Safe exploration in reproducing kernel Hilbert spaces

Research talk

Abdullah Tokmak[1,2]

[1] Aalto University, Espoo, Finland
[2] Uppsala University, Uppsala, Sweden

October 15, 2024

Aalto University
School of Electrical Engineering

UPPSALA UNIVERSITET

Motivational example

Introduction

Goal

Optimize control parameters of safety-critical real-world systems.

Introduction

Goal

Optimize control parameters of safety-critical real-world systems.

- Unknown reward function $f\colon \mathcal{A} \to \mathbb{R}$
- Control policy parameters $a \in \mathcal{A}$
- We require **sample efficiency** and **safety guarantees**

Introduction

Goal

Optimize control parameters of safety-critical real-world systems.

- Unknown reward function $f\colon \mathcal{A} \to \mathbb{R}$
- Control policy parameters $a \in \mathcal{A}$
- We require **sample efficiency** and **safety guarantees**

Solvable using classic reinforcement learning (RL)?

Classic RL struggles with both sample efficiency and safety guarantees.

Gaussian process (GP) regression

- GPs to model unknown reward function f from samples
- GP characterized by **kernel** k: **Mean prediction** μ_t, **standard deviation** σ_t

Gaussian process (GP) regression

- GPs to model unknown reward function f from samples
- GP characterized by **kernel** k: **Mean prediction** μ_t, **standard deviation** σ_t

Regularity assumption

The reward function f is a member of the **reproducing kernel Hilbert space (RKHS)** of the chosen kernel k, i.e., $f \in H_k$.

Gaussian process (GP) regression

- GPs to model unknown reward function f from samples
- GP characterized by **kernel** k: **Mean prediction** μ_t, **standard deviation** σ_t

Regularity assumption

The reward function f is a member of the **reproducing kernel Hilbert space (RKHS)** of the chosen kernel k, i.e., $f \in H_k$.

Regularity assumption

An upper bound B on the RKHS norm $\|f\|_k$, i.e., $B \geq \|f\|_k$, is known a priori.

Gaussian process (GP) regression

- GPs to model unknown reward function f from samples
- GP characterized by **kernel** k: **Mean prediction** μ_t, **standard deviation** σ_t

Regularity assumption

The reward function f is a member of the **reproducing kernel Hilbert space (RKHS)** of the chosen kernel k, i.e., $f \in H_k$.

Regularity assumption

An upper bound B on the RKHS norm $\|f\|_k$, i.e., $B \geq \|f\|_k$, is known a priori.

GP confidence intervals

$|f(a) - \mu_t(a)| \leq (B + \text{“data-term”})\,\sigma_t(a)$

Safe Bayesian optimization (BO) with GPs

Control policy optimization problem

$\max_{a \in \mathcal{A}} f(a) \quad \text{subject to } f(a) \geq h$

GP confidence intervals

$|f(a) - \mu_t(a)| \leq (B + \text{“data-term”})\, \sigma_t(a)$

[1] Y. Sui, A. Gotovos, J. W. Burdick, A. Krause, “Safe exploration for optimization with Gaussian processes,” 2015.

Safe Bayesian optimization (BO) with GPs

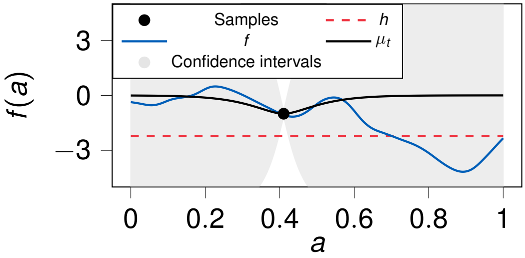
Control policy optimization problem

$\max_{a \in \mathcal{A}} f(a) \quad \text{subject to } f(a) \geq h$

GP confidence intervals

$|f(a) - \mu_t(a)| \leq (B + \text{“data-term”})\, \sigma_t(a)$

SafeOpt[1] $(t = 0, B = \|f\|_k)$



[1] Y. Sui, A. Gotovos, J. W. Burdick, A. Krause,“Safe exploration for optimization with Gaussian processes,” 2015.

Safe Bayesian optimization (BO) with GPs

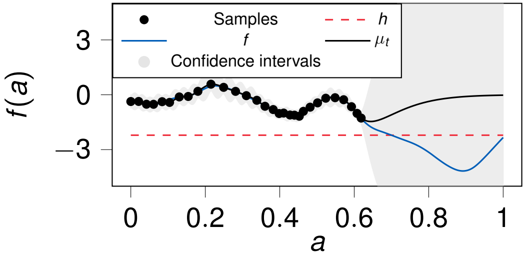
Control policy optimization problem

$\max_{a \in \mathcal{A}} f(a) \quad \text{subject to } f(a) \geq h$

GP confidence intervals

$|f(a) - \mu_t(a)| \leq (B + \text{"data-term"})\, \sigma_t(a)$

SAFEOPT[1] $(t = 30,\ B = \|f\|_k)$



[1] Y. Sui, A. Gotovos, J. W. Burdick, A. Krause, "Safe exploration for optimization with Gaussian processes," 2015.

Safe Bayesian optimization (BO) with GPs

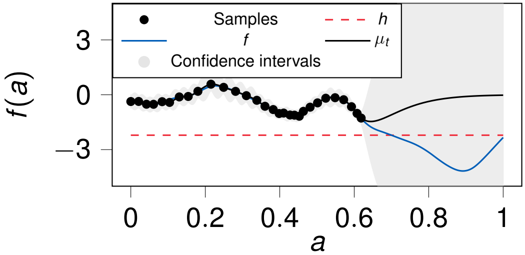
Control policy optimization problem

$\max_{a\in\mathcal{A}} f(a)$ subject to $f(a) \geq h$

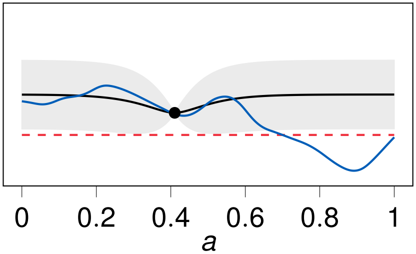
GP confidence intervals

$|f(a) - \mu_t(a)| \leq (B + \text{“data-term”})\, \sigma_t(a)$

SafeOpt[1] $(t = 30,\ B = \|f\|_k)$



SafeOpt[1] $(t = 0,\ B < \|f\|_k)$



[1] Y. Sui, A. Gotovos, J. W. Burdick, A. Krause,“Safe exploration for optimization with Gaussian processes,” 2015.

Safe Bayesian optimization (BO) with GPs

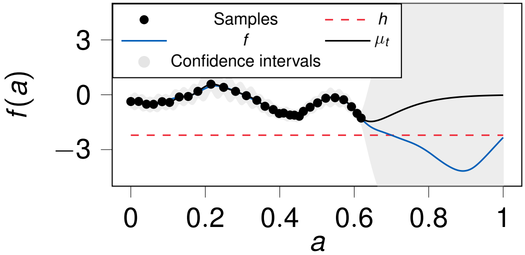
Control policy optimization problem

$\max_{a \in \mathcal{A}} f(a) \quad \text{subject to } f(a) \geq h$

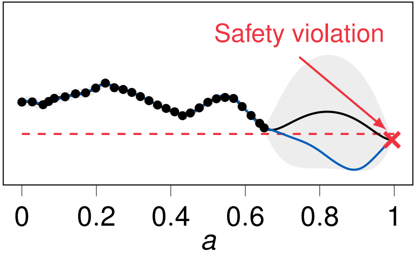
GP confidence intervals

$|f(a) - \mu_t(a)| \leq (B + \text{"data-term"})\, \sigma_t(a)$

SafeOpt[1] $(t = 30,\ B = \|f\|_k)$



SafeOpt[1] $(t = 30,\ B < \|f\|_k)$



[1] Y. Sui, A. Gotovos, J. W. Burdick, A. Krause,"Safe exploration for optimization with Gaussian processes," 2015.

RKHS norm assumption in safe BO

Regularity assumption

Most safe BO algorithms require an upper bound B on the RKHS norm ($B \geq \|f\|_k$) a priori.

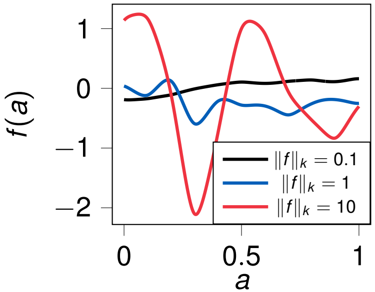
[2] Tokmak, Fiedler, Zeilinger, Trimpe, Köhler, "Automatic nonlinear MPC approximation with closed-loop guarantees," submitted to IEEE TAC, 2023.

RKHS norm assumption in safe BO

Regularity assumption

Most safe BO algorithms require an upper bound B on the RKHS norm ($B \geq \|f\|_k$) a priori.

- RKHS norm $\|f\|_k$ characterizes **"smoothness"** of function f



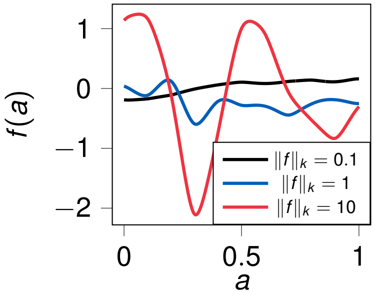
[2] Tokmak, Fiedler, Zeilinger, Trimpe, Köhler, "Automatic nonlinear MPC approximation with closed-loop guarantees," submitted to IEEE TAC, 2023.

RKHS norm assumption in safe BO

Regularity assumption

Most safe BO algorithms require an upper bound B on the RKHS norm ($B \geq \|f\|_k$) a priori.

- RKHS norm $\|f\|_k$ characterizes "**smoothness**" of function f
- Tight upper bound for **practicality**



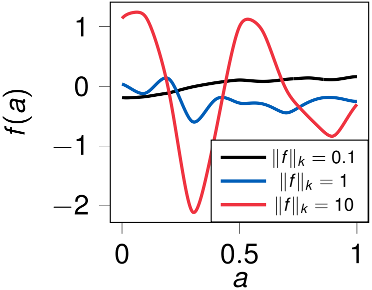
[2] Tokmak, Fiedler, Zeilinger, Trimpe, Köhler, "Automatic nonlinear MPC approximation with closed-loop guarantees," submitted to IEEE TAC, 2023.

RKHS norm assumption in safe BO

Regularity assumption

Most safe BO algorithms require an upper bound B on the RKHS norm ($B \geq \|f\|_k$) a priori.

- RKHS norm $\|f\|_k$ characterizes **“smoothness”** of function f
- Tight upper bound for **practicality**
- It is **unclear how to upper bound the RKHS norm** of unknown functions[2]



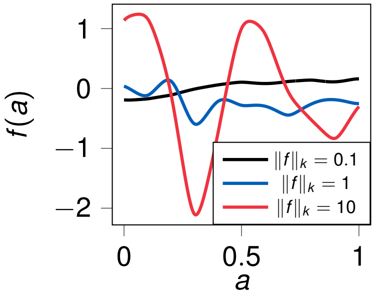
[2] Tokmak, Fiedler, Zeilinger, Trimpe, Köhler, “Automatic nonlinear MPC approximation with closed-loop guarantees,” submitted to IEEE TAC, 2023.

RKHS norm assumption in safe BO

Regularity assumption

Most safe BO algorithms require an upper bound B on the RKHS norm ($B \geq \|f\|_k$) a priori.

- RKHS norm $\|f\|_k$ characterizes "**smoothness**" of function f
- Tight upper bound for **practicality**
- It is **unclear how to upper bound the RKHS norm** of unknown functions[2]



Problem definition

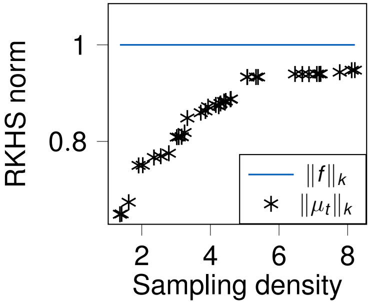
Develop a safe BO algorithm that over-estimates the RKHS norm $\|f\|_k$ with statistical guarantees.

[2] Tokmak, Fiedler, Zeilinger, Trimpe, Köhler, "Automatic nonlinear MPC approximation with closed-loop guarantees," submitted to IEEE TAC, 2023.

′

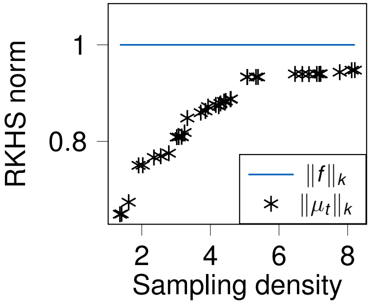
Predicting RKHS norms

- Increasing sampling density: $\mu_t \to f$ and $\|\mu_t\|_k \to \|f\|_k$ **from below**



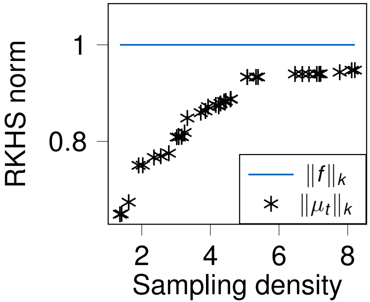
Predicting RKHS norms

- Increasing sampling density: $\mu_t \to f$ and $\|\mu_t\|_k \to \|f\|_k$ **from below**
- We require RKHS norm **over-estimation**



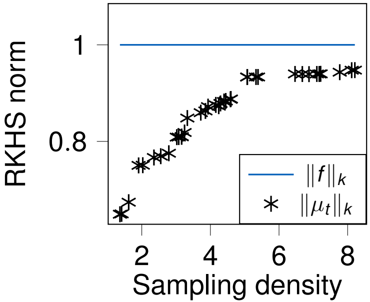
Predicting RKHS norms

- Increasing sampling density: $\mu_t \to f$ and $\|\mu_t\|_k \to \|f\|_k$ **from below**
- We require RKHS norm **over-estimation**
- **Extrapolate** B_t from inputs $\|\mu_t\|_k$ and sampling density
- **Training data** from toy examples
- Extrapolation: RNNs to exploit **sequential** nature of inputs



Predicting RKHS norms

- Increasing sampling density: $\mu_t \to f$ and $\|\mu_t\|_k \to \|f\|_k$ **from below**
- We require RKHS norm **over-estimation**
- **Extrapolate** B_t from inputs $\|\mu_t\|_k$ and sampling density
- **Training data** from toy examples
- Extrapolation: RNNs to exploit **sequential** nature of inputs

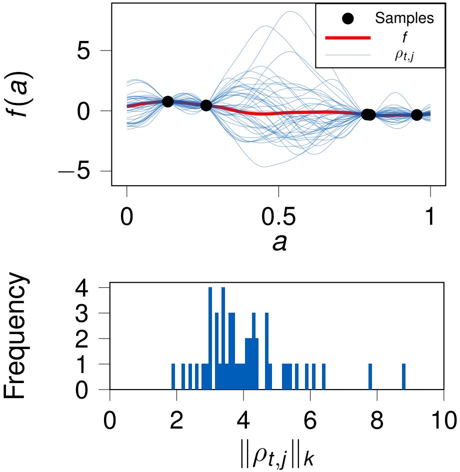


Theoretical guarantees instead of only heuristics

How do we get theoretical guarantees on the RKHS norm over-estimation?

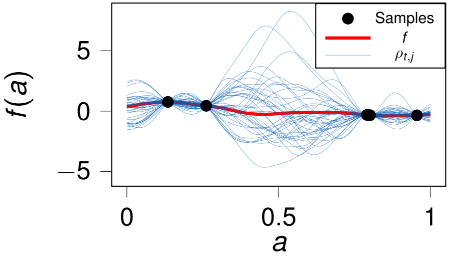
Random RKHS functions

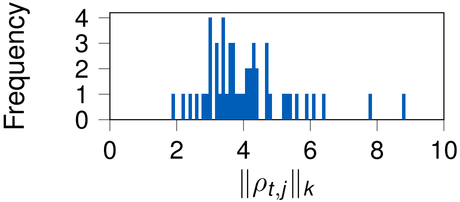
- Compute random RKHS functions $\rho_{t,j},\ j \in \{1, \ldots, \mathrm{m}\}$ with kernel k
- Random RKHS functions $\rho_{t,j}$ capture the behavior of reward function f



Random RKHS functions

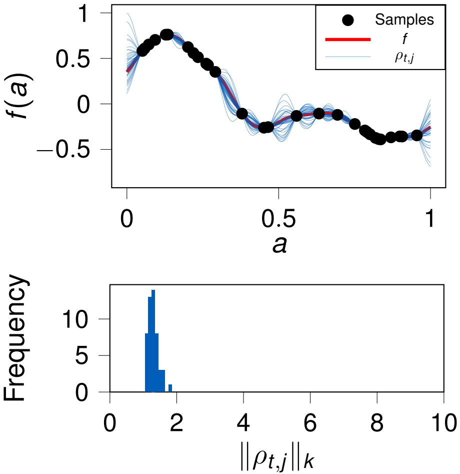
- Compute random RKHS functions $\rho_{t,j},\ j \in \{1, \ldots, \mathrm{m}\}$ with kernel k
- Random RKHS functions $\rho_{t,j}$ capture the behavior of reward function f
- Increasing sampling density: $\rho_{t,j}, \|\rho_{t,j}\|_k \to f, \|f\|_k$





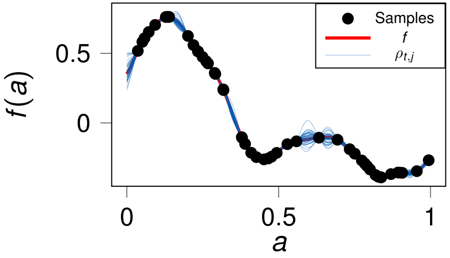
Random RKHS functions

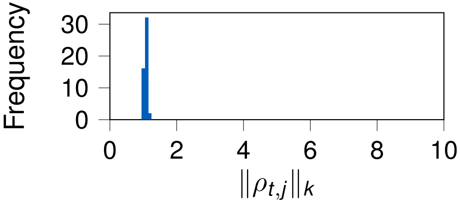
- Compute random RKHS functions $\rho_{t,j},\ j \in \{1, \ldots, \mathrm{m}\}$ with kernel k
- Random RKHS functions $\rho_{t,j}$ capture the behavior of reward function f
- Increasing sampling density: $\rho_{t,j}, \|\rho_{t,j}\|_k \to f, \|f\|_k$



Random RKHS functions

- Compute random RKHS functions $\rho_{t,j},\ j \in \{1, \ldots, \mathrm{m}\}$ with kernel k
- Random RKHS functions $\rho_{t,j}$ capture the behavior of reward function f
- Increasing sampling density: $\rho_{t,j}, \|\rho_{t,j}\|_k \to f, \|f\|_k$





Statistical guarantees

Regularity assumptions

- Reward function f is a member of the RKHS of kernel k
- $\|f\|_k \leq \lim_{s\to\infty} \frac{1}{s}\sum_{j=1}^{s} \|\rho_{t,j}\|_k$

[3] W. Hoeffding, "Probability inequalities for sums of bounded random variables," The Annals of Statistics, 1962

Statistical guarantees

Regularity assumptions

- Reward function f is a member of the RKHS of kernel k
- $\|f\|_k \leq \lim_{s\to\infty} \frac{1}{s} \sum_{j=1}^{s} \|\rho_{t,j}\|_k$

Theorem

Over-estimation of RKHS norm $B_t \geq \|f\|_k$ is probably approximate correct (PAC) $\forall t \geq 1$.

[3] W. Hoeffding, "Probability inequalities for sums of bounded random variables," The Annals of Statistics, 1962

Statistical guarantees

Regularity assumptions

- Reward function f is a member of the RKHS of kernel k
- $\|f\|_k \leq \lim_{s\to\infty} \frac{1}{s}\sum_{j=1}^{s} \|\rho_{t,j}\|_k$

Theorem

Over-estimation of RKHS norm $B_t \geq \|f\|_k$ is probably approximate correct (PAC) $\forall t \geq 1$.

Proof sketch

- $B_t \leftarrow \max\{\text{RNN prediction}, \frac{1}{m}\sum_{j=1}^{m} \|\rho_{t,j}\|_k + \text{“safety-term”}\}$
- Statistical guarantees through Hoeffding's inequality[3]

[3] W. Hoeffding, "Probability inequalities for sums of bounded random variables," The Annals of Statistics, 1962

RKHS norm over-estimation as optimization problem

Nontrivial optimization problem

Minimize $B_t \in \mathbb{R}_+$ subject to $B_t \geq \|f\|_k$.

RKHS norm over-estimation as optimization problem

Nontrivial optimization problem

Minimize $B_t \in \mathbb{R}_+$ subject to $B_t \geq \|f\|_k$.

- Optimization problem is **unsolvable** as constraint $\|f\|_k$ is unknown

RKHS norm over-estimation as optimization problem

Nontrivial optimization problem

Minimize $B_t \in \mathbb{R}_+$ subject to $B_t \geq \|f\|_k$.

- Optimization problem is **unsolvable** as constraint $\|f\|_k$ is unknown

Impractical optimization problem

Minimize $B_t \in \mathbb{R}_+$ subject to $B_t \geq \|\rho_t\|_k, \quad \forall \rho_t \in H_k$.

RKHS norm over-estimation as optimization problem

Nontrivial optimization problem

Minimize $B_t \in \mathbb{R}_+$ subject to $B_t \geq \|f\|_k$.

- Optimization problem is **unsolvable** as constraint $\|f\|_k$ is unknown

Impractical optimization problem

Minimize $B_t \in \mathbb{R}_+$ subject to $B_t \geq \|\rho_t\|_k, \quad \forall \rho_t \in H_k$.

- Optimization problem is **impractical** as $B_t = \infty$

RKHS norm over-estimation as optimization problem

Nontrivial optimization problem

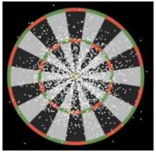
Minimize $B_t \in \mathbb{R}_+$ subject to $B_t \geq \|f\|_k$.

- Optimization problem is **unsolvable** as constraint $\|f\|_k$ is unknown

Impractical optimization problem

Minimize $B_t \in \mathbb{R}_+$ subject to $B_t \geq \|\rho_t\|_k, \quad \forall \rho_t \in H_k$.

- Optimization problem is **impractical** as $B_t = \infty$
- **Analogy**: Minimum radius on darts board that contains all points



https://www.3blue1brown.com/lessons/hyperdarts

RKHS norm over-estimation as optimization problem

Nontrivial optimization problem

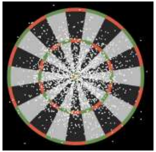
Minimize $B_t \in \mathbb{R}_+$ subject to $B_t \geq \|f\|_k$.

- Optimization problem is **unsolvable** as constraint $\|f\|_k$ is unknown

Impractical optimization problem

Minimize $B_t \in \mathbb{R}_+$ subject to $B_t \geq \|\rho_t\|_k, \quad \forall \rho_t \in H_k$.

- Optimization problem is **impractical** as $B_t = \infty$
- **Analogy**: Minimum radius on darts board that contains all points
- Can we get better performance with **statistical guarantees**?



https://www.3blue1brown.com/lessons/hyperdarts

RKHS norm over-estimation as optimization problem (2)

Chance-constrained optimization problem

Minimize $B_t \in \mathbb{R}_+$ subject to $B_t \geq \|f\|_k$ with high probability.

[4] M. C. Campi, S. Garatti, "Introduction to the scenario approach," SIAM, 2018.

RKHS norm over-estimation as optimization problem (2)

Chance-constrained optimization problem

Minimize $B_t \in \mathbb{R}_+$ subject to $B_t \geq \|f\|_k$ with high probability.

- Solve chance-constrained optimization problem using scenario approach[4] by fixing m i.i.d. scenarios
- Scenarios: random RKHS functions $\rho_{t,j},\ j \in \{1, \ldots, \mathrm{m}\}$

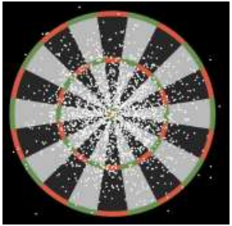
[4] M. C. Campi, S. Garatti, "Introduction to the scenario approach," SIAM, 2018.

RKHS norm over-estimation as optimization problem (2)

Chance-constrained optimization problem

Minimize $B_t \in \mathbb{R}_+$ subject to $B_t \geq \|f\|_k$ with high probability.

- Solve chance-constrained optimization problem using scenario approach[4] by fixing m i.i.d. scenarios
- Scenarios: random RKHS functions $\rho_{t,j},\ j \in \{1, \ldots, \mathrm{m}\}$



`https://www.3blue1brown.com/lessons/hyperdarts`

[4] M. C. Campi, S. Garatti, "Introduction to the scenario approach," SIAM, 2018.

RKHS norm over-estimation as optimization problem (2)

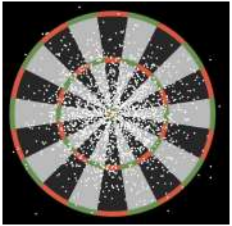
Chance-constrained optimization problem

Minimize $B_t \in \mathbb{R}_+$ subject to $B_t \geq \|f\|_k$ with high probability.

- Solve chance-constrained optimization problem using scenario approach[4] by fixing m i.i.d. scenarios
- Scenarios: random RKHS functions $\rho_{t,j},\ j \in \{1, \ldots, \mathrm{m}\}$

Scenario approach

Minimize $B_t \in \mathbb{R}_+$ subject to $B_t \geq \|\rho_{j,t}\|_k,\ j \in \{1, \ldots, \mathrm{m}\}$.



https://www.3blue1brown.com/lessons/hyperdarts

[4] M. C. Campi, S. Garatti, "Introduction to the scenario approach," SIAM, 2018.

RKHS norm over-estimation as optimization problem (3)

- Some random RKHS functions might be **outliers**, i.e., $\|\rho_{t,j}\|_k \gg \|f\|_k$
- Sampling-and-discarding scenario approach:[5] Trade **feasibility** for **performance**

[5] M. C. Campi, S. Garatti, "A sampling-and-discarding approach to chance-constrained optimization: Feasibility and optimality," Springer, 2011.

RKHS norm over-estimation as optimization problem (3)

- Some random RKHS functions might be **outliers**, i.e., $\|\rho_{t,j}\|_k \gg \|f\|_k$
- Sampling-and-discarding scenario approach:[5] Trade **feasibility** for **performance**
- Sort $\{\rho_{t,j}\}_{j=1}^{m}$ by **ascending RKHS norm** and **discard** r constraints $\{\rho_{t,j}\}_{j=m-r+1}^{m}$

[5] M. C. Campi, S. Garatti, "A sampling-and-discarding approach to chance-constrained optimization: Feasibility and optimality," Springer, 2011.

RKHS norm over-estimation as optimization problem (3)

- Some random RKHS functions might be **outliers**, i.e., $\|\rho_{t,j}\|_k \gg \|f\|_k$
- Sampling-and-discarding scenario approach:[5] Trade **feasibility** for **performance**
- Sort $\{\rho_{t,j}\}_{j=1}^m$ by **ascending RKHS norm** and **discard** r constraints $\{\rho_{t,j}\}_{j=m-r+1}^m$

Sampling-and-discarding scenario approach

Min. $B_t \in \mathbb{R}_+$ s.t. $B_t \geq \|\rho_{t,j}\|_k,\ j \in \{1, \dots, m-r\} \wedge B_t < \|\rho_{t,j}\|_k,\ j \in \{m-r+1, \dots, m\}$.

[5] M. C. Campi, S. Garatti, "A sampling-and-discarding approach to chance-constrained optimization: Feasibility and optimality," Springer, 2011.

RKHS norm over-estimation as optimization problem (3)

- Some random RKHS functions might be **outliers**, i.e., $\|\rho_{t,j}\|_k \gg \|f\|_k$
- Sampling-and-discarding scenario approach:[5] Trade **feasibility** for **performance**
- Sort $\{\rho_{t,j}\}_{j=1}^m$ by **ascending RKHS norm** and **discard** r constraints $\{\rho_{t,j}\}_{j=m-r+1}^m$

Sampling-and-discarding scenario approach

Min. $B_t \in \mathbb{R}_+$ s.t. $B_t \geq \|\rho_{t,j}\|_k,\ j \in \{1, \ldots, m-r\} \wedge B_t < \|\rho_{t,j}\|_k,\ j \in \{m-r+1, \ldots, m\}$.

Theorem

Over-estimation of RKHS norm $B_t \geq \|f\|_k$ is PAC $\forall t \geq 1$.

[5] M. C. Campi, S. Garatti, "A sampling-and-discarding approach to chance-constrained optimization: Feasibility and optimality," Springer, 2011.

RKHS norm over-estimation as optimization problem (3)

- Some random RKHS functions might be **outliers**, i.e., $\|\rho_{t,j}\|_k \gg \|f\|_k$
- Sampling-and-discarding scenario approach:[5] Trade **feasibility** for **performance**
- Sort $\{\rho_{t,j}\}_{j=1}^{m}$ by **ascending RKHS norm** and **discard** r constraints $\{\rho_{t,j}\}_{j=m-r+1}^{m}$

Sampling-and-discarding scenario approach

Min. $B_t \in \mathbb{R}_+$ s.t. $B_t \geq \|\rho_{t,j}\|_k,\ j \in \{1, \dots, m-r\} \wedge B_t < \|\rho_{t,j}\|_k,\ j \in \{m-r+1, \dots, m\}$.

Theorem

Over-estimation of RKHS norm $B_t \geq \|f\|_k$ is PAC $\forall t \geq 1$.

Proof sketch

- Sampling-and-discarding scenario approach: $B_t \leftarrow \|\rho_{t,m-r}\|_k$
- RNN introduces lower bound: $B_t \leftarrow \max\{\text{RNN prediction}, \|\rho_{t,m-r}\|_k\}$

[5] M. C. Campi, S. Garatti, "A sampling-and-discarding approach to chance-constrained optimization: Feasibility and optimality," Springer, 2011.

Contrasting both approaches

Assumption (Scenario approach)

RKHS norms $\|\rho_{t,j}\|_k$, $j \in \{1, \ldots, \mathrm{m}\}$ and $\|f\|_k$ are i.i.d. samples from the same (potentially unknown) probability space.

Assumption (Hoeffding's inequality)

$\|f\|_k \leq \lim_{s\to\infty} \frac{1}{s} \sum_{j=1}^{s} \|\rho_{t,j}\|_k$

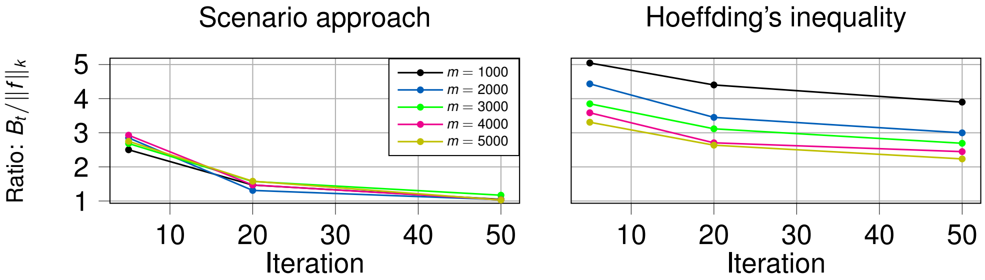
Contrasting both approaches

Assumption (Scenario approach)

RKHS norms $\|\rho_{t,j}\|_k$, $j \in \{1, \dots, \mathrm{m}\}$ and $\|f\|_k$ are i.i.d. samples from the same (potentially unknown) probability space.

Assumption (Hoeffding's inequality)

$\|f\|_k \leq \lim_{s \to \infty} \frac{1}{s} \sum_{j=1}^{s} \|\rho_{t,j}\|_k$



Contrasting both approaches

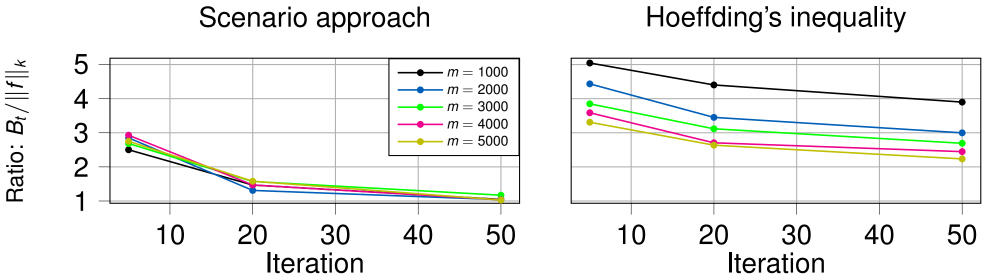
Assumption (Scenario approach)

RKHS norms $\|\rho_{t,j}\|_k$, $j \in \{1, \ldots, \mathrm{m}\}$ and $\|f\|_k$ are i.i.d. samples from the same (potentially unknown) probability space.

Assumption (Hoeffding's inequality)

$\|f\|_k \leq \lim_{s\to\infty} \frac{1}{s} \sum_{j=1}^{s} \|\rho_{t,j}\|_k$

$\Rightarrow$ Hoeffding assumption **interpretability**?



Safe BO with RKHS norm over-estimation

Problem definition

Develop a safe BO algorithm that estimates the RKHS norm $\|f\|_k$ with guarantees.

Safe BO with RKHS norm over-estimation

Problem definition

Develop a safe BO algorithm that estimates the RKHS norm $\|f\|_k$ with guarantees.

Theorem *(Safety)*

Safe BO algorithm with RKHS norm over-estimation ensures safety with high probability.

Proof sketch *(Safety)*

Combine safety proof of SafeOpt with RKHS norm over-estimation.

Local interpretation of the RKHS norm

- Safe exploration for optimization:
  Restricted to **sub-space** of domain

GP confidence intervals

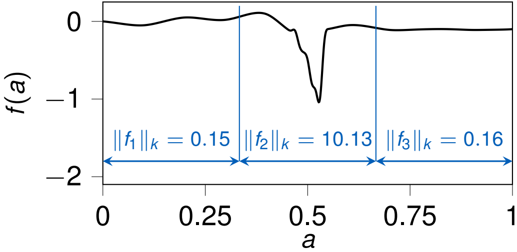
$$|f(a) - \mu_t(a)| \leq (B_t + \text{“data-term”})\, \sigma_t(a)$$

Local interpretation of the RKHS norm

- Safe exploration for optimization: Restricted to **sub-space** of domain
- Exploit **local "smoothness"** to allow for more **optimistic exploration**

GP confidence intervals

$$|f(a) - \mu_t(a)| \leq (B_t + \text{"data-term"})\, \sigma_t(a)$$

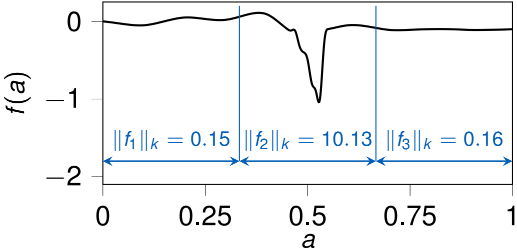


Local interpretation of the RKHS norm

- Safe exploration for optimization: Restricted to **sub-space** of domain
- Exploit **local "smoothness"** to allow for more **optimistic exploration**

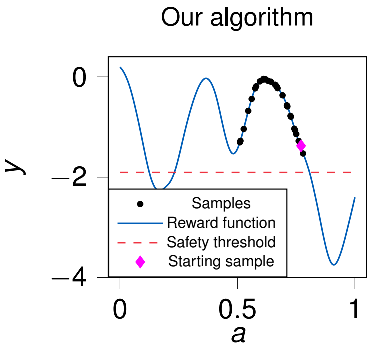
GP confidence intervals

$$|f(a) - \mu_t(a)| \leq (B_t + \text{"data-term"})\, \sigma_t(a)$$

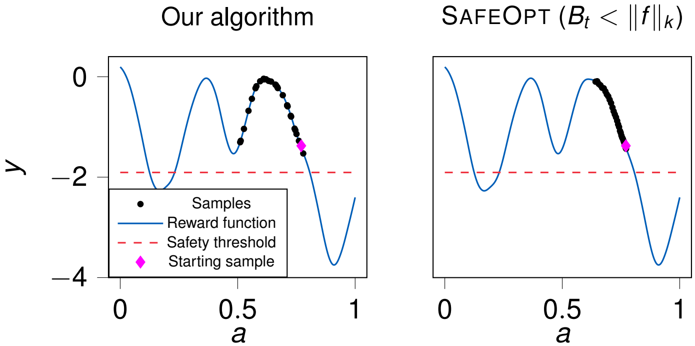


- **Adaptive** interpretation of locality: sub-domains around each sample
- **Significantly more scalable** through separate discretization in sub-domains

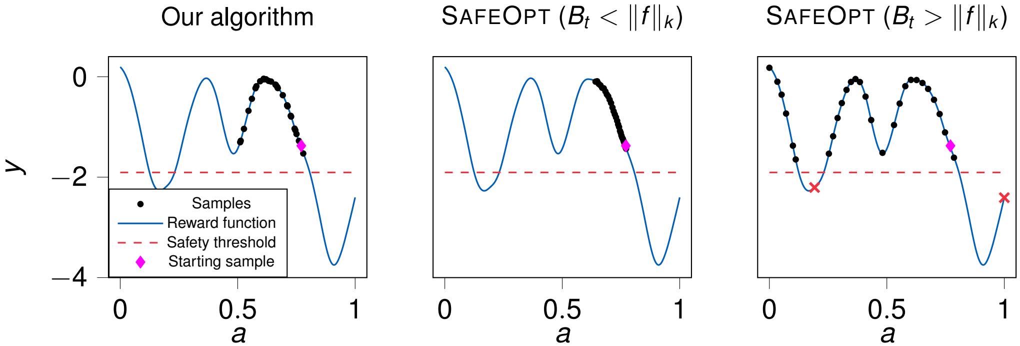
Numerical experiments



Numerical experiments

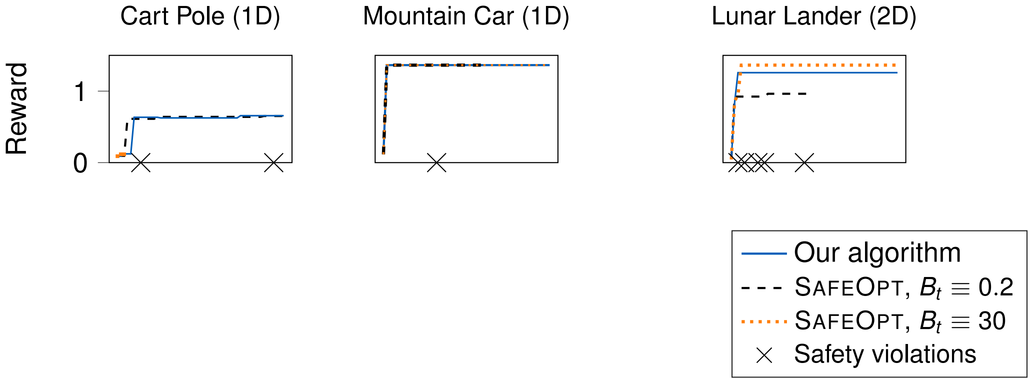


Numerical experiments

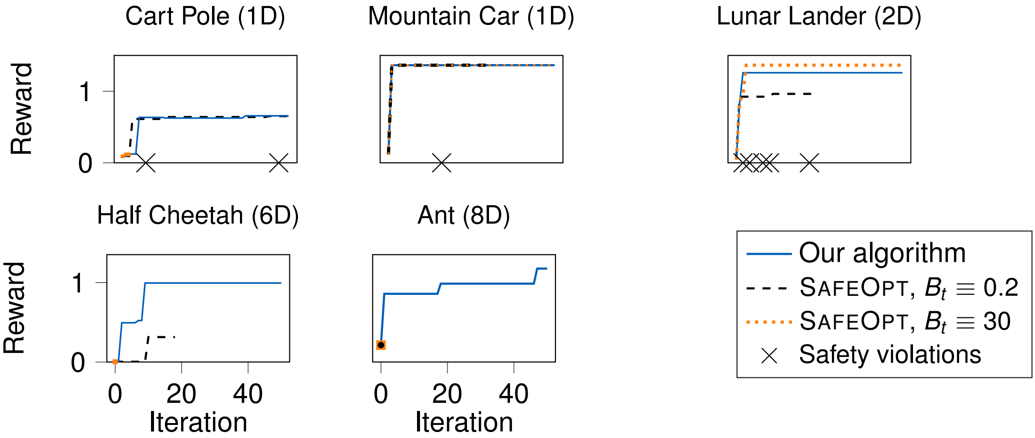


Safely fine-tuning RL policies

Safely fine-tuning RL policies



Safely fine-tuning RL policies



Hardware experiment

Limitations

Regularity assumption (Our approach)

RKHS norms $\|\rho_{t,j}\|_k$, $j \in \{1, \ldots, \mathrm{m}\}$ and $\|f\|_k$ are i.i.d. samples from the same (potentially unknown) probability space.

Limitations

Regularity assumption (Our approach)

RKHS norms $\|\rho_{t,j}\|_k$, $j \in \{1, \ldots, \mathrm{m}\}$ and $\|f\|_k$ are i.i.d. samples from the same (potentially unknown) probability space.

- **Frequentist setting**: Reward function f generated by *nature's* probability space

Limitations

Regularity assumption (Our approach)

RKHS norms $\|\rho_{t,j}\|_k$, $j \in \{1, \ldots, \mathrm{m}\}$ and $\|f\|_k$ are i.i.d. samples from the same (potentially unknown) probability space.

- **Frequentist setting**: Reward function f generated by *nature's* probability space
- By generating random RKHS functions, we **approximate nature's probability space**

Limitations

Regularity assumption (Our approach)

RKHS norms $\|\rho_{t,j}\|_k$, $j \in \{1, \ldots, \mathrm{m}\}$ and $\|f\|_k$ are i.i.d. samples from the same (potentially unknown) probability space.

- **Frequentist setting**: Reward function f generated by *nature's* probability space
- By generating random RKHS functions, we **approximate nature's probability space**
- **Mixing Bayesian** and **frequentist** methods by imposing a prior on f

Limitations

Regularity assumption (Our approach)

RKHS norms $\|\rho_{t,j}\|_k$, $j \in \{1, \ldots, \mathrm{m}\}$ and $\|f\|_k$ are i.i.d. samples from the same (potentially unknown) probability space.

- **Frequentist setting**: Reward function f generated by *nature's* probability space
- By generating random RKHS functions, we **approximate nature's probability space**
- **Mixing Bayesian** and **frequentist** methods by imposing a prior on f

Regularity assumption (SafeOpt)

Most safe BO algorithms require an upper bound B on the RKHS norm ($B \geq \|f\|_k$) a priori.

- In contrast to SafeOpt, we systematically **integrate data, adapt bounds** and cover a **rich set of functions**

Conclusions

Goal

Optimize control parameters of safety-critical real-world systems.

Problem definition

Develop a safe BO algorithm that estimates the RKHS norm $\|f\|_k$ with statistical guarantees.

Conclusions

Goal

Optimize control parameters of safety-critical real-world systems.

Problem definition

Develop a safe BO algorithm that estimates the RKHS norm $\|f\|_k$ with statistical guarantees.

Contributions

1. Abdullah Tokmak, Thomas B. Schön, Dominik Baumann, "**PACSBO: Probably approximately correct safe Bayesian optimization**," In: Springer Lecture Notes in Control and Information Sciences - Proceedings, 2024.
2. Abdullah Tokmak, Kiran G. Krishnan, Thomas B. Schön, Dominik Baumann, "**Safe exploration in reproducing kernel Hilbert spaces**," submitted to AISTATS 2025.

Conclusions

Goal

Optimize control parameters of safety-critical real-world systems.

Problem definition

Develop a safe BO algorithm that estimates the RKHS norm $\|f\|_k$ with statistical guarantees.

Contributions

1. Abdullah Tokmak, Thomas B. Schön, Dominik Baumann, "**PACSBO: Probably approximately correct safe Bayesian optimization**," In: Springer Lecture Notes in Control and Information Sciences - Proceedings, 2024.
2. Abdullah Tokmak, Kiran G. Krishnan, Thomas B. Schön, Dominik Baumann, "**Safe exploration in reproducing kernel Hilbert spaces**," submitted to AISTATS 2025.

Preprints