Online Submodular Coordination with Bounded Tracking Regret: Theory, Algorithm, and Applications to Multi-Robot Coordination

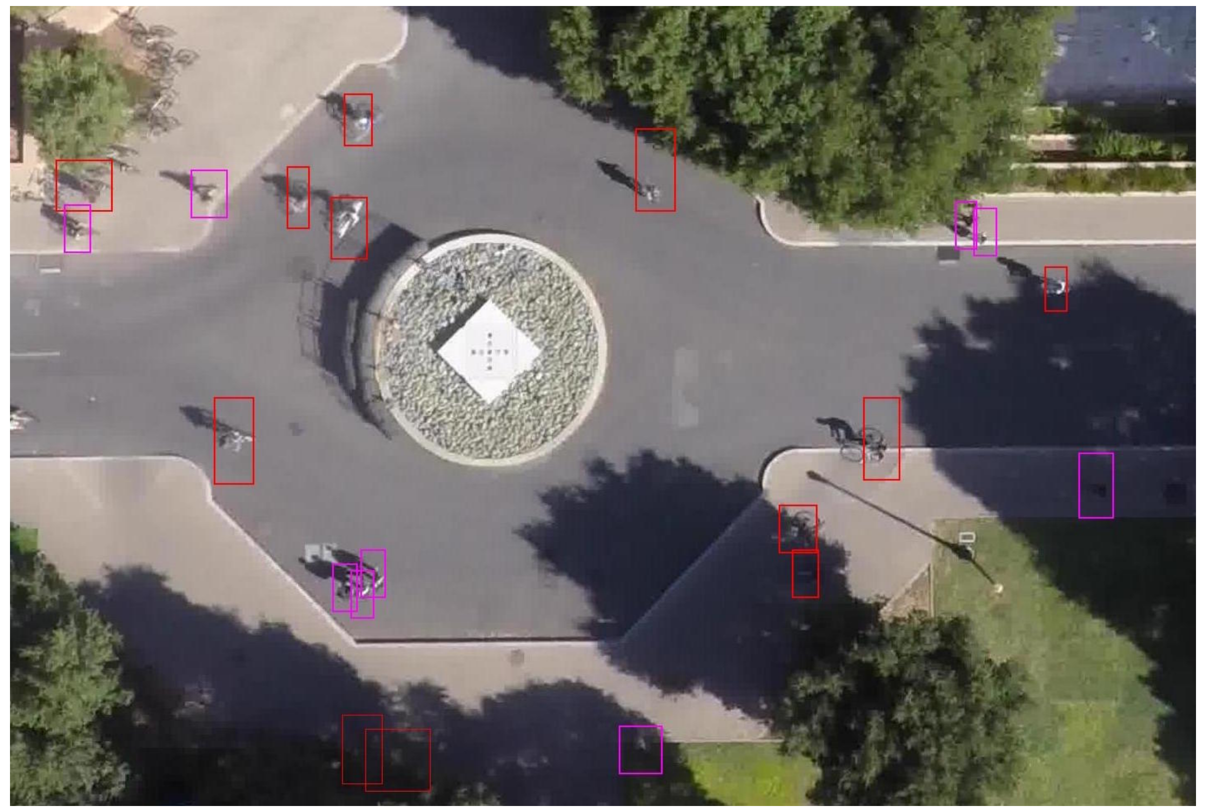
Zirui Xu, Hongyu Zhou, Vasileios Tzoumas



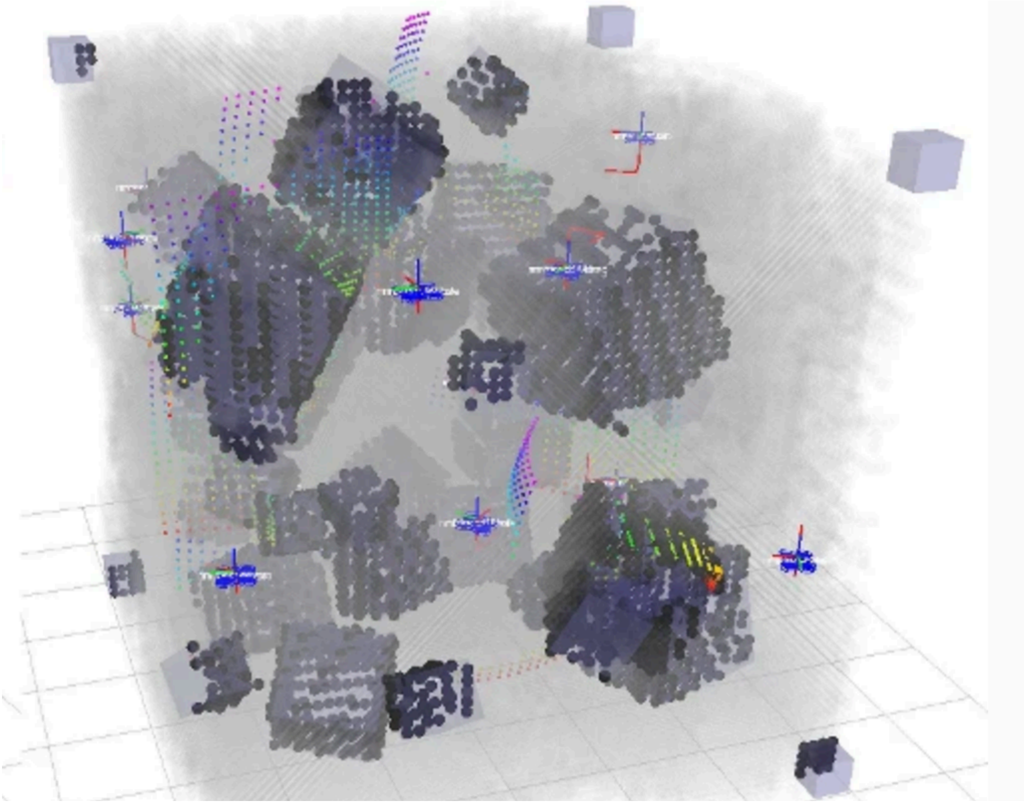
MICHIGAN ENGINEERING
UNIVERSITY OF MICHIGAN

Environmental Monitoring

Target Tracking



Visual Mapping

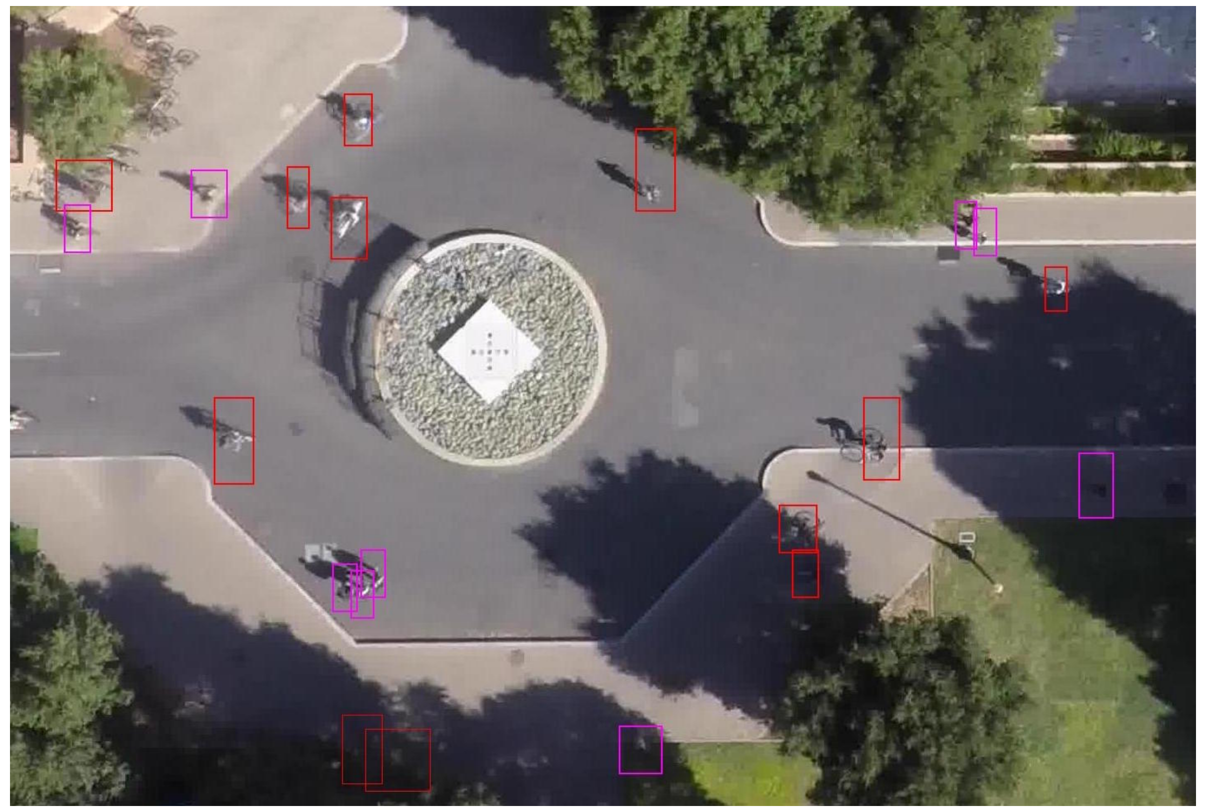


Goal: Robots need to coordinate their actions to complete complex tasks

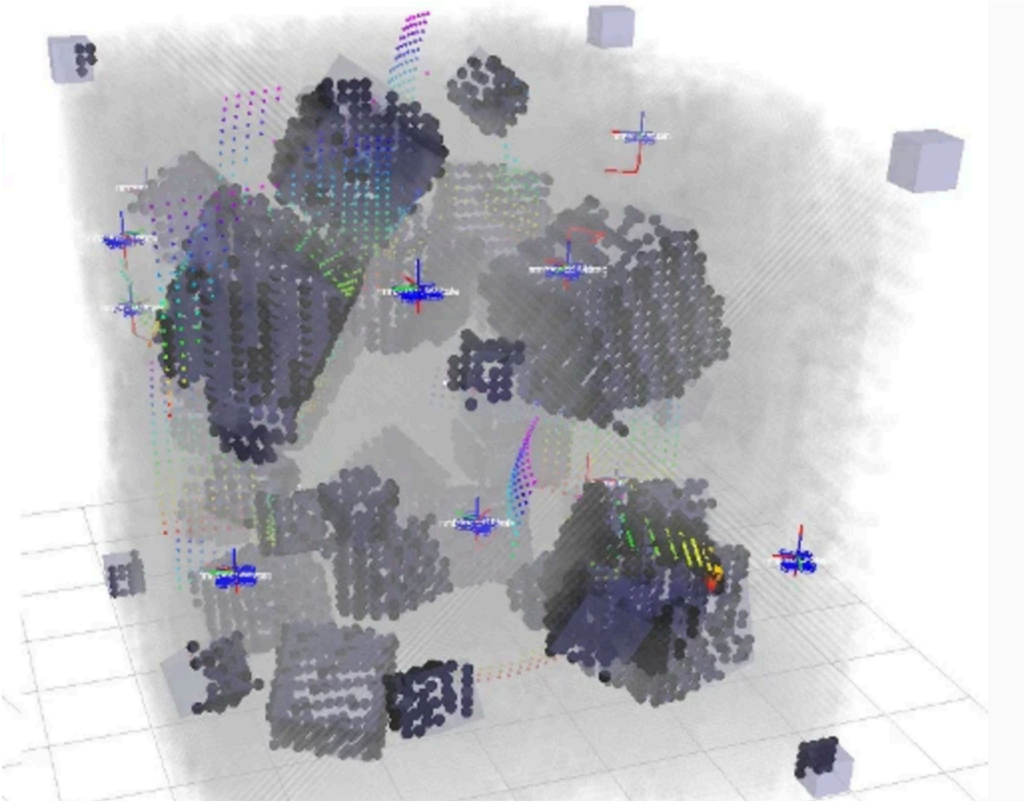
Multi-Robot Coordination Problems

Environmental Monitoring

Target Tracking



Visual Mapping



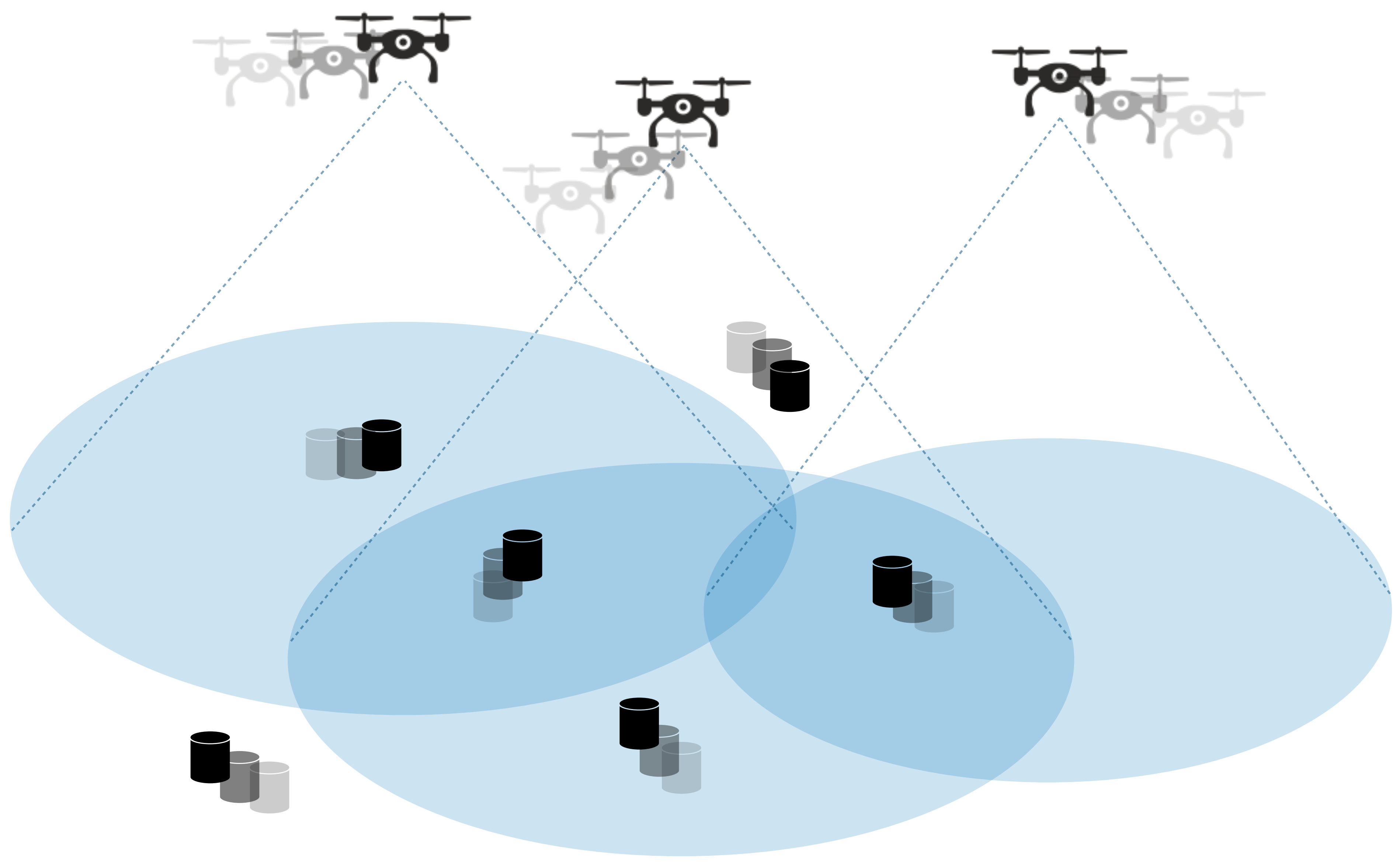
Goal: Robots need to coordinate their actions to complete complex tasks

Challenges: Such multi-robot information-gathering tasks are challenging because:

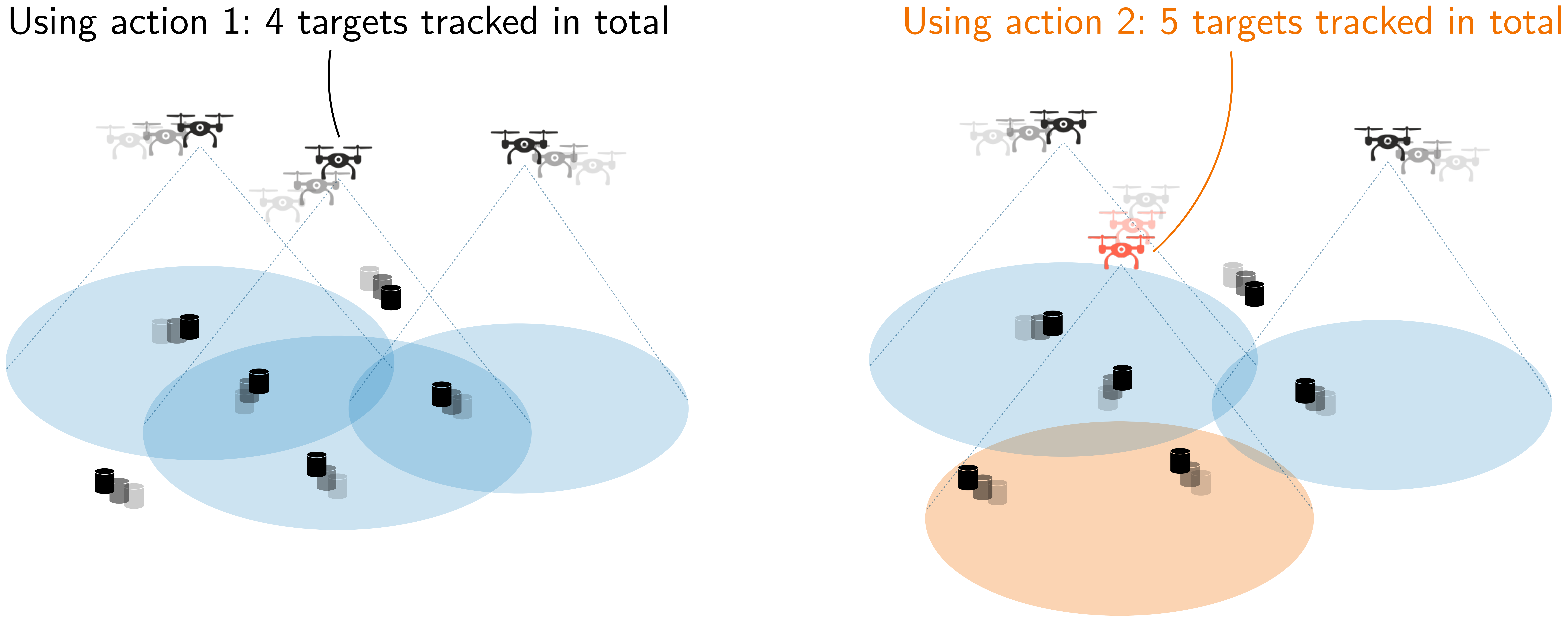
I. Actions of different robots have information overlap.

II. The environment can be unpredictable.

Goal: Maximize the number of moving targets tracked by the drones' field of view

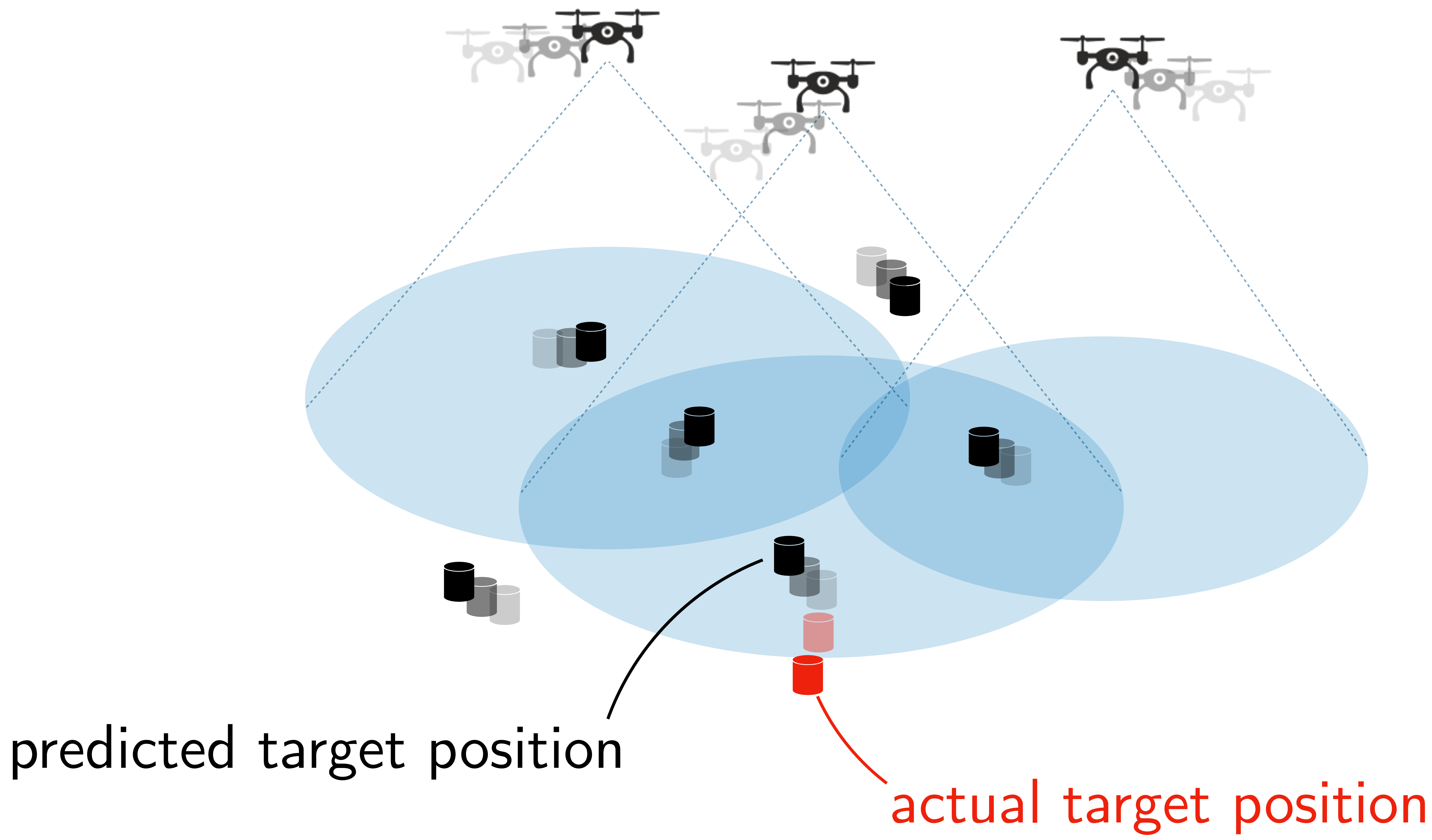


Challenge I (Utility Overlap): Compromises Actions' Effect



To overcome Challenge I: Robots need to coordinate to minimize utility overlap

Challenge II (Unpredictability): Compromises Ability to Evaluate Actions A Priori



To overcome Challenge II: Robots need to select actions based on past information only

Offline monotone submodular maximization with known environment:[1]

Given:

- robots $\mathcal{N}$
- finite action sets $\mathcal{V}_i$, $\forall i \in \mathcal{N}$
- set function $f\colon 2^{\prod_{i\in\mathcal{N}} \mathcal{V}_i} \mapsto \mathbb{R}$

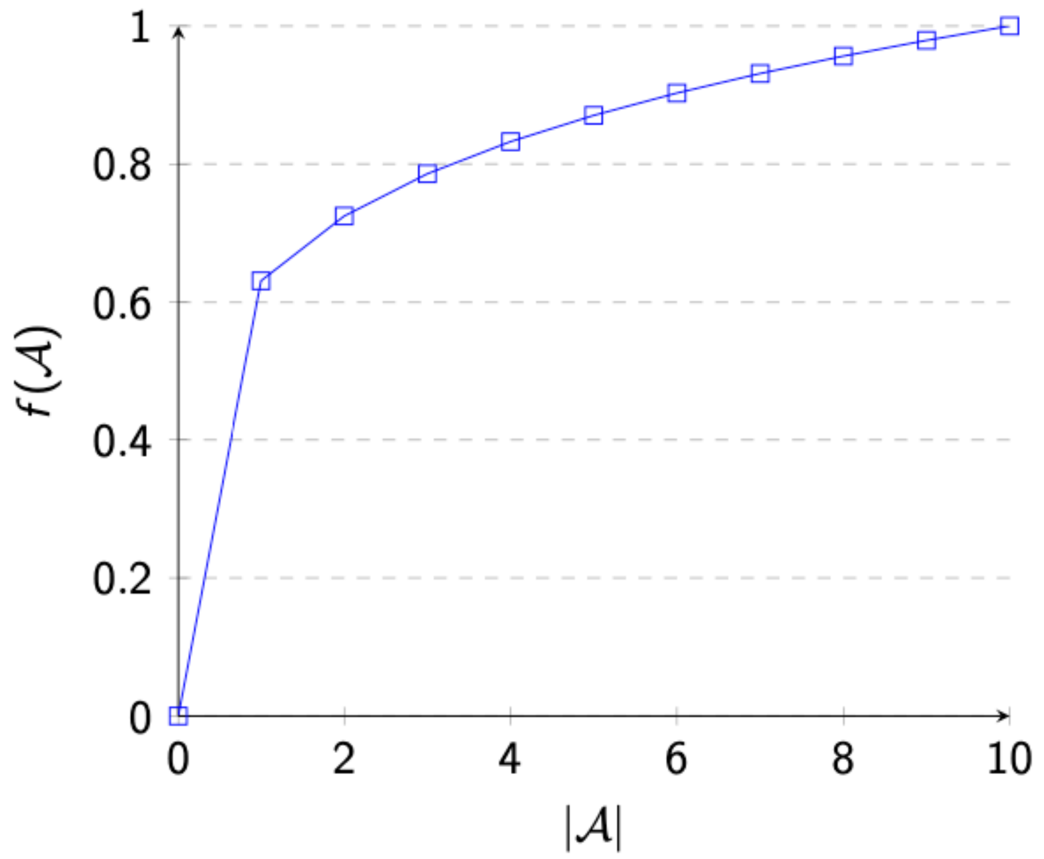
the robots $\mathcal{N}$ select actions $\{a_i\}_{i\in\mathcal{N}}$ to solve

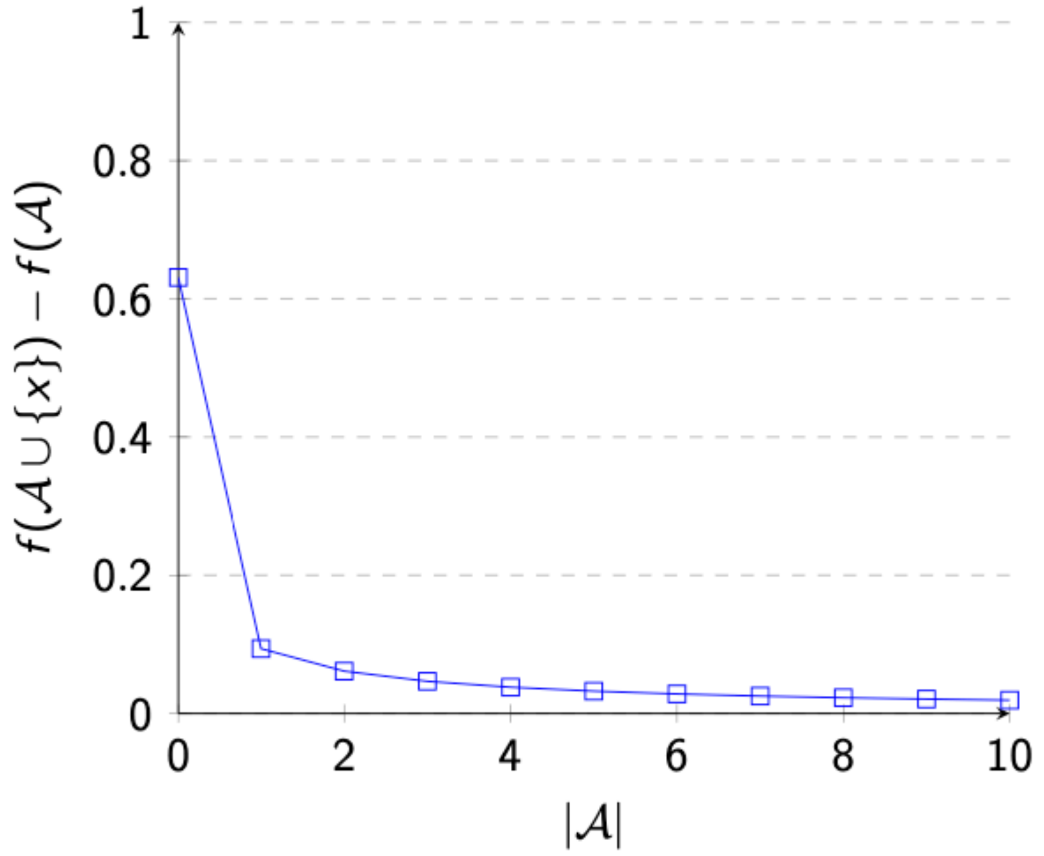
$$\max_{a_i \in \mathcal{V}_i,\, \forall i \in \mathcal{N}} \; f(\,\{a_i\}_{i\in\mathcal{N}}\,)$$

known a priori

non-decreasing

submodular
(due to information overlap)





[1]Atanasov; Bilmes; Bushnell; Calinescu; Chekuri; Clark; Corah; Gharesifard; Hassani; Hespanha; Iyer; Karbasi; Kia; Konda; Krause; Li; Marden; Martinez; Michael; Mirzasoleiman; Mokhtari; Pappas; Poovendran; Rezazadeh; Robey; Smith; Sundaram; Tokekar; ...

Problem (Online Submodular Coordination)

Given:

- time horizon T
- robots $\mathcal{N}$
- finite action sets $\mathcal{V}_i$, $\forall i \in \mathcal{N}$

at each time step $t \in [T]$, the robots $\mathcal{N}$ select actions $\{a_{i,\,t}\}_{i \in \mathcal{N}}$ online to solve

$$\max_{a_{i,\,t} \in \mathcal{V}_i,\, \forall i \in \mathcal{N}} \; f_t(\, \{a_{i,\,t}\}_{i \in \mathcal{N}} \,)$$

non-decreasing submodular

online feedback

where $\boxed{f_t}$ becomes known to the robots $\mathcal{N}$ only once $\{a_{i,\,t}\}_{i \in \mathcal{N}}$ are executed

Problem (Online Submodular Coordination)

Given:

- time horizon T
- robots $\mathcal{N}$
- finite action sets $\mathcal{V}_i$, $\forall i \in \mathcal{N}$

at each time step $t \in [T]$, the robots $\mathcal{N}$ select actions $\{a_{i,\,t}\}_{i \in \mathcal{N}}$ online to solve

$$\max_{a_{i,\,t} \in \mathcal{V}_i,\, \forall i \in \mathcal{N}} \; f_t(\,\{a_{i,\,t}\}_{i \in \mathcal{N}}\,)$$

non-decreasing submodular

online feedback

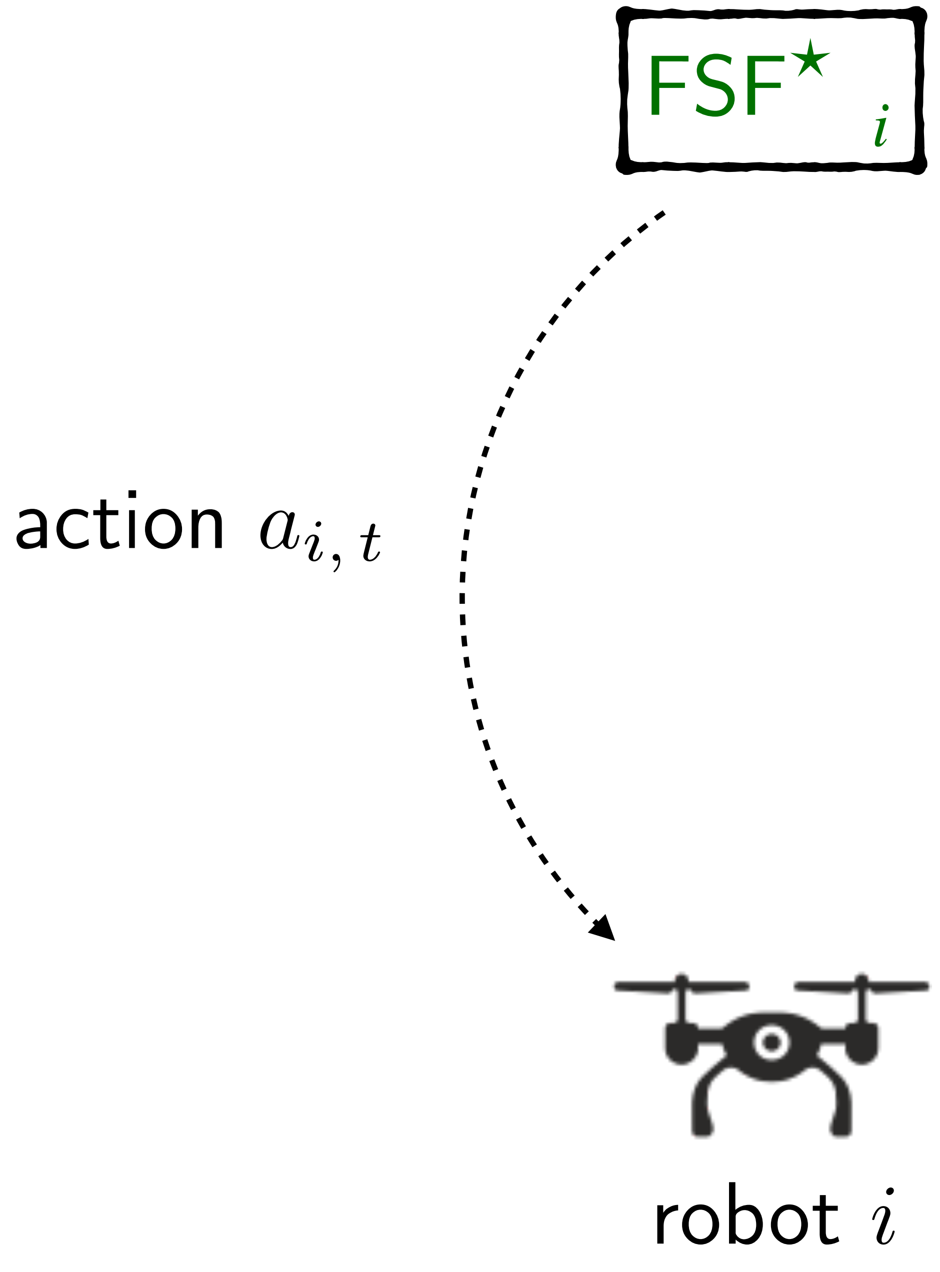
where $\boxed{f_t}$ becomes known to the robots $\mathcal{N}$ only once $\{a_{i,\,t}\}_{i \in \mathcal{N}}$ are executed

Difficulty: **NP-Hard** to achieve approximation bound better than 1/2 even when f_t is known a priori

Information Overlap	Unpredictability ✗	Unpredictability ✓
✓	Tokekar et al., IROS'14 Atanasov et al., ICRA'15 Gharesifard and Smith, TCNS'17 Grimsman et al., TCNS'18 Zhou et al., RAL'19 Corah and Michael, IROS'21 Schlotfeldt et al., TRO'21 Konda et al., ACC'22 Xu and Tzoumas, CDC'22	Our algorithm
✗	Nieto-Granda, IJRR'14 Alonso-Mora et al., ICRA'15 Wagner and Choset, Arti. Intel.'15 Mathew et al., TASE'15 Robin and Lacroix, AuRo'16 Gil et al., AuRo'17 Kemna et al., ICRA'17 Best et al., IJRR'19	Baykal et al., ICRA'17 Zhang and Hoi, AAAI'19 Lee et al., ICRA'21 Landgren et al., Automatica'21 Dahiya et al., TCNS'22

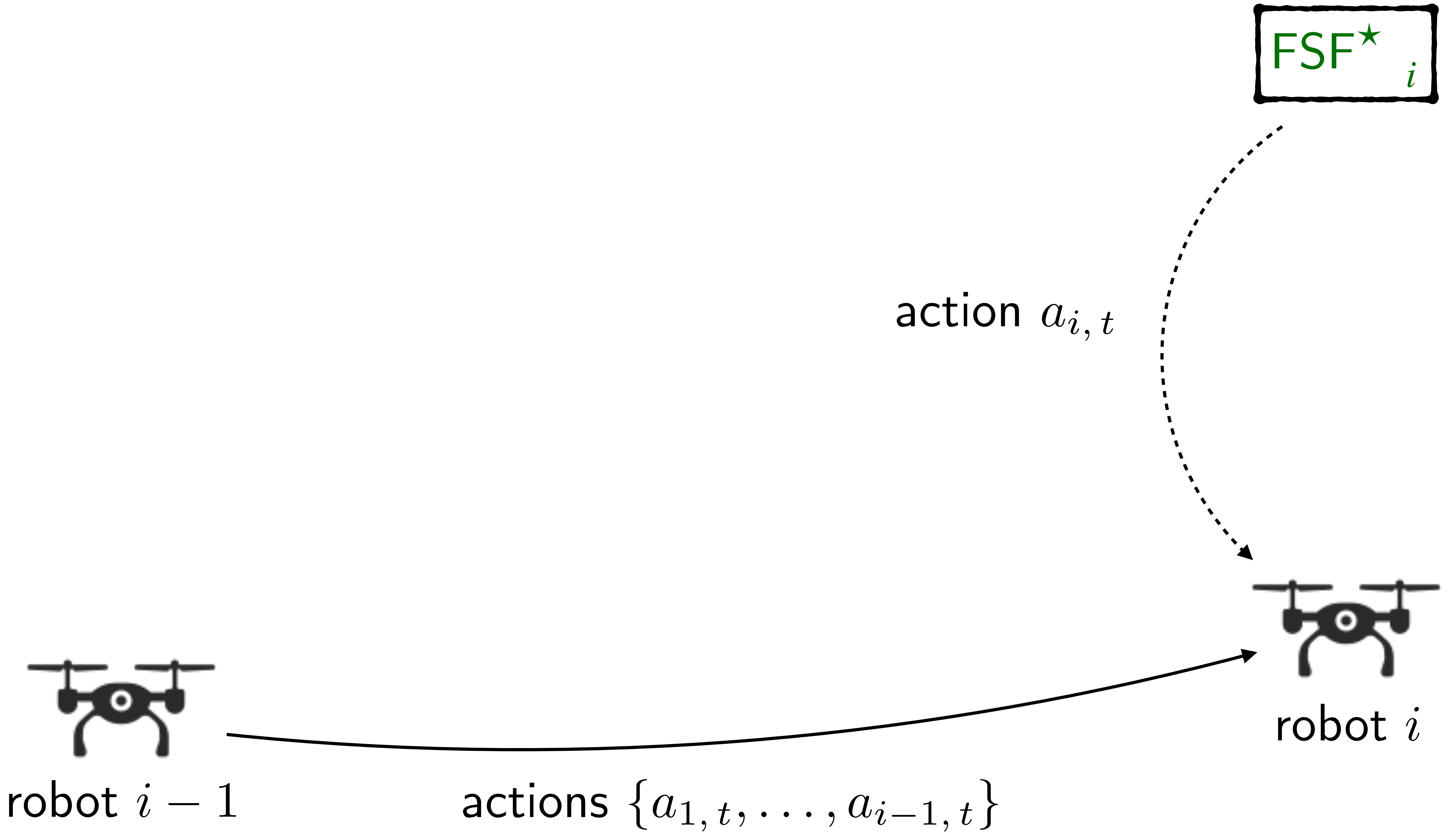
Our Algorithm: Online Sequential Greedy (OSG)

Step 1. Onboard algorithm FSF* samples next action $a_{i,\,t}$ from probability distribution $p_t^{(i)}$ that is computed based on past rewards

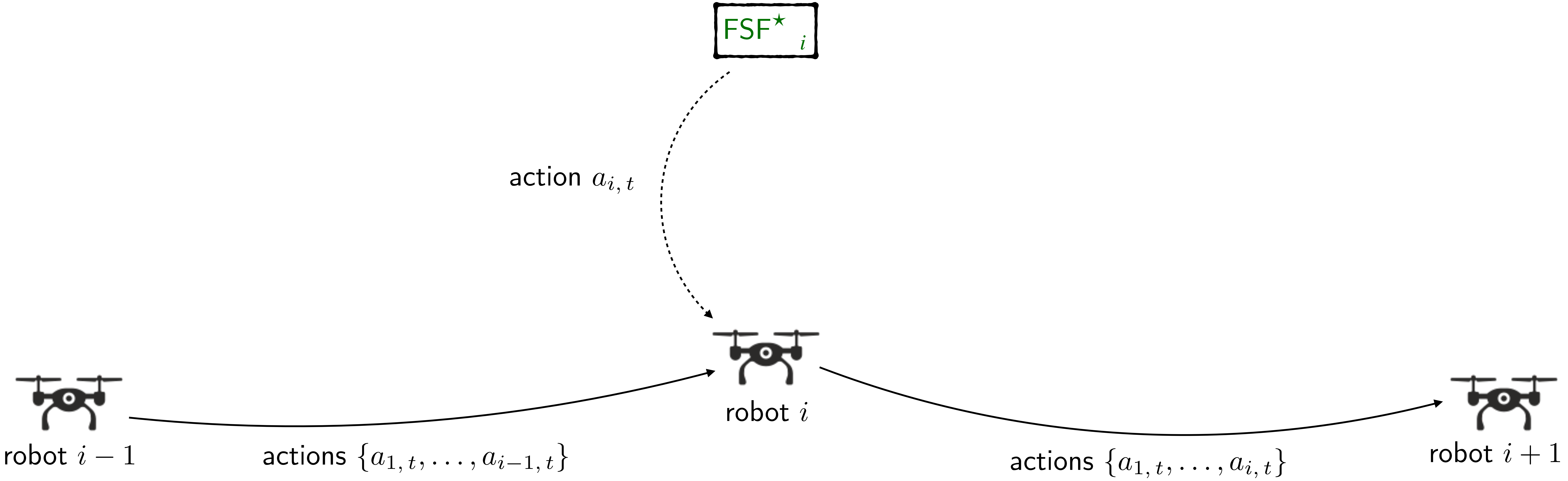


Our Algorithm: Online Sequential Greedy (OSG)

Step 2. Robot $i-1$ sends actions $\{a_{1,\,t}, \ldots, a_{i-1,\,t}\}$ to robot i

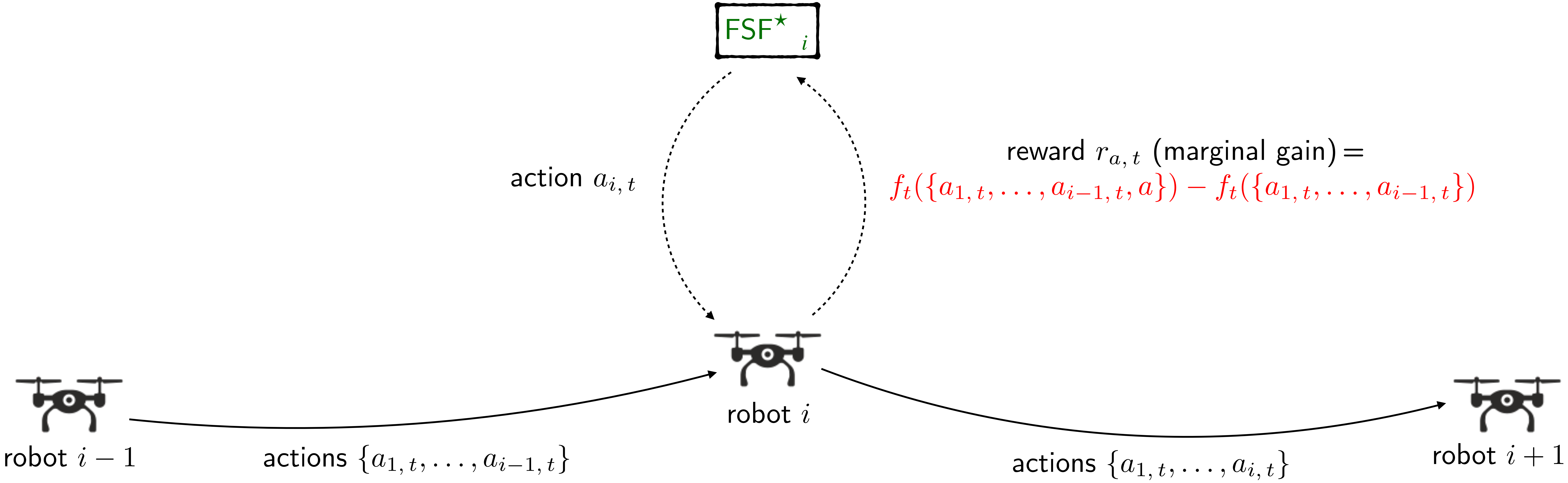


Step 3. Robot i sends actions $\{a_{1,\,t}, \ldots, a_{i,\,t}\}$ to robot $i+1$



Our Algorithm: Online Sequential Greedy (OSG)

Step 4. Robot i computes rewards $\{r_{a,\,t}\}_{a\,\in\,\mathcal{V}_i}$ (marginal gains) of all alternative actions $\mathcal{V}_i$ and feeds them into FSF* to update $p_t^{(i)}$



Computational complexity:

Theorem 1 (Sublinear Computational Complexity)

OSG requires each robot to perform $O(|\mathcal{V}_i|)$ function evaluation & $O(\log T)$ additions and multiplications per time step.

Performance Guarantees

Computational complexity:

Theorem 1 (Sublinear Computational Complexity)

OSG requires each robot to perform $O(|\mathcal{V}_i|)$ function evaluation & $O(\log T)$ additions and multiplications per time step.

Approximation performance:

Theorem 2 (Asymptotic 1/2 Near-Optimality*)

OSG instructs the robots to select actions that asymptotically achieve $1/2$ near-optimal coordination as the action-selection frequency increases.

Same performance as that of the near-optimal algorithm **Sequential Greedy** for predictable environments

*Achievable in environments that cannot adapt to the action-selection frequency of the algorithm

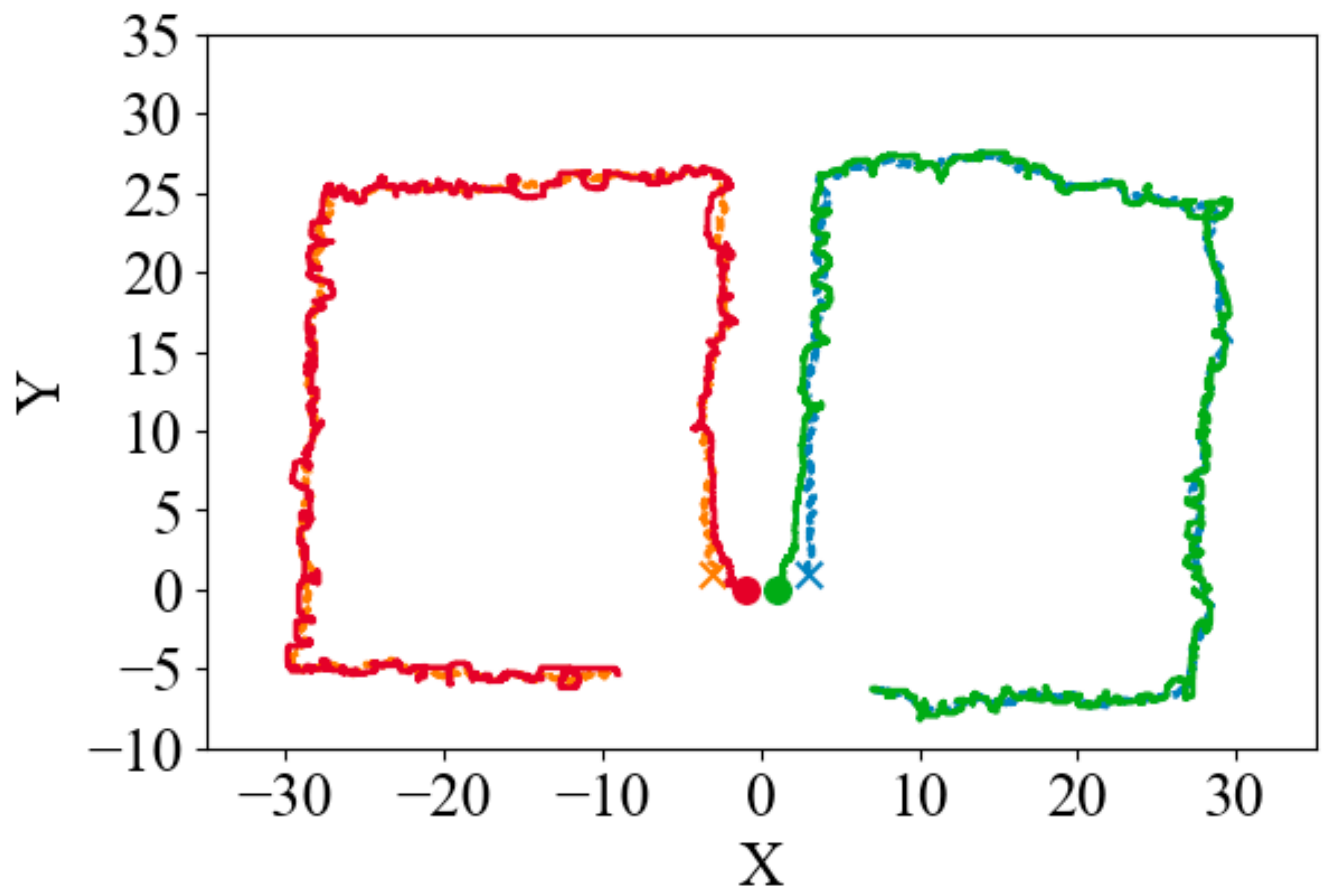
Goal: Minimize sum of distance from every target to its closest robot

Setup:

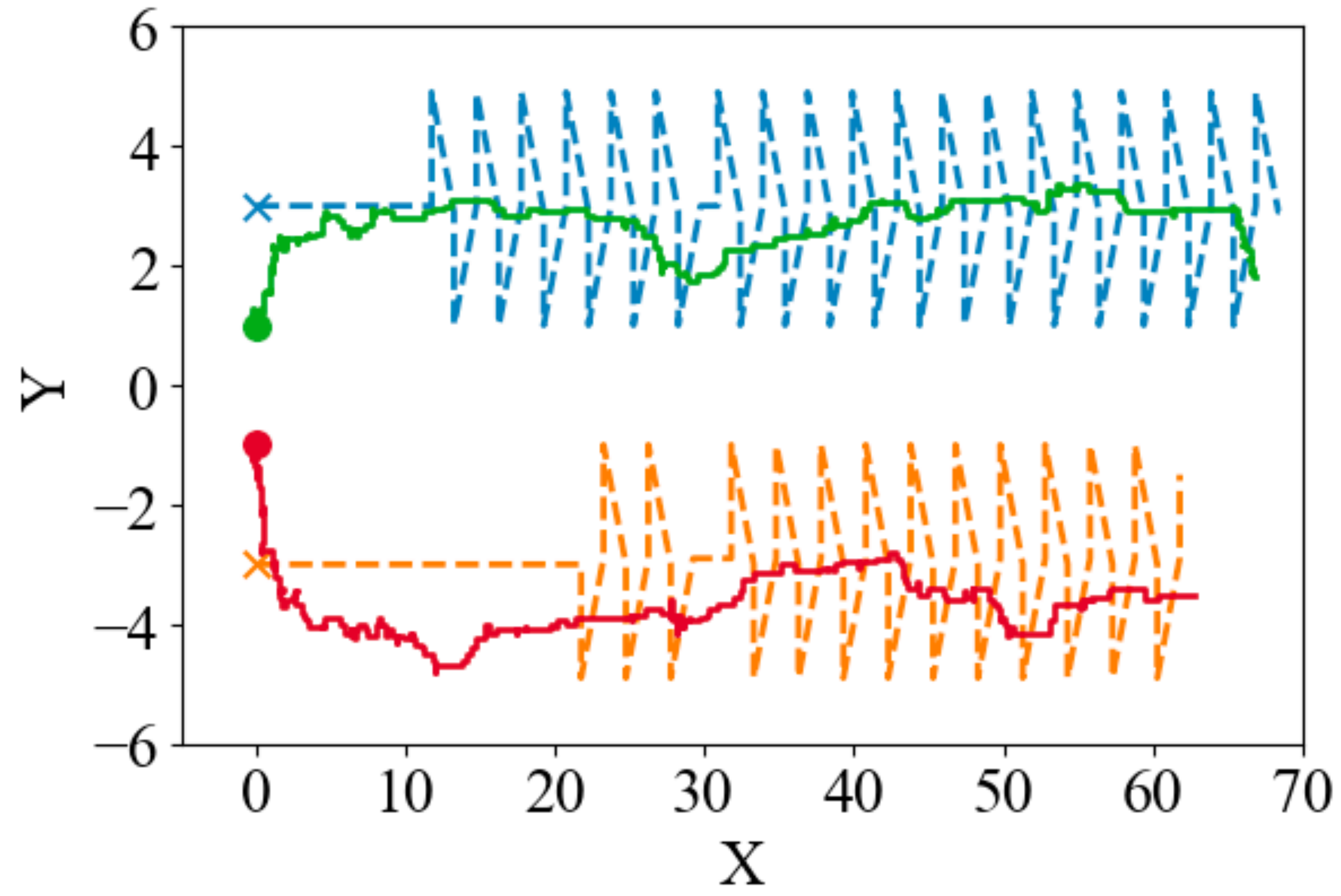
- The robots' field of view is unlimited
- The targets' future trajectories are unknown to the robots
- Two types of target trajectories:
  - Non-Adversarial: Pre-defined but corrupted with Gaussian noise
  - Adversarial: Adaptive to the robots' motion

Compared algorithm (SG-Heuristic): A heuristic algorithm selecting actions that 1/2 approximately

$$\max_{a_i \in \mathcal{V}_i,\, \forall i \in \mathcal{N}} f_{t-1}(\{a_i\}_{i \in \mathcal{N}})$$

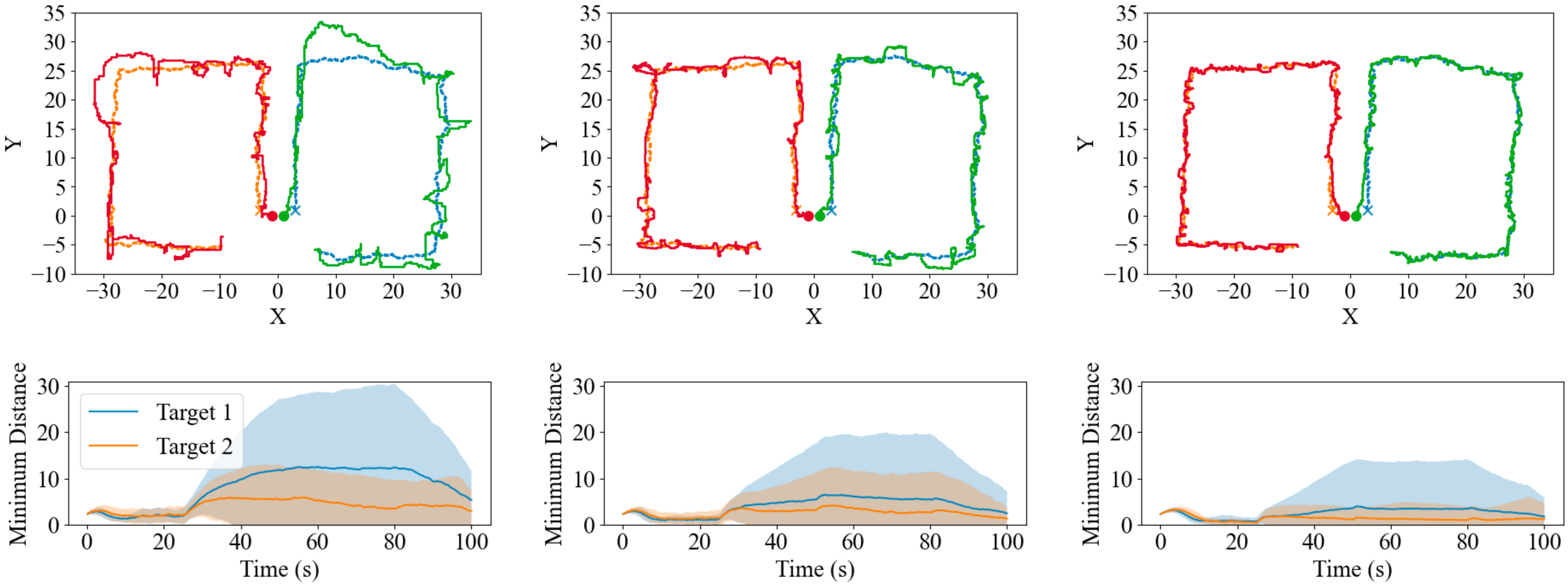


Non-Adversarial Target Motion



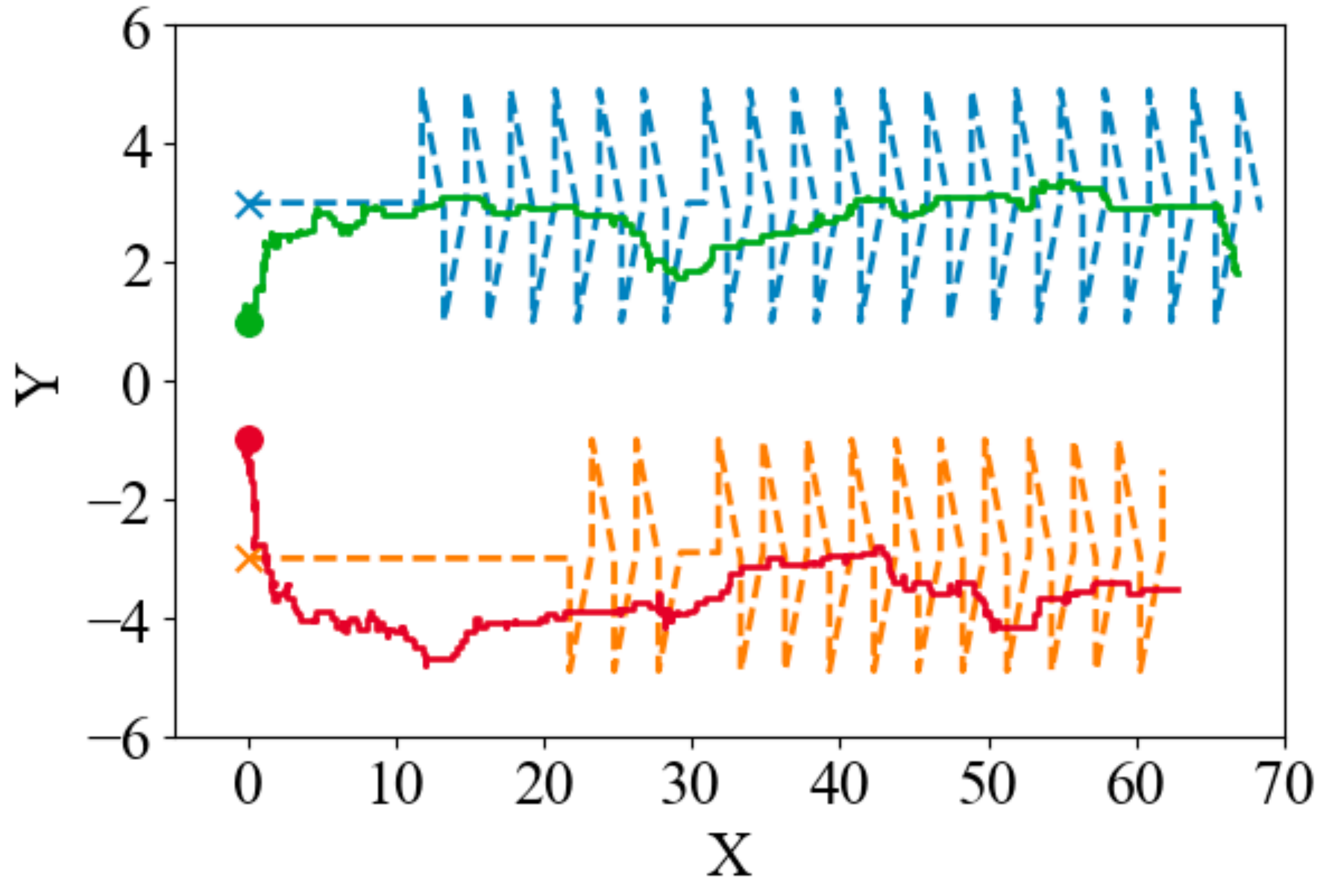
Adversarial Target Motion

Higher Action-Selection Frequency Improves Performance

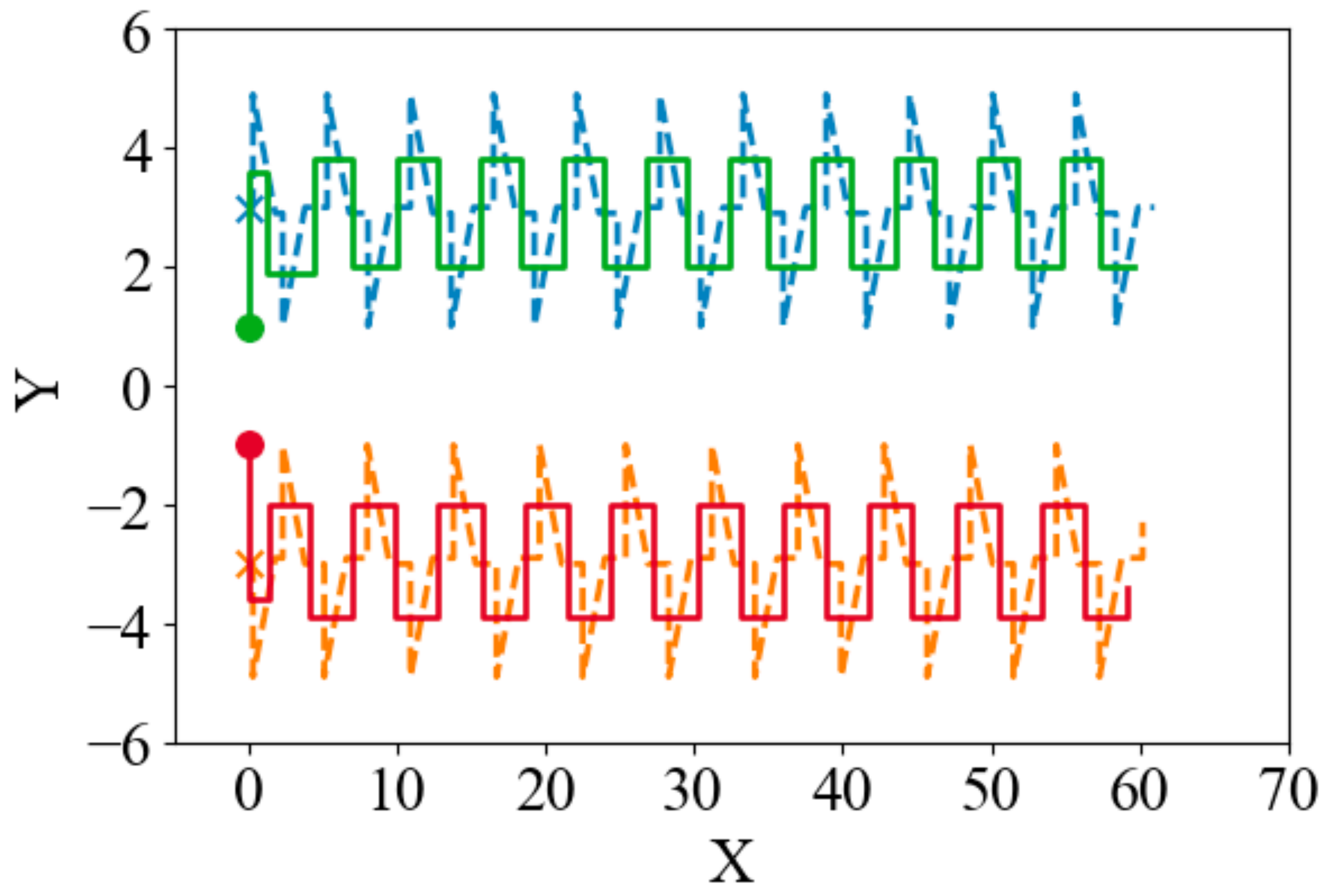


10Hz 20Hz 50Hz

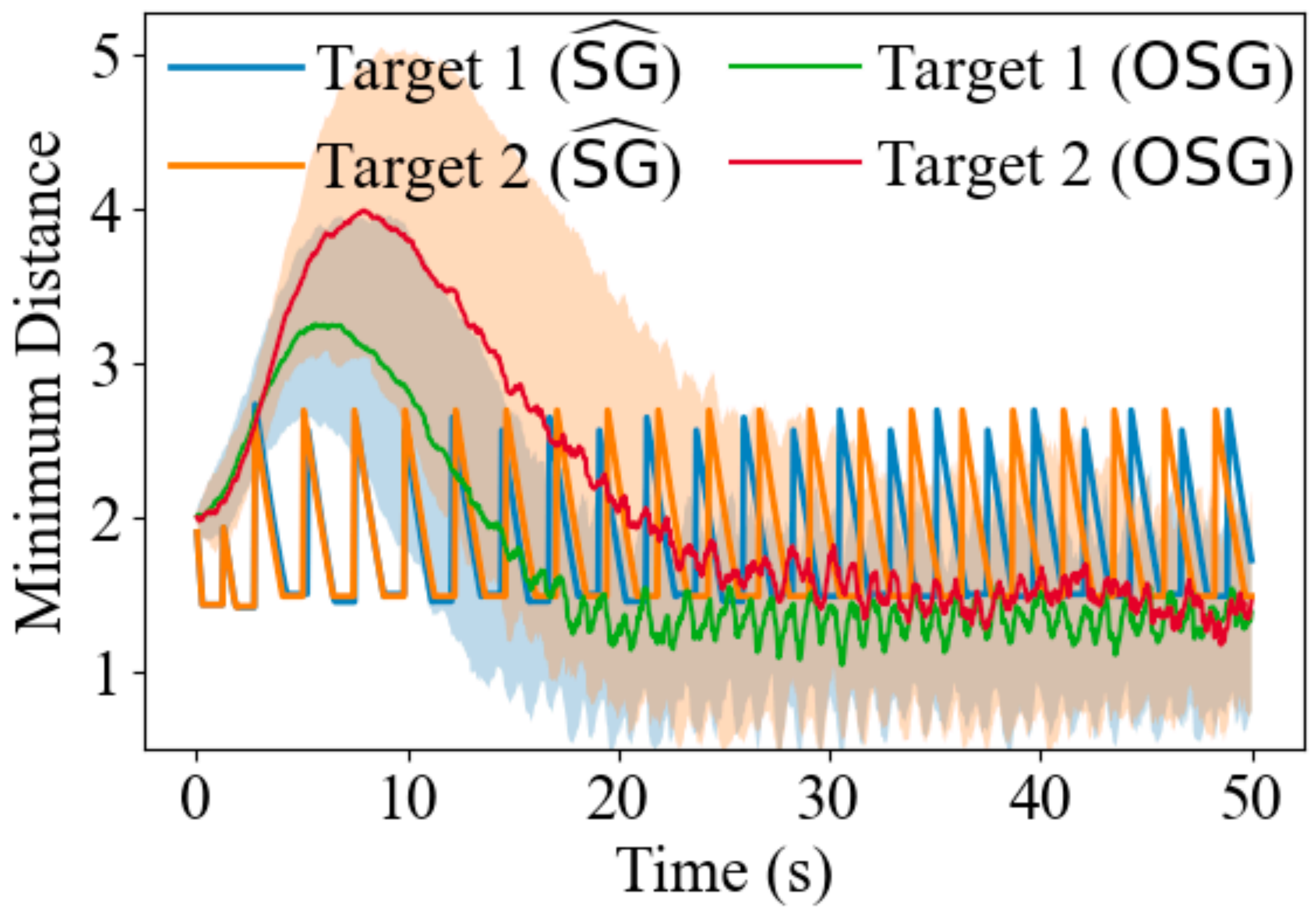
Total Minimum Target-Robot Distance: OSG with Different Action-Selection Frequencies in Non-Adversarial Scenarios



OSG



SG-Heuristic



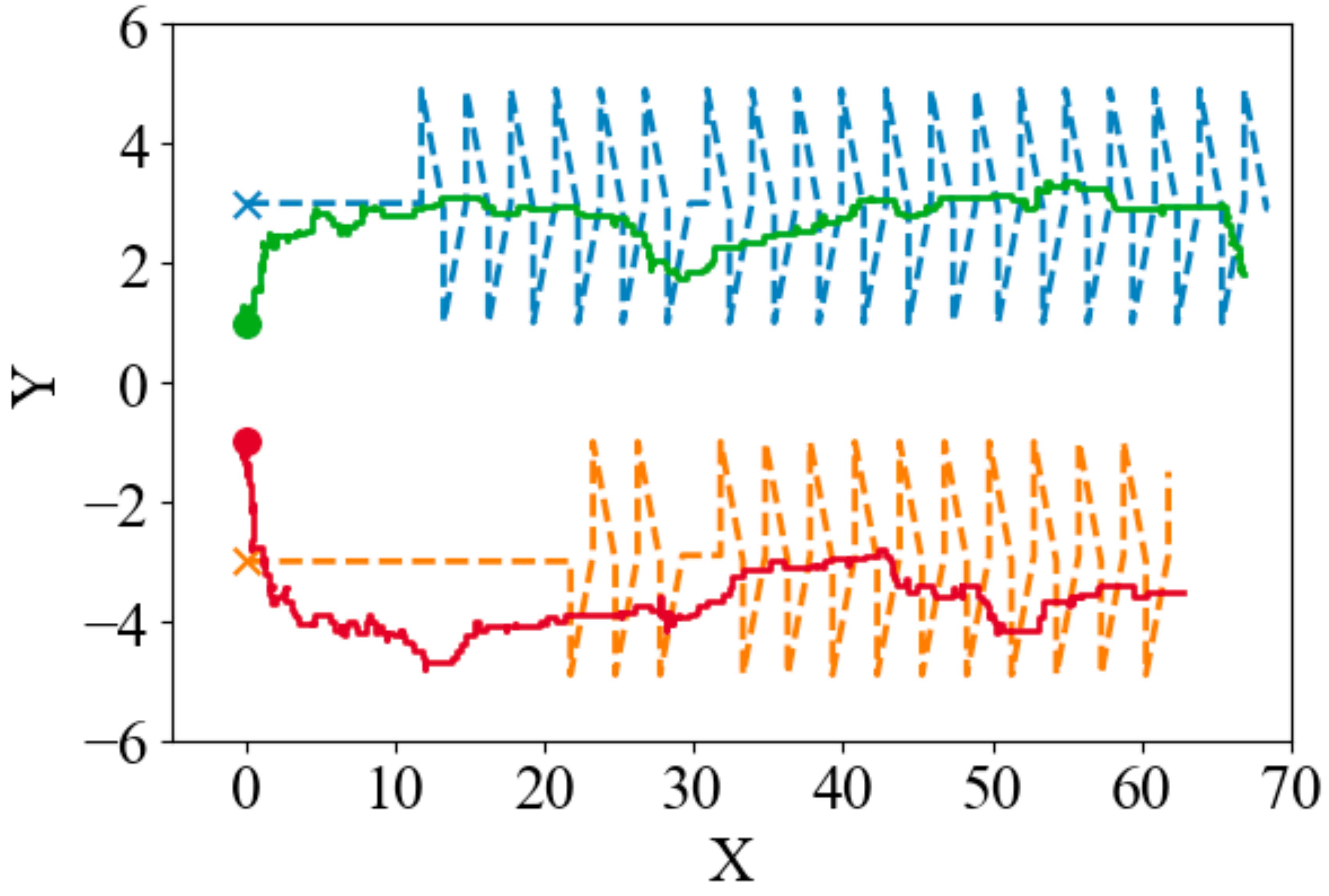
Minimum Distance: OSG vs. SG-Heuristic

Total Minimum Target-Robot Distance: OSG vs. SG-Heuristic in Adversarial Scenarios

Summary and Extensions

First coordination algorithm that achieves bounded suboptimality guarantee despite:

- *information overlap* among robot actions
- *completely unpredictable* future environment



Extensions:

- partial information feedback
- best of both worlds

Z. Xu, X. Lin, and V. Tzoumas, "Bandit submodular maximization for multi-robot coordination in unpredictable and partially observable environments," in Robotics: Science and Systems, 2023