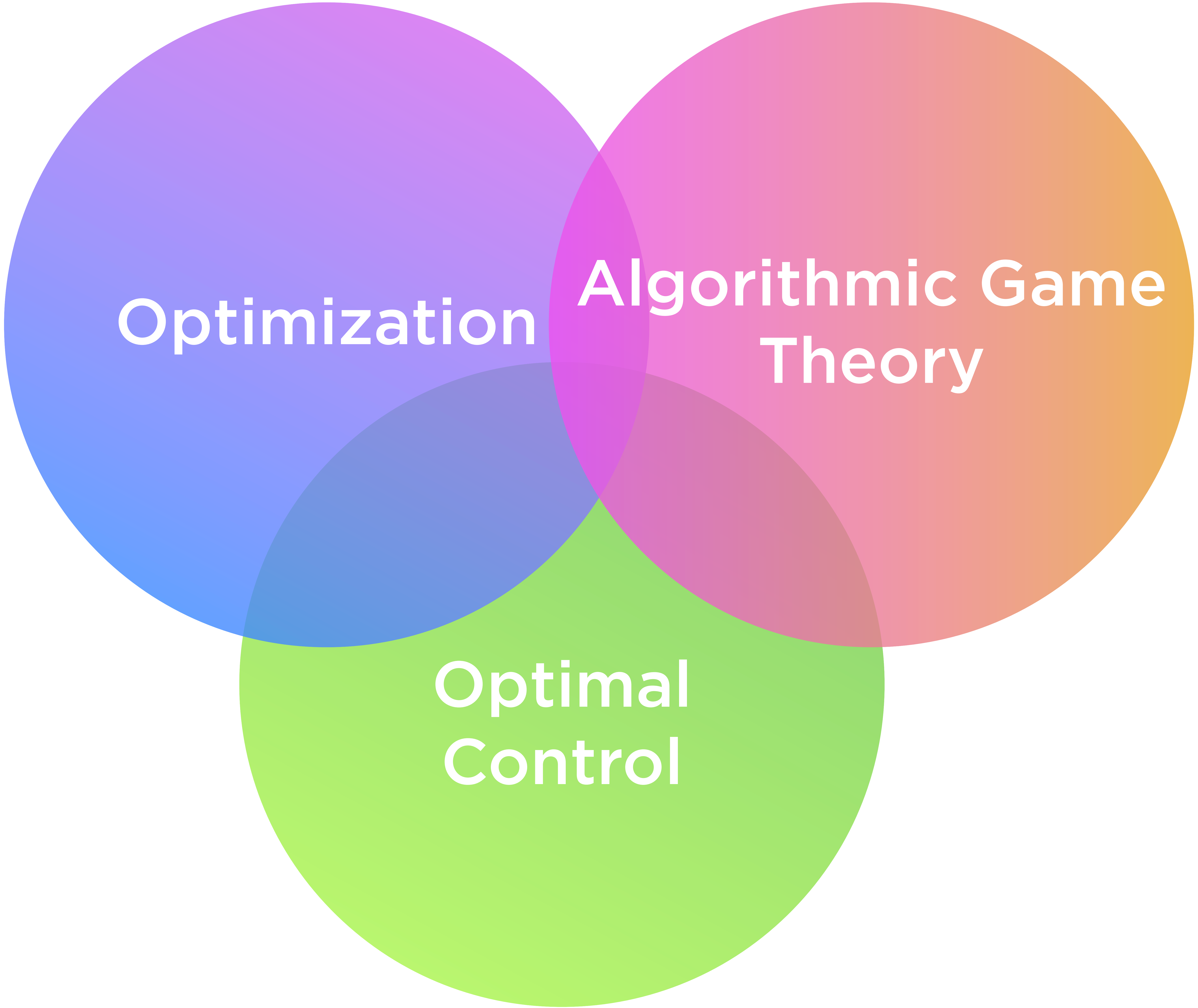
Who Plays First?

Haimin Hu*, ***Gabriele Dragotto****, Zixu Zhang, Kaiqu Liang, Bartolomeo Stellato, Jaime F. Fisac

PRINCETON UNIVERSITY


Optimization
Algorithmic Game Theory
Optimal Control

Regulating Airspace

is a complex FAA business…

Even more complex for UAVs...

Extremely complex when a myriad of ***autonomous*** UAVs interact

Preference and safety tradeoff?

Users

Want to **optimize their operations in real time**

- Avoid idle times,
- Match operations to their schedule,
- Plan trajectories
- Many players!

Regulator

Wants to (primarily) **ensure safety**

- **Collision-free** planning,
- Ensure fairness,
- Then, optimize efficiency.

Conflicting objectives

Wants to (primarily) **ensure safety**

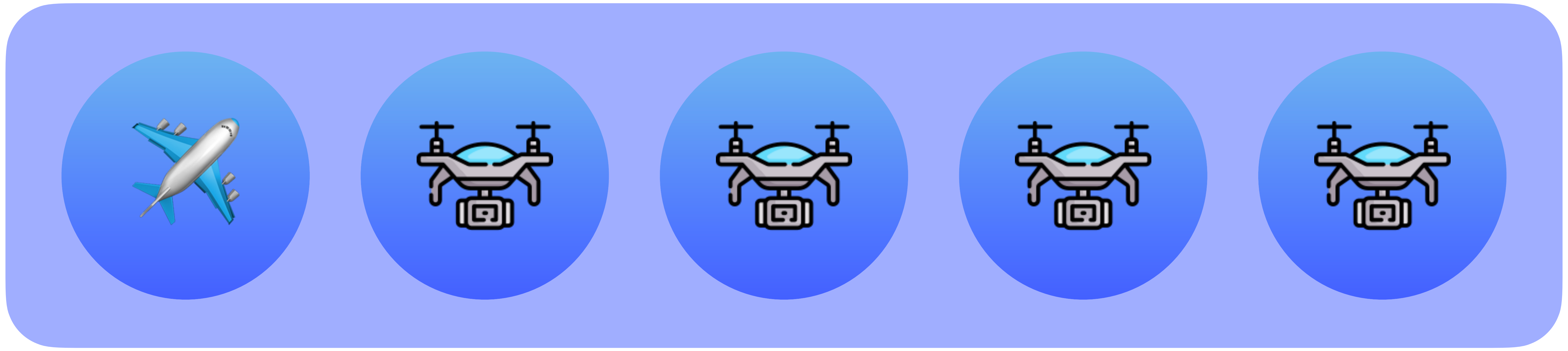
Set the **incentives**

Optimize their operations according to the incentives

The most natural incentive

How to determine the **order of play** (i.e., the order of trajectory planning) such that the **regulator optimizes its objectives**?

Who Plays First?



Problem abstraction

Enforces the order by setting $z_p = 1$ for some permutation $p \in \mathcal{P}$

Optimizes a cost function $f(x^1, \ldots, x^n) = f(x)$

Variables of all players but i

Play a strategy $x \in \mathcal{X}(z, x^{-i})$ factoring the order and the other players' strategies

Optimizes a cost function $g^i(x^i; x^{-i}, z)$

$$
\begin{aligned}
\min_{x,z,\hat{x}} \quad & f(x) \\
\text{subject to} \quad & (x_p^1, \ldots, x_p^n) \in \mathcal{S}(p), \qquad \forall p \in \mathcal{P} \\
& x = \sum_{p \in \mathcal{P}} z_p \hat{x}_p, \\
& z \in \{0,1\}^{|\mathcal{P}|}, \quad \sum_{p \in \mathcal{P}} z_p = 1
\end{aligned}
$$

"Blackbox"

Given a permutation

$$\text{subject to} \quad (x_p^1, \ldots, x_p^n) \in \mathcal{S}(p), \qquad \forall p \in \mathcal{P}$$

Explicit Blackbox

$\mathcal{S}(p)$ **is a linear system**

Given z, player i solves a **convex-quadratic equality-constrained Stackelberg game**

$$\min_{x^i} \quad g^i(x^i; x^{-i}, z) := \frac{1}{2}(x^i)^\top Q^i x^i + (c^i)^\top x^i + (x^{-i})^\top C^i x^i$$

Sequential "Stackelberg" (or n-level) game

$$x_{after}^{-i} \in OPT(x^{-i}, x^i)$$

"Followers" optimality

$$x^i \in \mathbb{R}^m$$

Implicit Blackbox

A blackbox can provide us the solution for the players

Given a permutation

$$\text{subject to} \quad (x_p^1, \ldots, x_p^n) \in \mathcal{S}(p), \qquad \forall p \in \mathcal{P}$$

Sequential Convex QP Game

$\mathcal{S}(p)$ **is a linear system**

Given z, player i solves a **convex-quadratic equality-constrained Stackelberg game**:

$$\begin{aligned} \min_{x^i} \quad & g^i(x^i; x^{-i}, z) := \frac{1}{2}(x^i)^\top Q^i x^i + (c^i)^\top x^i + (x^{-i})^\top C^i x^i \\ \text{subject to} \quad & A^i x^i + D^i x^{-i} - b^i = 0, \\ & x_{after}^{-i} \in OPT(x^{-i}, x^i) \\ & x^i \in \mathbb{R}^m \end{aligned}$$

"Followers" optimality

Blackbox

A blackbox can provide us the **solution for a given p**

STP converges to a Stackelberg equilibrium

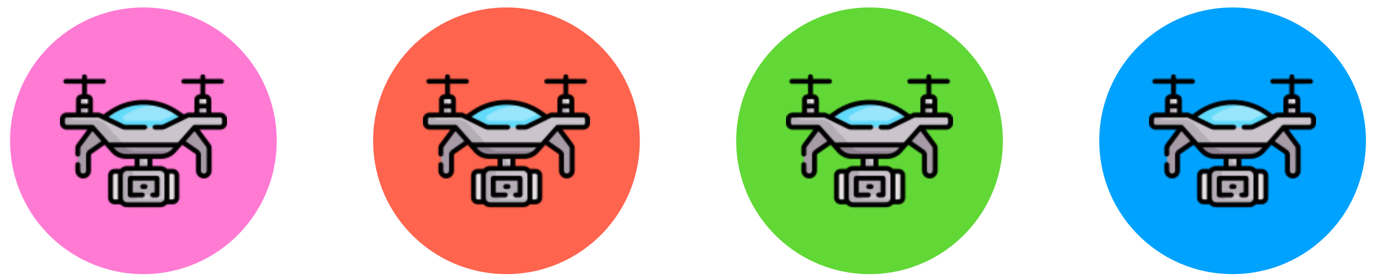
Sequential trajectory planning (STP), a popular multi-agent control approach, **produces a local Stackelberg equilibrium** in a single pass

Deterministic

Linear wrt n

Supports constraints

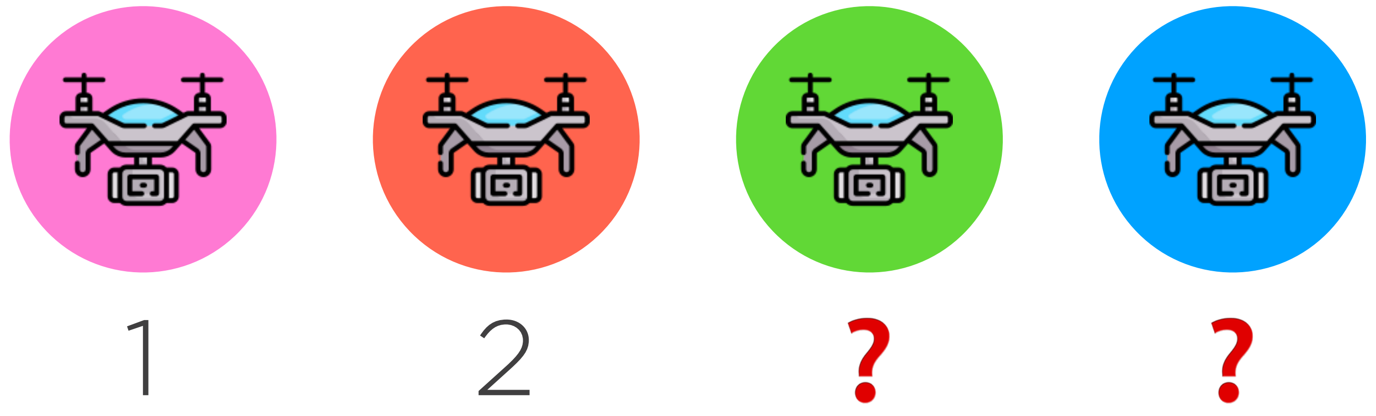
Two fundamental operations



1 **Partial Permutation**

A **partial assignment** $\tilde{p}$ of the order of play

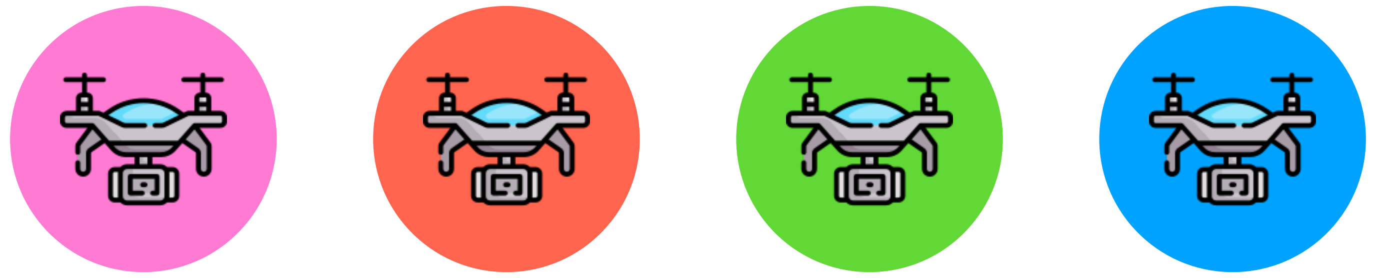
i.e., we know the first $k < (n-1)$ players playing



2 **Pricing $\tilde{p}$**

Assign an optimistic f-bound on $\tilde{p}$ **monotonically-increasing** w.r.t. any children

Pricing, in depth



2 **Pricing $\tilde{p}$**

Assign an optimistic f-bound on $\tilde{p}$ **monotonically-increasing** w.r.t. any children

Assuming the UAV i plays as the l-th player, it solves

$$
\begin{aligned}
\text{minimize} \quad & g^i(x^i; x^{-i}) := \textstyle\sum_{t=0}^{T} \ell^i(x_t^i, u_t^i) \\
\text{subject to} \quad & x_{t+1}^i = f^i(x_t^i, u_t^i) \\
& x_0^i = x_{init}^i \\
& x_t^i \in \mathcal{X}^i, \quad u_t^i \in U^i \\
& x_t^i \notin \mathcal{O}^i
\end{aligned}
$$

Trajectory Dynamics

Set of controls

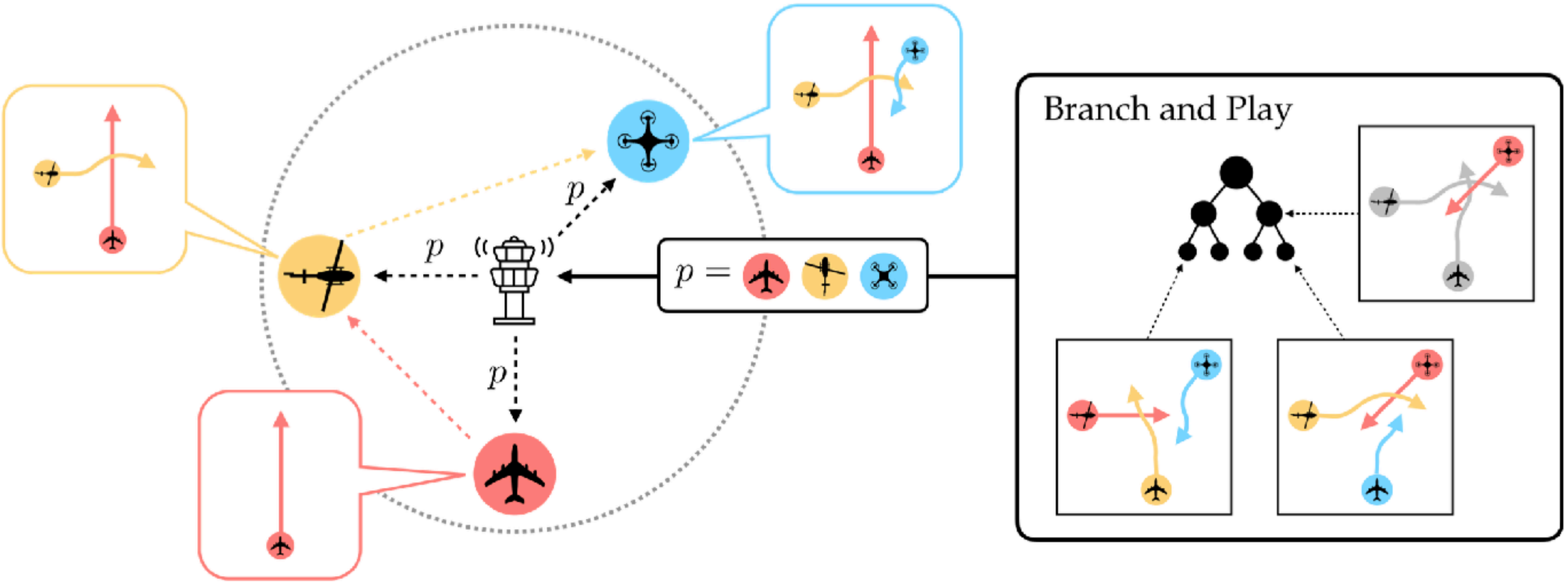
Obstacles and no-fly zones

The Branch-and-Play

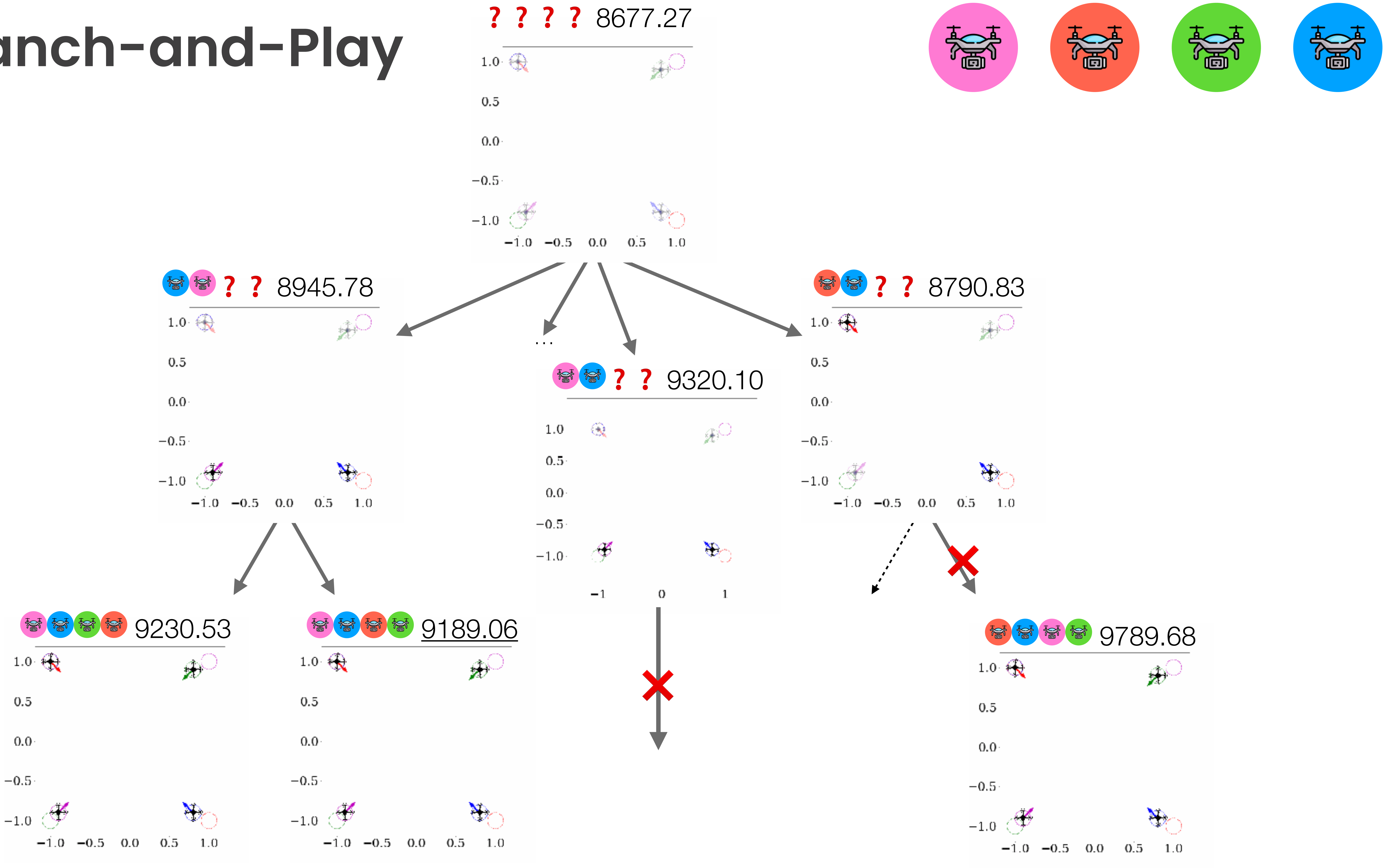
Branch-and-Play

Implicitly enumerate the space of permutations by:

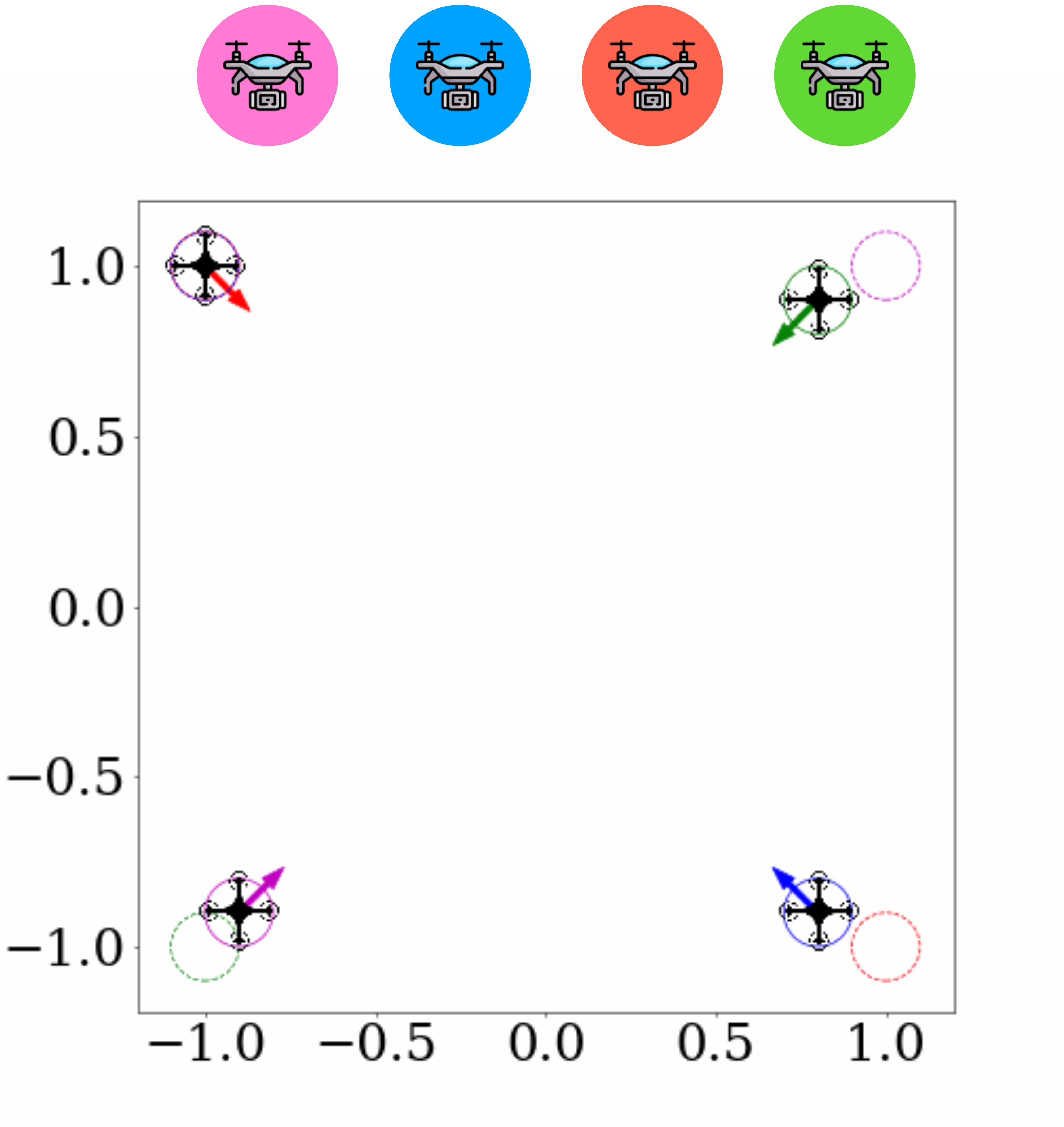
- **Pricing** of partial permutations
- **Pruning and exploration** strategies



The Branch-and-Play



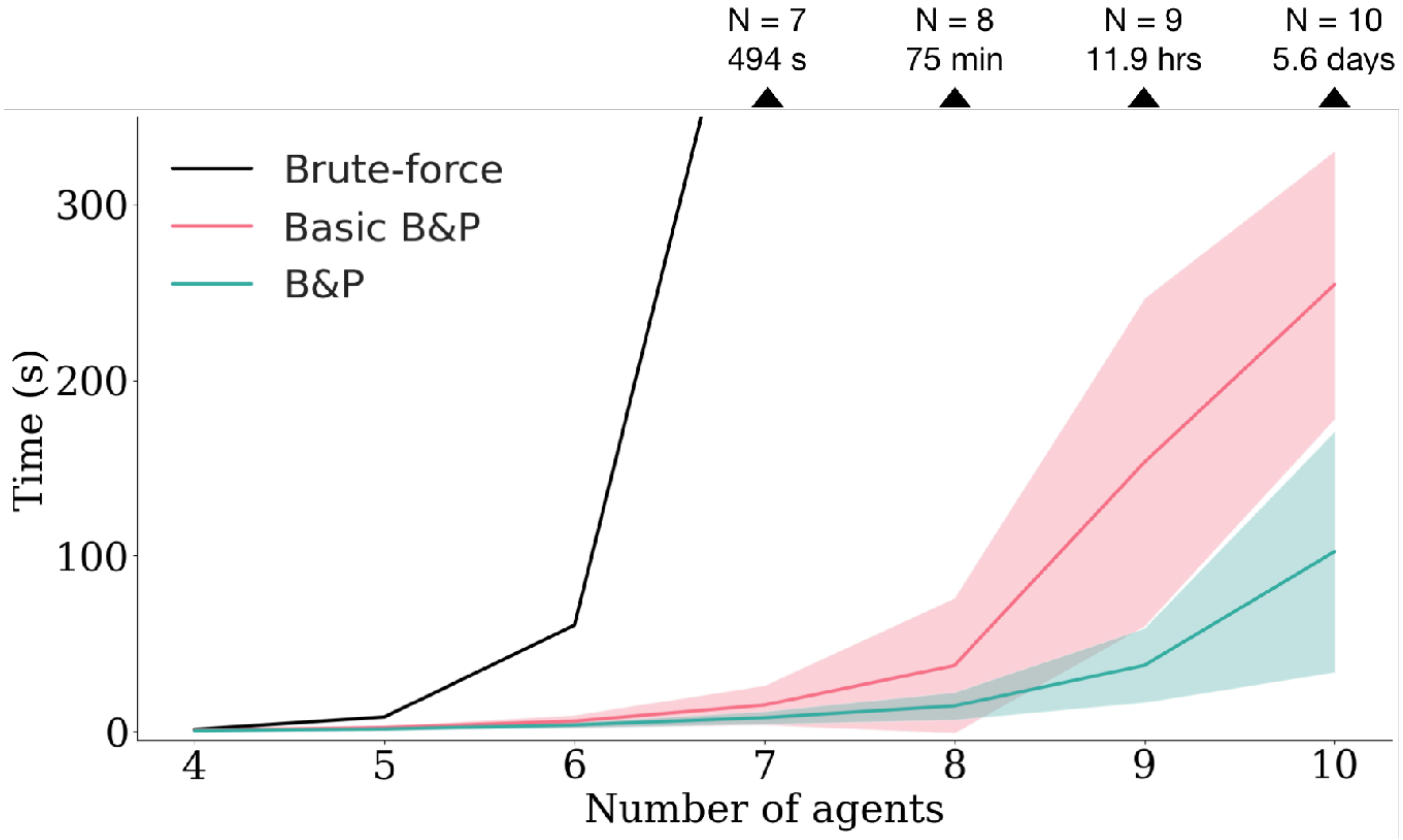
The Branch-and-Play



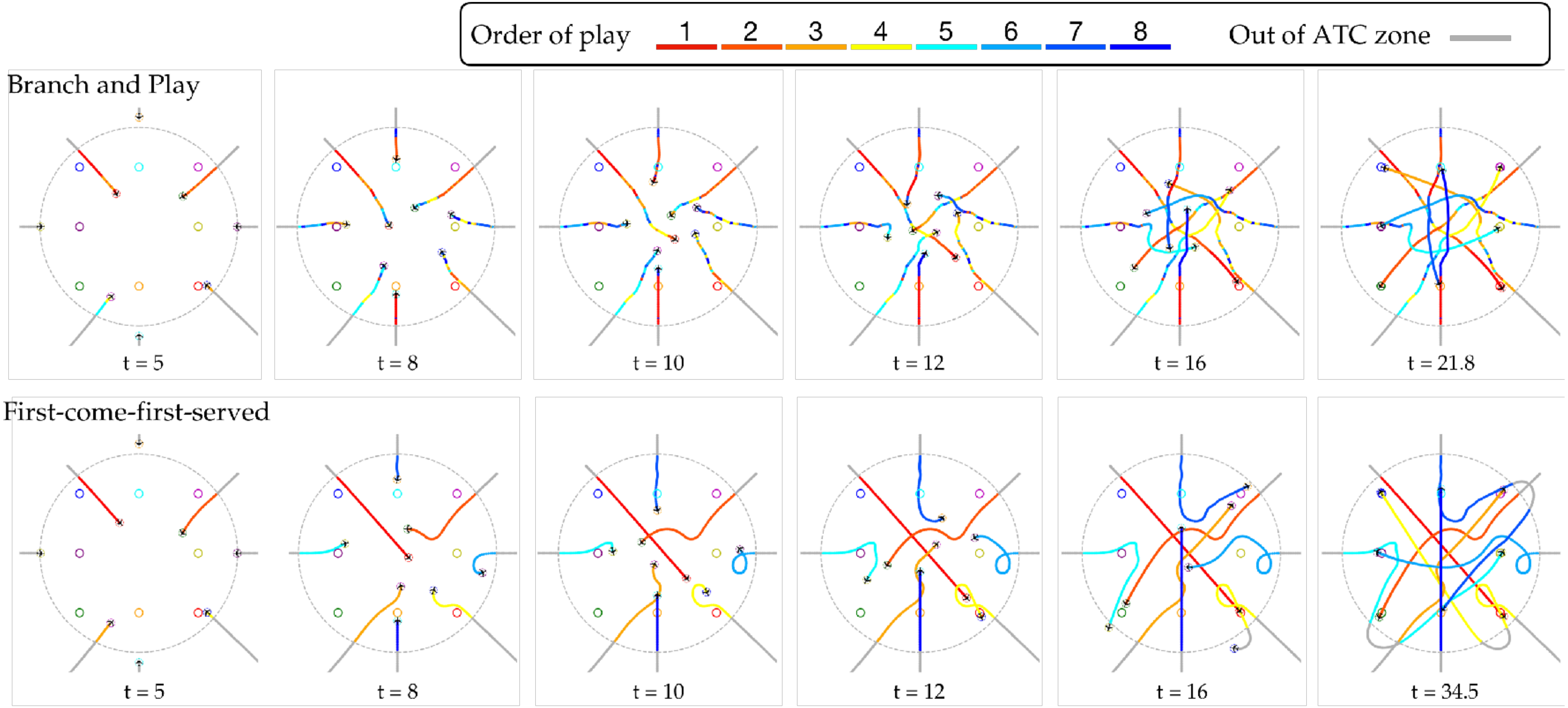
f-bound of <u>9189.06</u>

Experiments

BnP performs time-wise consistently in ATC



Visualizing ATC trajectories



Better flight times with ATC

Metric	N	FCFS	Randomized	Nash ILQ	B&P (ours)
Cost		1.24 ± 0.29	1.19 ± 0.23	3.61 ± 0.94	$\mathbf{1.04 \pm 0.24}$
Group (s)	4	25.76 ± 5.0	24.98 ± 4.06	44.50 ± 8.84	$\mathbf{20.41 \pm 2.72}$
T/O rate		1%	**0%**	36%	**0%**
Cost		1.73 ± 0.49	1.70 ± 0.43	5.40 ± 1.36	$\mathbf{1.41 \pm 0.3}$
Group (s)	5	29.27 ± 8.29	28.84 ± 6.74	48.34 ± 11.4	$\mathbf{20.5 \pm 6.04}$
T/O rate		6%	2%	64%	**1%**
Cost		2.43 ± 0.82	2.55 ± 1.09	8.73 ± 2.13	$\mathbf{2.0 \pm 0.51}$
Group (s)	6	31.83 ± 9.34	33.62 ± 12.05	50.9 ± 11.92	$\mathbf{22.18 \pm 7.74}$
T/O rate		9%	20%	96%	**2%**

Branch-and-Play **outperforms 3 baselines** in all three metrics

Swarm formation

Autonomous vehicles

Modular Algorithm — Adapts to a variety of applications. Enables **customizations**

Actionable Insights — Provides **interpretable interventions**

Solid Benchmark Results — Guarantees **tangible improvements** over the SOTA

www.dragot.to

Who Plays First? Optimizing the Order of Play in Stackelberg Games with Many Robots

Haimin Hu, Gabriele Dragotto, Zixu Zhang, Kaiqu Liang, Bartolomeo Stellato, Jaime F. Fisac

arXiv 2402.09246

DataX — Accelerating Scientific Discovery at Princeton

PRINCETON UNIVERSITY