0.766	0.713	0.742	0.690	0.655	0.728	0.714

Table 7: Tuning α when $K = 20$

Method	LinUCB			LinTS			LinIMED-1			LinIMED-2			LinIMED-3 ($C = 30$)			IDS		
α	0.75	0.8	0.85	0	0.05	0.1	0.1	0.15	0.2	0.05	0.1	0.15	0.05	0.1	0.15	0.3	0.35	0.4
CTR	0.708	0.754	0.713	0.517	0.711	0.646	0.648	0.668	0.595	0.658	0.668	0.651	0.697	0.717	0.649	0.643	0.688	0.606

Table 8: Tuning α when $K = 50$

Method	LinUCB			LinTS			LinIMED-1			LinIMED-2			LinIMED-3 ($C = 30$)			IDS		
α	0.85	0.9	0.95	0	0.05	0.1	0.05	0.1	0.15	0.05	0.1	0.15	0.05	0.1	0.15	0.3	0.35	0.4
CTR	0.721	0.754	0.745	0.487	0.674	0.588	0.682	0.729	0.594	0.687	0.729	0.594	0.689	0.705	0.594	0.684	0.739	0.695

Table 9: Tuning α when $K = 100$

We run these algorithms on the same dataset with different choices of α and we choose the best α with the corresponding largest reward.