When Demands Evolve Larger and Noisier: Learning and Earning in a Growing Environment — Supplementary Material

Anonymous Authors[1]

A. Proofs for Section 3

Proof of Lemma 1.

$$\begin{aligned}
\mathbb{P}\left(\left|\sum_{i=1}^{n} X_i\right| > \delta\right) &= \mathbb{P}\left(\sum_{i=1}^{n} X_i > \delta\right) + \mathbb{P}\left(\sum_{i=1}^{n} X_i < -\delta\right) \\
&\leqslant \inf_{\lambda\in\mathbb{R}^+} \frac{\exp\left(\lambda\sum_{i=1}^{n} X_i\right)}{\exp(\lambda\delta)} + \inf_{\lambda\in\mathbb{R}^-} \frac{\exp\left(\lambda\sum_{i=1}^{n} X_i\right)}{\exp(-\lambda\delta)} \\
&= \inf_{\lambda\in\mathbb{R}^+} \frac{\exp\left(\lambda\sum_{i=1}^{n} X_i\right)}{\exp(\lambda\delta)} + \inf_{\lambda\in\mathbb{R}^+} \frac{\exp\left(-\lambda\sum_{i=1}^{n} X_i\right)}{\exp(\lambda\delta)} \\
&\leqslant 2\inf_{\lambda\in\mathbb{R}_+} \exp\left(\frac{\lambda^2}{2}\sum_{i=1}^{n}\sigma_i^2 - \lambda\delta\right) = 2\exp\left(-\frac{\delta^2}{2\sum_{i=1}^{n}\sigma_i^2}\right).
\end{aligned}$$

□

Proof of Lemma 2. If $x \geqslant 0$, then

$$S_{x,t} = S_{x,t-1} + t^x \leqslant \int_{s=1}^{t} s^x ds + t^x = \frac{t^{x+1}-1}{x+1} + t^x = \frac{t^{x+1}-1}{x+1} + \max\{1, t^x\},$$

and

$$S_{x,t} \geqslant \int_{s=0}^{t} s^x ds > \int_{s=1}^{t} s^x = \frac{t^{x+1}-1}{x+1}.$$

If $-1 \leqslant x < 0$, then

$$S_{x,t} = 1 + \sum_{s=2}^{t} s^x \leqslant 1 + \int_{s=1}^{t} s^x dx = 1 + \frac{t^{x+1}-1}{x+1} \leqslant \frac{t^{x+1}-1}{x+1} + \max\{1, t^x\},$$

and

$$S_{x,t} = S_{x,t-1} + t^x \geqslant \int_{s=1}^{t} s^x ds + t^x > \frac{t^{x+1}-1}{x+1}.$$

□

A.1. Proofs for Section 3.1

Proof of Theorem 1. In the following proof, we assume that $\epsilon_t \sim \mathcal{N}(0, \sigma_t^2)$, where $\sigma_t = t^\alpha \sigma$. Given θ, the log density of history H_t is

$$f(H_t, \theta) = \sum_{s=1}^{t} -\frac{(d_s(p_s) - s^\gamma(a - bp_s))^2}{2\sigma_s^2}$$

Then the Fisher Information Matrix is

$$\mathcal{L}_t = \frac{\partial^2 f(H_t,\theta)}{\partial\theta^2} = \sum_{s=1}^{t} \begin{bmatrix} s^\gamma/\sigma_s \\ s^\gamma p_s/\sigma_s \end{bmatrix} \begin{bmatrix} \frac{s^\gamma}{\sigma_s} & \frac{s^\gamma p_s}{\sigma_s} \end{bmatrix} = \sigma^{-2}\sum_{s=1}^{t} s^{2\gamma-2\alpha} \begin{bmatrix} 1 \\ p_s \end{bmatrix} \begin{bmatrix} 1 & p_s \end{bmatrix}.$$

Let λ be an absolutely continuous density on Θ, taking positive values on the interior of Θ and zero on its boundary. Then the multivariate van Trees inequality (see, e.g., Gill et al., 1995; Keskin & Zeevi, 2014) implies that

$$\mathbb{E}_\lambda\{\mathbb{E}_\theta^\pi[(p_t-\phi(\theta)^2)]\} \geqslant \frac{(\mathbb{E}_\lambda[C(\theta)(\partial\phi/\partial\theta)^\top])^2}{\mathbb{E}_\lambda[C(\theta)\mathcal{L}_{t-1}^\pi C(\theta)^\top]+\tilde{\mathcal{L}}(\lambda)}, \tag{1}$$

where $\tilde{\mathcal{L}}(\lambda)$ is the Fisher information for the density λ, $\mathbb{E}_\lambda$ is the expectation operator with respect to λ, and we let $C(\theta) = \begin{bmatrix} -\phi(\theta) & 1 \end{bmatrix}$. Therefore,

$$\sum_{t=2}^{T} \mathbb{E}_\lambda\{\mathbb{E}_\theta^\pi[t^\gamma(p_t-\phi(\theta)^2)]\} \geqslant \sum_{t=2}^{T} \frac{t^\gamma(\mathbb{E}_\lambda[\phi(\theta)/(2b)])^2}{\mathbb{E}_\lambda[C(\theta)\mathbb{E}_\theta^\pi[\mathcal{L}_{t-1}]C(\theta)^\top]+\tilde{\mathcal{L}}(\lambda)}.$$

Thus,

$$\sup_{\theta\in\Theta}\mathbb{E}_\theta^\pi[\sum_{t=1}^{T} t^\gamma(p_t-\phi(\theta))^2] \geqslant \frac{\left(\frac{T^{\gamma+1}-1}{\gamma+1}\right)\inf_{\theta\in\Theta}(\phi(\theta)/(2b))^2}{\sigma^{-2}\sup_{\theta\in\Theta}\mathbb{E}_\theta^\pi[\sum_{t=1}^{T} t^{2\gamma-2\alpha}(p_t-\phi(\theta))^2]+\tilde{\mathcal{L}}(\lambda)}. \tag{2}$$

Part I: Fixed-time regret lower bound

Note that

$$\min\left\{\alpha+\frac{1}{2}, \frac{(\gamma+1)^2}{3\gamma-2\alpha+2}\right\} = \begin{cases} \alpha+\frac{1}{2}, & \text{if } \alpha\in[0,\frac{\gamma}{2}], \\ \frac{(\gamma+1)^2}{3\gamma-2\alpha+2}, & \text{if } \alpha\in(\frac{\gamma}{2},\gamma+\frac{1}{2}), \end{cases}$$

and we will prove the desired the result in four cases.

<u>*Case 1:*</u> If $\alpha\in[0,\frac{\gamma}{2}]$, then applying the following inequality into (2) yields the desired result:

$$\sup_{\theta\in\Theta}\mathbb{E}_\theta^\pi[\sum_{t=1}^{T} t^{2\gamma-2\alpha}(p_t-\phi(\theta))^2] \leqslant T^{\gamma-2\alpha}\sup_{\theta\in\Theta}\mathbb{E}_\theta^\pi[\sum_{t=1}^{T} t^\gamma(p_t-\phi(\theta))^2].$$

<u>*Case 2:*</u> If $\alpha\in(\frac{\gamma}{2},\gamma+\frac{1}{2})$, let $\eta=\frac{\gamma+1}{3\gamma-2\alpha+2}<1$ and $c>0$ be a constant such that $c<\frac{1}{2}$ and $cT^\eta\in\mathbb{Z}_+$. Then, using the notation that $\Delta_t=(p_t-\phi(\theta))^2$, we have

$$\begin{aligned}
&\mathbb{E}_\lambda\{\mathbb{E}_\theta^\pi[\sum_{t=1}^{T} t^\gamma\Delta_t]\} \\
&= \mathbb{E}_\lambda\{\mathbb{E}_\theta^\pi[\sum_{t=1}^{cT^\eta} t^\gamma\Delta_t]\} + \mathbb{E}_\lambda\{\mathbb{E}_\theta^\pi[\sum_{t=cT^\eta+1}^{\frac{T}{2}} t^{(2\alpha-\gamma)+(2\gamma-2\alpha)}\Delta_t]\} + \mathbb{E}_\lambda\{\mathbb{E}_\theta^\pi[\sum_{t=\frac{T}{2}+1}^{T} t^\gamma\Delta_t]\} \\
&\geqslant 0 + \frac{1}{2}(cT^\eta)^{2\alpha-\gamma}\mathbb{E}_\lambda\{\mathbb{E}_\theta^\pi[\sum_{t=cT^\eta+1}^{T} t^{2\gamma-2\alpha}\Delta_t]\} + \mathbb{E}_\lambda\{\mathbb{E}_\theta^\pi[\sum_{t=\frac{T}{2}+1}^{T}(t^\gamma-\frac{1}{2}(cT^\eta)^{2\alpha-\gamma}t^{2\gamma-2\alpha})\Delta_t]\} \\
&\geqslant \frac{1}{2}(cT^\eta)^{2\alpha-\gamma}\mathbb{E}_\lambda\{\mathbb{E}_\theta^\pi[\sum_{t=cT^\eta+1}^{T} t^{2\gamma-2\alpha}\Delta_t]\} + \frac{\left(\sum_{t=\frac{T}{2}+1}^{T} t^\gamma - \sum_{t=\frac{T}{2}+1}^{T}\frac{1}{2}t^{2\alpha-\gamma}t^{2\gamma-2\alpha}\right)\inf_{\theta\in\Theta}(\phi(\theta)/(2b))^2}{\sigma^{-2}\mathbb{E}_\lambda\{\mathbb{E}_\theta^\pi[\sum_{t=1}^{T} t^{2\gamma-2\alpha}(p_t-\phi(\theta))^2]\}+\tilde{\mathcal{L}}(\lambda)} \\
&\geqslant \frac{1}{2}(cT^\eta)^{2\alpha-\gamma}\mathbb{E}_\lambda\{\mathbb{E}_\theta^\pi[\sum_{t=cT^\eta+1}^{T} t^{2\gamma-2\alpha}\Delta_t]\} + \frac{\frac{1}{2}\int_{\frac{T}{2}}^{T} t^\gamma dt\,\inf_{\theta\in\Theta}(\phi(\theta)/(2b))^2}{\sigma^{-2}\mathbb{E}_\lambda\{\mathbb{E}_\theta^\pi[\sum_{t=1}^{T} t^{2\gamma-2\alpha}(p_t-\phi(\theta))^2]\}+\tilde{\mathcal{L}}(\lambda)} \\
&= \frac{1}{2}(cT^\eta)^{2\alpha-\gamma}y + \frac{c_0T^{\gamma+1}}{\sigma^{-2}(x+y)+\tilde{\mathcal{L}}(\lambda)},
\end{aligned}$$

where in the second inequality we utilize (1) and $cT^\eta \leqslant \frac{T}{2}$, and in the last equality we let

$$x = \mathbb{E}_\lambda\{\mathbb{E}_\theta^\pi[\sum_{t=1}^{cT^\eta} t^{2\gamma-2\alpha}\Delta_t]\},\ y = \mathbb{E}_\lambda\{\mathbb{E}_\theta^\pi[\sum_{cT^\eta+1}^{T} t^{2\gamma-2\alpha}\Delta_t]\},\ \text{and } c_0 = \frac{1-2^{-\gamma}}{2(\gamma+1)}\inf_{\theta\in\Theta}(\phi(\theta)/(2b))^2$$

Obviously, $x \leqslant (u-l)^2\sum_{t=1}^{cT^\eta} t^{2\gamma-2\alpha}$. If $y \geqslant (u-l)^2\sum_{t=1}^{cT^\eta} t^{2\gamma-2\alpha}$, we have

$$\mathbb{E}_\lambda\{\mathbb{E}_\theta^\pi[\sum_{t=1}^{T} t^\gamma\Delta_t]\} \geqslant \frac{1}{2}(u-l)^2(cT^\eta)^{2\alpha-\gamma}\sum_{t=1}^{cT^\eta} t^{2\gamma-2\alpha} = \Omega(T^\beta).$$

If $y < (u-l)^2\sum_{t=1}^{cT^\eta} t^{2\gamma-2\alpha}$, then we have

$$\mathbb{E}_\lambda\{\mathbb{E}_\theta^\pi[\sum_{t=1}^{T} t^\gamma\Delta_t]\} \geqslant \frac{c_0T^{\gamma+1}}{2\sigma^{-2}(u-l)^2\sum_{t=1}^{cT^\eta} t^{2\gamma-2\alpha} + \tilde{\mathcal{L}}(\lambda)} = \Omega(T^{\gamma+1-\eta(2\gamma-2\alpha+1)}) = \Omega(T^\beta).$$

Case 3: If $\alpha = \gamma + \frac{1}{2}$, then $2\gamma - 2\alpha = -1$. Thus,

$$\sum_{t=1}^{T} t^{2\gamma-2\alpha}(p_t - \phi(\theta))^2 \leqslant \sum_{t=1}^{T} t^{2\gamma-2\alpha}(u-l)^2 = O(\log T).$$

Then from (2), we can directly yield the result.

Case 4: If $\alpha > \gamma + \frac{1}{2}$, then $2\gamma - 2\alpha < -1$. Thus,

$$\sum_{t=1}^{T} t^{2\gamma-2\alpha}(p_t - \phi(\theta))^2 \leqslant \sum_{t=1}^{T} t^{2\gamma-2\alpha}(u-l)^2 < +\infty.$$

Then from (2), we can directly yield the result.

Part II: Any-time regret lower bound

Case 1: If $\alpha \in [0, \frac{\gamma}{2}]$, the result holds directly from Case 1 in Part I of our proof.

Case 2: If $\alpha \in (\frac{\gamma}{2}, \gamma + \frac{1}{2})$, suppose that the conclusion does not hold. Let $\beta = \alpha + \frac{1}{2}$. Then for any $\epsilon > 0$, there exists an *any-time* pricing policy π_ϵ, a constant $C_\epsilon \in (0, \epsilon)$, and $T_\epsilon = 2^{k_\epsilon}$ such that for any $T \geqslant T_\epsilon$, we have

$$\sup_{\theta\in\Theta}\mathbb{E}_\theta^{\pi_\epsilon}[\sum_{t=1}^{T} t^\gamma\Delta_t] \leqslant c_\epsilon T^\beta, \tag{3}$$

where $c_\epsilon = C_\epsilon/\inf_{\theta\in\Theta} b$. The critical step lies in estimating $\mathbb{E}_\theta^{\pi_\epsilon}[\sum_{t=1}^{T} t^{2\gamma-2\alpha}(p_t-\phi(\theta))^2]$. In fact, we have, $\forall\theta\in\Theta$,

$$\begin{aligned}
\mathbb{E}_\theta^{\pi_\epsilon}[\sum_{t=1}^{T} t^{2\gamma-2\alpha}\Delta_t] &\leqslant \mathbb{E}_\theta^{\pi_\epsilon}[\sum_{t=1}^{T_\epsilon} t^\gamma\Delta_t] + \sum_{k=\log_2 T_\epsilon}^{\lfloor\log_2 T\rfloor}\sum_{t=2^k}^{2^{k+1}-1} t^{2\gamma-2\alpha}\mathbb{E}_\theta^{\pi_\epsilon}[\Delta_t] \\
&\leqslant \mathbb{E}_\theta^{\pi_\epsilon}[\sum_{t=1}^{T_\epsilon} t^\gamma\Delta_t] + \sum_{k=\log_2 T_\epsilon}^{\lfloor\log_2 T\rfloor} 2^{k(\gamma-2\alpha)}\sum_{t=2^k}^{2^{k+1}-1} t^\gamma\mathbb{E}_\theta^{\pi_\epsilon}[\Delta_t] \\
&\leqslant (u-l)^2T_\epsilon^{\gamma+1} + \sum_{k=0}^{\lfloor\log_2 T\rfloor} 2^{k(\gamma-2\alpha)}c_\epsilon 2^{(k+1)\beta} \\
&= (u-l)^2T_\epsilon^{\gamma+1} + c_\epsilon 2^\beta\sum_{k=0}^{\lfloor\log_2 T\rfloor} 2^{(\gamma-2\alpha+\beta)k}
\end{aligned}$$

From our assumption, $\beta = \alpha + \frac{1}{2} > 2\alpha - \gamma$, then we have

$$c_\epsilon 2^\beta \sum_{k=0}^{\lfloor \log_2 T \rfloor} 2^{(\gamma-2\alpha+\beta)k} \leqslant c_\epsilon 2^\beta \frac{(2T)^{\gamma-2\alpha+\beta} - 1}{2^{\gamma-2\alpha+\beta} - 1}.$$

Thus we have

$$\sup_{\theta\in\Theta} \mathbb{E}_\theta^{\pi_\epsilon}[\sum_{t=1}^{T} t^{2\gamma-2\alpha}\Delta_t] \leqslant (u-l)^2 T_\epsilon^{\gamma+1} + c_\epsilon 2^\beta \frac{(2T)^{\gamma-2\alpha+\beta} - 1}{2^{\gamma-2\alpha+\beta} - 1}. \tag{4}$$

Together with (2), (3) and (4), we have

$$c_\epsilon T^\beta (u-l)^2 T_\epsilon^{\gamma+1} + c_\epsilon^2 \sigma^{-2} (2T)^\beta \frac{(2T)^{(\gamma-2\alpha+\beta)} - 1}{2^{\gamma-2\alpha+\beta} - 1} + c_\epsilon T^\beta \tilde{\mathcal{L}}(\lambda) \geqslant \frac{T^{\gamma+1} - 1}{\gamma+1} \inf_{\theta\in\Theta} (\phi(\theta)/(2\beta))^2$$

for all $T \geqslant T_\epsilon$. Let $\tilde{C} = \sigma^{-2} \frac{2^{\gamma+1}}{2^{\gamma+\frac{1}{2}-\alpha}-1}$ that depends only on α, γ and σ. Then dividing the both sides of the above formula by $T^{\gamma+1}$ and taking T to $+\infty$ yields

$$c_\epsilon^2 \tilde{C} \geqslant \frac{1}{\gamma+1} \inf_{\theta\in\Theta} (\phi(\theta)/(2b))^2 > 0.$$

However, when $\epsilon \to 0$, the left hand side turns into 0. A contradiction! Therefore, there must exists some constant $C > 0$ such that for any *any-time* pricing policy π,

$$\limsup_{T} \left\{ \sup_{\theta\in\Theta} \{R_\theta^\pi(T)\}/T^{\alpha+\frac{1}{2}} \right\} \geqslant C.$$

□

A.2. Proofs for Section 3.2

Proof of Lemma 3. Let $y = (y_1, y_2)$ such that $\|y\| = 1$. We have

$$\begin{aligned} y^\top \mathcal{J}_t y &= \sum_{s=1}^{t} s^{2\gamma-2\alpha}(y_1 + y_2 p_s)^2 = \sum_{s=1}^{t} s^{2\gamma-2\alpha}(y_1 + y_2 \bar{p}_t)^2 + y_2^2 \sum_{s=1}^{t} s^{2\gamma-2\alpha}(p_s - \bar{p}_t)^2 \\ &\geqslant \left(\frac{(y_1 + y_2\bar{p}_t)^2}{(u-l)^2} + y_2^2 \right) J_t \geqslant \frac{2}{(1+2u-l)^2} J_t. \end{aligned}$$

□

Proof of Lemma 4. It's easy to obtain that $\hat{\theta}_t - \theta = \mathcal{J}_t^{-1}\mathcal{M}_t$, where

$$\mathcal{M}_t = \sum_{s=1}^{t} s^{\gamma-\alpha}[\tilde{\epsilon}_s \; p_s\tilde{\epsilon}_s]^\top.$$

Under our assumption, we have

$$\mathbb{E}_\theta^\pi[\exp(z\tilde{\epsilon}_t)] \leqslant \exp(\frac{1}{2}\sigma^2 z^2), \forall z \in \mathbb{R}.$$

Now we define a series of martingales $\{Z_s^y\}$:

$$Z_s^y = \exp\{\frac{1}{\zeta\sigma^2}(y^\top \mathcal{M}_s - \frac{1}{2} y^\top \mathcal{J}_s y)\}, \forall s = 1, 2, ...,$$

where $\zeta = 1 \vee \delta$ and $\|y\| = \delta$. Then Z_s^y is integrable with respect to y for all s. Let $\mathcal{F}_s = \sigma(d_1, ..., d_s)$, then we get

$$\begin{aligned}
\mathbb{E}_\theta^\pi[Z_s^y|\mathcal{F}_{s-1}] &= \exp\{\frac{1}{\zeta\sigma^2}(y^\top \mathcal{M}_{s-1} - \frac{1}{2}y^\top \mathcal{J}_s y)\}\mathbb{E}_\theta^\pi[\exp\{\frac{1}{\zeta\sigma^2}y^\top(\mathcal{M}_s - \mathcal{M}_{s-1})\}|\mathcal{F}_{s-1}] \\
&\leqslant \exp\{\frac{1}{\zeta\sigma^2}(y^\top \mathcal{M}_{s-1} - \frac{1}{2}y^\top \mathcal{J}_s y)\}\exp\{\frac{1}{2\zeta\sigma^2}y^\top \begin{bmatrix} s^{2\gamma-2\alpha} & s^{2\gamma-2\alpha}p_s \\ s^{2\gamma-2\alpha}p_s & s^{2\gamma-2\alpha}p_s^2 \end{bmatrix} y\} \\
&= \exp\{\frac{1}{\zeta\sigma^2}(y^\top \mathcal{M}_{s-1} - \frac{1}{2}y^\top \mathcal{J}_{s-1} y)\} = Z_{s-1}^y.
\end{aligned}$$

Thus $(Z_s^y, \mathcal{F}_s)$ is a super-martingale for any $y \in \mathbb{R}^2$ with $\|y\| = \delta$.

Now we consider $\tilde{Z}_s = Z_s^{\omega_s}$ such that $\omega_s = \delta\mathcal{J}_s^{-1}\mathcal{M}_s/\|\mathcal{J}_s^{-1}\mathcal{M}_s\|$ for all s. Fix $m > 0$ and let $\xi \geqslant \delta$ be a positive real number to be determined later. Let $A = \{\|\mathcal{M}_t\| \leqslant \xi S_{2\gamma-2\alpha,t}\} \in \mathcal{F}_t$. Then we have

$$\begin{aligned}
\mathbb{P}_\theta^\pi(\|\hat{\theta} - \theta\| > \delta, J_t \geqslant m) &= \mathbb{P}_\theta^\pi(\|\mathcal{J}_t^{-1}\mathcal{M}_t\| > \delta, J_t \geqslant m) \\
&\leqslant \mathbb{P}_\theta^\pi(\|\mathcal{J}_t^{-1}\mathcal{M}_t\| > \delta, J_t \geqslant m, A) + \mathbb{P}_\theta^\pi(J_t \geqslant m, A^c).
\end{aligned} \tag{5}$$

For the first term, we have

$$\begin{aligned}
\mathbb{P}_\theta^\pi(\|\mathcal{J}_t^{-1}\mathcal{M}_t\| > \delta, J_t \geqslant m, A) &\leqslant \mathbb{P}_\theta^\pi(\omega_t^\top \mathcal{M}_t \geqslant \omega_t^\top \mathcal{J}_t\omega_t, J_t \geqslant m, A) \\
&\leqslant \mathbb{P}_\theta^\pi(\tilde{Z}_t \geqslant \exp(\frac{\omega_t^\top \mathcal{J}_t\omega_t}{2\zeta\sigma^2}), J_t \geqslant m, A) \\
&\leqslant \mathbb{P}_\theta^\pi(\tilde{Z}_t \geqslant \exp(\frac{\mu\delta^2 m}{2\zeta\sigma^2}), A) \quad \text{(Lemma 3)}
\end{aligned}$$

Note that

$$\begin{aligned}
&\left(\omega_t^\top \mathcal{M}_t - \frac{1}{2}\omega_t^\top \mathcal{J}_t\omega_t\right) - \left(y^\top \mathcal{M}_t - \frac{1}{2}y^\top \mathcal{J}_t y\right) \\
&\leqslant (\omega_t - y)^\top(\mathcal{M}_t - \mathcal{J}_t y) \\
&\leqslant \|\omega_t - y\| \left(\|\mathcal{M}_t\| + \|\mathcal{J}_t y\|\right) \\
&\leqslant \|\omega_t - y\| \left(\|\mathcal{M}_t\| + \sum_{s=1}^t s^{2\gamma-2\alpha}\|[1\ p_s]^\top(y_1 + p_s y_2)\|\right) \\
&\leqslant \|\omega_t - y\| \left(\xi S_{2\gamma-2\alpha,t} + (1+u^2)\delta S_{2\gamma-2\alpha,t}\right) \\
&\leqslant \|\omega_t - y\|(2+u^2)\xi S_{2\gamma-2\alpha,t}.
\end{aligned}$$

Thus, $\tilde{Z}_t \leqslant \exp\left(\frac{\xi(2+u^2)}{\zeta\sigma^2}\right) Z_t^y$ holds for any y that is within the $(1/S_{2\gamma-2\alpha,t})$-neighbourhood of ω_t. Therefore, considering a set of $\lceil \pi\delta S_{2\gamma-2\alpha,t}\rceil$ points evenly spaced on the circle $\{y \in \mathbb{R}^2 : \|y\| = \delta\}$ yields

$$\mathbb{P}_\theta^\pi(\|\mathcal{J}_t^{-1}\mathcal{M}_t\| > \delta, J_t \geqslant m, A) \leqslant \exp\left(\frac{\xi(2+u^2)}{\zeta\sigma^2}\right)(1 + \pi(1\vee\delta)S_{2\gamma-2\alpha,t})\exp\left(-\frac{\mu\delta^2 m}{2\zeta\sigma^2}\right), \tag{6}$$

where we use Markov Inequality for the super-martingale $(Z_t^y, \mathcal{F}_t)$.

For the second term, we have

$$
\begin{aligned}
\mathbb{P}_\theta^\pi(J_t \geqslant m, A^c) &= \mathbb{P}_\theta^\pi\left(J_t \geqslant m, \|\mathcal{M}_t\| > \xi S_{2\gamma-2\alpha,t}\right) \\
&\leqslant \mathbb{P}_\theta^\pi\left(J_t \geqslant m, \left|\sum_{s=1}^t s^{\gamma-\alpha}\epsilon_s\right| > \frac{\xi S_{2\gamma-2\alpha,t}}{\sqrt{2}}\right) + P_\theta^\pi\left(J_t \geqslant m, \left|\sum_{s=1}^t s^{\gamma-\alpha}\epsilon_s p_s\right| > \frac{\xi S_{2\gamma-2\alpha,t}}{\sqrt{2}}\right) \\
&\leqslant \mathbb{P}_\theta^\pi\left((u-l)^2\sum_{s=1}^t s^{2\gamma-2\alpha} \geqslant m, \left|\frac{\sum_{s=1}^t s^{\gamma-\alpha}\epsilon_s}{\sqrt{\sum_{s=1}^t s^{2\gamma-2\alpha}}}\right| > \frac{\xi S_{2\gamma-2\alpha,t}}{\sqrt{2\sum_{s=1}^t s^{2\gamma-2\alpha}}}\right) + \\
&\quad \mathbb{P}_\theta^\pi\left((u-l)^2\sum_{s=1}^t s^{2\gamma-2\alpha} \geqslant m, \left|\frac{\sum_{s=1}^t s^{\gamma-\alpha}\epsilon_s}{\sqrt{\sum_{s=1}^t s^{2\gamma-2\alpha}}}\right| > \frac{\xi S_{2\gamma-2\alpha,t}}{u\sqrt{2\sum_{s=1}^t s^{2\gamma-2\alpha}}}\right) \\
&\leqslant \mathbb{P}_\theta^\pi\left(\left|\frac{\sum_{s=1}^t s^{\gamma-\alpha}\epsilon_s}{\sqrt{\sum_{s=1}^t s^{2\gamma-2\alpha}}}\right| > \frac{\xi\sqrt{m}}{\sqrt{2}(u-l)}\right) + \mathbb{P}_\theta^\pi\left(\left|\frac{\sum_{s=1}^t s^{\gamma-\alpha}\epsilon_s}{\sqrt{\sum_{s=1}^t s^{2\gamma-2\alpha}}}\right| > \frac{\xi\sqrt{m}}{\sqrt{2}(u-l)u}\right) \\
&\leqslant 2\exp\left(-\frac{\xi^2 m}{4\sigma^2(u-l)^2}\right) + 2\exp\left(-\frac{\xi^2 m}{4\sigma^2(u-l)^2u^2}\right),
\end{aligned}
\tag{7}
$$

where in the last inequality we utilize Lemma 1. Now we let $\xi = 1 \vee \delta$, we can obtain in (6) and (7) that

$$
\begin{aligned}
&\mathbb{P}_\theta^\pi(\|\mathcal{J}_t^{-1}\mathcal{M}_t\| > \delta, J_t \geqslant m, A) \leqslant \exp\left(\frac{2+u^2}{\sigma^2}\right)(1+\pi)(1\vee\delta)S_{2\gamma-2\alpha,t}\exp\left(-\frac{\mu}{2\sigma^2}m(\delta\wedge\delta^2)\right), \\
&\mathbb{P}_\theta^\pi(J_t \geqslant m, A^c) \leqslant 4\exp\left(-\frac{1}{4\sigma^2(u-l)^2(u\wedge 1)^2}m(\delta\wedge\delta^2)\right).
\end{aligned}
$$

Now letting $k = \exp(\frac{2+u^2}{\sigma^2})(1+\pi)+4$ and $\rho = \frac{\mu}{2\sigma^2} \wedge \frac{1}{4\sigma^2(u-l)^2(u\wedge 1)^2}$ completes the proof. □

Proof of Theorem 2. Since $p_1 \neq p_2$, we have $J_t > 0$ for all $t \geqslant 2$. From Lemma 3, $\mathcal{J}_t$ is invertible for all $t \geqslant 2$, and thus ϑ_t exists. Therefore, for any $t \geqslant 2$,

$$
\begin{aligned}
&\mathbb{E}_\theta^\pi[(\phi(\theta)-\phi(\vartheta_t))^2] \\
&\leqslant K\mathbb{E}_\theta^\pi\|\theta-\vartheta_t\|^2 \leqslant K\mathbb{E}_\theta^\pi\|\theta-\hat{\theta}_t\|^2 \\
&= K\int_0^\infty \mathbb{P}_\theta^\pi(\|\theta-\hat{\theta}_t\|^2 > x, J_t \geqslant \kappa_0\sqrt{S_{2\gamma-2\alpha,t}})dx \\
&\leqslant \frac{(2\gamma-2\alpha+1)K\log t}{\rho\kappa_0\sqrt{S_{2\gamma-2\alpha,t}}} + K\int_{\frac{(2\gamma-2\alpha+1)\log t}{\rho\kappa_0\sqrt{S_{2\gamma-2\alpha,t}}}}^1 kS_{2\gamma-2\alpha,t}\exp(-\rho x\kappa_0\sqrt{S_{2\gamma-2\alpha,t}})dx + \\
&\quad K\int_1^\infty k\sqrt{x}S_{2\gamma-2\alpha,t}\exp(-\rho\sqrt{x}\kappa_0\sqrt{S_{2\gamma-2\alpha,t}})dx \quad \text{(Lemma 4)} \\
&\leqslant \frac{(2\gamma-2\alpha+1)K}{\rho\kappa_0\sqrt{S_{2\gamma-2\alpha,t}}}\log t + \frac{Kk}{\rho\kappa_0}t^{-(2\gamma-2\alpha+1)}\sqrt{S_{2\gamma-2\alpha,t}} + \\
&\quad \frac{2Kk}{\rho^3\kappa_0^3\sqrt{S_{2\gamma-2\alpha,t}}}\int_{\rho\kappa_0\sqrt{S_{2\gamma-2\alpha,t}}}^{+\infty} x^2\exp(-x)dx,
\end{aligned}
$$

where in the first inequality we bound $\sup_{\theta'\in\Theta}\left\|\frac{\partial\phi(\theta')}{\partial\theta'}\right\|^2$ by K. Notice that

$$
\int_{\rho\kappa_0\sqrt{S_{2\gamma-2\alpha,t}}}^{+\infty} x^2\exp(-x)dx \leqslant \int_0^{+\infty} x^2\exp(-x)dx = 2.
$$

Therefore, for all $t \geqslant 2$,

$$\begin{aligned}
&\mathbb{E}_\theta^\pi[(\phi(\theta)-\phi(\vartheta_t))^2] \\
&\leqslant \frac{\max\{2\gamma-2\alpha+1,1\}K}{\rho\kappa_0\sqrt{S_{2\gamma-2\alpha,t}}}(\log t + k + \frac{2k}{\rho\kappa_0}\int_{\rho^2\kappa_0^2\sqrt{S_{2\gamma-2\alpha,t}}}^{\infty} x^2\exp(-x)dx) \\
&\leqslant C_0 t^{\alpha-\gamma-\frac{1}{2}}\log t,
\end{aligned}$$

where C_0 is independent of t, but possibly dependent on α and γ. Therefore,

$$\begin{aligned}
R_\theta^\pi(T) &\leqslant \sum_{t=1}^{2} t^\gamma \mathbb{E}_\theta^\pi[(\phi(\theta)-p_t)^2] + b_{\max}\sum_{t=3}^{T} t^\gamma \mathbb{E}_\theta^\pi[(\phi(\theta)-\phi(\vartheta_{t-1}))^2] + b_{\max}\sum_{t=3}^{T} t^\gamma \mathbb{E}_\theta^\pi[(\phi(\vartheta_{t-1})-p_t)^2] \\
&\leqslant C_1 + b_{\max}C_0\left(\frac{3}{2}\right)^\gamma \sum_{t=1}^{T} t^{\alpha-\frac{1}{2}}\log T + b_{\max}\kappa_1 S_{\alpha-\frac{1}{2},T} \\
&= O\left(T^{\alpha+\frac{1}{2}}\log T\right).
\end{aligned}$$

□

Proof of Corollary 1. We only need to verify the two conditions in Theorem 2. First,

$$\begin{aligned}
J_t &= \sum_{s=1}^{t} s^{2\gamma-2\alpha}(p_s-\bar{p}_t)^2 = \sum_{s=2}^{t}\frac{s^{2\gamma-2\alpha}S_{s-1,2\gamma-2\alpha}}{S_{s,2\gamma-2\alpha}}(p_s-\bar{p}_{s-1})^2 \\
&\geqslant \sum_{s=2}^{t}\frac{s^{2\gamma-2\alpha}S_{s-1,2\gamma-2\alpha}}{S_{s,2\gamma-2\alpha}}\left(\kappa s^{\frac{\alpha-\gamma}{2}-\frac{1}{4}}\right)^2 \geqslant \frac{\kappa^2}{1+2^{2\gamma-2\alpha}}\sum_{s=2}^{t} s^{\gamma-\alpha-\frac{1}{2}} \\
&= \Omega(S_{\gamma-\alpha-\frac{1}{2},t}) = \Omega(t^{\gamma-\alpha+\frac{1}{2}}) \\
&= \Omega\left(\sqrt{t^{2\gamma-2\alpha+1}}\right) = \Omega\left(\sqrt{S_{2\gamma-2\alpha,t}}\right).
\end{aligned}$$

Second,

$$\sum_{s=0}^{t} s^\gamma(\phi(\vartheta_s)-p_{s+1})^2 \leqslant \sum_{s=1}^{t} s^\gamma\left(\kappa s^{\frac{\alpha-\gamma}{2}-\frac{1}{4}}\right)^2 = \kappa^2\sum_{s=1}^{t} s^{\alpha-\frac{1}{2}} = O(S_{\alpha-\frac{1}{2},t}).$$

□

A.3. Proofs for Section 3.3

Proof of Theorem 3. We have

$$\det(\mathcal{J}_t) = \lambda_1(\mathcal{J}_t)\lambda_2(\mathcal{J}_t) \leqslant \left(\frac{\lambda_1(\mathcal{J}_t)+\lambda_2(\mathcal{J}_t)}{2}\right)^2 = \frac{1}{4}\mathrm{tr}(\mathcal{J}_t)^2 = \frac{1}{4}\left(\lambda + (1+u^2)\sum_{s=1}^{t} s^{2\gamma-2\alpha}\right)^2. \tag{8}$$

For fixed T, we let $\delta = 1/\sum_{s=1}^{T} s^\gamma$, then

$$\frac{\det(\mathcal{J}_t)^{\frac{1}{2}}\det(\lambda I)^{-\frac{1}{2}}}{\delta} \leqslant \frac{1}{2}S_{\gamma,T}(1+S_{2\gamma-2\alpha,t}).$$

Thus, from Lemma 5 we have

$$\mathbb{P}(\exists t\in[T]:\theta\notin\mathcal{C}_t) = 1-\mathbb{P}(\theta\in\mathcal{C}_t,\forall t\in[T]) \leqslant \delta = 1/\sum_{s=1}^{T} s^\gamma.$$

Now we fix a and b, and let $p^* = \phi(\theta) = \frac{a}{2b}$ be the optimal price. Assume that $\theta \in \mathcal{C}_{t-1}$ happens for all $t \in (T_0, T]$, then we have

$$t^\gamma p^*(a - bp^*) - t^\gamma p_t(a - bp_t) \leqslant b_{\max} t^\gamma \|\phi(\theta) - \phi(\vartheta_t)\|_2^2 = t^\gamma O(\|\theta - \vartheta_t\|_2^2).$$

We let $\vartheta_t = (\alpha_t, \beta_t)$, and further let $\Delta\alpha_t = \alpha_t - a$ and $\Delta\beta_t = \beta_t - b$, then we have

$$\|\theta - \vartheta_t\|_{\mathcal{J}_{t-1}}^2 \leqslant \|\theta - \hat{\theta}_t\|_{\mathcal{J}_{t-1}}^2 + \|\hat{\theta}_t - \vartheta_t\|_{\mathcal{J}_{t-1}}^2 \leqslant 2w_{t-1}^2,$$

which is equivalent to

$$\lambda(\Delta\alpha_t^2 + \Delta\beta_t^2) + \sum_{s=1}^{t-1} s^{2\gamma-2\alpha}(\Delta\alpha_t + \Delta\beta_t p_s)^2 \leqslant 2w_{t-1}^2.$$

If $\Delta\beta_t = 0$, then

$$\|\theta - \vartheta_t\|^2 = \Delta\alpha_t^2 + \Delta\beta_t^2 = \Delta\alpha_t^2 \leqslant \frac{2w_{t-1}^2}{\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}} \leqslant O(\frac{1}{T^{\eta(2\gamma-2\alpha+1)}}).$$

Else, we let $\gamma_t = \frac{\Delta\alpha_t}{\Delta\beta_t}$. Then

$$\Delta\beta_t^2 \leqslant \frac{2w_{t-1}^2}{\sum_{s=1}^{t-1} s^{2\gamma-2\alpha}(\gamma_t + p_s)^2},$$

which means

$$\|\theta - \vartheta_t\|_2^2 = \Delta\beta_t^2(1 + \gamma_t^2) \leqslant \frac{2w_{t-1}^2(1 + \gamma_t^2)}{\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}(\gamma_t + p_s)^2}.$$

If $|\gamma_t| \leqslant 2u$, then

$$\begin{aligned}
\frac{1 + \gamma_t^2}{\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}(\gamma_t + p_t)^2} &\leqslant \frac{1 + 4u^2}{\sum_{s=1, s \text{ odd}}^{T_0} s^{2\gamma-2\alpha}(l_0 + \gamma_t)^2 + \sum_{s=1, s \text{ even}}^{T_0} s^{2\gamma-2\alpha}(u_0 + \gamma_t)^2} \\
&\leqslant \frac{(1 + 4u^2)\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}}{(u_0 - l_0)^2(\sum_{s=1, s \text{ odd}}^{T_0} s^{2\gamma-2\alpha})(\sum_{s=1, s \text{ even}}^{T_0} s^{2\gamma-2\alpha})} \quad \text{(Cauchy Inequality)} \\
&\leqslant \frac{4(1 + 4u^2)\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}}{(u_0 - l_0)^2\left((\sum_{s=1}^{T_0} s^{2\gamma-2\alpha})^2 - (T_0^{2\gamma-2\alpha})^2\right)} \\
&\leqslant \frac{4 + 16u^2}{(u_0 - l_0)^2\sum_{s=1}^{T_0-1} s^{2\gamma-2\alpha}} = O(\frac{1}{T^{\eta(2\gamma-2\alpha+1)}}).
\end{aligned}$$

If $|\gamma_t| > 2u$, then

$$\begin{aligned}
\frac{1 + \gamma_t^2}{\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}(\gamma_t + p_t)^2} &\leqslant \frac{1 + \gamma_t^2}{\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}(|\gamma_t| - u)^2} \\
&= \frac{1 + \frac{1}{\gamma_t^2}}{\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}(1 - \frac{u}{|\gamma_t|})^2} \leqslant \frac{1 + 4u^2}{\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}u^2} \leqslant O(\frac{1}{T^{\eta(2\gamma-2\alpha+1)}}).
\end{aligned}$$

Therefore,

$$\|\theta - \vartheta_t\|_2^2 \leqslant O(\frac{w_{t-1}^2}{T^{\eta(2\gamma-2\alpha+1)}}),\ \forall T_0 < t \leqslant T.$$

We have

$$\begin{aligned}
R_\theta^\pi(T) &\leqslant b_{\max}(u-l)^2 \sum_{t=1}^{T_0} t^\gamma + KT^\gamma \sum_{t=T_0+1}^{T} \|\theta - \vartheta_t\|_2^2 + \sum_{t=T_0+1}^{T} \frac{b_{\max}(u-l)^2 t^\gamma}{\sum_{s=1}^{T} s^\gamma} \\
&\leqslant O(T^{\eta(\gamma+1)}) + O(w_{T-1}^2 T^{\gamma+1-\eta(2\gamma-2\alpha+1)}) + O(1) \\
&= O(T^{\frac{(\gamma+1)^2}{3\gamma-2\alpha+2}} \log T).
\end{aligned}$$

□

A.4. Proofs for Section 3.4

Proof of Theorem 4. Let $\mathcal{J}_{t,i}$ denote the adjoint matrix of the ith diagonal element of $\mathcal{J}_t$. Then $\|e_i\|_{\mathcal{J}_t^{-1}} = \left(\frac{\det(\mathcal{J}_{t,i})}{\det(\mathcal{J}_t)}\right)^{\frac{1}{2}}$. To prove the theorem we need the following crucial lemma that characterizes the lower bound of the determinant of the information matrix.

Lemma. *Given $\alpha \geqslant 0$ and $\gamma > 0$, there exists some constant C such that for all T and corresponding T_0, we have*

$$\det(\mathcal{J}_{T_0}) \geqslant C \left(\sum_{s=1}^{T_0} s^{-2\alpha}\right) \left(\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}\right)^2. \tag{9}$$

Proof of Lemma. In the following proof, for simplicity of notation, we will ignore the T_0 in subscripts. Also, without loss of generality, we write l and u instead of l_0 and u_0. For any $x \geqslant -1$, $S_{x,t} \to +\infty$. Thus

$$\begin{aligned}
P_{x,t} &= l \sum_{s \text{ odd}} s^x + u \sum_{s \text{ even}} s^x = S_{x,t}\left(\frac{l+u}{2} + o(1)\right), \\
Q_{x,t} &= l^2 \sum_{s \text{ odd}} s^x + u^2 \sum_{s \text{ even}} s^x = S_{x,t}\left(\frac{l^2+u^2}{2} + o(1)\right).
\end{aligned}$$

Case 1: When $\alpha \leqslant \frac{1}{2}$, $2\gamma - 2\alpha > \gamma - 2\alpha > -2\alpha \geqslant -1$. Therefore, as $T_0 \to +\infty$, we have

$$\begin{aligned}
\det(\mathcal{J}_{T_0}) &\geqslant S_{-2\alpha}(S_{2\gamma-2\alpha}Q_{2\gamma-2\alpha} - P_{2\gamma-2\alpha}^2) - S_{2\gamma-2\alpha}P_{\gamma-2\alpha}^2 - Q_{2\gamma-2\alpha}S_{\gamma-2\alpha}^2 + 2S_{\gamma-2\alpha}P_{\gamma-2\alpha}P_{2\gamma-2\alpha} \\
&= S_{-2\alpha}\left(S_{2\gamma-2\alpha}^2\left(\frac{l^2+u^2}{2} + o(1)\right) - S_{2\gamma-2\alpha}^2\left(\frac{l+u}{2} + o(1)\right)^2\right) + \\
&\quad S_{2\gamma-2\alpha}S_{\gamma-2\alpha}^2\left(-\left(\frac{l+u}{2} + o(1)\right)^2 - \left(\frac{l^2+u^2}{2} + o(1)\right) + 2\left(\frac{l+u}{2} + o(1)\right)^2\right) \\
&= S_{2\gamma-2\alpha}(S_{-2\alpha}S_{2\gamma-2\alpha} - S_{\gamma-2\alpha}^2)\left(\left(\frac{u-l}{2}\right)^2 + o(1)\right) \\
&= S_{-2\alpha}S_{2\gamma-2\alpha}^2(1 - \frac{S_{\gamma-2\alpha}^2}{S_{-2\alpha}S_{2\gamma-2\alpha}})\left(\left(\frac{u-l}{2}\right)^2 + o(1)\right) \\
&\geqslant S_{-2\alpha}S_{2\gamma-2\alpha}^2\left(1 - \frac{\left(\frac{T_0^{\gamma-2\alpha+1}-1}{\gamma-2\alpha+1} + 1 + T_0^{\gamma-2\alpha}\right)^2}{\left(\frac{T_0^{1-2\alpha}-1}{1-2\alpha}\right)\left(\frac{T_0^{2\gamma-2\alpha+1}-1}{2\gamma-2\alpha+1}\right)}\right)\left(\left(\frac{u-l}{2}\right)^2 + o(1)\right) \\
&= \Omega(S_{-2\alpha}S_{2\gamma-2\alpha}^2).
\end{aligned}$$

Case 2: When $\alpha > \frac{1}{2}$, $S_{-2\alpha,t} < +\infty$, so we only need to prove that $\det\left(\mathcal{J}_{T_0}\right) \geqslant C\left(\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}\right)^2$. In fact, we have

$$\begin{aligned}
\det(\mathcal{J}_{T_0}) &= \det(\mathcal{J}_{T_0,0}) \det(\lambda + S_{-2\alpha} - [S_{\gamma-2\alpha}\ P_{\gamma-2\alpha}]\mathcal{J}_{T_0,0}^{-1}[S_{\gamma-2\alpha}\ P_{\gamma-2\alpha}]^\top) \\
&\geqslant \det(\mathcal{J}_{T_0,0})\lambda \\
&\geqslant \lambda\left(S_{2\gamma-2\alpha}Q_{2\gamma-2\alpha} - P_{2\gamma-2\alpha}^2\right) \\
&= \lambda S_{2\gamma-2\alpha}^2\left(\frac{l^2+u^2}{2} + o(1) - \left(\frac{l+u}{2} + o(1)\right)^2\right) = \Omega(S_{2\gamma-2\alpha}^2).
\end{aligned}$$

□

Let $\delta = 1/S_{\gamma,T}$, by utilizing (8), we have

$$\frac{\det(\mathcal{J}_t)^{\frac{1}{2}}\det(\lambda I)^{-\frac{1}{2}}}{\delta} \leqslant S_{\gamma,T}(1+\frac{S_{-2\alpha,t}}{\lambda})^{\frac{1}{2}}\det(\mathcal{J}_{t,0}/\lambda)^{\frac{1}{2}} \leqslant \frac{1}{2}S_{\gamma,T}(1+\frac{S_{-2\alpha,t}}{\lambda})^{\frac{1}{2}}(1+S_{2\gamma-2\alpha,t}).$$

In the following, we consider the case where $\theta \in \mathcal{C}_t$ for all t, which holds w.p. at least $1-\delta$.

From (9), for any $t > T_0$, we have

$$\begin{aligned}
|\theta(i) - \theta_t(i)|^2 &\leqslant \|e_i\|^2_{\mathcal{J}_t^{-1}} O(\log t) \leqslant \|e_i\|^2_{\mathcal{J}_{T_0}^{-1}} O(\log t) \\
&= \frac{\det\left(\mathcal{J}_{T_0,i}\right)}{\det\left(\mathcal{J}_{T_0}\right)} O(\log t) = O\left(\frac{\det(\mathcal{J}_{T_0,i})\log t}{\left(\sum_{s=1}^{T_0} s^{-2\alpha}\right)\left(\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}\right)^2}\right).
\end{aligned}$$

Also, we have

$$\begin{aligned}
\det(\mathcal{J}_{T_0,0}) &\leqslant O\left(\left(\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}\right)^2\right), \\
\det(\mathcal{J}_{T_0,1}) &\leqslant O\left(\left(\sum_{s=1}^{T_0} s^{-2\alpha}\right)\left(\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}\right)\right), \\
\det(\mathcal{J}_{T_0,2}) &\leqslant O\left(\left(\sum_{s=1}^{T_0} s^{-2\alpha}\right)\left(\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}\right)\right).
\end{aligned}$$

This implies that

$$\begin{aligned}
R_\theta^\pi(T) &\leqslant b_{\max}(u-l)^2\sum_{t=1}^{T_0} t^\gamma + b_{\max}T^\gamma \sum_{t=T_0+1}^{T} \|\phi(\theta) - \phi(\vartheta_t)\|_2^2 + \sum_{t=T_0+1}^{T}\frac{b_{\max}(u-l)^2 t^\gamma}{\sum_{s=1}^{T} s^\gamma} \\
&= O(T^{\frac{(\gamma+1)^2}{3\gamma-2\alpha+2}}) + b_{\max}T^\gamma \sum_{t=T_0+1}^{T}\sum_{i=1}^{3}\sup_{\theta'\in\Theta}\left(\frac{\partial\phi(\theta')}{\partial\theta'(i)}\right)^2 |\theta(i) - \vartheta_t(i)|^2 \\
&= O(T^{\frac{(\gamma+1)^2}{3\gamma-2\alpha+2}}) + T^\gamma\sum_{t=T_0+1}^{T} O\left(\frac{t^{-2\gamma}\log t}{\sum_{s=1}^{T_0} s^{-2\alpha}}\right) + O\left(\frac{\log t}{\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}}\right) + O\left(\frac{\log t}{\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}}\right) \\
&\leqslant O(T^{\frac{(\gamma+1)^2}{3\gamma-2\alpha+2}}) + T^\gamma\sum_{t=T_0+1}^{T} O\left(\frac{1}{\sum_{s=1}^{T_0} s^{2\gamma-2\alpha}}\right)\log T \\
&= O(T^{\frac{(\gamma+1)^2}{3\gamma-2\alpha+2}}\log T).
\end{aligned}$$

□

B. Proofs for Section 4

Proof of Theorem 5. In the following proof, we assume that the reward of each arm in the first period is a Gaussian variable with variance σ^2. We consider $\theta_1 = (\delta, 0, ..., 0)$ and $\theta_2 = (\delta, 0, ..., 0, 2\delta, 0, ..., 0)$ with $\delta > 0$ to be determined later, where

$$\theta_2(j) = \begin{cases} 2\delta, & \text{if } j = \arg\min_{i>1}\{\sum_{t=1}^T t^{2\gamma-2\alpha}\mathbb{P}^\pi_{\theta_1}(A_t = i)\} \triangleq i^*, \\ \theta_1(j), & \text{otherwise} \end{cases}$$

We have $R^\pi_{\theta_1}(T) = \mathbb{E}^\pi_{\theta_1}\left[\delta\sum_{t=1}^T \gamma^t \mathbb{1}\{A_t \neq 1\}\right]$ and $R^\pi_{\theta_2}(T) \geqslant \mathbb{E}^\pi_{\theta_2}\left[\delta\sum_{t=1}^T \gamma^t \mathbb{1}\{A_t = 1\}\right]$. Denoting $\mathrm{D}(\mathbb{P}, \mathbb{Q})$ as the KL divergence of two probability measures $\mathbb{P}$ and $\mathbb{Q}$, we thus have

$$\begin{aligned} R^\pi_{\theta_1}(T) + R^\pi_{\theta_2}(T) &\geqslant \delta\sum_{t=1}^T t^\gamma\left(\mathbb{P}^\pi_{\theta_1}(A_t \neq 1) + \mathbb{P}^\pi_{\theta_2}(A_t = 1)\right) \\ &\geqslant \frac{1}{2}\delta\sum_{t=1}^T t^\gamma \exp(-\mathrm{D}(\mathbb{P}^\pi_{\theta_1}, \mathbb{P}^\pi_{\theta_2})) \\ &= \frac{1}{2}\delta\sum_{t=1}^T t^\gamma \exp(-\mathbb{E}^\pi_{\theta_1}[\sum_{s=1}^T \mathrm{D}(\mathbb{P}_{\theta_1(A_s)}, \mathbb{P}_{\theta_2(A_s)})]) \\ &= \frac{1}{2}\delta\exp(-\sum_{t=1}^T \mathbb{P}^\pi_{\theta_1}(A_t = i^*)t^{2\gamma-2\alpha}\frac{4\delta^2}{\sigma^2})\sum_{t=1}^T t^\gamma \\ &\geqslant \frac{\delta}{2}\exp(-\frac{4\delta^2\sum_{t=1}^T t^{2\gamma-2\alpha}}{(K-1)\sigma^2})\sum_{t=1}^T t^\gamma, \end{aligned}$$

where for the first equality we utilize a divergence decomposition for general action spaces (see Exercise 15.8 in Lattimore & Szepesvári, 2019), and for the second equality we apply a known result on the KL divergence between two Gaussian distributions with same variance. Now we let $\delta = \frac{\sqrt{K-1}\sigma}{2\sqrt{\sum_{t=1}^T t^{2\gamma-2\alpha}}}$, then we can obtain that the regret is no less than

$$\frac{e^{-1}}{4}\sqrt{K-1}\sigma\frac{\sum_{t=1}^T t^\gamma}{\sqrt{\sum_{t=1}^T t^{2\gamma-2\alpha}}} = \Omega(\sqrt{K}T^{\alpha+\frac{1}{2}}).$$

□

Proof of Theorem 6. Let $\beta = 1 + \frac{(\gamma+1-\alpha)(\gamma+1)}{\alpha+\frac{1}{2}} \geqslant \frac{3}{2}$. Fix the total time periods as T. Set $T_0 = T^{\frac{\alpha+\frac{1}{2}}{\gamma+1}} < T$. We split $[1, T]$ as $[1, T_0]$ and $(T_0, T]$. For $[1, T_0]$, the regret is at most

$$\sum_{t=1}^{T_0} t^\gamma = T_0^{\gamma+1} = O(T^{\alpha+\frac{1}{2}})$$

Now we focus on what happened after T_0. From Lemma 1, we know that for any time period t, if $A_t = i$, then $r(i) \notin \mathcal{C}(i)$ w.p. at most $2\exp(-\beta\log t) = 2t^{-\beta}$. Thus $r(i) \in \mathcal{C}(i)$ holds for all time periods $t > T_0$ and all $i \in [K]$ such that $A_t = i$ w.p. at least

$$1 - \sum_{t=T_0+1}^T 2t^{-\beta} \geqslant 1 - \int_{t=T_0}^{+\infty} 2t^{-\beta}dt = 1 - \frac{2}{\beta-1}T_0^{-(\beta-1)}.$$

If this does not hold, the expected regret is at most

$$\frac{2}{\beta-1}T_0^{-(\beta-1)}\sum_{s=1}^T s^\gamma = O(T^{\gamma+1-\frac{(\beta-1)(\alpha+\frac{1}{2})}{\gamma+1}}) = O(T^\alpha).$$

Now we assume that $r(i) \in \mathcal{C}(i)$ always holds for all time periods $t > T_0$ and $i \in [K]$ such that $A_t = i$. Since for any activated arm, we construct a new confidence interval once after we pull it, thus by assumption all reward parameters lie in all the confidence intervals we construct. Therefore, the optimal arm i^* is forever activated. Suppose that in some period $\max\{T_0, 2K\} < t \leqslant T$, we pull an arm i that is not optimal. Then within the round that contains t, the gap between $r(i)$ and the reward of the optimal arm $r(i^*)$ is bounded by the sum of the lengths of confidence intervals in the last round. Here, a "round" means pulling all arms in the activation set once. If not, then arm i or i^* must be eliminated before time period t. Suppose when the last round ends, the time period is $t - K \leqslant \tilde{t} < t$. Thus

$$\Delta_t \leqslant \sigma 2\sqrt{2\beta}\left(\frac{1}{\sqrt{\sum_{s\leqslant \tilde{t}, A_s = i} s^{2\gamma-2\alpha}}} + \frac{1}{\sqrt{\sum_{s\leqslant \tilde{t}, A_s = i^*} s^{2\gamma-2\alpha}}}\right)\log T$$

Next, we seek to bound $\sum_{s\leqslant \tilde{t}, A_s = i} s^{2\gamma-2\alpha}$ and $\sum_{s\leqslant \tilde{t}, A_s = i^*} s^{2\gamma-2\alpha}$. Consider the activation set $\mathcal{A}$ before the deactivation procedure in the last round. For any arm $i_1 \in \mathcal{A}$ and $j_1 \notin \mathcal{A}$, since we pull each arm in a fixed order, and in the last round we pull i_1 but do not pull j_1, we have

$$\sum_{s\leqslant \tilde{t}, A_s = j_1} s^{2\gamma-2\alpha} \leqslant \sum_{s\leqslant \tilde{t}, A_s = i_1} s^{2\gamma-2\alpha}.$$

For any two arms $i_1, i_2 \in \mathcal{A}$, since in all rounds i_1 is pulled either before or after i_2, we have

$$\begin{aligned}
&\sum_{s\leqslant \tilde{t}, A_s = i_1} s^{2\gamma-2\alpha} - \sum_{s\leqslant \tilde{t}, A_s = i_2} s^{2\gamma-2\alpha} \\
&= \max_{s\leqslant \tilde{t}, A_s = i_1} s^{2\gamma-2\alpha} + \left(\sum_{s\leqslant \tilde{t}, A_s = i_1} s^{2\gamma-2\alpha} - \max_{s\leqslant \tilde{t}, A_s = i_1} s^{2\gamma-2\alpha}\right) - \sum_{s\leqslant \tilde{t}, A_s = i_2} s^{2\gamma-2\alpha} \\
&\leqslant \max_{s\leqslant \tilde{t}, A_s = i_1} s^{2\gamma-2\alpha} \leqslant \tilde{t}^{2\gamma-2\alpha}.
\end{aligned}$$

Thus,

$$\sum_{s\leqslant \tilde{t}, A_s = i} s^{2\gamma-2\alpha} \geqslant \frac{1}{|\mathcal{A}|}\left(\sum_{s\leqslant \tilde{t}, s\in\mathcal{A}} s^{2\gamma-2\alpha} - (|\mathcal{A}|-1)\tilde{t}^{2\gamma-2\alpha}\right) \geqslant \frac{1}{K}\sum_{s\leqslant \tilde{t}} s^{2\gamma-2\alpha} - \tilde{t}^{2\gamma-2\alpha}.$$

The inequality above is also valid when we replace i with i^*. Since i and i^* are still activated after the last round, then we must have $i, i^* \in \mathcal{A}$. Therefore,

$$\begin{aligned}
\Delta_t &\leqslant O\left(\frac{\sqrt{K}\log T}{\sqrt{\sum_{s\leqslant \tilde{t}} s^{2\gamma-2\alpha} - K\tilde{t}^{2\gamma-2\alpha}}}\right) \\
&= O\left(\frac{\sqrt{K}\log T}{\sqrt{\sum_{s\leqslant t} s^{2\gamma-2\alpha}}}\right)\left(\sqrt{\frac{\sum_{s\leqslant t} s^{2\gamma-2\alpha}}{\sum_{s\leqslant \tilde{t}} s^{2\gamma-2\alpha} - K\tilde{t}^{2\gamma-2\alpha}}}\right) \\
&\leqslant \sqrt{K} O(t^{\alpha-\gamma-\frac{1}{2}}\log T)\sqrt{\frac{1 + \frac{\sum_{\tilde{t}<s\leqslant t} s^{2\gamma-2\alpha}}{\sum_{s\leqslant \tilde{t}} s^{2\gamma-2\alpha}}}{1 - K\frac{\tilde{t}^{2\gamma-2\alpha}}{\sum_{s\leqslant \tilde{t}} s^{2\gamma-2\alpha}}}} \\
&\leqslant \sqrt{K} O(t^{\alpha-\gamma-\frac{1}{2}})\sqrt{\frac{1 + K\frac{t^{2\gamma-2\alpha}}{\sum_{s\leqslant t-K} s^{2\gamma-2\alpha}}}{1 - K\frac{t^{2\gamma-2\alpha}}{\sum_{s\leqslant t-K} s^{2\gamma-2\alpha}}}}\log T
\end{aligned}$$

Since $\lim_{t\to+\infty} \frac{t^{2\gamma-2\alpha}}{\sum_{s\leqslant t-K} s^{2\gamma-2\alpha}} = 0$, there exists T_1 such that for all $T_0 \geqslant T_1$, we have $\frac{T_0^{2\gamma-2\alpha}}{\sum_{s\leqslant T_0-K} s^{2\gamma-2\alpha}} \leqslant \frac{1}{2K}$. As a result, for all $T \geqslant T_1^{\frac{\gamma+1}{\alpha+\frac{1}{2}}}$, we have

$$R_\theta^\pi(T) \leqslant \sum_{t=1}^{T_0} t^\gamma + O(T^\alpha) + \sum_{t=T_0+1}^{T} t^\gamma \sqrt{K} O(t^{\alpha-\gamma-\frac{1}{2}}) \log T$$
$$\leqslant O(\sqrt{K} T^{\alpha+\frac{1}{2}} \log T).$$

□

References

Gill, R. D., Levit, B. Y., et al. Applications of the van trees inequality: a bayesian cramér-rao bound. *Bernoulli*, 1(1-2): 59–79, 1995.

Keskin, N. B. and Zeevi, A. Dynamic pricing with an unknown demand model: Asymptotically optimal semi-myopic policies. *Operations Research*, 62(5):1142–1167, 2014.

Lattimore, T. and Szepesvári, C. *Bandit algorithms*. Cambridge University Press (preprint), 2019.