Supplementary material for 'Adaptive Sensor Placement for Continuous Spaces'

James A. Grant, Alexis Boukouvalas, Ryan-Rhys Griffiths, David S. Leslie,
Sattar Vakili, Enrique Munoz de Cote

May 9, 2019

1 Regret bound proofs

Proof of Lemma 1

Define $A_{\min,t} = \bigcap_{A \in \mathcal{A}_t : A^* \subseteq A} A$ as the smallest interval (or union of intervals) in $\mathcal{A}_t$ containing the optimal interval (or union of intervals). It will be easier to bound the regret of $A_{\min,t}$ than A_t^* wrt A^*. We have, for $t \in \mathbb{N}$,

$$
\begin{aligned}
\delta(A_t^*) &= r(A^*) - r(A_t^*) \\
&\leq r(A^*) - r(A_{\min,t}) \\
&= \int_{A^*} (\lambda(x) - C)\,\mathrm{d}x - \int_{A_{\min,t}} (\lambda(x) - C)\,\mathrm{d}x \\
&= C|A_{\min,t} \setminus A^*| - \int_{A_{min,t} \setminus A^*} \lambda(x) dx \\
&\leq 2CU\Delta_t.
\end{aligned}
$$

Here, the final inequality holds since $2\Delta_t$ bounds the difference between the lengths of subintervals of $A_{min,t}$ and A_t^*, and there are U such subintervals. Since $\Delta_t = K_t^{-1} \leq \underline{K}^{-1} T^{-1/3}$ the result follows immediately.

Proof of Lemma 2

Consider the term inside the expectation

$$
\begin{aligned}
\sum_{t=1}^{T} U_{t,T}(A_t) - L_{t,T}(A_t) &= 2\Delta_T \sum_{t=1}^{T} \sum_{k:B_{k,T}\subseteq A_t} D_{k,T}(t-1) \\
&= 2\Delta_T \sum_{t=1}^{T} \sum_{k:B_{k,T}\subseteq A_t} \frac{2\log(t)}{\Delta_T N_{k,T}(t-1)} + \sqrt{\frac{6\lambda_{\max}\log(t)}{\Delta_T N_{k,T}(t-1)}} \\
&= 2\Delta_T \sum_{t=1}^{T} \sum_{k=1}^{K_T} \mathbb{I}\{B_{k,T}\subseteq A_t\}\left(\frac{2\log(t)}{\Delta_T \sum_{s=1}^{t-1}\mathbb{I}\{B_{k,T}\subseteq A_s\}} + \sqrt{\frac{6\lambda_{\max}\log(t)}{\Delta_T\sum_{s=1}^{t-1}\mathbb{I}\{B_{k,T}\subseteq A_s\}}}\right) \\
&\leq 2\Delta_T \sum_{k=1}^{K_T}\sum_{j=1}^{N_{k,T}} \frac{2\log(T)}{j\Delta_T} + \sqrt{\frac{6\lambda_{\max}\log(T)}{j\Delta_T}} \\
&\leq 2\Delta_T K_T \left(\sum_{j=1}^{T}\frac{2\log(T)}{j\Delta_T} + \sum_{j=1}^{T}\sqrt{\frac{6\lambda_{\max}\log(T)}{j\Delta_T}}\right) \\
&= 4K_T\log(T)\log(T+1) + \sqrt{24\lambda_{\max}K_T\log(T)}T^{1/2} \\
&\leq 4\overline{K}\log(T)\log(T+1)T^{1/3} + \sqrt{24\overline{K}\lambda_{\max}\log(T)}T^{2/3}
\end{aligned}
$$

where the penultimate line is due to $\Delta_T = K_T^{-1}$, and the final inequality is because $K_T \leq \overline{K}T^{1/3}$.

Proof of Lemma 3

We have the following, which holds for any round t

$$
\begin{aligned}
&P\Big(r(A_t) \notin [L_{t,T}(A_t), U_{t,T}(A_t)]\Big) \\
&\leq P\Big(r(A_t) \leq L_{t,T}(A_t)\Big) + P\Big(r(A_t) \geq U_{t,T}(A_t)\Big) \\
&= P\Big(\sum_{k:B_{k,T}\subseteq A_t} \psi_{k,T} \leq \sum_{k:B_{k,T}\subseteq A_t}\Big[\hat{\psi}_{k,T}(t-1) - D_{k,T}(t-1)\Big]\Big) \\
&\quad + P\Big(\sum_{k:B_{k,T}\subseteq A_t} \psi_{k,T} \geq \sum_{k:B_{k,T}\subseteq A_t}\Big[\hat{\psi}_{k,T}(t-1) + D_{k,T}(t-1)\Big]\Big) \\
&\leq \sum_{k:B_{k,T}\subseteq A_t}\Big[P\Big(\psi_{k,T} - \hat{\psi}_{k,T}(t-1) \leq -D_{k,T}(t-1)\Big) + P\Big(\psi_{k,T} - \hat{\psi}_{k,T}(t-1) \geq D_{k,T}(t-1)\Big)\Big] \\
&\leq \sum_{k=1}^{K_T} P\left(|\psi_{k,T} - \hat{\psi}_{k,T}(t-1)| \geq \frac{2\log(t)}{\Delta_T N_{k,T}(t-1)} + \sqrt{\frac{6\lambda_{\max}\log(t)}{\Delta_T N_{k,T}(t-1)}}\right) \\
&\leq \sum_{k=1}^{K_T}\sum_{s=1}^{t-1} P\left(|\psi_{k,T} - \hat{\psi}_{k,T}(t-1)| \geq \frac{2\log(t)}{\Delta_T N_{k,T}(t-1)} + \sqrt{\frac{6\lambda_{\max}\log(t)}{\Delta_T N_{k,T}(t-1)}} \;\middle|\; N_{k,T}(t-1) = s\right) \leq 2K_T t^{-2}.
\end{aligned}
$$

The final inequality is a direct application of Lemma 1 of [Grant et al., 2018] which in turn exploits Bernstein's Inequality for independent Poisson random variables.

2 Proof of optimality and efficiency of AS-IM

Proof of Theorem 1

Recall that the reward of an action is the sum of the weights of the intervals that comprise that action.

We prove the theorem by induction. Assume at least one initial I_n has a positive weight (otherwise the optimal action is to do no sensing). For $N = 1$ initial interval, which therefore has a positive weight, AS-IM simply returns this interval, which is optimal. For $N = 2$ initial intervals, with one positive weight, AS-IM returns the postitively-weighted interval, which is the optimal action. Now, assuming AS-IM returns the optimal action for $N \geq 1$, we prove that AS-IM returns the optimal action for $N + 2$ initial intervals. The result follows by induction.

Given $\mathcal{I} = \{I_n\}_{n=1}^{N+2}$, if the number of intervals in $\mathcal{I}$ with positive weight is not bigger than U, AS-IM returns all such intervals. This is the optimal action since all bins with positive reward can be covered without incurring the cost of any bins with negative reward; any other action either omits a positive-reward bin, or includes a negative-reward bin.

Similarly, consider the situation in which no interval satisfies the merging condition. Suppose that the optimal action A^* places a sensor on a sequence of intervals $I_m \cup \cdots \cup I_n$ with $n > m$. Clearly we must have $w(I_m) > 0$ and $w(I_n) > 0$ since otherwise the total weight could be increased by omitting the negatively-weighted end interval. But the fact that no interval can be merged implies that either $|w(I_{m+1})| > |w(I_m)|$ or $|w(I_{n-1})| > |w(I_n)|$. Hence removing either $I_m \cup I_{m+1}$ or $I_{n-1} \cup I_n$ from the sensor will improve the total weight. It follows that, under A^*, each sensor is allocated to a single interval, and allocating to the U highest-weight intervals, as specified by AS-IM, maximises the reward.

Now, assume that at least one interval is merged in AS-IM. Let I_n be the interval which minimises $|w(I_n)|$ and so is the first interval which is merged with its neighbours in AS-IM into a single interval $\tilde{I}_n = I_{n-1} \cup I_n \cup I_{n+1}$. Let $\tilde{A}^*$ be AS-IM's solution for the set of intervals $\tilde{\mathcal{I}} = \{I_1, \cdots, I_{n-2}, \tilde{I}_n, I_{n+2}, \cdots, I_{N+2}\}$. By induction, $\tilde{A}^*$ is optimal for $\tilde{I}$. We prove that A^*, the optimal solution for $\mathcal{I}$, is equal to $\tilde{A}^*$. To prove this, we consider different cases based on the sign of $w(I_n)$.

Case 1: $w(I_n) < 0$**.** First note that the optimal solution cannot include only one neighbour of I_n. If I_{n-1} were included but I_{n+1} were not, we could add both I_n and I_{n+1} and increase the overall weight (since I_n has the smallest absolute weight). Similarly, A^* can not include both I_{n-1} and I_{n+1} but not I_n; if so then A^* could be improved by (i) using a single sensor in place of the two that cover I_{n-1} and I_{n+1}, adding I_n to A^*, and (ii) redeploying the sensor we have saved to either split one existing sensor by removing a negative-weight I_m with $|w(I_m)| > |w(I_n)|$, or adding a new positive-weight I_m with $|w(I_m)| > |w(I_n)|$. The net outcome is an improved total weight. We have shown that A^* includes either all or none of $I_{n-1} \cup I_n \cup I_{n+1}$. Since A^* is optimal for $\mathcal{I}$, and the restriction to $\tilde{\mathcal{I}}$ does not prevent AS-IM from finding this optimal A^*, it follows that $\tilde{A}^* = A^*$.

Case 2: $w(I_n) > 0$**.** Under the optimal solution A^*, a sensor cannot have a negative-weighted interval as an end interval, since dropping the negative-weight interval only increases the total weight. Furthermore, a sensor cannot include I_n as an end interval of a series of intervals, since then the total weight could be improved by stopping sensing both I_n and its sensed neighbour. Thus if I_n is included in A^* then either a sensor is observing only I_n, or a single sensor observes all of $I_{n-2} \cup I_{n-1} \cup I_n \cup I_{n+1} \cup I_{n+2}$. As in Case 1, if a sensor is observing only I_n we can improve on A^* by redeploying this sensor to either sense a better interval, or stop sensing an interval which has a higher negative weight than is lost by stopping sensing I_n. So again, under A^*, I_n is either sensed with all its neighbours, or none of them are sensed. The same logic as in Case 1 ensures $\tilde{A}^* = A^*$.

Complexity: AS-IM requires sorting the N initial intervals. Noticing that there are at most N mergings, and assuming constant complexity for each merging, AS-IM offers an $O(N \log N)$ sample complexity. Since $N \leq K_t$, AS-IM has a sample complexity not bigger than $O(K_t \log K_t)$.

3 Discretisation error under linear and cubic root rates

The effect of the different rates on the unavoidable discretisation error is depicted in Figure 1. The regret for the linear rate is reduced at a faster rate than for the cubic root rate as the number of bins is increased at a much faster rate. However as we show in the main paper (Section 5.1) the other part of the regret due to error in action selection from the model forecast is much higher under the linear regret rate.

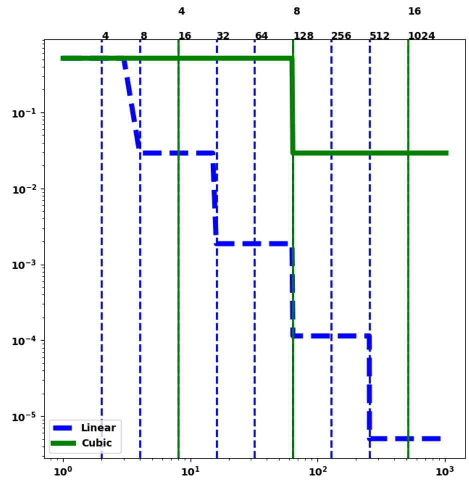

Figure 1: Instantaneous regret comparing linear and cube root rebinning rates. The vertical lines depict the rebinning times for the two different rate schedules. The time step (horizontal axis) and the regret (vertical axis) are both on a log scale. The number of bins for each rebinning rate are shown on the top horizontal axis.

4 Baselines used in the empirical study

In the paper we have compared the TS approach other approaches which we now describe in more details.

1. *UCB* approach, which is based on the FP-CUCB algorithm of [Grant et al., 2018] and requires the specification of an upper bound on the rate which we fix to the correct value in our experiments; in practise a conservative estimate is usually available. This is described in Algorithm 1.

2. A modified-UCB approach (*mUCB*) which has the same form as Algorithm 1 except $\lambda_{\max}$ is replaced with the empirical mean. Note this modification breaks the upper bound regret guarantee. The indices are :

$$\bar{\psi}_{k,t} = \hat{\psi}_{k,t}(t-1) + \frac{2\log(t)}{\Delta_t N_{k,t}(t-1)} + \sqrt{\frac{6\hat{\psi}_{k,t}(t-1)\log(t)}{\Delta_t N_{k,t}(t-1)}}, \;\; k \in [K_t]$$

where $\hat{\psi}_{k,t}(t-1) = \frac{H_{k,t}(t-1)}{\Delta_t N_{k,t}(t-1)}$.

Inputs: Upper bound $\lambda_{\max} \geq \max_{x \in [0,1]} \lambda(x)$
Initialisation Phase: For $t = 1$

- Select $A = [0, 1]$

Iterative Phase: For $t \geq 2$

- For each $k \in \{1, \ldots, K_t\}$, evaluate $H_{k,t}(t-1)$ and $N_{k,t}(t-1)$ and calculate an index

$$\bar{\psi}_{k,t} = \frac{H_{k,t}(t-1)}{\Delta_t N_{k,t}(t-1)} + \frac{2\log(t)}{\Delta_t N_{k,t}(t-1)} + \sqrt{\frac{6\lambda_{\max}\log(t)}{\Delta_t N_{k,t}(t-1)}}.$$

- Choose an action A_t that maximises $r(A)$ conditional on the true rate being given by the $\bar{\psi}_{k,t}$ values
- Observe the events in A_t

Algorithm 1: UCB

3. An ϵ-*Greedy* approach where with probability $1 - p_\epsilon$ an action A_t is selected that maximises $r(A)$ conditional on the rate being given by the empirical mean values $\hat{\psi}_{k,t}$. With probability p_ϵ, the action is instead chosen by sampling rates $\tilde{\psi}_{k,t}$ from independent $Gamma(\alpha, \beta)$ priors. In our experiments we fix $p_\epsilon = 0.01$.

References

[Grant et al., 2018] Grant, J.A., Leslie, D.S., Glazebrook, K., Szechtman, R. and Letchford, A. (2018): Adaptive policies for perimeter surveillance problems. arXiv preprint arXiv:1810.02176