# 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+ 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:

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

**Problem:** Underlying data distribution is unknown

# Related Works

Particle Gibbs based optimal control **[Lefringhausen+ 2024]**

$\Rightarrow$ Guarantees determined retroactively via scenario theory

Alternative approaches for sampling-based optimization:

- Wasserstein ambiguity [Hota+ 2019]
  - $\Rightarrow$ Constraints limited to affine functions
- Kernel embeddings **[Nemmour+ 2022]**
  [Thorpe+ 2022]

# Scenario Optimization

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

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

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

### Chance Constraints

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

$$\Downarrow$$

### Scenario Approach [Lefringhausen+ 2024]

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

$\Rightarrow$ Risk factor $\alpha$ not considered in optimization

# Maximum Mean Discrepancy (MMD) ambiguity sets

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

**Our Approach:** Replace distribution $P$ with the ambiguity set $\mathcal{P}$

## MMD ambiguity set

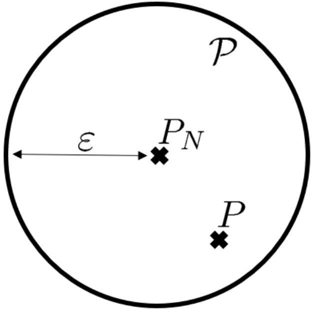
$$\mathcal{P} = \left\{ \tilde{P} : \mathsf{MMD}(\tilde{P}, P_N(\boldsymbol{\delta}^{[1:N]})) \leq \varepsilon \right\}.$$

Radius $\varepsilon$ obtained via bootstrap construction

$$\Downarrow$$

## Expanded Chance-Constraints

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



# Constraint Reformulation

## Feasible Region of chance constraint

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

$$\Downarrow$$

## Reformulated Feasible Region [Nemmour+ 2022]

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

# Hyperparameter Tuning

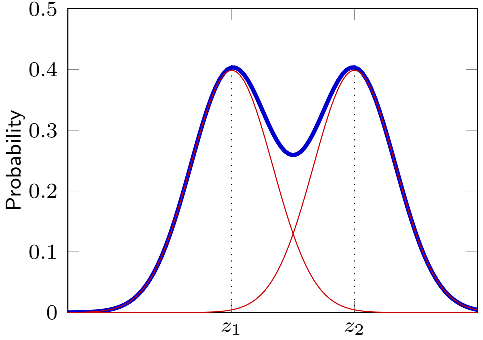
- Gaussian kernels

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

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

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

- Grid search to maximize sum of likelyhoods over test set



# Hyperparameter Tuning

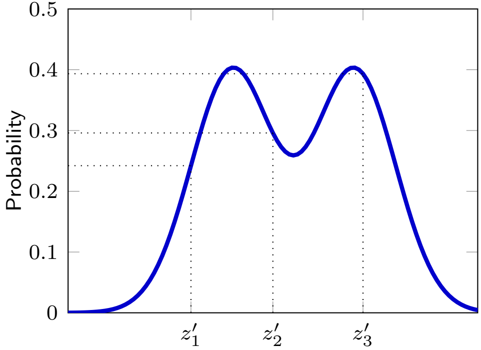
- Gaussian kernels

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

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

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

- Grid search to maximize sum of likelyhoods over test set



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

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

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

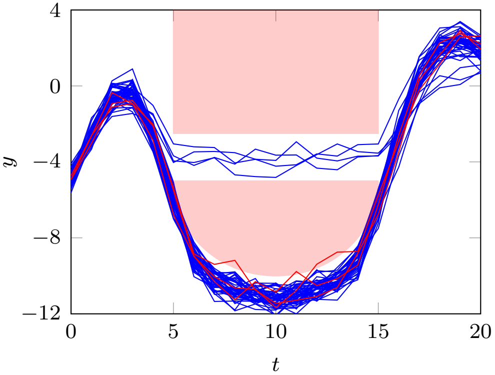
$\boldsymbol{A} \sim \mathcal{MN}(\mathbf{0}, \boldsymbol{Q}, 10\boldsymbol{I}_2)$ [Svensson+ 2017]

$\boldsymbol{x}_{\text{-}T} \sim \mathcal{N}([2, 2]^\mathsf{T}, \boldsymbol{I}_2)$

- 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: $N = 100$

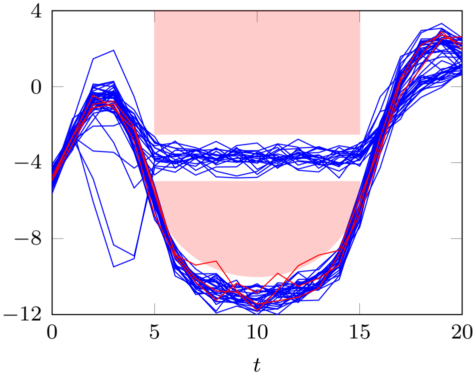
# Risk Tuning

Scenario Approach




Average Cost $J_H = 338.4$

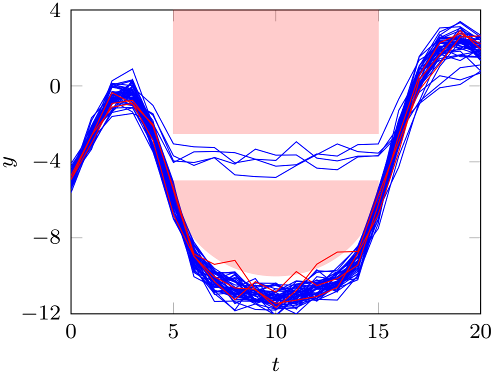
Kernel Approach ($\alpha = 0.1$)



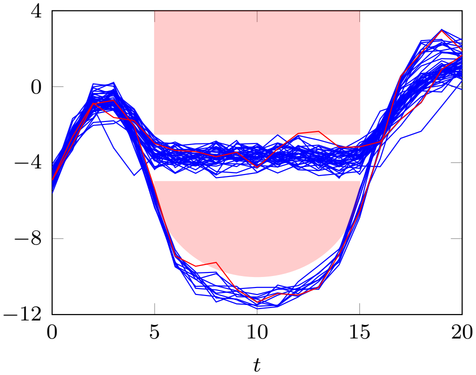

Average Cost $J_H = 255.9$

# Risk Tuning

Scenario Approach




Kernel Approach ($\alpha = 0.2$)

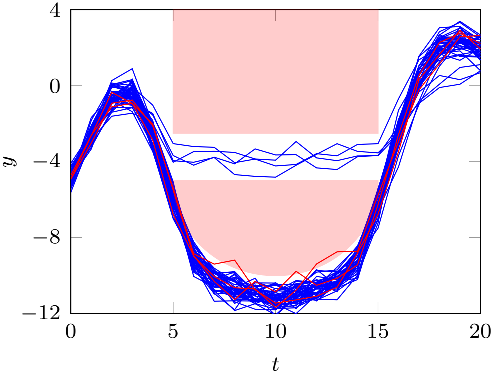



Average Cost $J_H = 338.4$

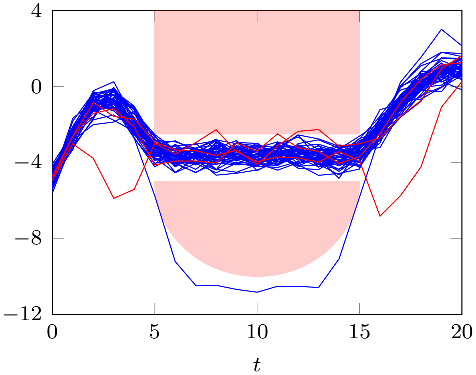
Average Cost $J_H = 129.5$

# Risk Tuning

Scenario Approach




Kernel Approach ($\alpha = 0.3$)




Average Cost $J_H = 338.4$

Average Cost $J_H = 69.9$

# Nonlinear System

**Nonlinear System**

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

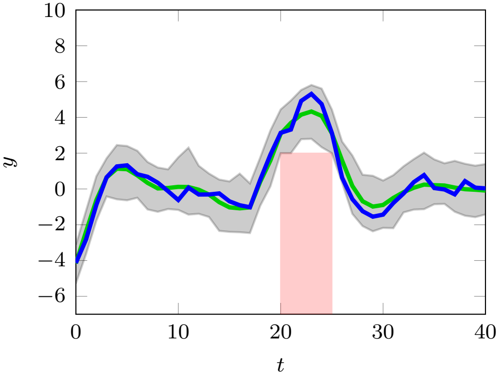
- Known system structure: $\boldsymbol{f}(\boldsymbol{x}, u) = \boldsymbol{A}\,[x_1, x_2, u, \cos(3x_1)x_2, \sin(2x_2)u]^\mathsf{T}$
- Input constraints $|u| \leq 5$
- Number of scenarios used for optimization: $N = 200$

Challenge: Numerical issues complicates solving OCPs with

- Low number of samples $N$
- Low Risk Factor $\alpha$
- High number of constraints

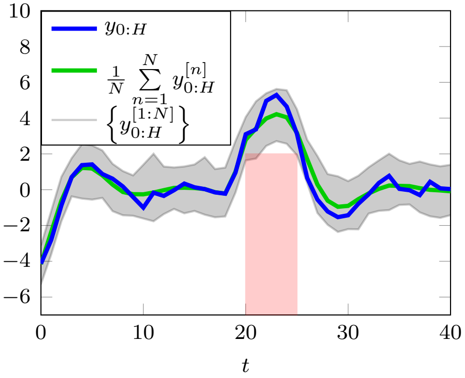
# Nonlinear Optimal Control

Scenario Approach




Cost $J_H = 29.56$

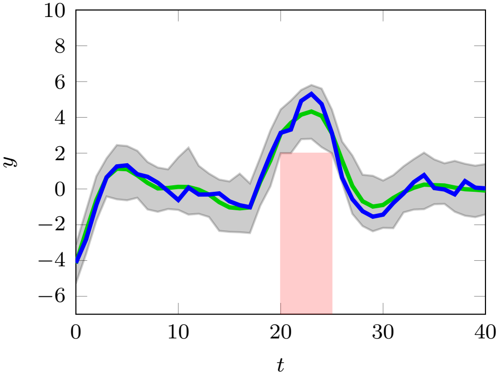
Kernel Approach ($\alpha = 0.2$)




Cost $J_H = 19.45$

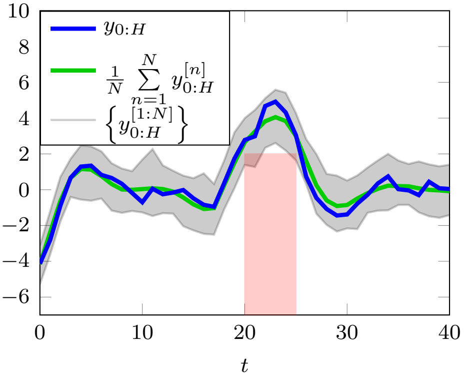
# Nonlinear Optimal Control

Scenario Approach




Cost $J_H = 29.56$

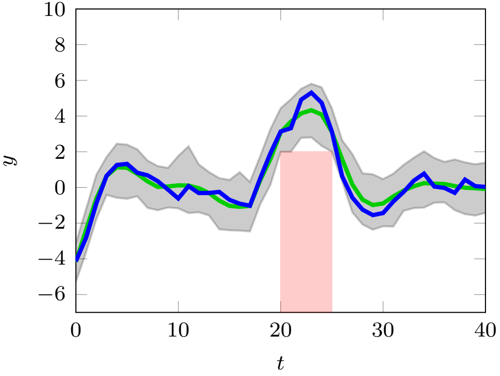
Kernel Approach ($\alpha = 0.4$)




Cost $J_H = 18.77$

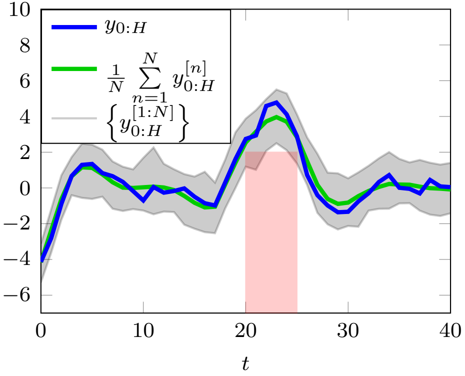
# Nonlinear Optimal Control

Scenario Approach




Cost $J_H = 29.56$

Kernel Approach ($\alpha = 0.6$)

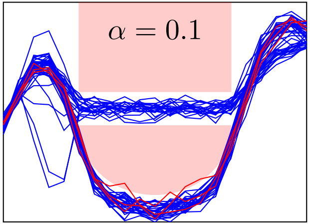



Cost $J_H = 16.2$

# Conclusion

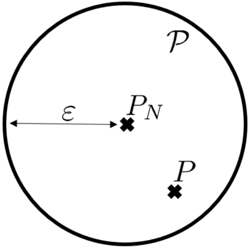
**Summary:**

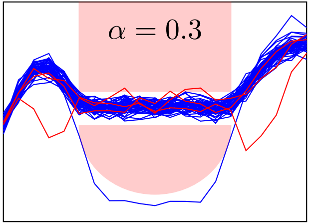
- Sampling from unknown system with particle Gibbs
- Ambiguity set around empirical distribution
- Optimization over ambuiguity set for robust solution




Kernel Embeddings allow for ...

- solving chance-constrained OCPs
- Choosing risk factor $\alpha$



# 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).

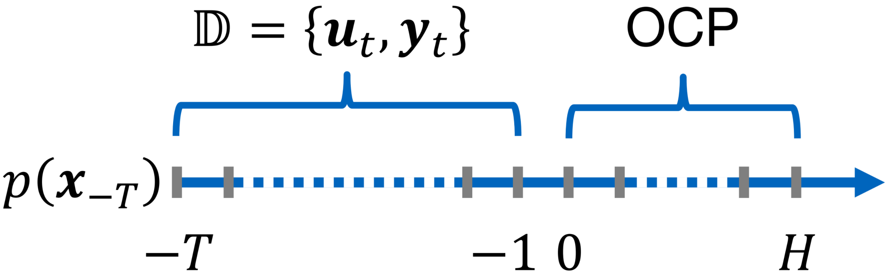
Yassine Nemmour, Heiner Kremer, Bernhard Schoelkopf and Jia-Jie Zhu. **Maximum Mean Discrepancy Distributionally Robust Nonlinear Chance-Constrained Optimization with Finite-Sample Guarantee**. In: *arXiv preprint arXiv:2204.11564* (2022).

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 $\{\boldsymbol{\theta}, \boldsymbol{x}_{-T:-1}\}^{[n]}$ from $p\left(\boldsymbol{\theta}, \boldsymbol{x}_{-T:-1} \mid \mathbb{D}\right)$ 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 $t = -1, ..., H$
3. Set $\boldsymbol{x}_0^{[n]} = \boldsymbol{f}_{\boldsymbol{\theta}^{[n]}}\left(\boldsymbol{x}_{-1}^{[n]}, \boldsymbol{u}_{-1}\right) + \boldsymbol{v}_{-1}^{[n]}$

**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. $\boldsymbol{K} \leftarrow$ *kernel*$(\boldsymbol{\delta}, \boldsymbol{\delta})$
2. **For** $m = 1, ..., B$
3. $\quad I \leftarrow N$ numbers from $\{1, \dots N\}$ 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. $\mathrm{MMD} \leftarrow$ *sort*$(\mathrm{MMD})$
8. $\varepsilon \leftarrow \mathrm{MMD}[$*ceil*$(B\beta)]$

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

$B = 1000, \beta = 0.95$

# Maximum Mean Discrepancy (MMD)

## Maximum Mean Discrepancy

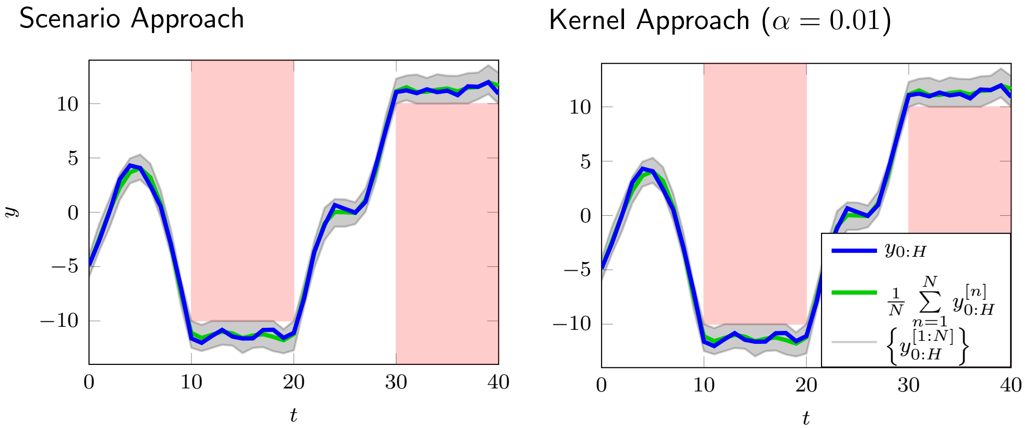
$$\begin{aligned}\text{MMD}(\tilde{P}, P_N) &= ||\mu_{\tilde{P}} - \mu_{P_N}||_{\mathcal{H}} \\ &= \mathsf{E}_{x,x' \sim \tilde{P}}[k(x,x')] + \mathsf{E}_{y,y' \sim P_N}[k(y,y')] - 2\mathsf{E}_{x \sim \tilde{P}, y \sim P_N}[k(x,y)]\end{aligned}$$

## (Biased) MMD estimator

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

# Optimal Control with Constrained Outputs

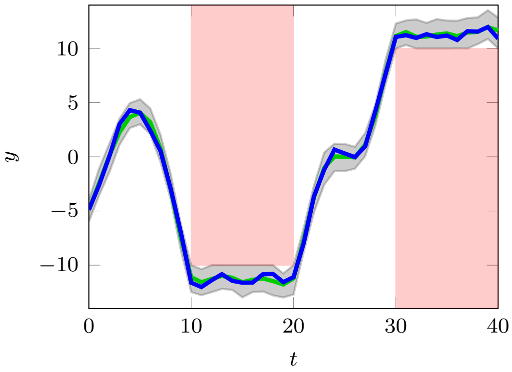
Number of scenarios used for optimization: $N = 200$



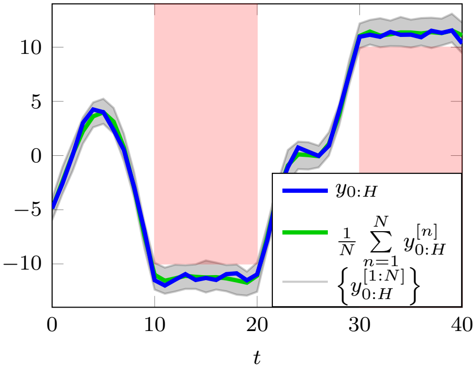
# Optimal Control with Constrained Outputs

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

Scenario Approach



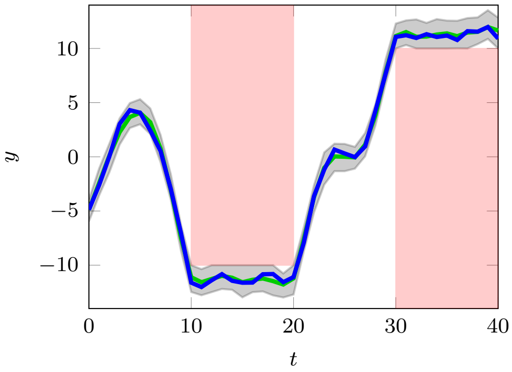

Kernel Approach ($\alpha = 0.2$)



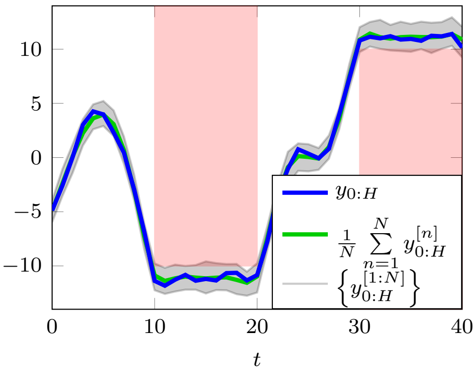
# Optimal Control with Constrained Outputs

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

Scenario Approach

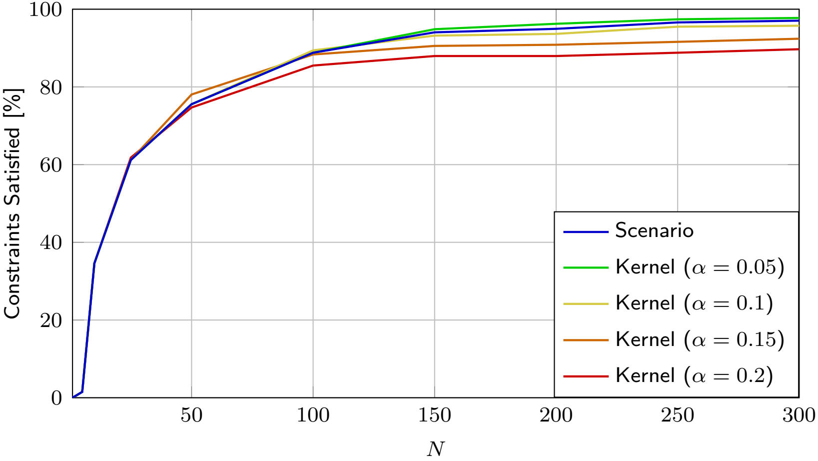



Kernel Approach ($\alpha = 0.5$)



# Successrate of Solution

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

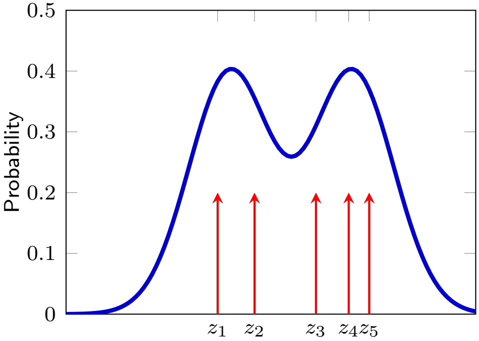


# Empirical Distribution

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

**Empirical Distribution**

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



# Constraint Reformulation

**Feasible Region of chance constraint**

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

$$\Downarrow$$

**Reformulated Feasible Region [Nemmour+ 2022]**

$$Z := \left\{ \boldsymbol{u}_{0:H} \in \mathcal{U}^{H+1} : \begin{array}{l} g_0 + \frac{1}{N} \sum_{n=1}^{N} g(\boldsymbol{\delta}^{[n]}) + \varepsilon ||g||_{\mathcal{H}} \leq \alpha \\ 1(h(\boldsymbol{u}_{0:H}, \boldsymbol{x}_{0:H}^{[n]}, \boldsymbol{y}_{0:H}^{[n]}) > 0) \leq g_0 + g(\boldsymbol{\delta}^{[n]}), \; n = 1, ..., N \\ g_0 \in \mathbb{R}, g \in \mathcal{H} \end{array} \right\}$$

# Value at Risk (VaR)

**VaR**

$$\mathsf{VaR}^{P}_{1-\alpha} := \inf \left\{ t' \in \mathbb{R} : P\left[ h(\boldsymbol{u}_{0:H}, \boldsymbol{x}_{0:H}, \boldsymbol{y}_{0:H}) \le t' \right] \ge 1 - \alpha \right\}.$$

**Reformulated Constraint**

$$\mathsf{VaR}^{P}_{1-\alpha} \le 0 \iff P\left[ h(\boldsymbol{u}_{0:H}, \boldsymbol{x}_{0:H}, \boldsymbol{y}_{0:H}) \le 0 \right] \ge 1 - \alpha.$$

# Conditional Value at Risk (CVaR)

**CVaR**

$$\mathsf{CVaR}^{P}_{1-\alpha} := \inf_{t\in\mathbb{R}} \mathsf{E}_{P}\left[[\tilde{h}(\boldsymbol{u}_{0:H},\boldsymbol{\delta}) + t']_{+} - t'\alpha\right]$$

**Reformulated Constraint**

$$\begin{aligned}\sup_{\tilde{P}\in\mathcal{P}} \mathsf{CVaR}^{\tilde{P}}_{1-\alpha} &= \sup_{\tilde{P}\in\mathcal{P}} \inf_{t\in\mathbb{R}} \mathsf{E}_{\tilde{P}}\left[[\tilde{h}(\boldsymbol{u}_{0:H},\boldsymbol{\delta}) + t']_{+} - t'\alpha\right] \\ &= \inf_{t\in\mathbb{R}} \sup_{\tilde{P}\in\mathcal{P}} \mathsf{E}_{\tilde{P}}\left[[\tilde{h}(\boldsymbol{u}_{0:H},\boldsymbol{\delta}) + t']_{+} - t'\alpha\right] \leq 0.\end{aligned}$$