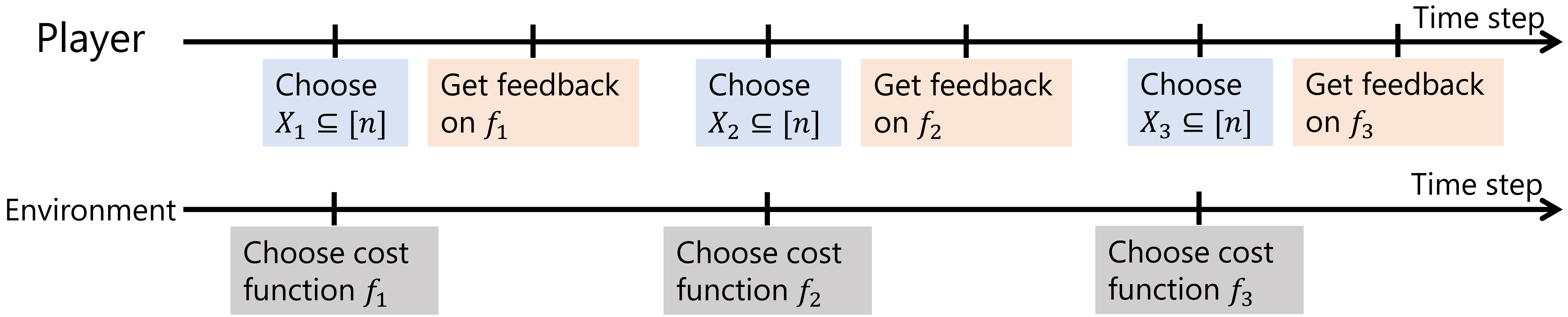
Revisiting Online Submodular Minimization: Gap-Dependent Regret Bounds, Best of Both Worlds and Adversarial Robustness

Shinji Ito (NEC Corporation)

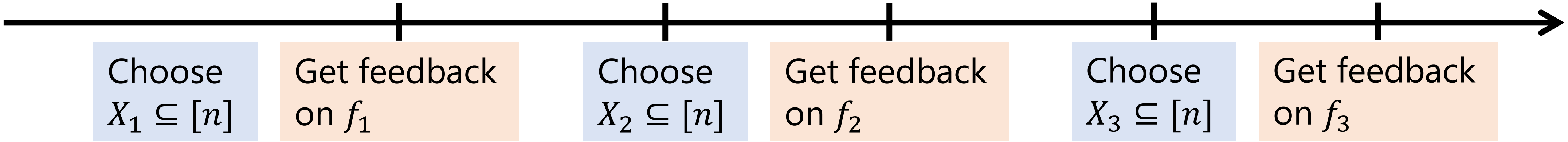
ICML2022

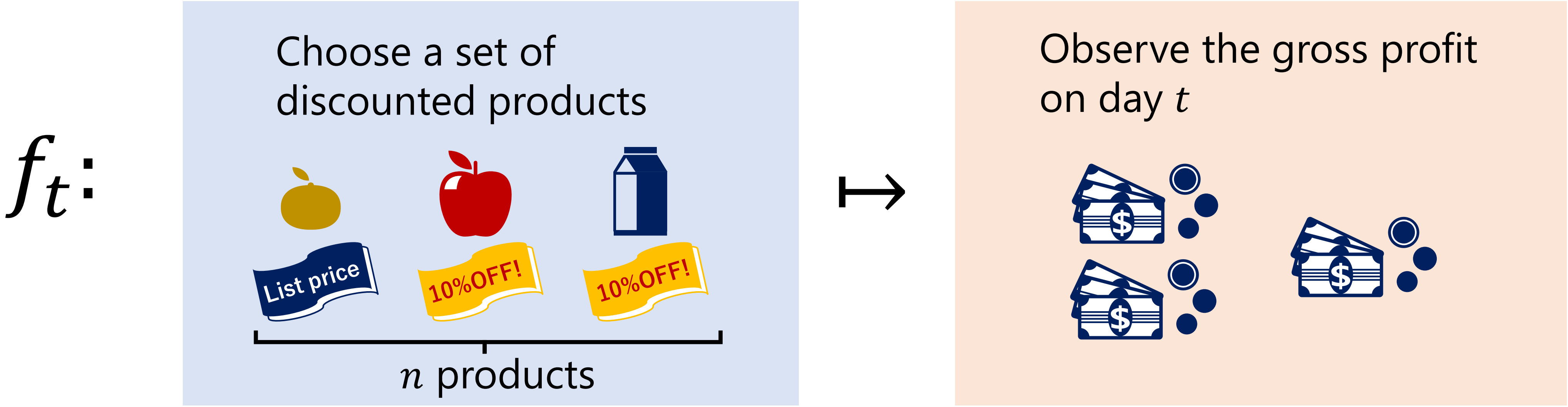
Online submodular minimization [Hazan & Kale, 2012]

- Given ground set $[n] = \{1,2,\dots,n\}$ is given
- For $t = 1,2,\dots,T$:
  - The environment chooses *submodular cost function* $f_t: 2^{[n]} \to [0,1]$
  - The player chooses *action* $X_t \subseteq [n]$ without knowing f_t
  - The player gets feedback on f_t
    - *Full-info setting*: $f_t(X)$ for any $X \subseteq [n]$ is observable
    - *Bandit setting*: only $f_t(X_t)$ is observable



Application example: multi-product price optimization





Under some assumptions on the demand model (substitute-goods property), the objective function f_t is submodular (cf. [Ito & Fujimaki, 2016])

Performance metric: Regret

Algorithm performance is evaluated by means of *regret* R_T defined as

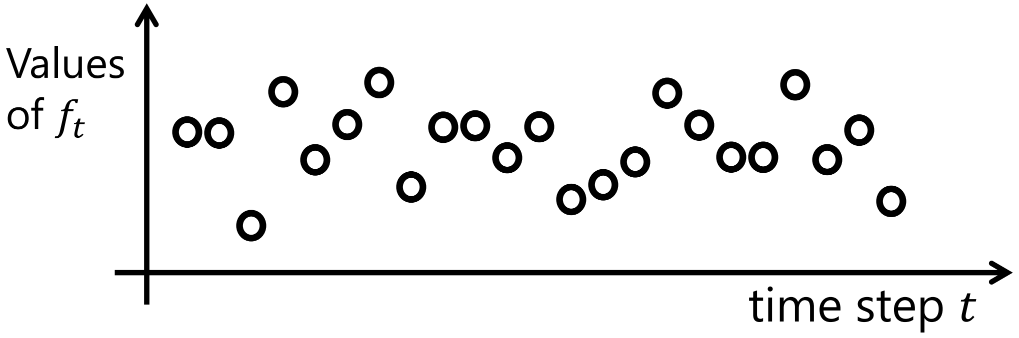
$$R_T = \mathbf{E}\left[\sum_{t=1}^{T} f_t(X_t)\right] - \min_{X^* \subseteq [n]} \mathbf{E}\left[\sum_{t=1}^{T} f_t(X^*)\right]$$

Cumulative costs incurred by **the player**

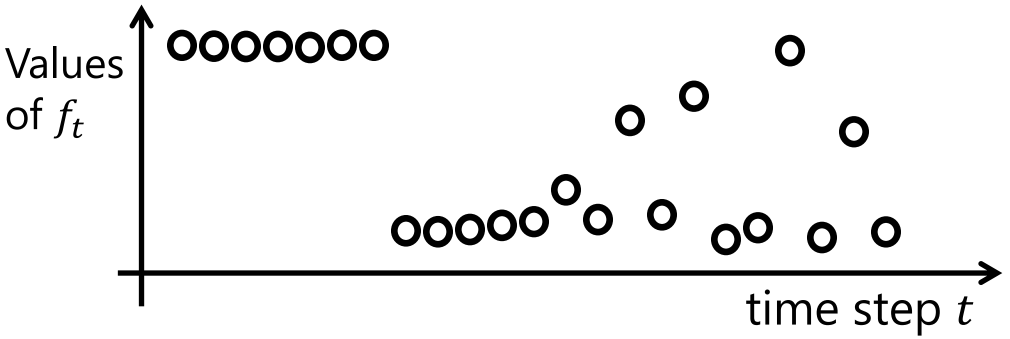
Cumulative costs for **the optimal fixed action** X^*

Two models for environments

1. **Stochastic (stationary) model**:
   cost functions f_t follow **i.i.d. distributions**

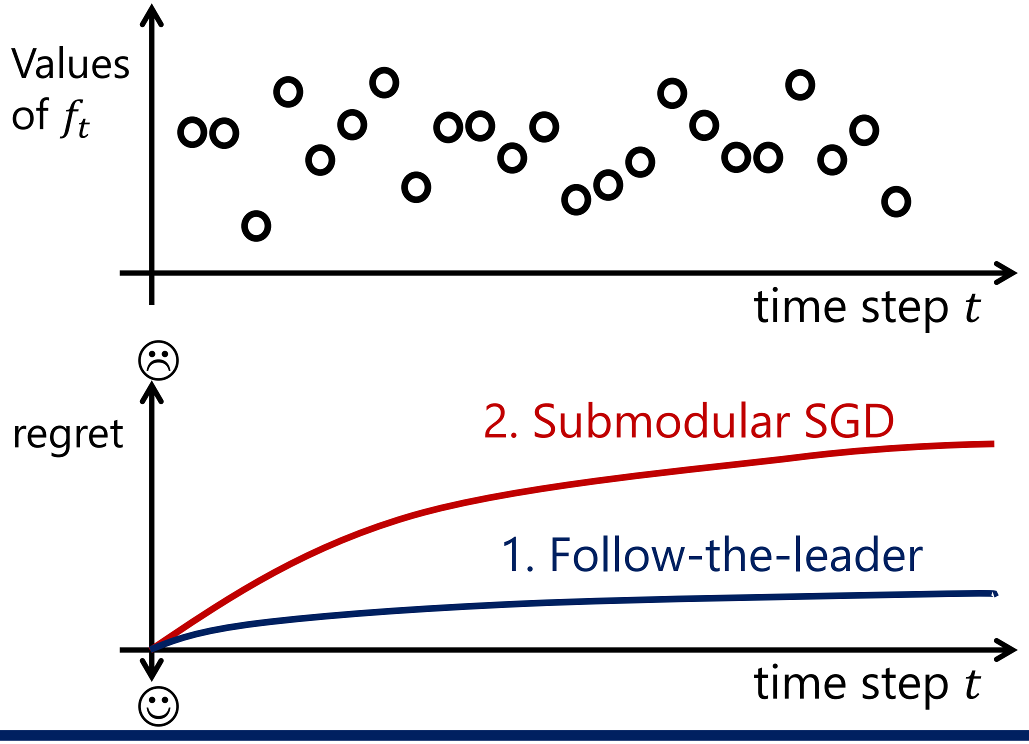


2. **Adversarial (non-stationary) model**:
   $\{f_t\}_{t=1}^{T}$ is an **arbitrary sequence**

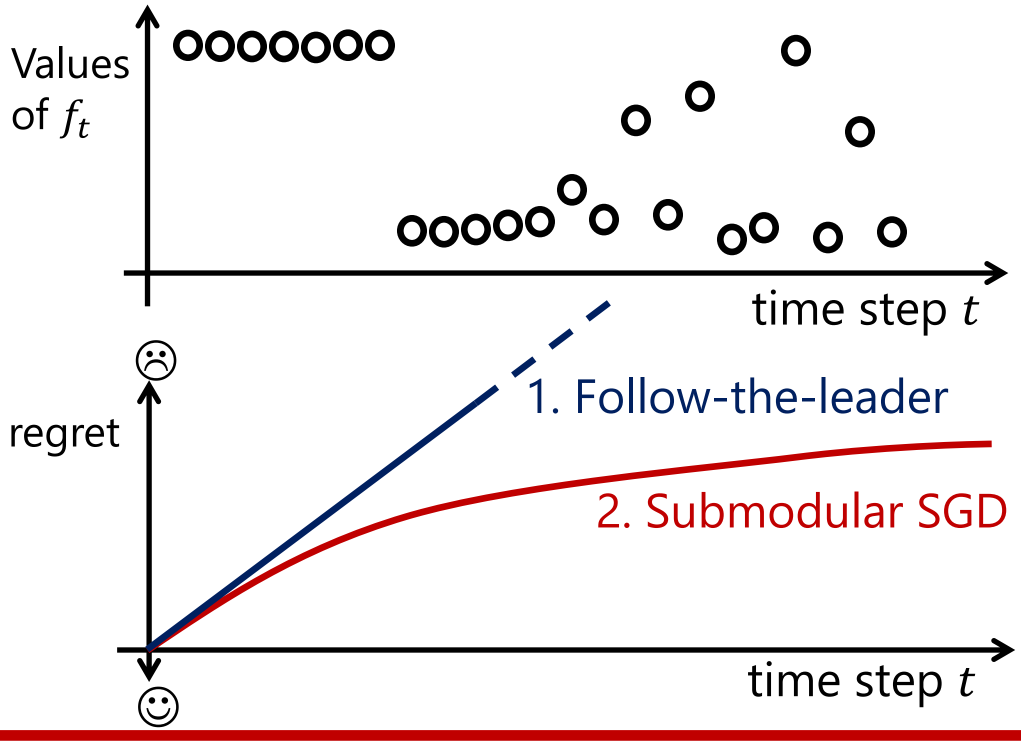


Two models for environments

1. **Stochastic (stationary) model**:
   cost functions f_t follow **i.i.d. distributions**

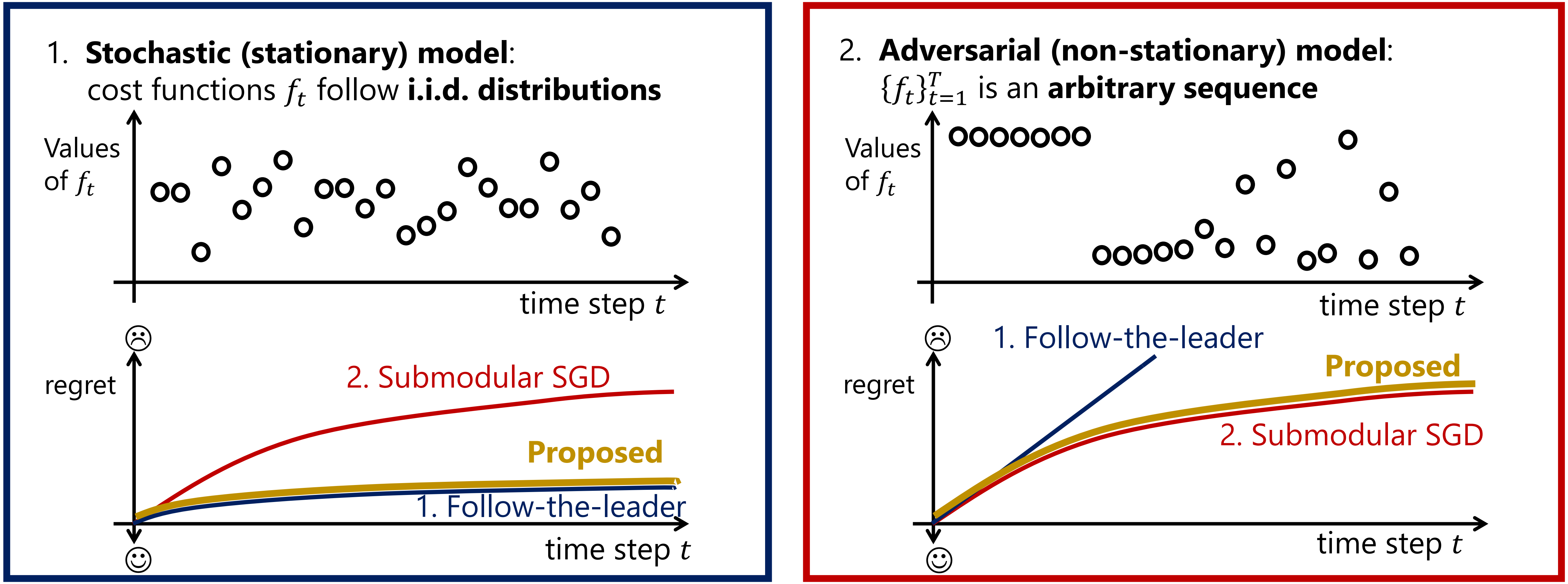


2. **Adversarial (non-stationary) model**:
   $\{f_t\}_{t=1}^{T}$ is an **arbitrary sequence**



- It is important to choose **the right algorithm for the environment.**
- However, it is difficult to know in advance which algorithm better matches the environment.

Two models for environments



- It is important to choose **the right algorithm for the environment.**
- However, it is difficult to know in advance which algorithm better matches the environment.
- **This study**: **best-of-both-worlds** algorithms that work well for both models

Summary of contributions

Regret bounds for full-information settings. Δ: suboptimality gap, C: corruption level

Algorithm	Stochastic	Adversarial	Stochastic + adversarial
Submodular SGD [Hazan&Kale, 2012]	$O(\sqrt{nT})$	$O(\sqrt{nT})$	$O(\sqrt{nT})$
Follow-the-leader	$O\left(\min\left\{\sqrt{nT}, \frac{n}{\Delta}\right\}\right)$	-	-
Proposed algo. 1 [This work]	$O\left(\min\left\{\sqrt{nT}, \frac{n}{\Delta}\right\}\right)$	$O(\sqrt{nT})$	$O\left(\min\left\{\sqrt{nT}, \frac{n}{\Delta} + \sqrt{\frac{Cn}{\Delta}}\right\}\right)$

Summary of contributions

Regret bounds for full-information settings. Δ: suboptimality gap, C: corruption level

Algorithm	Stochastic	Adversarial	Stochastic + adversarial
Submodular SGD [Hazan&Kale, 2012]	$O(\sqrt{nT})$	$O(\sqrt{nT})$	$O(\sqrt{nT})$
Follow-the-leader	$O\left(\min\left\{\sqrt{nT}, \frac{n}{\Delta}\right\}\right)$	-	-
Proposed algo. 1 [This work]	$O\left(\min\left\{\sqrt{nT}, \frac{n}{\Delta}\right\}\right)$	$O(\sqrt{nT})$	$O\left(\min\left\{\sqrt{nT}, \frac{n}{\Delta} + \sqrt{\frac{Cn}{\Delta}}\right\}\right)$

Best-of-both-worlds algorithm

Summary of contributions

Regret bounds for full-information settings. Δ: suboptimality gap, C: corruption level

Algorithm	Stochastic	Adversarial	Stochastic + adversarial
Submodular SGD [Hazan&Kale, 2012]	$O(\sqrt{nT})$	$O(\sqrt{nT})$	$O(\sqrt{nT})$
Follow-the-leader	$O\left(\min\left\{\sqrt{nT}, \frac{n}{\Delta}\right\}\right)$	-	-
Proposed algo. 1 [This work]	$O\left(\min\left\{\sqrt{nT}, \frac{n}{\Delta}\right\}\right)$	$O(\sqrt{nT})$	$O\left(\min\left\{\sqrt{nT}, \frac{n}{\Delta} + \sqrt{\frac{Cn}{\Delta}}\right\}\right)$

Best-of-both-worlds algorithm

Robust to adversarial corruption

Summary of contributions

Regret bounds for full-information settings. Δ: suboptimality gap, C: corruption level

Algorithm	Stochastic	Adversarial	Stochastic + adversarial
Submodular SGD [Hazan&Kale, 2012]	$O(\sqrt{nT})$	$O(\sqrt{nT})$	$O(\sqrt{nT})$
Follow-the-leader	$O\left(\min\left\{\sqrt{nT}, \frac{n}{\Delta}\right\}\right)$	-	-
Proposed algo. 1 [This work]	$O\left(\min\left\{\sqrt{nT}, \frac{n}{\Delta}\right\}\right)$	$O(\sqrt{nT})$	$O\left(\min\left\{\sqrt{nT}, \frac{n}{\Delta} + \sqrt{\frac{Cn}{\Delta}}\right\}\right)$

Regret bounds for bandit-feedback settings.

Model	Stochastic	Adversarial	Stochastic + adversarial
Bandit submodular SGD [Hazan&Kale, 2012]	$O(nT^{2/3})$	$O(nT^{2/3})$	$O(nT^{2/3})$
Proposed algo. 1 [This work]	$\tilde{O}\left(\min\left\{nT^{2/3}, \frac{n^3}{\Delta^2}\right\}\right)$	$\tilde{O}(nT^{2/3})$	$\tilde{O}\left(\min\left\{nT^{2/3}, \frac{n^3}{\Delta^2} + \left(\frac{C^2 n^3}{\Delta^2}\right)^{1/3}\right\}\right)$