PACSBO: Probably approximately correct safe Bayesian optimization

AI Day 2024

Abdullah Tokmak[1,2]

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

October 21, 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.

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.

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.

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

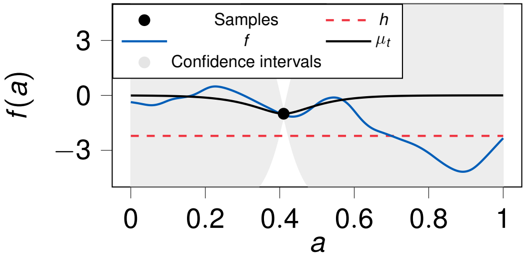
Control policy optimization problem

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

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

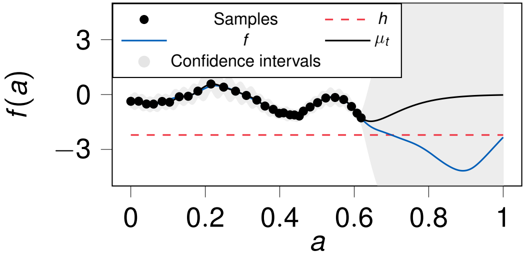
Control policy optimization problem

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

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

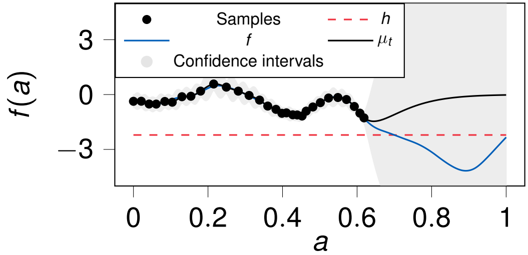
Control policy optimization problem

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

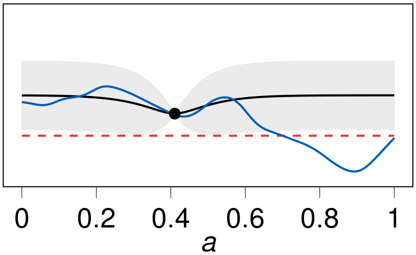
GP confidence intervals

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

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



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



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

Safe Bayesian optimization (BO) with GPs

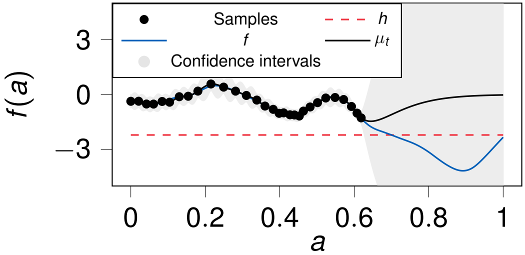
Control policy optimization problem

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

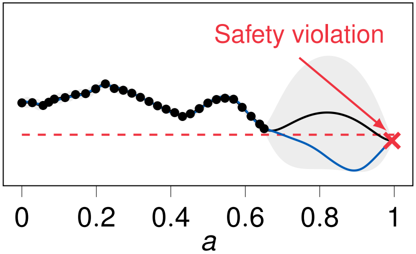
GP confidence intervals

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

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

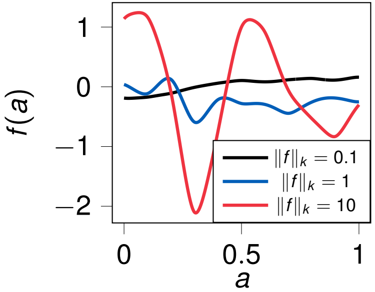
Most safe BO algorithms require an upper bound B on the RKHS norm ($B \geq \|f\|_k$) a priori.

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

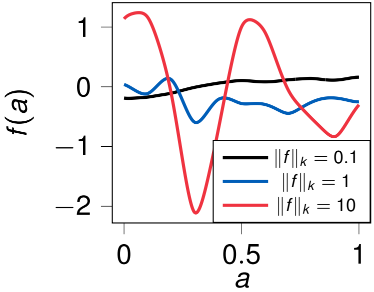


RKHS norm assumption in safe BO

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**

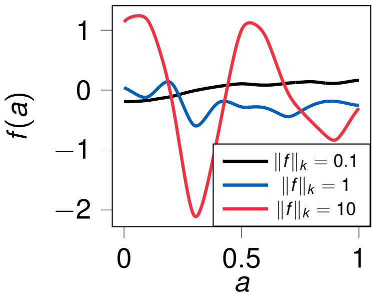


RKHS norm assumption in safe BO

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

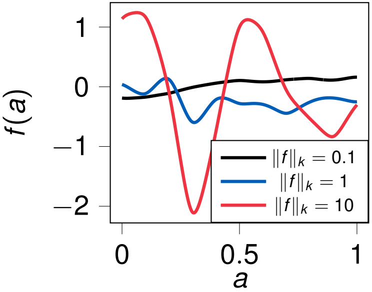


RKHS norm assumption in safe BO

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

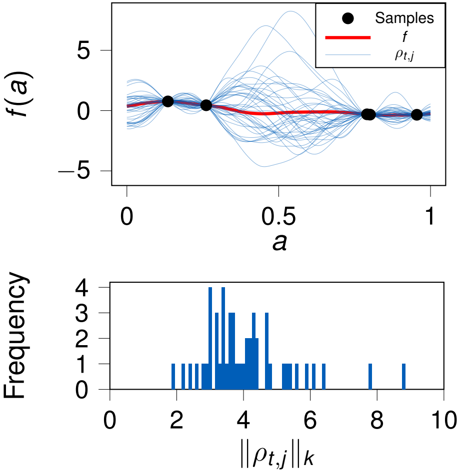


Problem definition

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

Random RKHS functions

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

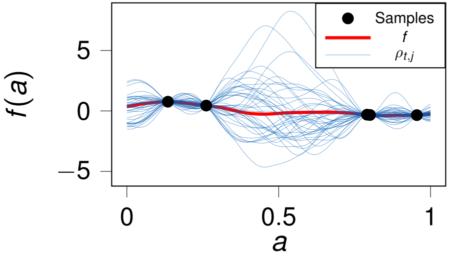


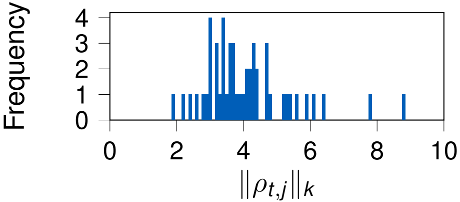
Random RKHS functions

- Compute random RKHS functions $\rho_{t,j},\ j \in \{1,\dots,\mathrm{m}\}$ with kernel k

- Random RKHS functions $\rho_{t,j}$ capture the behavior of reward function f

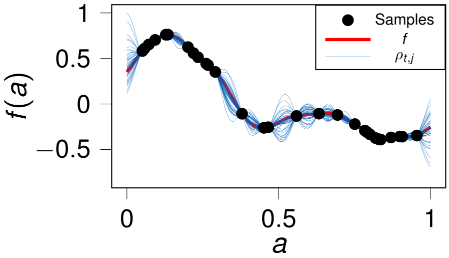
- Increasing sampling density: $\rho_{t,j}, \|\rho_{t,j}\|_k \to f, \|f\|_k$

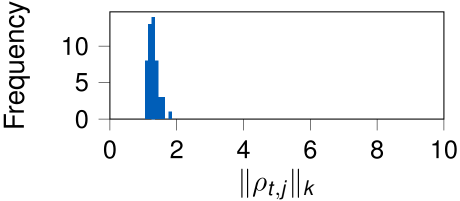




Random RKHS functions

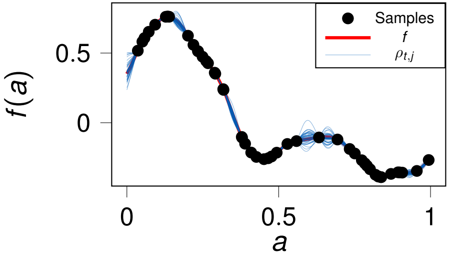
- Compute random RKHS functions $\rho_{t,j},\ j \in \{1, \ldots, \mathrm{m}\}$ with kernel k
- Random RKHS functions $\rho_{t,j}$ capture the behavior of reward function f
- Increasing sampling density: $\rho_{t,j}, \|\rho_{t,j}\|_k \rightarrow f, \|f\|_k$

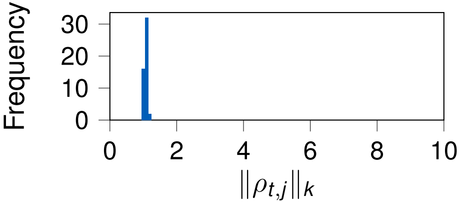




Random RKHS functions

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

[2] 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) $\forall t \geq 1$.

[2] 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

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

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

Local interpretation of the RKHS norm

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

GP confidence intervals

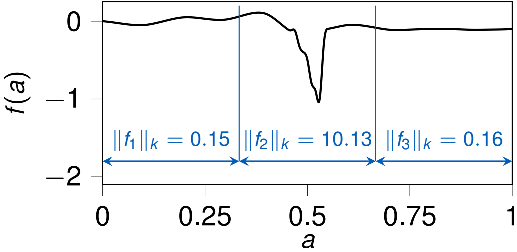
$$|f(a) - \mu_t(a)| \leq (B_t + \text{“data-term”})\, \sigma_t(a)$$

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

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

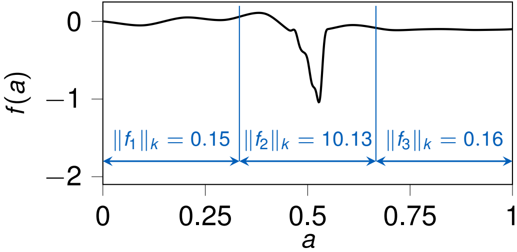


Local interpretation of the RKHS norm

- Safe exploration for optimization: Restricted to **sub-space** of domain
- Exploit **local “smoothness”** to allow for more **optimistic exploration**
- Implementation: Three sub-domains around the convex hull of samples

GP confidence intervals

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



Problem definition

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

PACSBO

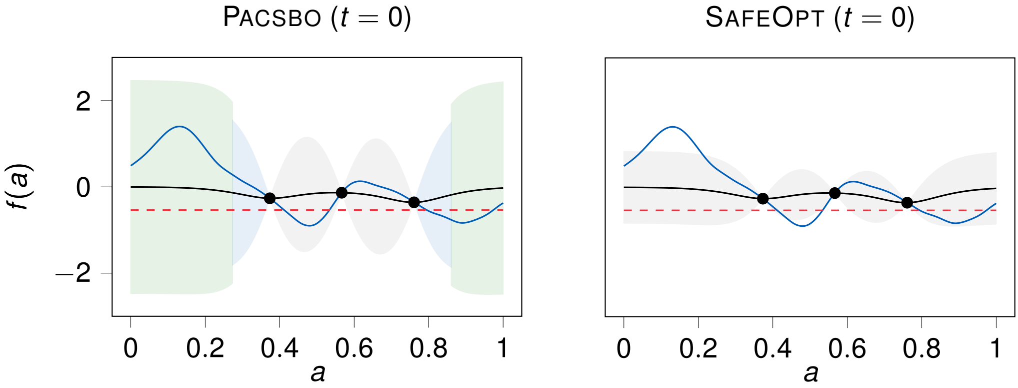
Problem definition

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

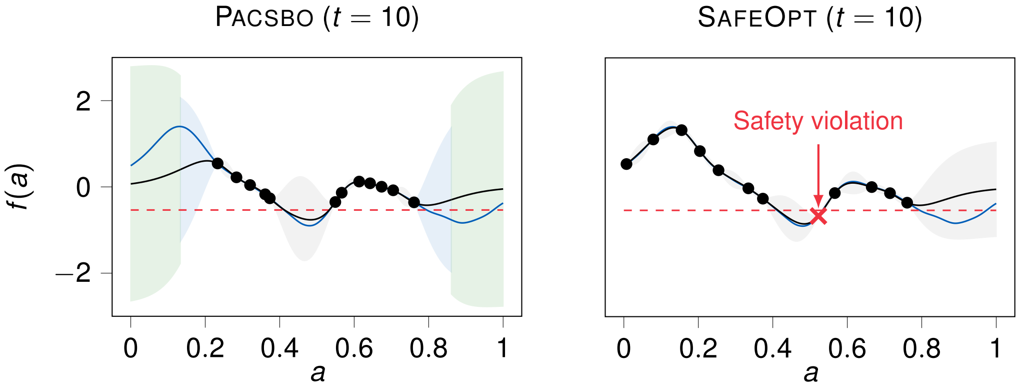
PACSBO: **P**robably **a**pproximately **c**orrect **s**afe **B**ayesian **o**ptimization

- Data-driven RKHS norm over-estimation with PAC bounds
- Local interpretation of the RKHS norm

Numerical experiments



Numerical experiments



Hardware experiment

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 *Symposium on Systems Theory in Data and Optimization*, 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 *Symposium on Systems Theory in Data and Optimization*, 2024.
2. Abdullah Tokmak, Kiran G. Krishnan, Thomas B. Schön, Dominik Baumann, "**Safe exploration in reproducing kernel Hilbert spaces**," submitted to *AISTATS 2025*.

Preprints