Kernel Embedding for Particle Gibbs-Based Optimal Control

L. Hochschwarzer
Final Report Master's Thesis
Supervisor: R. Lefringhausen

Chair of Information-oriented Control
Technical University of Munich

Motivation

[Xiloyannis, Chiaradia, Frisoli and Masia 2019]

Challenges:

- Unknown dynamics
- Latent states
- Safety

Problem Statement - System

Given: Dataset $\mathbb{D} = \{\boldsymbol{u}_t, \boldsymbol{y}_t\}_{t=\text{-}T:\text{-}1}$ from unknown system

$$\begin{aligned} \boldsymbol{x}_{t+1} &= \boldsymbol{f}\left(\boldsymbol{x}_t, \boldsymbol{u}_t\right) + \boldsymbol{v}_t, \qquad && \boldsymbol{v}_t \sim \boldsymbol{\mathcal{V}}, \\ \boldsymbol{y}_t &= \boldsymbol{g}\left(\boldsymbol{x}_t, \boldsymbol{u}_t\right) + \boldsymbol{w}_t, \qquad && \boldsymbol{w}_t \sim \boldsymbol{\mathcal{W}} \end{aligned}$$

Assumptions

- Known system structure

$$\begin{aligned} \boldsymbol{x}_{t+1} &= \boldsymbol{f}_{\boldsymbol{\theta}}\left(\boldsymbol{x}_t, \boldsymbol{u}_t\right) + \boldsymbol{v}_t, \qquad && \boldsymbol{v}_t \sim \boldsymbol{\mathcal{V}}_{\boldsymbol{\theta}}, \\ \boldsymbol{y}_t &= \boldsymbol{g}_{\boldsymbol{\theta}}\left(\boldsymbol{x}_t, \boldsymbol{u}_t\right) + \boldsymbol{w}_t, \qquad && \boldsymbol{w}_t \sim \boldsymbol{\mathcal{W}}_{\boldsymbol{\theta}}. \end{aligned}$$

- Known priors $p(\boldsymbol{\theta})$ and $p(\boldsymbol{x}_{\text{-}T})$

Problem Statement - Optimal Control Problem

Goal: Solve optimal control problem (OCP)

Stochastic OCP

$$\min_{\boldsymbol{u}_{0:H}} J_H(\boldsymbol{u}_{0:H})$$

subject to:

$$P\left[\max(\boldsymbol{h}(\boldsymbol{u}_{0:H}, \boldsymbol{x}_{0:H}, \boldsymbol{y}_{0:H})) \leq 0\right] \geq 1 - \alpha$$

Problem: Underlying data distribution is unknown

Related Works

Particle Gibbs based optimal control **[Lefringhausen, Srithasan, Lederer and Hirche 2024]**

$\Rightarrow$ Guarantees determined retroactively via scenario theory

Alternative approaches:

- Wasserstein ambiguity [Hota, Cherukuri and Lygeros 2019]
  - $\Rightarrow$ Constraints limited to affine functions
- Kernel embeddings **[Nemmour, Kremer, Schoelkopf and Zhu 2022]**
  [Thorpe, Lew, Oishi and Zhu 2022]

Particle Gibbs Scenarios

Particle Gibbs provides the scenarios $\boldsymbol{\delta}^{[1:N]} = \{\boldsymbol{\theta}, \boldsymbol{x}_0, \boldsymbol{v}_{0:H}, \boldsymbol{w}_{0:H}\}^{[1:N]}$

$\Rightarrow$ Scenarios **and** input $\boldsymbol{u}_{0:H}$ define the trajectory

Required: Reformulate chance-constraint problem with scenarios $\boldsymbol{\delta}^{[1:N]}$

Chance Constraints

$$P\left[\max(\boldsymbol{h}(\boldsymbol{u}_{0:H}, \boldsymbol{x}_{0:H}, \boldsymbol{y}_{0:H}) \leq 0\right] \geq 1 - \alpha$$

$$\Downarrow$$

Scenario Approach ([Lefringhausen+ 2024])

$$\max(\boldsymbol{h}(\boldsymbol{u}_{0:H}, \boldsymbol{x}_{0:H}^{[n]}, \boldsymbol{y}_{0:H}^{[n]})) \leq 0, \ \forall n = 1, \ldots, N$$

$\Rightarrow$ Risk factor α not considered in optimization

Maximum Mean Discrepancy (MMD) ambiguity sets

Required: Reformulate chance-constraint problem with scenarios $\boldsymbol{\delta}^{[1:N]}$
Approach: Replace distribution P with the ambiguity set $\mathcal{P}$

MMD ambiguity set

$$\mathcal{P} = \left\{ \tilde{P} : \mathrm{MMD}(\tilde{P}, P_N) \leq \varepsilon \right\}.$$

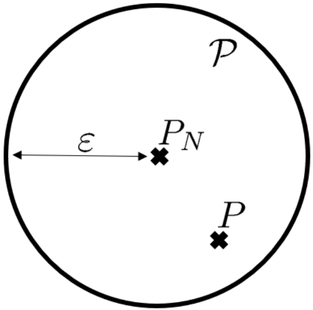
Radius ε obtained via bootstrap construction

$$\Downarrow$$

Expanded Chance-Constraints

$$\inf_{\tilde{P} \in \mathcal{P}} \tilde{P} \left[\tilde{h}(\boldsymbol{u}_{0:H}, \boldsymbol{\delta}) \leq 0 \right] \geq 1 - \alpha.$$

with $\tilde{h}(\boldsymbol{u}_{0:H}, \boldsymbol{\delta}) = \max(\boldsymbol{h}(\boldsymbol{u}_{0:H}, \boldsymbol{x}_{0:H}, \boldsymbol{y}_{0:H}))$.



Constraint Reformulation

Feasible Region of chance constraint

$$Z := \left\{ \boldsymbol{u}_{0:H} \in \mathcal{U}^{H+1} : \inf_{\tilde{P} \in \mathcal{P}} \tilde{P}\left[\tilde{h}(\boldsymbol{u}_{0:H}, \boldsymbol{\delta}) \leq 0 \right] \geq 1 - \alpha \right\}$$

$$\Downarrow$$

Reformulated Feasible Region [Nemmour+ 2022]

$$\hat{Z} := \left\{ \boldsymbol{u}_{0:H} \in \mathcal{U}^{H+1} : \begin{array}{l} g_0 + \dfrac{1}{N} \displaystyle\sum_{n=1}^{N} (\boldsymbol{K}\boldsymbol{\gamma})_n + \varepsilon \sqrt{\boldsymbol{\gamma}^{\mathsf{T}} \boldsymbol{K} \boldsymbol{\gamma}} \leq t' \alpha \\ [\tilde{h}(\boldsymbol{u}_{0:H}, \boldsymbol{\delta}^{[n]}) + t']_+ \leq g_0 + (\boldsymbol{K}\boldsymbol{\gamma})_n, \; n = 1, \ldots, N \\ g_0 \in \mathbb{R}, \boldsymbol{\gamma} \in \mathbb{R}^N, t' \in \mathbb{R} \end{array} \right\}$$

Hyperparameter Tuning

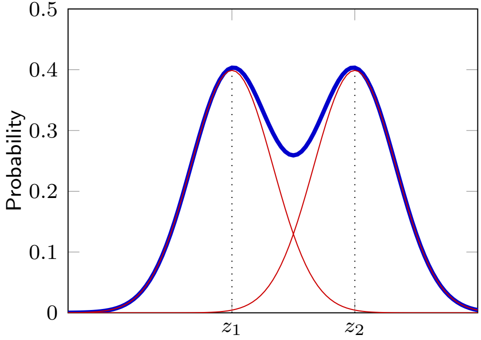
- Gaussian kernels

$$k(z, z') = \exp\left(-\frac{1}{2\sigma^2}(z - z')^2\right)$$

- Split scenarios into training set $\{z_i\}$ and test set $\{z'_j\}$
- Create likelyhood function

$$p(z) = \frac{1}{N_{\text{train}}} \sum_{i=1}^{N_{\text{train}}} \frac{1}{\sqrt{2\pi\sigma^2}} k(z, z_i)$$

- Maximize sum of likelyhoods over test set by selected a good σ



Hyperparameter Tuning

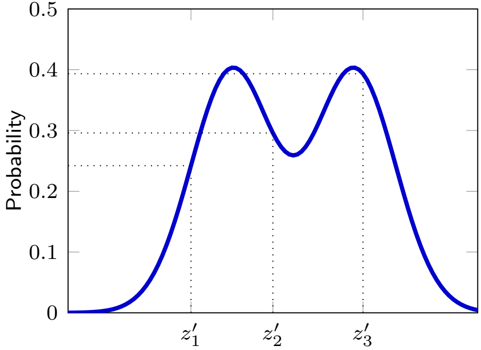
- Gaussian kernels

$$k(z, z') = \text{exp}\left(-\frac{1}{2\sigma^2}(z - z')^2\right)$$

- Split scenarios into training set $\{z_i\}$ and test set $\{z'_j\}$
- Create likelyhood function

$$p(z) = \frac{1}{N_{\text{train}}}\sum_{i=1}^{N_{\text{train}}}\frac{1}{\sqrt{2\pi\sigma^2}}k(z, z_i)$$

- Maximize sum of likelyhoods over test set by selected a good σ



Reformulated Optimal Control Problem

$$\min_{\boldsymbol{u}_{0:H}, g_0, \boldsymbol{\gamma}, t'} J_H(\boldsymbol{u}_{0:H})$$

subject to: $\forall n \in \mathbb{N}_{\leq N},\ \forall t \in \mathbb{N}^0_{\leq H},$

$$\left.\begin{aligned} \boldsymbol{x}_{t+1}^{[n]} &= \boldsymbol{f}_{\boldsymbol{\theta}^{[n]}}\left(\boldsymbol{x}_t^{[n]}, \boldsymbol{u}_t\right) + \boldsymbol{v}_t^{[n]} \\ \boldsymbol{y}_t^{[n]} &= \boldsymbol{g}_{\boldsymbol{\theta}^{[n]}}\left(\boldsymbol{x}_t^{[n]}, \boldsymbol{u}_t\right) + \boldsymbol{w}_t^{[n]} \end{aligned}\right\} \text{Dynamic Constraints}$$

$$\boldsymbol{u}_{0:H} \in \hat{Z}(g_0, \boldsymbol{\gamma}, t') \quad \Big\} \text{ Reformulated Chance Constraints}$$

Simulation Setup

- Unknown linear system: $\boldsymbol{f}(\boldsymbol{x},u) = \begin{bmatrix} 0.8x_1 - 0.5x_2 \\ 0.4x_1 + 0.5x_2 + u \end{bmatrix}$

$$\boldsymbol{v}_t \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{Q} = \begin{bmatrix} 0.03 & \text{-}0.004 \\ \text{-}0.004 & 0.01 \end{bmatrix}\right).$$

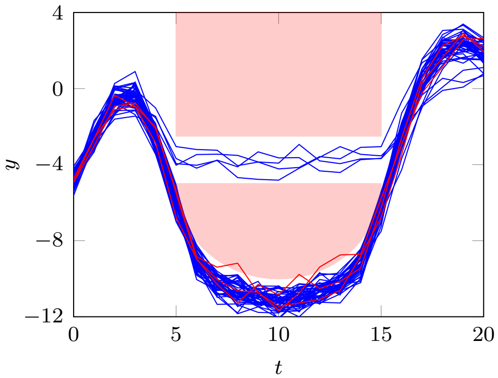
- Known system structure: $\boldsymbol{f}(\boldsymbol{x},u) = \boldsymbol{A}\,[x_1, x_2, u]^\mathsf{T}\,,\ \boldsymbol{v}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q})$
- Priors: $\boldsymbol{Q} \sim \mathcal{IW}(100\boldsymbol{I}_2, 10)$

  $\boldsymbol{A} \sim \mathcal{MN}(\boldsymbol{0}, \boldsymbol{Q}, 10\boldsymbol{I}_2)$ [Svensson+ 2017]

  $\boldsymbol{x}_{\text{-}T} \sim \mathcal{N}([2,2]^\mathsf{T}, \boldsymbol{I}_2)$
- Known measurement model $g(\boldsymbol{x},u) = x_1,\ w_t \sim \mathcal{N}(0, 0.1)$ (w.l.o.g.)
- Cost function $J_H = \sum_{t=0}^{H} u_t^2$
- Input constraints $|u| \leq 10$
- Number of scenarios used for optimization: $N = 100$

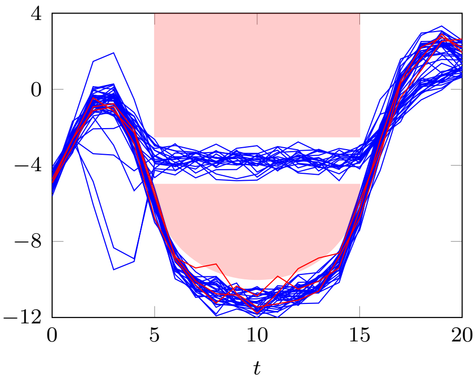
Risk Tuning

Scenario Approach



Average Cost $J_H = 338.4$

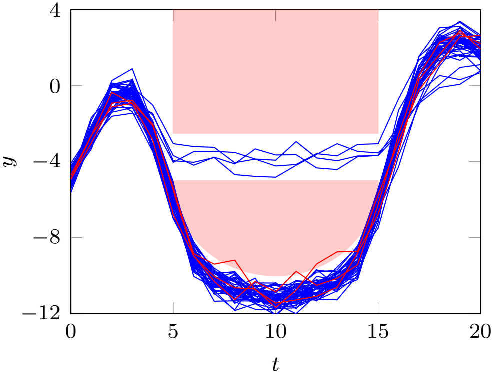
Kernel Approach ($\alpha = 0.1$)



Average Cost $J_H = 255.9$

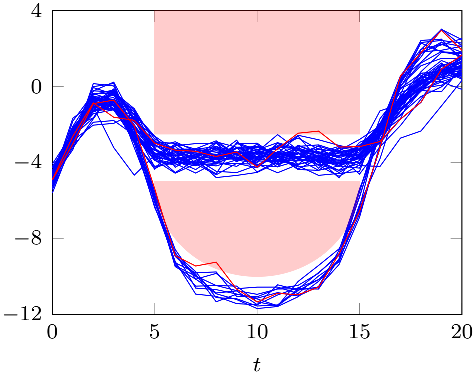
Risk Tuning

Scenario Approach



Average Cost $J_H = 338.4$

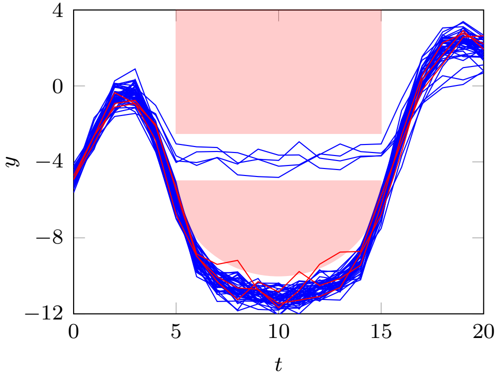
Kernel Approach ($\alpha = 0.2$)



Average Cost $J_H = 129.5$

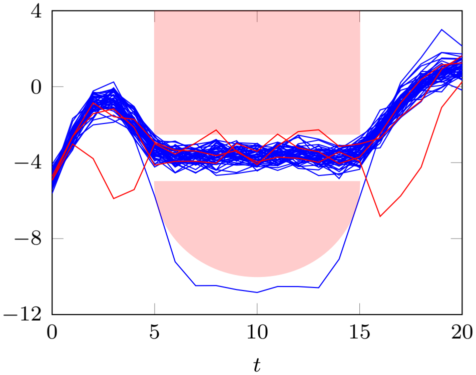
Risk Tuning

Scenario Approach



Average Cost $J_H = 338.4$

Kernel Approach ($\alpha = 0.3$)



Average Cost $J_H = 69.9$

Nonlinear System

Nonlinear System

$$\boldsymbol{f}(\boldsymbol{x}, u) = \begin{bmatrix} 0.8x_1 - 0.5x_2 + 0.1\cos(3x_1)x_2 \\ 0.4x_1 + 0.5x_2 + (1 + 0.3\sin(2x_2))u \end{bmatrix}$$

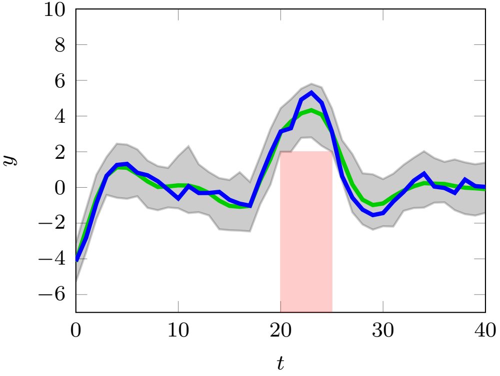
- Known system structure: $\boldsymbol{f}(\boldsymbol{x}, u) = \boldsymbol{A}\left[x_1, x_2, u, \cos(3x_1)x_2, \sin(2x_2)u\right]^\mathsf{T}$
- Input constraints $|u| \leq 5$
- Number of scenarios used for optimization: $N = 200$

Challenge: Numerical issues complicates solving OCPs with

- Low number of samples N
- Low Risk Factor α
- High number of constraints

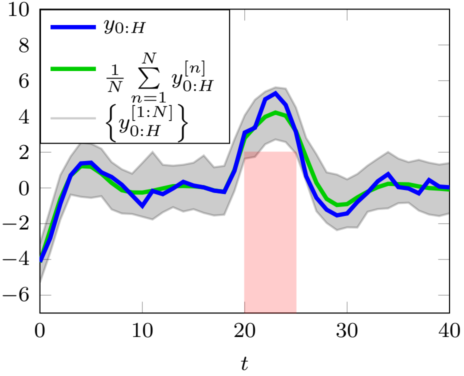
Optimal Control of Nonlinear Systems

Scenario Approach



Computation Time $= 32.8$s

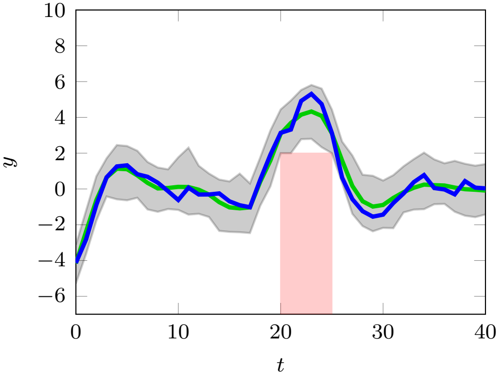
Kernel Approach ($\alpha = 0.2$)



Computation Time $= 176.1$s

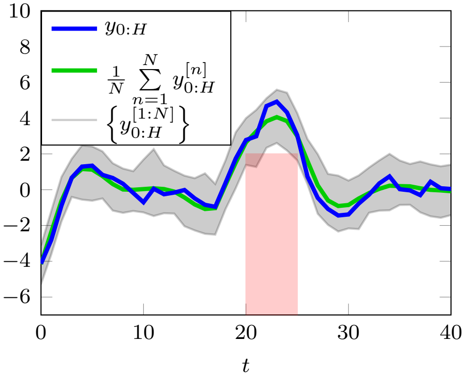
Optimal Control of Nonlinear Systems

Scenario Approach



Computation Time $= 32.8$s

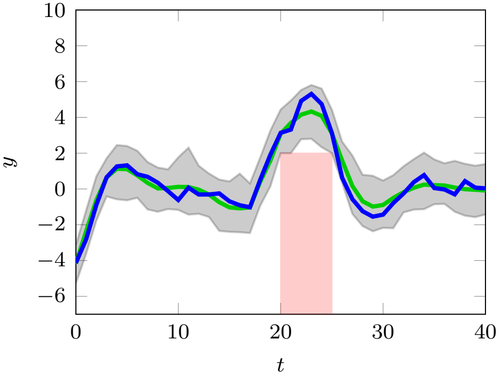
Kernel Approach ($\alpha = 0.4$)



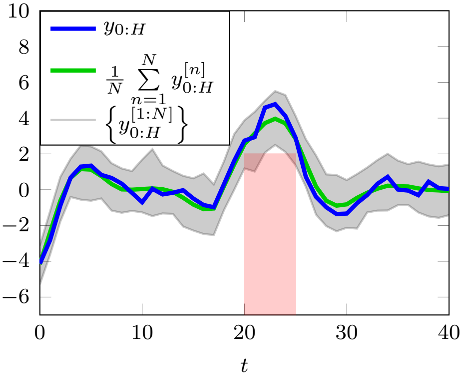
Computation Time $= 50.9$s

Optimal Control of Nonlinear Systems

Scenario Approach



Kernel Approach ($\alpha = 0.6$)



Computation Time $= 32.8$s

Computation Time $= 41.0$s

Conclusion

Summary:
The approach presented combines Particle Gibbs sampling with MMD-ambiguity sets using Kernel Embeddings

Kernel Embeddings allow for ...

- solving chance-constrained OCPs
- Choosing risk factor α

References

Ashish R Hota, Ashish Cherukuri and John Lygeros. **Data-Driven Chance Constrained Optimization under Wasserstein Ambiguity Sets**. In: *American Control Conference (ACC)* (2019).

Robert Lefringhausen, Supitsana Srithasan, Armin Lederer and Sandra Hirche. **Learning-Based Optimal Control with Performance Guarantees for Unknown Systems with Latent States**. In: *European Control Conference* (2024).

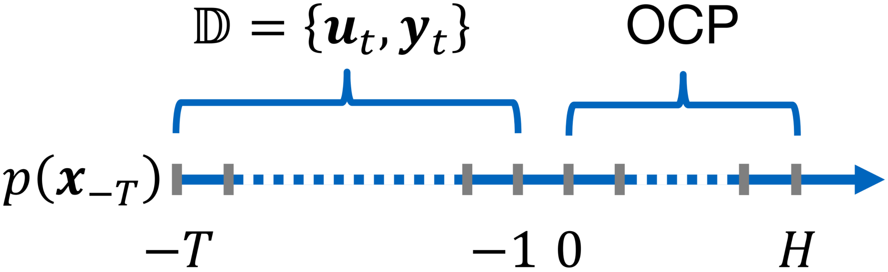
Yassine Nemmour, Heiner Kremer, Bernhard Schoelkopf and Jia-Jie Zhu. **Maximum Mean Discrepancy Distributionally Robust Nonlinear Chance-Constrained Optimization with Finite-Sample Guarantee**. In: *arXiv preprint arXiv:2204.11564* (2022).

Andreas Svensson and Thomas B Schoen. **A flexible state-space model for learning nonlinear dynamical systems**. In: *Automatica* 80 (2017), pp. 189–199.

Adam J. Thorpe, Thomas Lew, Meeko M. K. Oishi and Jia-Jie Zhu. **Data-Driven Chance Constrained Control using Kernel Distribution Embeddings**. In: *Learn. for Dynamics and Ctrl. Conf.* 168 (2022), pp. 790–802.

Michele Xiloyannis, Domenico Chiaradia, Antonio Frisoli and Lorenzo Masia. **Physiological and Kinematic Effects of a Soft Exosuit on Arm Movements**. In: *Journal of NeuroEngineering and Rehabilitation* 16 (2019).

Timeline



$\mathbb{D} = \{\boldsymbol{u}_t, \boldsymbol{y}_t\}$
OCP
$p(\boldsymbol{x}_{-T})$
$-T$
-1
0
H

Scenario Generation

Goal: Generate scenarios $\boldsymbol{\delta}^{[1:N]}$ using the observations $\mathbb{D}$

Algorithm: Scenario Generation

For $n = 1, ..., N$:

1. Sample $\{\boldsymbol{\theta}, \boldsymbol{x}_{\text{-}T:\text{-}1}\}^{[n]}$ from $p\left(\boldsymbol{\theta}, \boldsymbol{x}_{\text{-}T:\text{-}1} \mid \mathbb{D}\right)$ using PMCMC [Lefringhausen+ 2024].
2. Sample $\boldsymbol{v}_t^{[n]}$ from $\boldsymbol{\mathcal{V}}_{\boldsymbol{\theta}^{[n]}}$ and $\boldsymbol{w}_t^{[n]}$ from $\boldsymbol{\mathcal{W}}_{\boldsymbol{\theta}^{[n]}}$ for $t = \text{-}1, ..., H$
3. Set $\boldsymbol{x}_0^{[n]} = \boldsymbol{f}_{\boldsymbol{\theta}^{[n]}}\left(\boldsymbol{x}_{\text{-}1}^{[n]}, \boldsymbol{u}_{\text{-}1}\right) + \boldsymbol{v}_{\text{-}1}^{[n]}$

Output: Scenarios $\boldsymbol{\delta}^{[1:N]} = \{\boldsymbol{\theta}, \boldsymbol{x}_0, \boldsymbol{v}_{0:H}, \boldsymbol{w}_{0:H}\}^{[1:N]}$

Bootstrap Construction

Algorithm: Bootstrap MMD ambiguity set

1. $\boldsymbol{K} \leftarrow \textit{kernel}(\boldsymbol{\delta}, \boldsymbol{\delta})$
2. **For** $m = 1, ..., B$
3. $\quad I \leftarrow N$ numbers from $\{1, \ldots N\}$ with replacement
4. $\quad K_x \leftarrow \sum_{i,j=1}^{N} K_{ij},\ K_y \leftarrow \sum_{i,j \in I} K_{ij},\ K_{xy} \leftarrow \sum_{j \in I} \sum_{i=1}^{N} K_{ij}$
5. $\quad \mathrm{MMD}[m] \leftarrow \frac{1}{N^2} \left(K_x + K_y - 2K_{xy}\right)$
6. **End For**
7. $\mathrm{MMD} \leftarrow \textit{sort}(\mathrm{MMD})$
8. $\varepsilon \leftarrow \mathrm{MMD}[\textit{ceil}(B\beta)]$

Output: Gram matrix $\boldsymbol{K}$, Radius of MMD ambiguity set ε

$B = 1000, \beta = 0.95$

Maximum Mean Discrepancy (MMD)

Maximum Mean Discrepancy

$$\begin{aligned}\mathsf{MMD}(\tilde{P}, P_N) &= ||\mu_{\tilde{P}} - \mu_{P_N}||_{\mathcal{H}} \\ &= \mathsf{E}_{x,x'\sim\tilde{P}}[k(x,x')] + \mathsf{E}_{y,y'\sim P_N}[k(y,y')] - 2\mathsf{E}_{x\sim\tilde{P},y\sim P_N}[k(x,y)]\end{aligned}$$

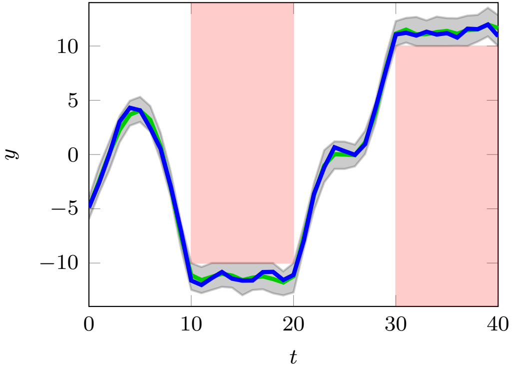
(Biased) MMD estimator

$$\widehat{\mathsf{MMD}}(P,Q) = \frac{1}{N^2}\sum_{i,j=1}^{N} k(\boldsymbol{\delta}^{[i]}, \boldsymbol{\delta}^{[j]}) + k(\tilde{\boldsymbol{\delta}}^{[i]}, \tilde{\boldsymbol{\delta}}^{[j]}) - 2k(\boldsymbol{\delta}^{[i]}, \tilde{\boldsymbol{\delta}}^{[j]})$$

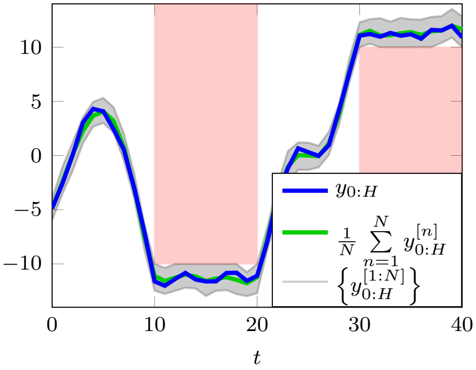
Optimal Control with Constrained Outputs

Number of scenarios used for optimization: $N = 200$

Scenario Approach



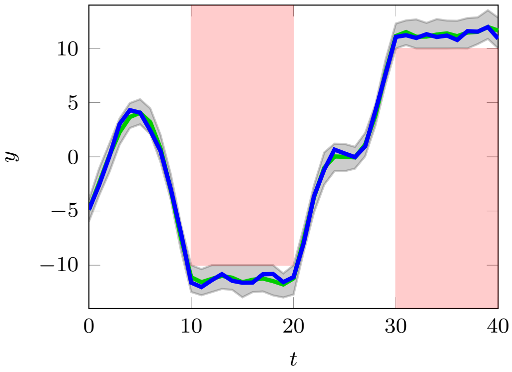
Kernel Approach ($\alpha = 0.01$)



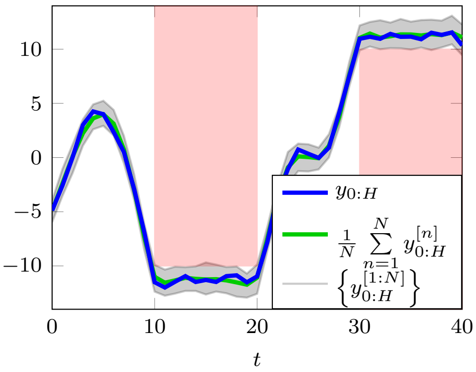
Optimal Control with Constrained Outputs

Number of scenarios used for optimization: $N = 200$

Scenario Approach



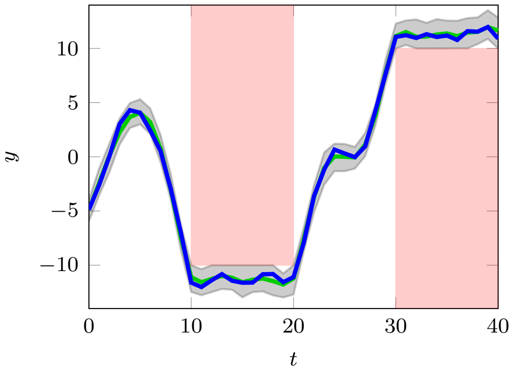
Kernel Approach ($\alpha = 0.2$)



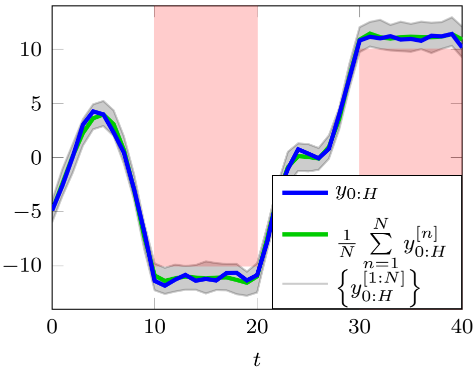
Optimal Control with Constrained Outputs

Number of scenarios used for optimization: $N = 200$

Scenario Approach

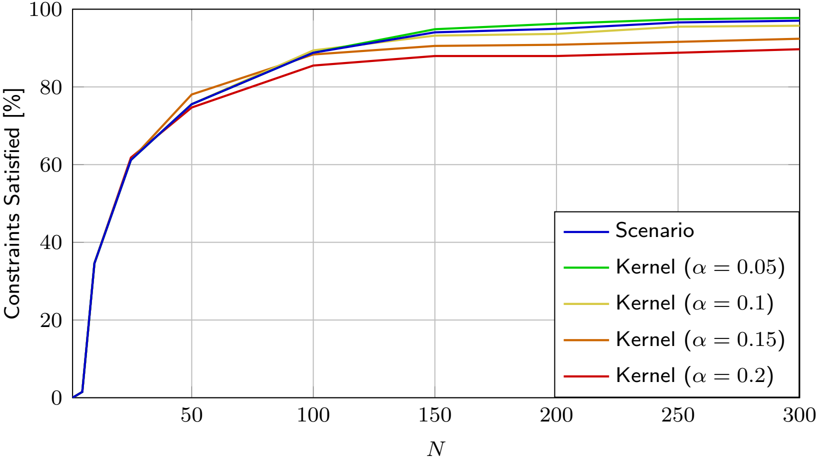


Kernel Approach ($\alpha = 0.5$)



Successrate of Solution

$N' = 2000$ scenarios used to test $u_{0:H}$



Empirical Distribution

Given: Samples $z_i, i = 1, ..., N$

Empirical Distribution

$$P_N(z) = \frac{1}{N} \sum_{i=1}^{N} \text{dirac}(z - z_i)$$