Online Continuous Submodular Maximization: From Full-Information to Bandit Feedback

Mingrui Zhang[1], Lin Chen[1], Hamed Hassani[2] and Amin Karbasi[1]

[1]Yale University and [2]University of Pennsylvania

1. Submodular Functions

- Submodularity is an intuitive diminishing returns property: a discrete set function $f : 2^V \to \mathbb{R}_+$ is **submodular** if for all sets $A \subseteq B \subseteq V$ and $x \notin B$, we have

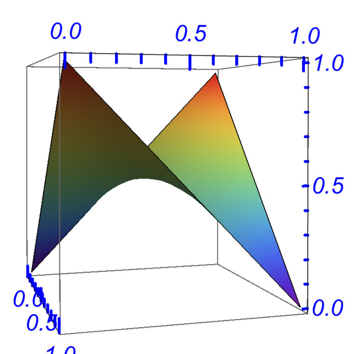
$$f(A \cup \{x\}) - f(A) \geq f(B \cup \{x\}) - f(B)$$

f(🍟🍔)-f(🍟)≥f(🍟🍕🍔)-f(🍟🍕)

Submodular function f is **monotone** if for all $A \subseteq B$: $f(A) \leq f(B)$

- Continuous analogue: a differentiable function $F : \mathcal{X} \to \mathbb{R}_{\geq 0}$ defined on a box $\mathcal{X} \triangleq \prod_{i=1}^{d} \mathcal{X}_i$, where each $\mathcal{X}_i$ is a closed interval of $\mathbb{R}_{\geq 0}$. We say that F is **Continuous DR-submodular** [1] if for every $x, y \in \mathcal{X}$ that satisfy $x \leq y$ and every $i \in [d] \triangleq \{1, \ldots, d\}$, we have

$$\frac{\partial F}{\partial x_i}(x) \geq \frac{\partial F}{\partial x_i}(y).$$



Example: $F(x_1, x_2) = x_1(1 - x_2) + (1 - x_1)x_2$.

2. Bandit Optimization

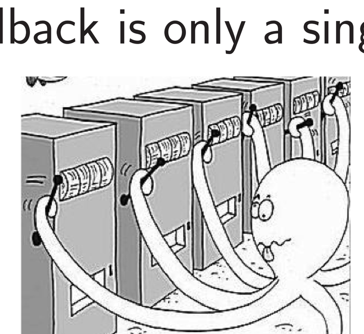
Online optimization is a repeated two-player game. At each round t:

- The learner chooses an action x_t from a convex set $\mathcal{K} \subseteq \mathbb{R}^n$;
- The adversary chooses a reward function F_t from $\mathcal{F}$, a family of real-valued functions;
- The learner receives a reward $F_t(x_t)$, and observes feedback.

The aim is to maximize the **regret**: the gap between her accumulated reward and the reward of the best single choice in hindsight

$$\mathcal{R}_T \triangleq \max_{x \in \mathcal{K}} \sum_{t=1}^{T} F_t(x) - \sum_{t=1}^{T} F_t(x_t).$$

In the **bandit** setting, the feedback is only a single real number $F_t(x_t)$.



Bandit Problems (source: http://www.primarydigit.com/blog/multi-arm-bandits-explorationexploitation-trade-off)

However, even in the offline scenario, continuous DR-submodular maximization problem cannot be approximated within a factor of $(1 - 1/e + \epsilon)$ for any $\epsilon > 0$ in polynomial time, unless $RP = NP$ [1]. Therefore, we consider the **$(1 - 1/e)$-regret**

$$\mathcal{R}_{1-1/e,T} \triangleq (1 - 1/e) \max_{x \in \mathcal{K}} \sum_{t=1}^{T} F_t(x) - \sum_{t=1}^{T} F_t(x_t).$$

3. Our Contribution

We studied the following three problems:

- `OCSM` : the Online Continuous DR-Submodular Maximization problem,
- `BCSM` : the Bandit Continuous DR-Submodular Maximization problem, and
- `RBSM` : the Responsive Bandit Submodular Maximization problem.

Comparison of previous and our proposed algorithms:

Setting	Algorithm	Stochastic gradient	# of grad. evaluations	$(1 - 1/e)$-regret
OCSM	Meta-FW[3]	No	$T^{1/2}$	$O(\sqrt{T})$
	VR-FW[2]	Yes	$T^{3/2}$	$O(\sqrt{T})$
	Mono-FW (this work)	Yes	1	$O(T^{4/5})$
BCSM	**Bandit-FW (this work)**	-	-	$O(T^{8/9})$
RBSM	**Responsive-FW (this work)**	-	-	$O(T^{8/9})$

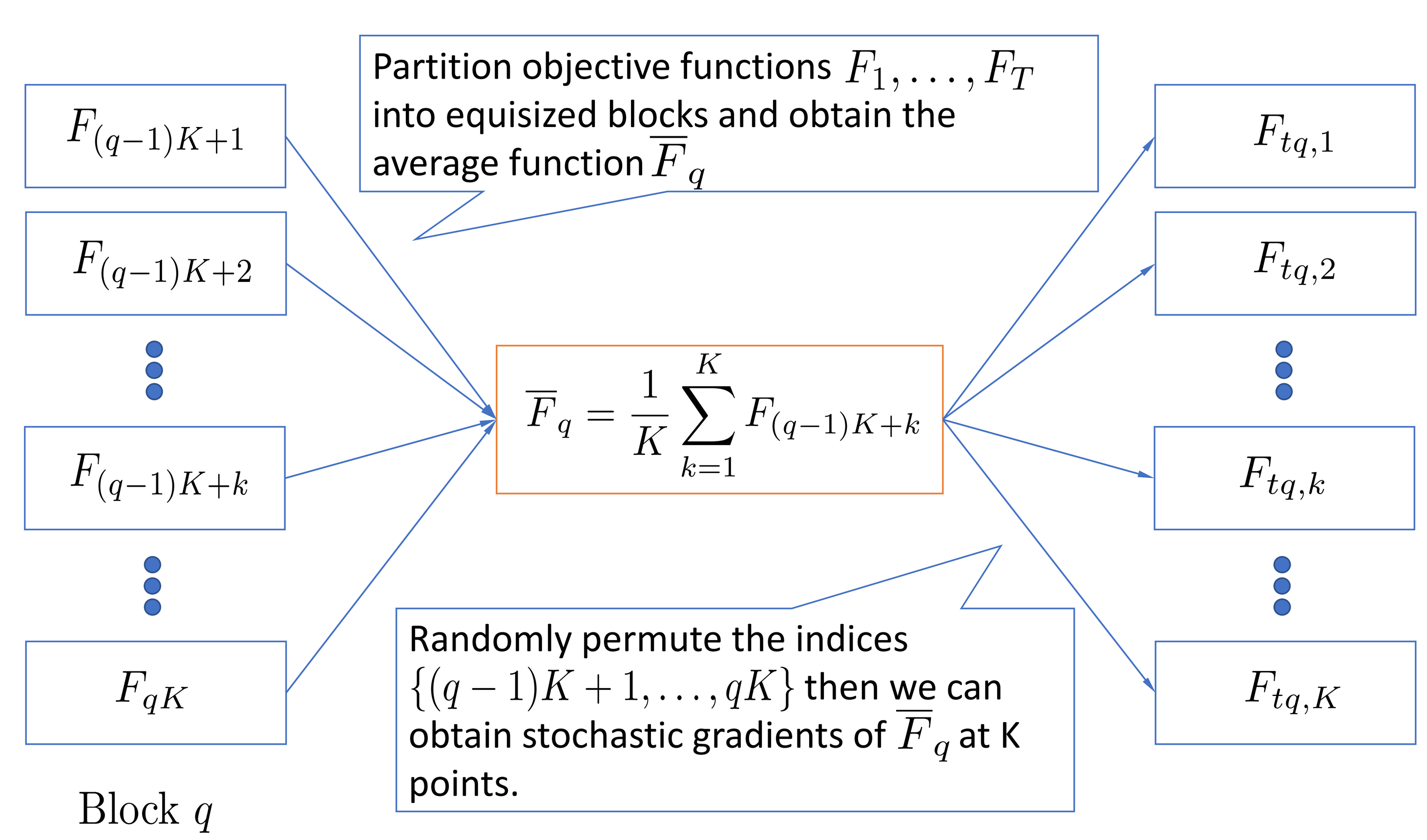
4. One-shot Online Continuous DR-Submodular Maximization

Offline **Frank-Wolfe (FW)** method for maximizing monotone continuous DR-submodular functions: at k-th iteration, solves a linear optimization problem

$$v^{(k)} \leftarrow \arg\max_{v \in \mathcal{K}} \langle v, \nabla F(x^{(k)}) \rangle$$

which is used to update $x^{(k+1)} \leftarrow x^{(k)} + \eta_k v^{(k)}$, where η_k is the step size. We extended it to online setting.

- **Obstacle 1:** To obtain $v_t^{(k)}$, we have to know the gradient in advance.
  Solution: Use K no-regret online linear maximization oracles $\{\mathcal{E}^{(k)}\}, k \in [K]$, with $\langle \cdot, d_t^{(k)} \rangle$ being the objective function, and $v_t^{(k)}$ being the output, where $d_t^{(k)}$ is an estimation of $\nabla F_t(x_t(k))$.
- **Obstacle 2:** To obtain a **one-shot** algorithm, we need to reduce K from $T^{3/2}$ [2, 3] to 1.
  Solution: Proposed the **blocking procedure** and the **permutation methods**.



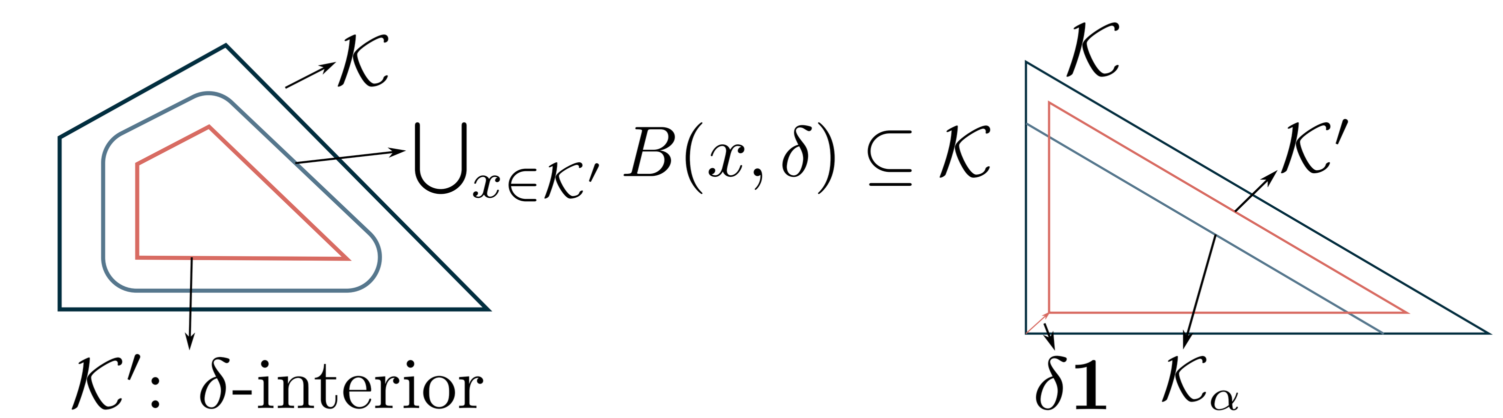
5. Bandit Continuous DR-Submodular Maximization

One-point Gradient Estimator: define the δ-smoothed version of a function F as $\hat{F}_\delta(x) \triangleq \mathbb{E}_{v \sim B^d}[F(x + \delta v)]$. Then we have

$$\nabla \hat{F}_\delta(x) = \mathbb{E}_{u \sim S^{d-1}} \left[\frac{d}{\delta} F(x + \delta u) u \right].$$

- **Obstacle:** The point $x + \delta u$ may fall outside of $\mathcal{K}$.
  Solution: Introduce the notion of **δ-interior**. A set $\mathcal{K}'$ is a δ-interior of $\mathcal{K}$ if it is a *subset* of

$$\operatorname{int}_\delta(\mathcal{K}) = \{x \in \mathcal{K} | \inf_{s \in \partial \mathcal{K}} d(x, s) \geq \delta\}.$$



(a) Example of δ-interior (b) Construct a δ-interior

Also define the **discrepancy** between $\mathcal{K}$ and $\mathcal{K}'$ by

$$d(\mathcal{K}, \mathcal{K}') = \sup_{x \in \mathcal{K}} d(x, \mathcal{K}'),$$

When F_t is Lipschitz and $d(\mathcal{K}, \mathcal{K}')$ is small, we can **approximate** the optimal total reward on $\mathcal{K}$ by that on $\mathcal{K}'$.

Lemma 1: Under some regularization assumptions, $\mathcal{K}' = (1 - \alpha)\mathcal{K} + \delta 1$ is a δ-interior of $\mathcal{K}$ with $d(\mathcal{K}, \mathcal{K}') \leq [\sqrt{d}(\frac{R}{r} + 1) + \frac{R}{r}]\delta$.

6. Bandit Submodular Set Maximization

- **Obstacle:** If there is a rounding scheme $\mathbf{round}_{\mathcal{I}} : [0, 1]^d \to \mathcal{I}$ satisfying the following unbiasedness condition

$$\mathbb{E}[f(\mathbf{round}_{\mathcal{I}}(x))] = F(x), \quad \forall x \in [0, 1]^d$$

for any submodular set function f within the matroid constraint $\mathcal{I}$ and its multilinear extension F, we can solve the discrete bandit problem by applying **Bandit-FW** on F_t. However, this kind of rounding schemes does **NOT** exist (**Lemma 2**).
Solution: Study the **responsive** model. If $X_t \notin \mathcal{I}$, we can still observe the function value $f_t(X_t)$ as feedback, while the received reward at round t is 0.

7. References

[1] An Bian, Baharan Mirzasoleiman, Joachim M. Buhmann, and Andreas Krause. Guaranteed non-convex optimization: Submodular maximization over continuous domains. In *AISTATS*, February 2017.

[2] Lin Chen, Christopher Harshaw, Hamed Hassani, and Amin Karbasi. Projection-free online optimization with stochastic gradient: From convexity to submodularity. In *ICML*, pages 813–822, 2018.

[3] Lin Chen, Hamed Hassani, and Amin Karbasi. Online continuous submodular maximization. In *AISTATS*, pages 1896–1905, 2018.