Tightening Regret Lower and Upper Bounds in Restless Rising Bandits

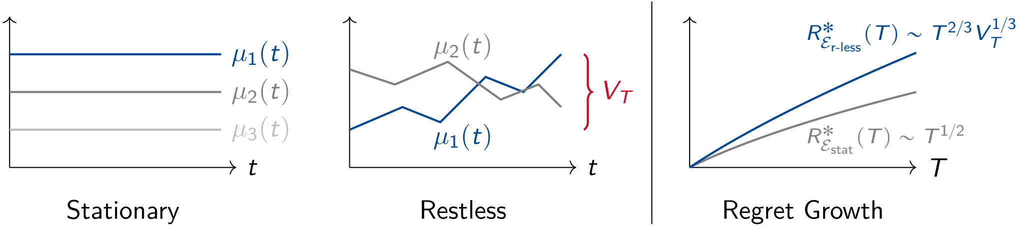
Cristiano Migali Marco Mussi Gianmarco Genalti Alberto Maria Metelli

Politecnico di Milano

The Thirty-Ninth Annual Conference on Neural Information Processing Systems (NeurIPS 2025)

Setting

- In the **standard** *Multi-Armed Bandit* setting, the expected rewards μ_i of the K available actions do not change with time.
- *Restless* (Tekin et al., 2012) (a.k.a. *non-stationary*) *bandits* **relax** this **assumption**, **increasing** the **complexity** of the setting.
- This is captured by the **order of growth** of the *minimax* **regret** $R^*_{\mathcal{E}}(T)$.



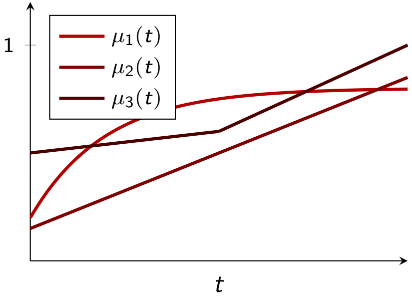
- This work studies the **restless rising** and **restless rising concave** classes of bandits.

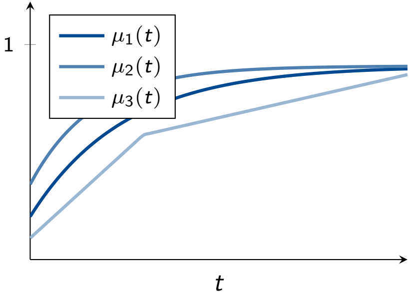
Non-Decreasing:

$$\gamma_i(t) \coloneqq \mu_i(t+1) - \mu_i(t) \geqslant 0$$

Non-Decreasing: $\gamma_i(t) \geqslant 0$

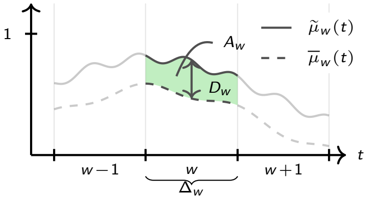
Concave: $\gamma_i(t) \geqslant \gamma_i(t+1)$





General Recipe

$$R^*(T) = \Omega\left(\sum_{w=1}^{W}\left(1-\sqrt{\frac{D_w}{K}}\right)A_w\right)$$

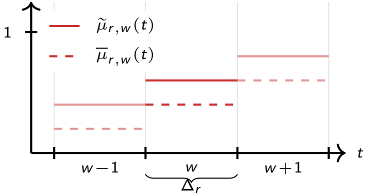


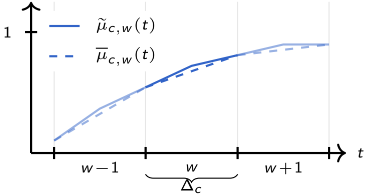
Restless Rising

$$R^*(T) = \Omega\left(T^{\frac{2}{3}}\min\{1, V_T\}^{\frac{1}{3}}\right)$$

Restless Rising Concave

$$R^*(T) = \Omega\left(T^{\frac{3}{5}}\min\{1, V_T\}^{\frac{1}{5}}\right)$$





Algorithm for Restless Rising Concave Bandits

Algorithm $\texttt{RC-BE}(\alpha)$.

Let $w \leftarrow 1$ be the **window index**

RESTART:

Let $\mathcal{A} \leftarrow [\![K]\!]$ be the set of **alive arms**,

$d \leftarrow 1$ the **round index** in the **current window**,

$\hat{S}_i \leftarrow 0$ for $i \in [\![K]\!]$ the **cumulative reward**

While $d \leqslant \Delta_w^{(\alpha)} := \lceil w^\alpha \rceil$:

If $|\mathcal{A}| > 1$: **If** $|\mathcal{A}| > 1$:

Play each arm in $\mathcal{A}$ once, **Play** each arm in $\mathcal{A}$ once,

increment d (**stop** if $d > \Delta_w^{(\alpha)}$) **increment** d (**stop** if $d > \Delta_w^{(\alpha)}$)

Remove from $\mathcal{A}$ all i

s.t. $\hat{S}_i + B_w^{(\alpha)} < \hat{S}^* := \max_{i \in [\![K]\!]} \hat{S}_i$

If $\mathcal{A} = \{\hat{i}^*\}$ **If** $\mathcal{A} = \{\hat{i}^*\}$

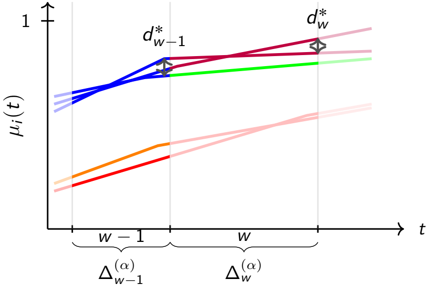
Commit to the remaining arm $\hat{i}^*$ **Commit** to the remaining arm $\hat{i}^*$

Increment d **Increment** d

Increment w

GOTO RESTART

- $\texttt{RC-BE}(\alpha)$ adapts $\texttt{BE}$ (Jia et al., 2023) to the **rising concave** setting.



REGRET UPPER BOUND

$$R^*(T) = \widetilde{\mathcal{O}}\left(T^{\frac{7}{11}} V_T^{\frac{2}{11}}\right)$$

Conclusions

Setting	Lower Bound
Rising Rising	$\Omega\left(T^{1/2}\right)$ (Lattimore et al., 2020) $\Rightarrow\ \boldsymbol{\Omega\left(T^{2/3}V_T^{1/3}\right)}$ $\Rightarrow\ \boldsymbol{\Omega\left(T^{2/3}V_T^{1/3}\right)}$ **(This Work)** **(This Work)**
Rising Concave Rising Concave	$\Omega(T^{1/2})$ (Lattimore et al., 2020) $\Rightarrow\ \boldsymbol{\Omega\left(T^{3/5}V_T^{1/5}\right)}$ $\Rightarrow\ \boldsymbol{\Omega\left(T^{3/5}V_T^{1/5}\right)}$ **(This Work)** **(This Work)**

References

- Besbes, Omar, Yonatan Gur, and Assaf Zeevi (2014). "Stochastic multi-armed-bandit problem with non-stationary rewards". In: *Advances in Neural Information Processing Systems (NIPS)*.
- Jia, Su et al. (2023). "Smooth Non-stationary Bandits". In: *International Conference on Machine Learning (ICML)*.
- Lattimore, Tor and Csaba Szepesvári (2020). *Bandit Algorithms*. Cambridge University Press.
- Metelli, Alberto M. et al. (2022). "Stochastic Rising Bandits". In: *International Conference on Machine Learning (ICML)*.
- Tekin, Cem and Mingyan Liu (2012). "Online Learning of Rested and Restless Bandits". In: *IEEE Transactions on Information Theory* 58.8, pp. 5588–5611.