软件开发环境国家重点实验室
State Key Laboratory of Software Development Environment

北京航空航天大學
BEIHANG UNIVERSITY

北京大學
PEKING UNIVERSITY

NANYANG TECHNOLOGICAL UNIVERSITY

Empirical Study on Robustness and Resilience in Cooperative Multi-Agent Reinforcement Learning

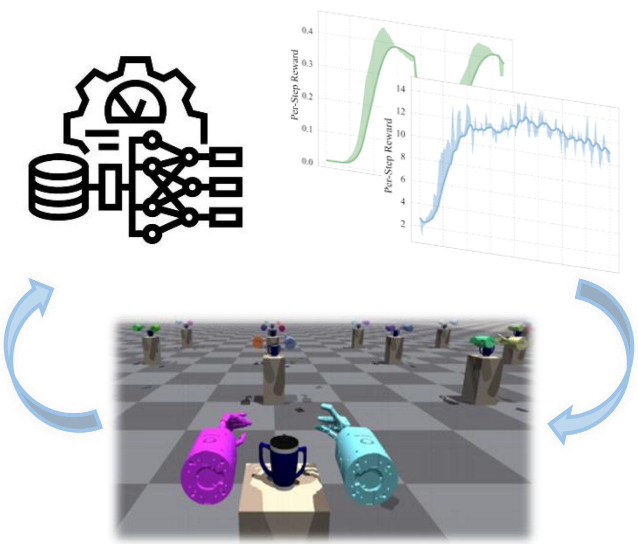
Paper ID 8858

Simin Li, Zihao Mao, Hanxiao Li, Zonglei Jing, Zhuohang Bian, Jun Guo, Li Wang
Zhuoran Han, Ruixiao Xu, Xin Yu, Chengdong Ma, Yuqing Ma, Bo An
Yaodong Yang, Weifeng Lv, Xianglong Liu

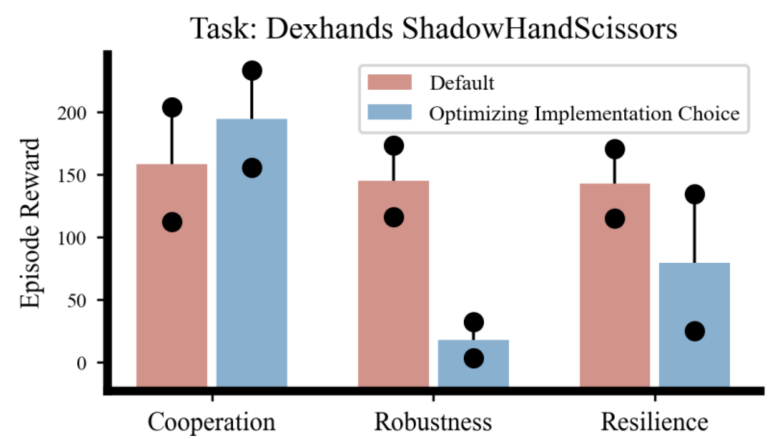
Introduction

Background

- Hyperparameter tuning is a common practice to maximize cooperative performance in cooperative Multi-Agent Reinforcement Learning (MARL).
- However, policies tuned for cooperation often fail under real-world uncertainties such as noise, delays, or perturbations.



Optimizing hyperparameters to improve cooperation performance in MARL



The model gets significantly less robust and resilient when uncertainty occurs

Introduction

Introducing Robustness and Resilience

To build trustworthy MARL systems, we need a deeper understanding of:

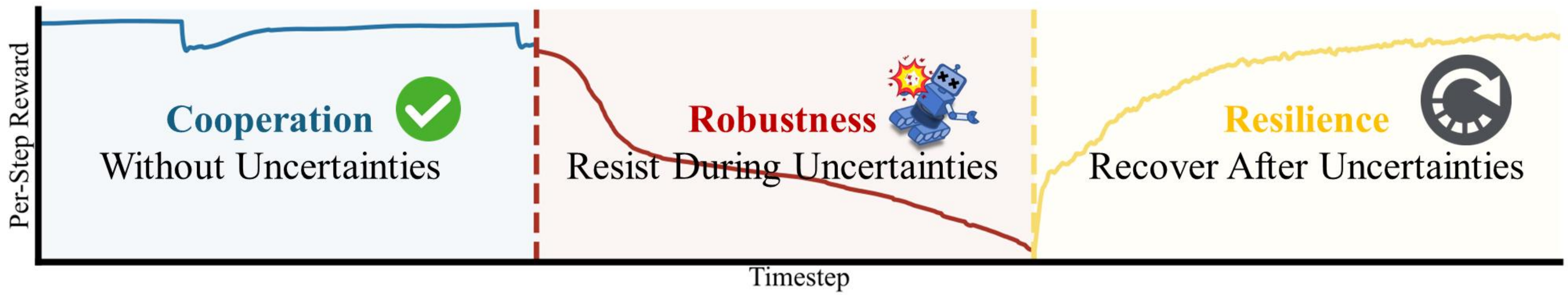
- *Robustness*: The stability under uncertainties

$$J^{\text{robust}}(\pi) = \mathbb{E}_{u\sim\mathcal{U}}\left[\mathbb{E}_{s_0\sim\rho_0}\mathbb{E}_{\pi,u}\left[\sum_{t=0}^{\infty}\gamma^t r_t \,\middle|\, s_0\right]\right].$$

- *Resilience*: the ability to recover from disrupted states

$$J^{\text{resilience}}(\pi) = \mathbb{E}_{u\sim\mathcal{U}}\left[\mathbb{E}_{s_u\sim\rho_u}\mathbb{E}_{\pi}\left[\sum_{t=t_u}^{\infty}\gamma^t r_t \,\middle|\, s_0 = s_u\right]\right].$$

where $\pi \in \Pi$ is a fixed policy, $u \in \mathcal{U}$ represents a perturbation induced by uncertainties, ρ_0 is the initial state distribution conditioned on normal states, ρ_u is the initial state distribution conditioned on perturbed states



Relation between cooperation, robustness, and resilience under uncertainty. Cooperative MARL is trained without perturbations, but must be robust and resilient when they occur.

Introduction

Contributions

In this work:

- We present a large-scale empirical study comprising over 82,620 experiments to evaluate cooperation, robustness, and resilience in MARL across 3 algorithms, 4 real-world environments, 13 uncertainty types, and 15 hyperparameters.

- We identify 4 main findings and recommended practices for handing different scenarios in MARL.

- By optimizing hyperparameters only, we observe substantial improvement in cooperation, robustness and resilience across all MARL backbones, with the phenomenon also generalizing to robust MARL methods across these backbones.

Experiment Procedure

Experiment Setups

- 3 Algorithms:
  - MADDPG, MAPPO, HAPPO
- 4 Real-World Environments:
  - Dexterous Hand Manipulation
  - Quadrotor Swarm Control
  - Intelligent Traffic Control
  - Active Voltage Control
- 15 Hyperparameters
- 13 Types of Uncertainties spanning:
  - Observation, Action, Environment uncertainties
  - Applied to single/all agents

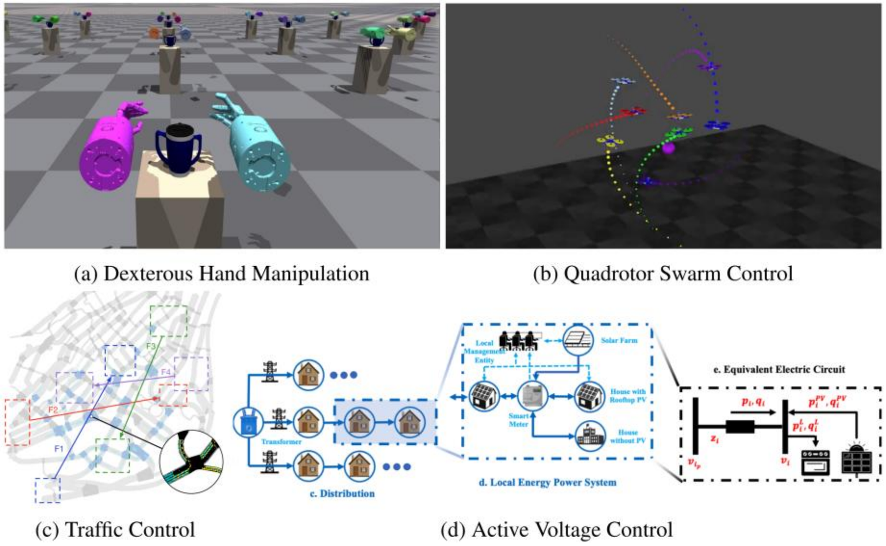
General hyperparameters	
Choices	Choice Range
Network Hidden Size	{64, **128**, 256}
Discount Factor (γ)	{0.9, 0.95, **0.99**}
Activation Function	{**ReLU**, Leaky_ReLU, SELU, Sigmoid, Tanh}
Initialization Method	{**Orthorgonal**, Xavier}
Neural Network Type	{**MLP**, RNN}
Learning Rate (LR)	{5e-5, **5e-4**, 5e-3}
Critic Learning Rate	{5e-5, **5e-4**, 5e-3}
Feature Normalization	{**True**, False}
Share Parameters	{True, **False**}
Early Stop	{True, **False**}
MADDPG Specific hyperparameters	
TD Steps (N Step)	{5, **20**, 25}
Exploratory noise	{0.001, 0.01, **0.1**, 0.5, 1}
MAPPO/HAPPO Specific hyperparameters	
Entropy Coefficient	{0.0001, 0.001, **0.01**, 0.1}
Use GAE	{**True**, False}
Use PopArt	{**True**, False}

Environment	Task Type	Control Mode	Episode Len.	Engine/Source	Data Source	Challenge
DexHand	Multi-robot Manipulation	Continuous	$\sim$80	Isaac Gym	Real-world Robots	Precise Control
Quads	Multi-robot Navigation	Continuous	$\sim$1600	OpenAI Gym	Real-world Robots	Long-range Task Assignment
Traffic	Network Control	Discrete	$\sim$1000	SUMO	Real-world Traffic	Long-range Control
Voltage	Network Control	Continuous	$\sim$200	IEEE Standard	Real-world Power Grid	Complex and Noisy Dynamics

Experiment Procedure

Procedure

- ***Step 1*** - train all algorithms and 18 tasks among the 4 environments, using default hyperparameters.
- ***Step 2*** - for each hyperparameter, vary one setting at a time to create a set of cooperative models.
- ***Step 3*** - Under 13 types of uncertainties, fix all cooperative models (1 cooperative baseline) and measure their robustness (13 evaluations) and resilience (13 evaluations).



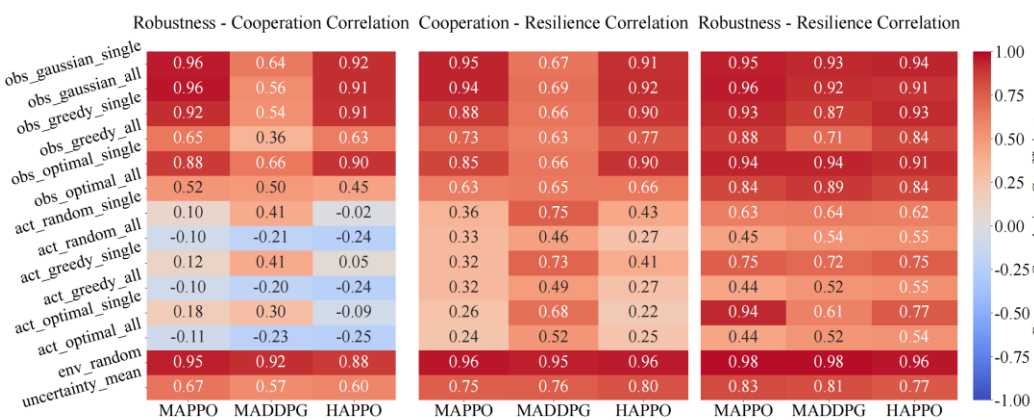
(a) Dexterous Hand Manipulation
(b) Quadrotor Swarm Control
(c) Traffic Control
(d) Active Voltage Control

5 (random seeds) × 27 (uncertainty settings) × 18 (tasks) × 34 (hyperparameters) = 82620 experiments.

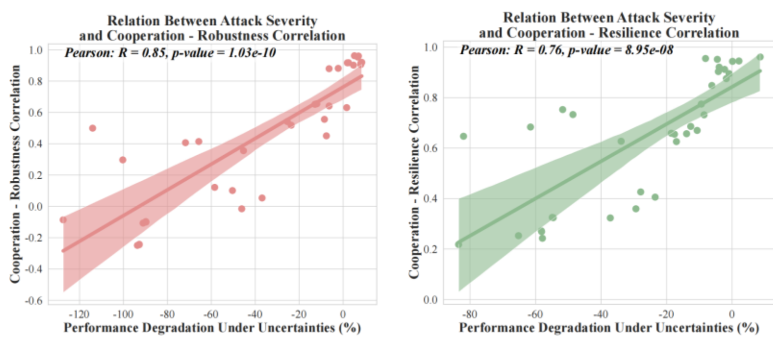
Experiment Results and Takeaways

1. Are Cooperation, Robustness, and Resilience Correlated?

- Cooperation improves robustness and resilience under mild uncertainty, but this correlation weakens as attack severity increases. The phenomenon holds across most uncertainty types, agent scopes, and attack strategies.
- MADDPG is preferable for action uncertainties, while MAPPO and HAPPO are better suited for observation uncertainties.



Correlation between cooperation, robustness, and resilience under uncertainty types and algorithms.

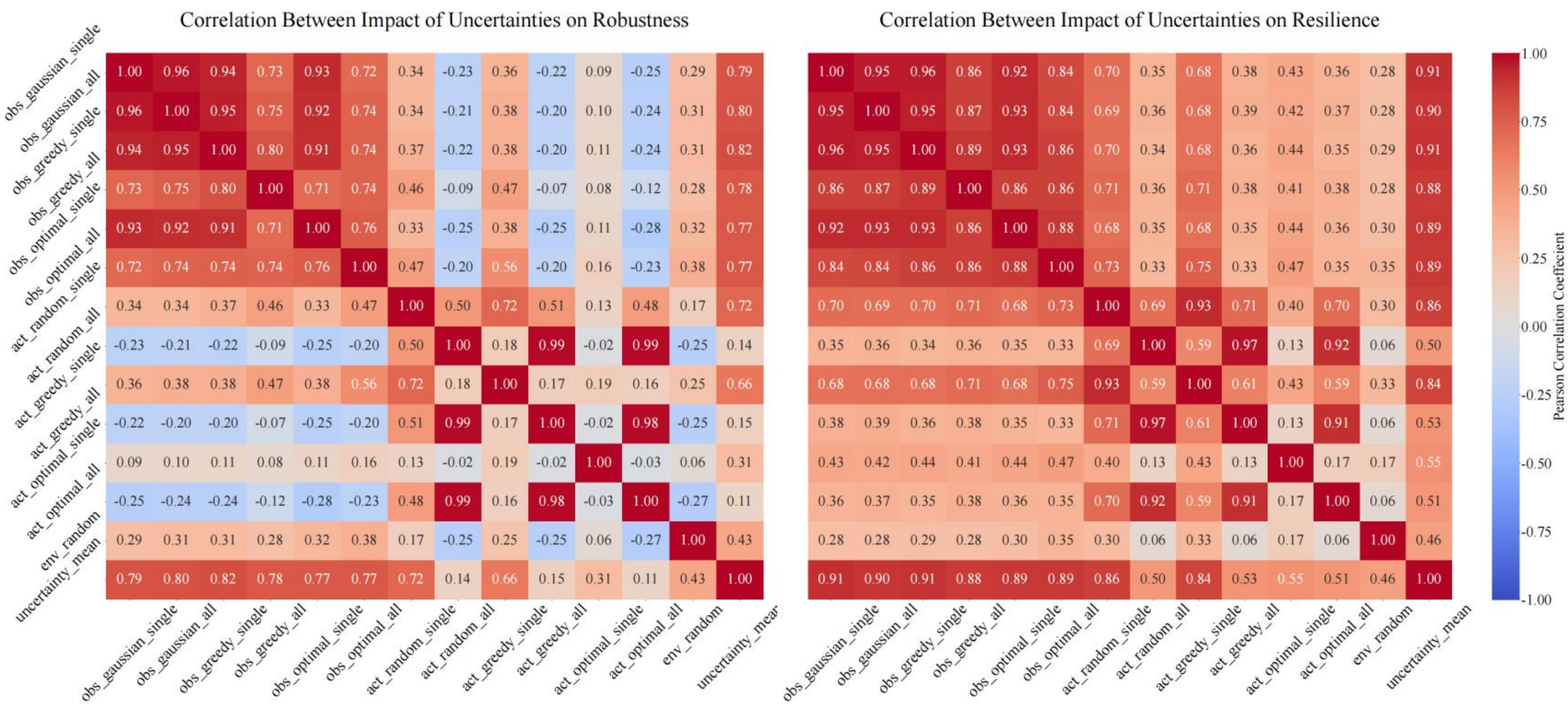


The relation between attack severity and cooperation – robustness/resilience correlation

Experiment Results and Takeaways

2. Uncertainty Diversity Matters

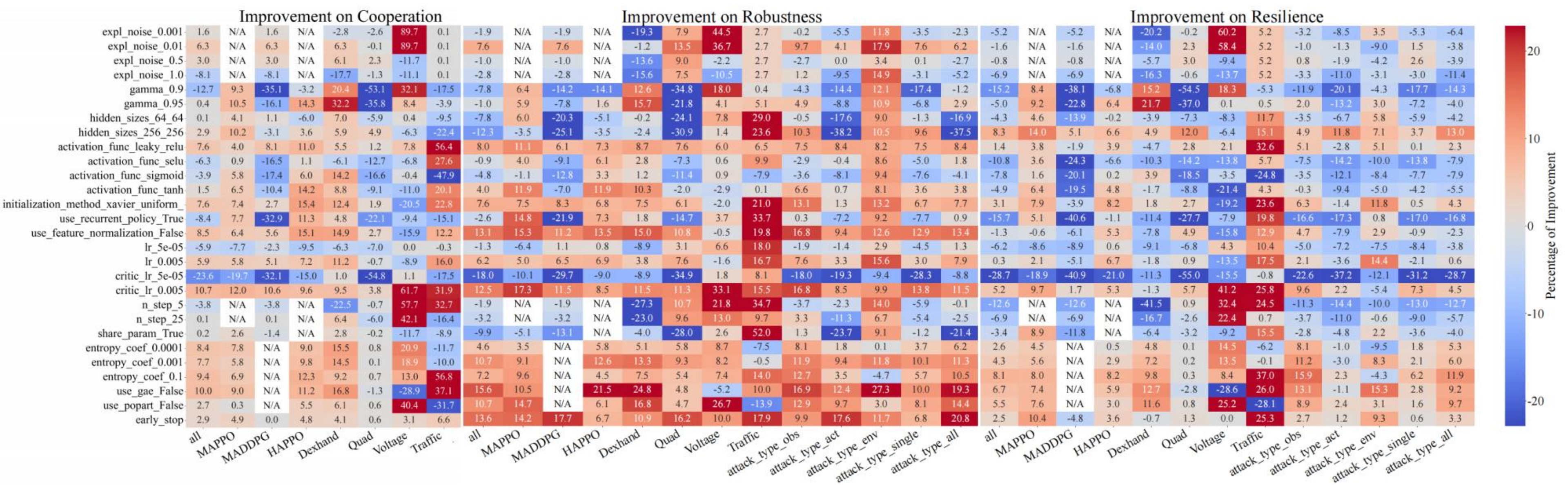
- Robustness and resilience in MARL can not generalize across uncertainty modalities (observations, actions, environments) or agent scopes (applied to individual or all agents).
- Trustworthy MARL systems must therefore evaluate against diverse types of uncertainty, and account for both individual and group-level perturbations.



Experiment Results and Takeaways

3. What hyperparameters are effective?

- Hyperparameter tuning plays a critical role in robustness and resilience.
- Surprisingly, common practices such as parameter sharing, GAE, and PopArt can hurt performance under uncertainty, while techniques like early stopping, critic-dominant learning rates, and Leaky ReLU consistently improve it.

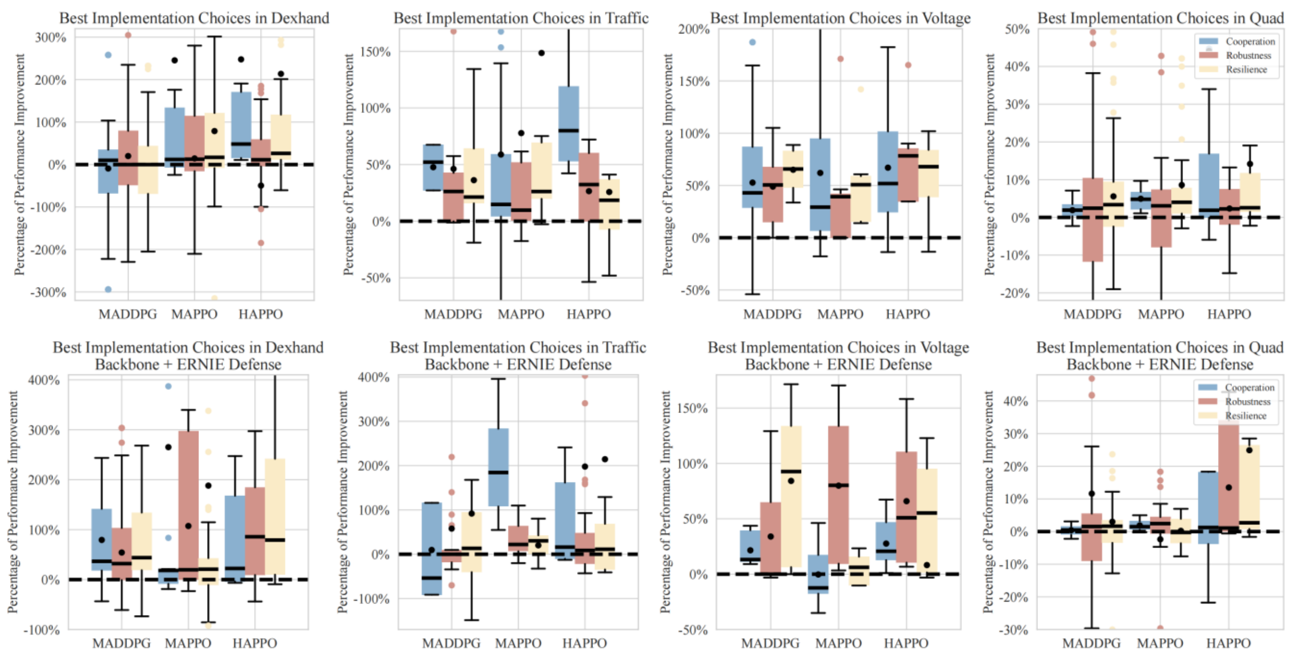


Percentage change in cooperation, robustness, and resilience caused by varying hyperparameters.

Experiment Results and Takeaways

4. Improving Robustness and Resilience

- Robustness and resilience can be significantly improved through hyperparameters alone.
- By combining best set of hyperparameters, we observe an average improvement of 52.60% in cooperation, 34.78% in robustness and 60.34% in resilience.
- This same set of hyperparameters generalizes to robust MARL methods on the same backbones, yielding average improvements of 89.43% on cooperation, 65.83% on robustness, 82.96% on resilience.



Thanks For Your Interest!