Bayesian optimization and scenario programming for safe control parameter tuning

Abdullah Tokmak[1,2]

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

January 15, 2026

Aalto University
School of Electrical
Engineering

UPPSALA
UNIVERSITET

Personal introduction: Abdullah Tokmak

- 3rd year Ph.D. student at **Aalto University** with **Dominik Baumann** and **Thomas Schön** (Uppsala University)
- **Research interests:** Safe Bayesian optimization, kernel methods, uncertainty quantification

- M.Sc. and B.Sc. degrees in mechanical engineering at **RWTH Aachen University**;
  Working with **Sebastian Trimpe** on learning-based control
- Master's thesis at **ETH Zurich** on approximate MPC with **Melanie Zeilinger and Johannes Köhler**

Motivational example

Introduction

Goal

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

- Unknown **black-box** reward function $f : \mathcal{A} \subseteq \mathbb{R}^n \mapsto \mathbb{R}$
- Control policy parameters $a \in \mathcal{A}$; **parametrized** policy
- Regularity: Function f member of **RKHS** H_k of kernel k with RKHS norm $\|f\|_k$
- Safety-critical real-world systems: **Safety** and **sample-efficiency**

Solvable using classical RL?

Vanilla RL struggles with both **sample-efficiency** and **safety guarantees**.

Kernels and Gaussian processes (GPs)

- GPs to **predict reward function** f from samples with **uncertainty quantification**
  - At iteration $t \geq 1$: Control parameter $a_t \in \mathcal{A}$, noisy reward $y_t := f(a_t) + \epsilon_t$
  - Until iteration $t \geq 1$: $a_{1:t} := [a_1, \ldots, a_t]^\top$, $y_{1:t} := [y_1, \ldots, y_t]^\top$
- GPs characterized by **kernel** k, posterior mean μ_t, posterior standard deviation σ_t:

$$\mu_t(a) := k_t(a)^\top (K_t + \eta I_t)^{-1} y_{1:t}$$

$$\sigma_t(a) := \sqrt{1 - k_t(a)^\top (K_t + \eta I_t)^{-1} k_t(a)}$$

- **Notation**: Regularization constant $\eta > 0$, covariance matrix K_t, covariance vector $k_t(a) := [k(a_1, a), \ldots, k(a_t, a)]^\top$, identity matrix I_t

Theorem (Confidence intervals)[1]

Suppose:

- Sufficiently regular reward function f (RKHS norm)
- GP posterior with noisy observations and same kernel k

Then:

$$\begin{aligned}
|f(a) - \mu_t(a)| &= |f(a) - k_t(a)^\top (K_t + \eta I_t)^{-1} y_{1:t}| \\
&\leq |k_t(a)^\top (K_t + \eta I_t)^{-1} f_{1:t}| + |k_t(a)^\top (K_t + \eta I_t)^{-1} \epsilon_{1:t}| \\
&\leq |\|f\|_k \sigma_t(a) + \frac{1}{\eta} \epsilon_{1:t}^\top K_t (K_t + \eta I_t)^{-1} \epsilon_{1:t} \sigma_t(a) \\
&=: \left(\|f\|_k + \frac{1}{\eta} \epsilon_{1:t}^\top \Xi_t \epsilon_{1:t} \right) \sigma_t(a) \qquad \texttt{\\\\ RKHS term + noise term}
\end{aligned}$$

[1] Chowdhury et al., "On kernelized multi-armed bandits," ICML 2017.

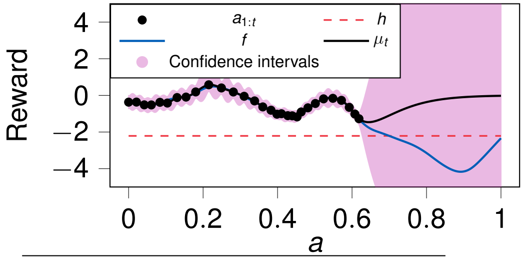
Safe Bayesian optimization (BO) with GPs

- Safety critical real-world system
  - GPs: Confidence intervals ⇒ **safety**
  - BO: Exploration/exploitation ⇒ **efficient**
- **Episodic**: Experiment ↻ Parameter acquisition

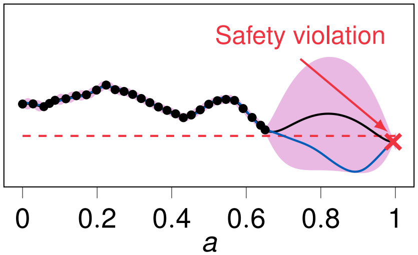
Safe policy optimization

$$\max_{a \in \mathcal{A}} f(a) \text{ s.t. } f(a_t) \geq h, \ \forall t \geq 1$$

SAFEOPT[2] (correct confidence intervals)



SAFEOPT[2] (wrong confidence intervals)



[2] Y. Sui et al., "Safe exploration for optimization with Gaussian processes," ICML 2015.

Correctness of confidence intervals

$$|f(a) - \mu_t(a)| \leq \left(\|f\|_k + \frac{1}{\eta} \epsilon_{1:t}^\top \Xi_t \epsilon_t \right) \sigma_t(a)$$

```
\\ RKHS term + noise term
```

State of the art safe BO algorithms require:

- Guess $B \geq \|f\|_k$ *a priori*
- Assume homoscedastic R sub-Gaussian noise

$$\mathbb{P}_\epsilon \left[|f(a) - \mu_t(a)| \leq \left(B + \frac{R}{\sqrt{\eta}} \sqrt{\log \det \left(\frac{1}{\eta} K_t + I_t \right) - 2 \log(\nu)} \right) \sigma_t(a) \right] \geq 1 - \nu$$

Problem definition

How can we have *(i)* more general noise model and *(ii)* estimate the RKHS norm?

Bounding observation noise with high probability

- *R*-**sub-Gaussian**: Family of distributions whose tail decay is at most $\mathcal{N}(0, R^2)$
- **Homoscedastic**: The noise does not vary with the input $a \in \mathcal{A}$

Classic safe BO: Observation noise ϵ_t is homoscedastic and R-sub-Gaussian.

Our assumption: Observation noise ϵ_t is defined on a probability space, from which we can sample.

- Our assumption unifies different noise models
- Generating samples from the distribution $\Rightarrow$ **statistical bounds**

PAC bounds via the scenario approach[3]

- Data-driven decision making and **uncertainty quantification** tool
- Let $\epsilon \in \mathcal{E}$ be unknown and $\tilde{\epsilon}_j \in \mathcal{E}, j \in [1, \dots, m]$ be i.i.d. *scenarios*

Nontrivial convex optimization problem

$$\min_{x \in \mathbb{R}^d} f_c(x) \quad \text{s.t.} \quad g_c(x, \epsilon) \leq 0$$

Convex scenario program

$$x^\star := \arg\min_{x \in \mathbb{R}^d} f_c(x) \quad \text{s.t.} \quad g_c(x, \tilde{\epsilon}_j) \leq 0, \forall j \in [1, m]$$

Generalization guarantees of the scenario approach

$$\mathbb{P}^m[\mathbb{P}[\epsilon \in \mathcal{E} : g_c(x^\star, \epsilon) \leq 0] \geq 1 - \nu] \geq 1 - \sum_{i=0}^{d} \binom{m}{i} \nu^i (1 - \nu)^{m-i}$$

[3] Campi and Garatti, "Introduction to the scenario approach," SIAM, 2018.

Scenario approach to bound observation noise

- Upper-bound (unknown) observation noise: $\min_{\bar{\epsilon}_t \in \mathbb{R}_{\geq 0}} \bar{\epsilon}_t \quad \text{s.t.} \quad \bar{\epsilon}_t \geq \epsilon_t$
- Scenario program with discarded constraints:

$$\bar{\epsilon}_t := \min_{\bar{\epsilon}_t \in \mathbb{R}_{\geq 0}} \bar{\epsilon}_t \quad \text{s.t.} \quad \bar{\epsilon}_t \geq |\tilde{\epsilon}_{j,t}|, \quad \forall j \in \mathrm{Quantile}(m, \nu, \kappa)$$

$$\Downarrow$$

$$\mathbb{P}^m_\epsilon[t \geq 1 : \mathbb{P}_\epsilon[\bar{\epsilon}_t < |\epsilon_t|] > \nu] \leq \kappa$$

- Inductive conformal prediction yields **exactly same guarantees**[4]
  - **One-dimensional** optimization problem, where risk is bounded by beta distribution
  - No separation of **learning** and **validation** needed

[4] O'Sullivan et al., "Bridging conformal prediction and scenario optimization," CDC, 2025.

Resulting confidence intervals

- For **fixed iteration** t: $\mathbb{P}^m_\epsilon[t \geq 1 : \mathbb{P}_\epsilon[\bar{\epsilon}_t < |\epsilon_t|] > \nu] \leq \kappa$
- Iteration-dependent $\kappa_t := \frac{6\kappa}{\pi^2 t^2}$ and Boole's inequality for simulaneous bounds
- $\mathbb{P}^m_\epsilon[\forall t{:}\mathbb{P}_\epsilon[\bar{\epsilon}_t < |\epsilon_t|] > \nu] \leq \sum_{t=1}^\infty \mathbb{P}^m_\epsilon[t \geq 1{:}\mathbb{P}_\epsilon[\bar{\epsilon}_t < |\epsilon_t|] > \nu] \leq \sum_{t=1}^\infty \kappa_t = \sum_{t=1}^\infty \frac{6\kappa}{\pi^2 t^2} = \kappa$

$$
\begin{aligned}
|f(a) - \mu_t(a)| &\leq \left(\|f\|_k + \frac{1}{\eta}\epsilon_{1:t}^\top \Xi_t \epsilon_{1:t}\right)\sigma_t(a) && \texttt{\\\\ RKHS term + noise term} \\
&\leq \left(\|f\|_k + \sqrt{\frac{\lambda_{\max}(\Xi_t)}{\eta}}\|\epsilon_{1:t}\|_2\right)\sigma_t(a) && \texttt{\\\\ deterministically} \\
&\leq \left(\|f\|_k + \sqrt{\frac{\lambda_{\max}(\Xi_t)}{\eta}}\|\bar{\epsilon}_{1:t}\|_2\right)\sigma_t(a) && \texttt{\\\\ with high probability}
\end{aligned}
$$

Safe BO: Safety and optimality[5]

Theorem: Safe BO under general noise models

Suppose:

- Start exploration from safe (sub-optimal) policy parameter
- Execute safe BO with confidence intervals we derived using scenario approach

Then, with confidence $1 - \kappa$, the following holds with probability $1 - \nu$:

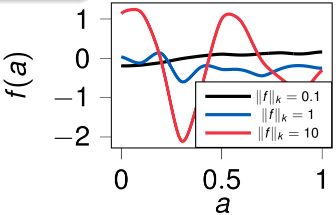
- We find the **reachable optimum** after at most $t^\star$ iterations
- We are **safe** at each iteration $t \in [1, t^\star]$

[5] Tokmak et al., "Safe Bayesian optimization across noise models via scenario programming," IEEE L-CSS, 2026.

RKHS norm assumption in safe BO

$$|f(a) - \mu_t(a)| \leq \left(B + \sqrt{\frac{\lambda_{\max}(\Xi_t)}{\eta}}\|\bar{\epsilon}_{1:t}\|_2\right)\sigma_t(a)$$

- RKHS norm bound $B \geq \|f\|_k$ characterizes **"smoothness"** of function **unknown** reward function f
- Upper bound for **safety**, tightness for **practicality**
- It is **unclear how to bound/guess RKHS norm** of unknown functions[6]



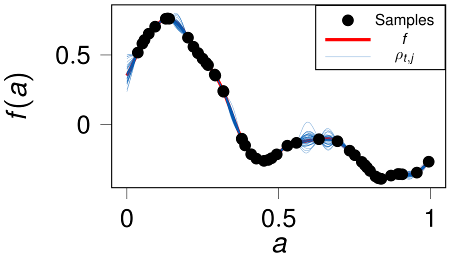
Problem definition

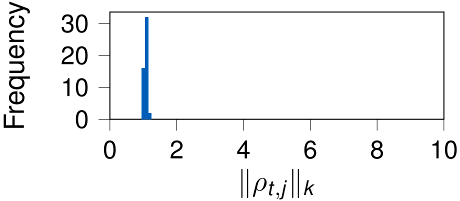
Instead of guessing $B \geq \|f\|_k$ *a priori*, derive **data-driven RKHS norm over estimation** $B_t \geq \|f\|_k$ with statistical guarantees.

[6] Tokmak et al.,"Automatic nonlinear MPC approximation with closed-loop guarantees," IEEE TAC, 2025.

Random RKHS functions[7,8]

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





[7] Tokmak et al., “PACSBO: Probably appproximately correct safe Bayesian optimization,” SysDO, 2024.

[8] Tokmak et al., “Safe exploration in reproducing kernel Hilbert spaces,” AISTATS, 2025.

Scenario approach for RKHS norm estimation

- Assume: Reward f and random RKHS functions $\rho_{t,j}$ are from same probability space
- Implementation: $\rho_{t,j}(x) = \sum_{i=1}^{N} \alpha_i k(x_i, x)$, where α_i and x_i are random
- Mapping α_i, x_i to RKHS norms is deterministic and measurable
  $\Rightarrow$ RKHS norms $\|f\|_k$ and $\|\rho_{t,j}\|_k$ are from the same *induced* probability space

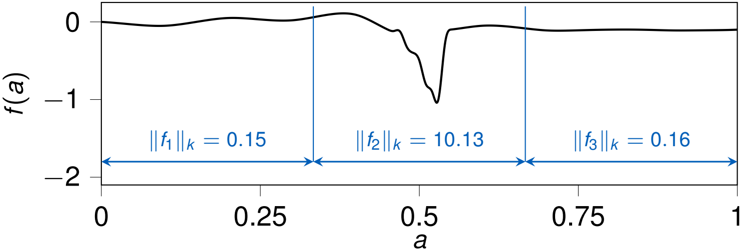
$$\min_{B_t \in \mathbb{R}_{\geq 0}} B_t \quad \text{s.t.} \quad B_t \geq \|\rho_{t,j}\|_k, \forall j \in \mathbf{Quantile}(m, \nu, \kappa) \qquad \texttt{\\\\ convex optimization}$$

$$\mathbb{P}_f^m[\mathbb{P}_f[f \in H_k : B_t \geq \|f\|_k] \geq 1 - \nu] \geq 1 - \kappa \qquad \texttt{\\\\ RKHS norm over-estimation}$$

$$\mathbb{P}_{f,\epsilon}^m[\mathbb{P}_{f,\epsilon}[f \in H_k : |f(a) - \mu_t(a)| \leq (B_t + \text{noise})\, \sigma_t(a)] \geq 1 - \nu] \geq 1 - \kappa \quad \texttt{\\\\ conf. int.}$$

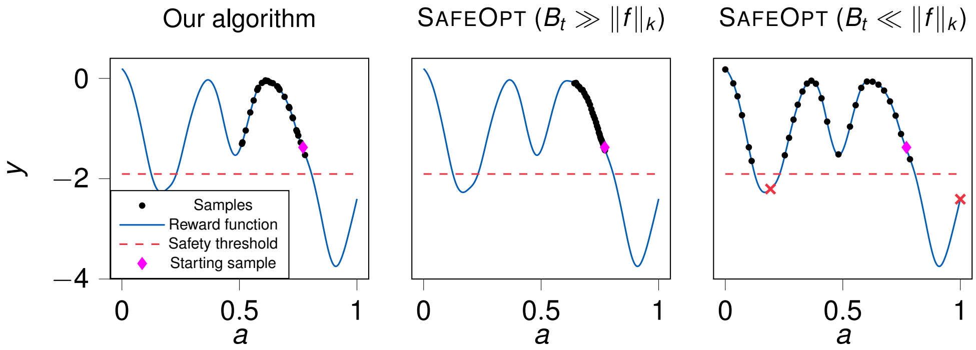
Local interpretation of the RKHS norm

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



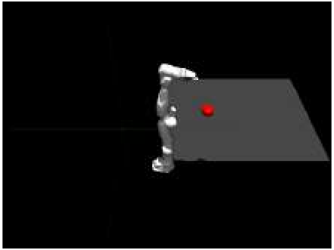
- **Adaptive** interpretation of locality: sub-domains around each sample
- **Significantly more scalable** through separate discretization in sub-domains

Numerical experiment



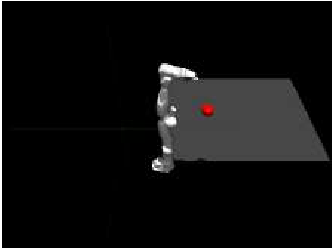
Franka simulator (initial policy)

- Operational space impedance controller K for a set-point tracking task
- We obtain K from solving LQR problem and tune entries in Q and R
- Reward function encourages reacing target quickly with small inputs
- Constraint function requires that the distance to the target descreases sufficiently



Franka simulator (final policy)

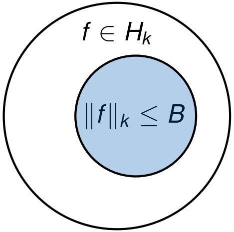
- Operational space impedance controller K for a set-point tracking task
- We obtain K from solving LQR problem and tune entries in Q and R
- Reward function encourages reacing target quickly with small inputs
- Constraint function requires that the distance to the target descreases sufficiently

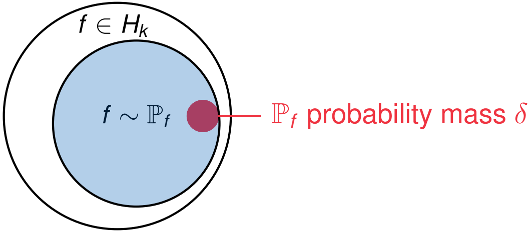


Motivational example revisited

Motivational example: Final policy

Frequentist vs. Bayesian viewpoints





- Worst case/agnostic view on f
- $\mathbb{P}_\epsilon$ is randomness over noise
- $\mathbb{P}_\epsilon$[safety violation] $\leq \delta$ is *robust* w.r.t f

- RKHS norm estimation $\Rightarrow \mathbb{P}_f$ over f
- *Impossible*: $\mathbb{P}_f[\text{safety violation} | f = \mathfrak{f}] \leq \delta$
- $\mathbb{P}_{\epsilon,f}[\text{safety violation}] \leq \delta$, $\mathbb{P}_{\epsilon,f} := \mathbb{P}_\epsilon \otimes \mathbb{P}_f$

Although associated function spaces similar, the resulting **guarantees** are **different**!

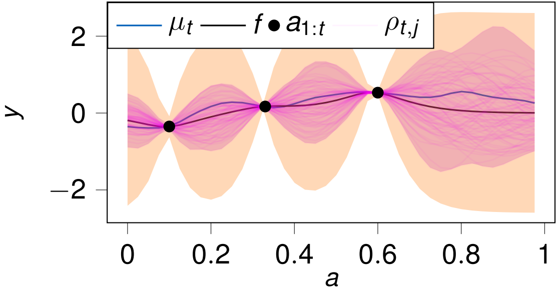
A Bayesian view on uncertainty tubes

Regularity assumption with RKHS norm estimation

Asm. 1: Reward f is a member of the RKHS of kernel k

Asm. 2: Reward f and random RKHS functions $\rho_{t,j}$ are i.i.d. samples of same prob. space

Can we work only under Assumption 2 and **drop Assumption 1?**



- Bounds seem tighter in a clean Bayesian setting
- Guarantees based on **compression**, which generalizes scenario theory[9]

[9] Campi and Garatti, "Compression, generalization and learning," JMLR, 2023.

Future work

Near-term perspectives

- Overlooked applications of safe Bayesian optimization
  - Optimizing radio resource management (Nokia Bell Labs)
  - Safe BO with noise oracles from radar data (Finnish industrial partners)
  - Experiment design in quantum variational algorithms
- Safe control parameter tuning in distributed multi-agent systems
- Bayesian confidence tube via scenario programming (with T. Karvonen, S. Garatti)

Long-term perspectives

- Can we have a more sophisticated definition of safety?
- Is the complete model-free approach too conservative?

Conclusions

Recap

Goal

Safe and **sample-efficient** control policy parameter optimization.

Core contributions

- RKHS norm estimation and general noise models
- Integration into safe BO algorithms for parameter tuning

Selected works

- **A. Tokmak**, T. B. Schön, D. Baumann, "PACSBO: Probably approximately correct safe Bayesian optimization," Syposium on Systems Theory in Data and Optimization, 2024.
- **A. Tokmak**, K. Krishnan, T. B. Schön, D. Baumann, "Safe exploration in reproducing kernel Hilbert spaces," International Conference on Artificial Intelligence and Statistics, 2025.
- **A. Tokmak**, T. B. Schön, D. Baumann, "Safe Bayesian optimization across noise models via scenario programming," IEEE Control Systems Letters, 2026.