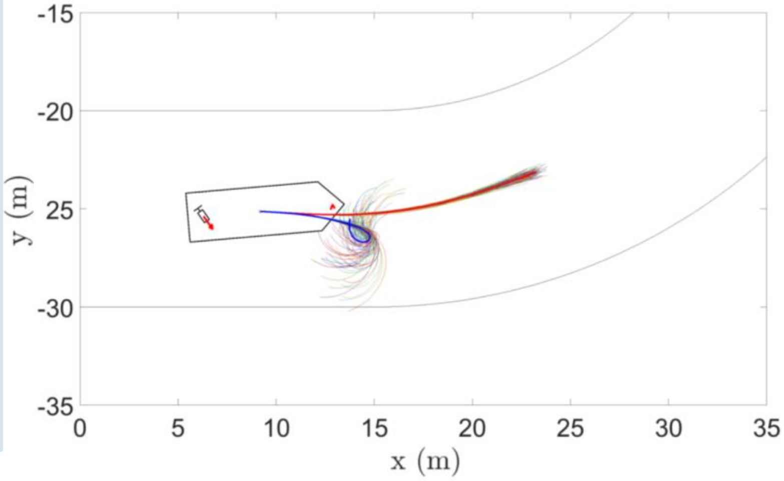
Feature-Based Nonlinear Model Predictive Path Integral (MPPI) Control with Application to Maritime Systems

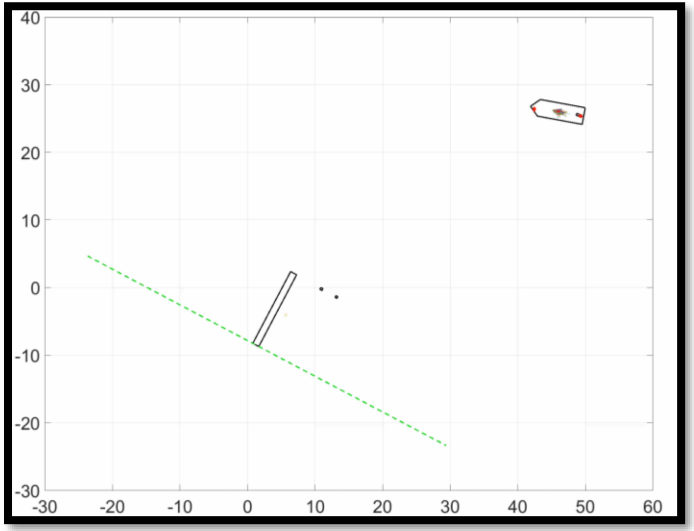


Hannes Homburger

Furuta Pendulum

Self-Balancing MonoChair2

Research Vessel Solgenia



Sample-based MPPI control leads to desired behavior in different applications.
How can the algorithm be improved?

A brief history of MPPI

[Fleming and Mitter, 1982] Posterior inference in a certain class of diffusion processes can be mapped onto a stochastic optimal control problem

[Kappen, 2005] Path integral (PI) control problems → Value function can be transformed into a linear partial differential equation → Existence of closed-form solution via Feynman-Kac path integral

[Todorov, 2009] The optimal control can be estimated using Monte Carlo sampling

[Thijssen and Kappen, 2015] Lemma: Solution of PI OCP is the optimal sampler for Monte Carlo method

[Williams et al., 2018] Using the PI Framework in MPC → MPPI Control

Based on this framework …

A lot of research on MPPI

The International Journal of Robotics Research
1–21
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/02783649211047890
journals.sagepub.com/home/ijr
SAGE

Constrained stochastic optimal control with learned importance sampling: A path integral approach

Jan Carius[1], René Ranftl[2], Farbod Farshidian[1] and Marco Hutter[1]

IEEE ROBOTICS AND AUTOMATION LETTERS. PREPRINT VERSION. ACCEPTED JULY, 2022

Autonomous Navigation of *AGVs* in Unknown Cluttered Environments: *log-MPPI* Control Strategy

Ihab S. Mohamed[1], Kai Yin[2], and Lantao Liu[1]

2022 IEEE 61st Conference on Decision and Control (CDC)
December 6-9, 2022. Cancún, Mexico

Path Integral Methods with Stochastic Control Barrier Functions

Chuyuan Tao†, Hyung-Jin Yoon*, Hunmin Kim†, Naira Hovakimyan†, and Petros Voulgaris*

2022 IEEE International Conference on Robotics and Automation (ICRA)
May 23-27, 2022. Philadelphia, PA, USA

Trajectory Distribution Control for Model Predictive Path Integral Control using Covariance Steering

Ji Yin, Zhiyuan Zhang, Evangelos Theodorou, Panagiotis Tsiotras

2022 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)

Entropy Regularised Deterministic Optimal Control: From Path Integral Solution to Sample-Based Trajectory Optimisation

Tom Lefebvre, Guillaume Crevecoeur

Temporal Difference Learning for Model Predictive Control

Nicklas A Hansen, Hao Su, Xiaolong Wang *Proceedings of the 39th International Conference on Machine Learning*, PMLR 162:8387-8406, 2022.

IEEE ROBOTICS AND AUTOMATION LETTERS. PREPRINT VERSION. ACCEPTED JULY, 2022

Smooth Model Predictive Path Integral Control without Smoothing

Taekyung Kim, Gyuhyun Park, Kiho Kwak, Jihwan Bae, and Wonsuk Lee

arXiv > cs > arXiv:2208.02439

Computer Science > Robotics

[Submitted on 4 Aug 2022]

MPPI-IPDDP: Hybrid Method of Collision-Free Smooth Trajectory Generation for Autonomous Robots

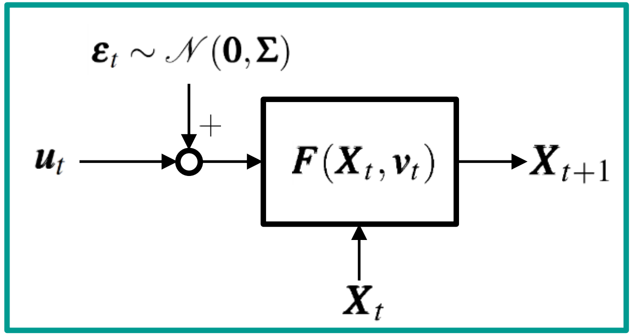
Min-Gyeom Kim, Kwang-Ki K. Kim

… many recently published papers further develop MPPI

Path Integral Control Problem

Hochschule Konstanz
Department of Electrical Engineering
and Information Technology

Dynamics:



For the nonlinear discrete time stochastic system

$$\boldsymbol{X}_{t+1} = \boldsymbol{F}(\boldsymbol{X}_t, \boldsymbol{v}_t), \quad \forall t \in \{0, 1, \ldots, T-1\},$$
$$\boldsymbol{v}_t = \boldsymbol{u}_t + \boldsymbol{\varepsilon}_t, \ \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$$

a *PI control problem* is given if the following assumptions hold:

$$\min_{\mathbf{U}} \mathbb{E}_{\mathbb{Q}} \left\{ \phi(\boldsymbol{X}_T) + \sum_{t=0}^{T-1} [C(\boldsymbol{X}_t) + \boldsymbol{u}_t^\top \boldsymbol{R} \boldsymbol{u}_t] | \boldsymbol{X}_0 = \boldsymbol{x}_0 \right\} \quad \text{with} \quad \boldsymbol{R} = \lambda \boldsymbol{\Sigma}^{-1}, \ \lambda \in \mathbb{R}^+,$$

$$\mathbb{E}_{\mathbb{Q}}[\phi(\boldsymbol{V}, \boldsymbol{U}, \boldsymbol{\Sigma})] = \int \phi(\boldsymbol{V}, \boldsymbol{U}, \boldsymbol{\Sigma}) q(\boldsymbol{V}|\boldsymbol{U}, \boldsymbol{\Sigma}) d\boldsymbol{V}, \quad q(\boldsymbol{V}|\boldsymbol{U}, \boldsymbol{\Sigma}) = \prod_{k=0}^{N-1} [(2\pi)^N |\boldsymbol{\Sigma}|]^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \boldsymbol{\varepsilon}_k^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}_k\right) \quad \text{and} \quad \boldsymbol{\varepsilon}_k = \boldsymbol{v}_k - \boldsymbol{u}_k$$

MPPI Control

Solution:

$$u_t^* = \mathbb{E}_{\mathbb{Q}} \left\{ \omega(\mathbf{V}) \nu_t \right\}, \quad t = 0, \ldots, T-1$$

Algorithm:

Optimize a control sequence:

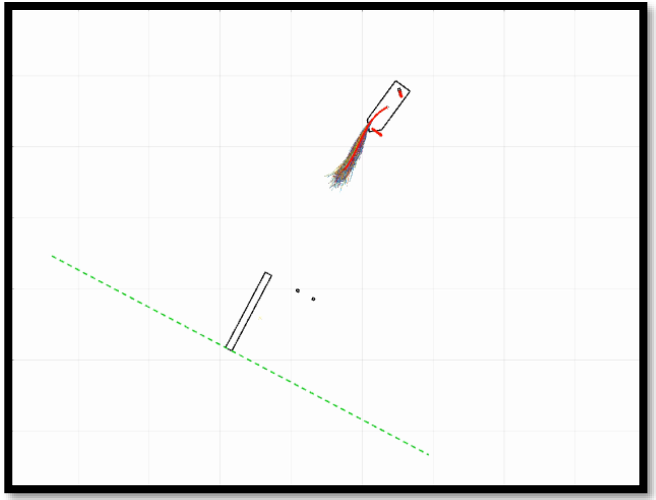
1. Get an initial control sequence and a recent state estimate
2. Sample trajectories by simulation using the recent state estimate and disturbed inputs
3. Calculate the *path costs* for each sampled trajectory
4. Calculate the weight for each sampled trajectory (high cost → low weight)
5. Improve the initial control sequence by the weighted inference of the randomly disturbed input samples

6. Apply the first component of the improved control sequence, shift the sequence and go to 1

Hochschule Konstanz
Department of Electrical Engineering
and Information Technology

Exploration Problem

- Using an **infinite number of samples** the MPPI algorithm would find the **global optimal control sequence**
- **Number of samples is limited** despite **excellent parallelizability**
- Samples are **spread „around“** the **previous optimal solution**



How to sample trajectories in low-cost areas?

Enlargement of the explored state-space

Ideas:

[Wiliams et al. 2018]

Draw more samples

[Yin et al. 2021]

Covariance Steering

[Kusomoto et al. 2019]

Distribution Learning

This approach:

We can use knowledge about the system
to define features and learn expert behavior w.r.t. these features.

Feature-Based Extension

A feature is implicitly defined by a second optimal control problem by choosing feature costs based on system knowledge:

Differ in state cost functions

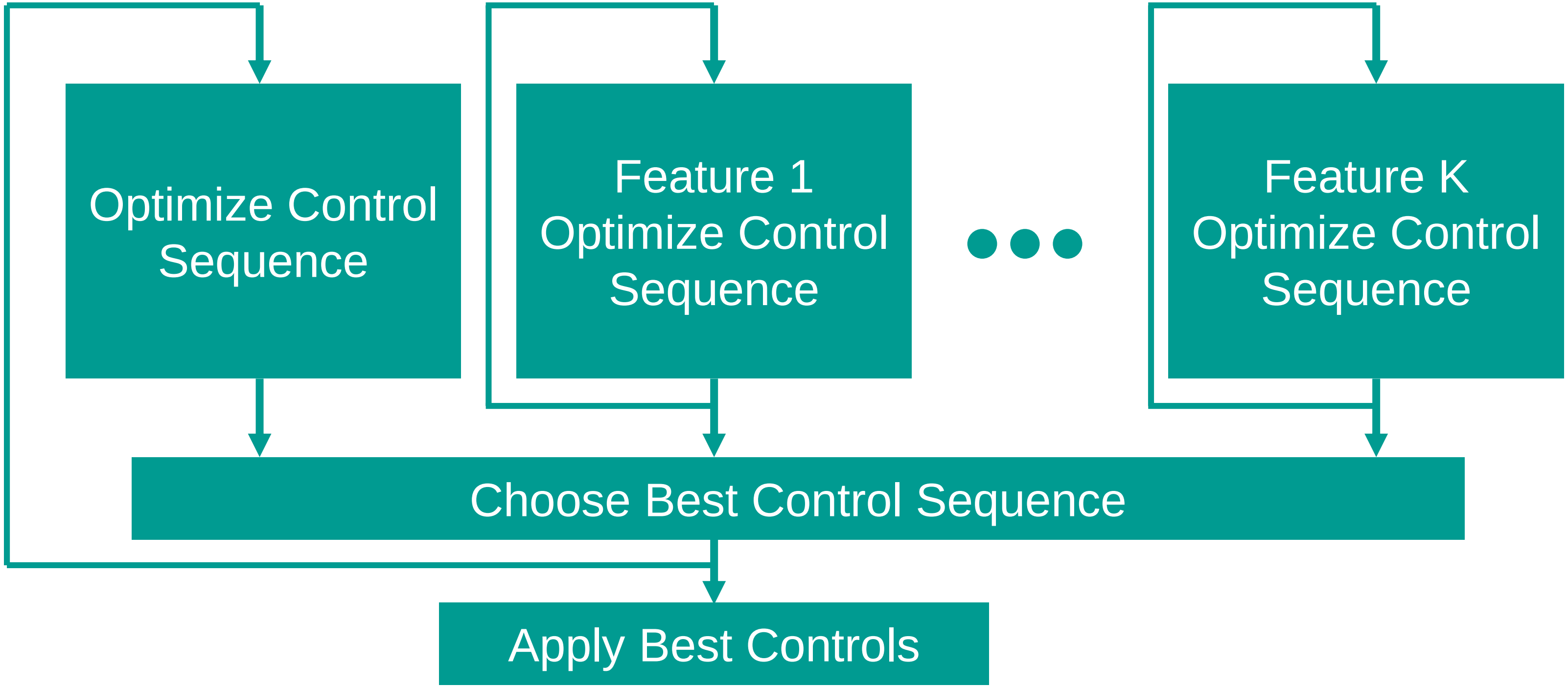
$$\boldsymbol{U}^F = \underset{\boldsymbol{U}}{\operatorname{argmin}}\ \mathbb{E}_{\mathbb{Q}} \left\{ \phi_F(\boldsymbol{X}_T) + \sum_{t=0}^{T-1} [C_F(\boldsymbol{X}_t) + \boldsymbol{u}_t^\top \boldsymbol{R} \boldsymbol{u}_t] \middle| \boldsymbol{X}_0 = \boldsymbol{x}_0 \right\}$$

$$\boldsymbol{X}_{t+1} = \boldsymbol{F}(\boldsymbol{X}_t, \boldsymbol{v}_t) \text{ with } \boldsymbol{v}_t \sim \mathcal{N}(\boldsymbol{u}_t, \boldsymbol{\Sigma}),\ \boldsymbol{R} = \lambda \boldsymbol{\Sigma}^{-1}$$

same dynamics

same noise

Feature-Based Algorithm



The control sequence is chosen that causes the lowest main costs.

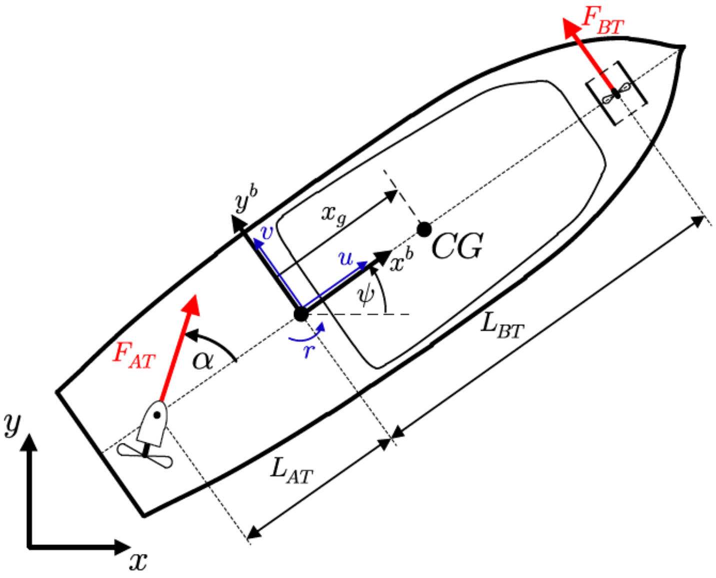
Vessel Dynamics

- 3DOF model describing the planar dynamics
  roll, pitch and heave motion is neglected

- System state is given by $\boldsymbol{x} = (\boldsymbol{\eta}^\top \ \ \boldsymbol{\nu}^\top \ \ \boldsymbol{a}^\top \ \ \boldsymbol{w}^\top)^\top$

- Dynamics is given by

$$\dot{\boldsymbol{w}} = \boldsymbol{u}$$
$$\dot{\boldsymbol{a}} = \boldsymbol{d}(\boldsymbol{a}, \boldsymbol{w})$$
$$\boldsymbol{M}\dot{\boldsymbol{\nu}} + \boldsymbol{C}_{RB}(\boldsymbol{\nu})\boldsymbol{\nu} + \boldsymbol{N}(\boldsymbol{\nu})\boldsymbol{\nu} = \boldsymbol{\tau}_c(\boldsymbol{a}, \boldsymbol{\nu}) + \boldsymbol{\tau}_d$$
$$\dot{\boldsymbol{\eta}} = \boldsymbol{J}(\psi)\boldsymbol{\nu},$$

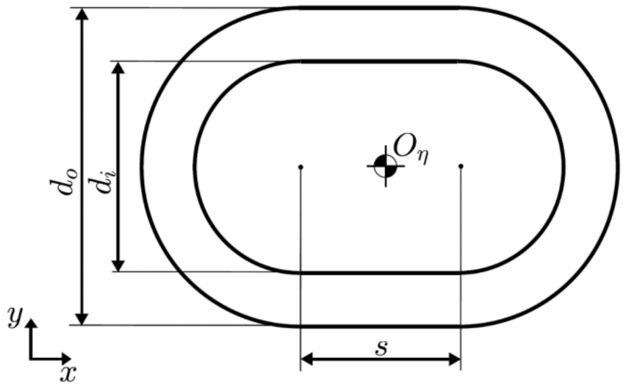


Hochschule Konstanz
Department of Electrical Engineering
and Information Technology

Lagrange Costs for Tracking

$$C(\boldsymbol{x}) = C_{\text{pos}}(\boldsymbol{\eta}) + C_{\text{vel}}(\boldsymbol{v}) + C_{\text{col}}(\boldsymbol{\eta})$$

$$C_{\text{pos}}(\boldsymbol{\eta}) = c \sum_{k_x=-N/2}^{N/2} \sum_{k_y=-N/2}^{N/2} \zeta(x + k_x\Delta, y + k_y\Delta)$$

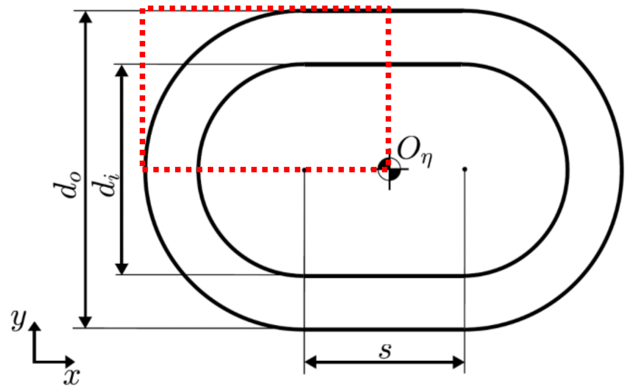


"Stay on the track"

$$C_{\text{vel}}(\boldsymbol{v}) = c_u(u_{\text{ref}} - u)^2$$

"Hold the desired velocity"

$$C_{\text{col}}(\boldsymbol{\eta}) = \begin{cases} c_{\text{col}} & \text{if } \boldsymbol{\eta} \in \boldsymbol{\eta}_C \\ 0 & \text{else} \end{cases}$$



"Don't collide with obstacles"

Inequality Constraints

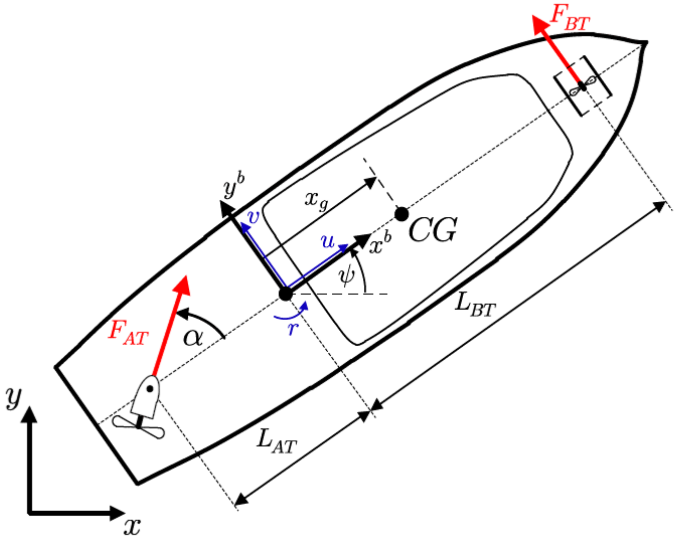
Due to the inequality constraints on the actuators‘ states

$$\boldsymbol{h}(\boldsymbol{x},\boldsymbol{u}) = \begin{pmatrix} |n_{\mathrm{AT}}| - |n_{\mathrm{AT,max}}| \\ |n_{\mathrm{BT}}| - |n_{\mathrm{BT,max}}| \end{pmatrix} \leq \boldsymbol{0}$$

we add a penalty term

$$C_{\mathrm{MPPI}}(\boldsymbol{x}) = C(\boldsymbol{x}) + c_{\mathrm{ineq}} \max[\boldsymbol{0}, \boldsymbol{h}(\boldsymbol{a})]$$

Note: The maximal velocity of the azimuth thruster‘s orientation is already a part of the actuator dynamics in the system dynamics

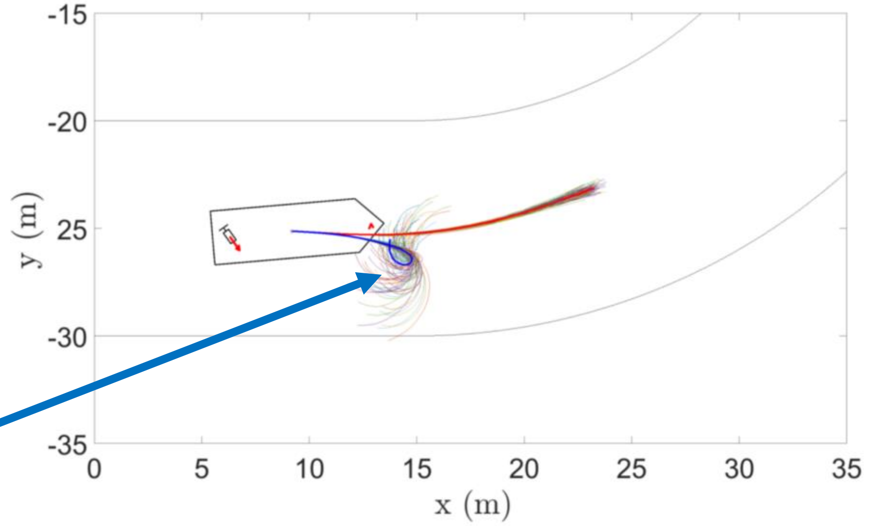


Method is derivative free → no differentiability requirements

Feature Definition

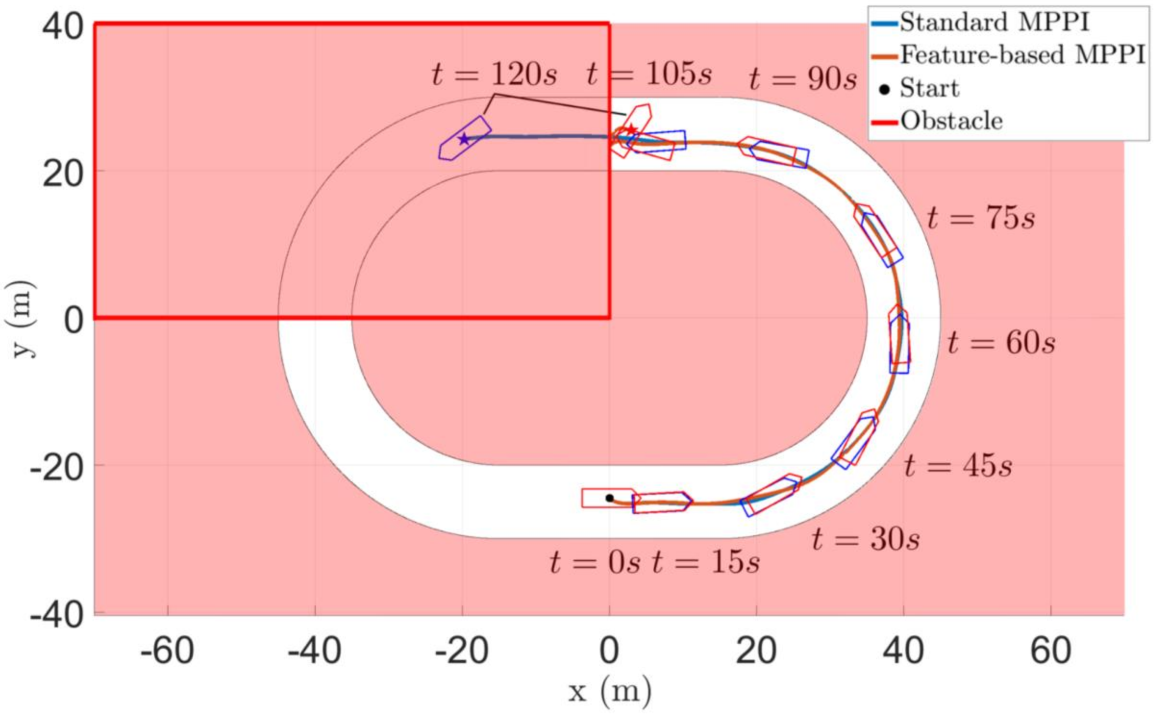
For collision avoidance we define a emergency break feature with the costs:

$$C_1(\boldsymbol{x}) = \phi_1(\boldsymbol{x}) = \boldsymbol{v}^\top \boldsymbol{Q}\boldsymbol{v}, \quad \boldsymbol{Q} = \mathrm{diag}(c_u, c_v, c_r)$$



While processing, the controller learns an optimal braking control sequence

Resulting Trajectories

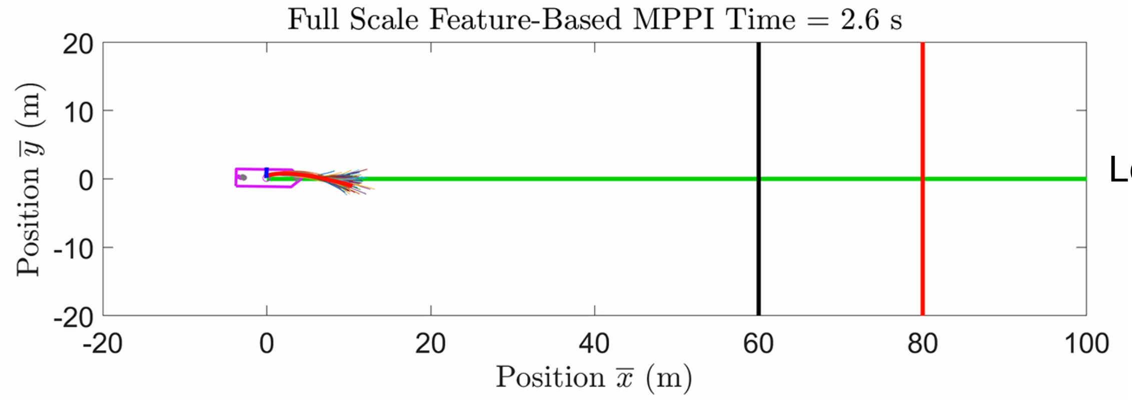


While normal processing both controllers show nearly same behavior

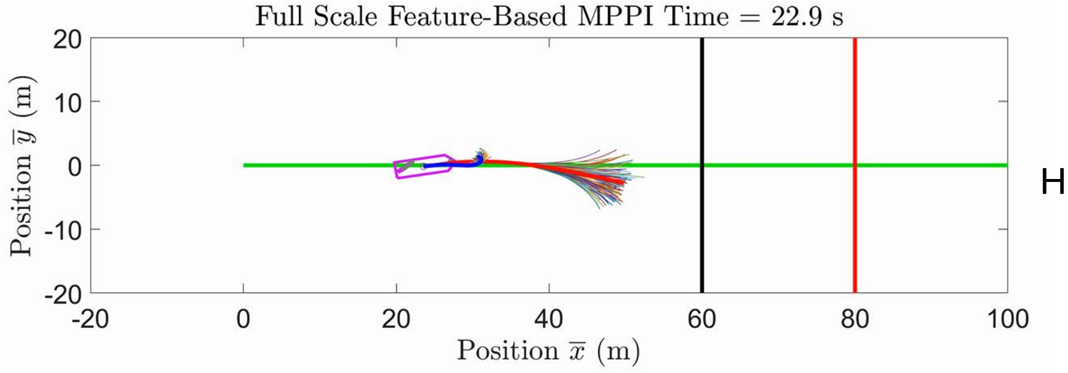
However, next to the obstacle while standard MPPI causes a collision, feature-based MPPI prevents a collision

And in real-world…

Full Scale Experiment

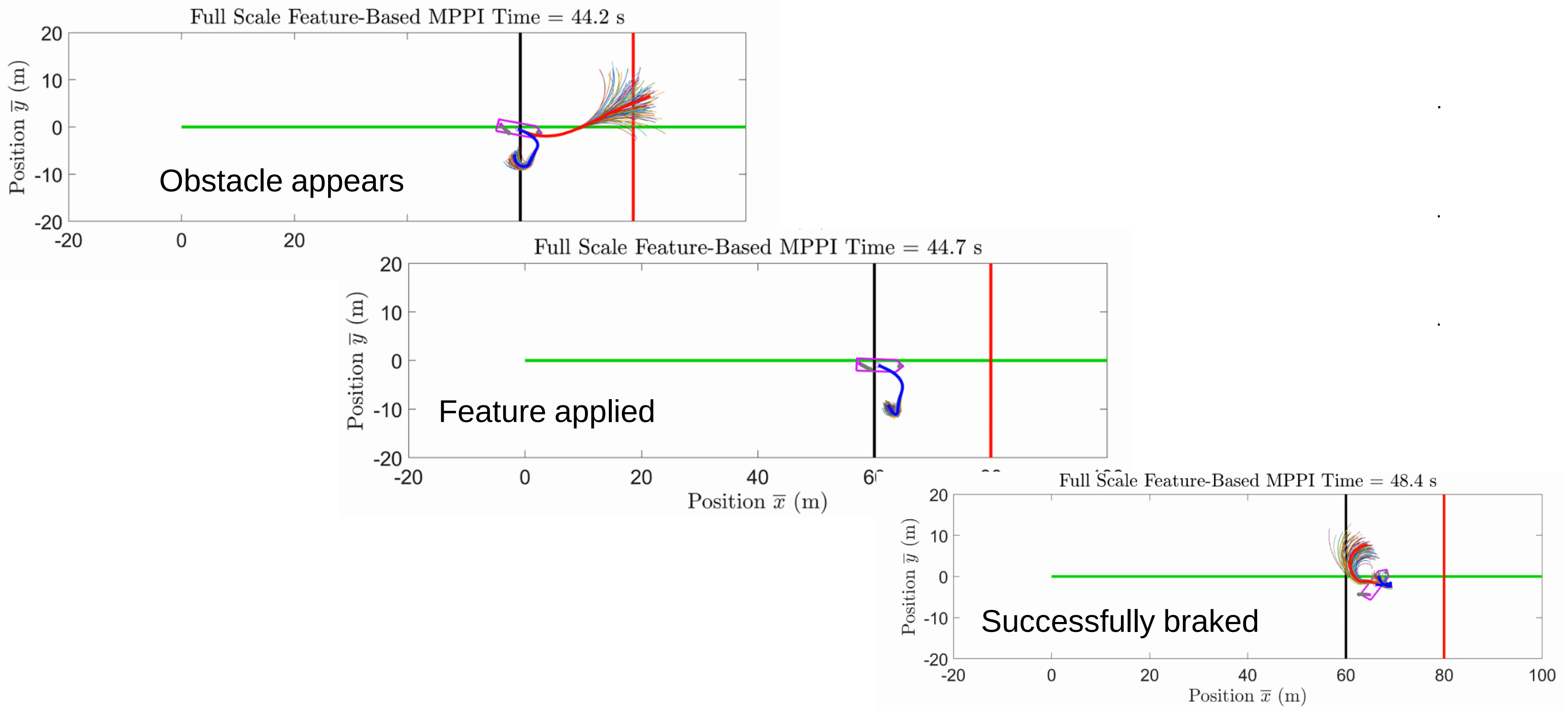


Low Speed



High Speed

Full Scale Experiment



Full Scale Experiment

Hochschule Konstanz
Department of Electrical Engineering
and Information Technology

Summary

- Feature-based MPPI control is applied to control an autonomous surface vessel
- Validation in simulation and full-scale experiments

Future Work

- Systematic comparison of benefits and drawbacks of different embedded solvers (acados, GRAMPC, and MPPI) within different applications

+ Don‘t need derivative information

− Don‘t use derivative information

References

[Kappen, 2005] Kappen, H. J. Linear theory for control of non-linear stochastic systems. *Physical Review letters*, 95:200201.

[Thijssen and Kappen, 2015] Thijssen, S. and Kappen, H. J. Path integral control and state-dependent feedback. *Phys. Rev. E*, 91:032104.

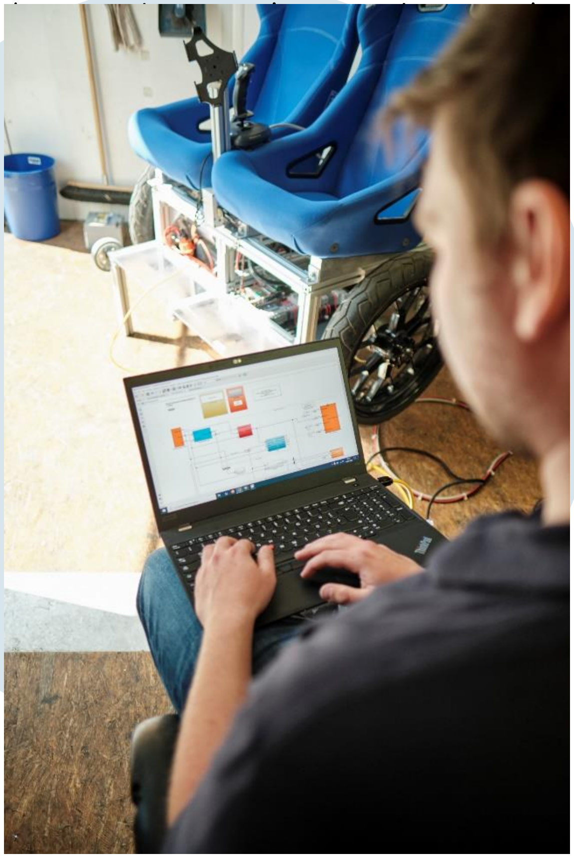
[Todorov, 2009] Todorov, E. (2009). Efficient computation of optimal actions. *Proceedings of the National Academy of Sciences*, 106:11478–11483.

[Wiliams et al. 2017] G. Williams, N. Wagener, B. Goldfain, P. Drews, J. Rehg, B. Boots, and E. Theodorou, Information Theoretic MPC for Model-Based Reinforcement Learning. In IEEE International Conference on Robotics and Automation (ICRA), Singapore, 2017.

[Yin et al. 2021] J. Yin, Z. Zhang, E. Theodorou and P. Tsiotras, Improving Model Predictive Path Integral using Covariance Steering. In arXiv preprints, arXiv:2109.12147, 2021.

[Kusumoto et. al 2019] R. Kusumoto, L. Palmieri, M. Spies, A. Csiszar, K. Arras, Informed Information Theoretic Model Predictive Control, 2019 International Conference on Robotics and Automation (ICRA), Palais des congres de Montreal, Montreal, Canada, 2019.

[Kappen and Ruiz, 2016] Kappen, H. J. and Ruiz, G. Adaptive importance sampling for control and inference, Journal of Statistical Physics, 2016.



Thank You!
Any Questions?

Hannes Homburger

Institute of System Dynamics
University of Applied Sciences Konstanz
Alfred-Wachtel-Straße 8
78462 Konstanz
Germany

Mail: hhomburg@htwg-konstanz.de