Safe model-based learning for robot control

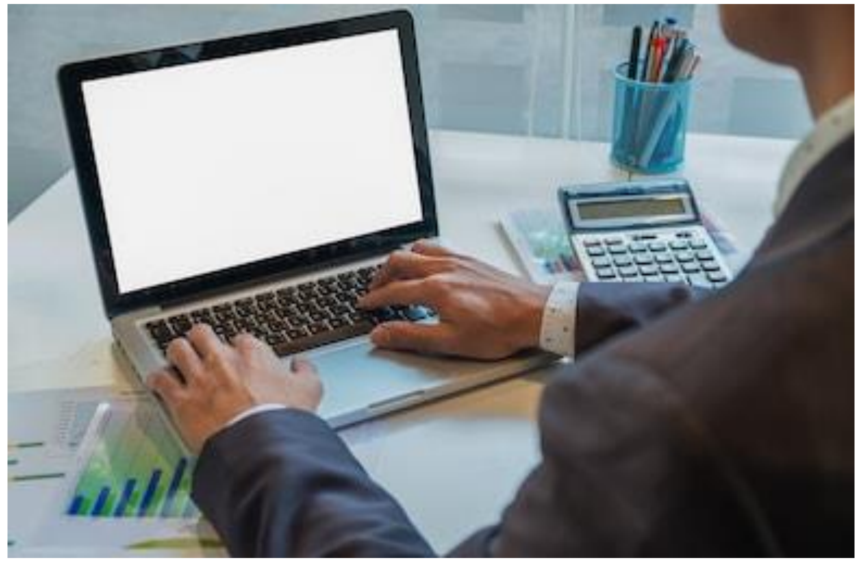
Breaking your robot is only fun in simulation

Felix Berkenkamp, Andreas Krause, Angela P. Schoellig

@LCCC Workshop on Learning and Adaptation for Sensorimotor Control – Lund University

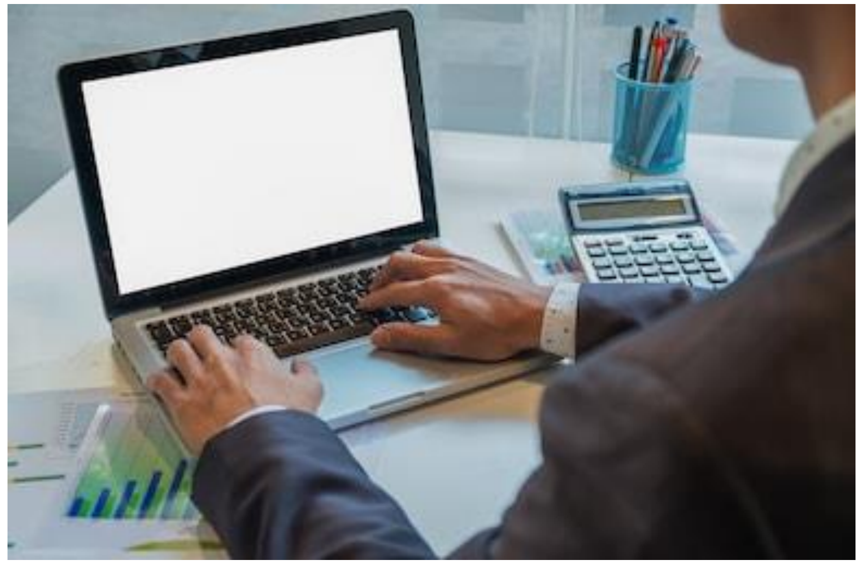
October 2018

The Promise of Robotics = Physical Interaction



Virtual world of data & information.

The Promise of Robotics = Physical Interaction



Virtual world of data & information.

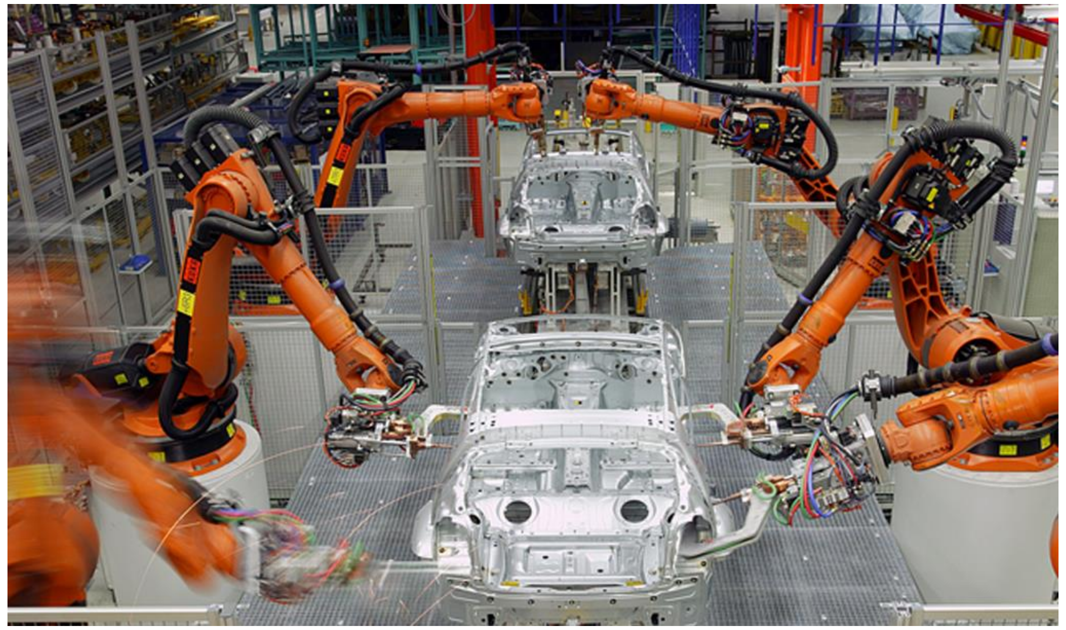
Virtual world

Real world

Exponential increase in complexity!

The Real World Is Complex | Robots Today… and Tomorrow

Dedicated Environments



Manually programmed.
Based on a-priori knowledge.

Robots are limited by our under-standing of the system/environment.

Human-centered Environments




Unknown, unpredictable and changing
Need safe and high-performance behavior

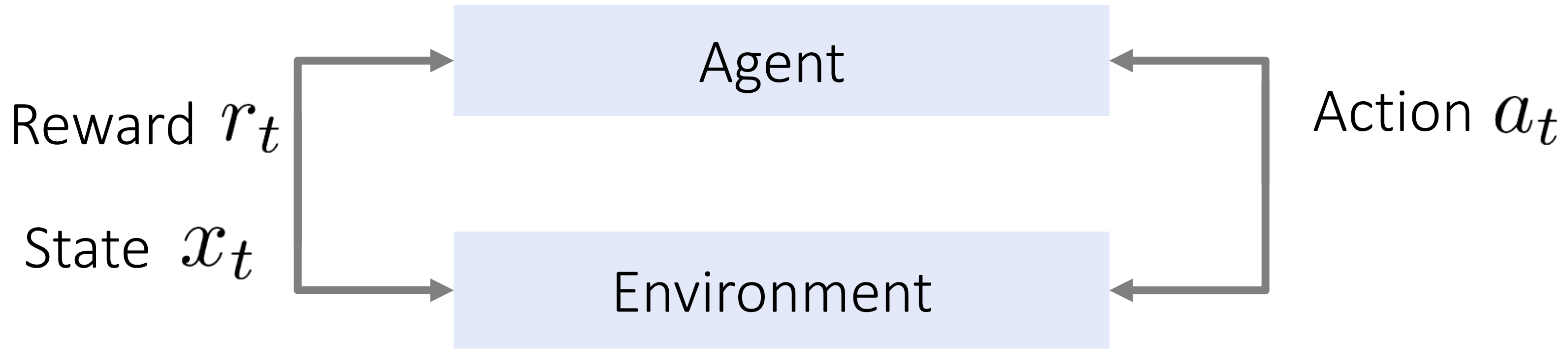
Robots must safely **learn and adapt**

Characteristics of Robot Learning

Robots are **feedback** systems

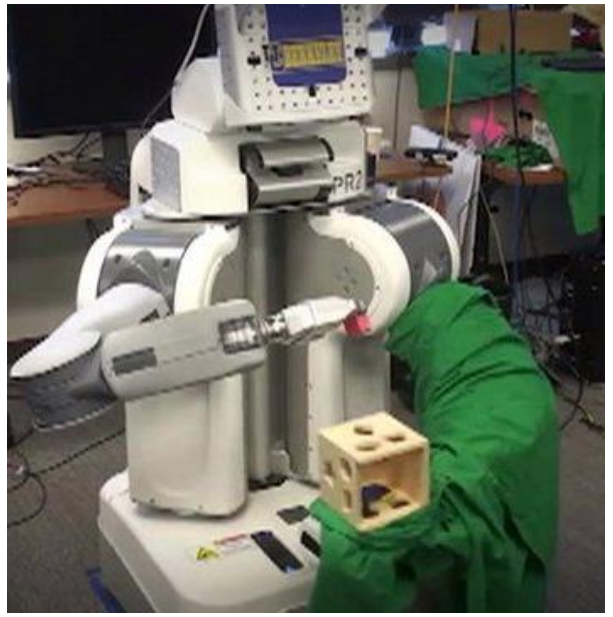
Strict safety requirements

Resource constraints (data, payload, communication)



Reinforcement Learning: An Introduction
R. Sutton, A.G. Barto, 1998

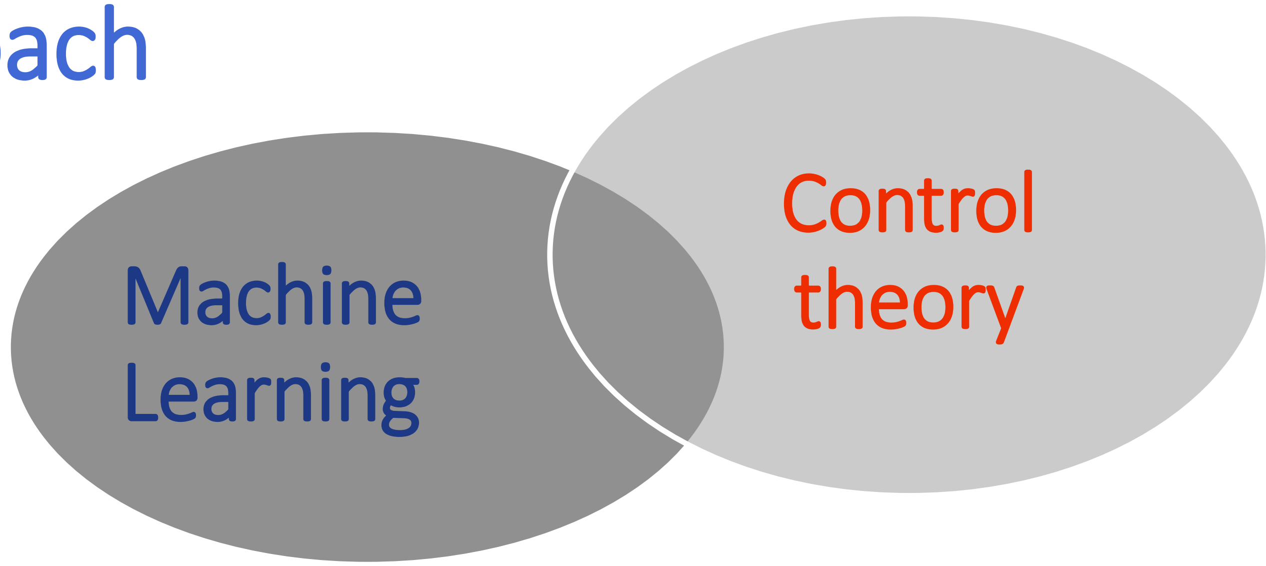
Results to date have been limited to learning **single tasks**, and demonstrated in **simulation or lab settings**.



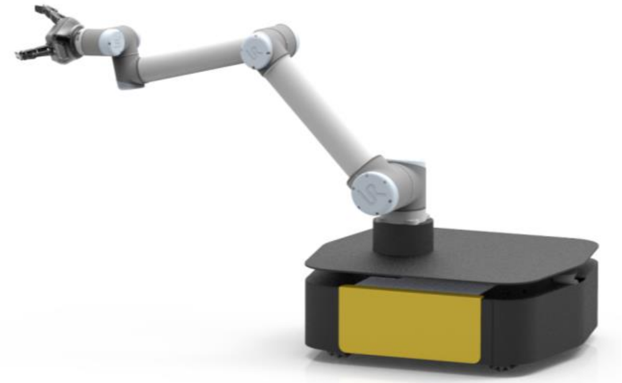
NEXT CHALLENGE: realistic application scenarios
— safety, data efficiency, online learning —

Work at the Dynamic Systems Lab (Prof. Schoellig)

Approach



= science of feedback
(stability, performance, robustness)

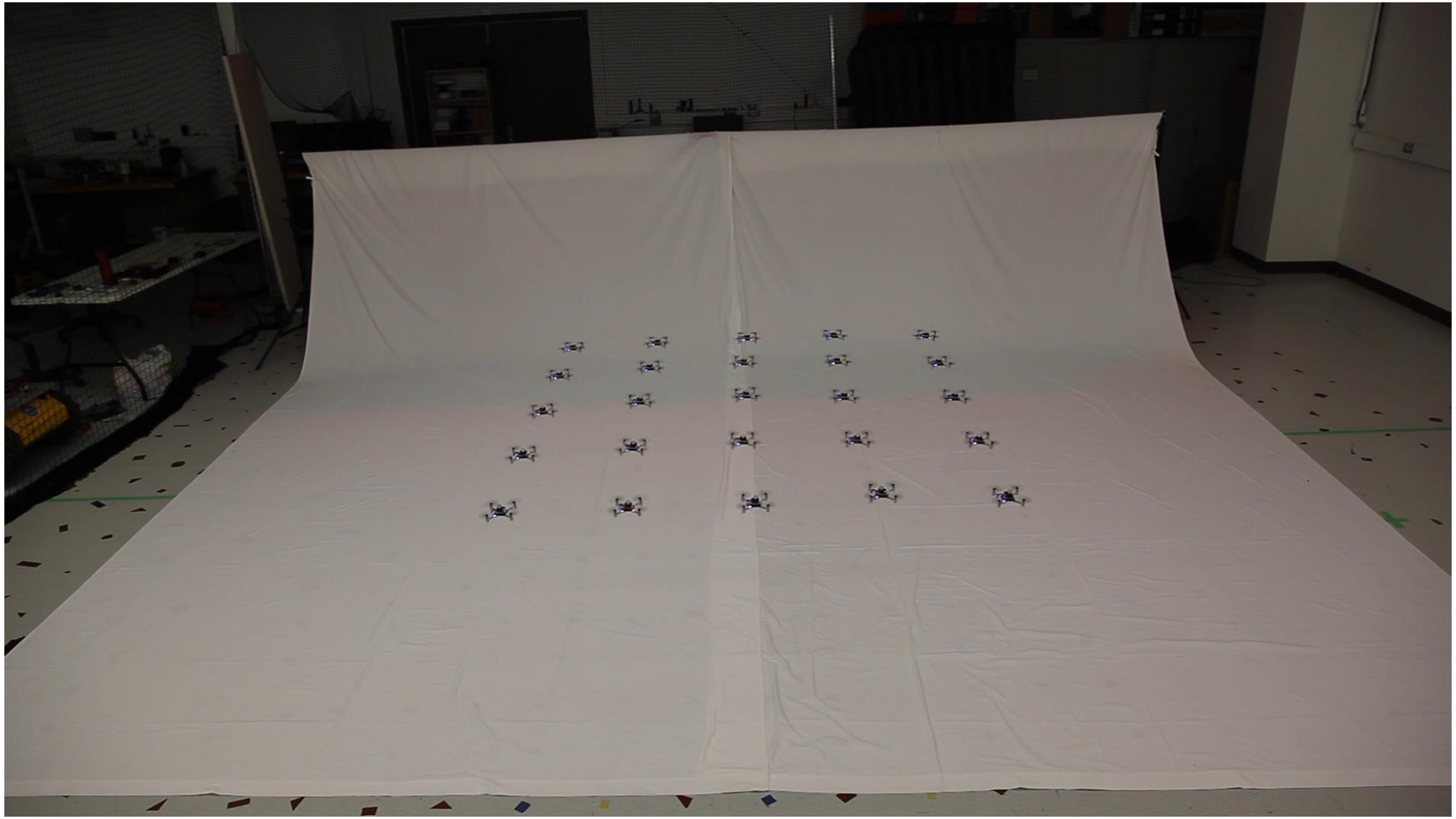


Research Characteristics

Algorithms that run on real robots.

- Data efficiency
- Online adaptation and learning
- **Safety guarantees** during learning **in a closed-loop system**

Performance and Safety: Fast Swarm Flight



Safety: Off-Road Driving

Prerequisites for safe reinforcement learning

Understand model and learning dynamics

Define safety, analyze a model for safety

Algorithm to safely acquire data

Safe Model-based Reinforcement Learning

Overview

Understand model and learning dynamics

Define safety, analyze a model for safety

Algorithm to safely acquire data

Safe Model-based Reinforcement Learning

Learning a model

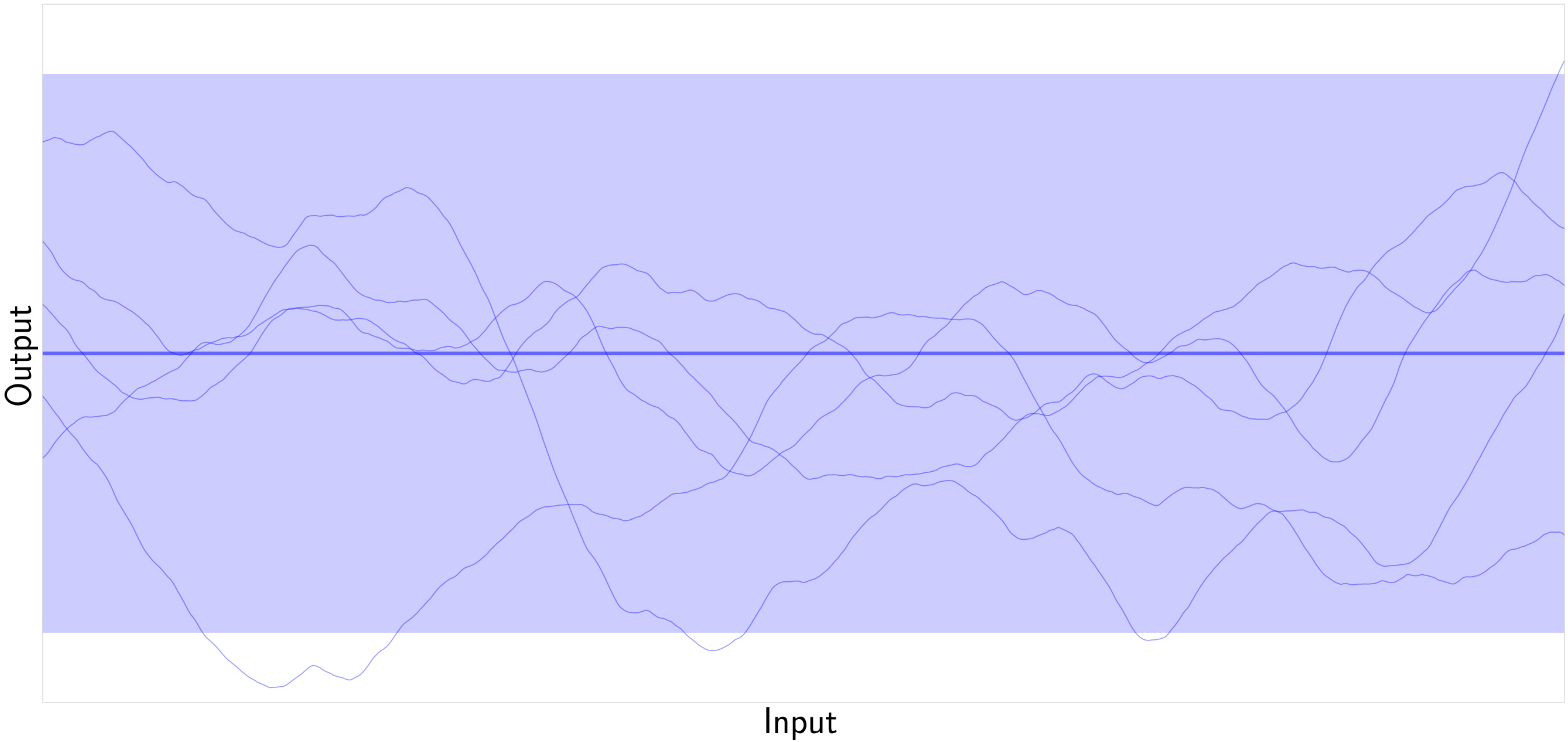
Dynamics

$$x_{t+1} = \underbrace{f(x_t, u_t)}_{a\ priori\ \text{model}} + \underbrace{g(x_t, u_t)}_{\text{unknown model}}$$

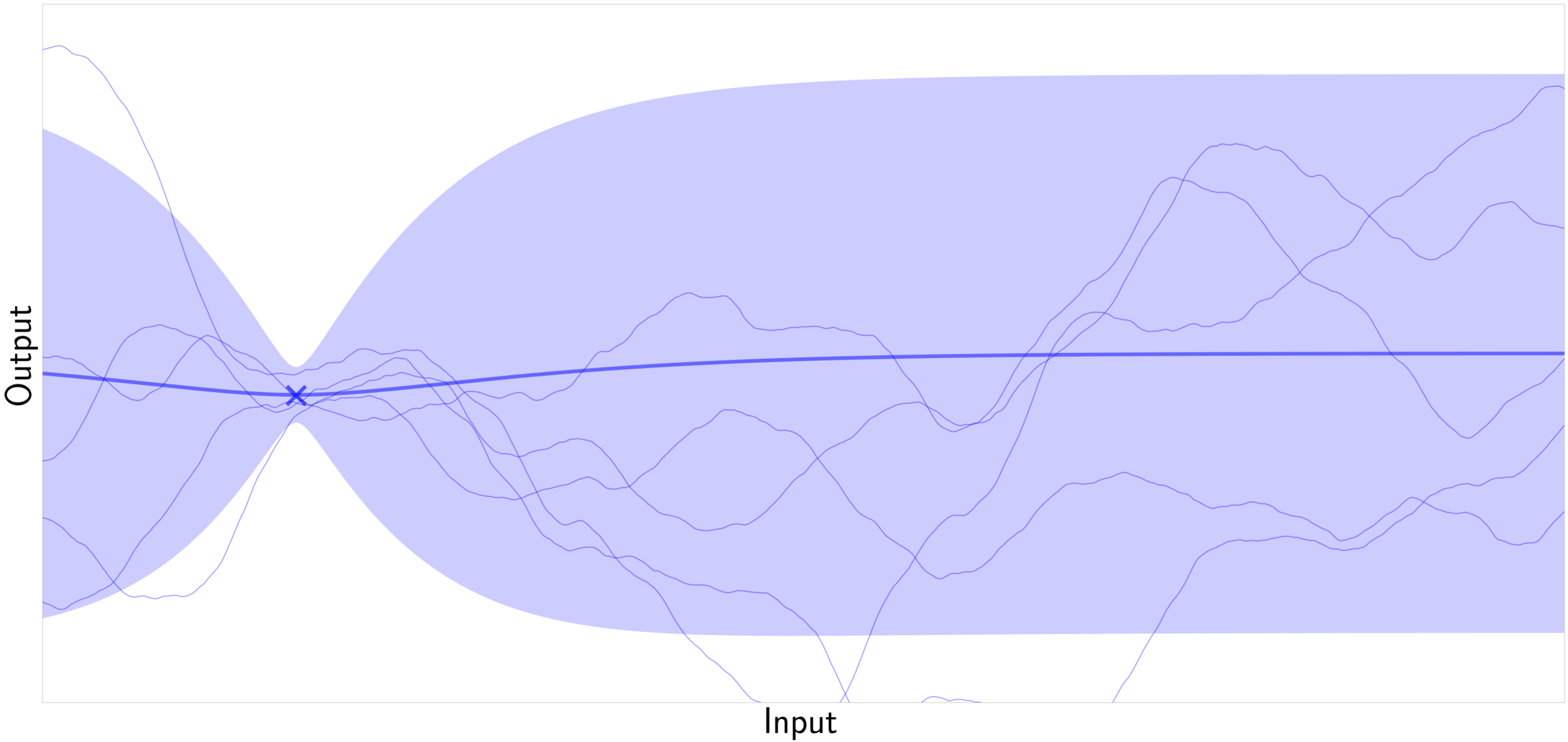
Model error must decrease with measurements

Need to quantify model error

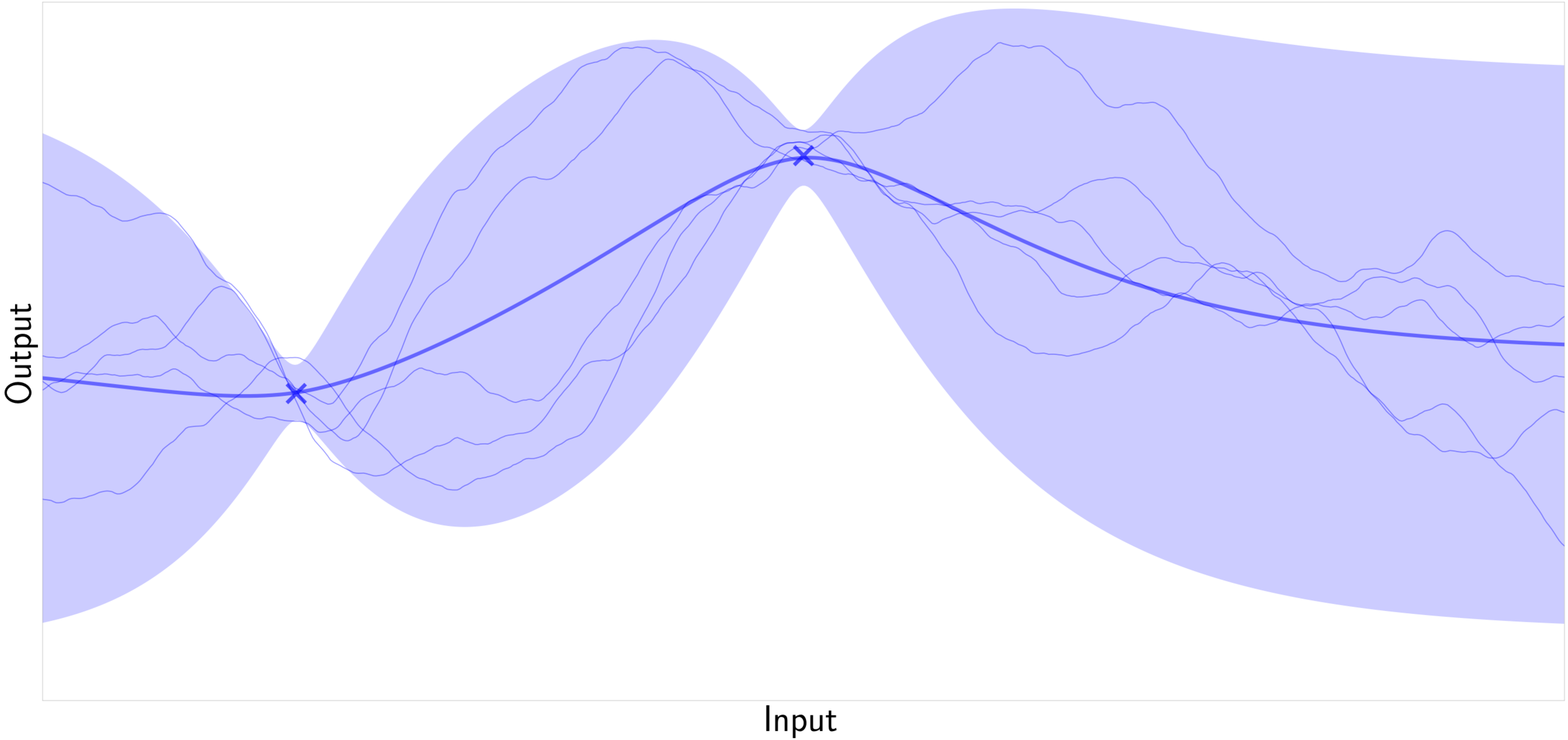
Gaussian process



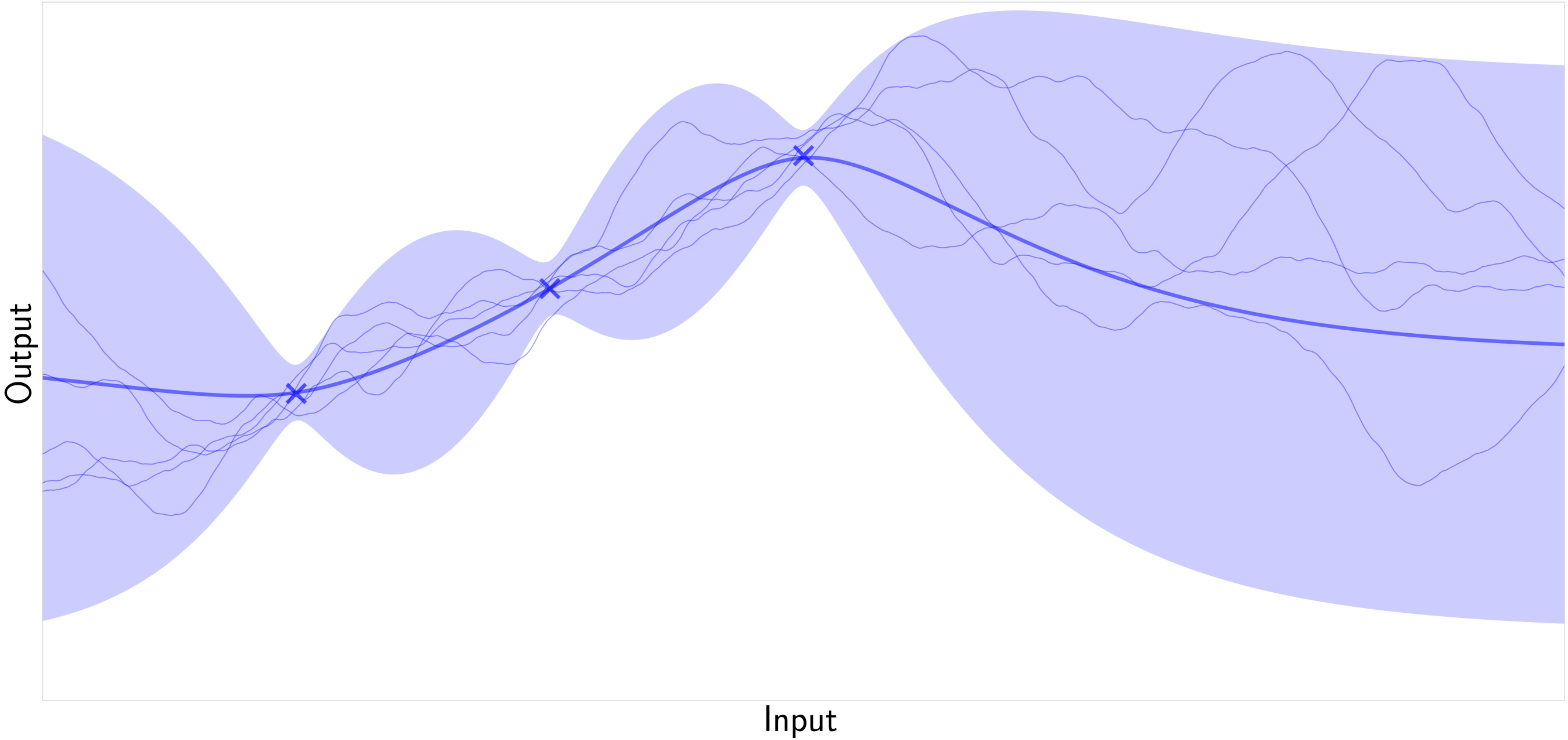
Gaussian process



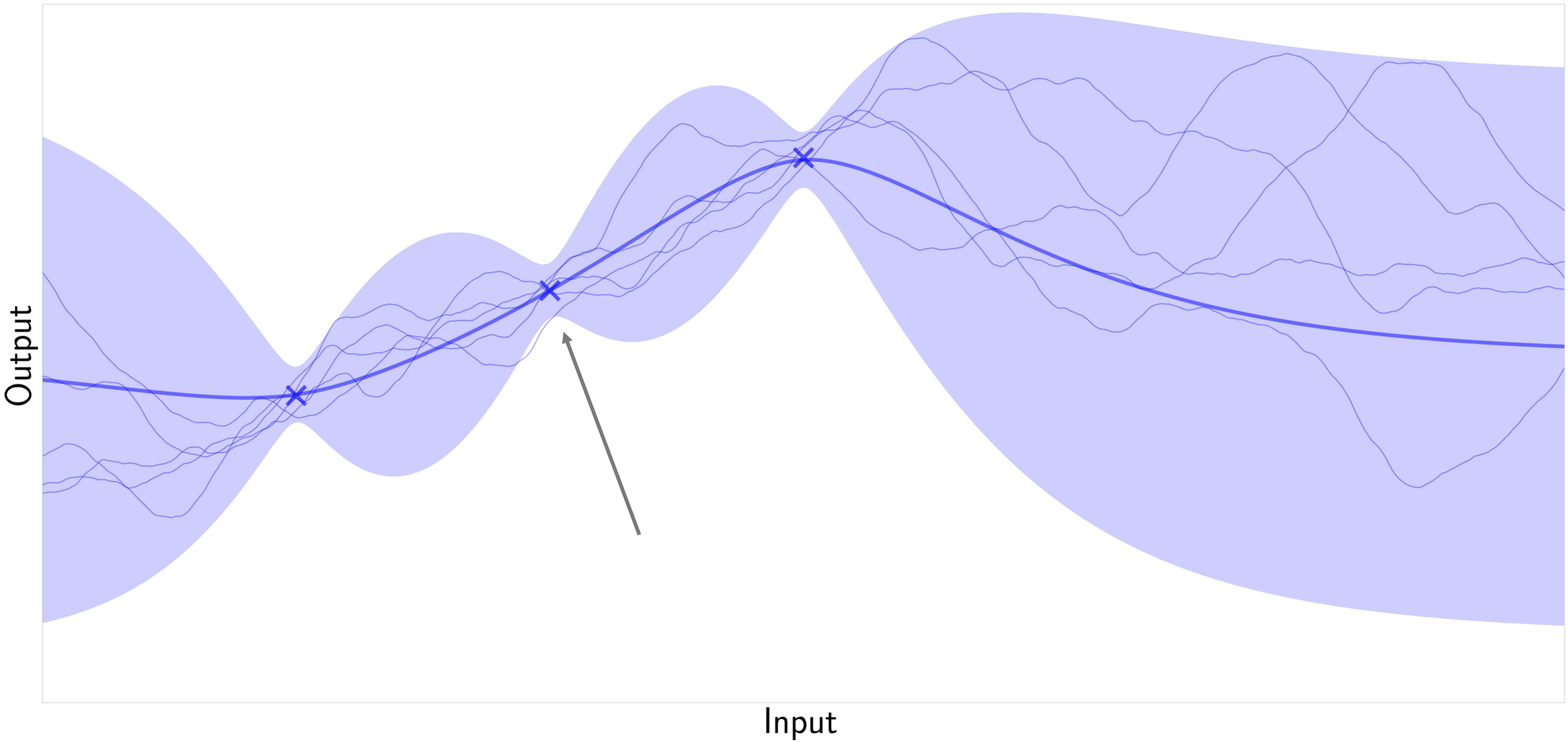
Gaussian process



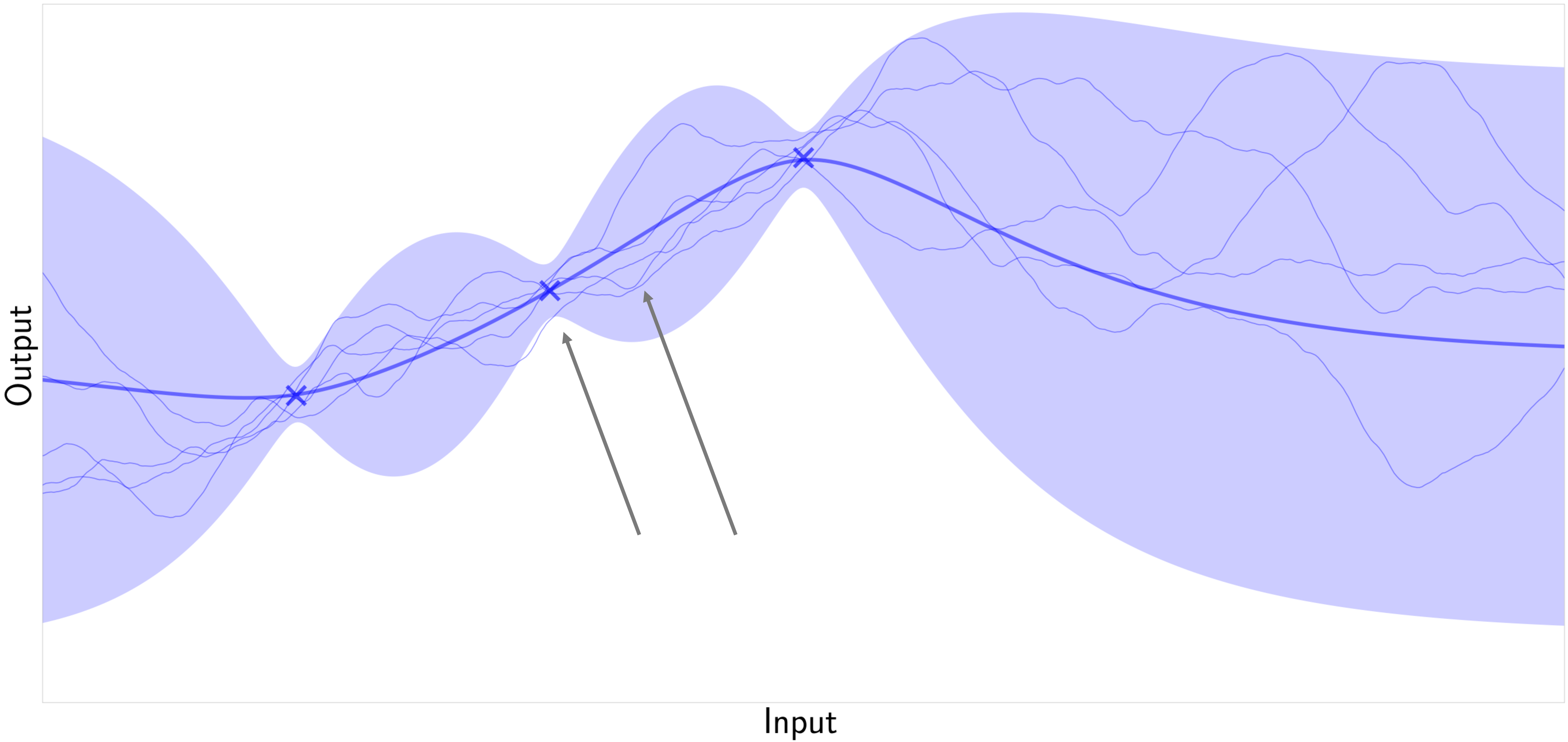
Gaussian process



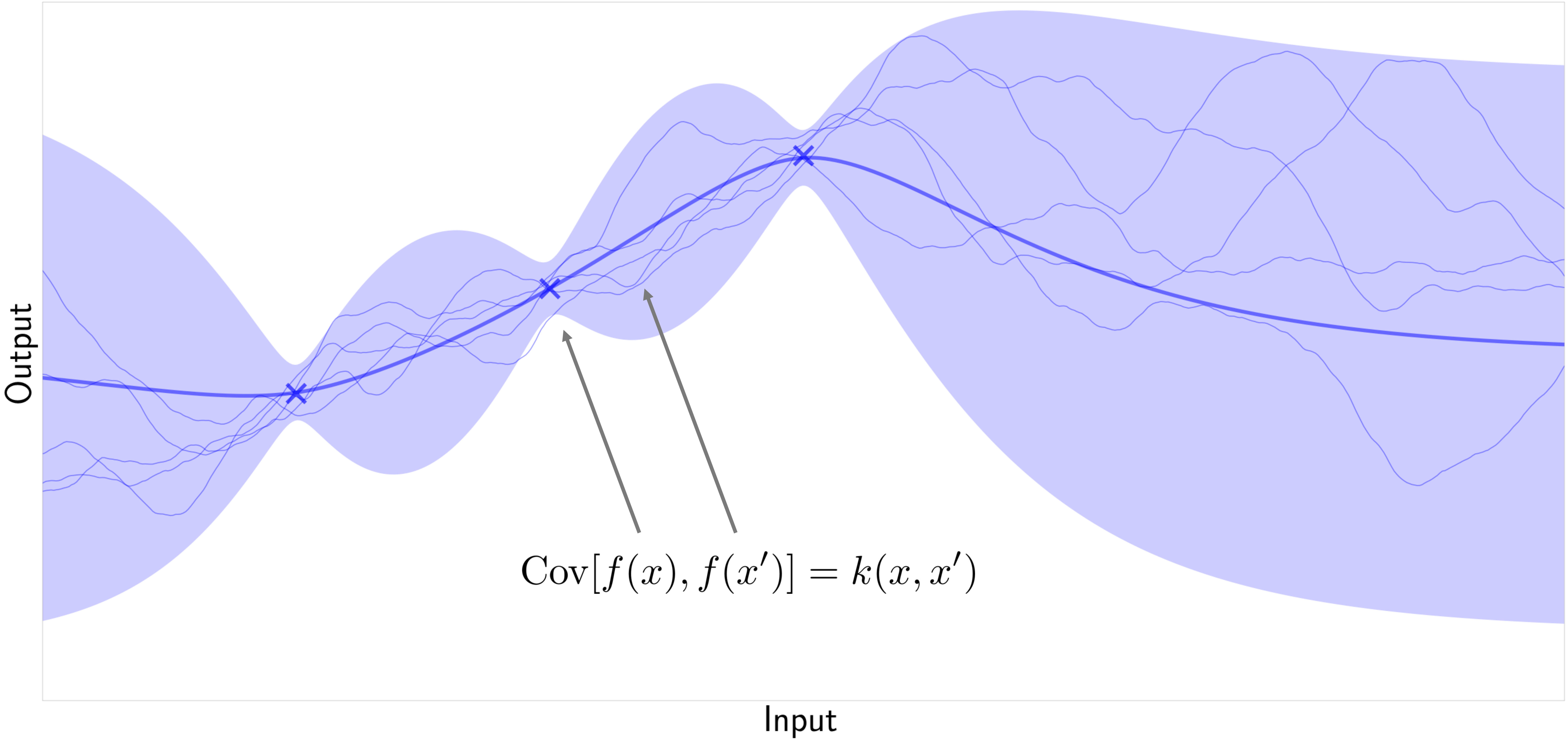
Gaussian process



Gaussian process



Gaussian process

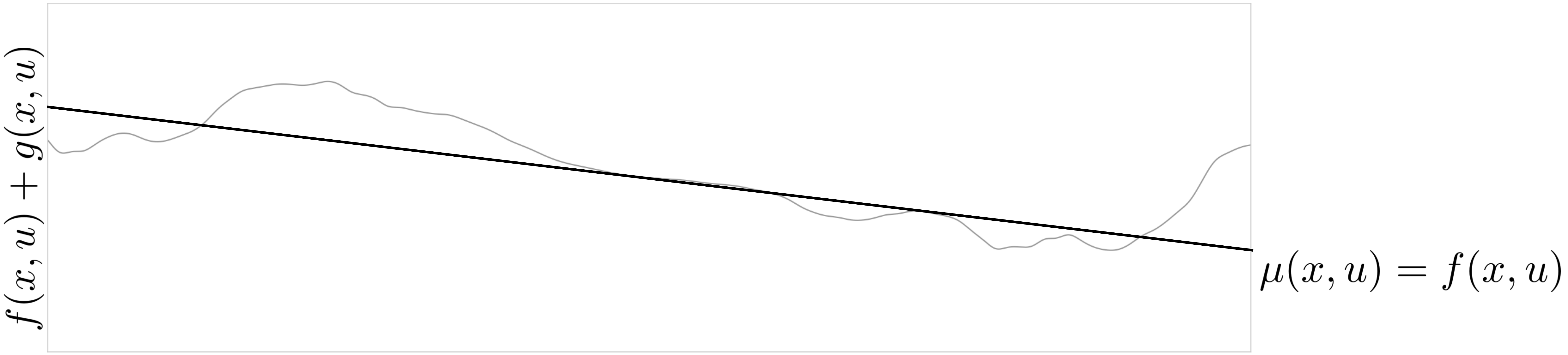


$$\mathrm{Cov}[f(x), f(x')] = k(x, x')$$

A Bayesian dynamics model

Dynamics

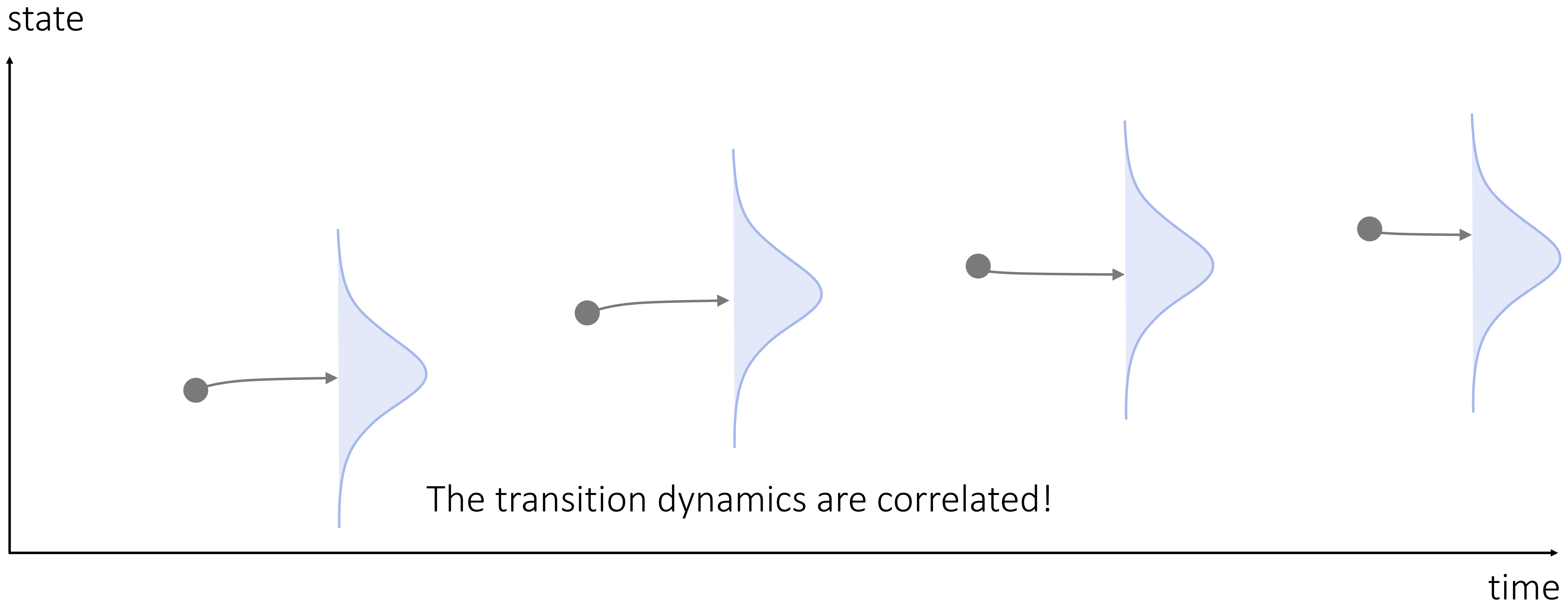
$$x_{t+1} = \underbrace{f(x_t, u_t)}_{a\ priori\ \text{model}} + \underbrace{g(x_t, u_t)}_{\text{unknown model}}$$



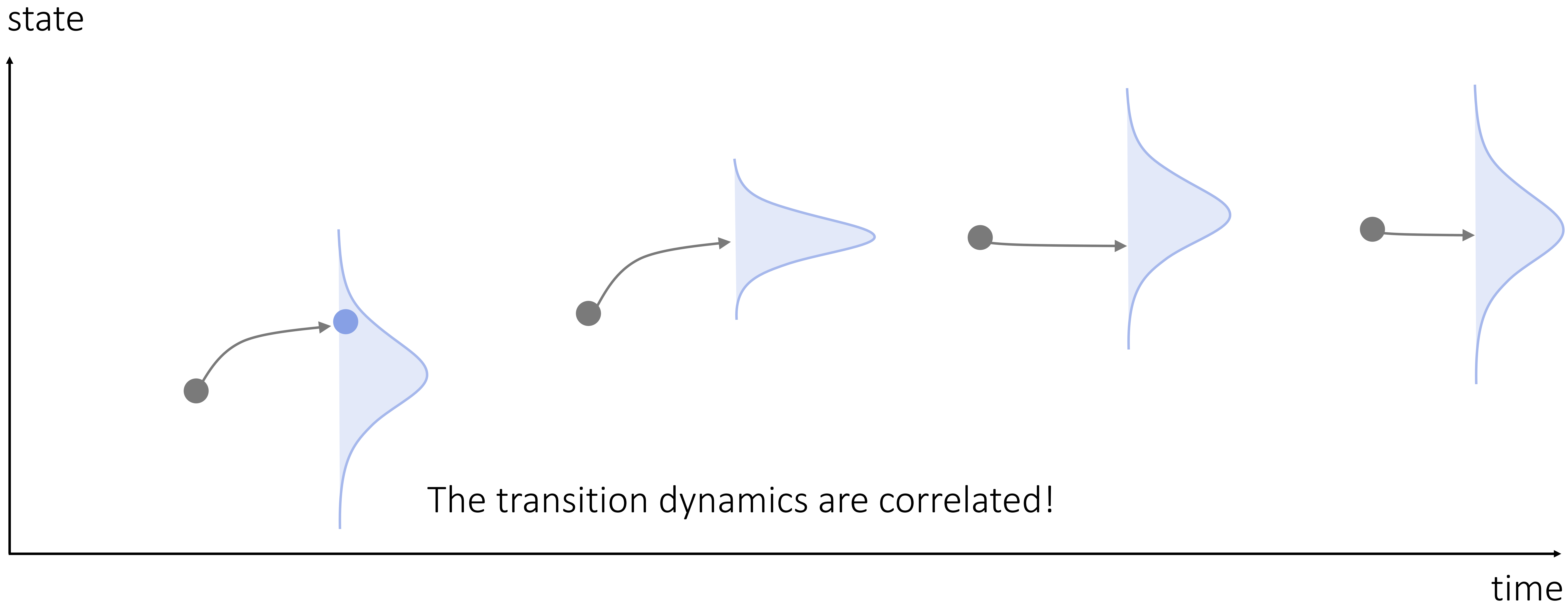
Online Learning of Linearly Parameterized Control Problems
Y. Abbasi-Yadkori, PhD thesis 2012

On Kernelized Multi-armed Bandits
S.R. Chowdhury, A. Gopalan, ICML 2017

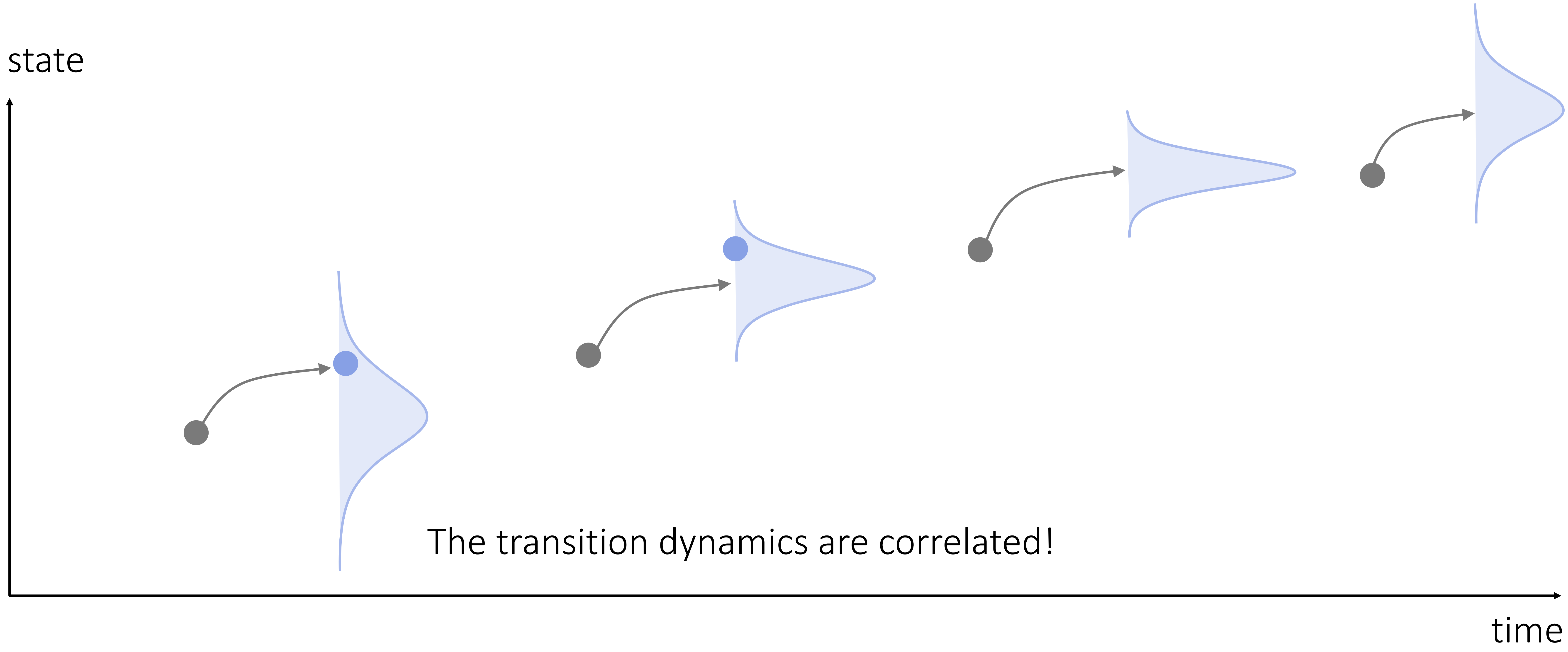
Samples from the Gaussian process prior



Samples from the Gaussian process prior



Samples from the Gaussian process prior



Overview

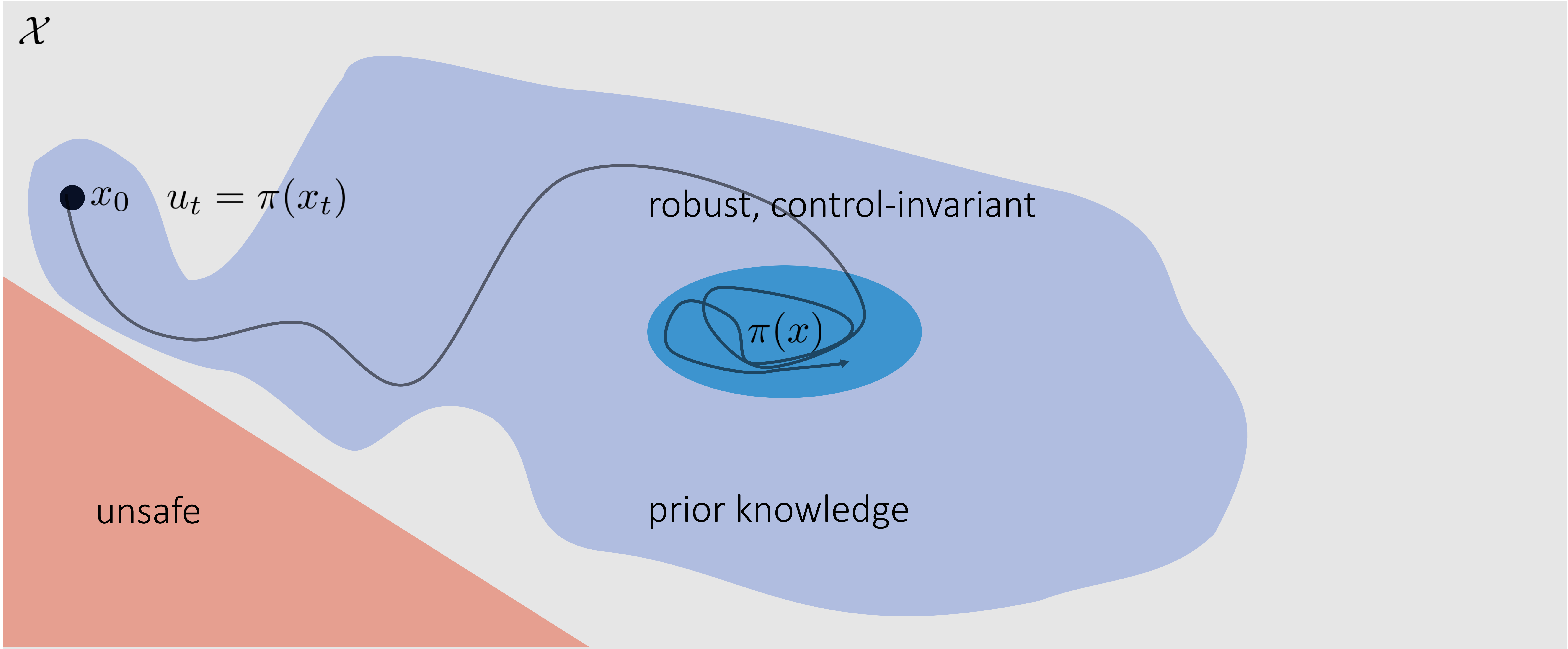
Understand model and learning dynamics

Define safety, analyze a model for safety

Algorithm to safely acquire data

Safe Model-based Reinforcement Learning

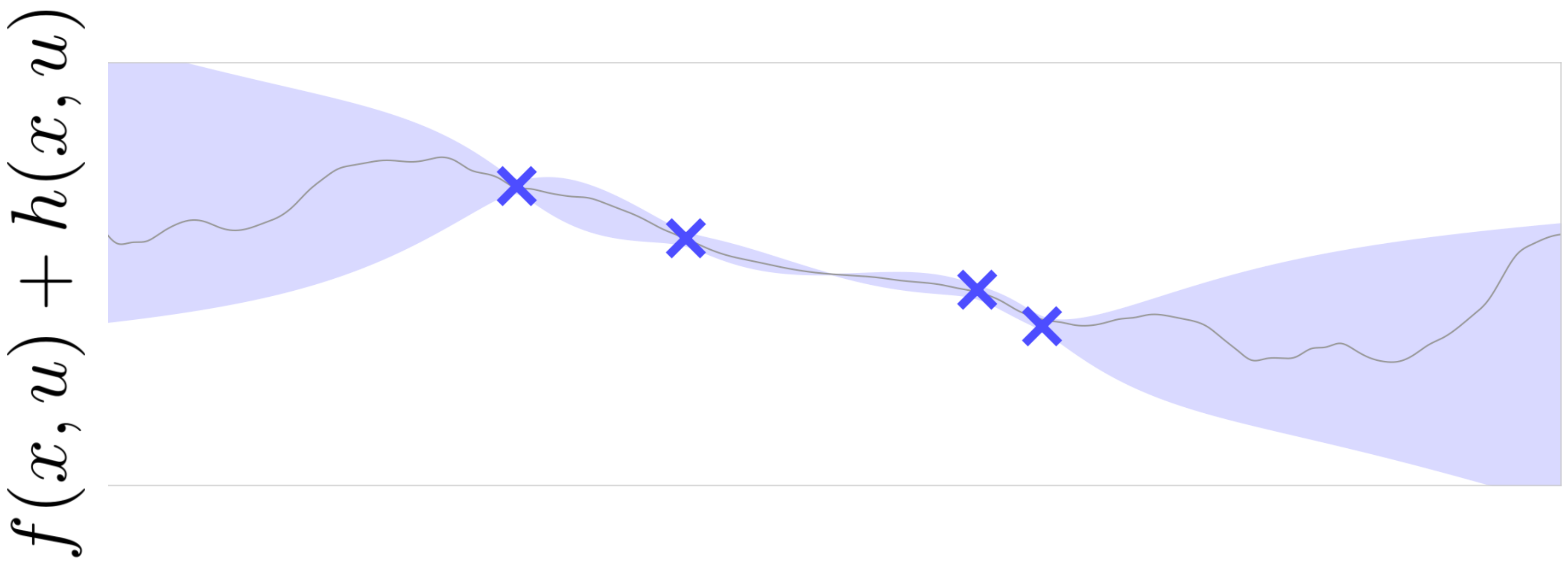
Safety definition



Safety for learned models

Dynamics

$$x_{t+1} = \underbrace{f(x_t, u_t)}_{a\ priori\ \text{model}} + \underbrace{h(x_t, u_t)}_{\text{unknown model}}$$



$+$

Policy

$$u_t = \pi(x_t)$$

Stability?

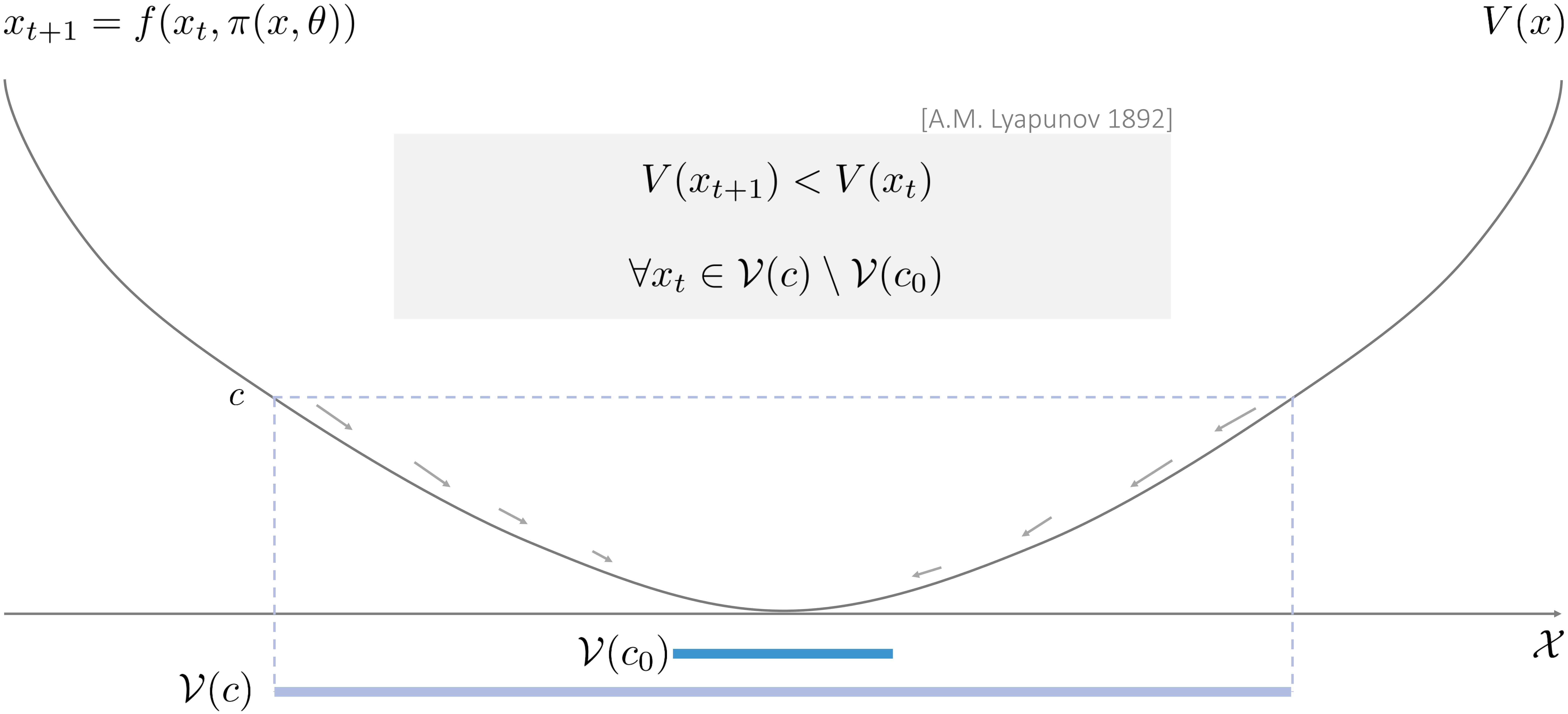
Lyapunov functions

$$x_{t+1} = f(x_t, \pi(x, \theta))$$

[A.M. Lyapunov 1892]

$$V(x_{t+1}) < V(x_t)$$

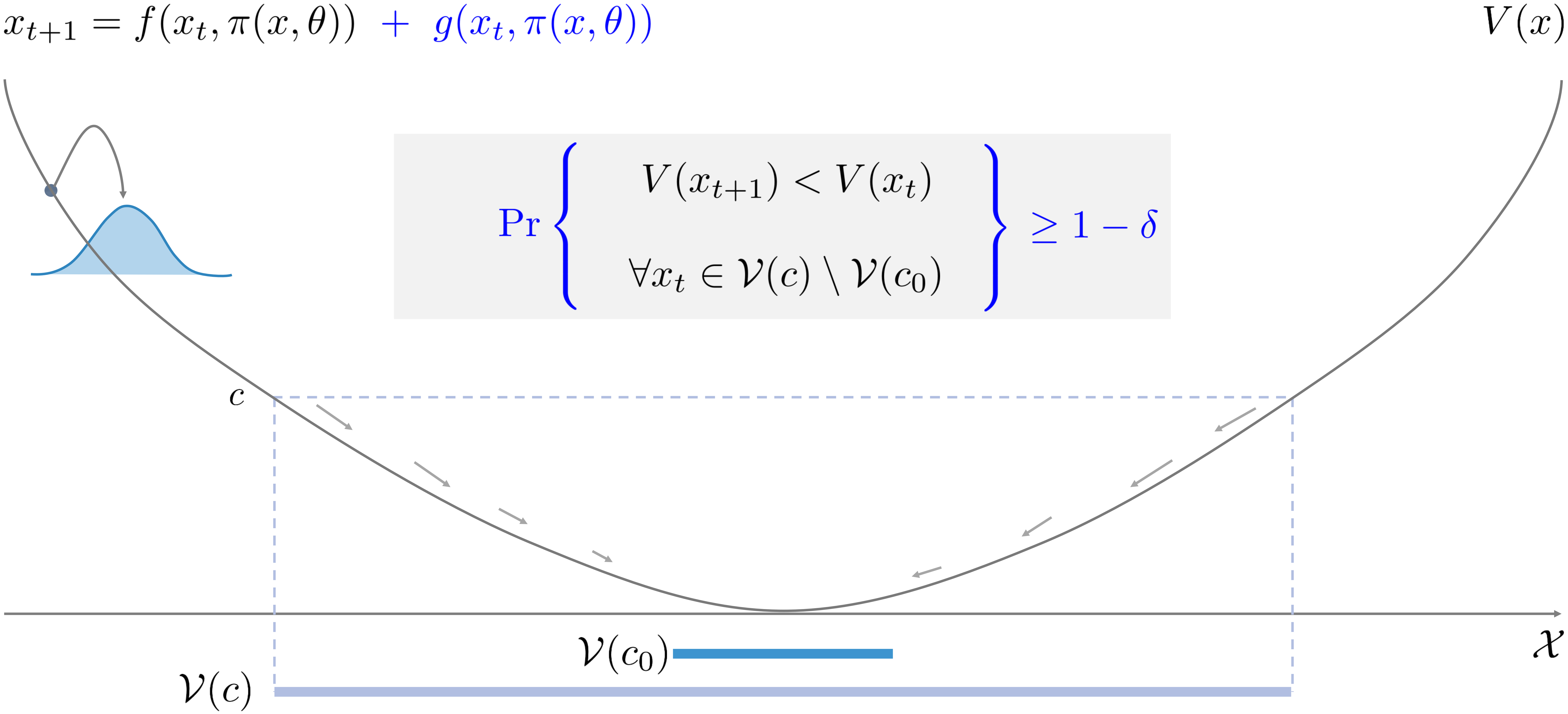
$$\forall x_t \in \mathcal{V}(c) \setminus \mathcal{V}(c_0)$$



Lyapunov functions

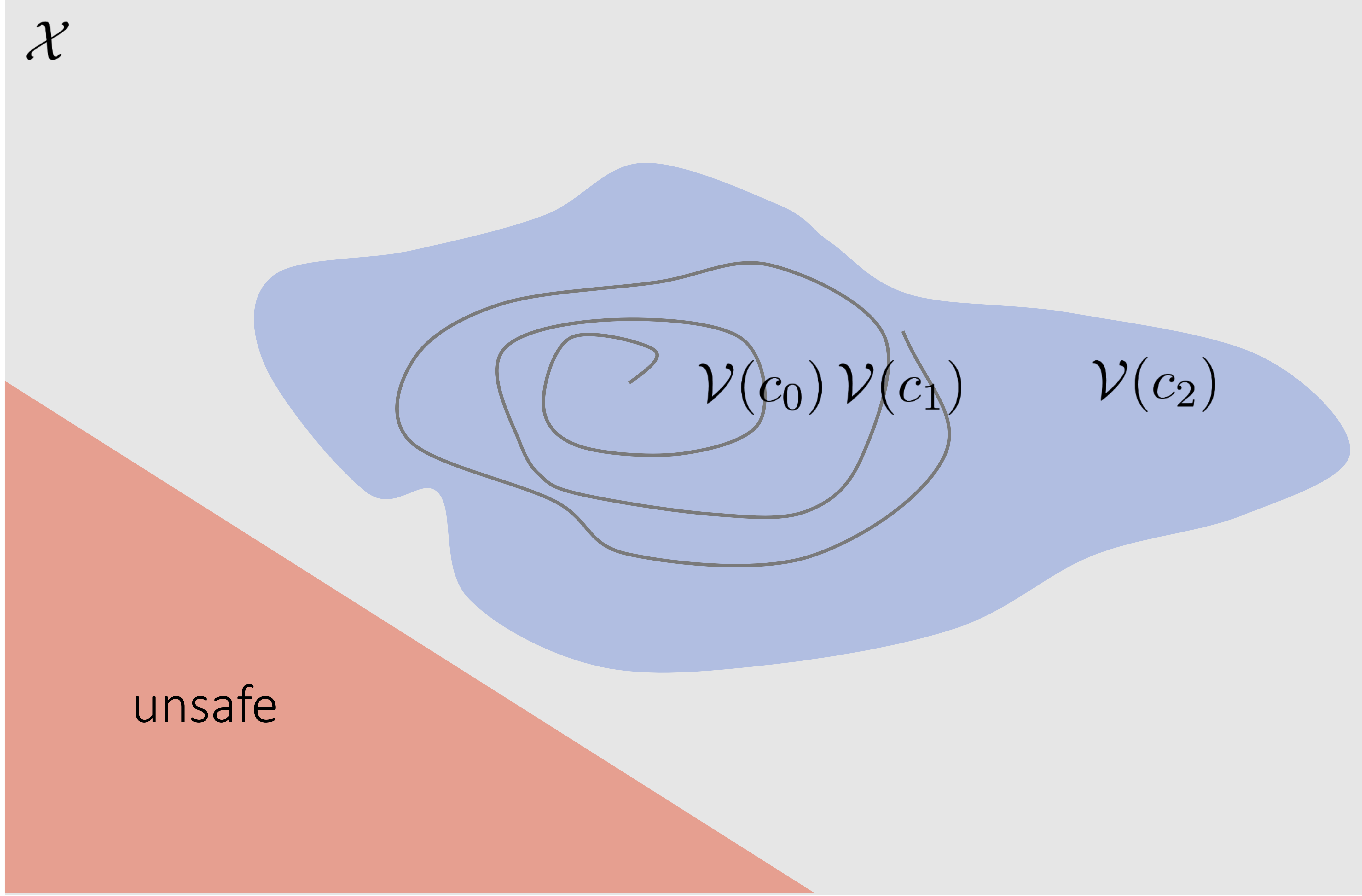
$$x_{t+1} = f(x_t, \pi(x,\theta)) \; + \; g(x_t, \pi(x,\theta))$$

$$\Pr \left\{ \begin{array}{c} V(x_{t+1}) < V(x_t) \\ \forall x_t \in \mathcal{V}(c) \setminus \mathcal{V}(c_0) \end{array} \right\} \geq 1 - \delta$$



Region of attraction

Safe Model-based Reinforcement Learning with Stability Guarantees
F. Berkenkamp, M. Turchetta, A.P. Schoellig, A. Krause, NIPS, 2017

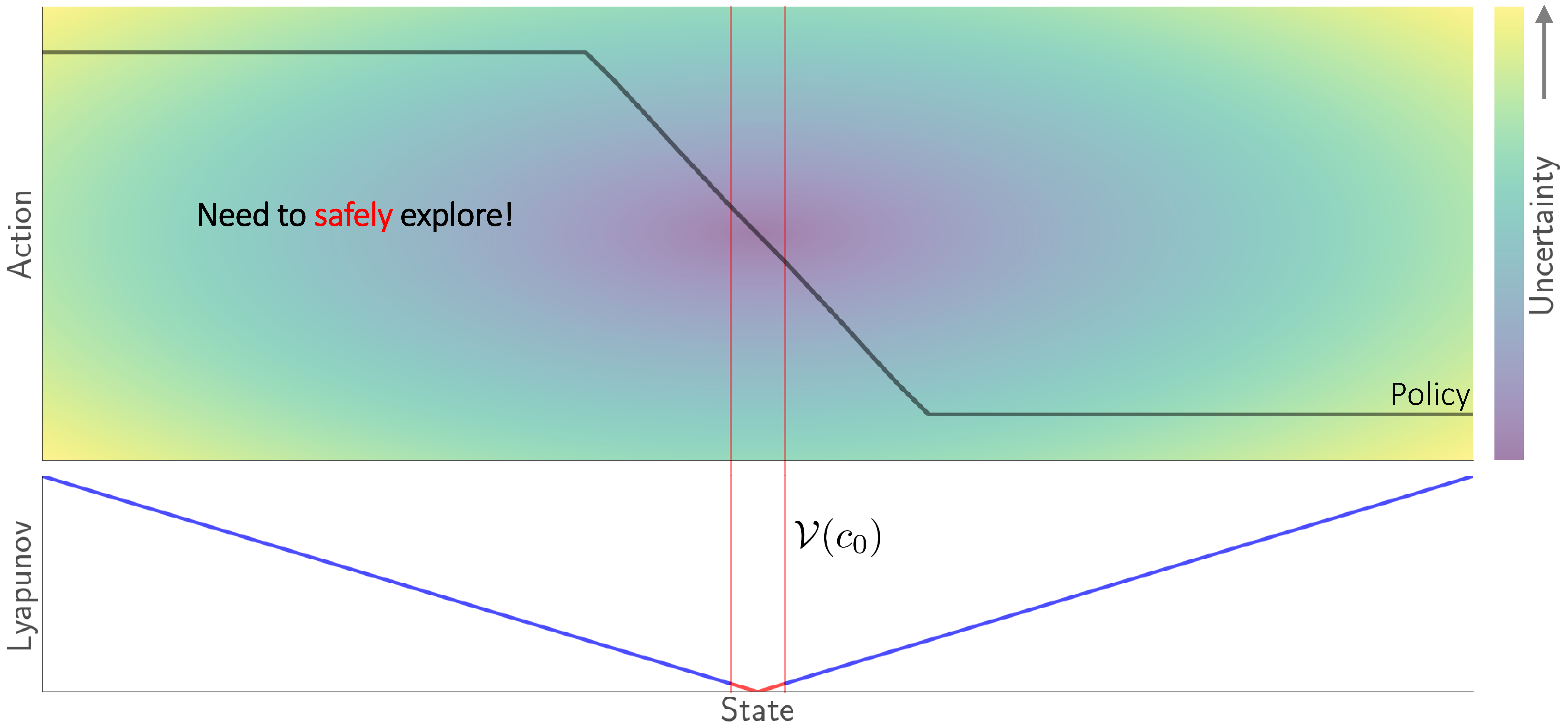


Initial safe policy π

Theorem (informally):

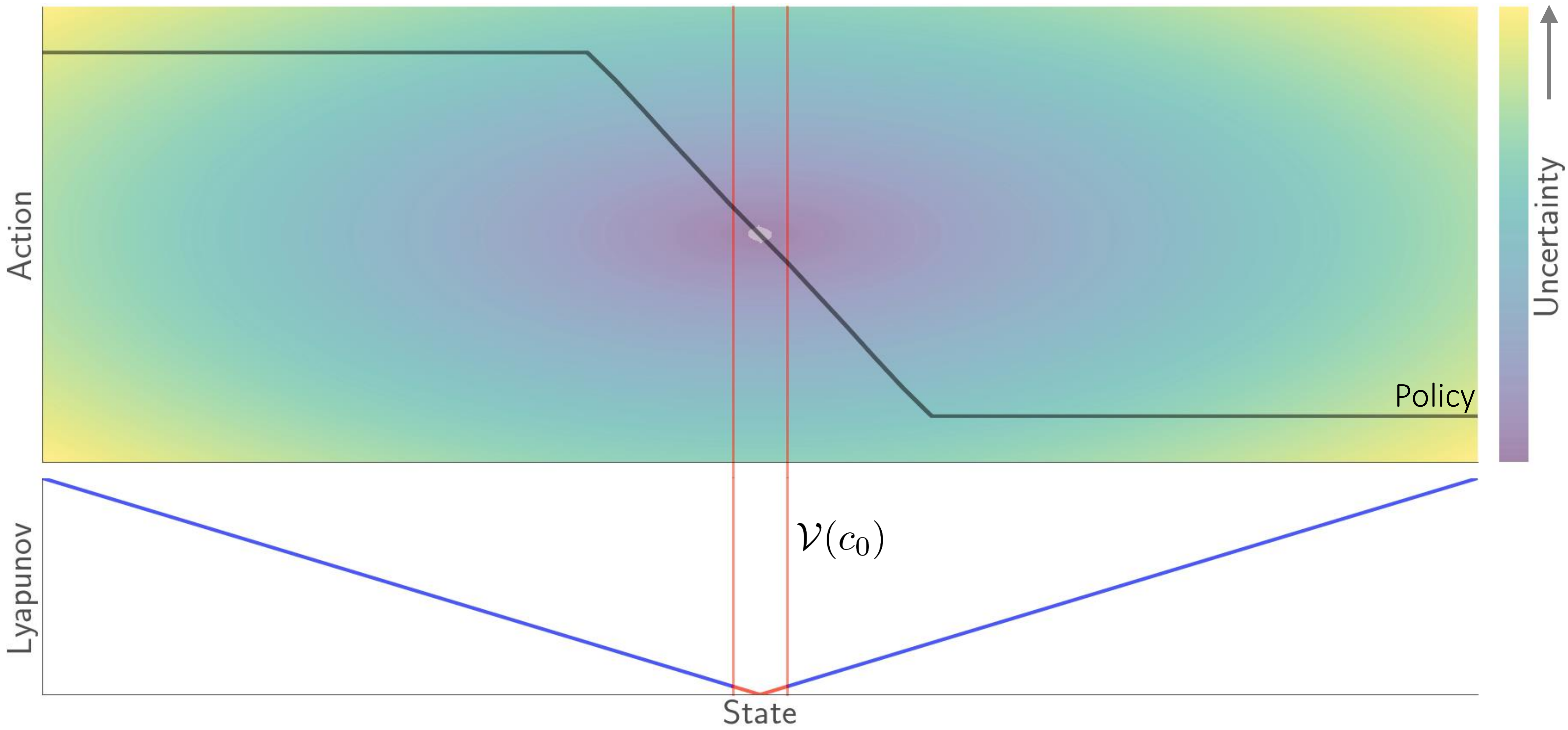
Under suitable conditions can identify (near-)maximal subset of X on which π is stable, while never leaving the safe set

Illustration of safe learning



Safe Model-based Reinforcement Learning with Stability Guarantees
F. Berkenkamp, M. Turchetta, A.P. Schoellig, A. Krause, NIPS, 2017

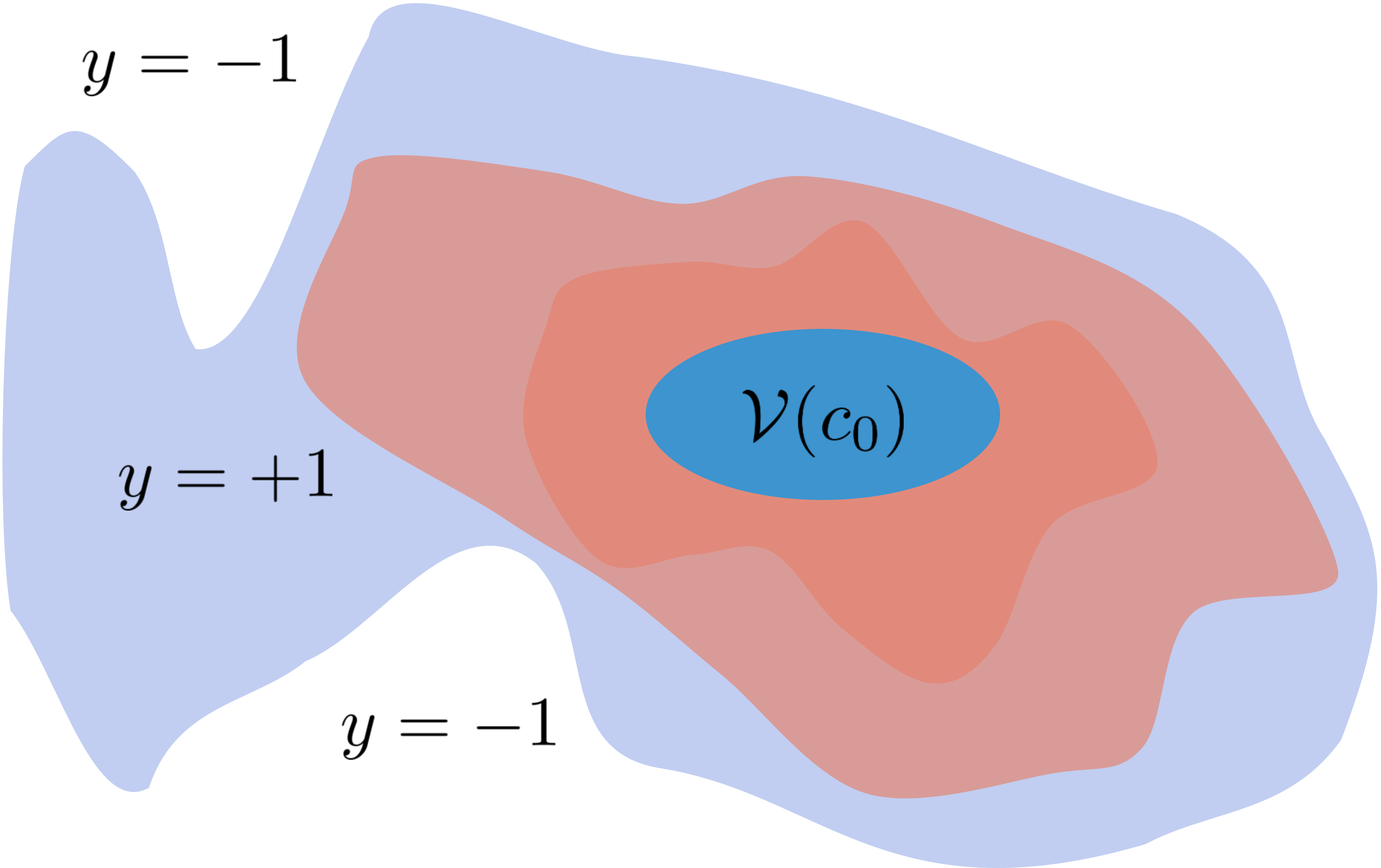
Illustration of safe learning



Safe Model-based Reinforcement Learning with Stability Guarantees
F. Berkenkamp, M. Turchetta, A.P. Schoellig, A. Krause, NIPS, 2017

Lyapunov function

Finding the right Lyapunov function is difficult!



$$V(x) = \phi_\theta(x)^{\mathrm{T}} \phi_\theta(x)$$

Weights - positive-definite
Nonlinearities - trivial nullspace

Decision boundary $V(x) = 1$

$$V(x_{t+1}) < V(x_t)$$

$$\forall x_t \in \mathcal{V}(c) \setminus \mathcal{V}(c_0)$$

The Lyapunov Neural Network: Adaptive Stability Certification for Safe Learning of Dynamic Systems
S.M. Richards, F. Berkenkamp, A. Krause, CoRL 2018

Overview

Understand model and learning dynamics

Define safety, analyze a model for safety

Algorithm to safely acquire data

Safe Model-based Reinforcement Learning

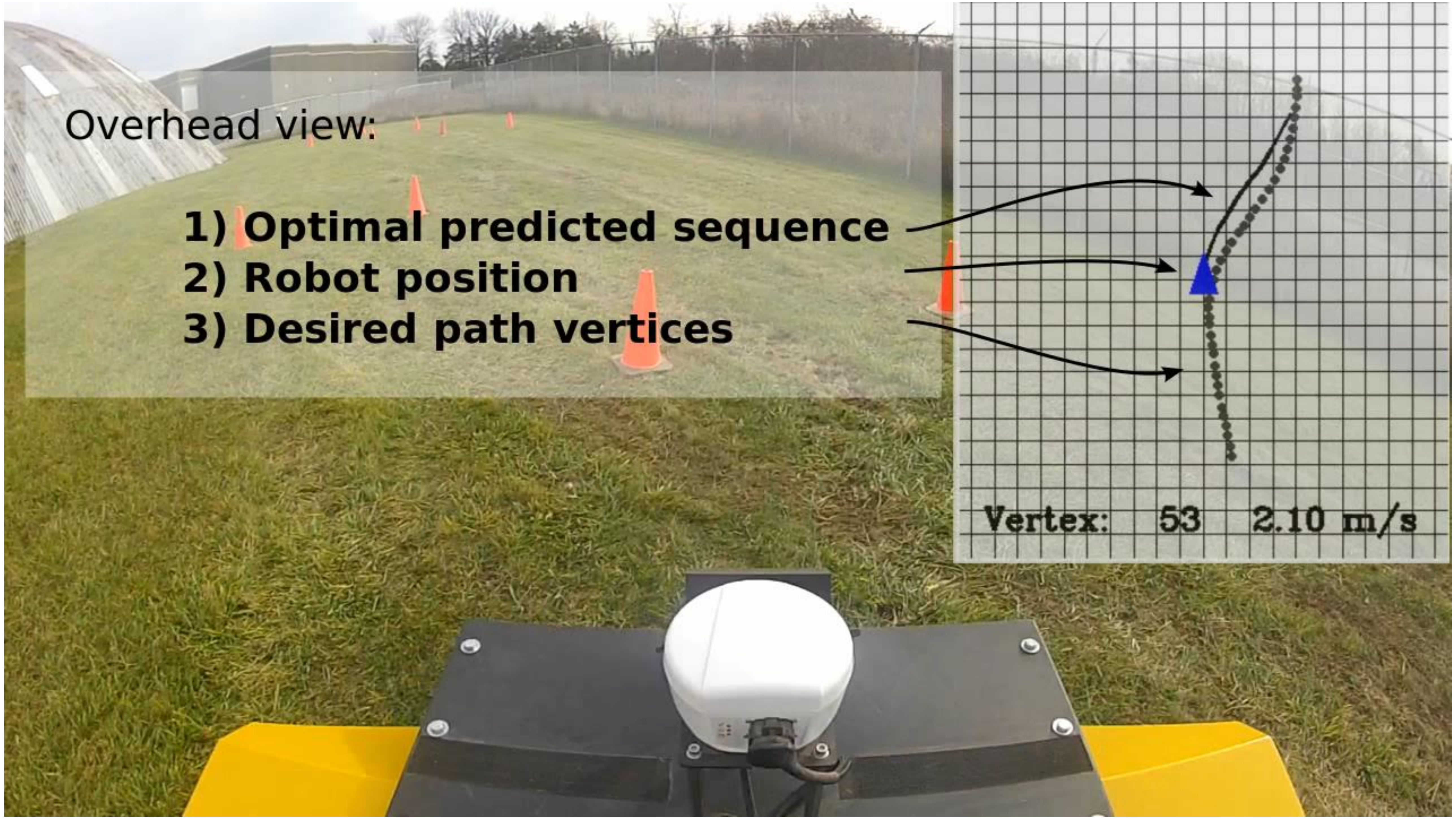
Model predictive control

$$
\begin{aligned}
\underset{\{u_0, u_1, \dots, u_{N-1}\}}{\text{minimize}} \quad & \sum_{k=0}^{N-1} J(x_k, u_k) + J_N(x_N) && \text{mission objective} \\
\text{subject to} \quad & x_0 = \overline{x}_0 && \text{system state} \\
& x_{k+1} = f(x_k, u_k) && \text{system dynamics} \\
& x_k \in \mathcal{X}_k && \text{state constraints} \\
& u_k \in \mathcal{U}_k && \text{input constraints}
\end{aligned}
$$

Makes decisions based on predictions about the future

Includes input / state constraints

Model predictive control on a robot



Robust constrained learning-based NMPC enabling reliable mobile robot path tracking
C.J. Ostafew, A.P. Schoellig, T.D. Barfoot, IJRR, 2016

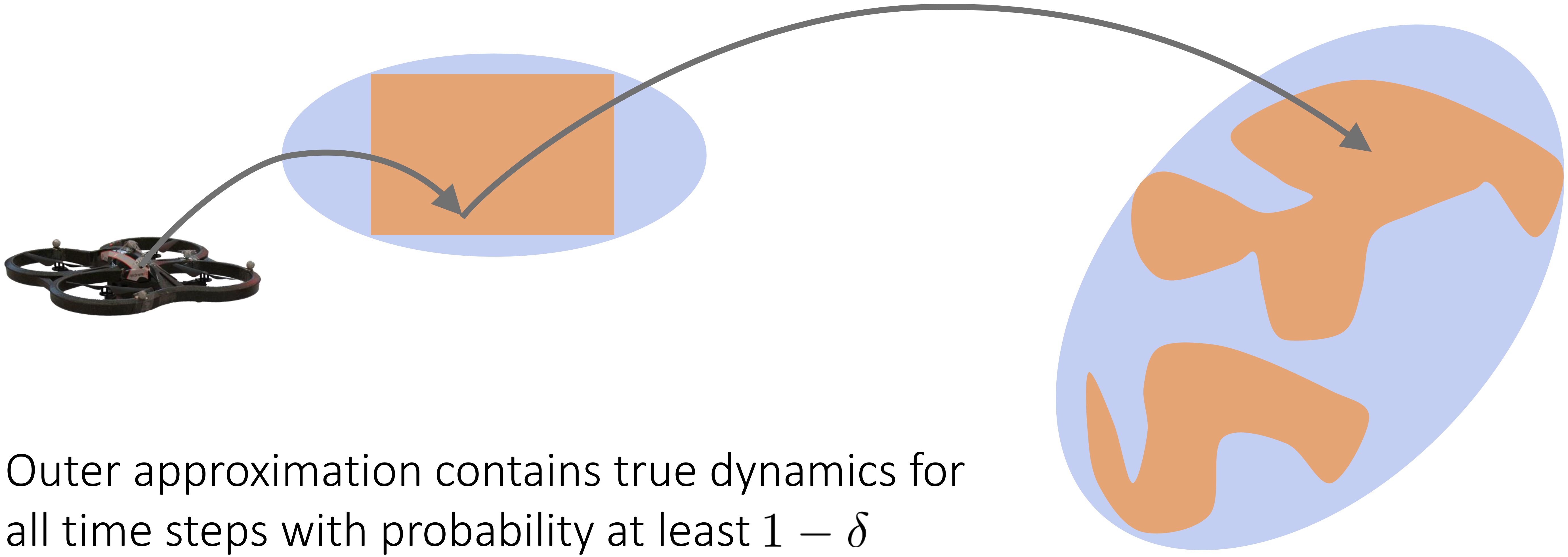
Video at https://youtu.be/3xRNmNv5Efk

Model predictive control

$$
\begin{aligned}
\underset{\{u_0, u_1, \dots, u_{N-1}\}}{\text{minimize}} \quad & \sum_{k=0}^{N-1} J(x_k, u_k) + J_N(x_N) && \text{mission objective} \\
\text{subject to} \quad & x_0 = \overline{x}_0 && \text{system state} \\
& x_{k+1} = f(x_k, u_k) + g(x_k, u_k) && \text{system dynamics} \\
& x_k \in \mathcal{X}_k && \text{state constraints} \\
& u_k \in \mathcal{U}_k && \text{input constraints}
\end{aligned}
$$

Problem: True dynamics $f(x,u) + g(x,u)$ are unknown!

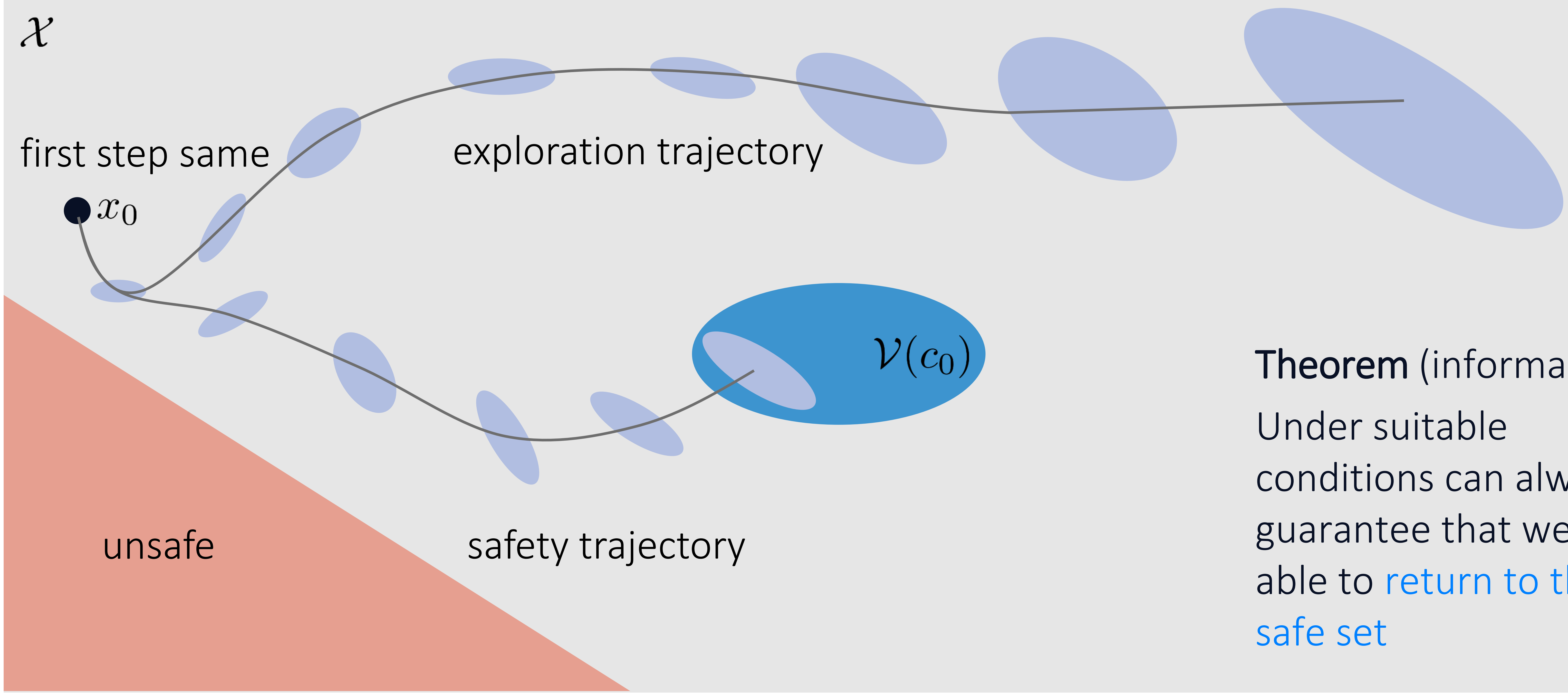
Forward-propagating uncertainty



Outer approximation contains true dynamics for all time steps with probability at least $1 - \delta$

Learning-based Model Predictive Control for Safe Exploration
T. Koller, F. Berkenkamp, M. Turchetta, A. Krause, CDC, 2018

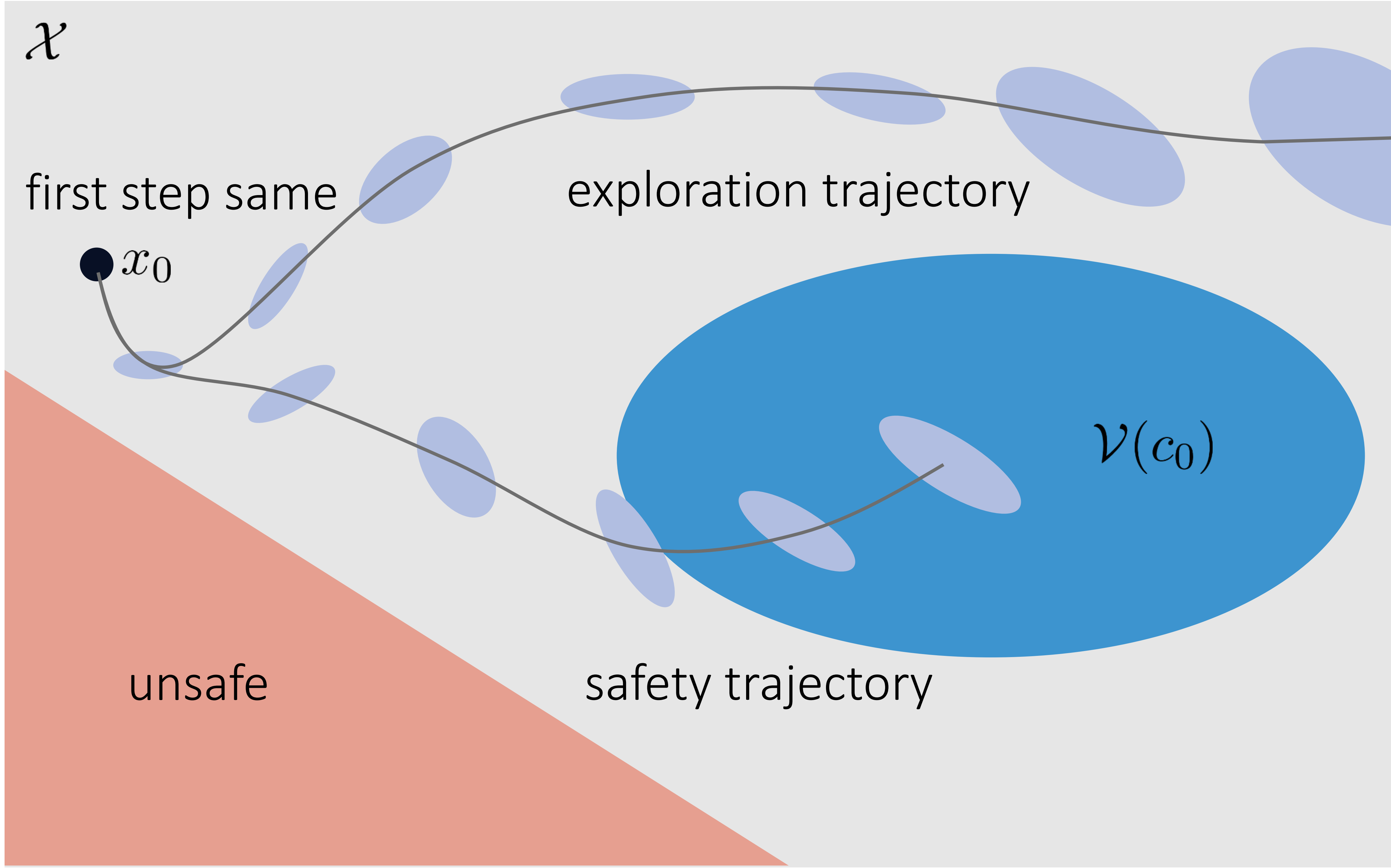
Safe model-based learning framework



Theorem (informally):

Under suitable conditions can always guarantee that we are able to return to the safe set

Safe model-based learning framework



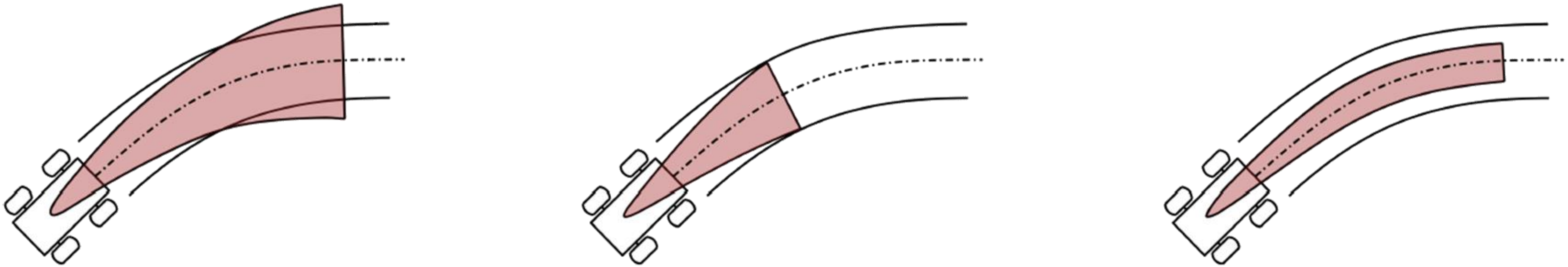
Exploration limited by size of the safe set!

How should we collect data for a control task?

Optimizing expected performance

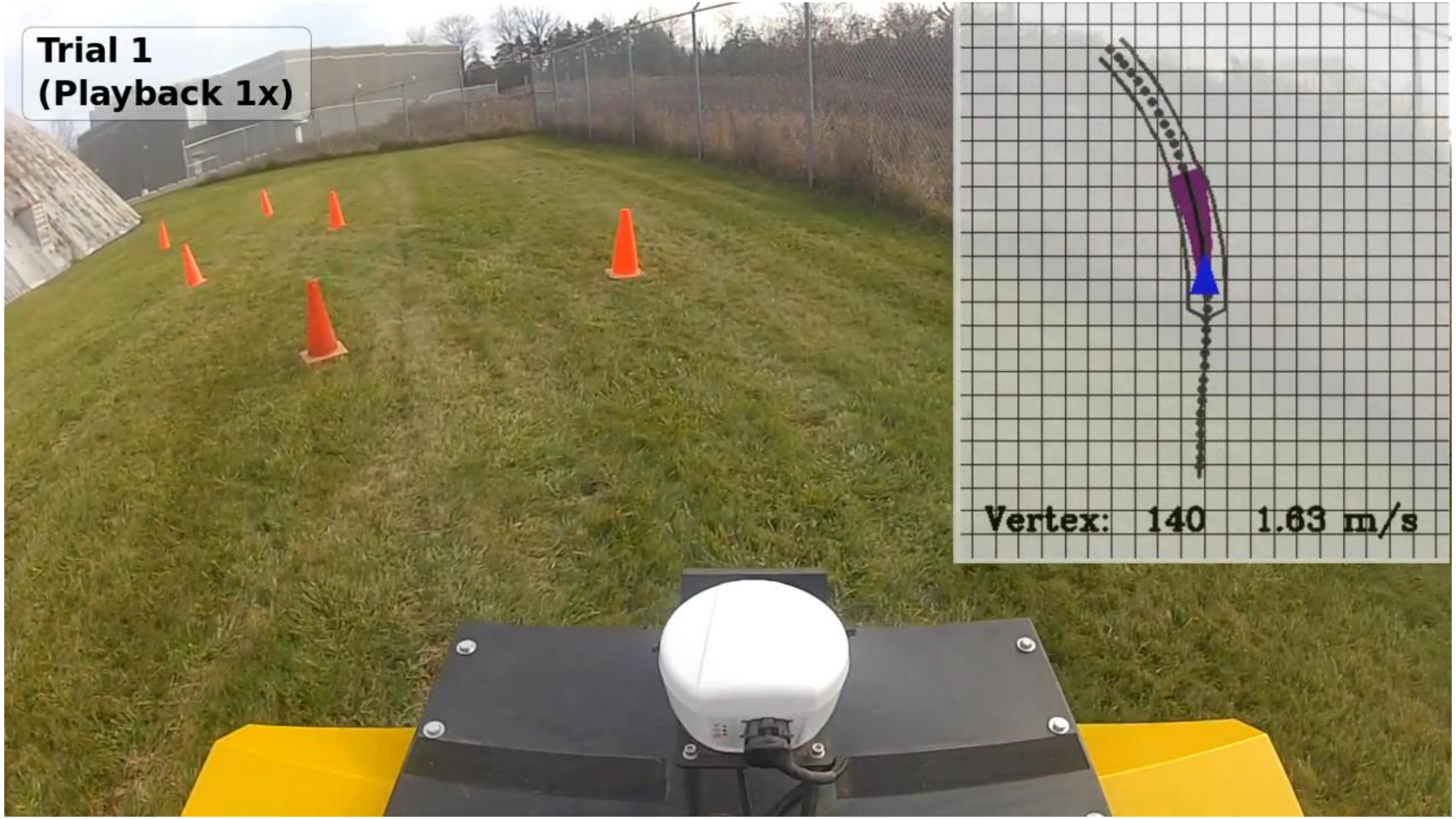
We design our cost functions to be helpful for optimization

Exploration objective: $$\underset{\{u_0, u_1, \ldots, u_{N-1}\}}{\text{minimize}} \mathrm{E}\left[\sum_{k=0}^{N-1} J(x_k, u_k) + J_N(x_N)\right]$$



Driving too fast | Slow down for safety | Faster driving after learning

Example



Video at https://youtu.be/3xRNmNv5Efk

Robust constrained learning-based NMPC enabling reliable mobile robot path tracking
C.J. Ostafew, A.P. Schoellig, T.D. Barfoot, IJRR, 2016

Summary and Outlook

Understand model and learning dynamics

Define safety, analyze a model for safety

Algorithm to safely acquire data

Gaussian processes

Lyapunov stability

Model predictive control

Safe Model-based Reinforcement Learning

https://berkenkamp.me

www.dynsyslab.org

Thanks To…

My Team – Industrial Partners – Funding Agencies

You Tube

Follow US!

www.dynsyslab.org

My outstanding collaborators at **U of T** (Tim Barfoot) and **ETH** (Andreas Krause, Raffaello D’Andrea and the whole FMA team).