A Bayesian Approach for Bandit Online Optimization with Switching Cost (Supplementary Material)

Zai Shi[1] **Jian Tan**[1] **FeiFei Li**[1]

[1]Alibaba Group, Hangzhou, Zhejiang, China

1 PROOF OF THEOREM 1

Defining $v_t^* = \arg\min_{x\mathcal{X}} f(x) + c(x, x_{t-1})$ and $x_0^* = x_0$, we have

$$
\begin{aligned}
&\sum_{t=1}^{T}(f(x_t) + c(x_t, x_{t-1}) - f(v_t^*) - c(v_t^*, x_{t-1})) \\
&\geq \sum_{t=1}^{T}(f(x_t) + c(x_t, x_{t-1}) - f(x_t^*) - c(x_t^*, x_{t-1})) && (1) \\
&\geq \sum_{t=1}^{T}(f(x_t) + c(x_t, x_{t-1}) - f(x_t^*) - \eta c(x_t^*, x_{t-1}^*) - \eta c(x_{t-1}^*, x_{t-1})) && (2) \\
&= \sum_{t=1}^{T}(f(x_t) + c(x_t, x_{t-1}) - f(x_t^*) - \eta c(x_t^*, x_{t-1}^*) - \eta c(x_t^*, x_t)) + \eta c(x_T^*, x_T) - \eta c(x_0^*, x_0) \\
&\geq \sum_{t=1}^{T}(f(x_t) + c(x_t, x_{t-1}) - f(x_t^*) - \eta c(x_t^*, x_{t-1}^*) - \eta c(x_t^*, x_t)) - \eta D && (3) \\
&\geq \sum_{t=1}^{T}(f(x_t) + c(x_t, x_{t-1}) - f(x_t^*) - \eta c(x_t^*, x_{t-1}^*) - \eta^2 c(x_t^*, x^*) - \eta^2 c(x_t, x^*)) - \eta D && (4) \\
&\geq \sum_{t=1}^{T}((1-\eta^2/\lambda) f(x_t) + c(x_t, x_{t-1}) - (1+\eta^2/\lambda) f(x_t^*) + 2\eta^2/\lambda f(x^*) - \eta c(x_t^*, x_{t-1}^*)) - \eta D && (5) \\
&\geq \sum_{t=1}^{T}((1-\eta^2/\lambda) f(x_t) + (1-\eta^2/\lambda) c(x_t, x_{t-1}) - (1+\eta^2/\lambda) f(x_t^*) - \eta c(x_t^*, x_{t-1}^*)) - \eta D && (6)
\end{aligned}
$$

where (1) is from the definition of v_t^*, (2) is from Assumption 3, (3) is from Assumption 4, (4) is from Assumption 3, (5) is from Assumption 2 and (6) is from $f(x) \geq 0$ and $c(x, y) \geq 0$ within $\mathcal{X}$.

Meanwhile, following a similar approach to the proof of Lemma 1 based on Lemma 2 of the main paper, we have

$$
\sum_{t=1}^{T}(f(x_t) + c(x_t, x_{t-1}) - f(v_t^*) - c(v_t^*, x_{t-1})) = O\left(B\sqrt{T\gamma_T'} + \sqrt{T\gamma_T'(\gamma_T' + \log(1/\delta))}\right) \quad (7)
$$

with probability at least $1 - \delta$, where γ_T' is the maximal information gain related to kernel k' and the domain $\mathcal{X} \times \mathcal{X}$ in $2d$-dimension.

Accepted for the 39th Conference on Uncertainty in Artificial Intelligence (UAI 2023).

Since $f(x) \geq 0$ and $c(x, y) \geq 0$ within $\mathcal{X}$, by combining (6) and (7) with $\eta^2/\lambda < 1$ we have

$$\sum_{l=1}^{T} (f(x_l) + c(x_l, x_{l-1})) - \sum_{l=1}^{T} \psi \left(f(x_l^*) + c(x_l^*, x_{l-1}^*) \right) = O\left(B\sqrt{T\gamma_T'} + \sqrt{T\gamma_T'(\gamma_T' + \log(1/\delta))} \right) = \tilde{O}(T^{g(2d)})$$

with probability at least $1 - \delta$, where $\psi = \max\{\frac{1+\eta^2/\lambda}{1-\eta^2/\lambda}, \frac{\eta}{1-\eta^2/\lambda}\}$.

2 PROOF OF COROLLARY 1

From the definition of competitive ratio, we have

$$\begin{aligned} CR &= \frac{\sum_{t=1}^{T} f_t(x_t) + c(x_t, x_{t-1})}{\sum_{t=1}^{T} f_t(x_t^*) + c(x_t^*, x_{t-1}^*)} \\ &\leq \frac{\sum_{t=1}^{T} \psi \left(f(x_t^*) + c(x_t^*, x_{t-1}^*) \right) + \tilde{O}(T^{g(2d)})}{\sum_{t=1}^{T} f_t(x_t^*) + c(x_t^*, x_{t-1}^*)} && (8) \\ &\leq \psi + \frac{\tilde{O}(T^{g(2d)})}{TC} && (9) \\ &= \psi + \tilde{O}(T^{g(2d)-1}), \end{aligned}$$

where (8) is from Theorem 1 and $f(x) \geq 0, c(x, y) \geq 0$ within $\mathcal{X}$, (9) is from Assumption 4 and the assumption in Corollary 1.

3 PROOF OF THEOREM 2

For epoch m, define $v_m^* = \arg\min_{x \in \mathcal{X}} f(x) + 2\eta c(x, v_m)$ and $x_0^* = x_0$, where v_m is the pivot point of epoch m defined in Algorithm 3. For epoch m, the time horizon is from $l = \frac{m(m-1)}{2} + 1$ to $l = \frac{m(m-1)}{2} + m$. Defining $t = \frac{m(m-1)}{2} + m$ and $x^* = \arg\min_{x \in \mathcal{X}} f(x)$, we have

$$\sum_{l=\frac{m(m-1)}{2}+1}^{t-1} (f(x_l) + 2\eta c(x_l, v_m) - f(v_m^*) - 2\eta c(v_m^*, v_m)) + f(x_t) - f(x^*) \tag{10}$$

$$\geq \sum_{l=\frac{m(m-1)}{2}+1}^{t} (f(x_l) + c(x_l, x_{l-1})) - \eta c(x_t, v_m) - \eta c(x_{\frac{m(m-1)}{2}}, v_m) - \sum_{l=\frac{m(m-1)}{2}+1}^{t-1} (f(v_m^*) + 2\eta c(v_m^*, v_m)) - f(x^*) \tag{11}$$

$$\geq \sum_{l=\frac{m(m-1)}{2}+1}^{t} (f(x_l) + c(x_l, x_{l-1})) - \sum_{l=\frac{m(m-1)}{2}+1}^{t-1} (f(x_l^*) + 2\eta c(x_l^*, v_m)) - 2\eta D - f(x_t^*) \tag{12}$$

$$\begin{aligned} &\geq \sum_{l=\frac{m(m-1)}{2}+1}^{t} (f(x_l) + c(x_l, x_{l-1})) - \sum_{l=\frac{m(m-1)}{2}+1}^{t-1} \left(f(x_l^*) + 2\eta^2 (c(x_l^*, x_{l-1}^*) + \eta c(x_{l-1}^*, x^*) + \eta c(v_m, x^*)) \right) \\ &- 2\eta D - f(x_t^*) \end{aligned} \tag{13}$$

$$\begin{aligned} &\geq \sum_{l=\frac{m(m-1)}{2}+1}^{t} (f(x_l) + c(x_l, x_{l-1})) - \sum_{l=\frac{m(m-1)}{2}+1}^{t} \left((1 + 2\eta^3/\lambda) f(x_l^*) + 2\eta^2 c(x_l^*, x_{l-1}^*) \right) - 2\eta^3 c(x_{\frac{m(m-1)}{2}}, x^*) \\ &+ 2\eta^2 c(x_t^*, x_{t-1}^*) + 2\eta^3 c(x_{t-1}^*, x^*) + 2\eta^3 c(x_t^*, x^*) - 2\eta^3/\lambda(m-1)(f(v_m) - 2f(x^*)) - 2\eta D \end{aligned} \tag{14}$$

$$\begin{aligned} &\geq \sum_{l=\frac{m(m-1)}{2}+1}^{t} (f(x_l) + c(x_l, x_{l-1})) - \sum_{l=\frac{m(m-1)}{2}+1}^{t} \left((1 + 2\eta^3/\lambda) f(x_l^*) + 2\eta^2 c(x_l^*, x_{l-1}^*) \right) \\ &- (2\eta + 2\eta^3) D - 2\eta^3/\lambda(m-1)(f(v_m) - f(x^*)) \end{aligned} \tag{15}$$

where (11) is from $\eta c(x_l, v_m) + \eta c(x_{l-1}, v_m) \geq c(x_l, x_{l-1})$ using Assumption 3 , (12) is from Assumption 4 and the definition of v_m^* and x^*, (13) is from using Assumption 3 for $c(x_l^*, v_m)$ and $c(x_{l-1}^*, v_m)$, (14) is from Assumption 2 and (15) is from Assumption 2 and 4.

Now we want to bound $(m-1)(f(v_m) - f(x^*))$ in (15) for $m > 1$ (it is 0 when $m = 1$). By the definition of v_m in Algorithm 3, we have

$$f(v_m) + \varepsilon_{v_m} - f(x_{\frac{n(n+1)}{2}}) - \varepsilon_{\frac{n(n+1)}{2}} \leq 0, \forall n \in \{1, ..., m-1\}$$

where ε_{v_m} is the observation noise when observing $f(v_m)$ and $\varepsilon_{x_{\frac{n(n+1)}{2}}}$ is the observation noise when observing $f(x_{\frac{n(n+1)}{2}})$. Note that $x_{\frac{n(n+1)}{2}}$ is the point chosen by UE in epoch n. Therefore, for any $\epsilon > 0$ and $m > 1$, if

$$\sum_{n=1}^{m-1} [\varepsilon_{\frac{n(n+1)}{2}} - \varepsilon_{v_m}] \leq \epsilon,$$

we have

$$(m-1)(f(v_m) - f(x^*)) - \sum_{n=1}^{m-1} (f(x_{\frac{n(n+1)}{2}}) - f(x^*)) \leq \epsilon$$

Then for $m > 1$ we have

$$\begin{aligned}
&\mathbb{P}[(m-1)(f(v_m) - f(x^*)) - \sum_{n=1}^{m-1} (f(x_{\frac{n(n+1)}{2}}) - f(x^*)) \leq \epsilon] \\
&\geq \mathbb{P}[\sum_{n=1}^{m-1} [\varepsilon_{\frac{n(n+1)}{2}} - \varepsilon_{v_m}] \leq \epsilon] \geq 1 - \exp(-\frac{\epsilon^2}{4(m-1)R})
\end{aligned}$$

since the observation noise is R-subGaussian from Assumption 1. As a result,

$$(m-1)(f(v_m) - f(x^*)) - \sum_{n=1}^{m-1} (f(x_{\frac{n(n+1)}{2}}) - f(x^*)) \leq \sqrt{4(m-1)R \log \frac{M-1}{\delta}} \tag{16}$$

with probability at least $1 - \delta/(M-1)$.

For UE across m epochs, we can regard them as running IGP-UCB on $f(x)$ for m iterations. Therefore, we have

$$\sum_{n=1}^{m-1} (f(x_{\frac{n(n+1)}{2}}) - f(x^*)) = O\left(B\sqrt{(m-1)\gamma_{m-1}} + \sqrt{(m-1)\gamma_{m-1}(\gamma_{m-1} + \log(1/\delta))}\right) \tag{17}$$

holds together for all choices of $m \in \{2, ..., M+1\}$ with probability at least $1 - \delta$ from Lemma 1. Combined with (16), we have

$$(m-1)(f(v_m) - f(x^*)) \leq \sqrt{4(m-1)R \log \frac{M-1}{\delta}} + O\left(B\sqrt{(m-1)\gamma_{m-1}} + \sqrt{(m-1)\gamma_{m-1}(\gamma_{m-1} + \log(1/\delta))}\right) \tag{18}$$

holds together for all choices of $m \in \{2, ..., M\}$ with probability at least $1 - 2\delta$, which finishes the bound of $(m-1)(f(v_m) - f(x^*))$.

Meanwhile, LE in epoch m can be regarded as running IGP-UCB on $f(x) + 2\eta c(x, v_m)$ for $m - 1$ iterations. Then from Lemma 1,

$$\begin{aligned}
&\sum_{l=\frac{m(m-1)}{2}+1}^{\frac{m(m+1)}{2}-1} (f(x_l) + 2\eta c(x_l, v_m) - f(v_m^*) - 2\eta c(v_m^*, v_m)) \\
&= O\left(B\sqrt{(m-1)\gamma_{m-1}} + \sqrt{(m-1)\gamma_{m-1}(\gamma_{m-1} + \log(M/\delta))}\right)
\end{aligned} \tag{19}$$

with probability at least $1-\delta/M$. From (17) and (19), we have

$$\sum_{m=1}^{M}[\sum_{l=\frac{m(m-1)}{2}+1}^{\frac{m(m+1)}{2}-1}(f(x_l)+2\eta c(x_l,v_m)-f(v_m^*)-2\eta c(v_m^*,v_m))+f(x_{\frac{m(m+1)}{2}})-f(x^*)]$$
$$=\sum_{m=1}^{M}O\left(B\sqrt{m\gamma_m}+\sqrt{m\gamma_m(\gamma_m+\log(M/\delta))}\right) \tag{20}$$

with probability at least $1-2\delta$, which gives the bound for the telescoping of (10) from $m=1$ to M. Combining it with (15) and (18), we have

$$\sum_{m=1}^{M}[\sum_{l=\frac{m(m-1)}{2}+1}^{\frac{m(m-1)}{2}+m}(f(x_l)+c(x_l,x_{l-1}))-\sum_{l=\frac{m(m-1)}{2}+1}^{\frac{m(m-1)}{2}+m}\left((1+2\eta^3/\lambda)f(x_l^*)+2\eta^2c(x_l^*,x_{l-1}^*)\right)]$$
$$=\sum_{t=1}^{T}(f(x_l)+c(x_l,x_{l-1}))-\sum_{t=1}^{T}\left((1+2\eta^3/\lambda)f(x_l^*)+2\eta^2c(x_l^*,x_{l-1}^*)\right)$$
$$=\sum_{m=1}^{M}O\left(B\sqrt{m\gamma_m}+\sqrt{m\gamma_m(\gamma_m+\log(M/\delta))}\right)$$

with probability at least $1-3\delta$ since (20) and (18) both use the event (17). From Assumption 2 and 4, we know that f and c are positive within $\mathcal{X}$. Set $\psi=\max\{1+2\eta^3/\lambda,2\eta^2\}$ and recall that $B\sqrt{m\gamma_m}+\sqrt{m}\gamma_m=m^{g(d)}$. Then since f and c are both positive within $\mathcal{X}$ and $T=1+2+...+M=\frac{M(M+1)}{2}$, we have

$$\sum_{l=1}^{T}(f(x_l)+c(x_l,x_{l-1}))-\sum_{l=1}^{T}\psi\left(f(x_l^*)+c(x_l^*,x_{l-1}^*)\right)=\tilde{O}(T^{(g(d)+1)/2}).$$

with probability at least $1-3\delta$.

4 EXPERIMENT DETAILS OF SECTION 6

4.0.1 Robot Pushing Problem

The original 14-dimensional robot pushing problem was first tested in Wang et al. [2018] without switching cost, where the authors implemented the simulation of pushing two objects with two robot hands in the Box2D physics engine. The original code is available at https://github.com/zi-w/Ensemble-Bayesian-Optimization. In this problem, we need to choose 14 control parameters that determine the location and rotation of the robot hands, pushing speed, moving direction and pushing time. The lower limit of these parameters is $[-5,-5,-10,-10,2,0,-5,-5,-10,-10,2,0,-5,-5]$ and the upper limit is $[5,5,10,10,30,2\pi,5,5,10,10,30,2\pi,5,5]$. Denote the initial positions of the objects i_0,i_1, the ending positions of the objects e_0,e_1 and the goal locations of the objects g_0,g_1, respectively. Then the reward is defined to be $r=||g_0-i_0||+||g_1-i_1||-||g_0-e_0||-||g_1-e_1||$, which is the progress made towards pushing the objects to the goal.

4.0.2 Lunar Lander Problem

The original 12-dimensional lunar lander problem in Eriksson et al. [2019] is to learn a controller for a lunar lander to minimize fuel consumption and the distance to a landing target, while also preventing crashes. The original code is available in the OpenAI gym: https://github.com/openai/gym. The 12 controllable parameters of the lander include its angle and position, their respective time derivatives, and so on. We use $[0,2]$ as the limit for each of these parameters.

5 RUNNING TIME RESULTS OF EXPERIMENTS IN SECTION 6

In this section, we show the running time of AS and GS for experiments in Section 6. For each algorithm, we run its 10 tests concurrently in a sever with 2.7GHz Intel(R) Xeon(R) Platinum CPU including 16 processors, and 30GB RAM. We list the maximum, minimum and the average running time of 10 tests of each algorithm in Table 1. For fairness, we just use the naive implementation of each algorithm without any empirical acceleration techniques proposed in previous BO works.

		Robot Pushing	Lunar Lander
Greedy Search	Avg	107993.2	49316.1
	Max	127680	67624
	Min	94107	31143
Alternating Search	Avg	55984.5	27973.2
	Max	67893	30840
	Min	44272	14680

Table 1: Running time of GS and AS for robot pushing and lunar lander problem. The time unit is second.

References

David Eriksson, Michael Pearce, Jacob Gardner, Ryan D Turner, and Matthias Poloczek. Scalable global optimization via local bayesian optimization. *Advances in neural information processing systems*, 32, 2019.

Zi Wang, Clement Gehring, Pushmeet Kohli, and Stefanie Jegelka. Batched large-scale bayesian optimization in high-dimensional spaces. In *International Conference on Artificial Intelligence and Statistics*, pages 745–754. PMLR, 2018.