Learning Model Predictive Control for Iterative Tasks

A safe data-driven control framework

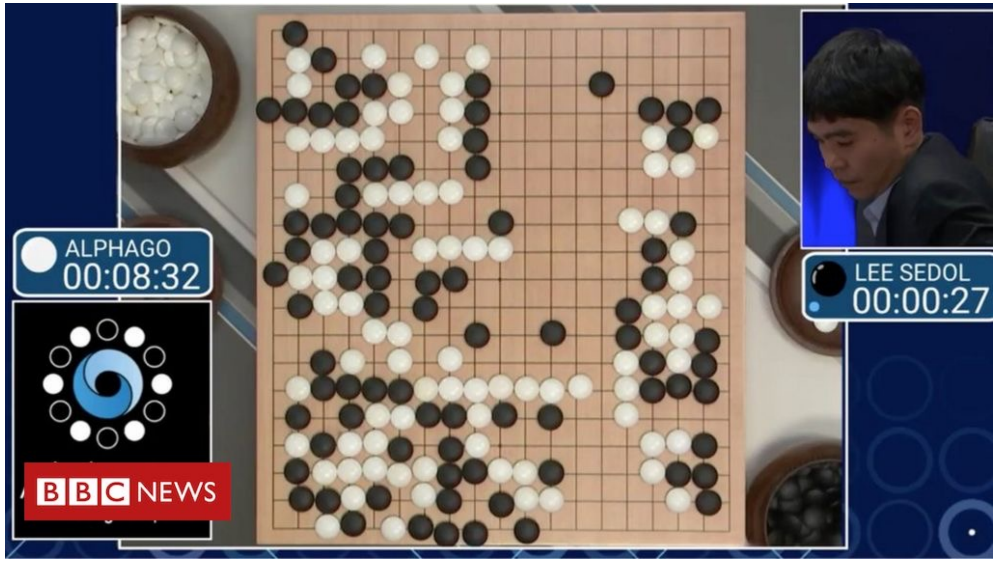
Ugo Rosolia

Research Scientist @Amazon

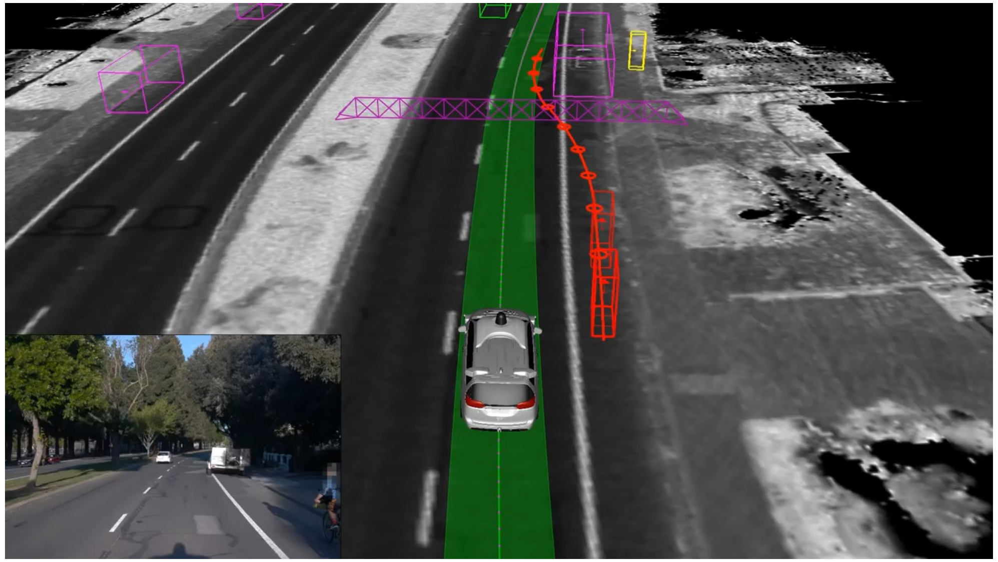
December, 2021

Success Stories from AI

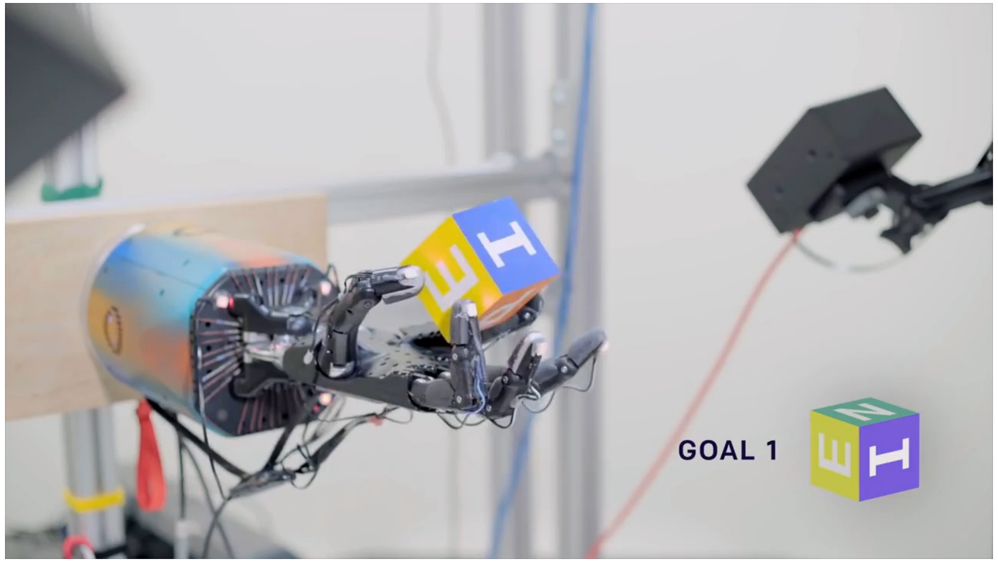
Alpha GO



Waymo's Perception Module



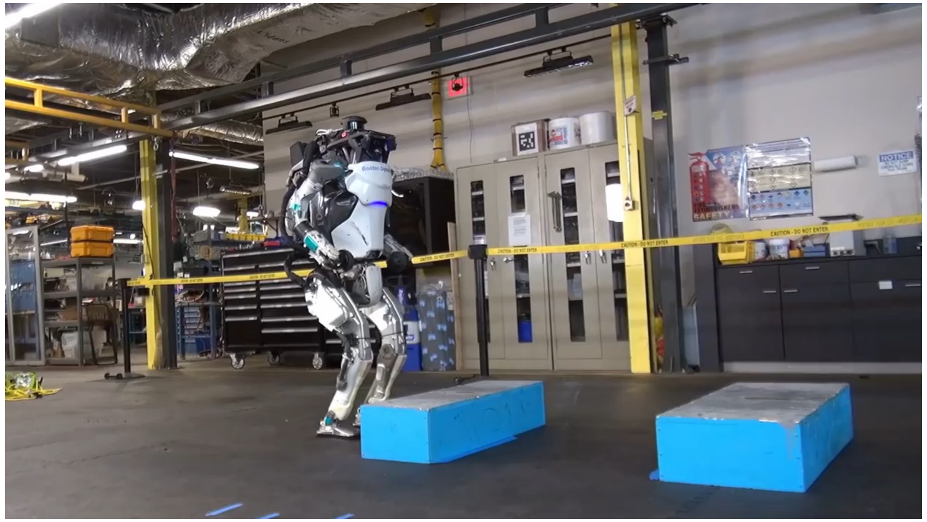
OpenAI



Google

Success Stories from Control Theory

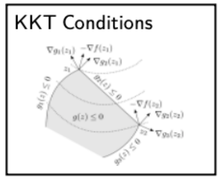
Boston Dynamics



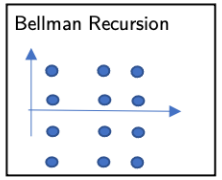
Stanford Dynamic Design Lab

Standard Control Pipeline

Optimal Trajectory

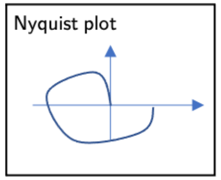


Optimization

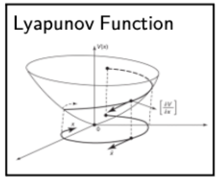


Dynamic Programming

Trajectory Tracking

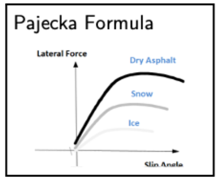


Frequency Domain

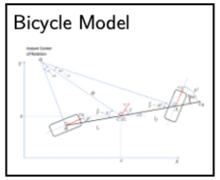


Nonlinear Control

System Identification

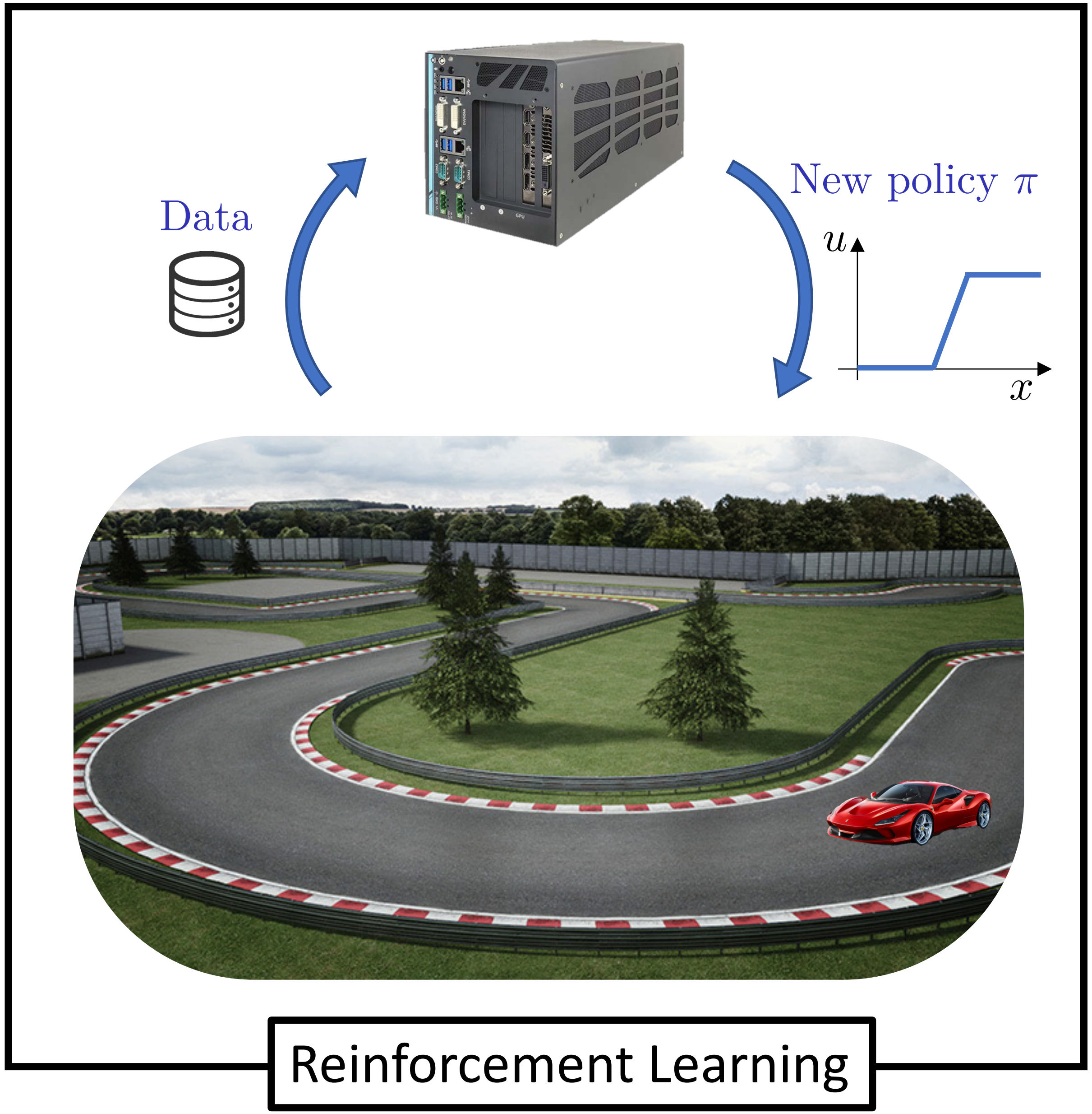


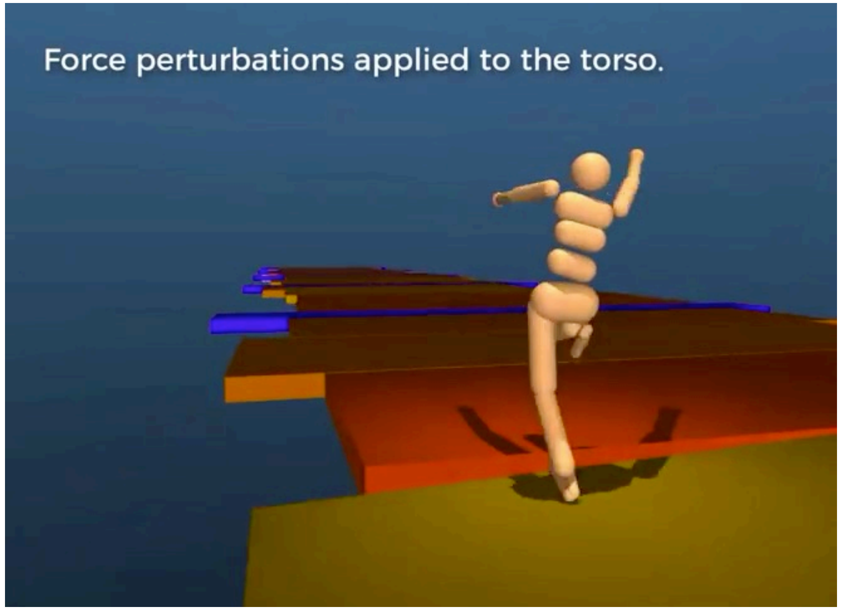
Tire Dynamics



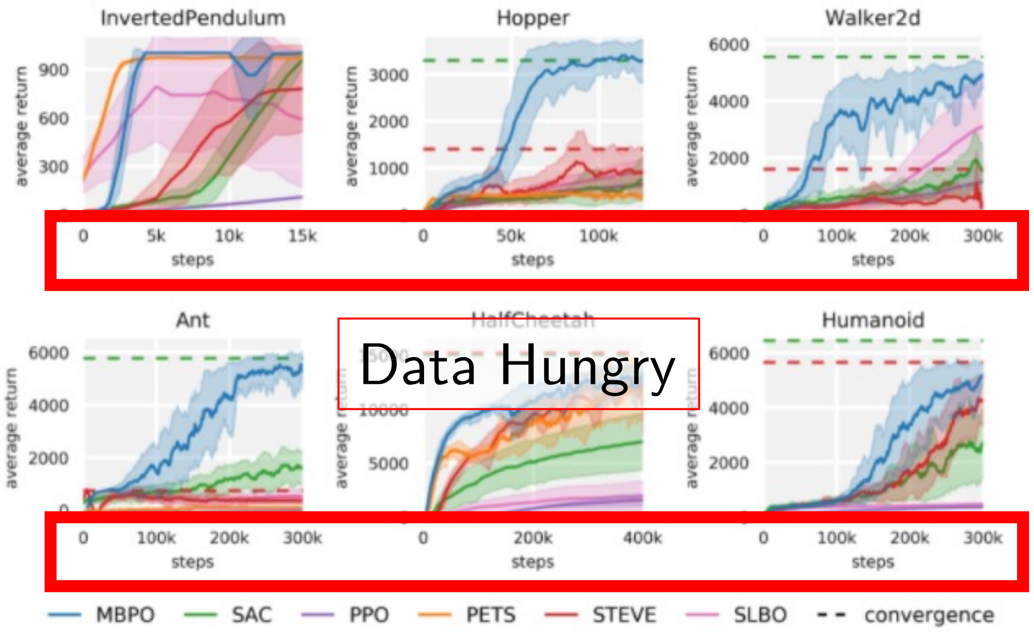
Vehicle Dynamics

Can we simplify the control design?



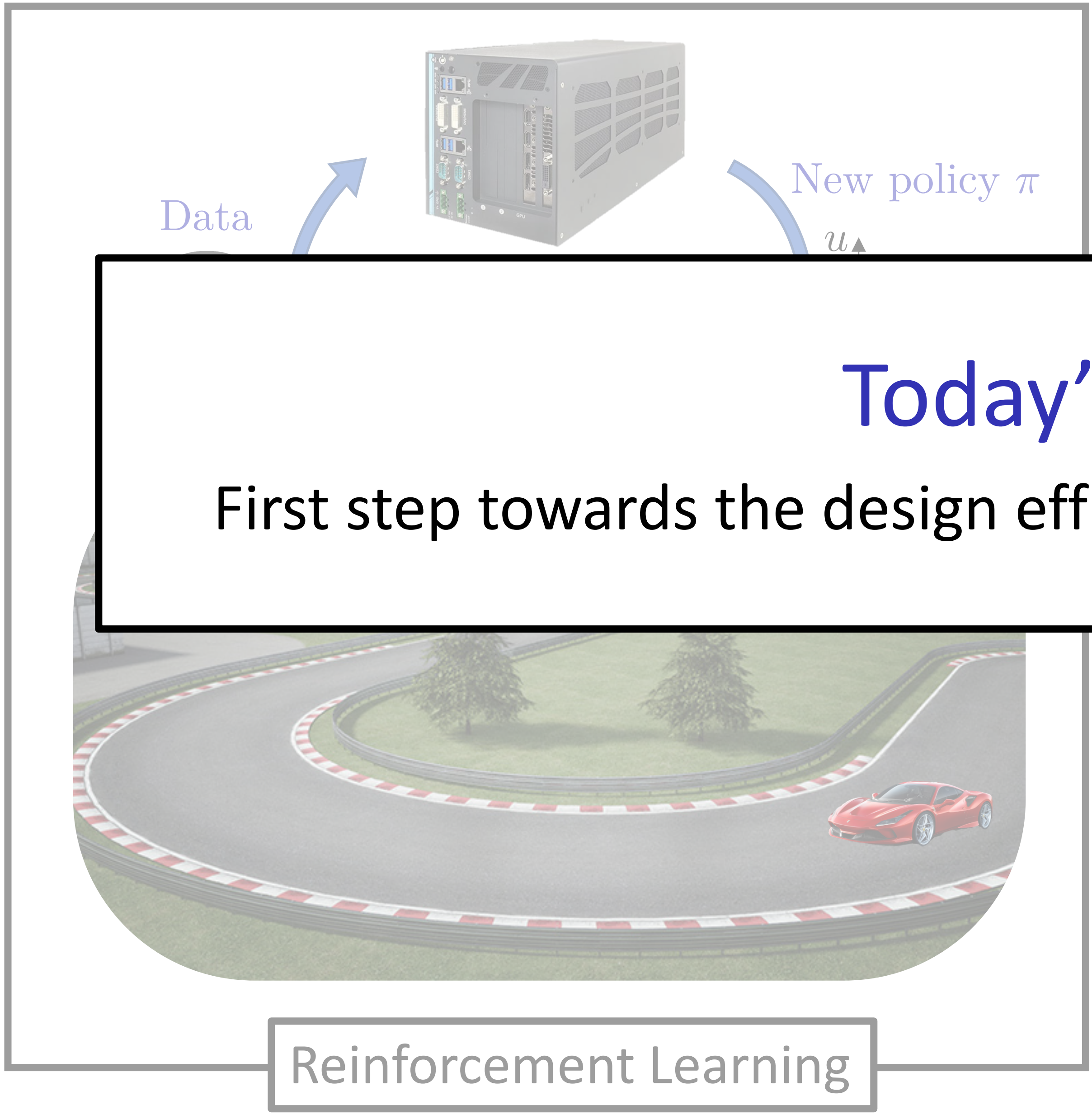


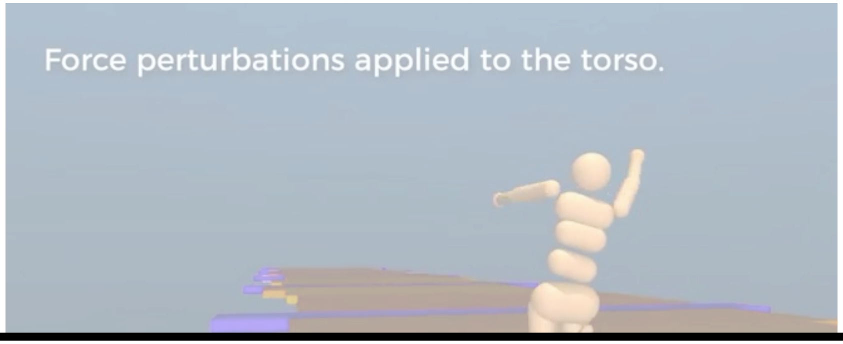
DeepMind



M. Janner, J. Fu, M. Zhang, and S. Levine. "When to trust your model: Model-based policy optimization." arXiv preprint arXiv:1906.08253 (2019).

Can we simplify the control design?

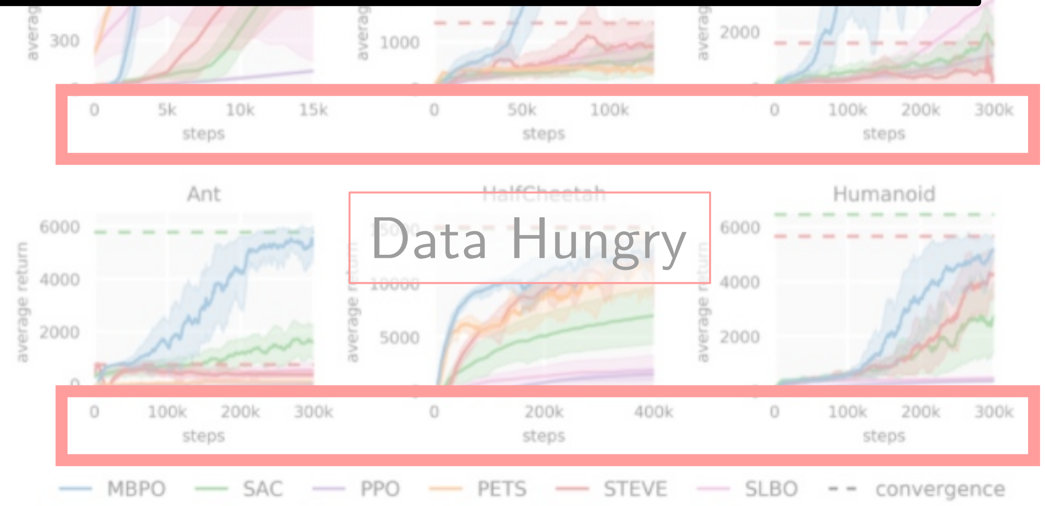




DeepMind

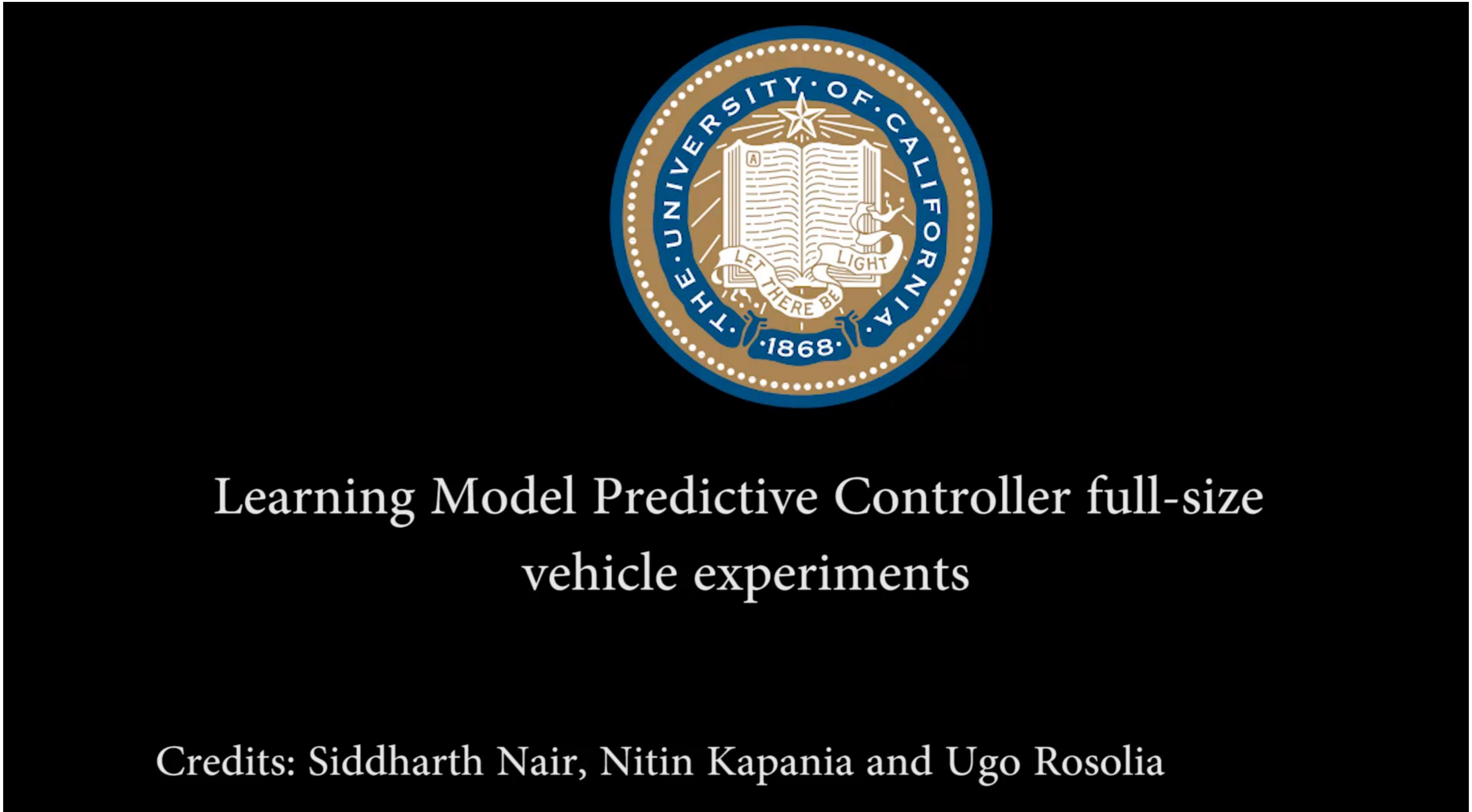
Today's goals:

First step towards the design efficient model-based RL framework

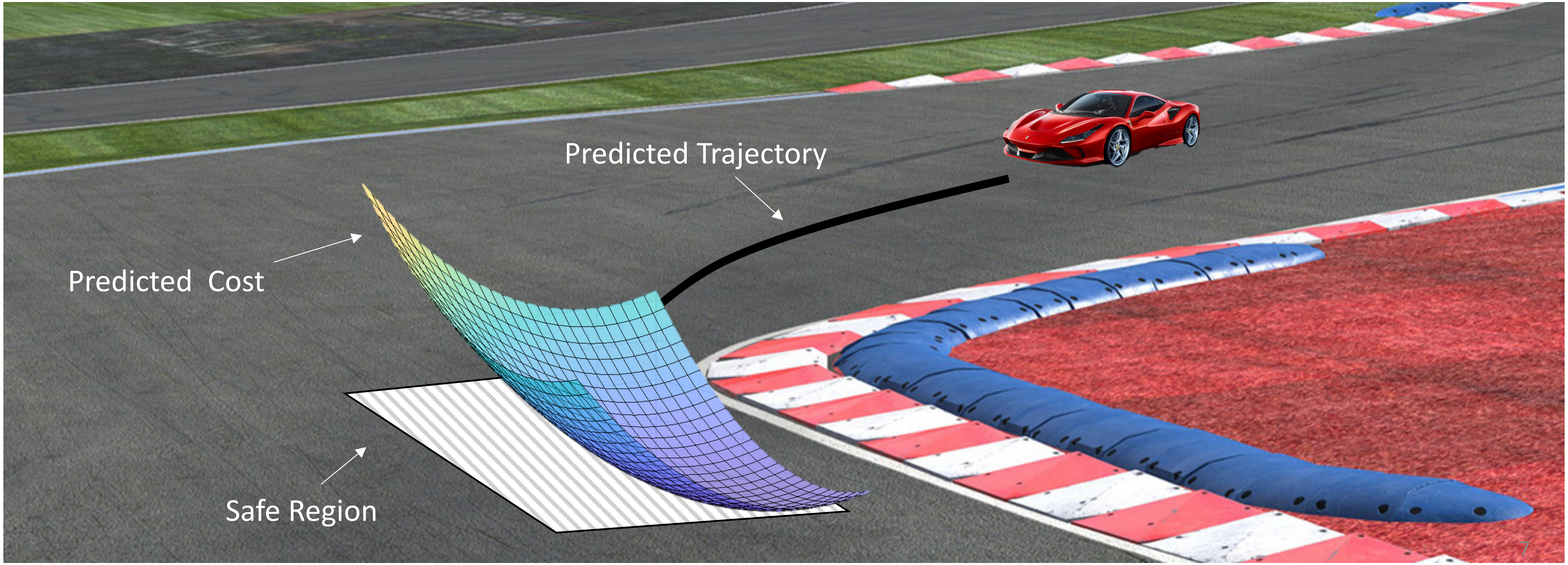


M. Janner, J. Fu, M. Zhang, and S. Levine. "When to trust your model: Model-based policy optimization." arXiv preprint arXiv:1906.08253 (2019).

Today's Example



How to compute control actions?



- Predicted trajectory given by Prediction Model
- Safe region estimated by the Safe Set
- Predicted cost estimated by Value Function

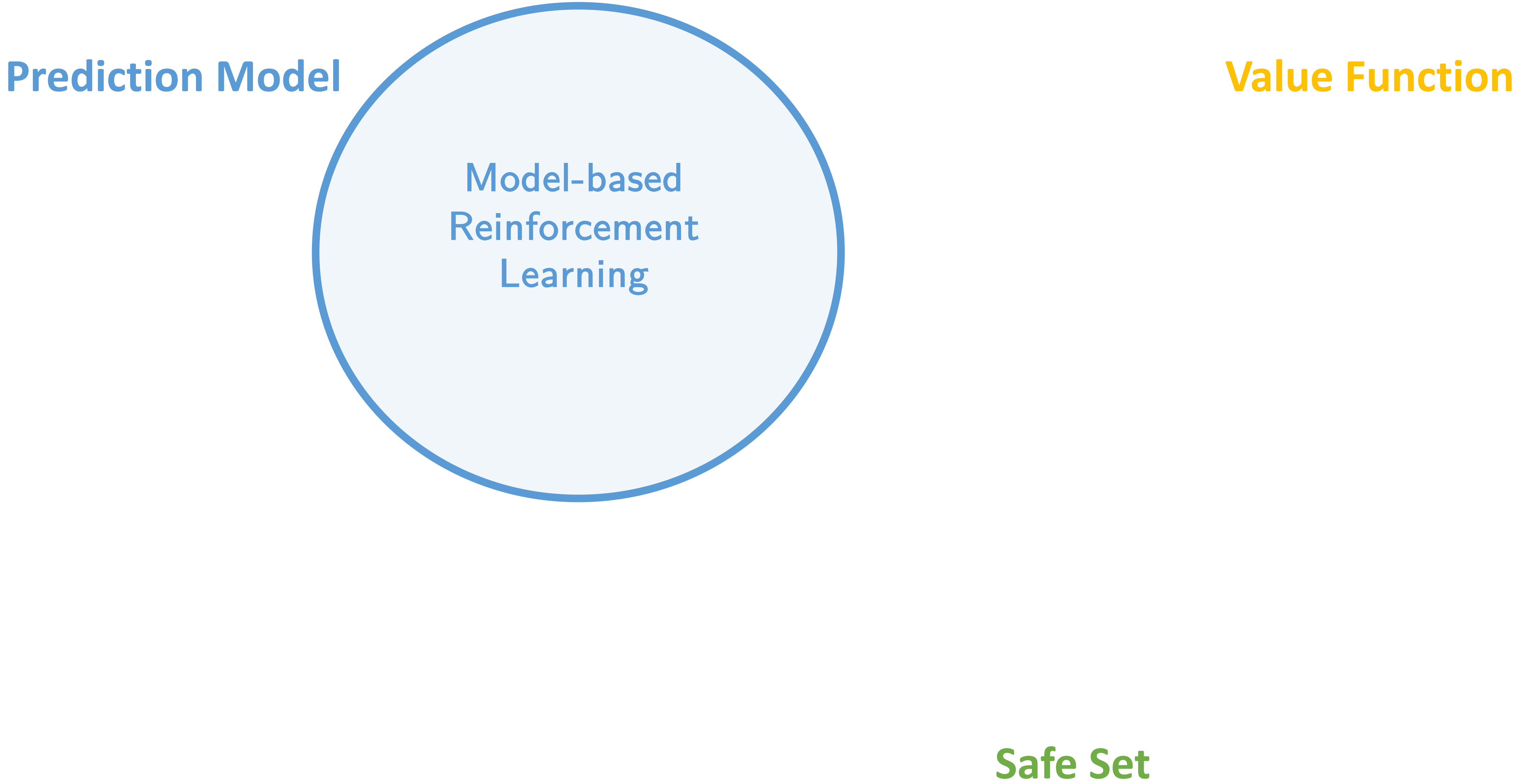
Three key components to learn

Prediction Model

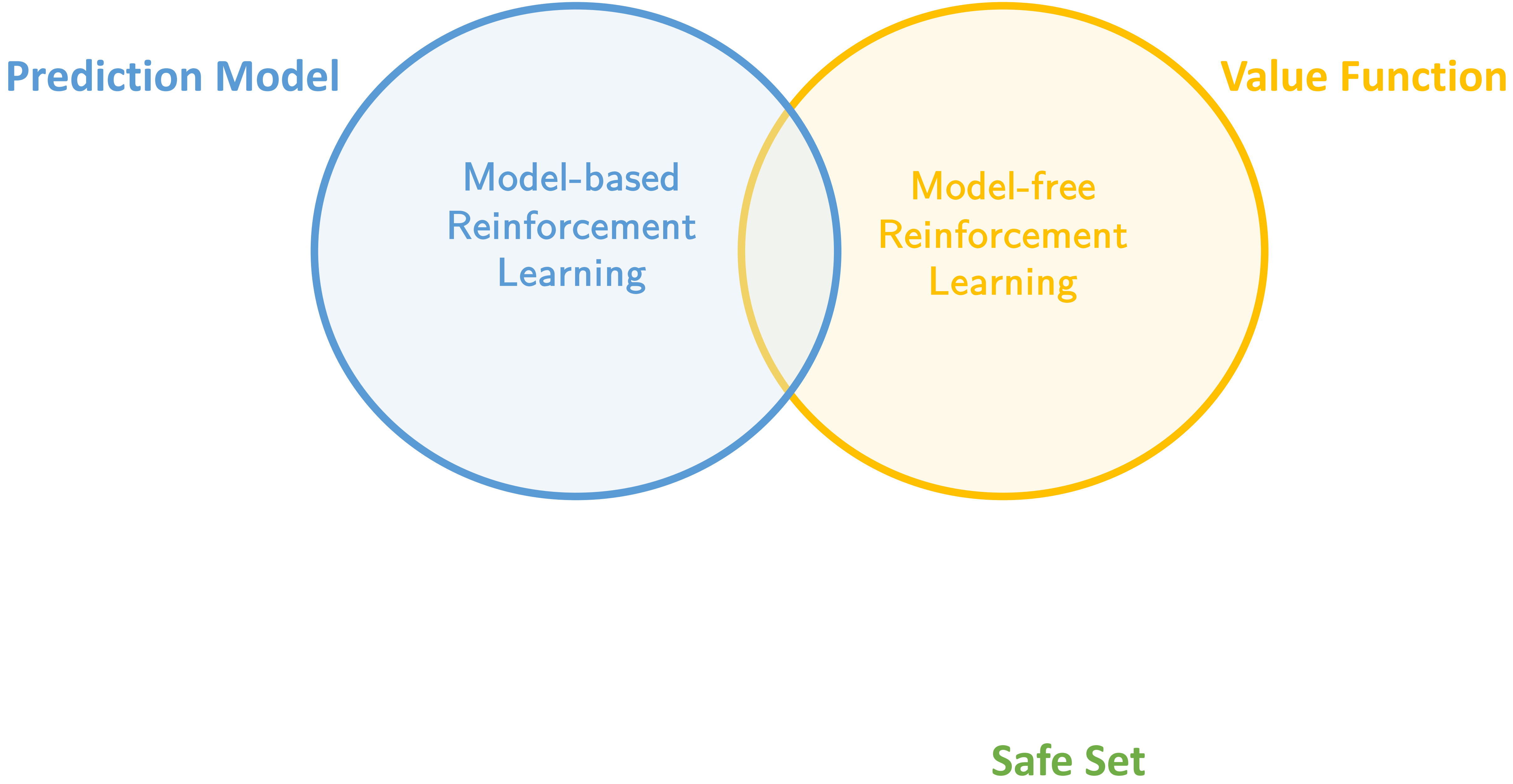
Value Function

Safe Set

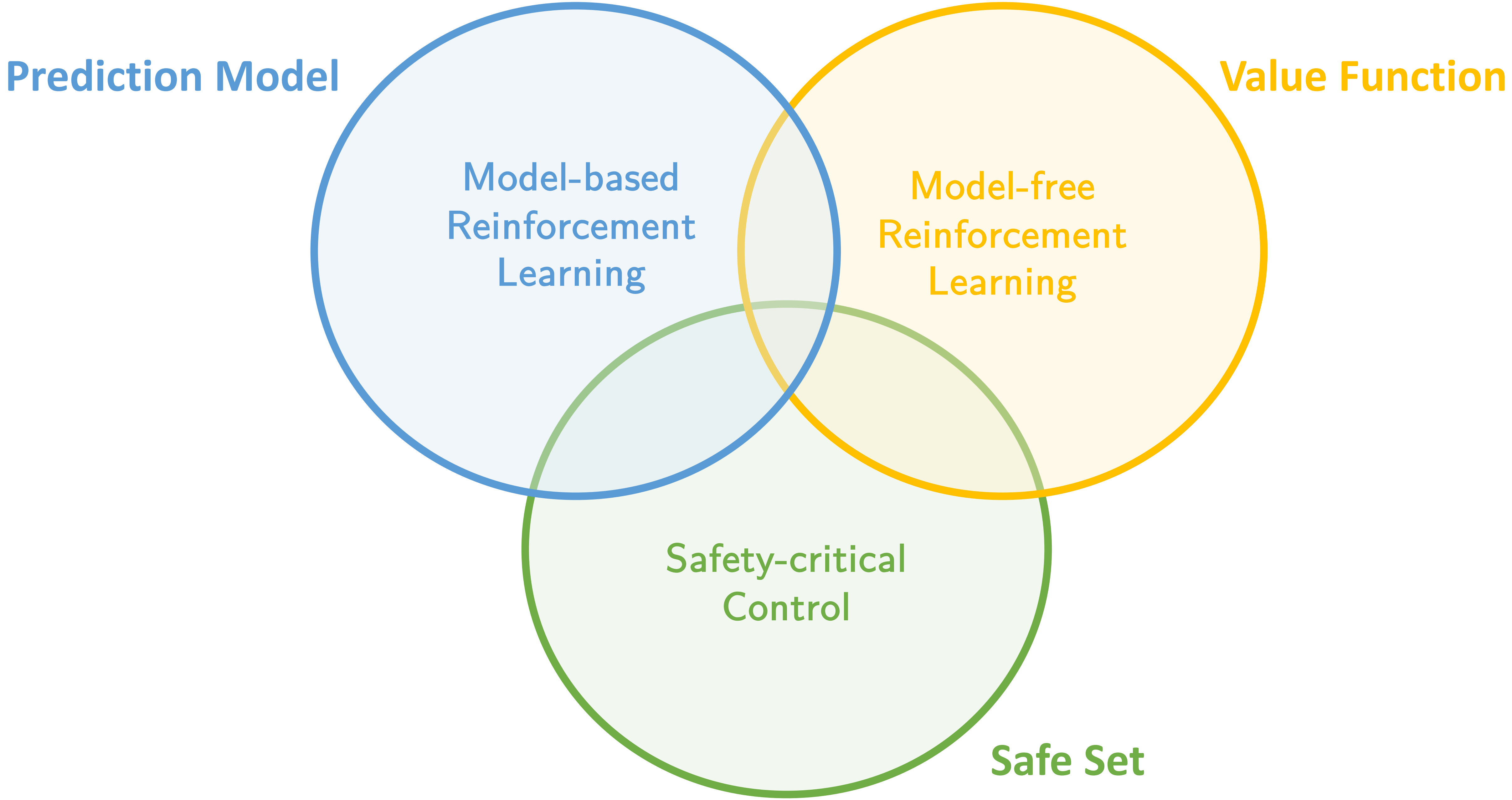
Three key components to learn



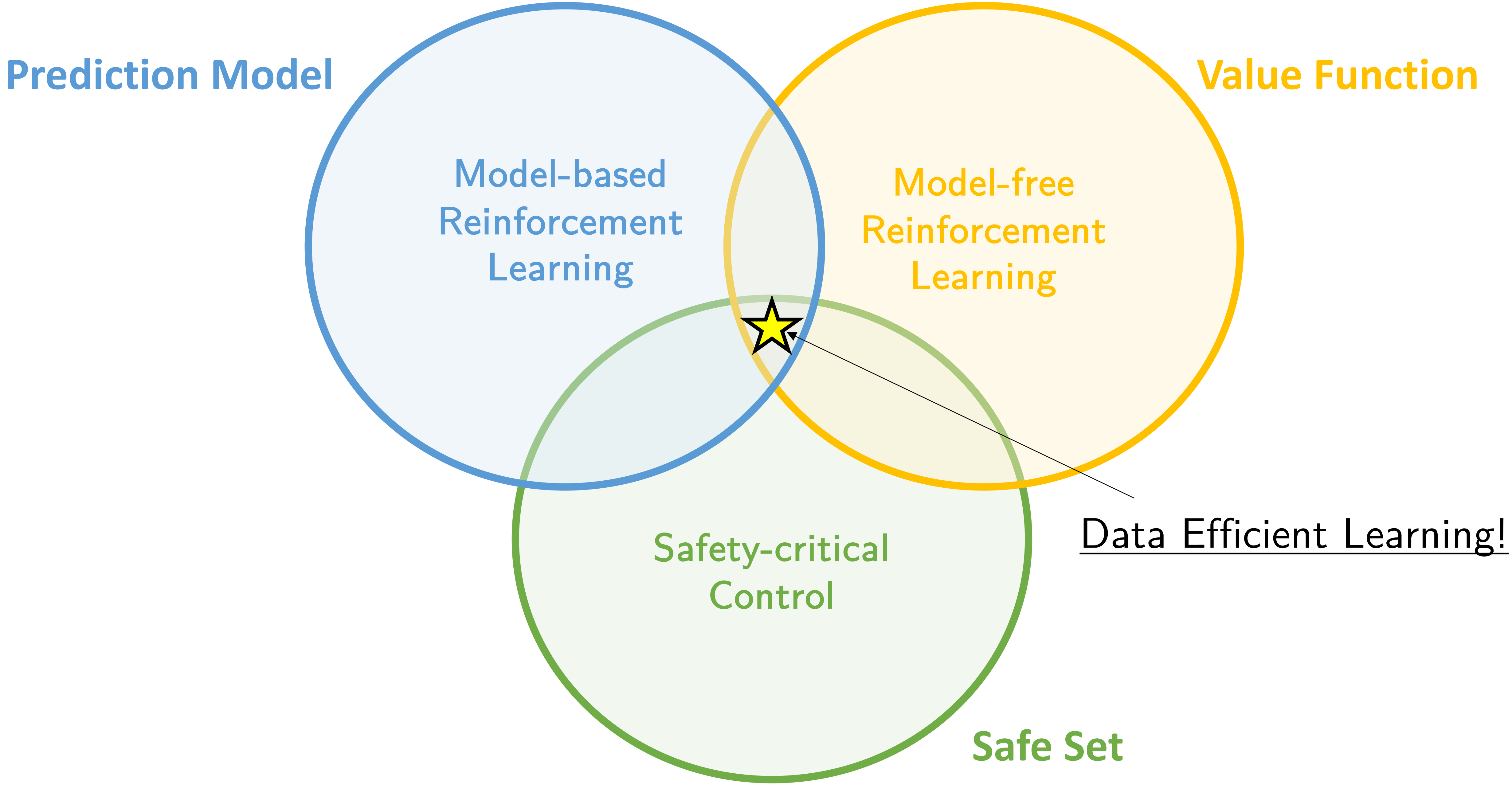
Three key components to learn



Three key components to learn



Three key components to learn



Problem Formulation

Minimum Time Control Problem

$$\min_{T,\mathbf{u}} \quad \boxed{T} \quad \text{Control objective}$$

$$\boxed{x_0 = x_s, \;\; x_T = \mathcal{X}_F} \quad \text{Start \& end position}$$

System dynamics
System constraints

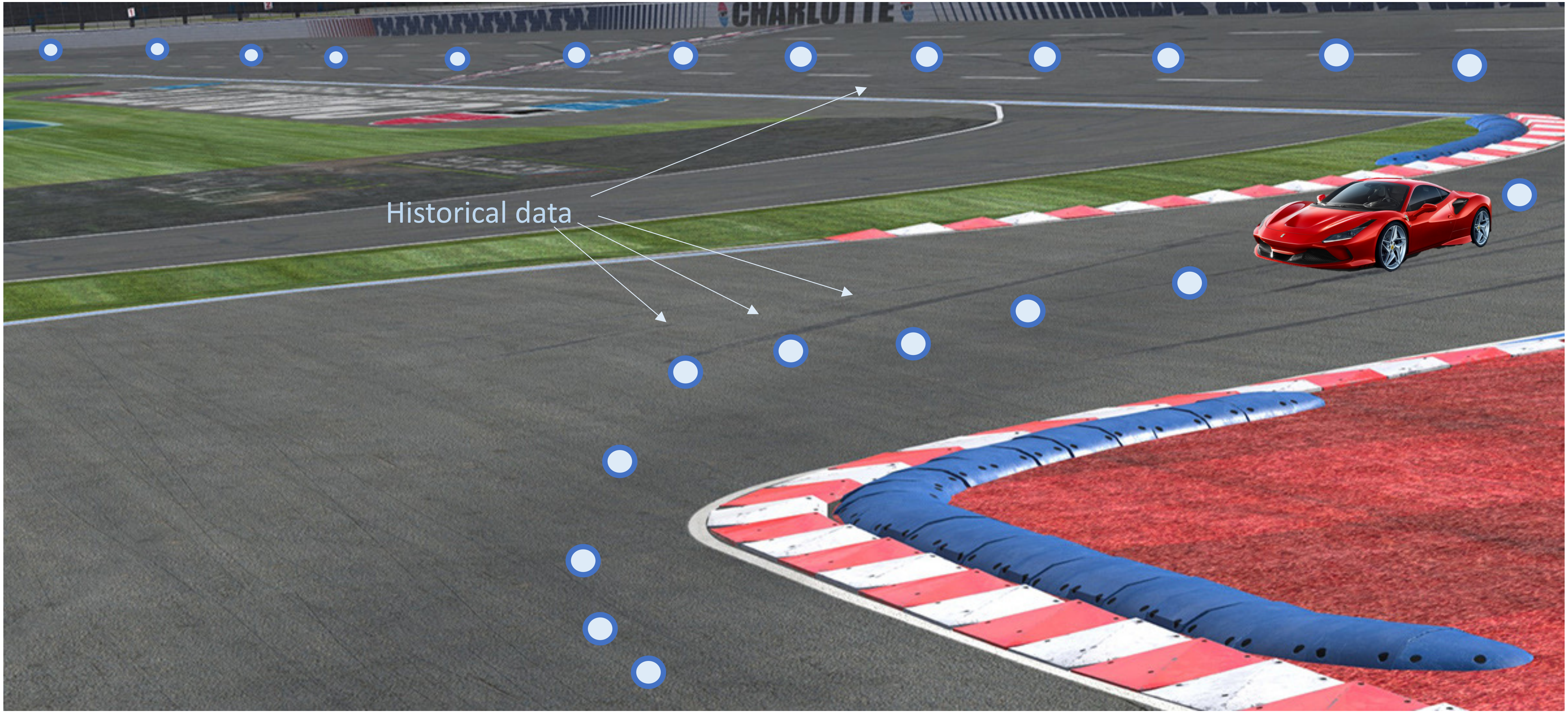
$$\boxed{x_{k+1} = f(x_k, u_k), \quad \forall k \in \{0, \ldots, T-1\}}$$

Safety constraints

$$\boxed{x_k \in \mathcal{X}, \;\; u_k \in \mathcal{U}, \quad \forall k \in \{0, \ldots, T-1\}}$$

Key Assumption

We are given a first feasible trajectory and/or controller



Learning Model Predictive Controller

At time t of iteration j solve the following Constrained Finite Time Optimal Control Problem (CFTOCP)

$$J^{\text{LMPC},j}_{t\rightarrow t+N}(x^j_t) = \min_{u^j_{t|t},\dots,u^j_{t+N-1|t}} \sum_{k=t}^{t+N-1} h(x^j_{k|t}, u^j_{k|t}) + V^{j-1}(x^j_{t+N|t}, {\color{red}x})$$

s.t.

$$x^j_{k+1|t} = A^j_{k|t} x^j_{k|t} + B^j_{k|t} u^j_{k|t} + C^j_{k|t}$$

$$x^j_{t|t} = x^j_t,$$

$$x^j_{k|t} \in \mathcal{X},\; u^j_{k|t} \in \mathcal{U},\; \forall k \in [t, \cdots, t+N-1]$$

$$x^j_{t+N|t} \in \mathcal{CS}^{j-1}({\color{red}x}),$$

Learning Model Predictive Controller

At time t of iteration j solve the following Constrained Finite Time Optimal Control Problem (CFTOCP)

$$J^{\mathrm{LMPC},j}_{t\to t+N}(x^j_t) = \min_{u^j_{t|t},\ldots,u^j_{t+N-1|t}} \sum_{k=t}^{t+N-1} h(x^j_{k|t}, u^j_{k|t}) + V^{j-1}(x^j_{t+N|t}, x)$$

s.t.

$$x^j_{k+1|t} = A^j_{k|t} x^j_{k|t} + B^j_{k|t} u^j_{k|t} + C^j_{k|t}$$

$$x^j_{t|t} = x^j_t,$$

$$x^j_{k|t} \in \mathcal{X},\ u^j_{k|t} \in \mathcal{U},\ \forall k \in [t, \cdots, t+N-1]$$

$$x^j_{t+N|t} \in \mathcal{CS}^{j-1}(x),$$

Prediction Model

Value Function

Safe Set

Learning Model Predictive Controller

At time t of iteration j solve the following Constrained Finite Time Optimal Control Problem (CFTOCP)

$$J^{\mathrm{LMPC},j}_{t\rightarrow t+N}(x^j_t) = \min_{u^j_{t|t},\dots,u^j_{t+N-1|t}} \sum_{k=t}^{t+N-1} h(x^j_{k|t}, u^j_{k|t}) + V^{j-1}(x^j_{t+N|t}, x)$$

s.t.

$$x^j_{k+1|t} = A^j_{k|t} x^j_{k|t} + B^j_{k|t} u^j_{k|t} + C^j_{k|t}$$

$$x^j_{t|t} = x^j_t,$$

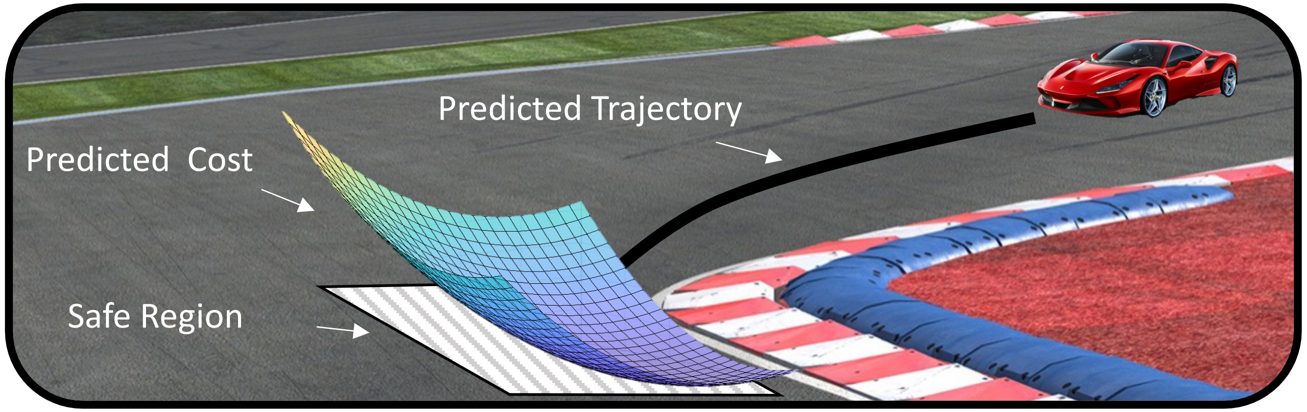
$$x^j_{k|t} \in \mathcal{X},\ u^j_{k|t} \in \mathcal{U},\ \forall k \in [t, \cdots, t+N-1]$$

$$x^j_{t+N|t} \in \mathcal{CS}^{j-1}(x),$$

Prediction Model

Value Function

Safe Set



Learning Model Predictive Controller

At time t of iteration j solve the following Constrained Finite Time Optimal Control Problem (CFTOCP)

$$J^{\mathrm{LMPC},j}_{t\to t+N}(x^j_t) = \min_{u^j_{t|t},\dots,u^j_{t+N-1|t}} \sum_{k=t}^{t+N-1} h(x^j_{k|t}, u^j_{k|t}) + V^{j-1}(x^j_{t+N|t}, {\color{red}x})$$

s.t.

$$x^j_{k+1|t} = A^j_{k|t} x^j_{k|t} + B^j_{k|t} u^j_{k|t} + C^j_{k|t}$$

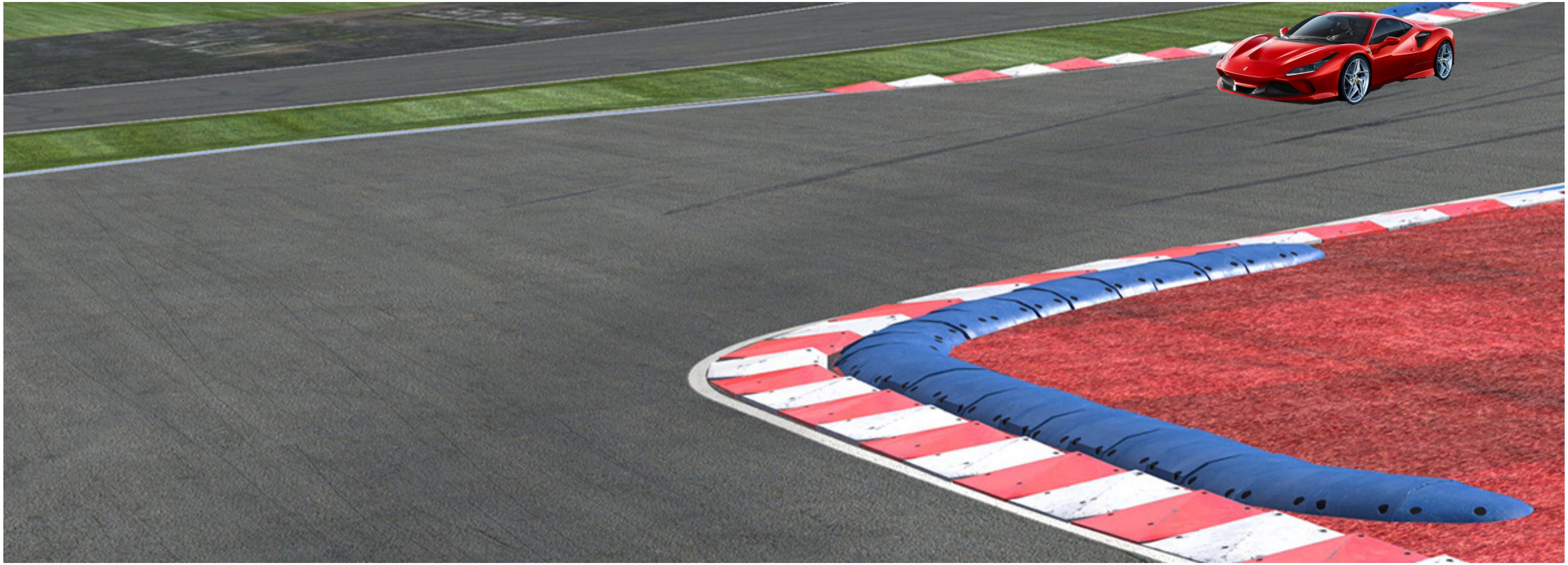
$$x^j_{t|t} = x^j_t,$$

$$x^j_{k|t} \in \mathcal{X},\ u^j_{k|t} \in \mathcal{U},\ \forall k \in [t, \cdots, t+N-1]$$

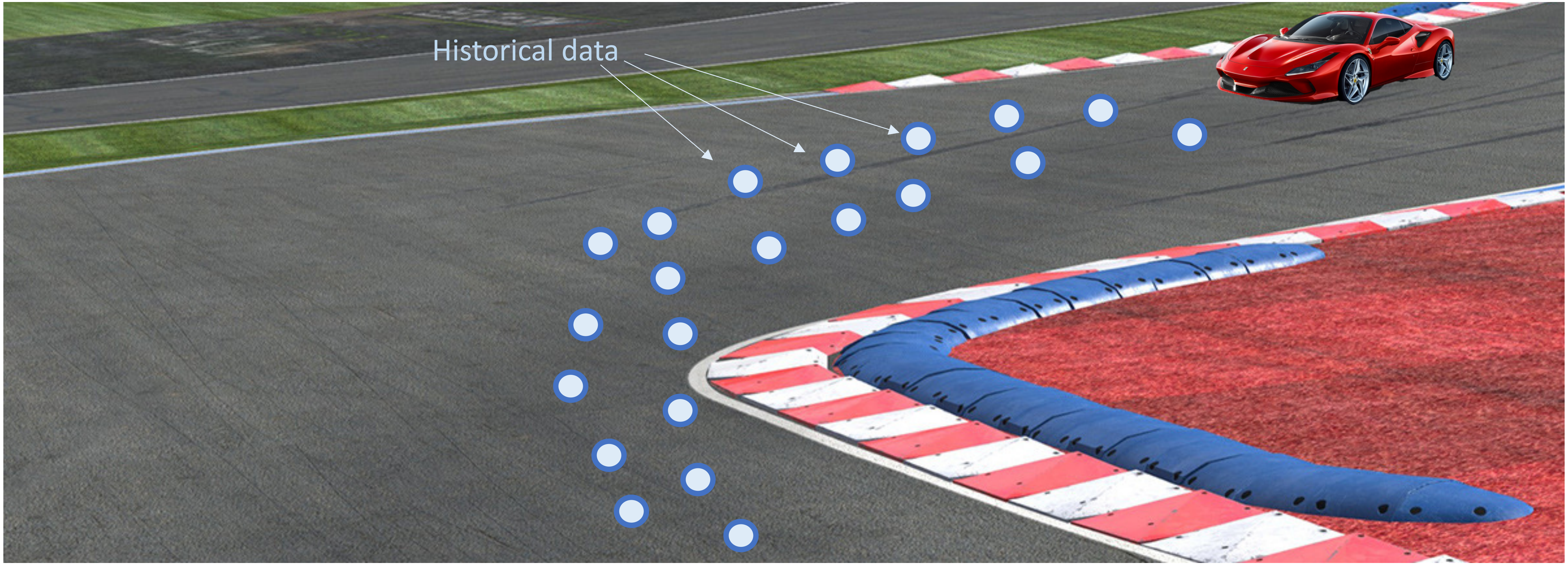
$$x^j_{t+N|t} \in \boxed{\mathcal{CS}^{j-1}({\color{red}x})},$$

Safe Set

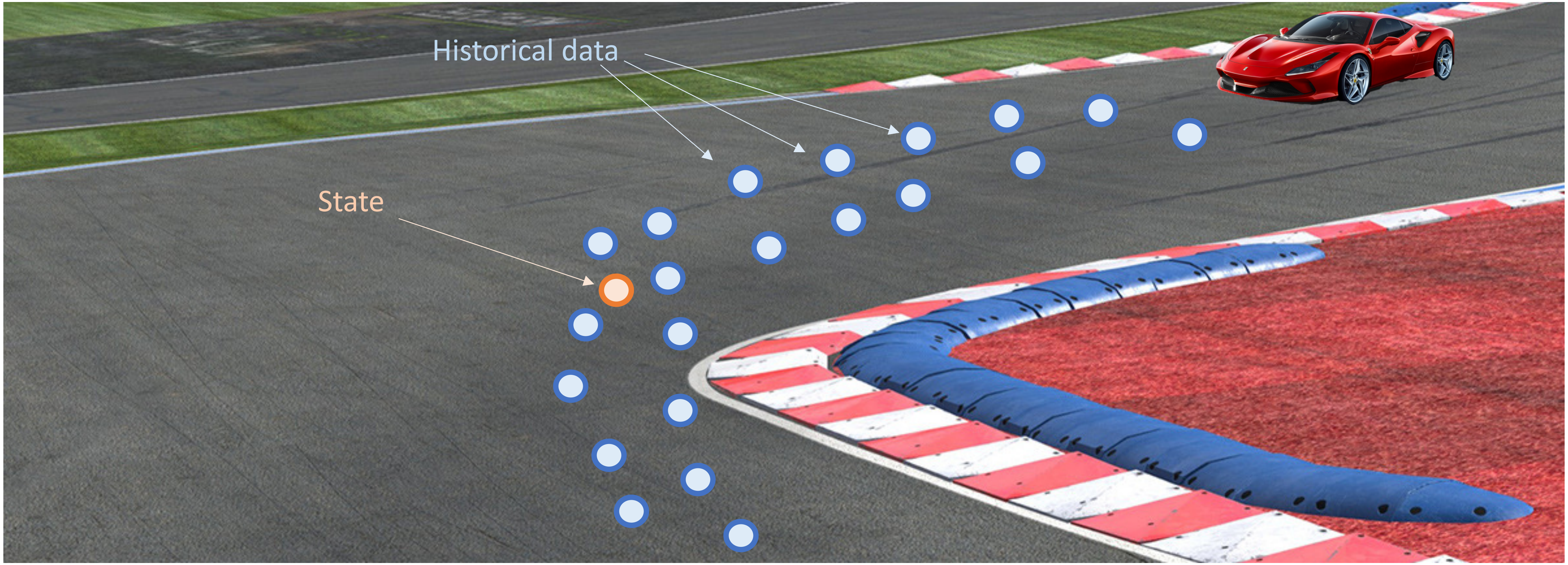
Safe Set Local Approximations



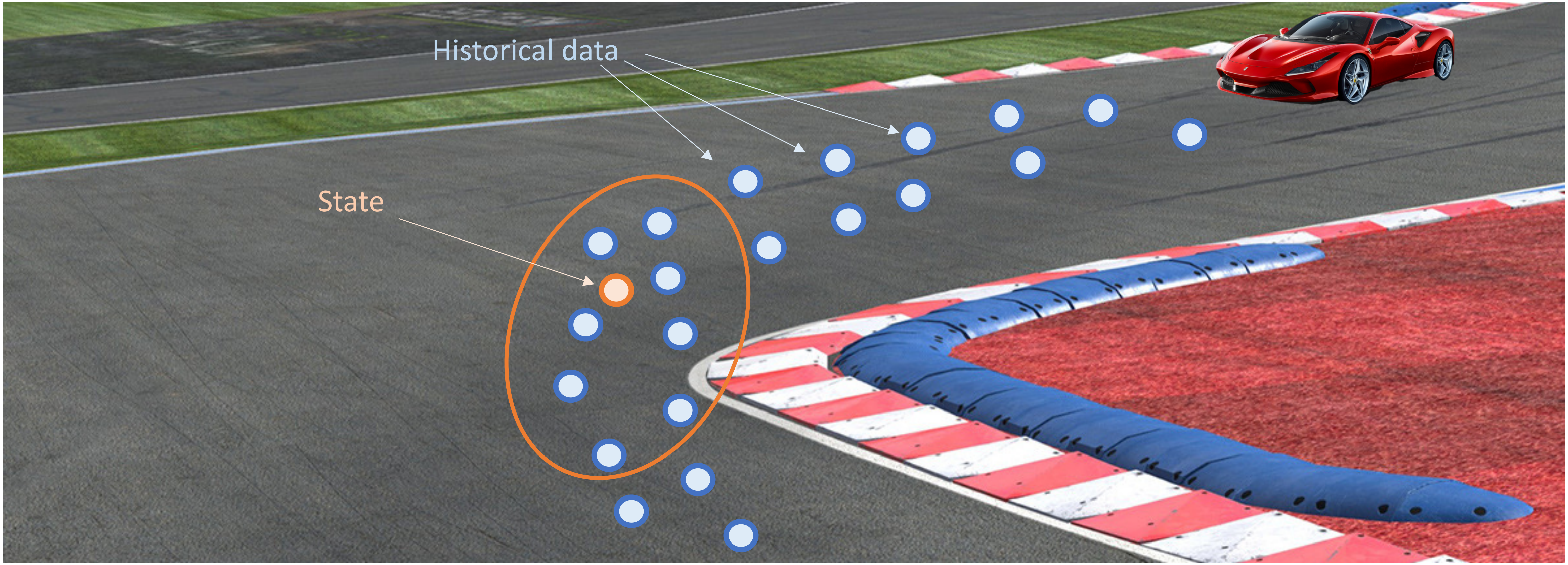
Safe Set Local Approximations



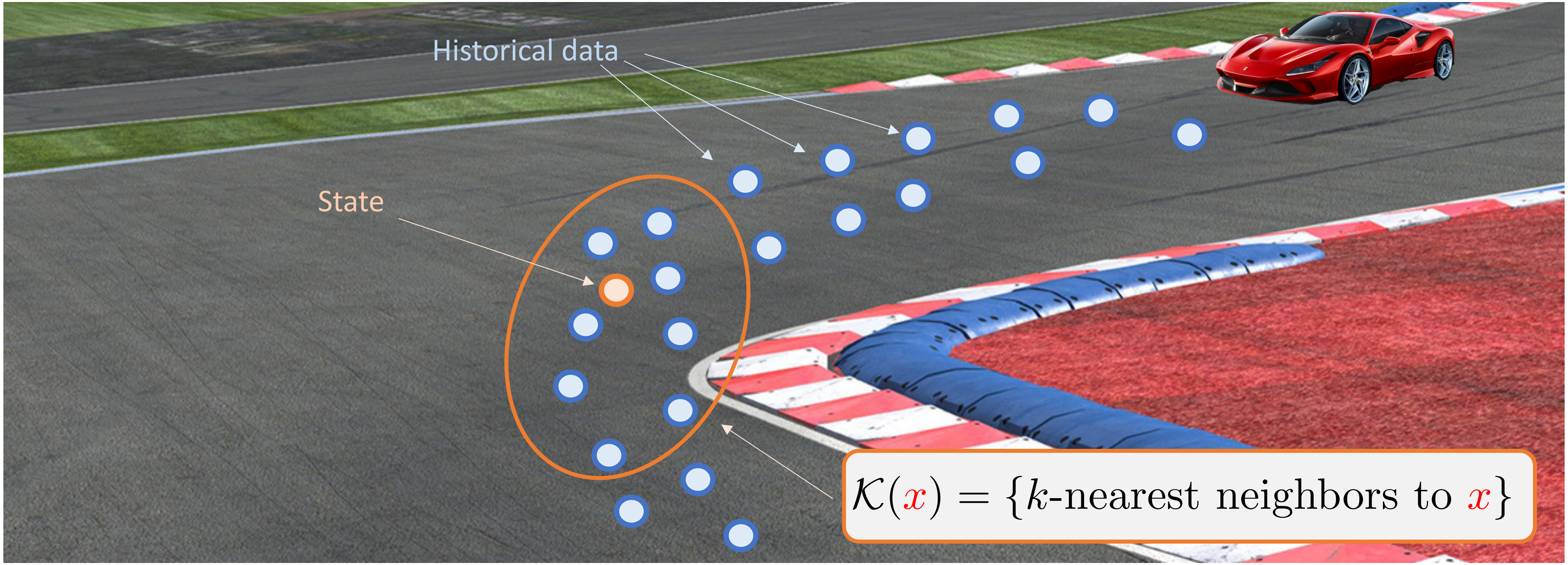
Safe Set Local Approximations



Safe Set Local Approximations

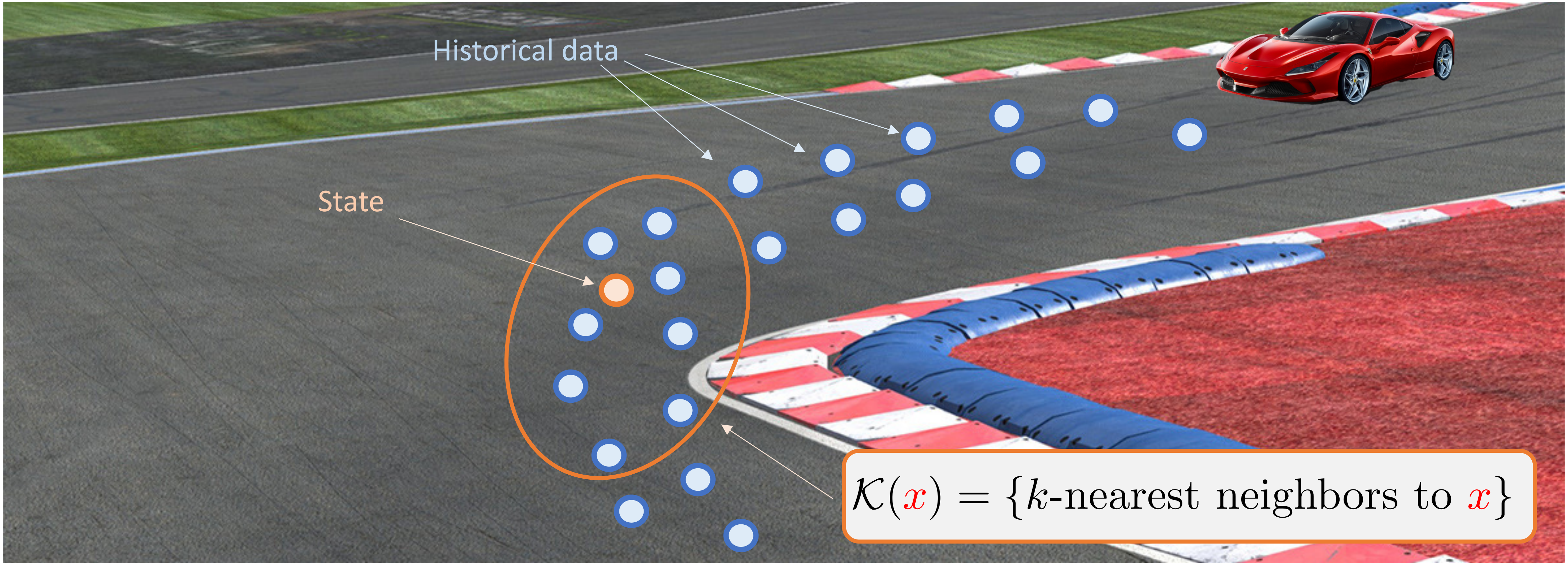


Safe Set Local Approximations



$$\mathcal{K}(x) = \{k\text{-nearest neighbors to } x\}$$

Safe Set Local Approximations



$\mathcal{K}(x) = \{k\text{-nearest neighbors to } x\}$

Local convex safe set approximation:

$$\mathcal{CS}^j(x) = \text{conv}\left(\cup_{x_t^j \in \mathcal{K}(x)} x_t^j \right)$$

Learning Model Predictive Controller

At time t of iteration j solve the following Constrained Finite Time Optimal Control Problem (CFTOCP)

$$J^{\mathrm{LMPC},j}_{t\rightarrow t+N}(x^j_t) = \min_{u^j_{t|t},\ldots,u^j_{t+N-1|t}} \sum_{k=t}^{t+N-1} h(x^j_{k|t}, u^j_{k|t}) + V^{j-1}(x^j_{t+N|t}, x)$$

s.t.

$$x^j_{k+1|t} = A^j_{k|t} x^j_{k|t} + B^j_{k|t} u^j_{k|t} + C^j_{k|t}$$

$$x^j_{t|t} = x^j_t,$$

$$x^j_{k|t} \in \mathcal{X},\ u^j_{k|t} \in \mathcal{U},\ \forall k \in [t, \cdots, t+N-1]$$

$$x^j_{t+N|t} \in \boxed{\mathcal{CS}^{j-1}(x),}$$

Safe Set

where $x = g(\text{Previous Optimal Trajectory})$

Learning Model Predictive Controller

At time t of iteration j solve the following Constrained Finite Time Optimal Control Problem (CFTOCP)

$$J^{\mathrm{LMPC},j}_{t\to t+N}(x^j_t) = \min_{u^j_{t|t},\dots,u^j_{t+N-1|t}} \sum_{k=t}^{t+N-1} h(x^j_{k|t}, u^j_{k|t}) + \boxed{V^{j-1}(x^j_{t+N|t}, {\color{red}x})}$$

s.t.

$$x^j_{k+1|t} = A^j_{k|t} x^j_{k|t} + B^j_{k|t} u^j_{k|t} + C^j_{k|t}$$

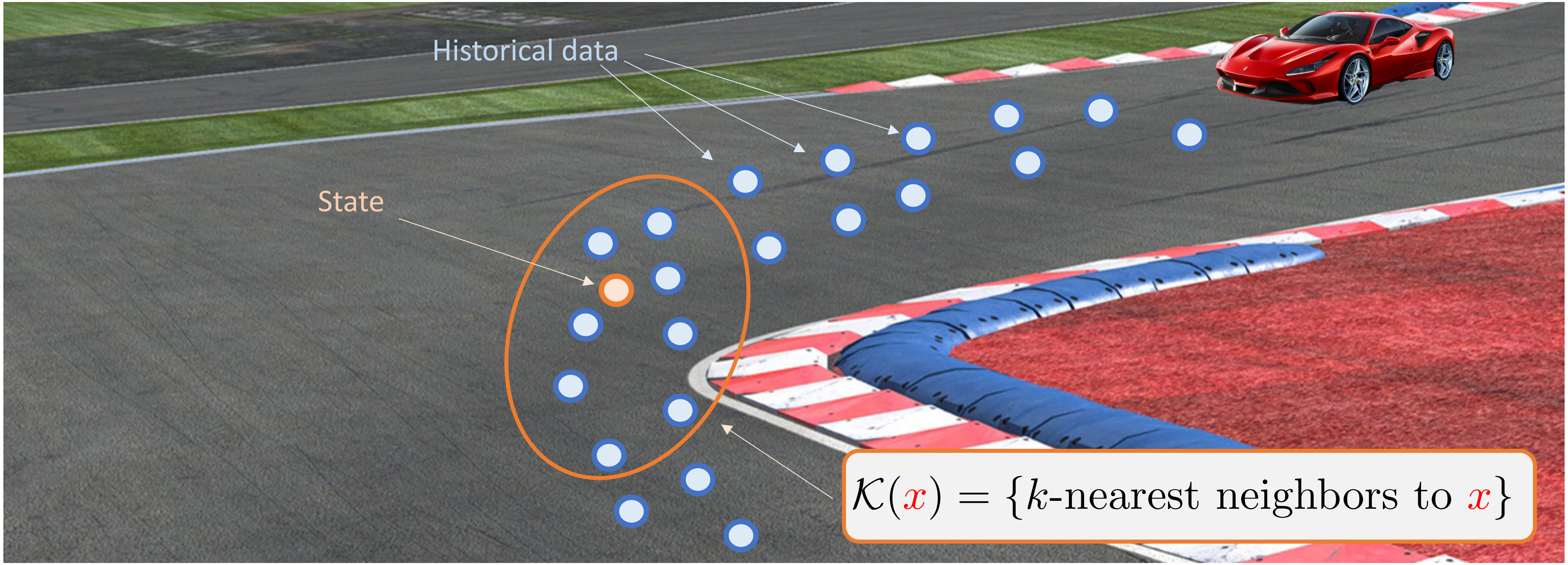
$$x^j_{t|t} = x^j_t,$$

$$x^j_{k|t} \in \mathcal{X},\ u^j_{k|t} \in \mathcal{U},\ \forall k \in [t, \cdots, t+N-1]$$

$$x^j_{t+N|t} \in \ \mathcal{CS}^{j-1}({\color{red}x}),$$

Value Function

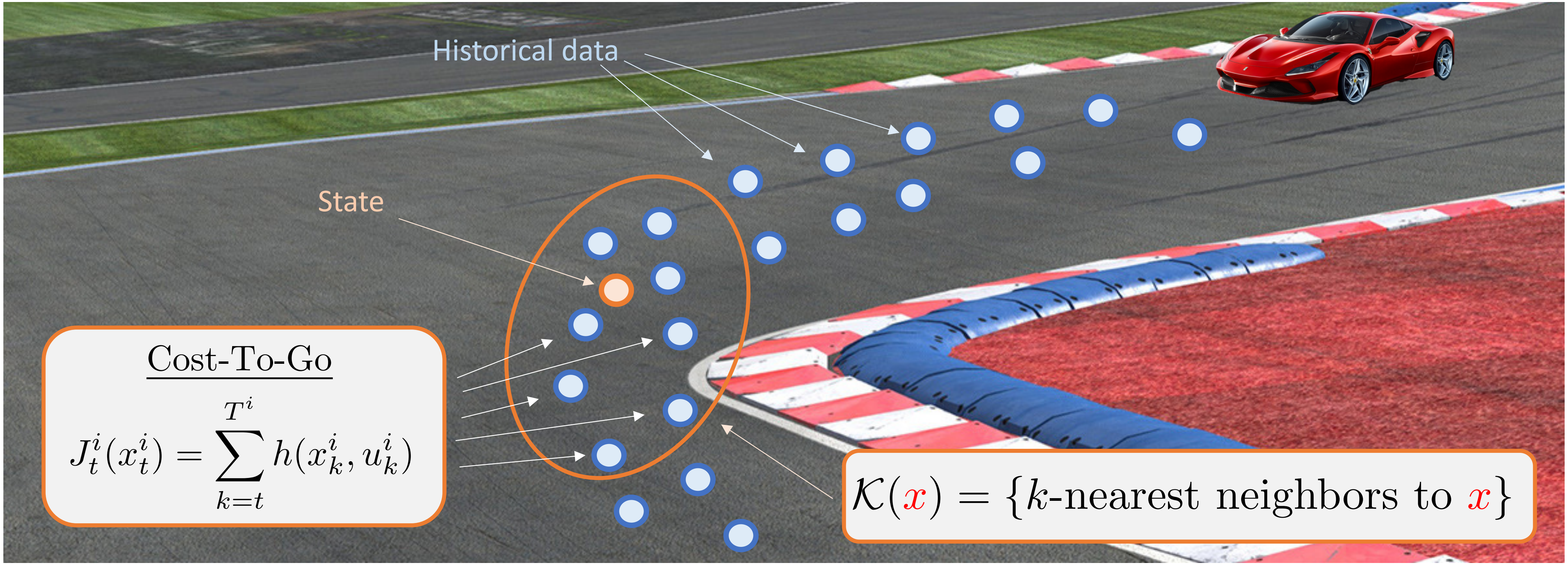
Value Function Local Approximations



Local convex safe set approximation:

$$\mathcal{CS}^j(x) = \mathrm{conv}\left(\cup_{x_t^j \in \mathcal{K}(x)} x_t^j \right)$$

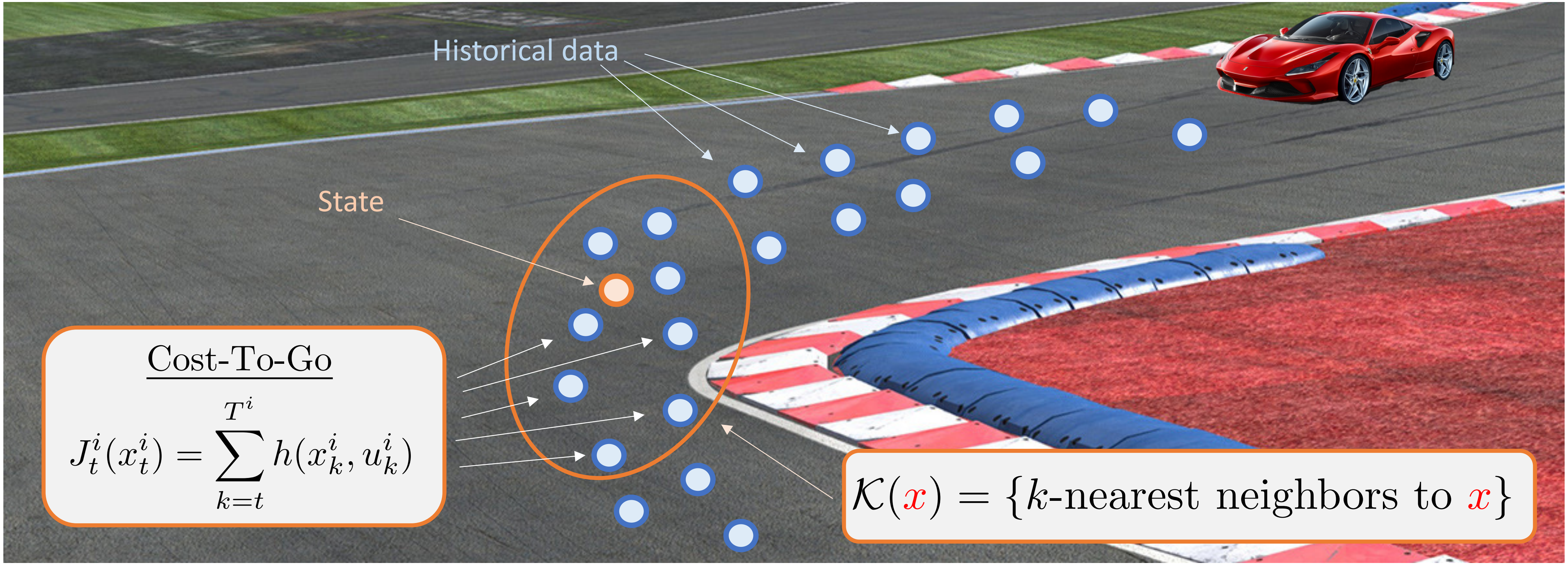
Value Function Local Approximations



Local convex safe set approximation:

$$\mathcal{CS}^j(x) = \text{conv}\left(\cup_{x_t^j \in \mathcal{K}(x)} x_t^j \right)$$

Value Function Local Approximations



Local value function approximation:

$$V^j(x, x) = \text{ Interpolation of the cost-to-go } J_t^i(x_t^i) = \sum_{k=t}^{T^i} h(x_k^i, u_k^i)$$

Learning Model Predictive Controller

At time t of iteration j solve the following Constrained Finite Time Optimal Control Problem (CFTOCP)

$$J^{\mathrm{LMPC},j}_{t\rightarrow t+N}(x^j_t) = \min_{u^j_{t|t},\ldots,u^j_{t+N-1|t}} \sum_{k=t}^{t+N-1} h(x^j_{k|t}, u^j_{k|t}) + V^{j-1}(x^j_{t+N|t}, x)$$

s.t.

$$x^j_{k+1|t} = A^j_{k|t} x^j_{k|t} + B^j_{k|t} u^j_{k|t} + C^j_{k|t}$$

$$x^j_{t|t} = x^j_t,$$

$$x^j_{k|t} \in \mathcal{X},\ u^j_{k|t} \in \mathcal{U},\ \forall k \in [t, \cdots, t+N-1]$$

$$x^j_{t+N|t} \in \ \mathcal{CS}^{j-1}(x),$$

Prediction Model

U. Rosolia, and F. Borrelli. "Learning how to autonomously race a car: a predictive control approach." IEEE Transactions on Control Systems Technology 28, no. 6 (2019): 2713-2719.

Learning Model Predictive Controller

At time t of iteration j solve the following Constrained Finite Time Optimal Control Problem (CFTOCP)

$$J^{\mathrm{LMPC},j}_{t\rightarrow t+N}(x^j_t) = \min_{u^j_{t|t},\dots,u^j_{t+N-1|t}} \sum_{k=t}^{t+N-1} h(x^j_{k|t}, u^j_{k|t}) + V^{j-1}(x^j_{t+N|t}, x)$$

s.t.

$$x^j_{k+1|t} = A^j_{k|t} x^j_{k|t} + B^j_{k|t} u^j_{k|t} + C^j_{k|t}$$

$$x^j_{t|t} = x^j_t,$$

$$x^j_{k|t} \in \mathcal{X},\ u^j_{k|t} \in \mathcal{U},\ \forall k \in [t,\cdots,t+N-1]$$

$$x^j_{t+N|t} \in \mathcal{CS}^{j-1}(x),$$

Prediction Model

Value Function

Safe Set

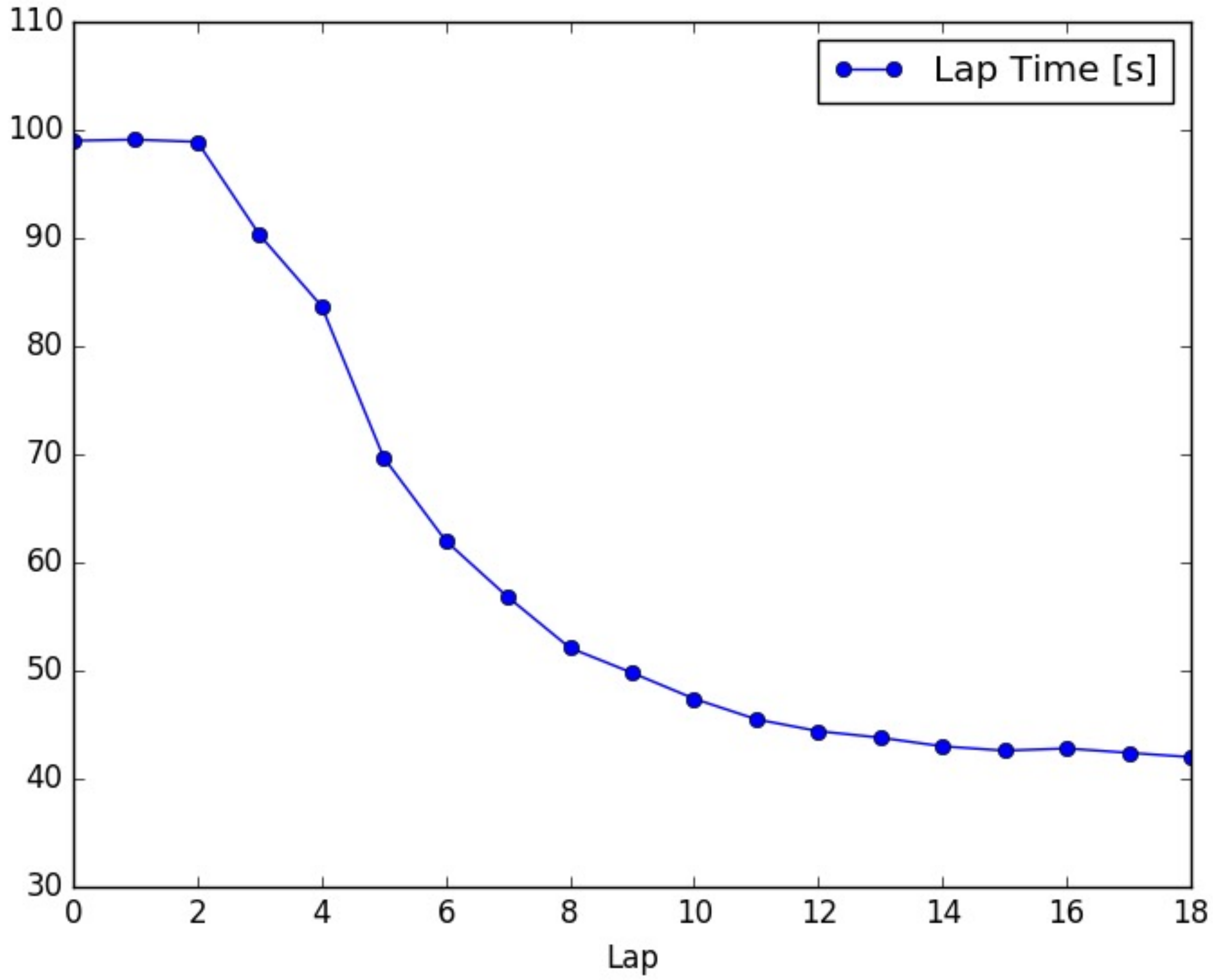
Learning Model Predictive Controller full-size vehicle experiments

Credits: Siddharth Nair, Nitin Kapania and Ugo Rosolia

Learning Model Predictive Controller full-size vehicle experiments

Credits: Siddharth Nair, Nitin Kapania and Ugo Rosolia

Lap Time



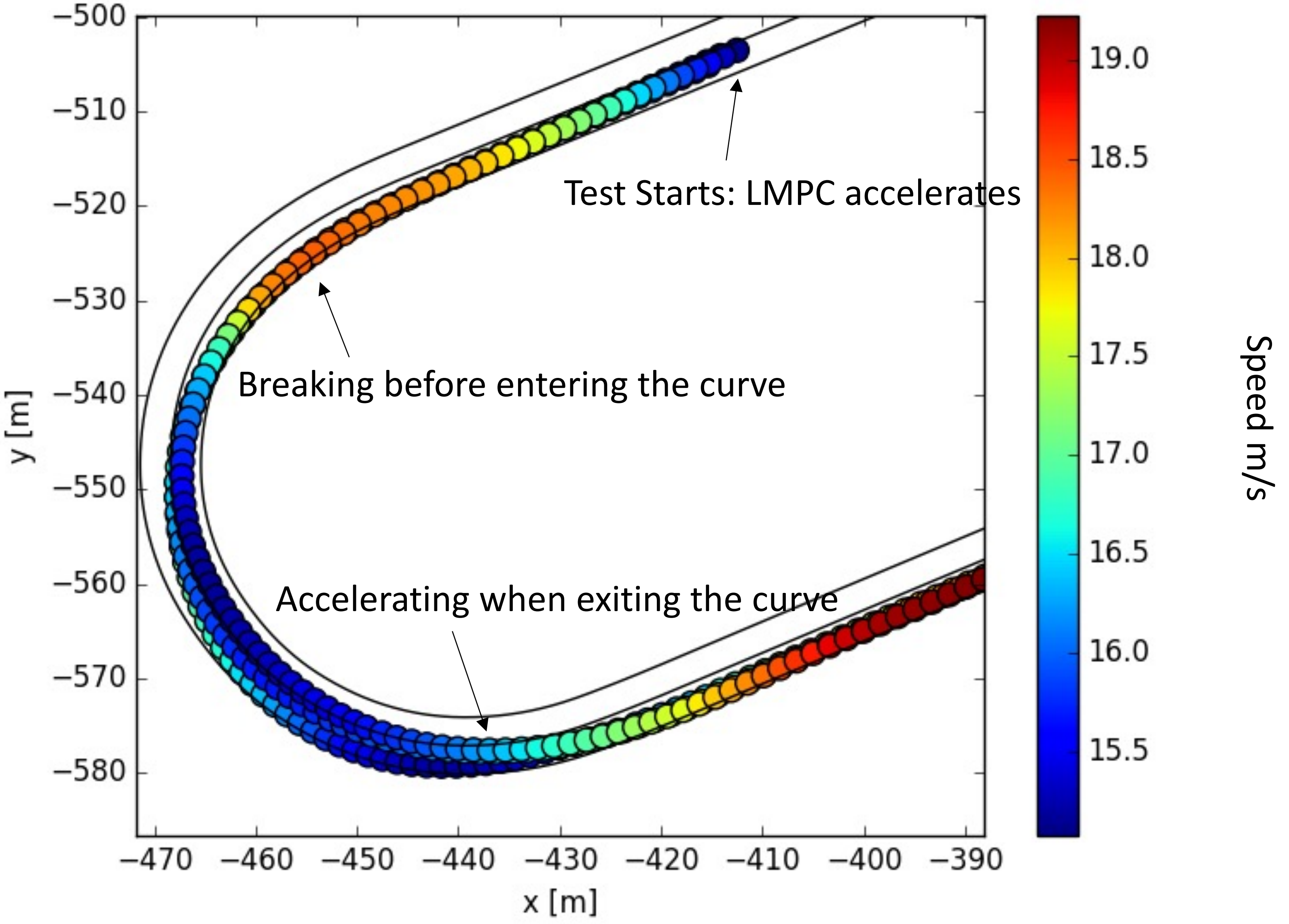
The control policy is constructed using ~1k data points (last 2 laps)

The control action is computed using ~100 data points

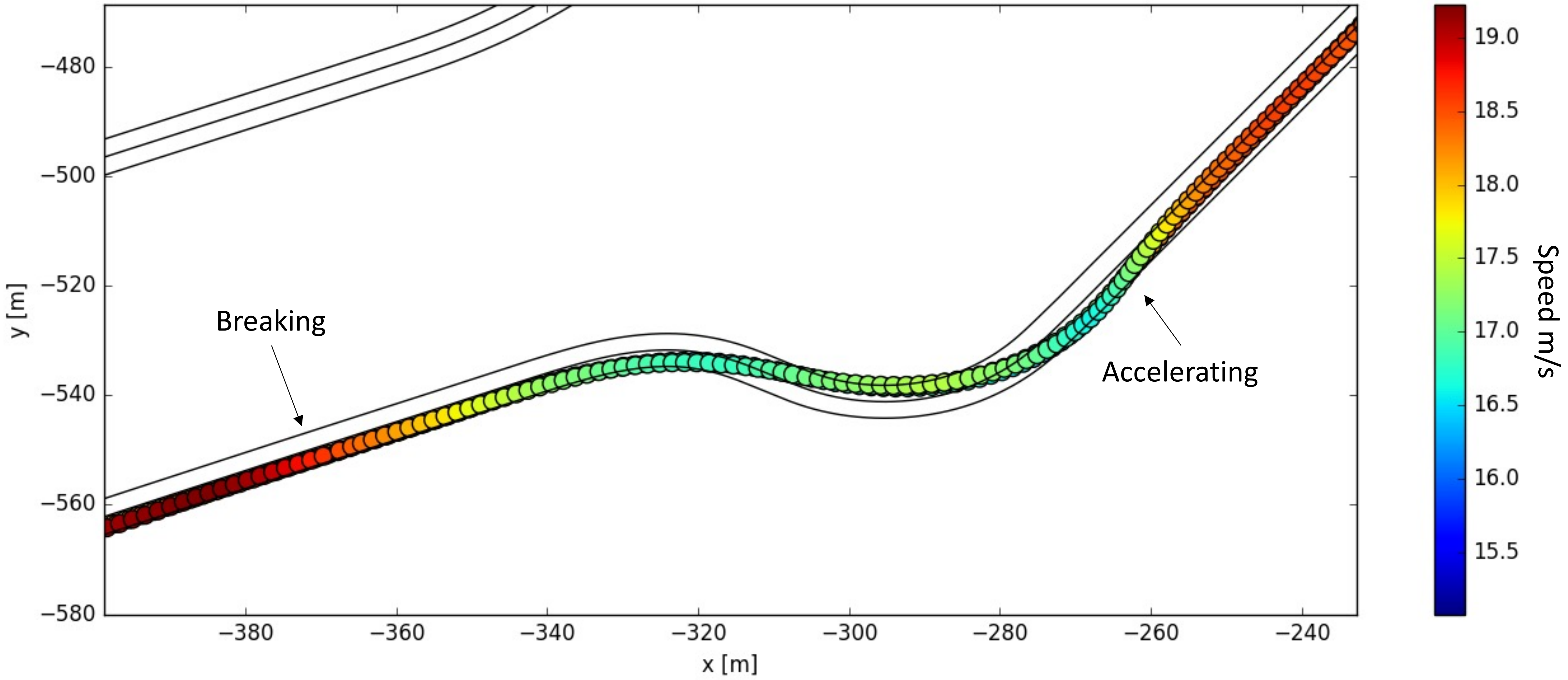
Learning Model Predictive Controller full-size vehicle experiments

Credits: Siddharth Nair, Nitin Kapania and Ugo Rosolia

Velocity Profile at Convergence (Curve 1)

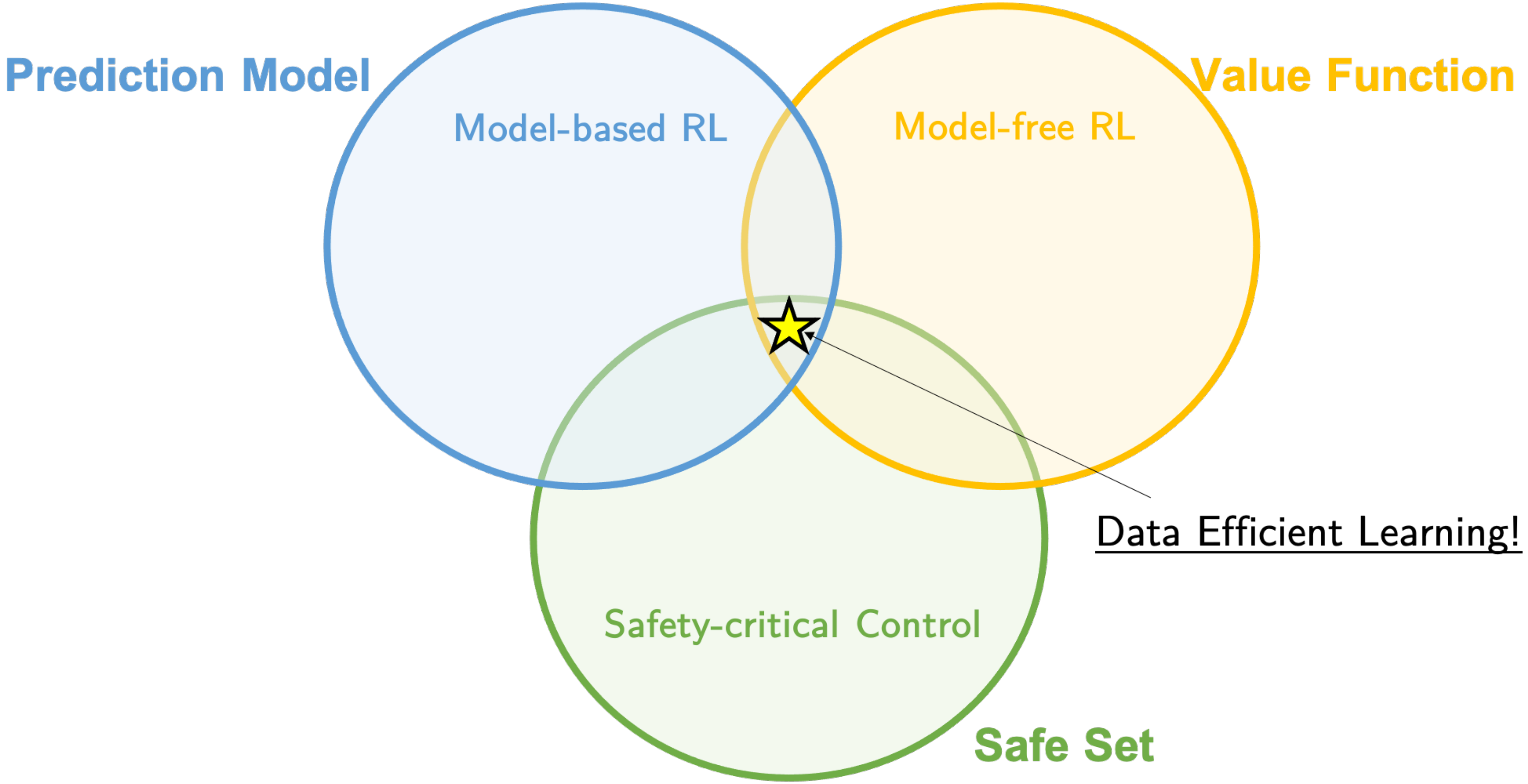


Velocity Profile at Convergence (Chicane)



The key components

- Predicted trajectory given by prediction model
- Predicted cost estimated by value function
- Safe region estimated by the safe set

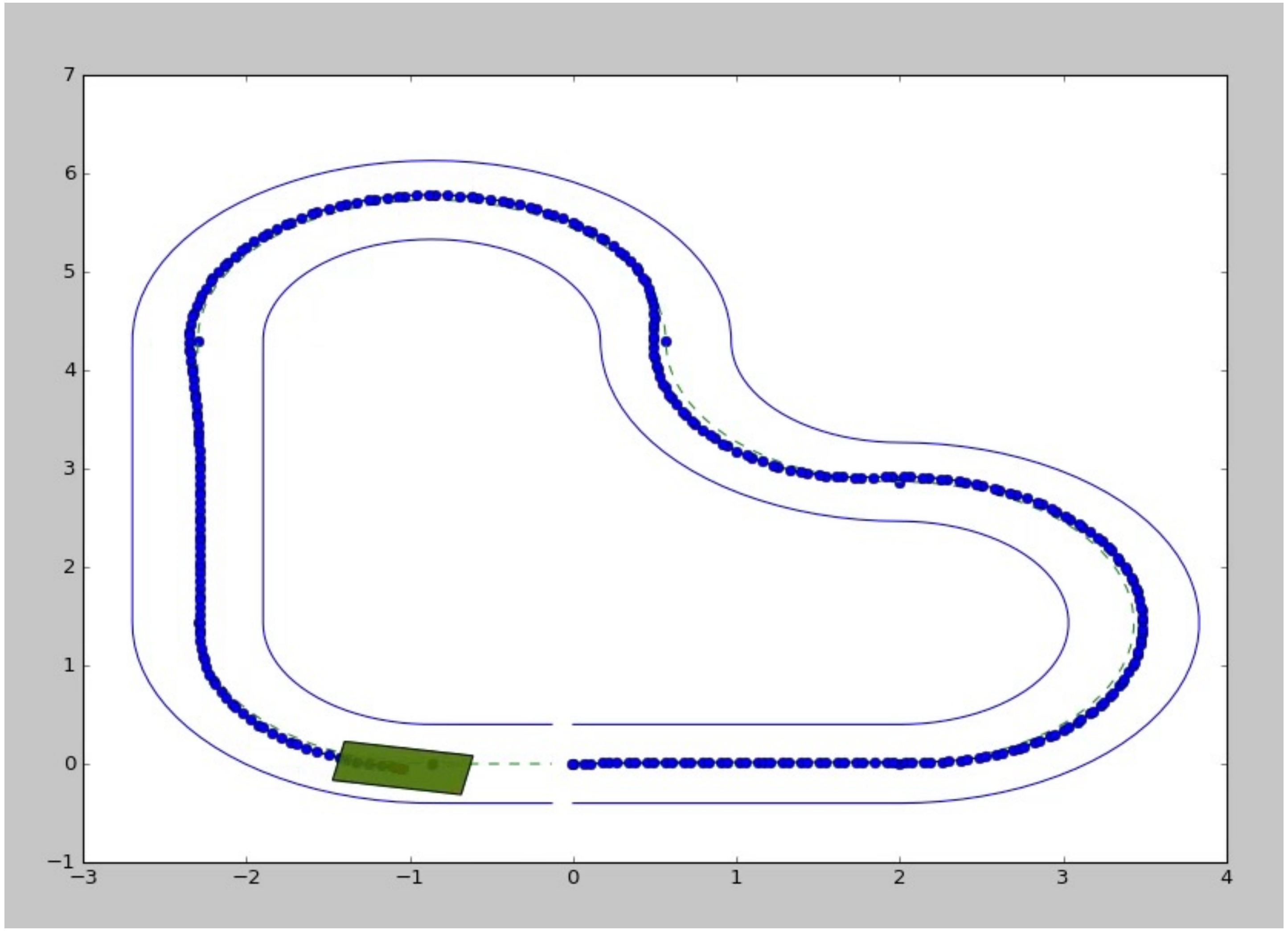


Do you need the safe set?

Do you need the safe set? – Yes

LMPC without the safe set

The controller extrapolates the value function on the Vx dimension

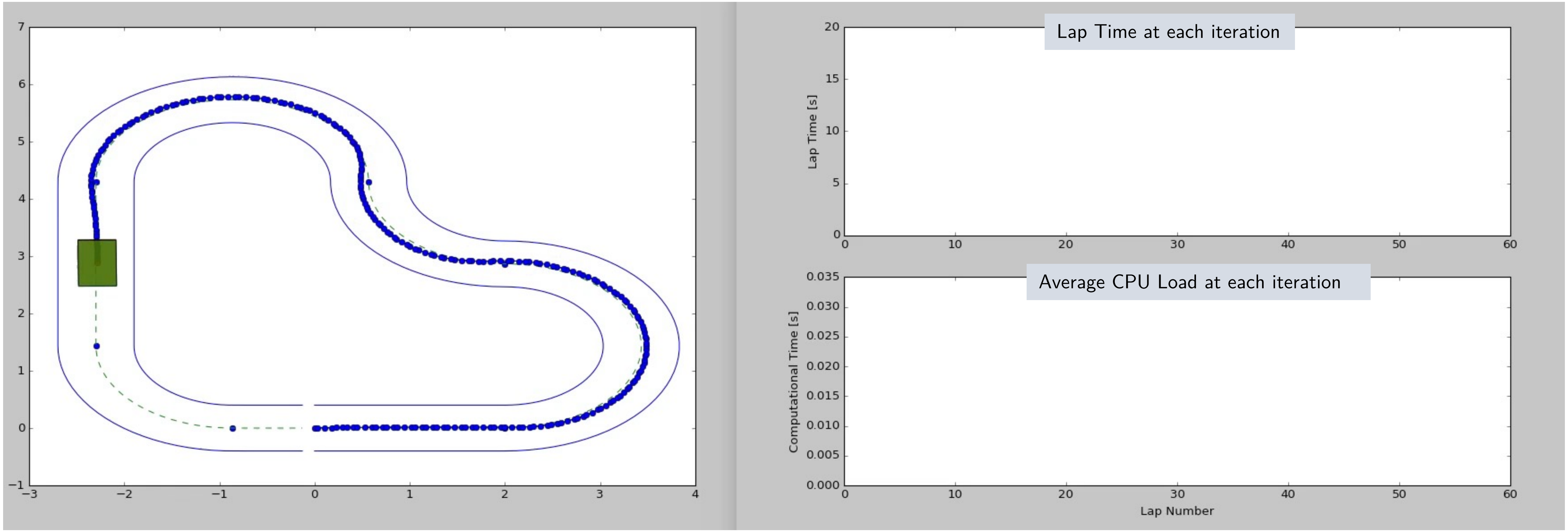


Do you need to Predict to Learn?

Do you need to Predict to Learn? Yes

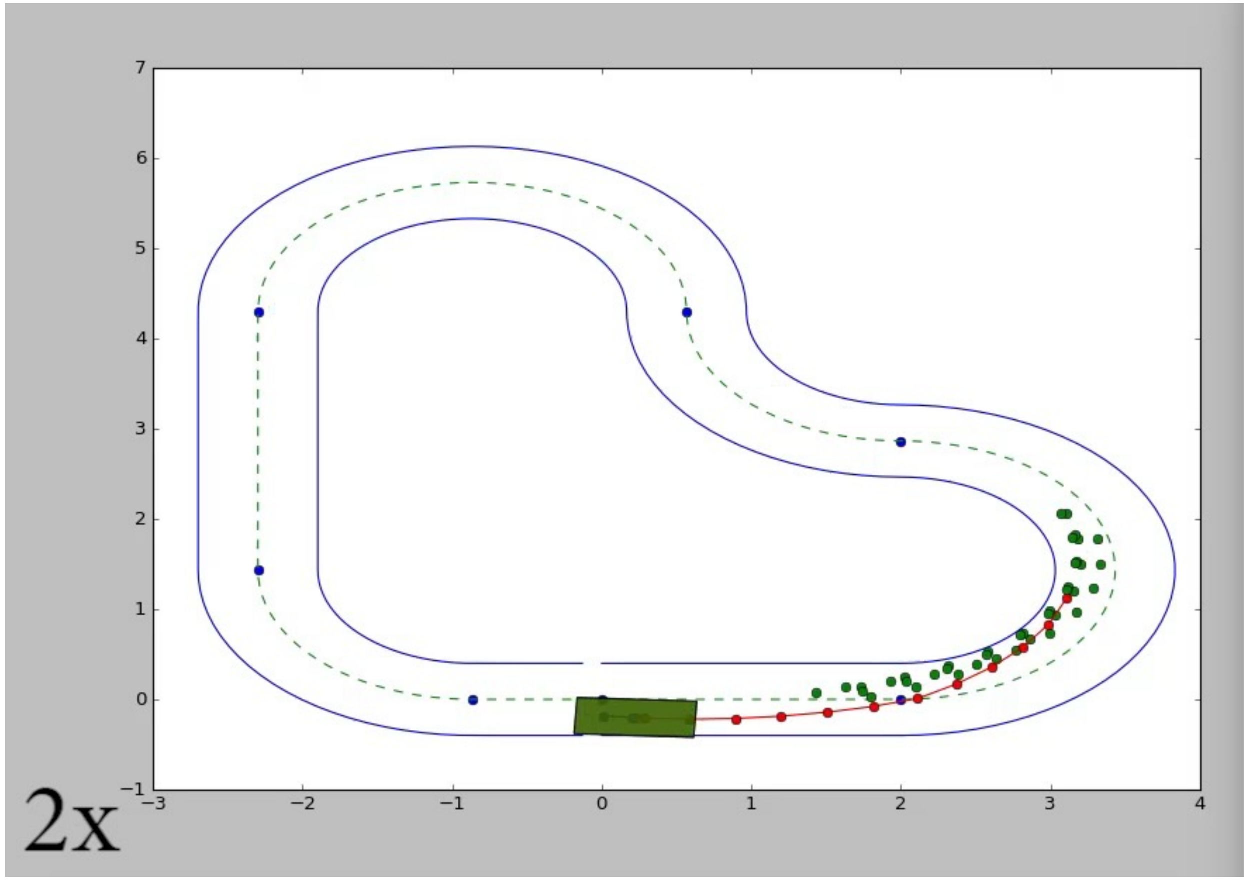
When the LMPC horizon is $N = 1$ the controller

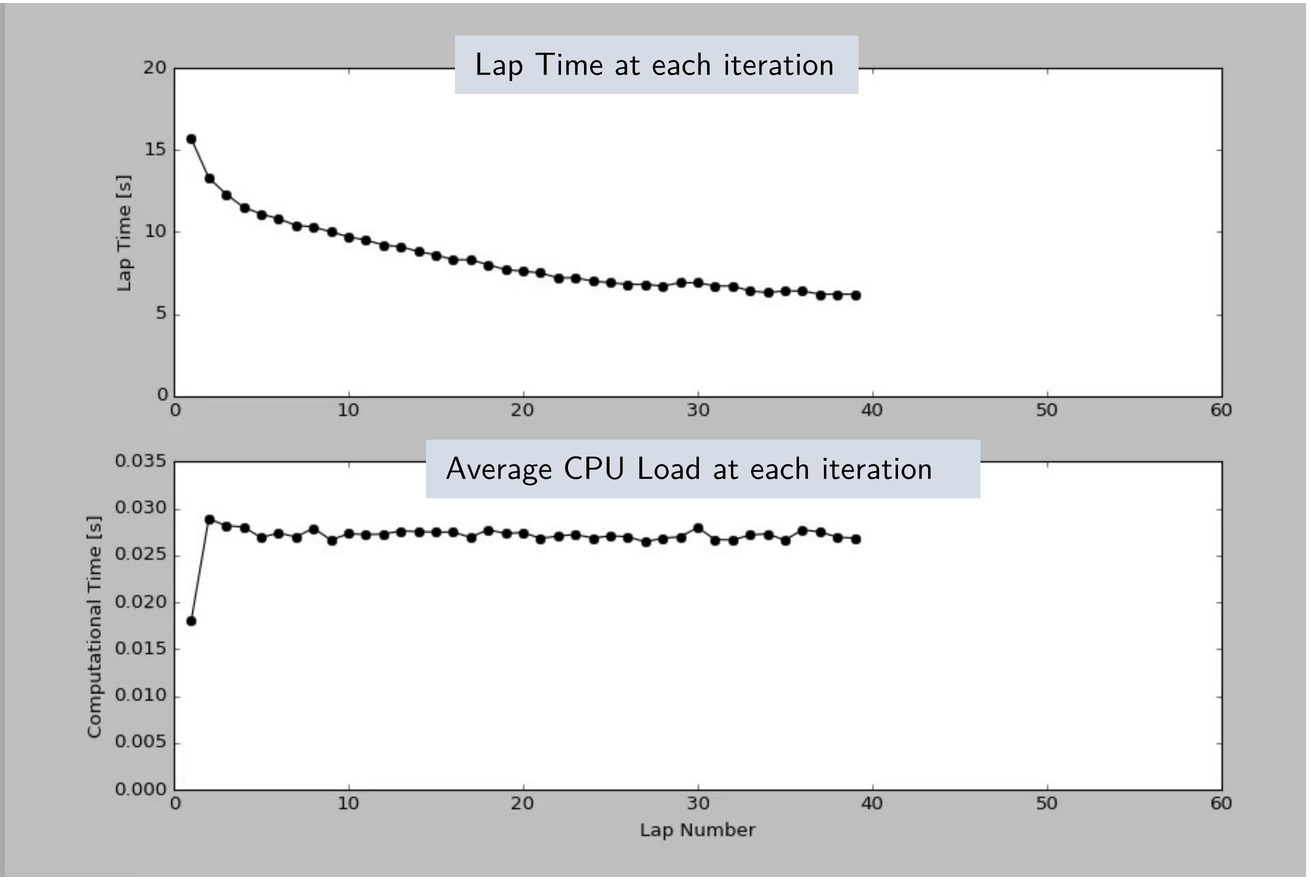
- solves the Bellman equation using the value function approximation
- does not explore the state space as it cannot plan outside the safe set



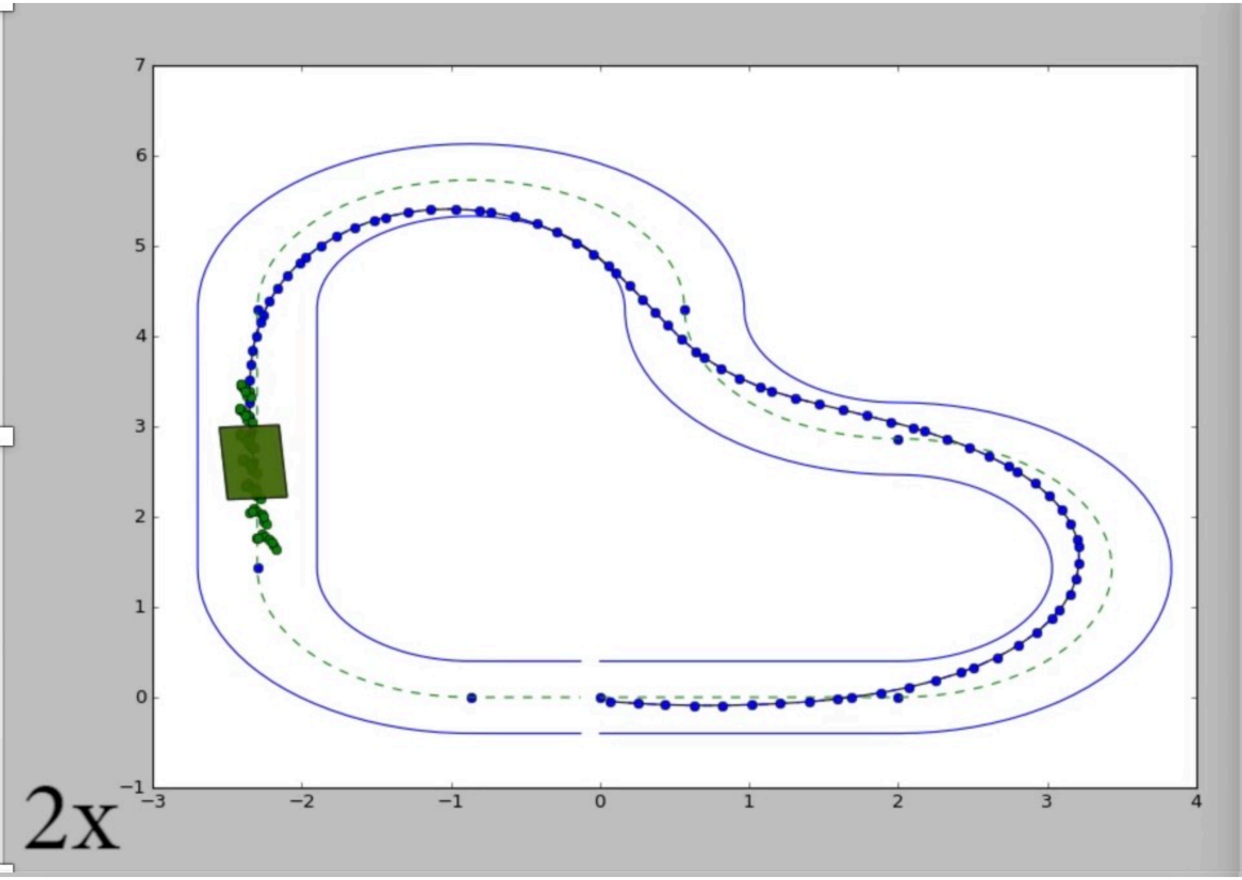
Do you need to Predict at Convergence?

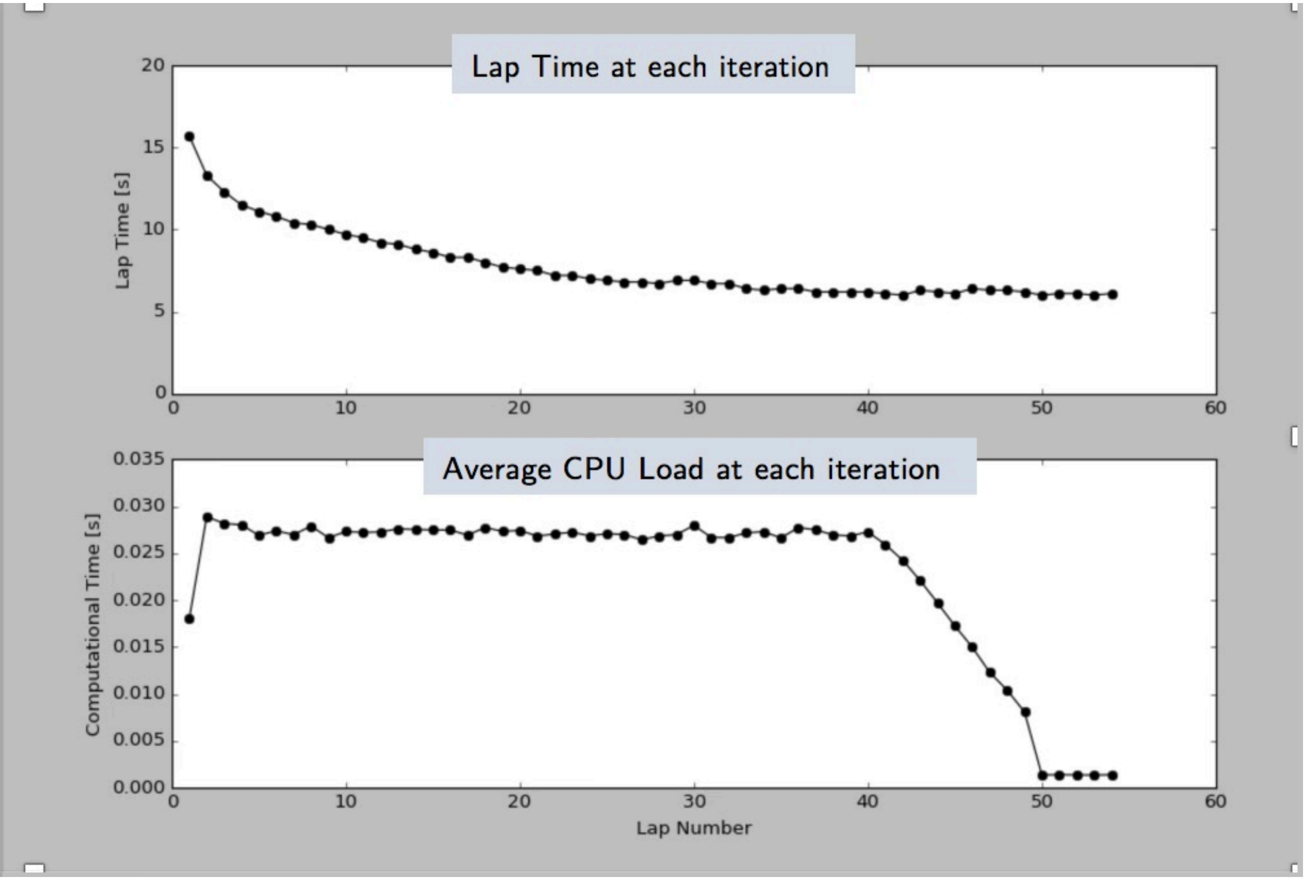
Do you need to Predict at Convergence? No





Do you need to Predict at Convergence? No





Value Function Approximation

$$[\lambda_0^{0,*}, \ldots, \lambda_i^{j,*}] = \arg\min_{\lambda_i^j \in [0,1]} \quad \sum_i \sum_j J_i^j \lambda_i^j$$

$$\text{s.t} \quad \sum_i \sum_j x_i^j \lambda_i^j = x(t),$$

$$\sum_i \sum_j \lambda_i^j = 1$$

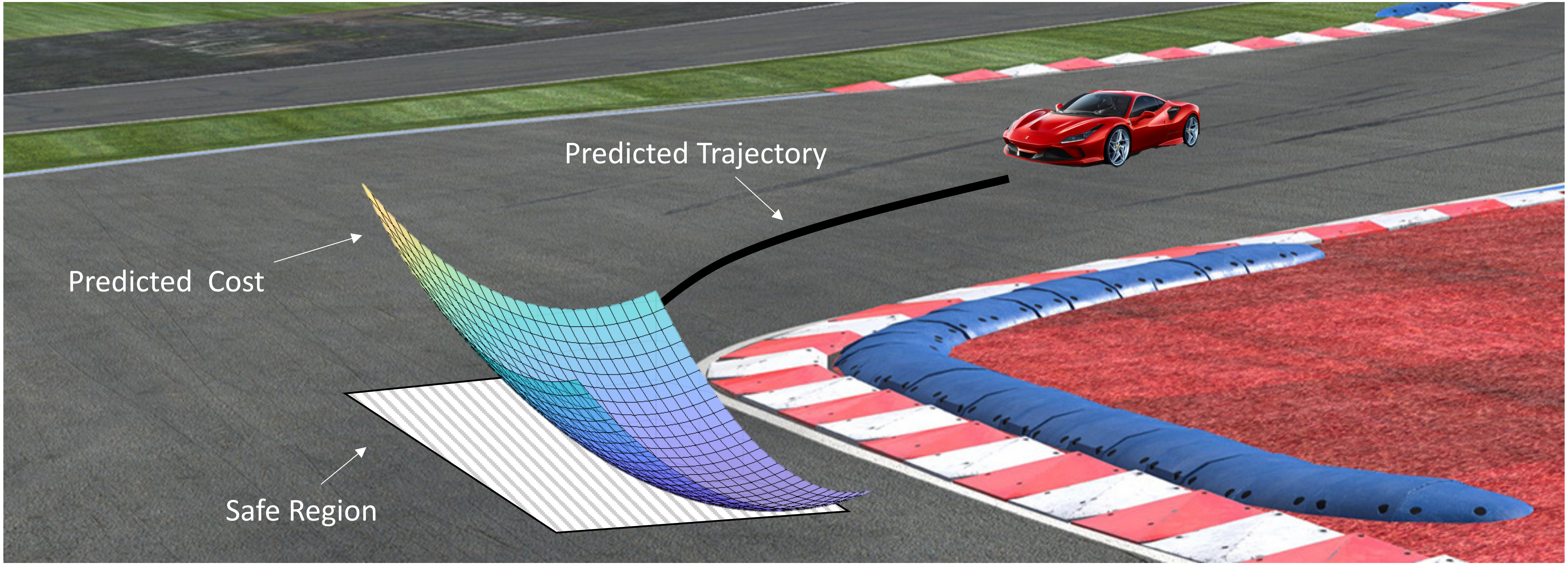
Control Policy

Stored Data

$$\pi\big(x(t)\big) = \sum_i \sum_j u_i^j \lambda_i^{j,*}$$

The key components

- Predicted trajectory given by prediction model
- Predicted cost estimated by value function
- Safe region estimated by the safe set

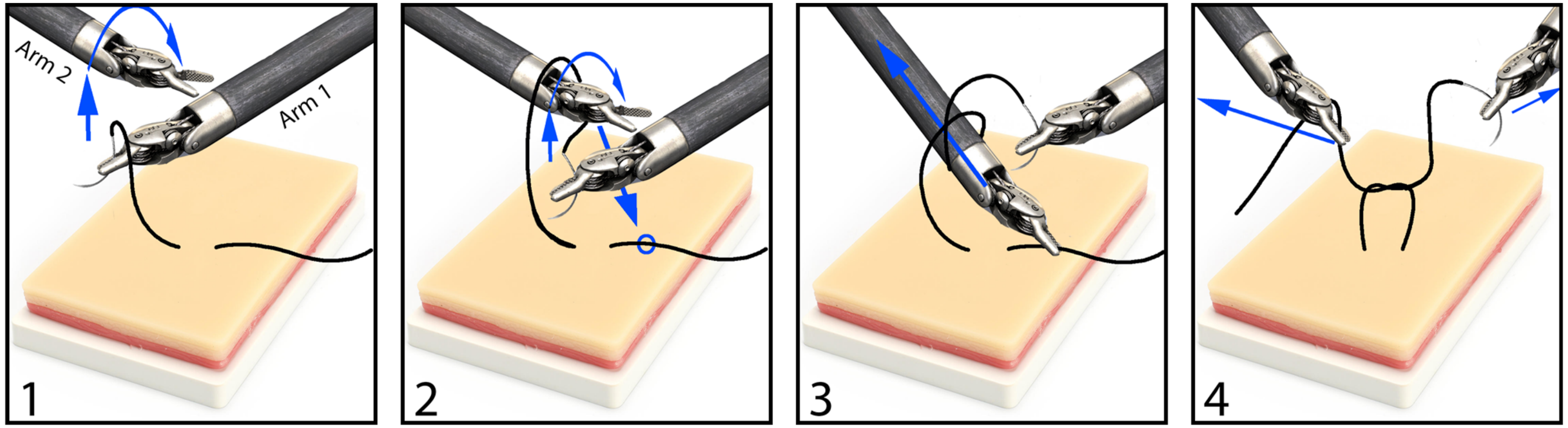


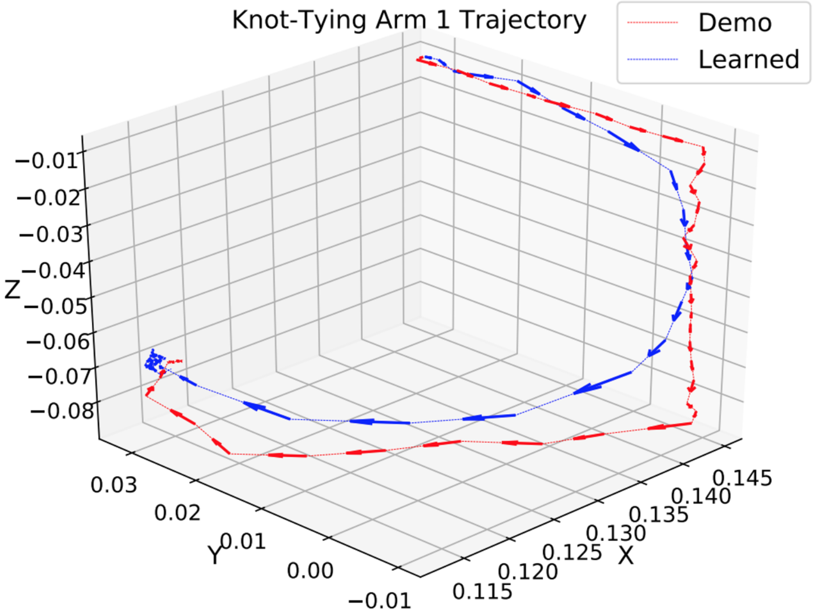
What about more complicated systems?

SAVED: Surgical Knot Tying

Brijen

Ashwin





- Safe Set constructed using non-parametric estimation
- Model ensemble and input sampling strategies for MPC
- Knot tying task on real surgical robot with inefficient demos (red)
- Constraints: stay within 1 cm tube of reference trajectory
- SAVED successfully smooths + optimizes demos

“Safety Augmented Value Estimation from Demonstrations (SAVED): Safe Deep Model-Based RL for Sparse Cost Robotic Tasks.”, B. Thananjeyan*, A. Balakrishna*, U. Rosolia, F. Li, R. McAllister, J. E. Gonzalez, S. Levine, F. Borrelli, K. Goldberg *IEEE Robotics and Automation Letters (RA-L)* (2020)

*= equal contribution

Adjustable Boundary Conditions

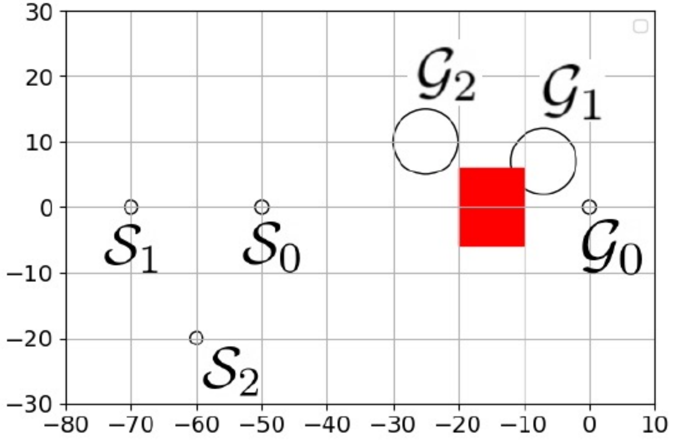
Brijen

Ashwin

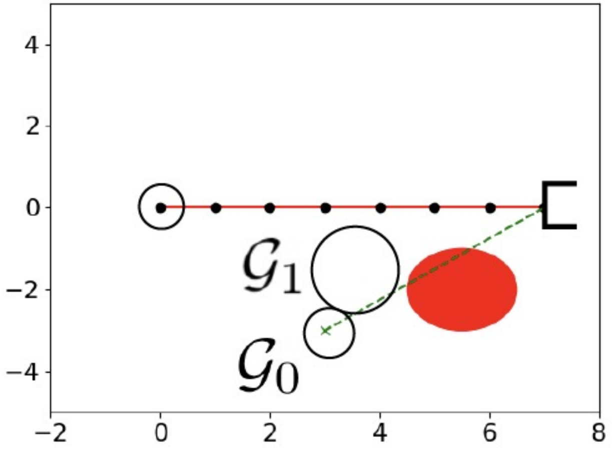
- Analysis for nonlinear stochastic systems
- Results for expected cost at each execution of the control task
- Safe set shaping for generalization to different initial and terminal conditions
- Exploration strategies to systematically pick the initial condition (domain expansion)

$$\mathcal{S}_i = \text{starting set}, \; \mathcal{G}_i = \text{goal set}$$

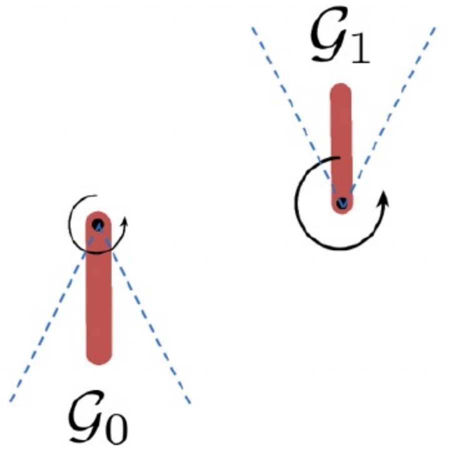
Point Navigation



2-D 7-Link Arm Reacher



Inverted Pendulum

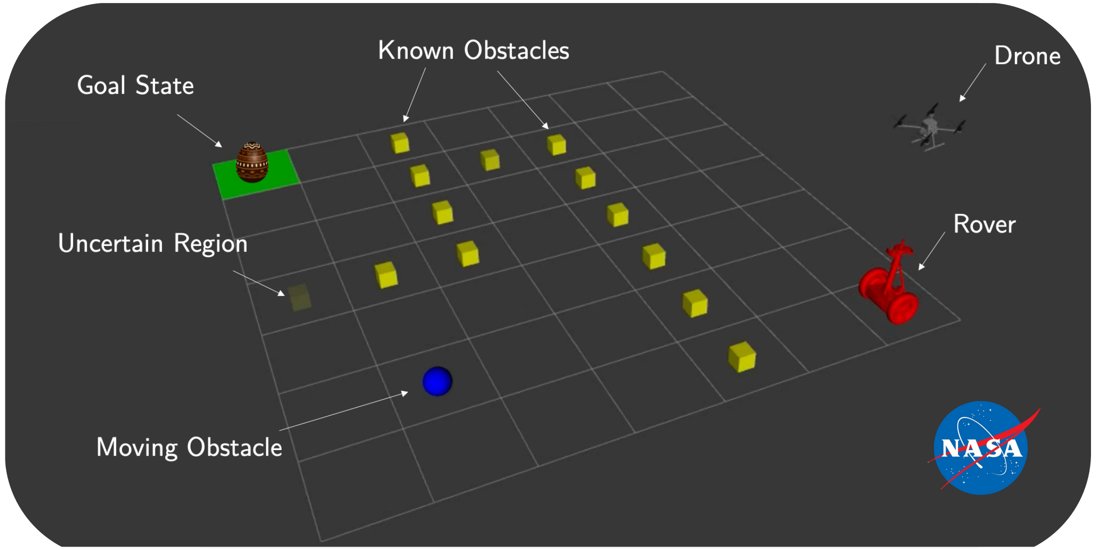


"ABC-LMPC: Safe Sample-Based Learning MPC for Stochastic Nonlinear Dynamical Systems with Adjustable Boundary Conditions." B. Thananjeyan*, A. Balakrishna*, U. Rosolia, J. E. Gonzalez, A. D. Ames, and K. Goldberg. *The 14th international Workshop on the Algorithm Foundation of Robotics (WAFR)*, (2020).

*= equal contribution

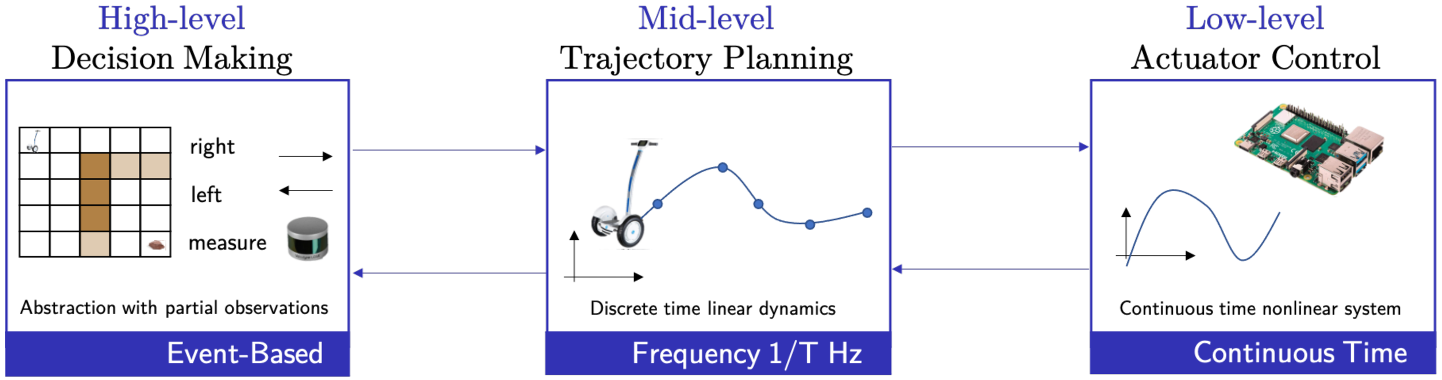
What is next?

- Partial Observability

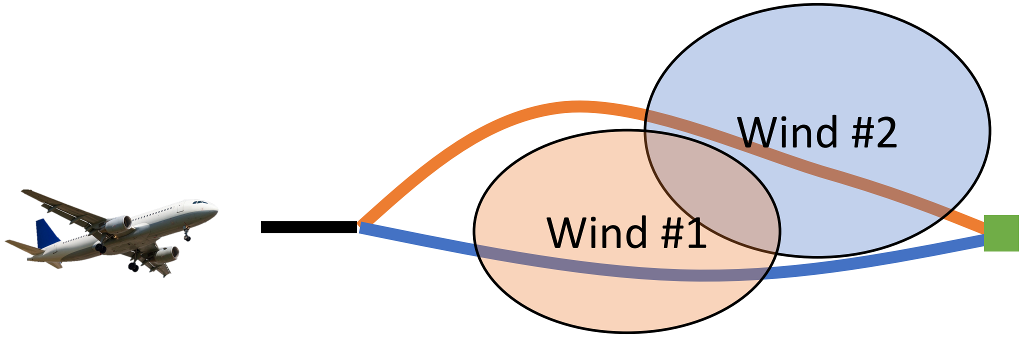


- Multi-agent systems

- Hierarchy + Learning

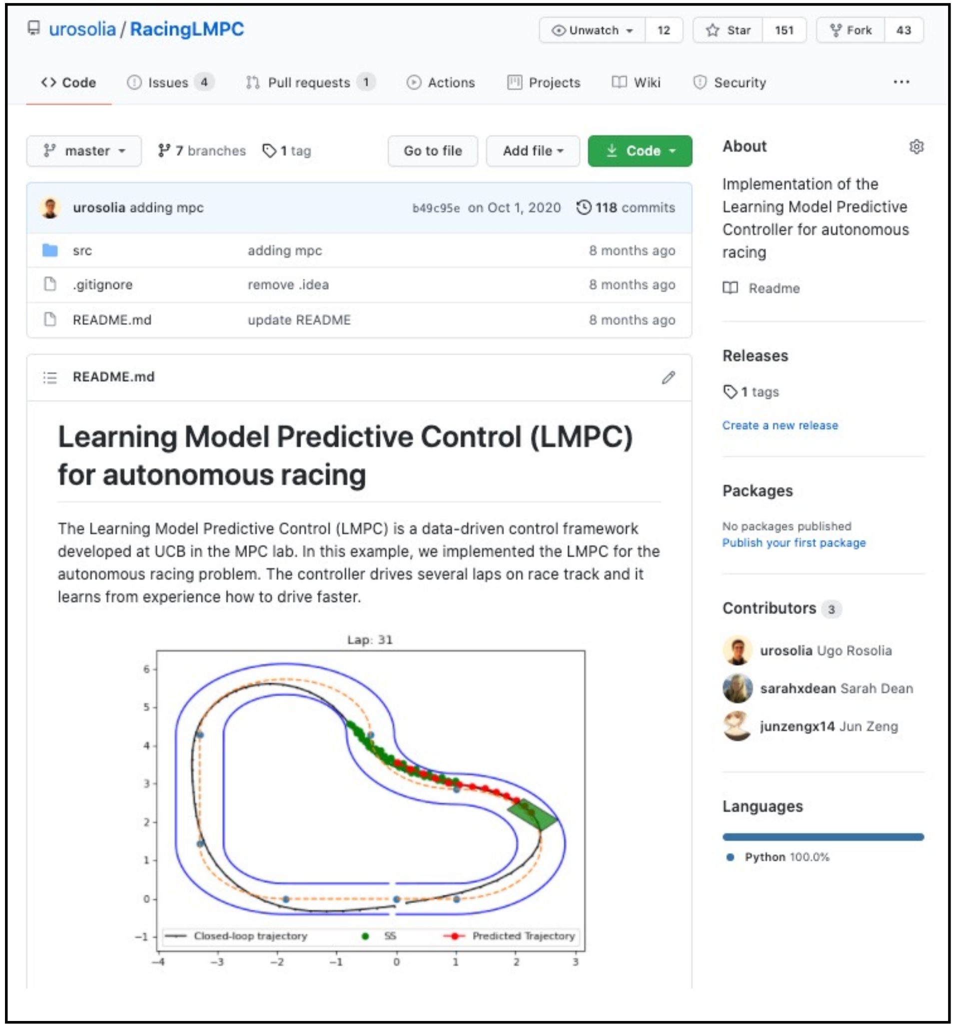


- Optimize over strategies, not trajectories



Questions?

Code available online



Course material online