A Regret Minimization Approach to Multi-Agent Control

Udaya Ghai, Udari Madhushani, Naomi Leonard, Elad Hazan

What can be controlled by independent agents?

Can we reduce complex systems to independent agents?

Independent regret minimizing

Independent actuators

- Limited computation
- Change/add motors
- Fault tolerance
- Noise robustness

Related Work

- Multi-Agent RL (MARL) mostly empirical. Recent game-theoretic guarantees [Jin et al., 2021, Golowich et al. ‘22]

- Distributed/decentralized control **centrally precomputed** stabilizing policies [Wang and Davidson, 1973, Wang et al., 2020]

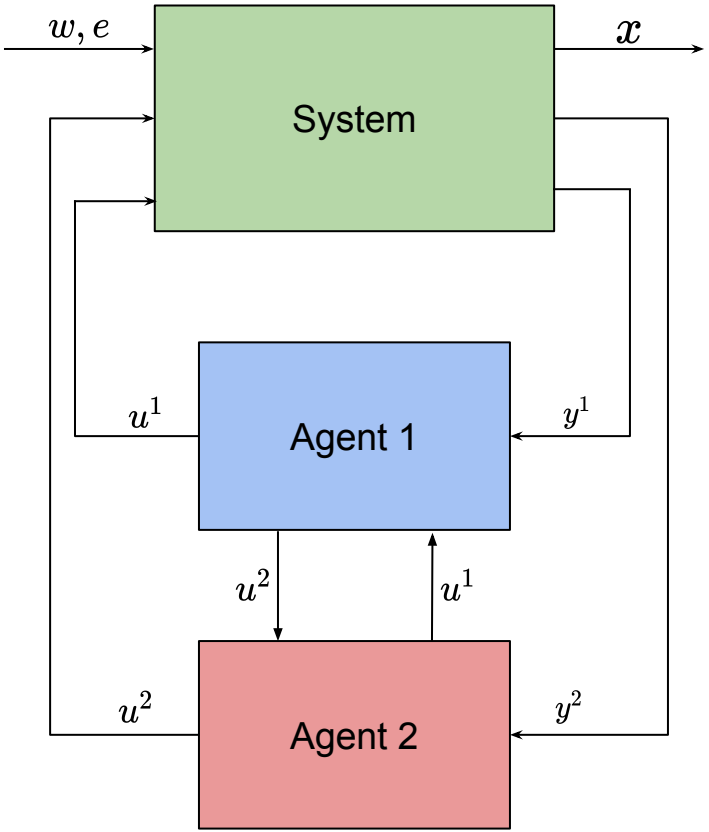
- Multi-agent extension of nonstochastic control framework **= instance optimal** guarantees [Agarwal et al., 2019].

- Our focus: provable **global** guarantees from **independent** regret-minimizing agents

Multi-Agent Control(MAC) Setting

$$y_t^i = g_i(x_t) + e_t^i, \quad x_{t+1} = f(x_t, u_t) + w_t,$$

- Local observation y_t^i and control u_t^i
- Agents see joint control $u_t = [u_t^1, \ldots, u_t^k]$
- **Adversarial** disturbances w_t and e_t^i
- **Cooperative** = single convex cost $c_t(x_t, u_t)$



Multi-Agent Policy Regret

- Single agent policy class $\prod_i$
- Joint comparator policy class $\prod = \prod_1 \times \cdots \times \prod_k$
- Multi-agent regret vs. **counterfactual trajectory** from the **best joint policy** in hindsight

$$\mathcal{R}_T(\mathcal{C}) = \frac{1}{T}\sum_{t=1}^{T} c_t(x_t, u_t) - \inf_{\pi \in \Pi} \frac{1}{T}\sum_{t=1}^{T} c_t(x_t^{\pi}, u_t^{\pi})$$

Our Results

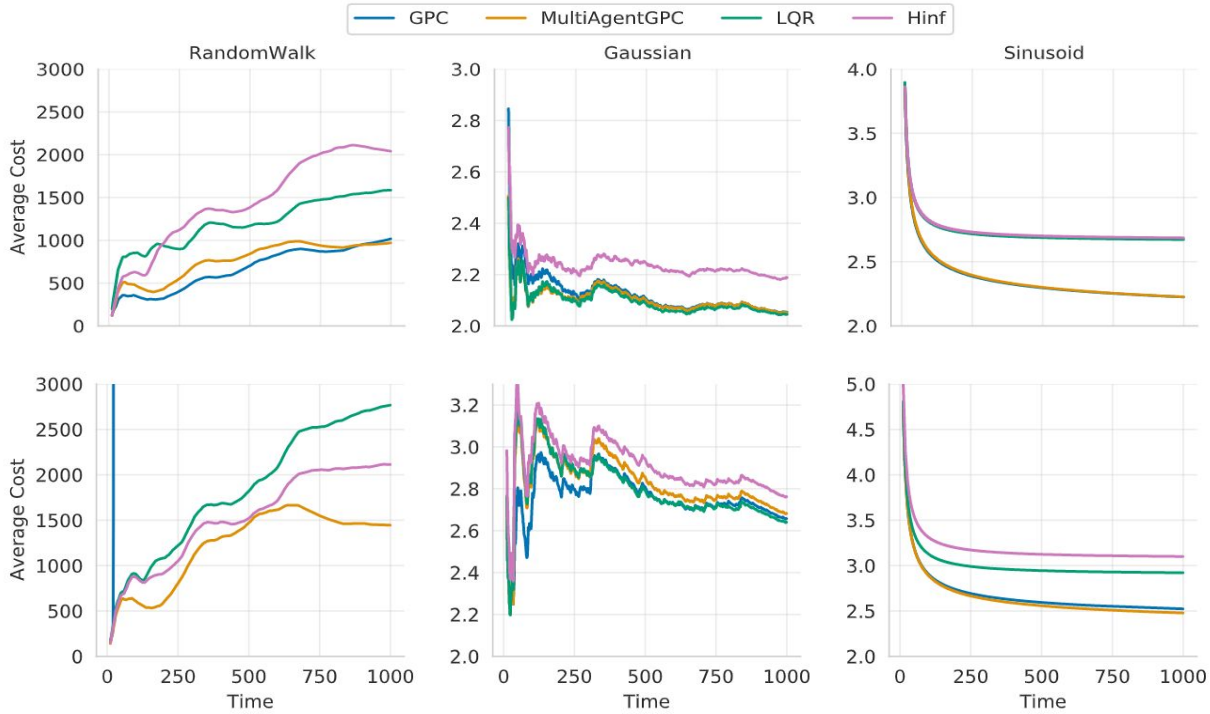
Reduction: Multi-Agent control → single-agent regret minimization $\mathcal{A}_i$ where

1. Single agent regret $\mathcal{R}_T(\mathcal{A}_i)$
2. Approximately policy evaluation
3. Agent policies are independent
4. Joint cost function is convex

Our meta-algorithm, has multi-agent regret bounded

$$\mathcal{R}_T(\mathcal{C}) \leq \sum_{i=1}^{k} \mathcal{R}_T(\mathcal{A}_i) + \tilde{O}(\frac{1}{T} + \epsilon)$$

Synthetic experiments



ADMIRE overactuated airplane, before/after deactivating fourth control

Future Directions

- Can similar regret metrics and algorithmic ideas be used in MARL?
- Can the information model of the setting be relaxed?
- To what extent can multi-agent regret minimizing agents provide robustness?