

# Revisiting Follow-the-Perturbed-Leader with Unbounded Perturbations in Bandit Problems

Jongyeong Lee<sup>1,\*</sup>, Junya Honda<sup>2,3</sup>, Shinji Ito<sup>4,3</sup>, and Min-hwan Oh<sup>5</sup>

<sup>1</sup>KIST, <sup>2</sup>Kyoto University, <sup>3</sup>RIKEN AIP, <sup>4</sup>The University of Tokyo,  
<sup>5</sup>Seoul National University

NeurIPS2025

\* JL was affiliated with Seoul National University at the time of submission.

# Best-of-both-worlds in bandit problems

---

2

- A policy that achieves (near) optimal performance in
  - Adversarial bandit: loss is chosen by an adversary.
  - Stochastic bandit: loss is generated from an unknown fixed distribution.without knowing the underlying environments.
- For example, in multi-armed bandits (MAB),
  - goal is to achieve  $O(\sqrt{KT})$  adversarial regret
  - $O(\log T)$  stochastic regret

# Recent advances in BOBW literature

---

## Follow-the-Regularized-Leader (FTRL)

- Successful framework in BOBW
- Play an arm w.p.  $w_t \in \Delta^{K-1}$ 

$$w_t = \operatorname{argmin}_w \{w^\top \hat{L}_t + \Phi(w)/\eta_t\}$$
- $\Phi \geq 0$ : regularization function
  - Tsallis entropy: BOBW in MAB [Zimmert+2021]
  - Tsallis + Shannon: BOBW in semi-bandits and graph bandits. [Zimmert+2019, Ito+2022]
- Have to solve optimization problems.

## Follow-the-Perturbed-Leader (FTPL)

- Relatively less studied in BOBW
- Play an arm  $i_t \in \{1, \dots, K\}$ 

$$i_t = \operatorname{argmin}_j \{\hat{L}_{t,j} - r_{t,j}/\eta_t\}$$
- $r_{t,j} \sim D$ : random perturbation
  - Frechet-type: BOBW in MAB
    - Perturbations over  $[0, \infty)$ . [Honda+2023, Lee+2024]
  - Frechet: BOBW in semi-bandits
    - $O(\log T)$  stochastic regret [Zhan+2025]
- Simple, but limited understanding

# Relationship between FTPL and FTRL

4

- When  $K = 2$ ,<sup>\*</sup> FTRL  $\rightarrow$  FTPL is possible. [Abernethy+2016]
- FTRL with Tsallis entropy is successful in BOBW literature.
  - When  $K = 2$ ,  
FTRL with Tsallis entropy = FTPL with a Frechet-type perturbation.
  - FTPL with certain Frechet-type can achieve BOBW in MAB.
  - But in general, there may be no FTPL for given FTRL. [Hofbauer+2002]

**Can we leverage such relationships  
to obtain BOBW results of FTPL?**

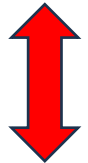
<sup>\*</sup>) Strictly speaking, 1d optimization.

# Revisiting perturbations on the real line

---

5

- Such relationship has been extensively studied for perturbations with support  $R$ . [Hofbauer+2002, Feng+2017, Fosgerau+2020]
- Hybrid regularizer: sum of Tsallis, Shannon, log-barrier...
- Hybrid perturbations: how to design?
  - Sum of perturbations? Generated from mixed distribution?
  - Need more understanding on FTPL in BOBW literature.



# In this paper...

---

1. Extend BOBW to asymmetric Frechet-type perturbations
  - whose support is the real line
  - a bridge between BOBW in bandits and discrete choice theory
  - Include Laplace-Pareto perturbation, a kind of hybrid perturbations.
2. If perturbations corresponds to Tsallis entropy exists,  
it would be symmetric Frechet-type over  $R$ !  
(when  $K = 2$ )

# In this paper...

---

7

- Then, can we obtain BOBW result for symmetric Frechet?

## 3. Partially yes

😊  $K = 2$ : we obtain BOBW results. (as equivalence suggests!)

😞  $K \geq 3$ : we find a counterexample that current analysis fails.

- Conjecture;

there exists an adversary that incurs at least  $O(\sqrt{KT \log K})$  regret.

# Unbounded, correlated perturbations (discrete choice theory literature)

## Unbounded, i.i.d. perturbations (online learning literature)

- Gumbel
  - EXP3
  - MNL model
  - FTRL with Shannon entropy

- Gaussian
  - unbounded hazard
  - near-optimal sto. regret (Kim and Tewari, 2019)
  - **conjectured linear adv. regret** (Abernethy et al. 2015)

## Bounded hazard function (Abernethy et al., 2015, Kim and Tewari, 2019)

- Gamma, Weibull
  - semi-infinite
  - near-optimal adv. regret

### Results in Section 5:

- Symmetric Fréchet-type
  - BOBW if  $K = 2$
  - near-optimal adv. regret (bounded hazard)
  - **our conjecture for  $K \geq 3$ : cannot be optimal (at most near-optimal)**

## Fréchet-type perturbation (BOBW literature)

### Our BOBW results (Section 4) - unbounded, **asymmetric**

-> connection to discrete choice theory

- Laplace-Pareto
  - BOBW
  - hybrid-type -> connection to FTRL with hybrid regularizer

### Previous BOBW results – semi-infinite

- Fréchet, Pareto
  - semi-infinite
  - BOBW (Honda et al. 2023, Lee et al. 2024)