Unsupervised Multiple Proactive Behavior Learning of Mobile Robots for Smooth and Safe Navigation

Arthicha Srisuchinnawong[a], Jonas Bæch[b], Marek Piotr Hyzy[c], Evangelos Boukas[c], Tsampikos Kounalakis[b], & Poramate Manoonpong[a]

[a] University of Southern Denmark, Odense, Denmark, & Vidyasirimedhi Institute of Science and Technology, Rayong, Thailand,
[b] Danish Technological Institute, Odense, Denmark, and [c] Technical University of Denmark, Lyngby, Denmark.

Motivation & Objective

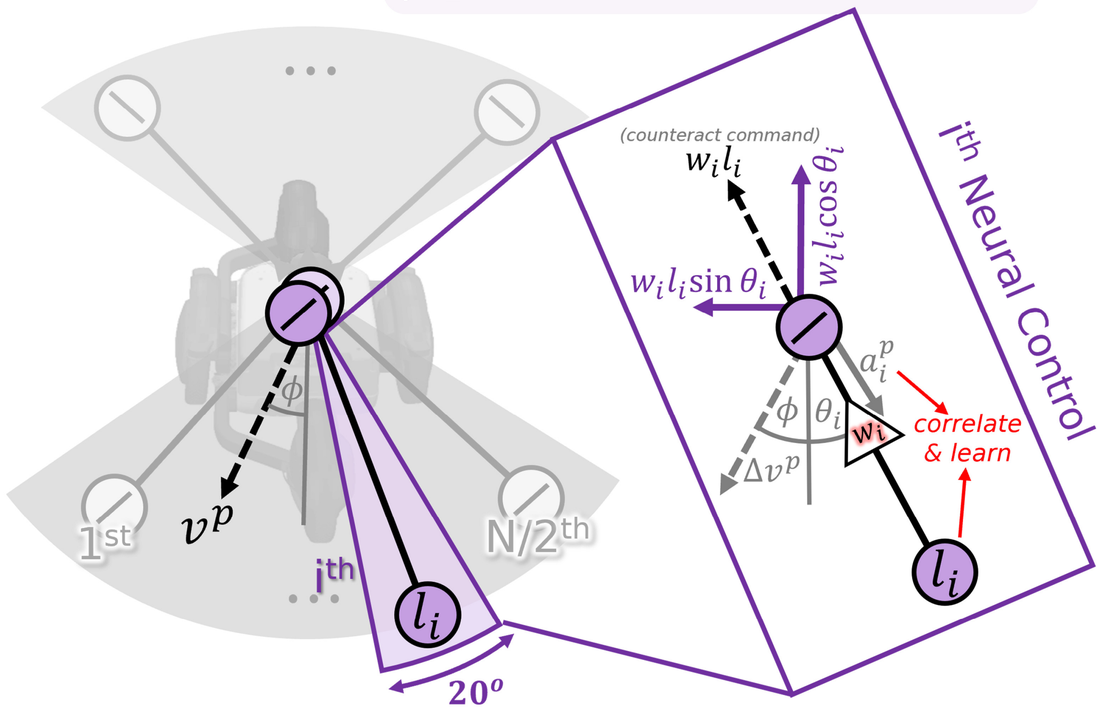
Most motion planners[1] require:

- **high update rates** but have computational constraints.
- **environment model** *(model-based prediction).*
- long **training times** and **reward** functions *(reinforcement learning).*
- **targets** or human demonstrations *(supervised learning).*

Multiple Proactive Behavior Learning (MPL)

- **Plug-and-play** MPL module.
- Combines **any model-based planning with model-free MPL**.

$$\begin{bmatrix} v \\ \omega \end{bmatrix} = \underbrace{\begin{bmatrix} v^p \\ \omega^p \end{bmatrix}}_{\text{planner}} + \underbrace{\sum_{i=1}^{N} \left(\begin{bmatrix} -w_i[t] & l_i[t] \end{bmatrix} \begin{bmatrix} \cos\theta_i[t] \\ \sin\theta_i[t] \end{bmatrix} \right)}_{\text{MPL command}}$$



- Proactive control with **unsupervised correlation-based learning**[2].
- Uses the obstacle detection feedback $(l_i[t], \theta_i[t])$ and the planner acceleration $(a_i^p = f(v^p, \omega^p))$.
- Operates **without additional targets & reward** functions.

$$w_i[t+1] = ReLU\left(\alpha w_i[t] - \eta\, l_i[t]\, \Delta a_i^p[t]\right)$$

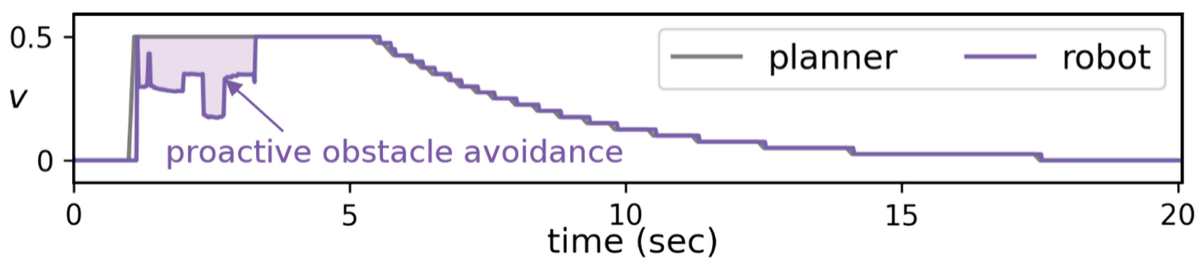
Model-based Motion Planning (low update rate)

Model-free Proactive Control with Unsupervised Learning (high update rate)

Video | GitHub | Project

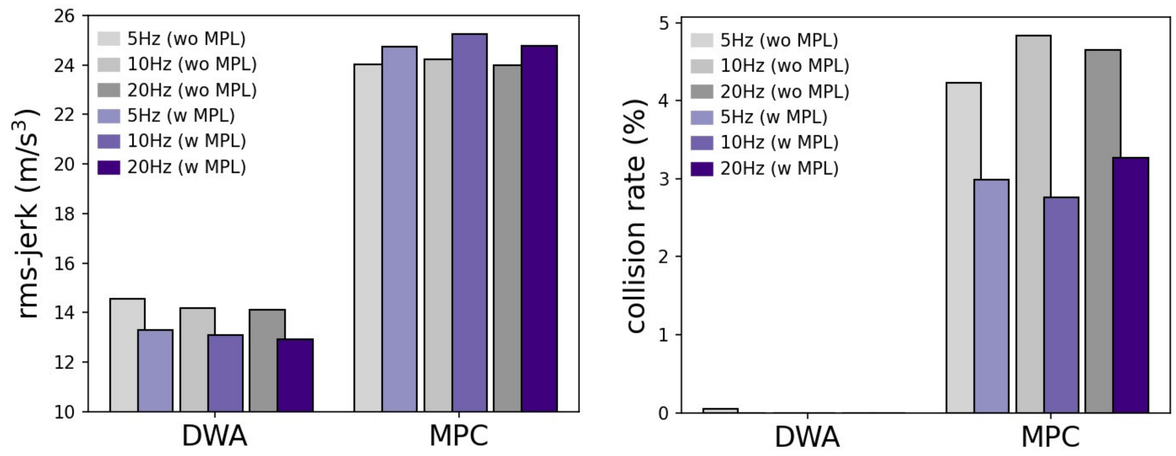
Experimental Results

- Produces **proactive commands**.
- Modifies the final commands *(the purple highlighted area).*



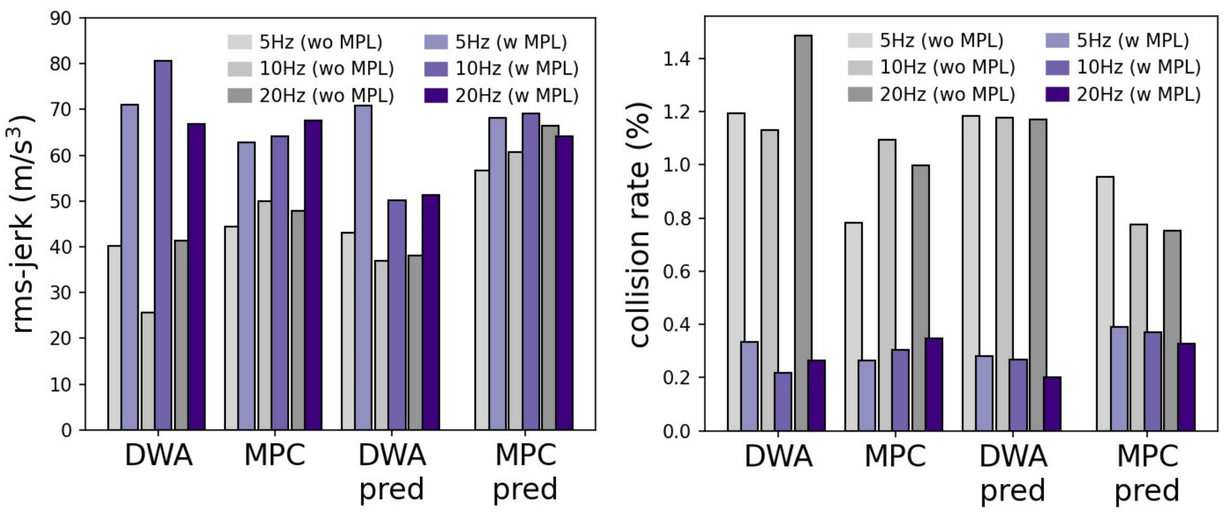
In a narrow corridor:

- The convergence time ≈ **1 minute**.
- **Improves smoothness by 10%** *(Dynamic Window Approach: DWA).*
- **Reduces collisions by 30%** *(Model Predictive Control: MPC).*



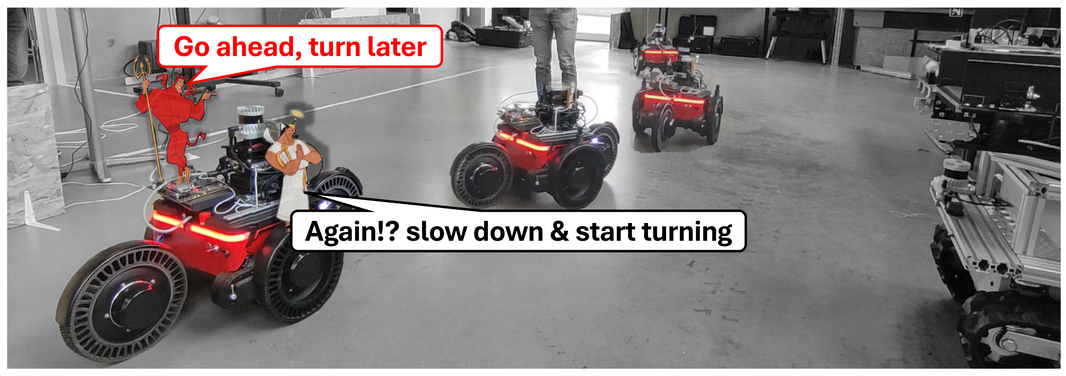
In a highly dynamic environment:

- The convergence time ≈ **2 minutes**.
- **Trades smoothness** for a **70% reduction in collisions**.



In a physical system:

- **Enhances robustness** and success rate



Conclusion

- Autonomously **enhances & balances multiple proactive strategies**.
- **No world model, long training times, reward functions, and targets**.
- **Applies unsupervised learning** to train a model-free neural control such that it cooperates with existing model-based control.

[1] Lu et al. "*A review of mobile robot motion planning methods: from classical motion planning workflows to reinforcement learning-based architectures.*" (2023).
[2] Porr et al. "*Strongly improved stability and faster convergence of temporal sequence learning by using input correlations only.*" (2006).

We also thank ***Odense Robotics***, ***Capra Robotics***, ***NCC***, ***Hecto Drone***, ***Buildcode***, ***Desupervised***, ***Lorenz Technology***, ***Meili Robots***, & ***Robstruct*** for their supports and discussions.