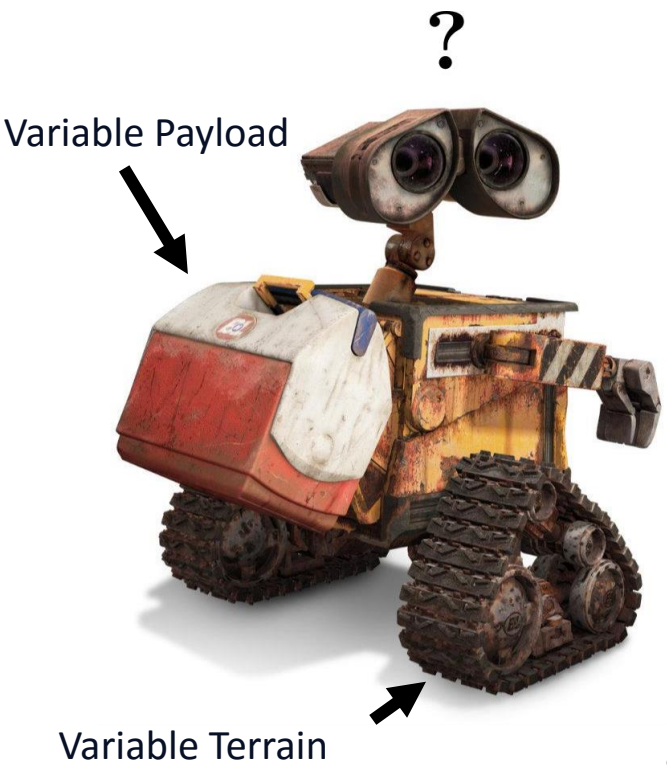
Experience-based Model Selection to Enable **Long-term, Safe** Control for Repetitive Tasks Under **Changing Conditions**

Chris McKinnon and Angela P. Schoellig

IROS 2018

Institute for Aerospace Studies
UNIVERSITY OF TORONTO

Our Goal is to Control a System with Multi-modal Dynamics



Dynamic Model:

$$\dot{\mathbf{x}} = \underbrace{\mathbf{f}(\mathbf{x}, \mathbf{u})}_{\text{Unicycle Model}} + \underbrace{\mathbf{g}^{\mathbf{c}}(\mathbf{a})}_{\text{Gaussian Process (GP)}}$$

Mode

Feature

Control Input:

$$\mathbf{u} = \mathbf{argmax}\ \mathbf{Reward}(\mathbf{x}, \mathbf{u})$$

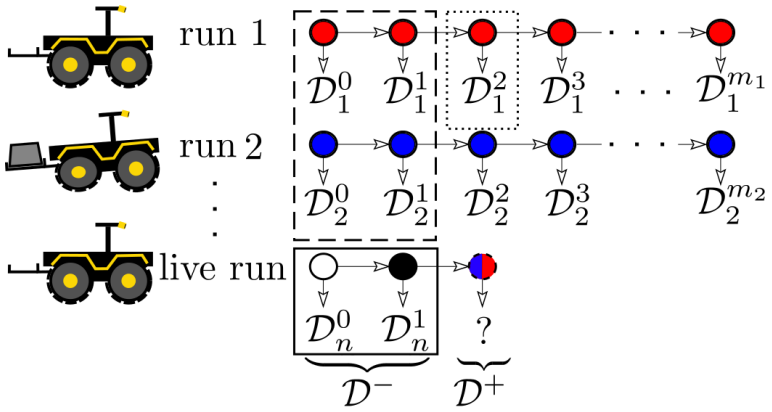
Constraints:

$\mathbf{x} \in \mathcal{X}$ Maximum Tracking Error

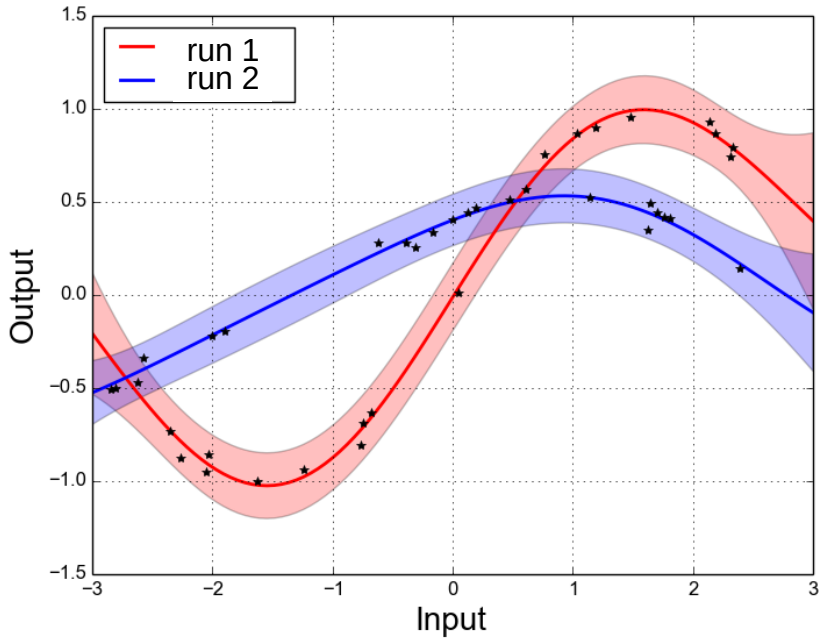
$\mathbf{u} \in \mathcal{U}$ Actuator Constraints

- We use a **Robust MPC** to be robust to model errors.
- **The goal of this work is to construct an accurate model for $\mathbf{g}^{\mathbf{c}}(\mathbf{a})$ with realistic bounds on model error.**

The Robot is Performing a Repetitive Path Following Task



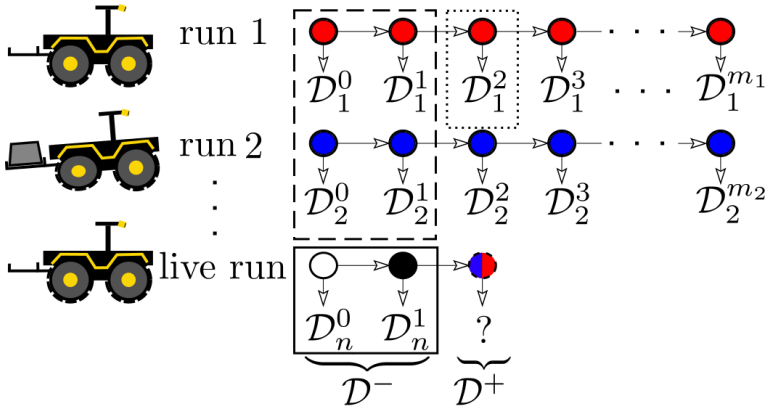
Step 1: Fit a local GP to each run in memory

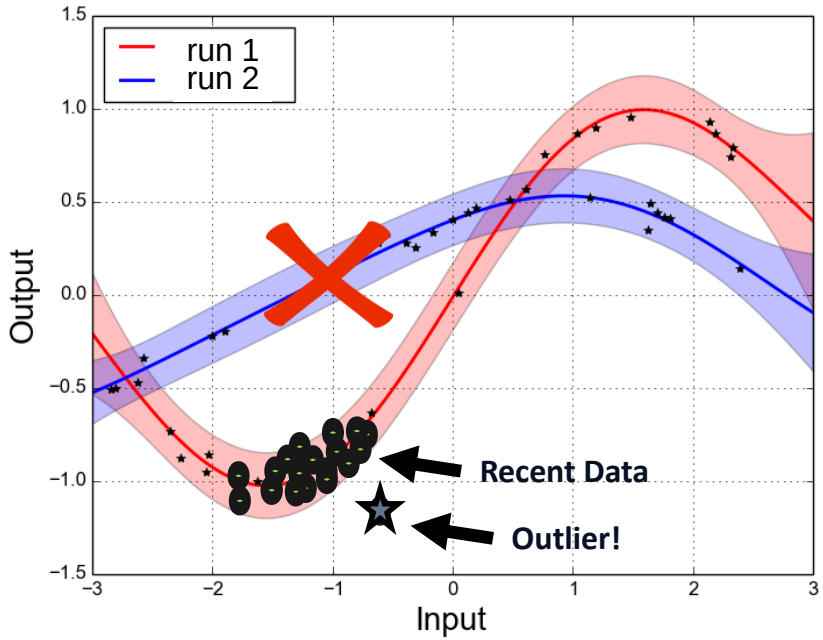


- Each run, the robot stores a new set of experiences.
- Our goal is to choose useful experiences from past runs to construct a local GP for control that **models the robot dynamics over the upcoming section of the path**.

Step 2: Check if we have experience that is safe to use

The robot may encounter **novel operating conditions** where it is **dangerous to use past experience!**





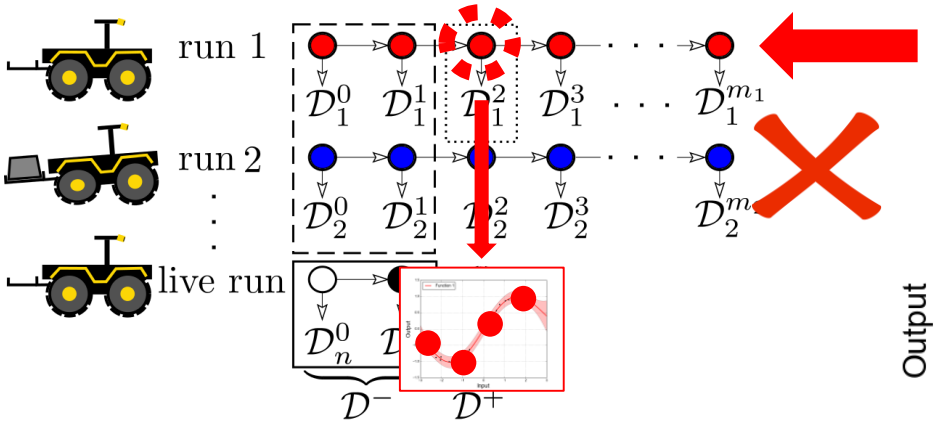
Compute the probability that outliers occurred by chance

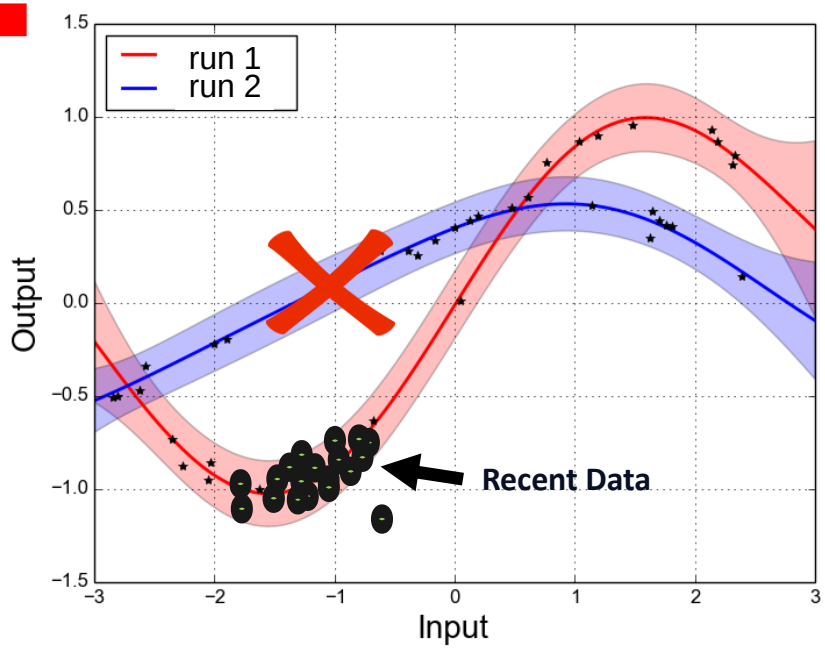
$$p(N_{\text{out}} \text{ or more}) = \sum_{x=N_{\text{out}}}^{m_n} B(x, m_n, p)$$

Reject if:

$$p(N_{\text{out}} \text{ or more}) < \alpha$$

Step 3: Construct a GP using data from the most similar runs





The most similar run is the run with the local GP that best explains recent data

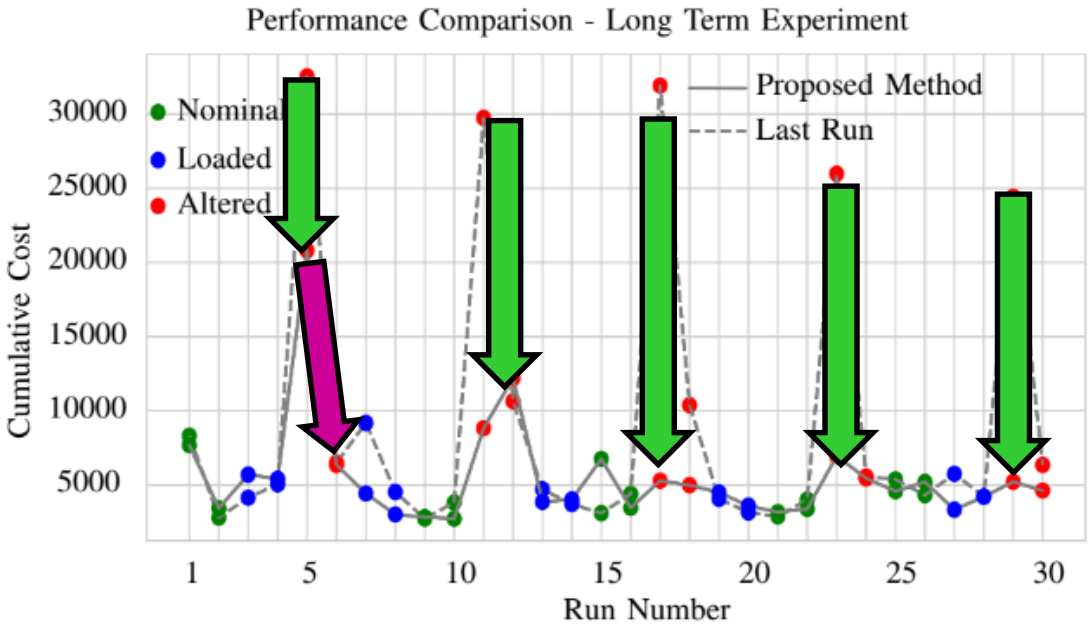
$$p(c = i|\mathcal{D}_n^-, \hat{g}_i^-) \propto p(\mathcal{D}_n^-|\hat{g}_i^-)p(c = i)$$

Likelihood of recent data given local GP from run *i*

Prior for run *i*

Experimental Results: Long-term autonomy in changing conditions

Experiments were conducted using a **900 kg ground robot** over **2.5 km of driving**.



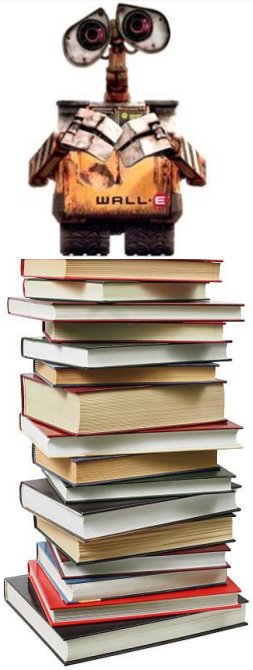
Modes

Nominal

- The proposed method i**mproves significantly after just one run** in each mode
- The proposed method can search **up to 300 previous runs** for relevant experience
- This enables **truly long-term safe learning**.

Thanks!

DYNAMIC SYSTEMS LAB



Loaded

Nominal

Station # 1