Safe Exploration for Optimization with Gaussian Processes

Yanan Sui
Caltech

Alkis Gotovos
ETH Zurich

Joel W. Burdick
Caltech

Andreas Krause
ETH Zurich

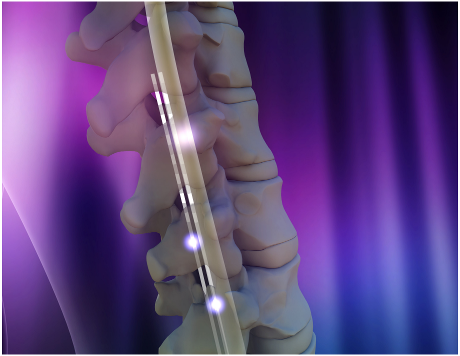
ICML Lille
International Conference on Machine Learning

Better safe than sorry

youtube.com/user/mattessons

Therapeutic spinal cord stimulation

girardgibbs.com



sjm.com

- Find electrode configurations that maximize muscle activity
- Bad configurations may cause pain or have negative effects on treatment

Goal

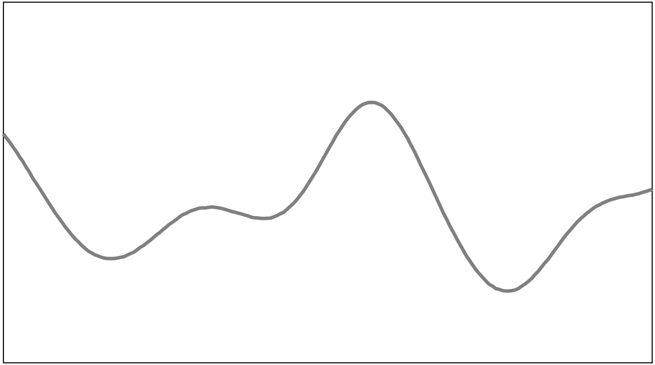
Optimize an unknown reward function via sequential sampling

AND

remain “safe” throughout the process

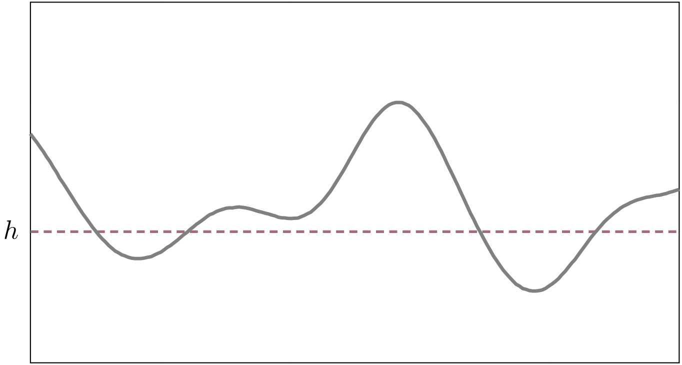
Problem statement

- Finite decision set D
- Unknown reward function $f : D \to \mathbb{R}$

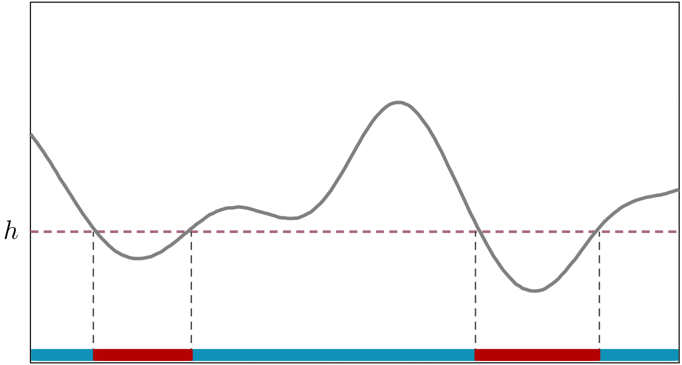


Problem statement

- Finite decision set D
- Unknown reward function $f : D \to \mathbb{R}$
- Safety threshold $h \in \mathbb{R}$

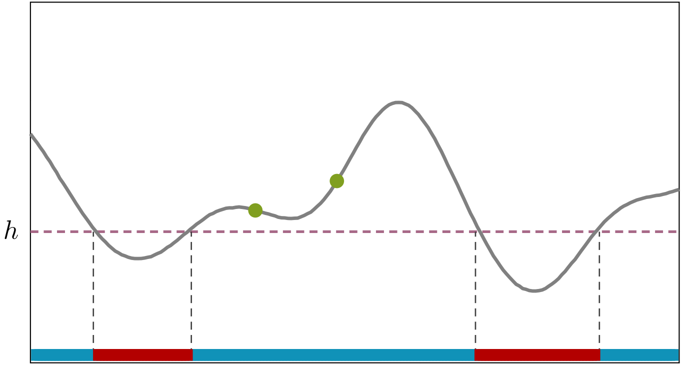


- Finite decision set D
- Unknown reward function $f : D \to \mathbb{R}$
- Safety threshold $h \in \mathbb{R}$



Problem statement

- Finite decision set D
- Unknown reward function $f : D \to \mathbb{R}$
- Safety threshold $h \in \mathbb{R}$
- Seed set S_0 of safe decisions ($\forall x \in S_0,\ f(x) \geq h$)



Sequential sampling

- For $t = 1, 2, \ldots$
  - select $x_t \in D$
  - observe $f(x_t) + n_t$

Problem statement

Sequential sampling

- For $t = 1, 2, \ldots$
  - select $x_t \in D$
  - observe $f(x_t) + n_t$

Goal

- Find $x^* \in \text{argmax}_{x \in D} f(x)$
- Remain safe: $\forall t \geq 1,\ f(x_t) \geq h$

Sequential sampling

- For $t = 1, 2, \ldots$
  - select $x_t \in D$
  - observe $f(x_t) + n_t$

Goal

- Find $x^* \in \text{argmax}_{x \in D} f(x)$
- $\boxed{\text{Remain safe: } \forall t \geq 1,\ f(x_t) \geq h}$

Related work

- Bayesian optimization: function evaluation is expensive

Related work

- Bayesian optimization: function evaluation is expensive
- Various proposed criteria, e.g.,
  - Expected improvement [Mockus et al., 1974]
  - UCB [Auer, 2002] [Srinivas et al., 2010]

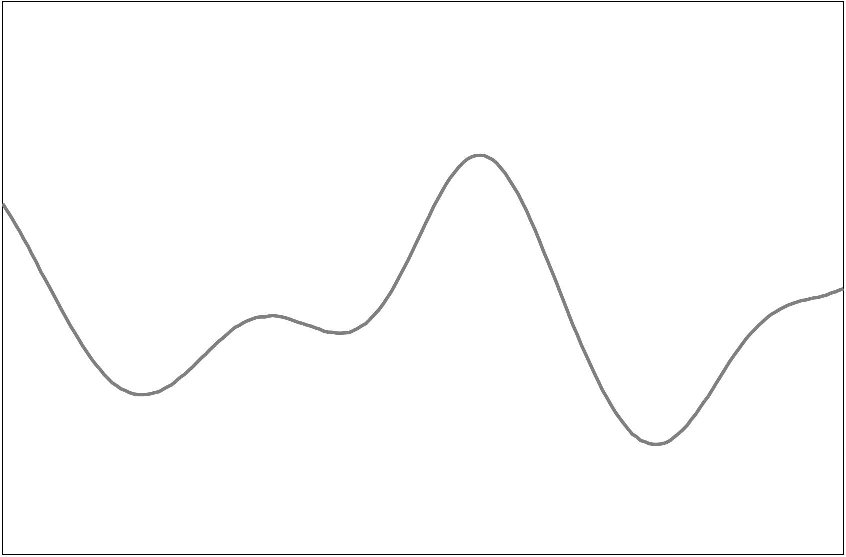
Related work

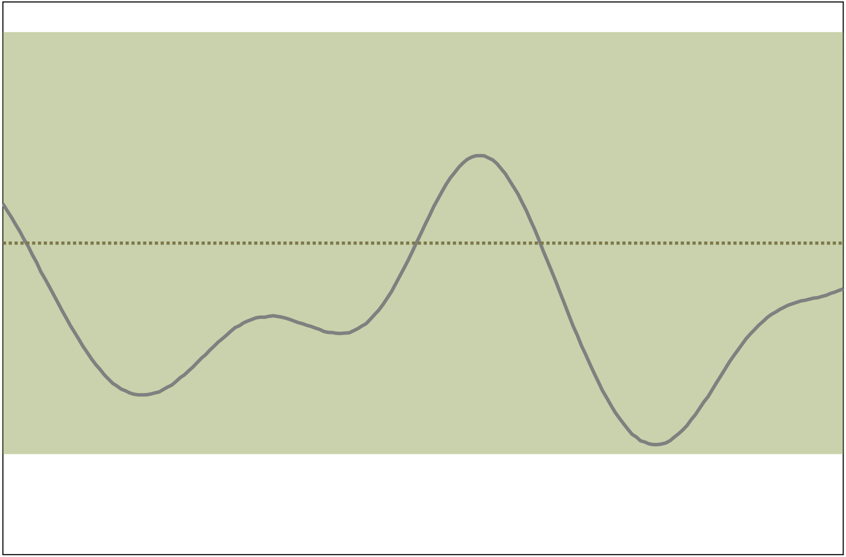
- Bayesian optimization: function evaluation is expensive
- Various proposed criteria, e.g.,
  - Expected improvement [Mockus et al., 1974]
  - UCB [Auer, 2002] [Srinivas et al., 2010]
- Related variants
  - Level set estimation [Gotovos et al., 2013]
  - Bayesian optimization with constraints [Gardner et al., 2014]

Related work

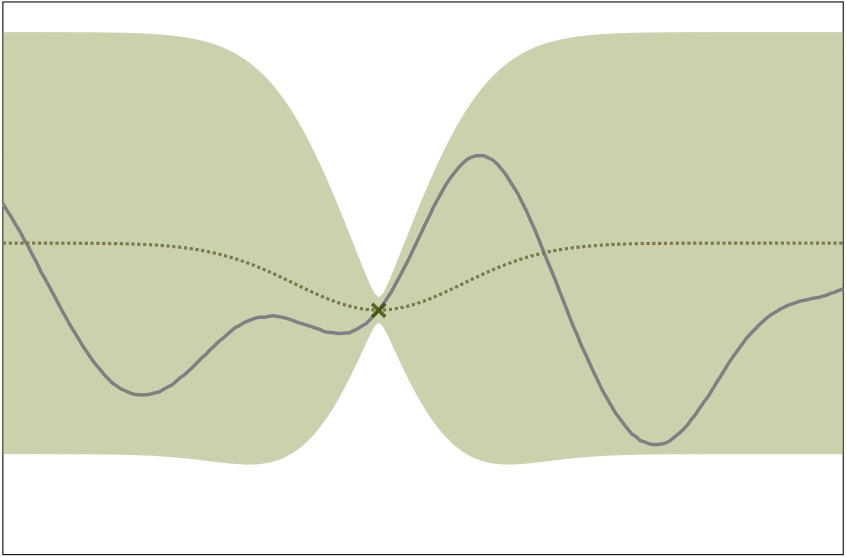
- Bayesian optimization: function evaluation is expensive
- Various proposed criteria, e.g.,
  - Expected improvement [Mockus et al., 1974]
  - UCB [Auer, 2002] [Srinivas et al., 2010]
- Related variants
  - Level set estimation [Gotovos et al., 2013]
  - Bayesian optimization with constraints [Gardner et al., 2014]
- Gaussian processes popular for modeling the unknown function

Gaussian process regression

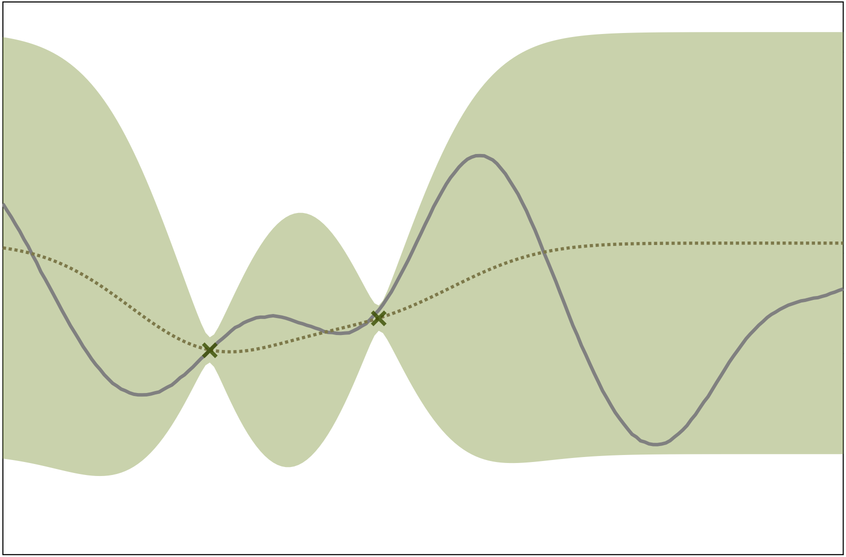




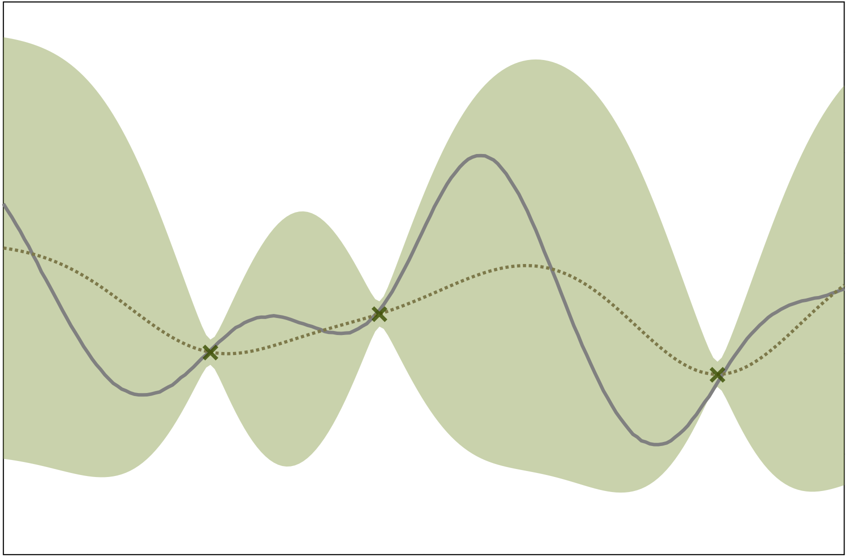
Gaussian process regression



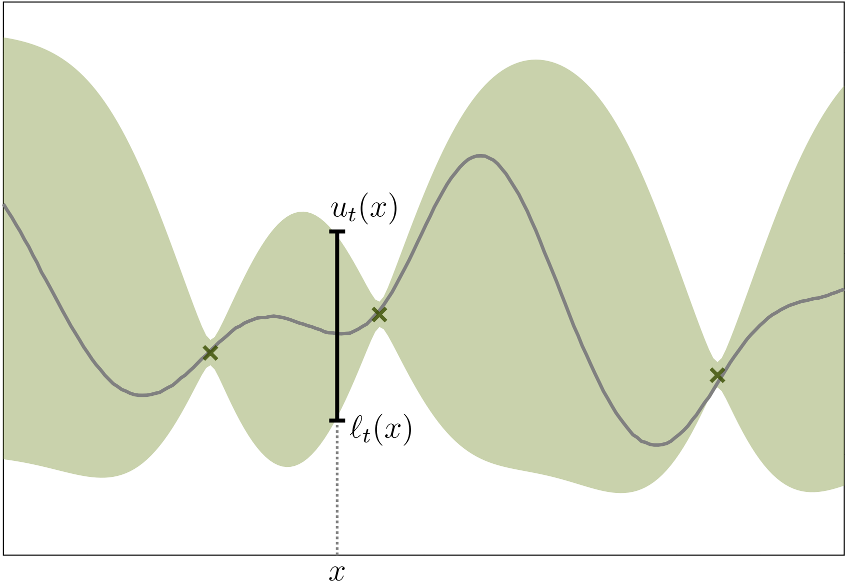
Gaussian process regression



Gaussian process regression

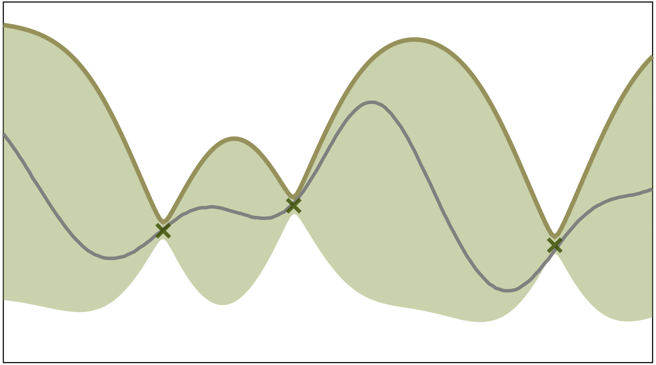


Gaussian process regression

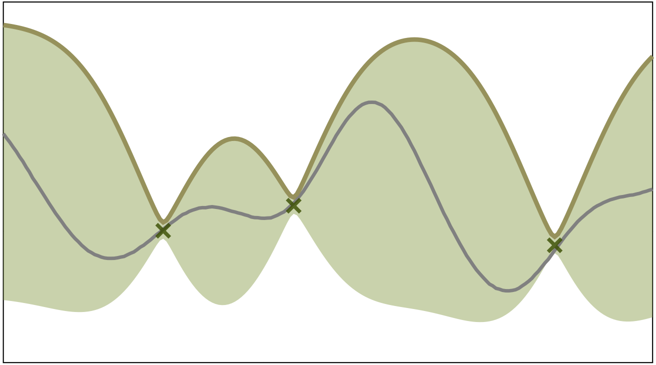


- Use upper confidence bounds for optimistic sampling

- Use upper confidence bounds for optimistic sampling
- $x_t = \text{argmax}_{x \in D}\, u_t(x)$

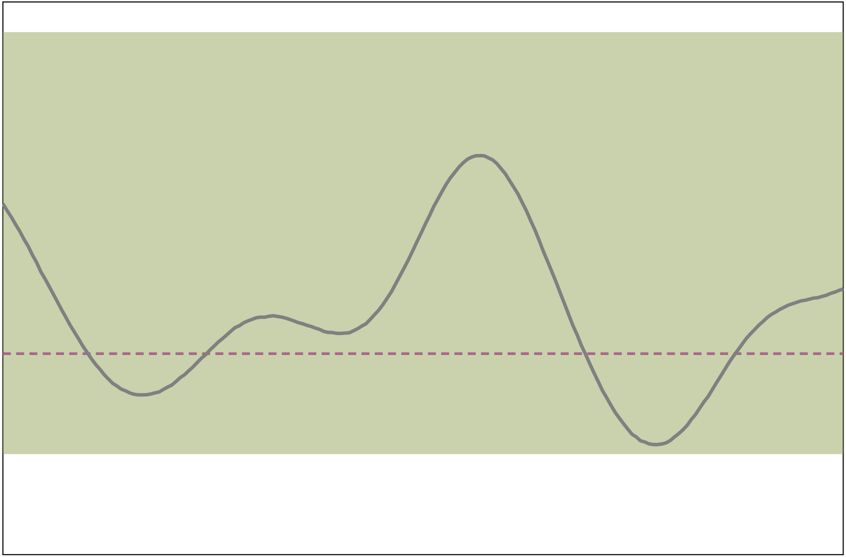


- Use upper confidence bounds for optimistic sampling
- $x_t = \text{argmax}_{x \in D}\, u_t(x)$

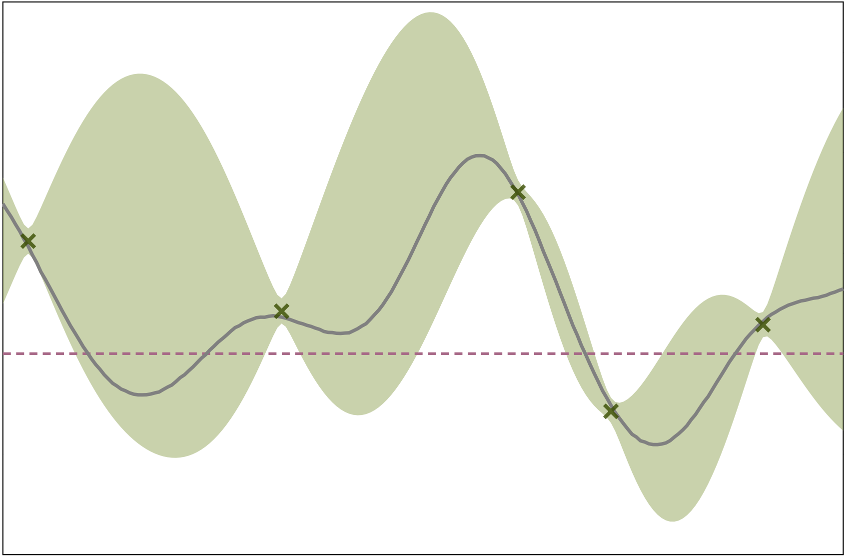


- Sublinear regret under suitable conditions on f [Srinivas et al., 2010]

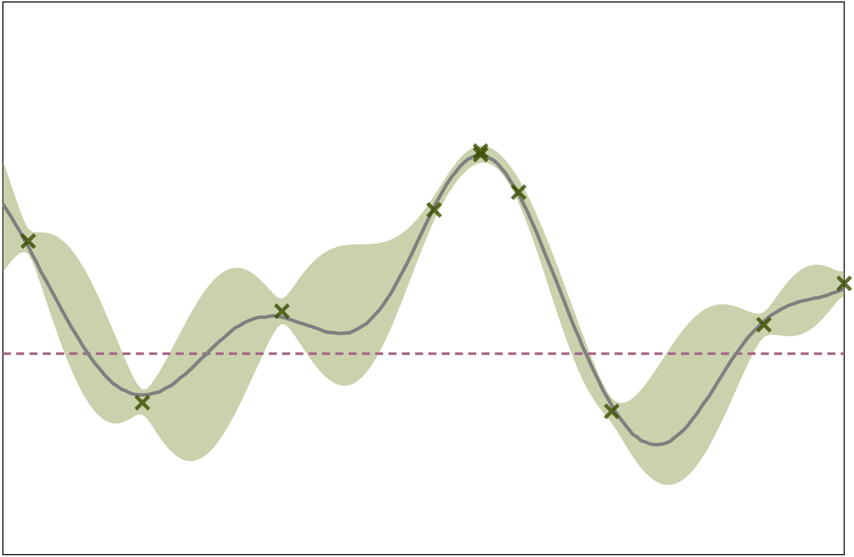
GP-UCB example ($t = 0$)



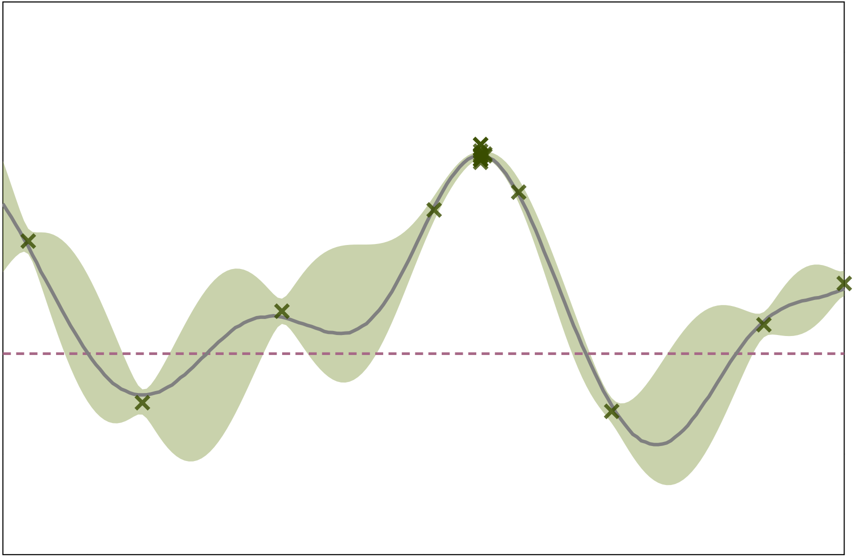
GP-UCB example ($t = 5$)



GP-UCB example ($t = 10$)



GP-UCB example ($t = 20$)



Certifying safety

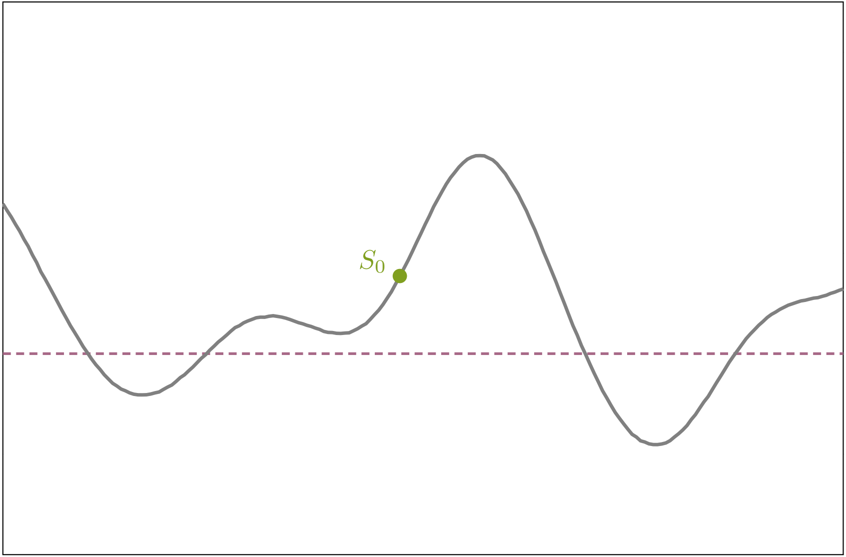
- Assume that f is L-Lipschitz continuous w.r.t. a metric d

Certifying safety

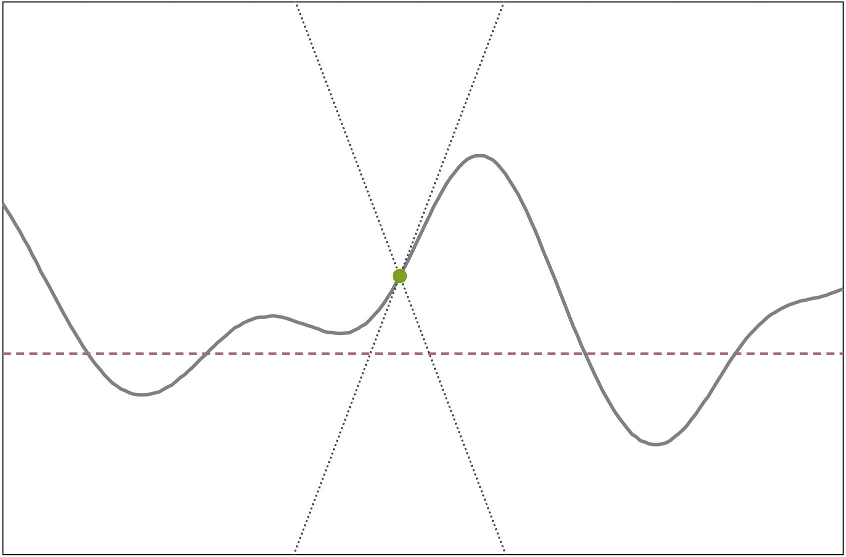
- Assume that f is L-Lipschitz continuous w.r.t. a metric d
- If for some safe x we know $f(x)$, then a safety certificate for x' is

$$f(x) - L\,d(x, x') \geq h$$

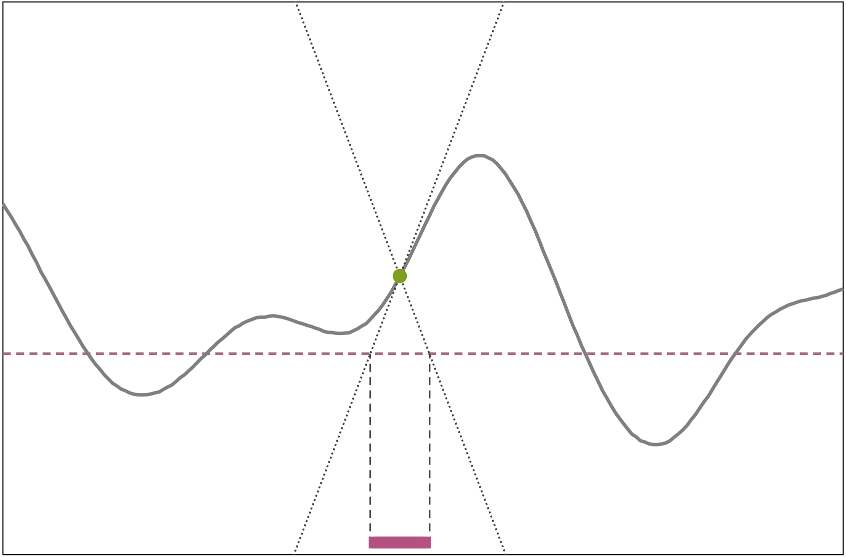
Certifying safety



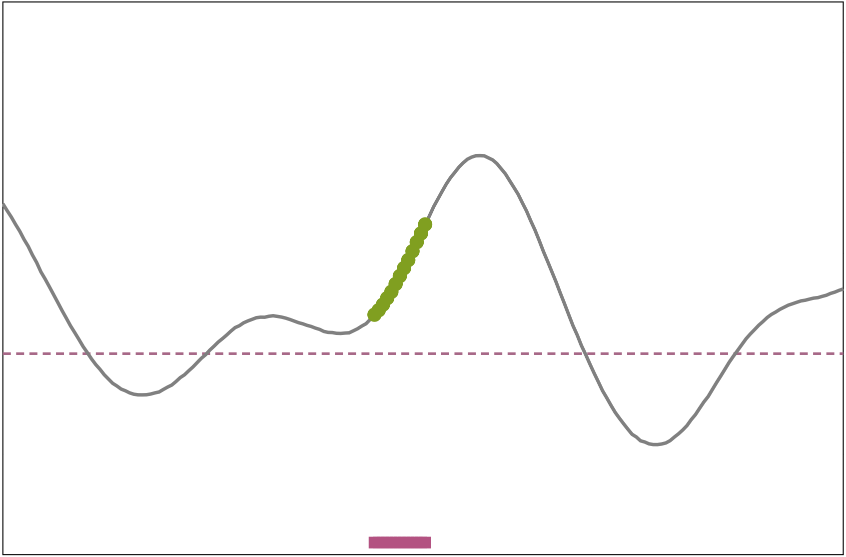
Certifying safety



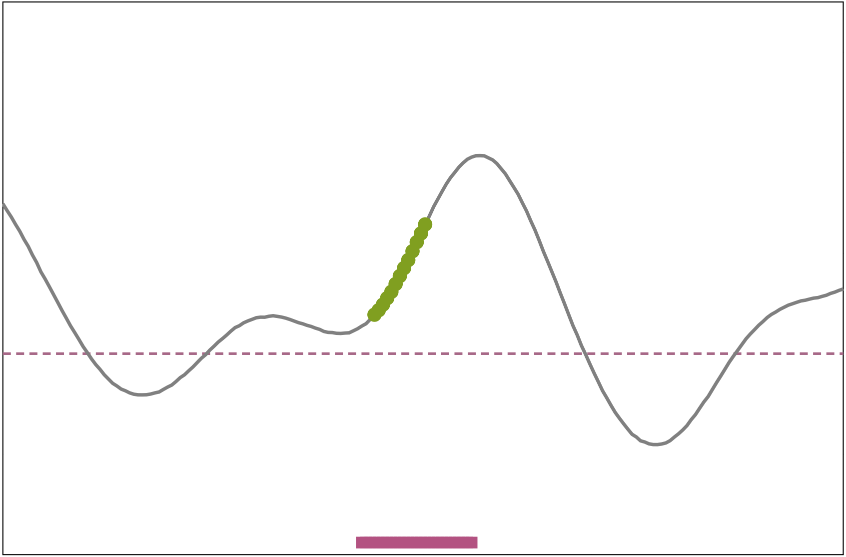
Certifying safety



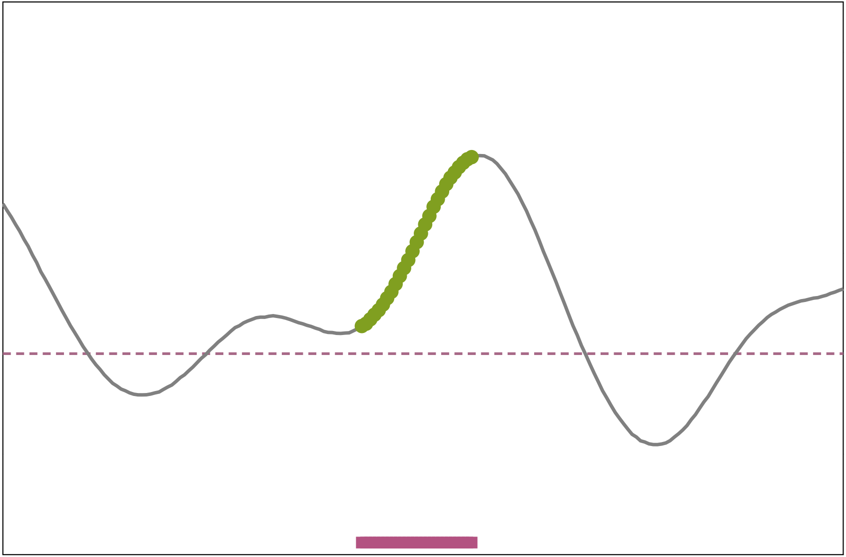
Certifying safety



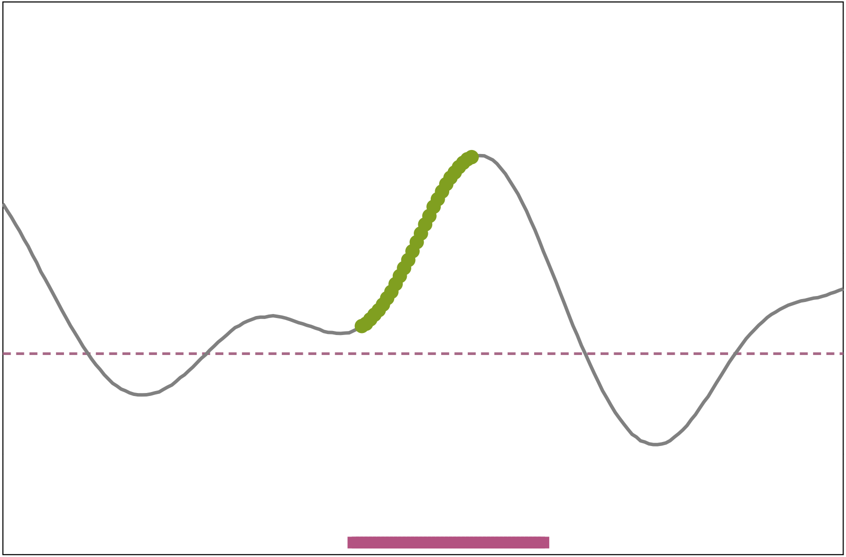
Certifying safety



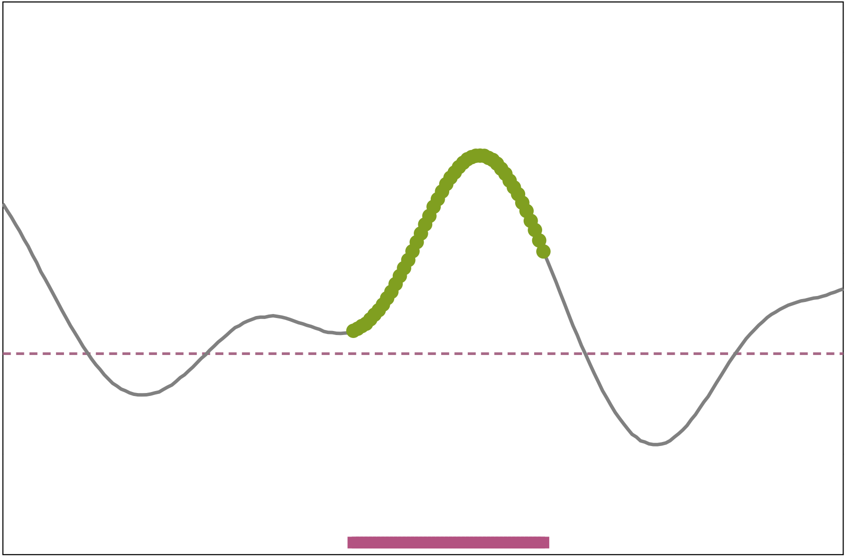
Certifying safety



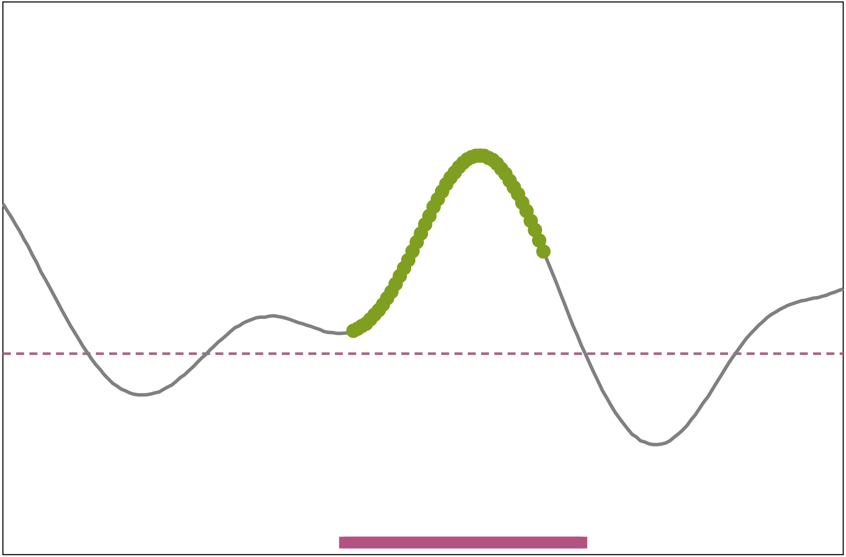
Certifying safety



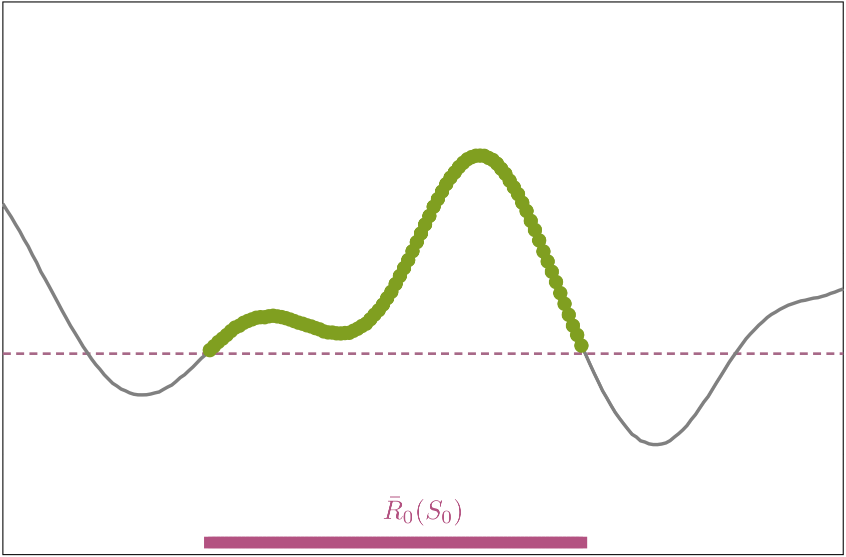
Certifying safety



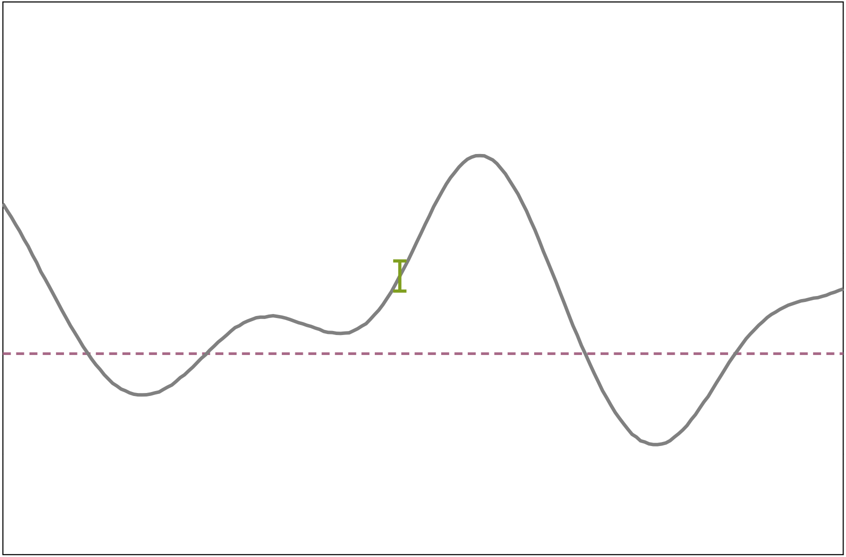
Certifying safety



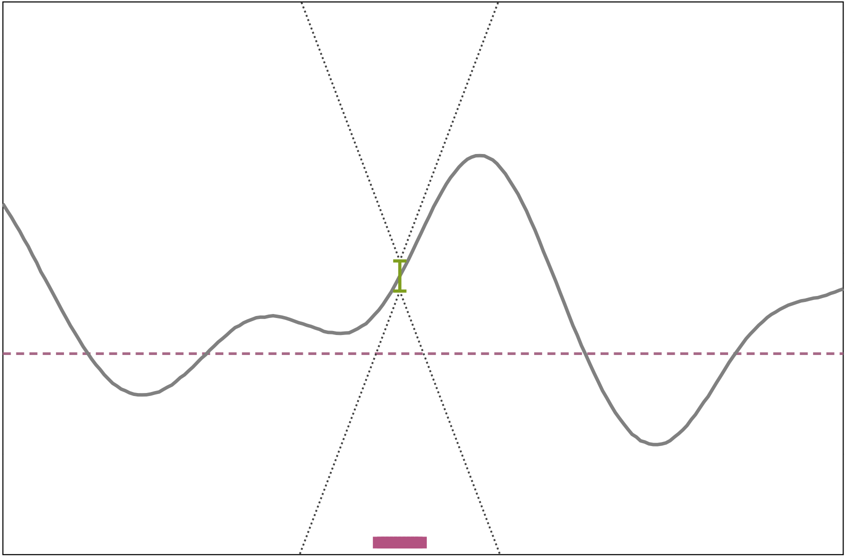
Certifying safety



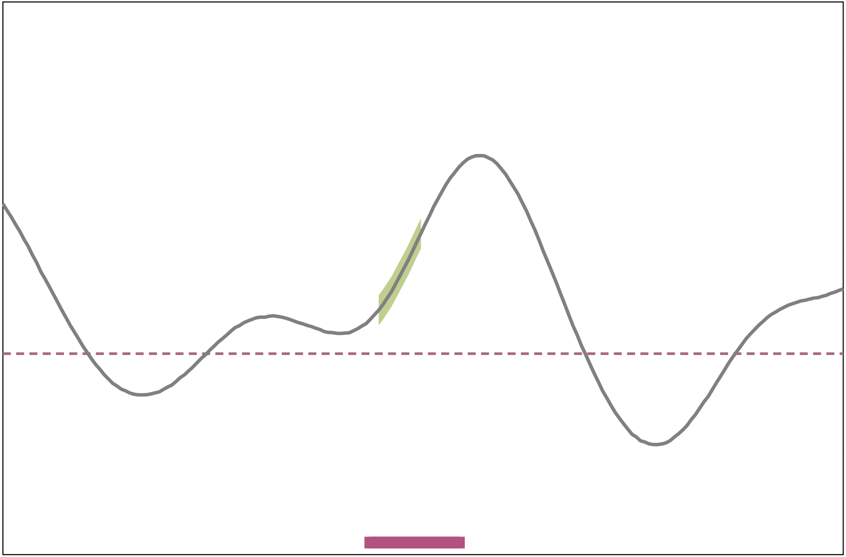
Certifying safety



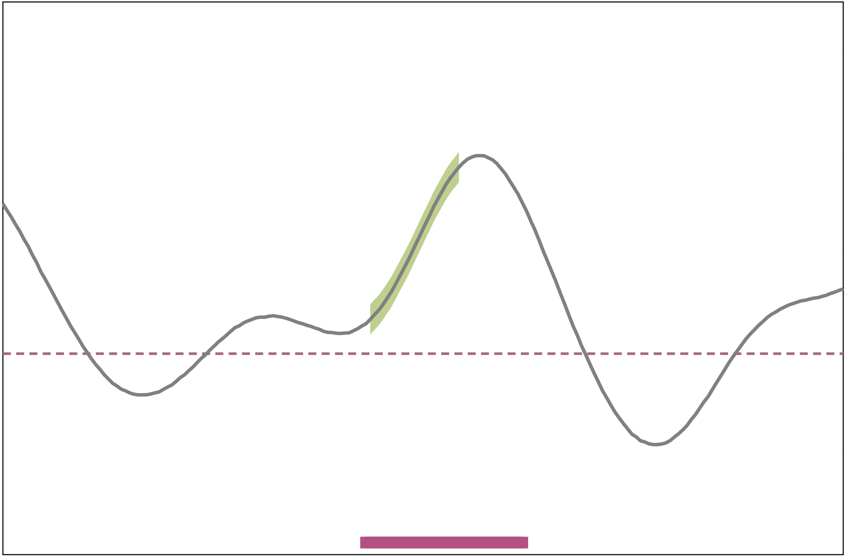
Certifying safety



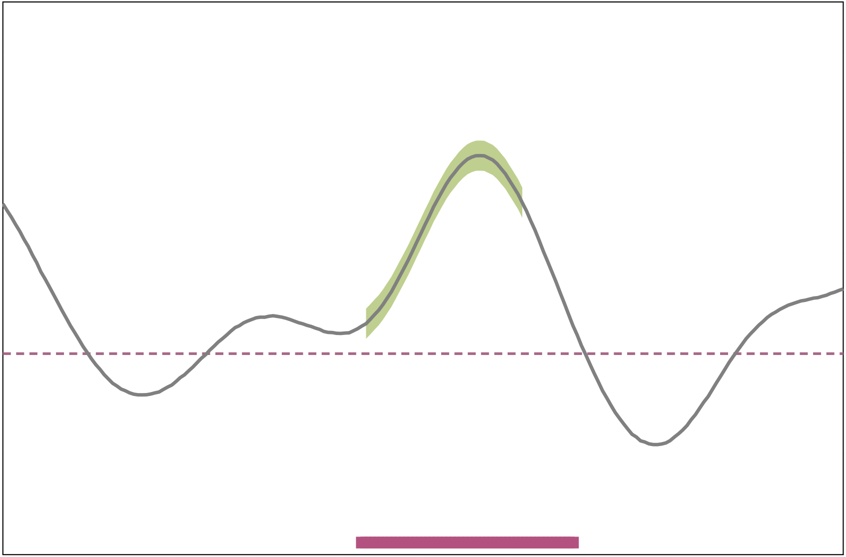
Reachability



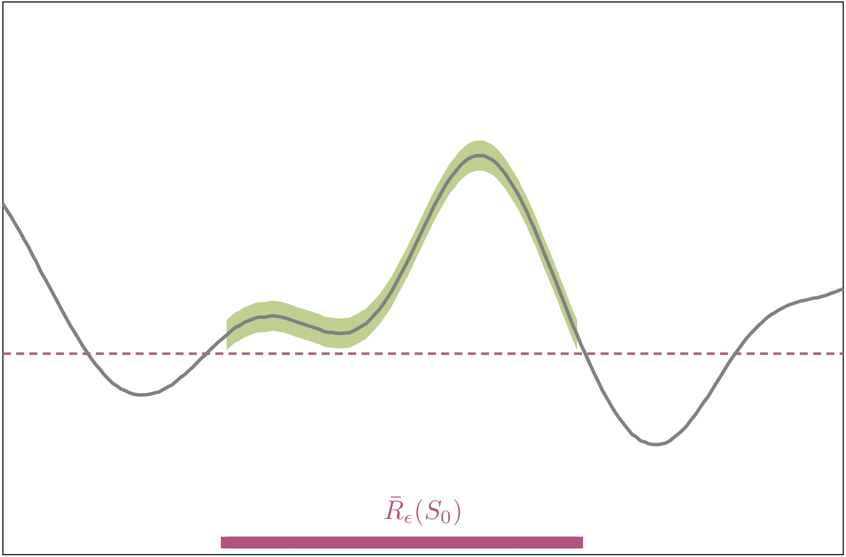
Reachability



Reachability



Reachability



Reconsidering optimization

- Initial goal of finding $f^* = \max_{x \in D} f(x)$ is unrealistic

- Initial goal of finding $f^* = \max_{x \in D} f(x)$ is unrealistic
- Instead, aim for the ϵ-reachable maximum

$$f^*_\epsilon = \max_{x \in \bar{R}_\epsilon(S_0)} f(x)$$

- Initial goal of finding $f^* = \max_{x \in D} f(x)$ is unrealistic
- Instead, aim for the ϵ-reachable maximum

$$f^*_\epsilon = \max_{x \in \bar{R}_\epsilon(S_0)} f(x)$$

- Smaller $\epsilon \rightarrow$ stricter goal $\rightarrow$ need more samples

First attempt: Safe-UCB

- Keep set S_t of certified safe points (starting with S_0)

First attempt: Safe-UCB

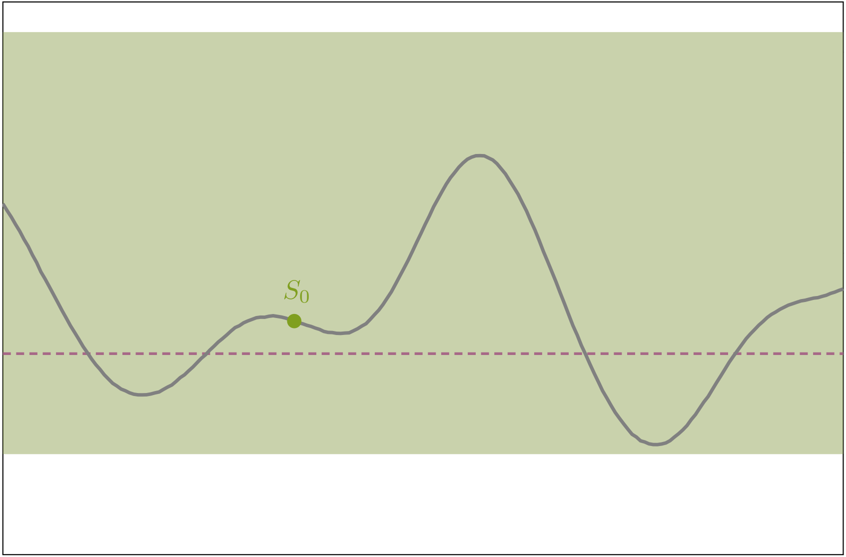
- Keep set S_t of certified safe points (starting with S_0)

- Use Lipschitz property with GP lower bounds to certify safety

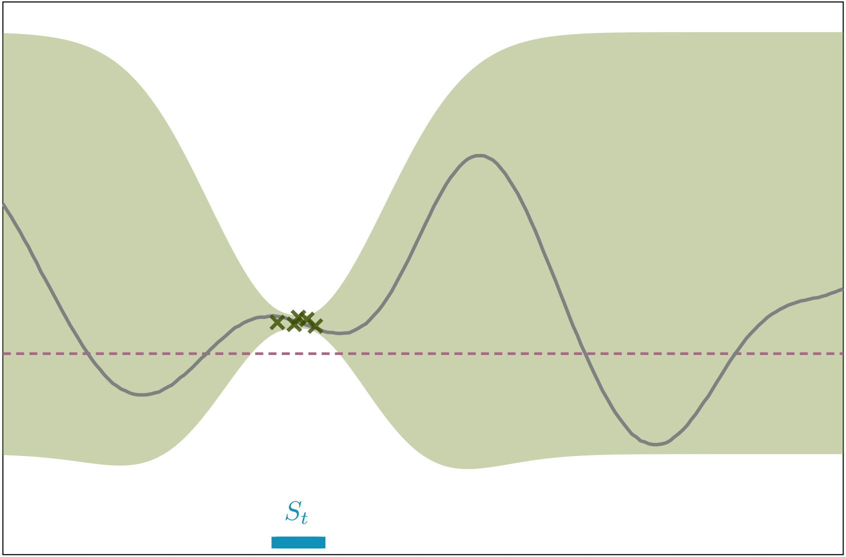
First attempt: Safe-UCB

- Keep set S_t of certified safe points (starting with S_0)
- Use Lipschitz property with GP lower bounds to certify safety
- $x_t = \text{argmax}_{x \in S_t} u_t(x)$

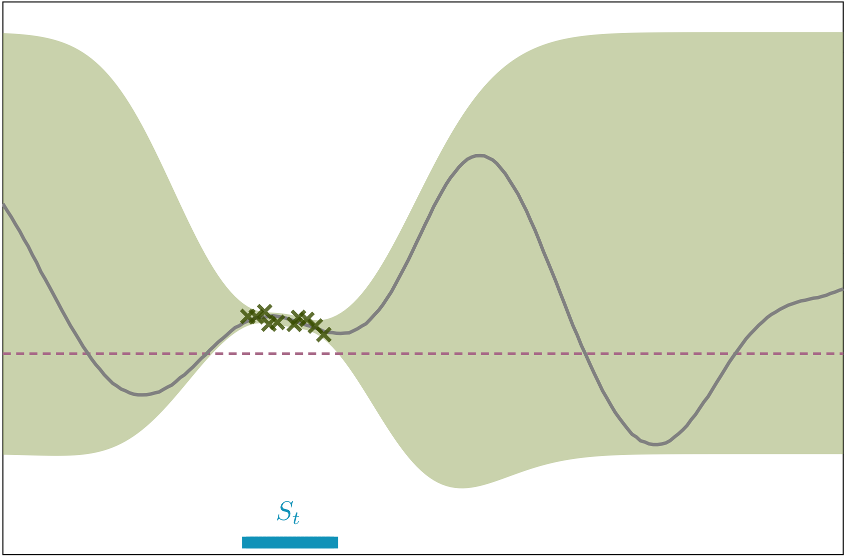
Safe-UCB example ($t = 0$)



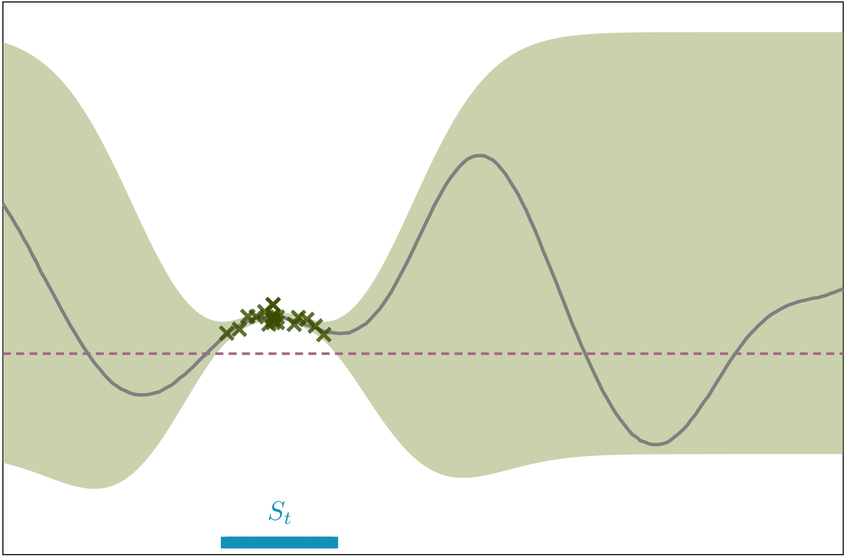
Safe-UCB example ($t = 5$)



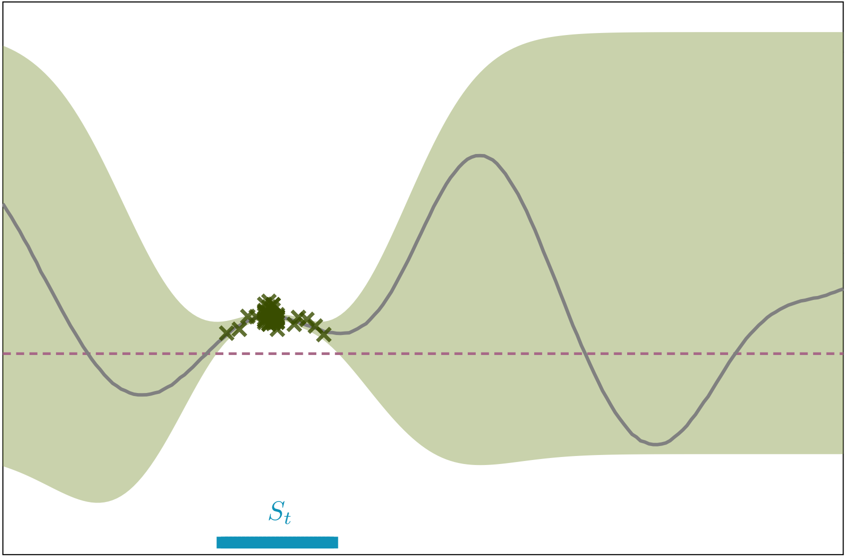
Safe-UCB example ($t = 10$)



Safe-UCB example ($t = 20$)



Safe-UCB example ($t = 50$)



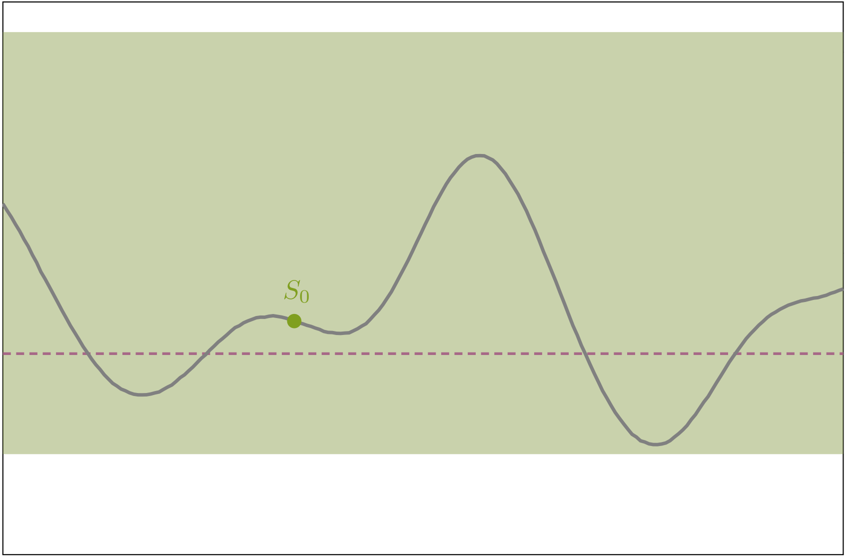
- Encourage expansion of $S_t \rightarrow$ keep set $G_t \subseteq S_t$ of potential expanders

- Encourage expansion of $S_t \rightarrow$ keep set $G_t \subseteq S_t$ of potential expanders
- Encourage locating the maximum within $S_t \rightarrow$ keep set $M_t \subseteq S_t$ of potential maximizers

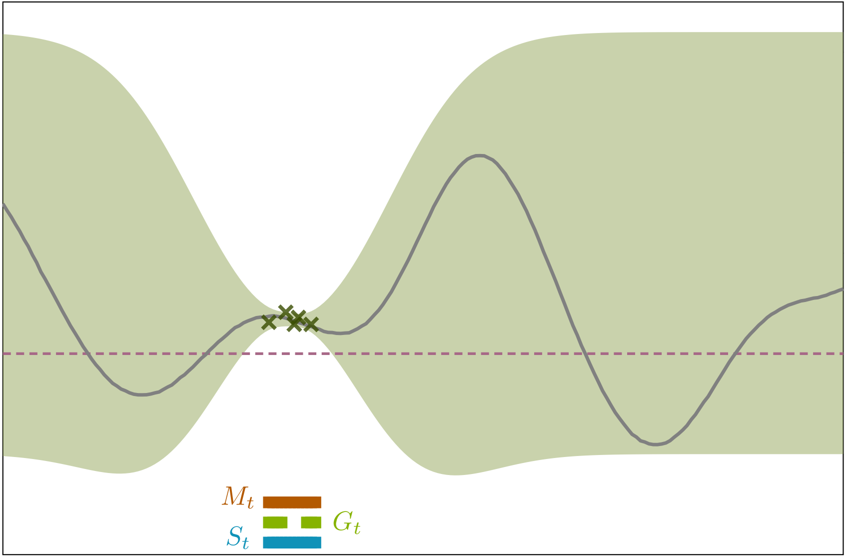
SafeOpt

- Encourage expansion of $S_t \rightarrow$ keep set $G_t \subseteq S_t$ of potential expanders
- Encourage locating the maximum within $S_t \rightarrow$ keep set $M_t \subseteq S_t$ of potential maximizers
- Pick most uncertain point within $G_t \cup M_t$

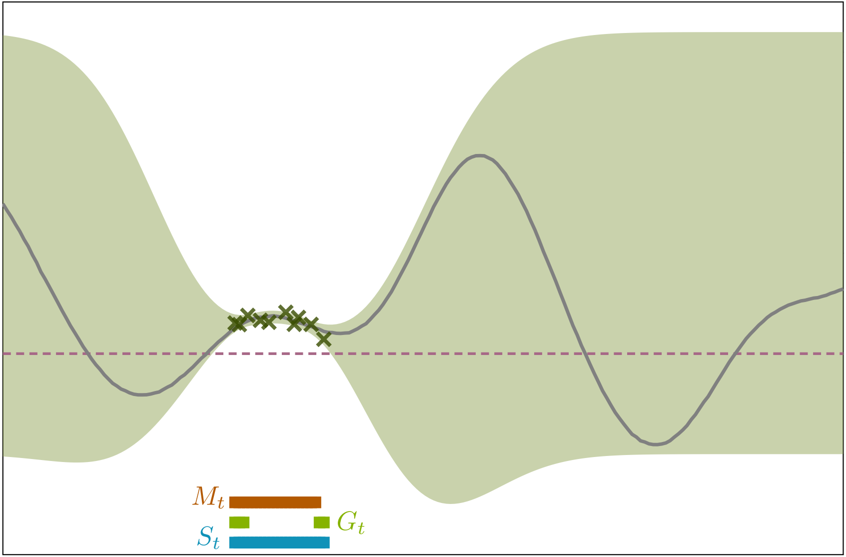
SafeOpt example ($t = 0$)



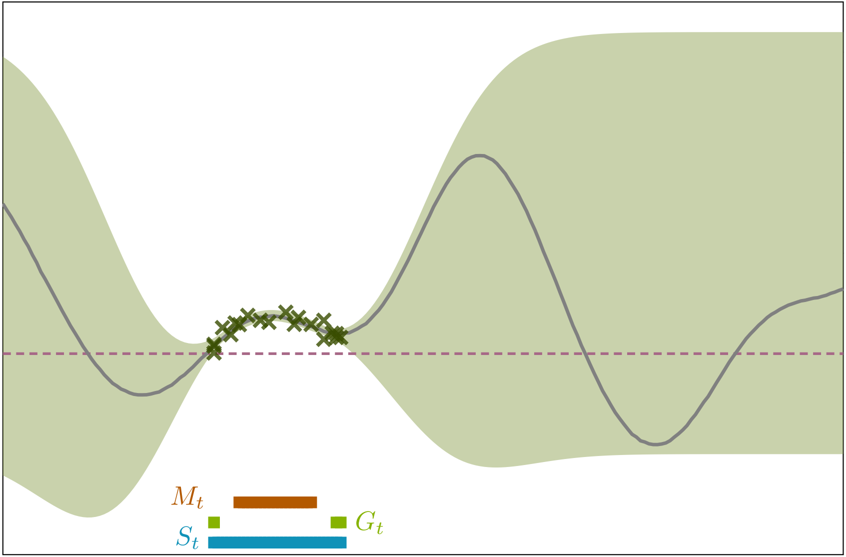
SafeOpt example ($t = 5$)



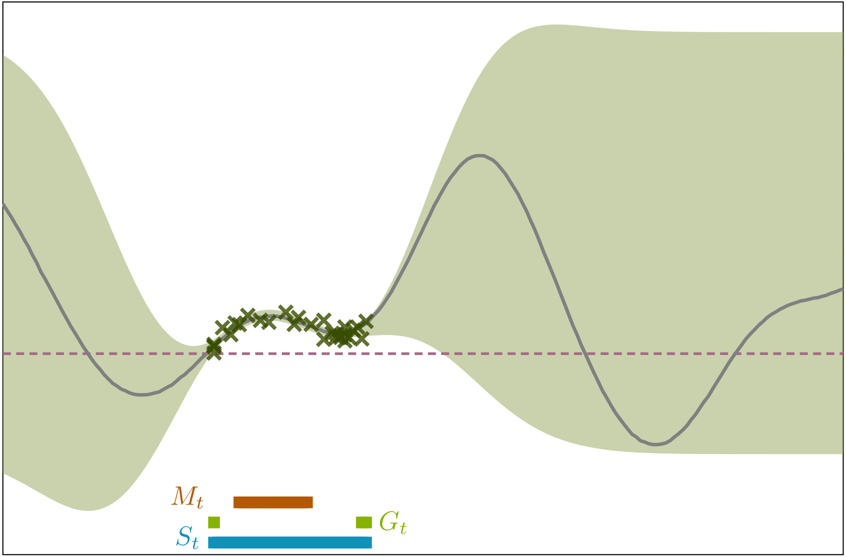
SafeOpt example ($t = 10$)

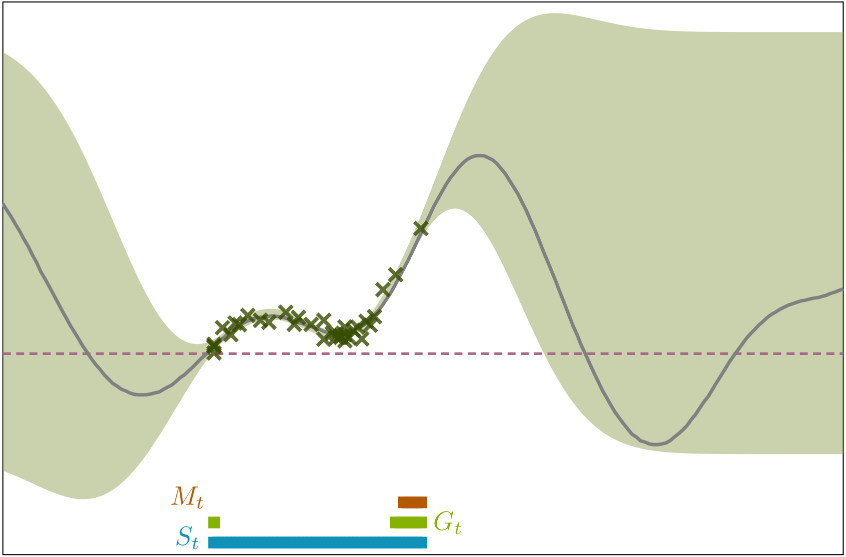


SafeOpt example ($t = 20$)

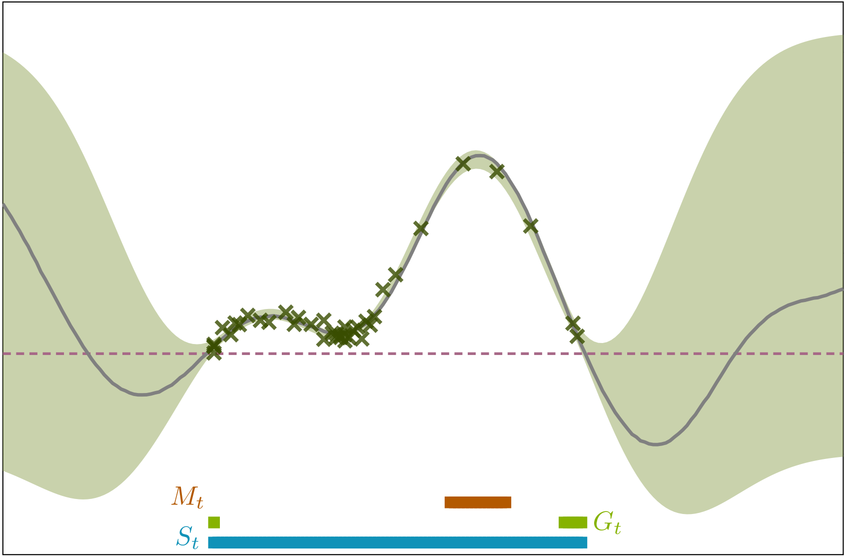


SafeOpt example ($t = 30$)

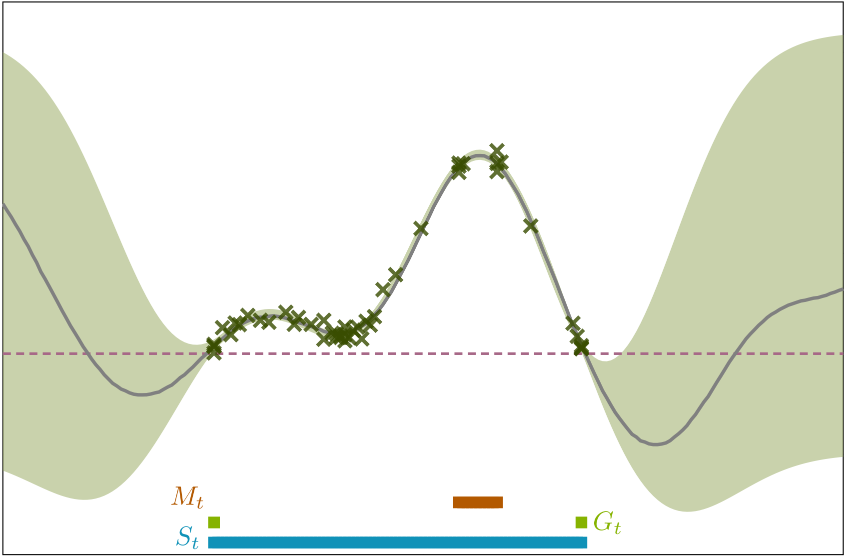



M_t
G_t
S_t

SafeOpt example ($t = 40$)



SafeOpt example ($t = 50$)



Input: sample set D,
kernel k,
Lipschitz constant L,
seed set S_0,
safety threshold h

for $t = 1, 2, \ldots$ **do**
  Update S_t, G_t, and M_t
  $x_t \leftarrow \text{argmax}_{x \in G_t \cup M_t} (u_t(x) - \ell_t(x))$
  $y_t \leftarrow f(x_t) + n_t$
  Update GP estimates
end for

Theorem

Assumptions

- f has bounded norm in the RKHS defined by k
- f is L-Lipschitz continuous
- n_t is a uniformly bounded martingale difference sequence

Theorem

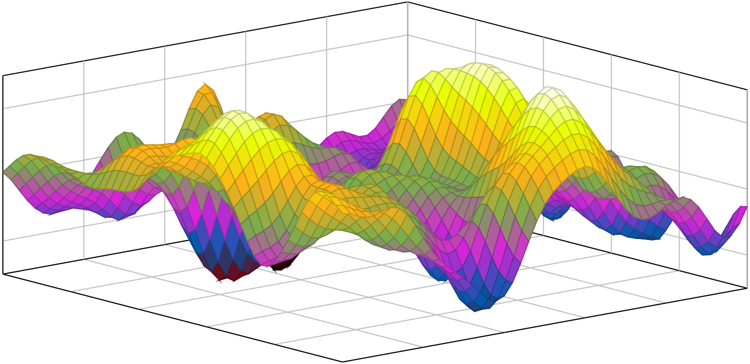
Assumptions

- f has bounded norm in the RKHS defined by k
- f is L-Lipschitz continuous
- n_t is a uniformly bounded martingale difference sequence

Under suitable scaling of the GP confidence intervals, the following jointly hold w.h.p.

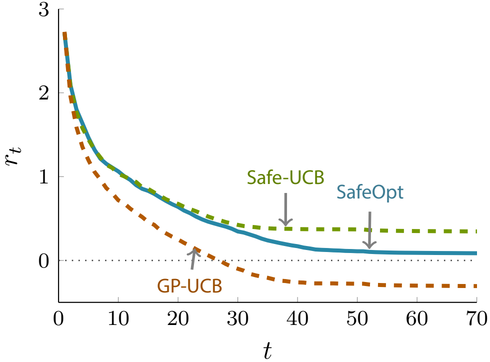
- $\forall t \geq 1$, $f(x_t) \geq h$
- $\forall t \geq t^*$, $f(\hat{x}_t) \geq f^*_\epsilon - \epsilon$

Experiment 1: Synthetic



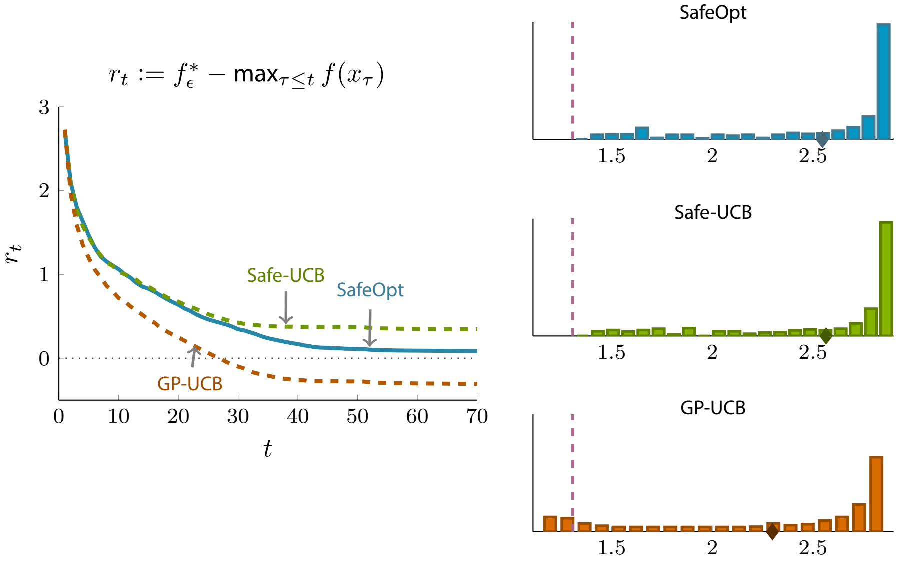
- Draw 100 random 2-D functions from GP prior (sq. exponential kernel)
- Use random singleton seed set S_0 per function
- Run 100 iterations of each algorithm

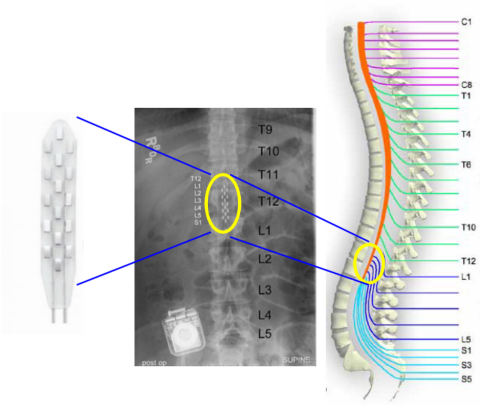
$$r_t := f_\epsilon^* - \max_{\tau \le t} f(x_\tau)$$

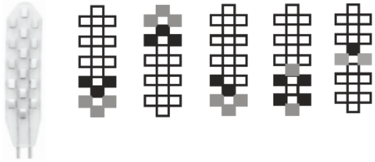


Experiment 1: Synthetic

$$r_t := f_\epsilon^* - \max_{\tau \le t} f(x_\tau)$$

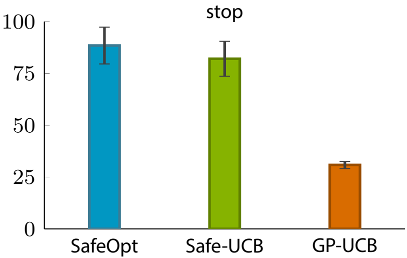




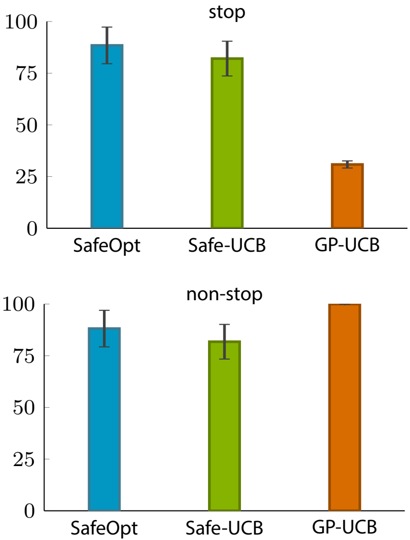


- Electrode configurations are represented by points in $\mathbb{R}^4$
- Fit sq. exponential ARD kernel
- Run 300 iterations of each algorithm

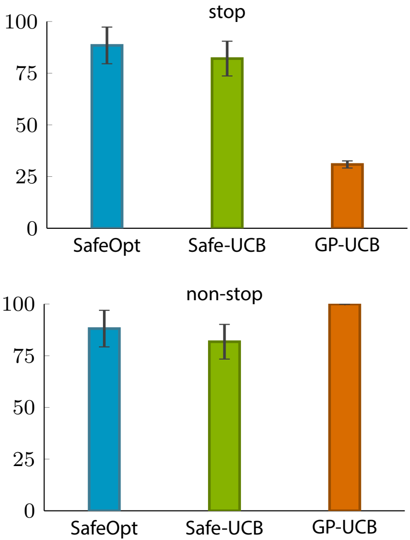
Experiment 2: Spinal cord therapy

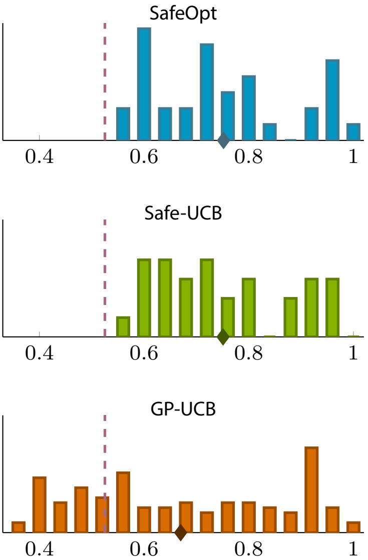


Experiment 2: Spinal cord therapy



Experiment 2: Spinal cord therapy





Conclusion

Recap

- We formulated safe optimization using the concept of reachability
- We proposed SafeOpt, an algorithm with theoretical guarantees

Conclusion

Recap

- We formulated safe optimization using the concept of reachability
- We proposed SafeOpt, an algorithm with theoretical guarantees

What we skipped here

- Rigorous theoretical setup and analysis
- Another application: safe movie recommendation