LEMURS: LEarning distributed MUlti-Robot interactionS

Eduardo Sebastián, Thai Duong, Nikolay Atanasov, Eduardo Montijano and Carlos Sagüés

Departamento de Informática e Ingeniería de Sistemas,
Universidad de Zaragoza

Department of Electrical and Computer Engineering
University of California, San Diego

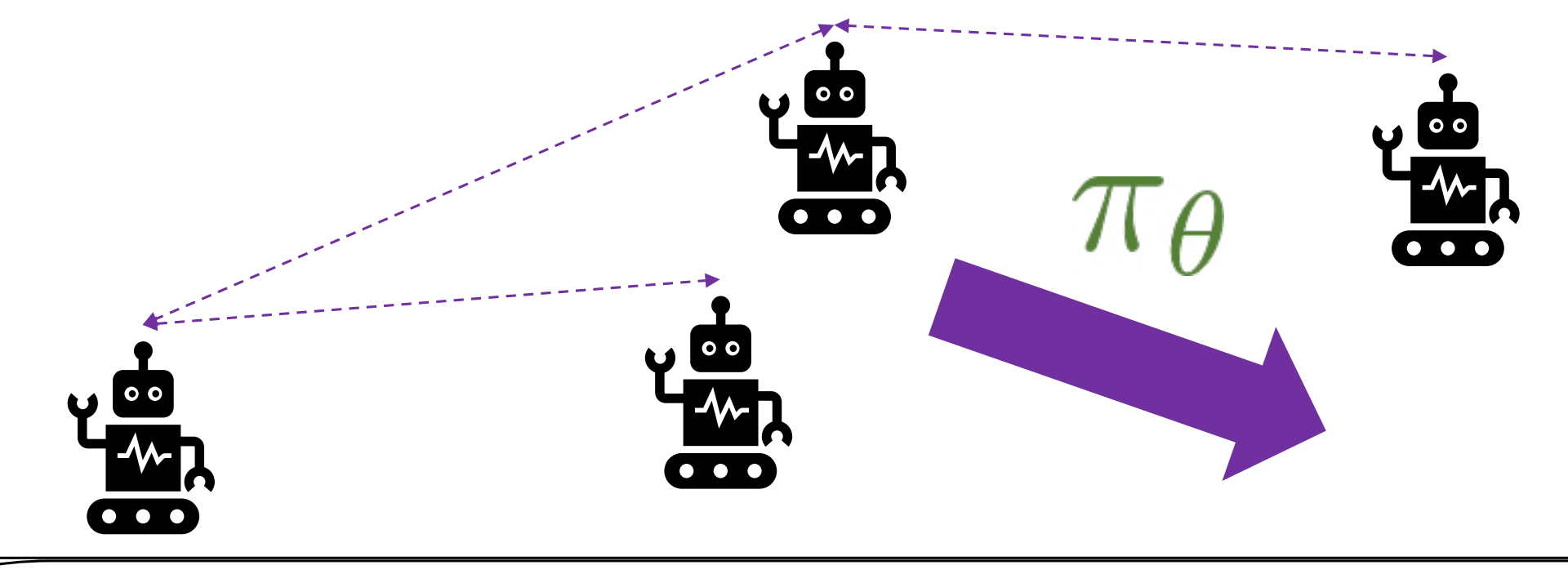
SCAN ME

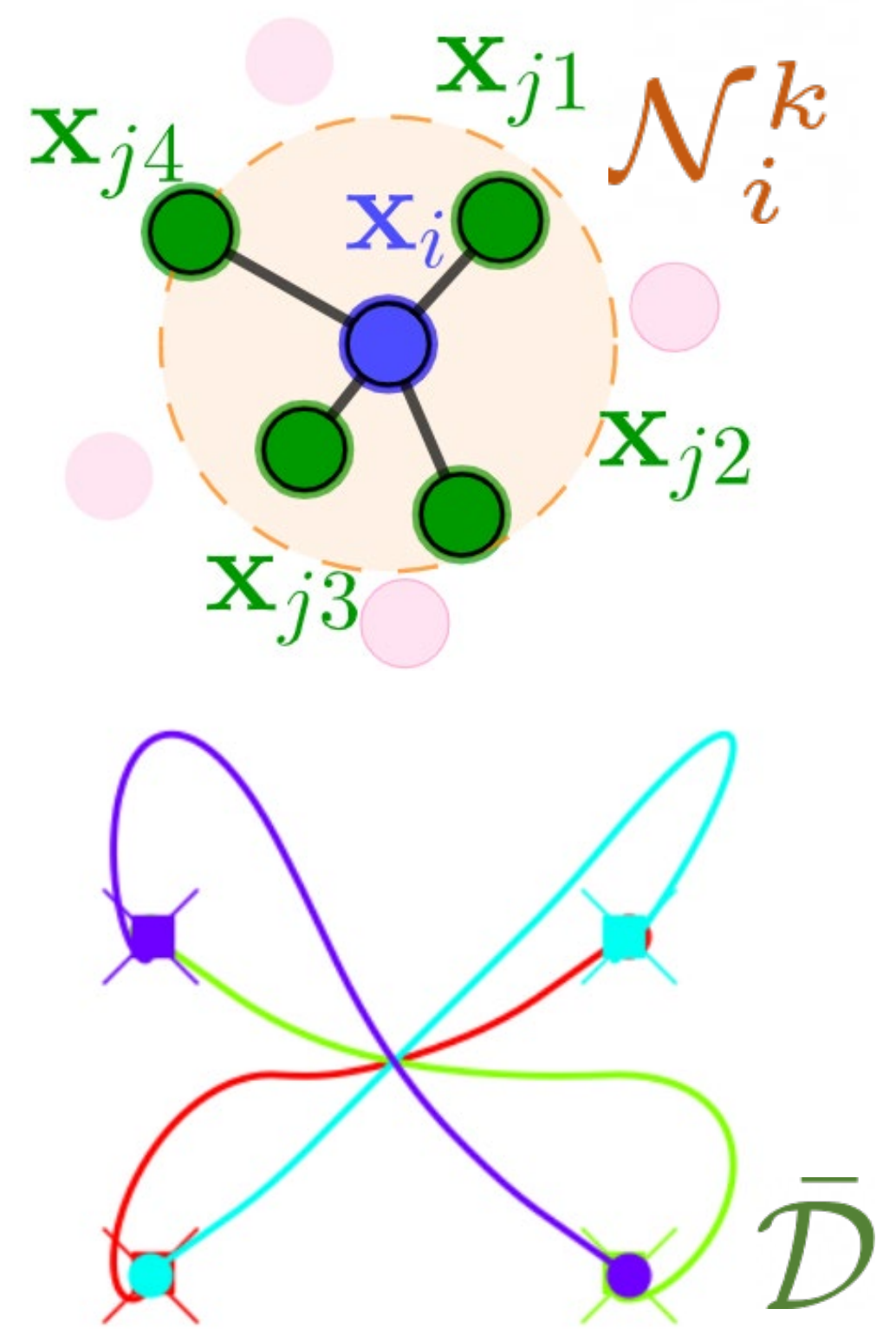
LEMURS learns scalable, distributed and stable multi-robot control policies that replicate cooperative tasks from demonstrations

I. Problem Statement

from task demonstrations

to multi-robot policies





known robot dynamics

$$\dot{\mathbf{x}}_i(t) = \mathbf{f}_i(\mathbf{x}_i(t), \mathbf{u}_i(t))$$

distributed control policy

$$\mathbf{u}_i(t) = \pi_\theta\left(\mathbf{x}_{\mathcal{N}_i^k}(t)\right)$$

loss to optimize

$$\min_\theta \mathcal{L}(\mathcal{D}, \bar{\mathcal{D}})$$

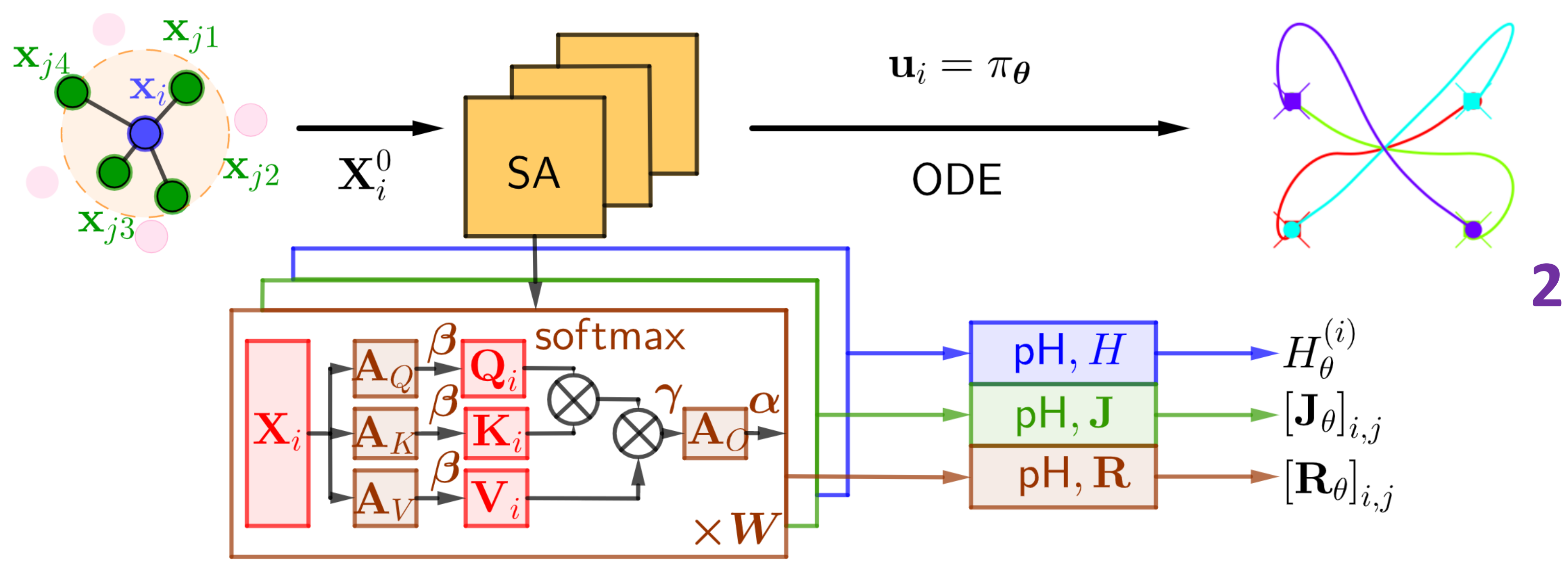
state and **input** of robot i

neighbors of i

control policy

learned and **demonstrated** trajectories

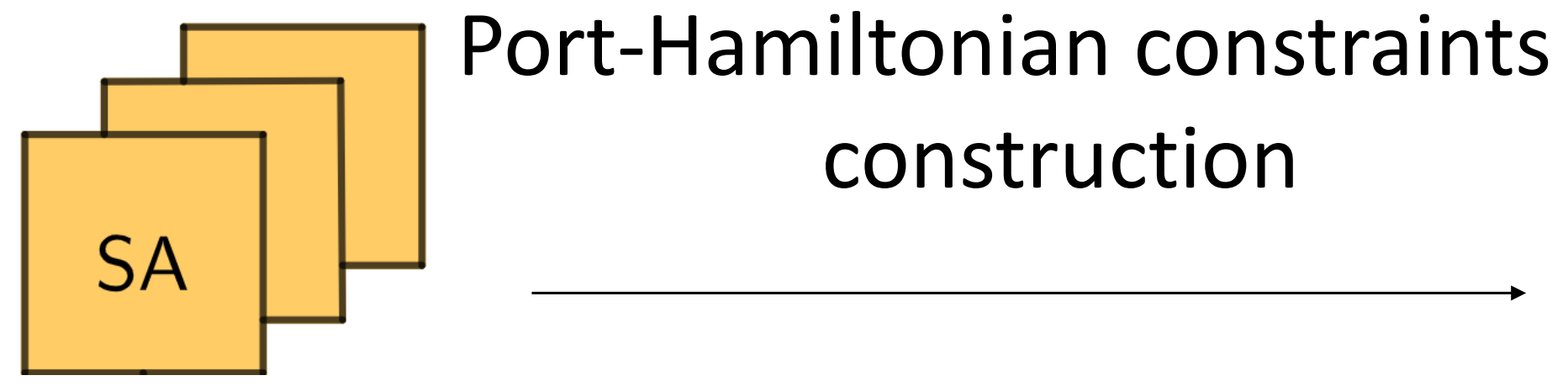
II. LEMURS



1. **Port-Hamiltonian modelling**: stable and distributed

$$\dot{\mathbf{x}} = (\mathbf{J}_s(\mathbf{x}) - \mathbf{R}_s(\mathbf{x}))\frac{\partial H_s(\mathbf{x})}{\partial \mathbf{x}} + \mathbf{F}_s(\mathbf{x})\mathbf{u}$$

2. **Self-attention**: time-varying team sizes



Port-Hamiltonian constraints by construction

$$[\mathbf{R}_\theta]_{ij} \quad [\mathbf{J}_\theta]_{ij} \quad H_\theta^{(i)}$$

3. **Neural Ordinary Differential Equations**: continuous-time dynamics. Backpropagation using the same ODE solver

Backward propagation

$$(\mathbf{x}_i(t_0), \boxed{\partial\mathcal{L}/\partial\theta}) = \text{ODESolver}(\mathbf{x}_i(t_f), \mathbf{f}_i, \theta)$$

Forward propagation

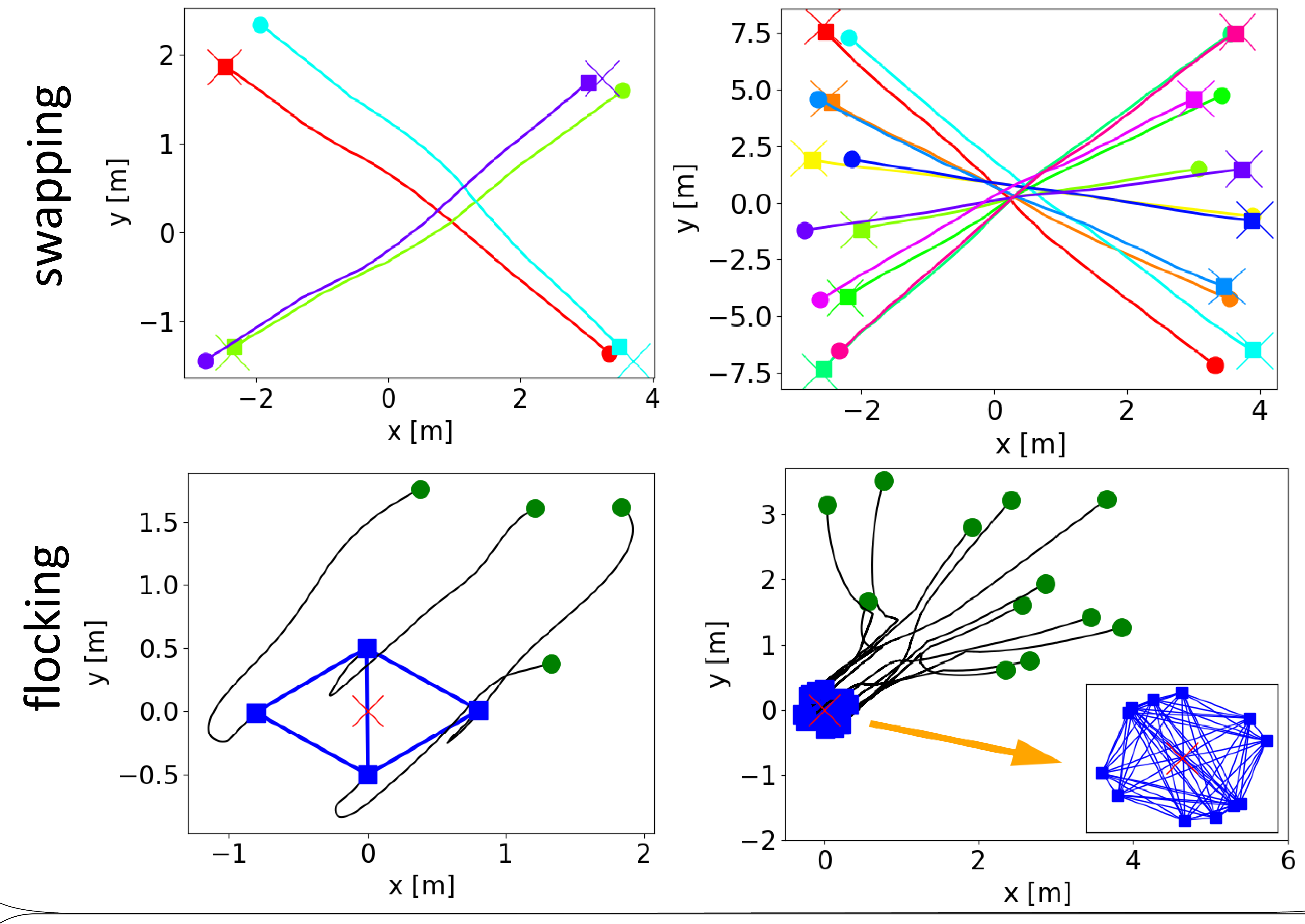
$$\mathbf{x}_i(t_f) = \text{ODESolver}(\mathbf{x}_i(t_0), \mathbf{f}_i, \theta)$$

This work has been supported by NSF CCF-2112665 (TILOS), by the ONR Global grant N62909-19-1-2027 and via Spanish projects PID2021-125514NB-I00, PID2021-124137OBI00 and TED2021-130224BI00 funded by MCIN/AEI/10.13039/501100011033, by ERDF A way of making Europe and by the European Union NextGenerationEU/PRTR, DGA T45-20R, and Spanish grant FPU19-05700 and EST22/00253.

III. Results

demonstration
only 4 agents

learned policy
scales to many robots!



IV. Takeaways

LEMURS learns robot interactions from demonstrations using self-attention and Hamiltonian-based neural ODE networks.

LEMURS advances the state of the art by learning control policies that **generalize to increasing numbers of robots and time-varying communications**, from just state-only trajectories of few robots.