Safely Learning Controlled Stochastic Dynamics

Workshop on data-driven control and analysis of dynamical systems - ENSEEIHT - 30/09/2025

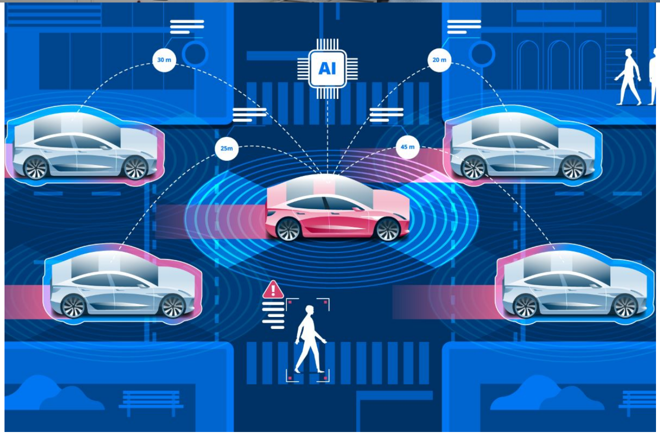
Riccardo Bonalli

Laboratoire des Signaux et Systèmes
CNRS and Université Paris-Saclay

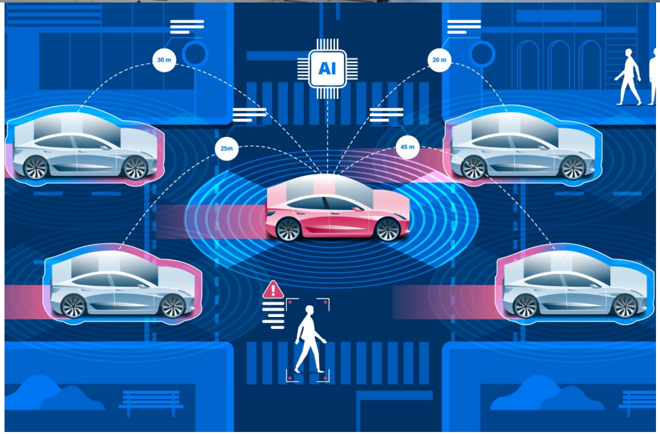
Contents

Contents

Challenges in controlling autonomous systems

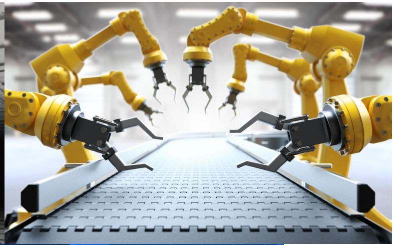


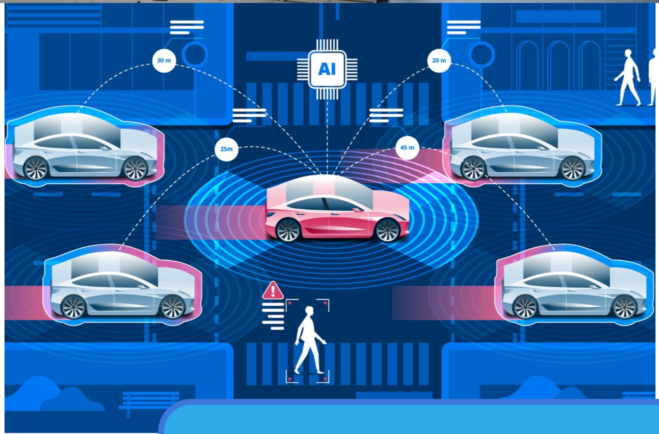
Challenges in controlling autonomous systems



Though unwanted events might happen…

Challenges in controlling autonomous systems





Though unwanted events might happen…

Let us use risk-averse control!

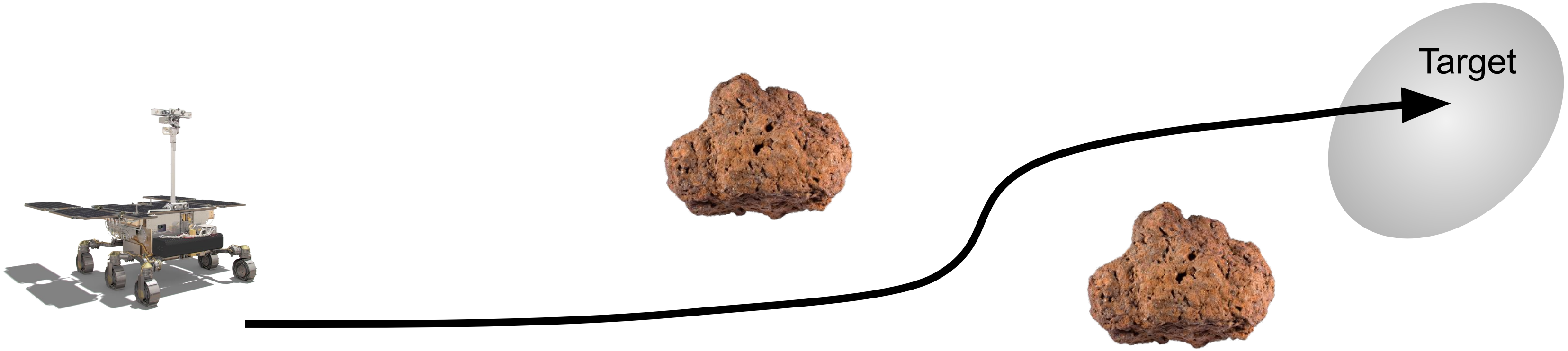
Mathematically modeling risk-averse control

$$dX(t) = b(X(t), u(t, X(t)))dt$$

Mathematically modeling risk-averse control

$$dX(t) = b(X(t), u(t, X(t)))dt$$

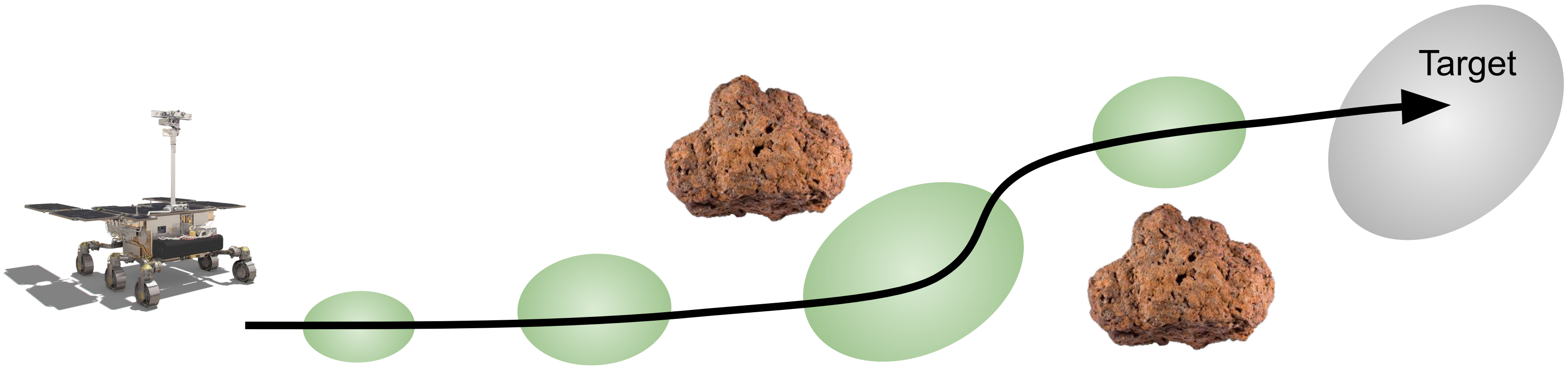
$$g(X(t)) \geq 0,\ t \in [0, T] \quad \text{and} \quad g_T(X(T)) \geq 0$$



Mathematically modeling risk-averse control

$$dX(t) = b(X(t), u(t, X(t)))dt$$

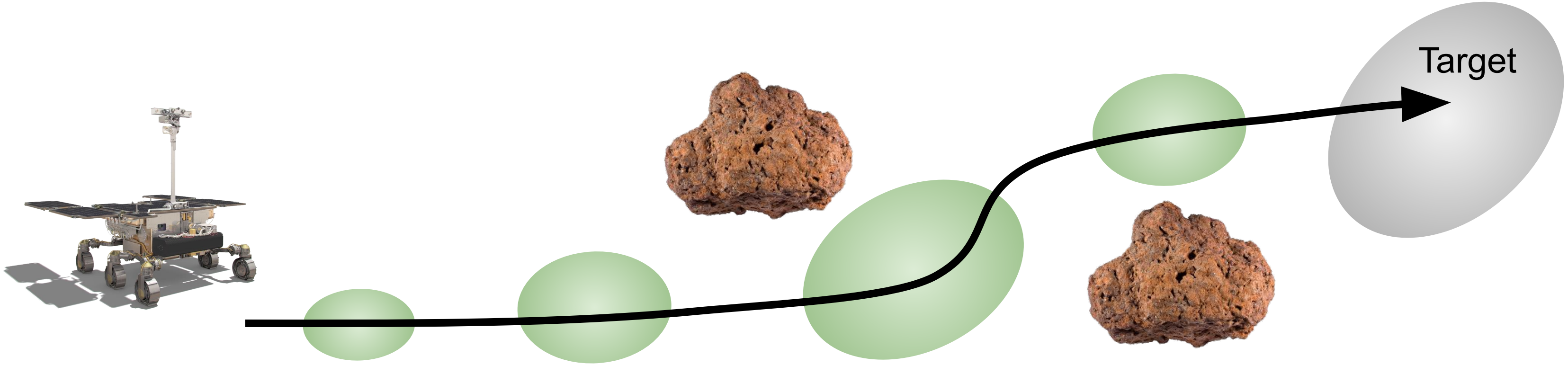
$$g(X(t)) \geq 0,\ t \in [0, T] \quad \text{and} \quad g_T(X(T)) \geq 0$$



Mathematically modeling risk-averse control

$$dX(t) = b(X(t), u(t, X(t)))dt$$
$$+ a(X(t), u(t, X(t)))dW(t)$$

$$g(X(t)) \geq 0, \ t \in [0, T] \quad \text{and} \quad g_T(X(T)) \geq 0$$

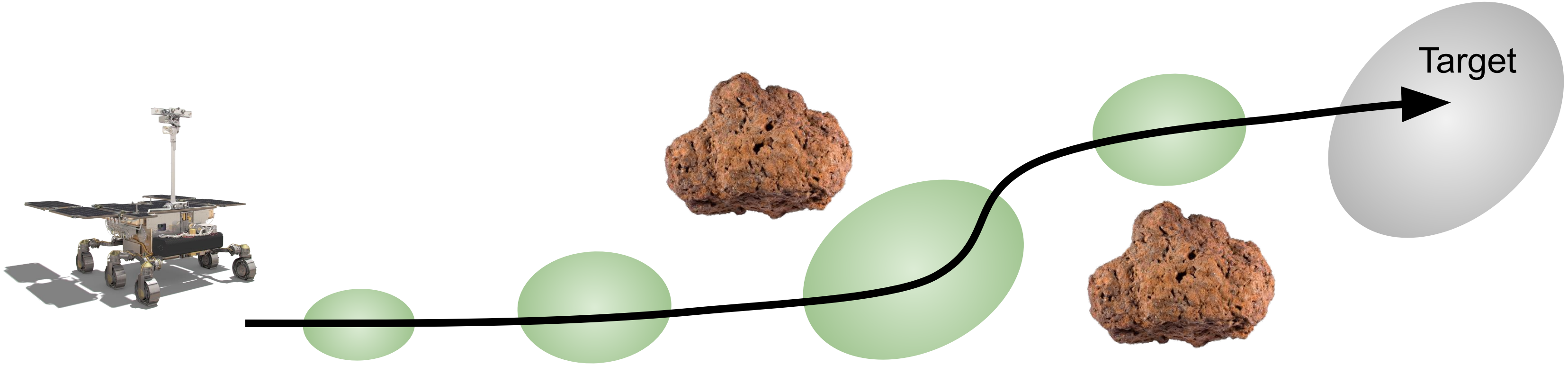


Mathematically modeling risk-averse control

$$dX(t) = b(X(t), u(t, X(t)))dt$$
$$+ a(X(t), u(t, X(t)))dW(t)$$

$$\cancel{g(X(t)) \geq 0,\ t \in [0, T] \quad \text{and} \quad g_T(X(T)) \geq 0}$$

$$\inf_{t \in [0,T]} \mathbb{P}(g(X(t)) \geq 0) \geq 1 - \varepsilon \quad \text{and} \quad \mathbb{P}(g(X_T(T)) \geq 0) \geq 1 - \varepsilon$$



Mathematically modeling risk-averse control

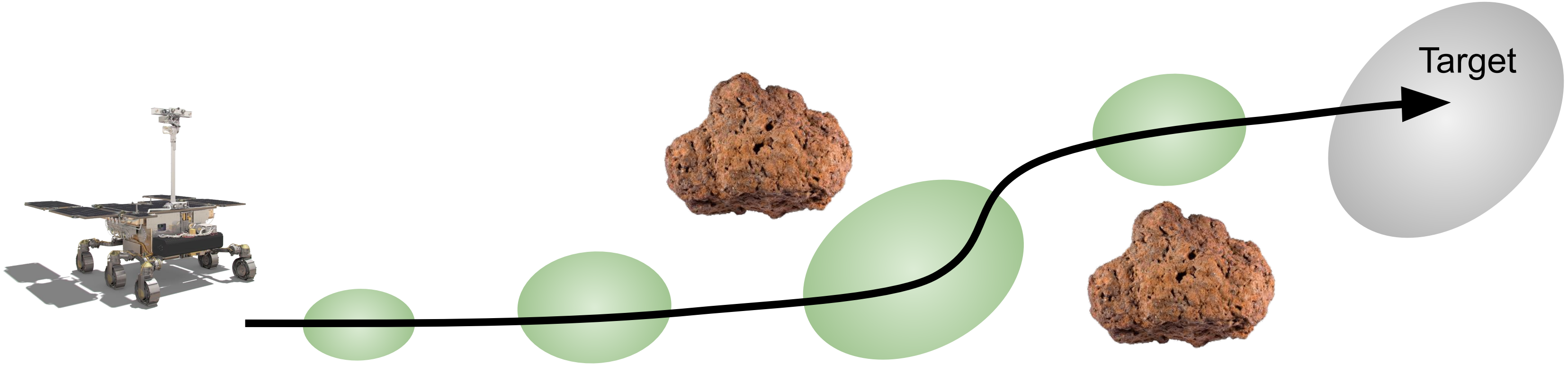
$$dX(t) = b(X(t), u(t, X(t)))dt$$
$$+ a(X(t), u(t, X(t)))dW(t)$$

$$\cancel{g(X(t)) \geq 0,\ t \in [0, T] \quad \text{and} \quad g_T(X(T)) \geq 0}$$

Problem

Under unknown drifts and diffusions, how to learn the space of safe controls with guarantees?

$$\inf_{t \in [0,T]} \mathbb{P}(g(X(t)) \geq 0) \geq 1 - \varepsilon \quad \text{and} \quad \mathbb{P}(g(X_T(T)) \geq 0) \geq 1 - \varepsilon$$



Several approaches have already been proposed…

Here are a few comments

- Much of the literature focuses on “discrete”, i.e., discrete-time or -state systems
- Bayesian methods provide some of the strongest high-confidence safety guarantees

Several approaches have already been proposed…

Here are a few comments

- Much of the literature focuses on “discrete”, i.e., discrete-time or -state systems
- Bayesian methods provide some of the strongest high-confidence safety guarantees

- For “continuous” systems, offline learning and online adaptation for nonlinear systems
- Lyapunov theory offers formal certification but often requires full dynamics knowledge

Several approaches have already been proposed…

Here are a few comments

- Much of the literature focuses on “discrete”, i.e., discrete-time or -state systems
- Bayesian methods provide some of the strongest high-confidence safety guarantees

- For “continuous” systems, offline learning and online adaptation for nonlinear systems
- Lyapunov theory offers formal certification but often requires full dynamics knowledge

- Weak point: prior models of the system dynamics or safety functions are required

Several approaches have already been proposed…

Here are a few comments

- Much of the literature focuses on “discrete”, i.e., discrete-time or -state systems
- Bayesian methods provide some of the strongest high-confidence safety guarantees

- For “continuous” systems, offline learning and online adaptation for nonlinear systems
- Lyapunov theory offers formal certification but often requires full dynamics knowledge

- Weak point: prior models of the system dynamics or safety functions are required

Presentation based on a collaboration with L. Brogat-Motte and A. Rudi, work recently accepted at NeurIPS 2025

Contents

Formal problem setting

- Control parametrization

$$u_\theta \,:\, [0, T_{\max}] \times \mathbb{R}^n \to \mathbb{R}^d \qquad \theta \in D \subset \mathbb{R}^m, \text{ where } D \text{ is a compact subset of } \mathbb{R}^m$$

Formal problem setting

- Control parametrization

$$u_\theta \; : \; [0, T_{\max}] \times \mathbb{R}^n \; \to \; \mathbb{R}^d \qquad \theta \in D \subset \mathbb{R}^m, \text{ where } D \text{ is a compact subset of } \mathbb{R}^m$$

- Safety functions

$$X_u \quad \text{solution to} \quad \mathrm{d}X(t) = b(X(t), u(t, X(t)))\,\mathrm{d}t + a(X(t), u(t, X(t)))\,\mathrm{d}W(t)$$

$X_u(0) \sim p_0.$ Define $\quad s(\theta, t) \triangleq \mathbb{P}\left(g(X_{u_\theta}(t)) \geq 0\right) \quad$ and $\quad s^\infty(\theta, T) \triangleq \inf_{t \in [0,T]} s(\theta, t)$

Formal problem setting

- Control parametrization

$$u_\theta \,:\, [0, T_{\max}] \times \mathbb{R}^n \to \mathbb{R}^d \qquad \theta \in D \subset \mathbb{R}^m, \text{ where } D \text{ is a compact subset of } \mathbb{R}^m$$

- Safety functions

$$X_u \quad \text{solution to} \quad \mathrm{d}X(t) = b(X(t), u(t, X(t)))\,\mathrm{d}t + a(X(t), u(t, X(t)))\,\mathrm{d}W(t)$$

$X_u(0) \sim p_0$. Define $s(\theta, t) \triangleq \mathbb{P}\left(g(X_{u_\theta}(t)) \geq 0\right)$ and $s^\infty(\theta, T) \triangleq \inf_{t \in [0,T]} s(\theta, t)$

- Learning problem

Collect data $(\theta_k, X_{u_{\theta_k}}(w_i^k, t_l))_{k \in \{1,\dots,K\},\, i \in \{1,\dots,Q\},\, l \in \{1,\dots,M_k\}}$ to “maximally cover” D,

while $s^\infty(\theta_k, T_k) \geq 1 - \varepsilon, \quad \text{for each } k \in [\![1, K]\!]$ (here $(T_k)_{k=1}^K = (t_{M_k})_{k=1}^K \leq T_{\max}$)

Formal problem setting

- Control parametrization

$$u_\theta \; : \; [0, T_{\max}] \times \mathbb{R}^n \to \mathbb{R}^d \qquad \theta \in D \subset \mathbb{R}^m, \text{ where } D \text{ is a compact subset of } \mathbb{R}^m$$

- Safety functions

$$X_u \quad \text{solution to} \quad \mathrm{d}X(t) = b(X(t), u(t, X(t)))\,\mathrm{d}t + a(X(t), u(t, X(t)))\,\mathrm{d}W(t)$$

$X_u(0) \sim p_0$. Define $\; s(\theta, t) \triangleq \mathbb{P}\left(g(X_{u_\theta}(t)) \geq 0\right) \;$ and $\; s^\infty(\theta, T) \triangleq \inf_{t \in [0,T]} s(\theta, t)$

- Learning problem

Collect data $\;(\theta_k, X_{u_{\theta_k}}(w_i^k, t_l))_{k \in \{1,\ldots,K\},\, i \in \{1,\ldots,Q\},\, l \in \{1,\ldots,M_k\}}\;$ to “maximally cover” D,

while $\; s^\infty(\theta_k, T_k) \geq 1 - \varepsilon, \quad \text{for each } k \in [\![1, K]\!] \;$ (here $(T_k)_{k=1}^K = (t_{M_k})_{k=1}^K \leq T_{\max}$)

In short: learning safe controls under constraints for the safety function!

Learning through resetting

- In practice, one explores by starting wandering around a known safe region (initial distribution)

Learning through resetting

- In practice, one explores by starting wandering around a known safe region (initial distribution)
- To enable "well-posed" learning, i.e., collecting iid samples, we introduce a **reset mechanism**

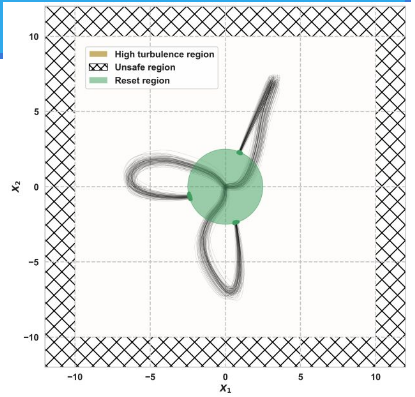
We let a function $h : \mathbb{R}^n \to \mathbb{R}$ define a region in the state space from which resets (to the initial distribution) are feasible.

Learning through resetting

- In practice, one explores by starting wandering around a known safe region (initial distribution)
- To enable "well-posed" learning, i.e., collecting iid samples, we introduce a **reset mechanism**

We let a function $h : \mathbb{R}^n \to \mathbb{R}$ define a region in the state space from which resets (to the initial distribution) are feasible. Specifically:

$$h(X_{u_\theta}(t)) \geq 0 \implies \left\{ \exists u,\ \widehat{T} \geq t : \ \mathbb{P}(X_u|_{X_u(t)=X_{u_\theta}(t)}(\widehat{T}) \in \operatorname{supp} p_0) = 1 \right.$$



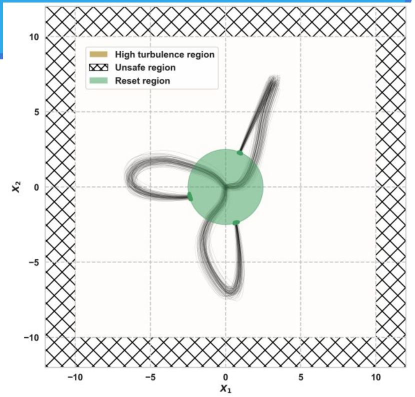
Learning through resetting

- In practice, one explores by starting wandering around a known safe region (initial distribution)
- To enable "well-posed" learning, i.e., collecting iid samples, we introduce a **reset mechanism**

We let a function $h : \mathbb{R}^n \to \mathbb{R}$ define a region in the state space from which resets (to the initial distribution) are feasible. Specifically:

$$h(X_{u_\theta}(t)) \geq 0 \implies \left\{ \exists u,\ \widehat{T} \geq t :\ \mathbb{P}(X_u|_{X_u(t)=X_{u_\theta}(t)}(\widehat{T}) \in \operatorname{supp} p_0) = 1 \right.$$

Not that strong in practice: $\widehat{T}$ may be a stopping time!



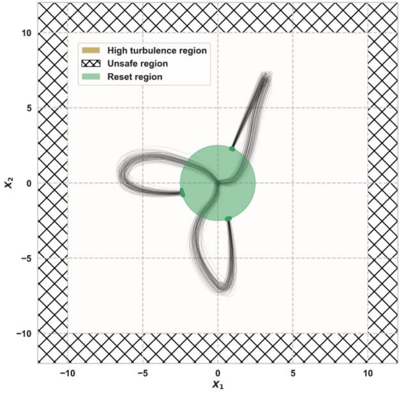
Learning through resetting

- In practice, one explores by starting wandering around a known safe region (initial distribution)
- To enable “well-posed” learning, i.e., collecting iid samples, we introduce a **reset mechanism**

We let a function $h : \mathbb{R}^n \to \mathbb{R}$ define a region in the state space from which resets (to the initial distribution) are feasible. Specifically:

$$h(X_{u_\theta}(t)) \geq 0 \implies \left\{ \exists u,\ \widehat{T} \geq t :\ \mathbb{P}(X_u|_{X_u(t)=X_{u_\theta}(t)}(\widehat{T}) \in \operatorname{supp} p_0) = 1 \right.$$

Not that strong in practice: $\widehat{T}$ may be a stopping time!



- By defining the reset function $r(\theta, t) \triangleq \mathbb{P}\left(h(X_{u_\theta}(t)) \geq 0\right)$,
we now learn subject to the **safety and reset constraints**

$$r(\theta_k, T_k) \geq 1 - \xi \quad \text{and} \quad s^\infty(\theta_k, T_k) \geq 1 - \varepsilon, \quad \text{for each } k \in [\![1, K]\!]$$

Learning through resetting

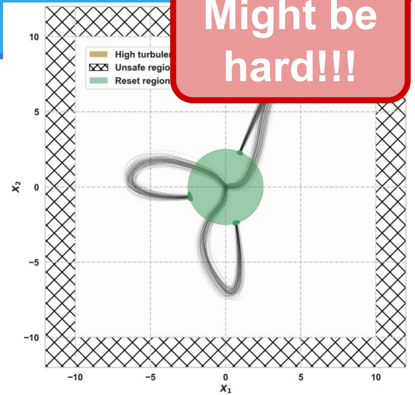
- In practice, one explores by starting wandering around a known safe region (initial distribution)
- To enable “well-posed” learning, i.e., collecting iid samples, we introduce a **reset mechanism**

We let a function $h : \mathbb{R}^n \to \mathbb{R}$ define a region in the state space from which resets (to the initial distribution) are feasible. Specifically:

$$h(X_{u_\theta}(t)) \geq 0 \implies \begin{cases} \exists u,\ \widehat{T} \geq t :\ \mathbb{P}(X_u|_{X_u(t)=X_{u_\theta}(t)}(\widehat{T}) \in \operatorname{supp} p_0) = 1 \\ \text{and} \quad \mathbb{P}\left(\inf_{s\in[t,\widehat{T}]} g(X_u(s)) \geq 0 \right) \geq 1-\varepsilon \end{cases}$$

Not that strong in practice: $\widehat{T}$ may be a stopping time!

Might be hard!!!



- By defining the reset function $r(\theta,t) \triangleq \mathbb{P}\left(h(X_{u_\theta}(t)) \geq 0\right)$,
we now learn subject to the **safety and reset constraints**

$$r(\theta_k, T_k) \geq 1-\xi \quad \text{and} \quad s^\infty(\theta_k, T_k) \geq 1-\varepsilon, \quad \text{for each } k \in [\![1, K]\!]$$

With that in mind…

Formal (ultimate) goal

For given precisions ε, ξ > 0, collect data to best estimate (learn)

$$\Gamma \triangleq \left\{(\theta, T) \in D \times [0, T_{\max}] \;\middle|\; s^{\infty}(\theta, T) \geq 1 - \varepsilon \quad \text{and} \quad r(\theta, T) \geq 1 - \xi\right\}$$

with **rates of convergence**.

Assumptions and algorithm - part I

Assumption (A1) (Initial safe controls). *For $\varepsilon \in [0,1]$, a non-empty set $S_0 \subset D \times [0, T_{\max}]$ is provided such that*

$$s(\theta, t) \geq 1 - \varepsilon \quad \textit{for all} \quad (\theta, t) \in S_0.$$

Assumptions and algorithm - part I

Assumption (A1) (Initial safe controls). *For* $\varepsilon \in [0,1]$, *a non-empty set* $S_0 \subset D \times [0, T_{\max}]$ *is provided such that*

$$s(\theta, t) \geq 1 - \varepsilon \quad \textit{for all} \quad (\theta, t) \in S_0.$$

Assumption (A2) (Initial resetting controls). *For* $\xi \in [0,1]$, *a non-empty set* $R_0 \subset D \times [0, T_{\max}]$ *is provided such that*

$$r(\theta, t) \geq 1 - \xi \quad \textit{for all} \quad (\theta, t) \in R_0.$$

Assumptions and algorithm - part I

Assumption (A1) (Initial safe controls). *For $\varepsilon \in [0,1]$, a non-empty set $S_0 \subset D \times [0, T_{\max}]$ is provided such that*

$$s(\theta, t) \geq 1 - \varepsilon \quad \textit{for all} \quad (\theta, t) \in S_0.$$

Assumption (A2) (Initial resetting controls). *For $\xi \in [0,1]$, a non-empty set $R_0 \subset D \times [0, T_{\max}]$ is provided such that*

$$r(\theta, t) \geq 1 - \xi \quad \textit{for all} \quad (\theta, t) \in R_0.$$

Let denote $p : (\theta, t, x) \in D \times [0, T_{\max}] \times \mathbb{R}^n \mapsto p_\theta(t, x)$,

where $p_\theta(t, x)$ is the density of the state $X_{u_\theta}(t)$ under control u_θ

Assumptions and algorithm - part I

Assumption (A1) (Initial safe controls). *For $\varepsilon \in [0,1]$, a non-empty set $S_0 \subset D \times [0, T_{\max}]$ is provided such that*

$$s(\theta, t) \geq 1 - \varepsilon \quad \textit{for all} \quad (\theta, t) \in S_0.$$

Assumption (A2) (Initial resetting controls). *For $\xi \in [0,1]$, a non-empty set $R_0 \subset D \times [0, T_{\max}]$ is provided such that*

$$r(\theta, t) \geq 1 - \xi \quad \textit{for all} \quad (\theta, t) \in R_0.$$

Let denote $p : (\theta, t, x) \in D \times [0, T_{\max}] \times \mathbb{R}^n \mapsto p_\theta(t, x)$,

where $p_\theta(t, x)$ is the density of the state $X_{u_\theta}(t)$ under control u_θ

Assumption (A3) (Smoothness of system dynamics). *The map p lies in the Sobolev space $H^\nu(\mathbb{R}^{n+m+1})$ with $\nu > n/2$, where n and m denote the state and control parameter dimensions, respectively. Moreover,* $\sup_{x \in \mathbb{R}^n} \left\| p(\cdot, \cdot, x) \right\|_{H^\nu(\mathbb{R}^{m+1})} < +\infty$*,* $\sup_{(\theta,t) \in D \times [0, T_{\max}]} \| p(\theta, t, \cdot) \|_{H^\nu(\mathbb{R}^n)} < \infty$.

Assumptions and algorithm - part I

Assumption (A1) (Initial safe controls). *For* $\varepsilon \in [0,1]$, *a non-empty set* $S_0 \subset D \times [0, T_{\max}]$ *is provided such that*

$$s(\theta, t) \geq 1 - \varepsilon \quad \textit{for all} \quad (\theta, t) \in S_0.$$

Assumption (A2) (Initial resetting controls). *For* $\xi \in [0,1]$, *a non-empty set* $R_0 \subset D \times [0, T_{\max}]$ *is provided such that*

$$r(\theta, t) \geq 1 - \xi \quad \textit{for all} \quad (\theta, t) \in R_0.$$

Let denote $p : (\theta, t, x) \in D \times [0, T_{\max}] \times \mathbb{R}^n \mapsto p_\theta(t, x)$,

where $p_\theta(t,x)$ is the density of the state $X_{u_\theta}(t)$ under control u_θ

Assumption (A3) (Smoothness of system dynamics). *The map* p *lies in the Sobolev space* $H^\nu(\mathbb{R}^{n+m+1})$ *with* $\nu > n/2$, *where* n *and* m *denote the state and control parameter dimensions, respectively. Moreover,* $\sup_{x \in \mathbb{R}^n} \left\| p(\cdot, \cdot, x) \right\|_{H^\nu(\mathbb{R}^{m+1})} < +\infty$, $\sup_{(\theta,t) \in D \times [0, T_{\max}]} \|p(\theta, t, \cdot)\|_{H^\nu(\mathbb{R}^n)} < \infty$.

Assumptions and algorithm - part I

Assumption (A1) (Initial safe controls). *For $\varepsilon \in [0,1]$, a non-empty set $S_0 \subset D \times [0, T_{\max}]$ is provided such that*

$$s(\theta, t) \geq 1 - \varepsilon \quad \textit{for all} \quad (\theta, t) \in S_0.$$

Assumption (A2) (Initial resetting controls). *For $\xi \in [0,1]$, a non-empty set $R_0 \subset D \times [0, T_{\max}]$ is provided such that*

$$r(\theta, t) \geq 1 - \xi \quad \textit{for all} \quad (\theta, t) \in R_0.$$

Hinges upon (local) stochastic controllability

Let denote $p : (\theta, t, x) \in D \times [0, T_{\max}] \times \mathbb{R}^n \mapsto p_\theta(t, x)$,

where $p_\theta(t, x)$ is the density of the state $X_{u_\theta}(t)$ under control u_θ

Assumption (A3) (Smoothness of system dynamics). *The map p lies in the Sobolev space $H^\nu(\mathbb{R}^{n+m+1})$ with $\nu > n/2$, where n and m denote the state and control parameter dimensions, respectively. Moreover,* $\sup_{x \in \mathbb{R}^n} \left\| p(\cdot, \cdot, x) \right\|_{H^\nu(\mathbb{R}^{m+1})} < +\infty$, $\sup_{(\theta,t) \in D \times [0, T_{\max}]} \| p(\theta, t, \cdot) \|_{H^\nu(\mathbb{R}^n)} < \infty$.

Assumptions and algorithm - part I

Assumption (A1) (Initial safe controls). *For* $\varepsilon \in [0,1]$*, a non-empty set* $S_0 \subset D \times [0, T_{\max}]$ *is provided such that*

$$s(\theta, t) \geq 1 - \varepsilon \quad \textit{for all} \quad (\theta, t) \in S_0.$$

Assumption (A2) (Initial resetting controls). *For* $\xi \in [0,1]$*, a non-empty set* $R_0 \subset D \times [0, T_{\max}]$ *is provided such that*

$$r(\theta, t) \geq 1 - \xi \quad \textit{for all} \quad (\theta, t) \in R_0.$$

Hinges upon (local) stochastic controllability

Let denote $p : (\theta, t, x) \in D \times [0, T_{\max}] \times \mathbb{R}^n \mapsto p_\theta(t, x)$,

where $p_\theta(t,x)$ is the density of the state $X_{u_\theta}(t)$ under control u_θ

Assumption (A3) (Smoothness of system dynamics). *The map* p *lies in the Sobolev space* $H^\nu(\mathbb{R}^{n+m+1})$ *with* $\nu > n/2$*, where* n *and* m *denote the state and control parameter dimensions, respectively. Moreover,* $\sup_{x \in \mathbb{R}^n} \left\| p(\cdot, \cdot, x) \right\|_{H^\nu(\mathbb{R}^{m+1})} < +\infty$, $\sup_{(\theta,t) \in D \times [0, T_{\max}]} \left\| p(\theta, t, \cdot) \right\|_{H^\nu(}$

Model-free settings can be considered as well!

Assumptions and algorithm - part II

- Initialization (N = 0)

$$\Gamma_0 \triangleq \left\{(\theta, t, T) \in D \times [0, T_{\max}]^2 \,\middle|\, t \le T,\ (\theta, t') \in S_0 \text{ for all } t' \in [0, T],\ (\theta, T) \in R_0\right\}$$

Assumptions and algorithm - part II

- Initialization (N = 0)

$$\Gamma_0 \triangleq \left\{ (\theta, t, T) \in D \times [0, T_{\max}]^2 \,\middle|\, t \leq T,\ (\theta, t') \in S_0 \text{ for all } t' \in [0, T],\ (\theta, T) \in R_0 \right\}$$

Think of it as “observation” time to decrease uncertainty: we may observe the system not exclusively at states from which it is resettable

Assumptions and algorithm - part II

- At iteration N: 5 steps

    1. Collect Q iid samples $(X_{u_{\theta_N}}(w_i^N, t_N))_{i \in [\![1,Q]\!]}$

Showed later why these are safe!
Note the necessity of resetting…

Assumptions and algorithm - part II

- At iteration N: 5 steps

Showed later why these are safe!
Note the necessity of resetting…

1. Collect Q iid samples $(X_{u_{\theta_N}}(w_i^N, t_N))_{i \in [\![1,Q]\!]}$

2. Build the kernel density estimator $\hat{p}_{\theta_N, t_N}(x) \triangleq \frac{1}{Q} \sum_{i=1}^{Q} \rho_R(x - X_{u_{\theta_N}}(w_i^N, t_N))$ (B. and Rudi FoCM 2025)

Assumptions and algorithm - part II

- At iteration N: 5 steps

Showed later why these are safe!
Note the necessity of resetting…

1. Collect Q iid samples $(X_{u_{\theta_N}}(w_i^N, t_N))_{i \in [\![1,Q]\!]}$

2. Build the kernel density estimator $\hat{p}_{\theta_N,t_N}(x) \triangleq \frac{1}{Q}\sum_{i=1}^{Q} \rho_R(x - X_{u_{\theta_N}}(w_i^N, t_N))$ (B. and Rudi FoCM 2025)

   and compute safety/reset: $\hat{s}_{\theta_N,t_N} \triangleq \int_{\{x\in\mathbb{R}^n : g(x)\geq 0\}} \hat{p}_{\theta_N}(t_N, x)\, dx, \quad \hat{r}_{\theta_N,t_N} \triangleq \int_{\{x\in\mathbb{R}^n : h(x)\geq 0\}} \hat{p}_{\theta_N}(t_N, x)\, dx$

Assumptions and algorithm - part II

- At iteration N: 5 steps

Showed later why these are safe!
Note the necessity of resetting…

1. Collect Q iid samples $(X_{u_{\theta_N}}(w_i^N, t_N))_{i\in[\![1,Q]\!]}$

2. Build the kernel density estimator $\hat{p}_{\theta_N,t_N}(x) \triangleq \frac{1}{Q}\sum_{i=1}^{Q} \rho_R(x - X_{u_{\theta_N}}(w_i^N, t_N))$ (B. and Rudi FoCM 2025)

   and compute safety/reset: $\hat{s}_{\theta_N,t_N} \triangleq \int_{\{x\in\mathbb{R}^n : g(x)\geq 0\}} \hat{p}_{\theta_N}(t_N, x)\, dx, \quad \hat{r}_{\theta_N,t_N} \triangleq \int_{\{x\in\mathbb{R}^n : h(x)\geq 0\}} \hat{p}_{\theta_N}(t_N, x)\, dx$

   Denote $\hat{P}(\cdot) \triangleq (\hat{p}_{\theta_i,t_i}(\cdot))_{i=1}^{N}, \ \hat{S} \triangleq (\hat{s}_{\theta_i,t_i})_{i=1}^{N}, \ \hat{R} \triangleq (\hat{r}_{\theta_i,t_i})_{i=1}^{N}$

Assumptions and algorithm - part II

- At iteration N: 5 steps

Showed later why these are safe!
Note the necessity of resetting…

1. Collect Q iid samples $(X_{u_{\theta_N}}(w_i^N, t_N))_{i \in [\![1,Q]\!]}$

2. Build the kernel density estimator $\hat{p}_{\theta_N,t_N}(x) \triangleq \frac{1}{Q}\sum_{i=1}^{Q} \rho_R(x - X_{u_{\theta_N}}(w_i^N, t_N))$ (B. and Rudi FoCM 2025)

   and compute safety/reset: $\hat{s}_{\theta_N,t_N} \triangleq \int_{\{x\in\mathbb{R}^n : g(x)\geq 0\}} \hat{p}_{\theta_N}(t_N, x)\, dx, \quad \hat{r}_{\theta_N,t_N} \triangleq \int_{\{x\in\mathbb{R}^n : h(x)\geq 0\}} \hat{p}_{\theta_N}(t_N, x)\, dx$

   Denote $\hat{P}(\cdot) \triangleq (\hat{p}_{\theta_i,t_i}(\cdot))_{i=1}^N,\ \hat{S} \triangleq (\hat{s}_{\theta_i,t_i})_{i=1}^N,\ \hat{R} \triangleq (\hat{r}_{\theta_i,t_i})_{i=1}^N$

3. Estimate safety/reset functions as

   $\hat{s}_N(\theta,t) \triangleq \hat{S}(K + N\lambda I)^{-1}k(\theta,t)$ and $\hat{r}_N(\theta,t) \triangleq \hat{R}(K + N\lambda I)^{-1}k(\theta,t)$

Assumptions and algorithm - part II

- At iteration N: 5 steps

Showed later why these are safe!
Note the necessity of resetting…

1. Collect Q iid samples $(X_{u_{\theta_N}}(w_i^N, t_N))_{i \in [\![1,Q]\!]}$

2. Build the kernel density estimator $\hat{p}_{\theta_N,t_N}(x) \triangleq \frac{1}{Q}\sum_{i=1}^{Q} \rho_R(x - X_{u_{\theta_N}}(w_i^N, t_N))$ (B. and Rudi FoCM 2025)

   and compute safety/reset: $\hat{s}_{\theta_N,t_N} \triangleq \int_{\{x\in\mathbb{R}^n : g(x)\geq 0\}} \hat{p}_{\theta_N}(t_N, x)\, dx$, $\quad \hat{r}_{\theta_N,t_N} \triangleq \int_{\{x\in\mathbb{R}^n : h(x)\geq 0\}} \hat{p}_{\theta_N}(t_N, x)\, dx$

   Denote $\hat{P}(\cdot) \triangleq (\hat{p}_{\theta_i,t_i}(\cdot))_{i=1}^N$, $\hat{S} \triangleq (\hat{s}_{\theta_i,t_i})_{i=1}^N$, $\hat{R} \triangleq (\hat{r}_{\theta_i,t_i})_{i=1}^N$

3. Estimate safety/reset functions as

   $\hat{s}_N(\theta,t) \triangleq \hat{S}(K + N\lambda I)^{-1}k(\theta,t)$ and $\hat{r}_N(\theta,t) \triangleq \hat{R}(K + N\lambda I)^{-1}k(\theta,t)$

   with $k(\theta,t) \triangleq (k((\theta,t),(\theta_i,t_i)))_{i=1}^N$, $K \triangleq (k((\theta_i,t_i),(\theta_j,t_j)))_{i,j=1}^N$

Assumptions and algorithm - part II

- At iteration N: 5 steps

4. Define the LCBs $\mathrm{LCB}^{\mathrm{s}}_{\mathrm{N}}(\theta, T) \triangleq \inf_{t \in [0,T]} \big(\hat{s}_N(\theta, t) - \beta^s_N \sigma_N(\theta, t)\big)$

$$\mathrm{LCB}^{\mathrm{r}}_{\mathrm{N}}(\theta, T) \triangleq \hat{r}_N(\theta, T) - \beta^r_N \sigma_N(\theta, T)$$

Assumptions and algorithm - part II

- At iteration N: 5 steps

4. Define the LCBs $\mathrm{LCB}^{\mathrm{s}}_{\mathrm{N}}(\theta, T) \triangleq \inf_{t \in [0,T]} \left(\hat{s}_N(\theta, t) - \beta^s_N \sigma_N(\theta, t) \right)$

   $\mathrm{LCB}^{\mathrm{r}}_{\mathrm{N}}(\theta, T) \triangleq \hat{r}_N(\theta, T) - \beta^r_N \sigma_N(\theta, T)$ with predictive uncertainty:

$$\sigma^2_N(\theta, t) \triangleq k((\theta, t), (\theta, t)) - k(\theta, t)^*(K + N\lambda I)^{-1} k(\theta, t).$$

Assumptions and algorithm - part II

- At iteration N: 5 steps

Unlike the SoA, our estimator $\hat{p}_{\theta_N, t_N}$ enables proving these are bounded through iterations!

4. Define the LCBs $\mathrm{LCB}^{\mathrm{s}}_{\mathrm{N}}(\theta, T) \triangleq \inf_{t \in [0,T]} \left(\hat{s}_N(\theta, t) - \beta^s_N \sigma_N(\theta, t)\right)$

   $\mathrm{LCB}^{\mathrm{r}}_{\mathrm{N}}(\theta, T) \triangleq \hat{r}_N(\theta, T) - \beta^r_N \sigma_N(\theta, T)$ with predictive uncertainty:

$$\sigma^2_N(\theta, t) \triangleq k((\theta, t), (\theta, t)) - k(\theta, t)^*(K + N\lambda I)^{-1} k(\theta, t).$$

Assumptions and algorithm - part II

- At iteration N: 5 steps

Unlike the SoA, our estimator $\hat{p}_{\theta_N, t_N}$ enables proving these are bounded through iterations!

Unlike Bayesian, βs proportional to density estimation error, not to noise variance!

4. Define the LCBs $\mathrm{LCB}^{\mathrm{s}}_{\mathrm{N}}(\theta, T) \triangleq \inf_{t \in [0,T]} \left(\hat{s}_N(\theta, t) - \beta^s_N \sigma_N(\theta, t)\right)$

$\mathrm{LCB}^{\mathrm{r}}_{\mathrm{N}}(\theta, T) \triangleq \hat{r}_N(\theta, T) - \beta^r_N \sigma_N(\theta, T)$ with predictive uncertainty:

$$\sigma^2_N(\theta, t) \triangleq k((\theta, t), (\theta, t)) - k(\theta, t)^*(K + N\lambda I)^{-1} k(\theta, t).$$

Assumptions and algorithm - part II

- At iteration N: 5 steps

Unlike the SoA, our estimator $\hat{p}_{\theta_N,t_N}$ enables proving these are bounded through iterations!

4. Define the LCBs $\mathrm{LCB}^{\mathrm{s}}_{\mathrm{N}}(\theta,T) \triangleq \inf_{t\in[0,T]} \left(\hat{s}_N(\theta,t) - \beta^s_N \sigma_N(\theta,t)\right)$

$\mathrm{LCB}^{\mathrm{r}}_{\mathrm{N}}(\theta,T) \triangleq \hat{r}_N(\theta,T) - \beta^r_N \sigma_N(\theta,T)$ with predictive uncertainty:

Unlike Bayesian, βs proportional to density estimation error, not to noise variance!

$$\sigma^2_N(\theta,t) \triangleq k((\theta,t),(\theta,t)) - k(\theta,t)^*(K+N\lambda I)^{-1}k(\theta,t).$$

5. Final step: define the safe-resettable feasible set

$$\Gamma_N = \Gamma_0 \cup \left\{(\theta,t,T)\in D\times[0,T_{\max}]^2 \mid t\le T,\ \mathrm{LCB}^{\mathrm{s}}_{\mathrm{N}}(\theta,T)\ge 1-\varepsilon,\ \mathrm{LCB}^{\mathrm{r}}_{\mathrm{N}}(\theta,T)\ge 1-\xi\right\}$$

Assumptions and algorithm - part II

- At iteration N: 5 steps

Unlike the SoA, our estimator $\hat{p}_{\theta_N, t_N}$ enables proving these are bounded through iterations!

Unlike Bayesian, βs proportional to density estimation error, not to noise variance!

4. Define the LCBs $\mathrm{LCB}^{\mathrm{s}}_{\mathrm{N}}(\theta, T) \triangleq \inf_{t \in [0,T]} \left(\hat{s}_N(\theta, t) - \beta^s_N \sigma_N(\theta, t) \right)$

$\mathrm{LCB}^{\mathrm{r}}_{\mathrm{N}}(\theta, T) \triangleq \hat{r}_N(\theta, T) - \beta^r_N \sigma_N(\theta, T)$ with predictive uncertainty:

$$\sigma^2_N(\theta, t) \triangleq k((\theta, t), (\theta, t)) - k(\theta, t)^* (K + N\lambda I)^{-1} k(\theta, t).$$

5. Final step: define the safe-resettable feasible set

$$\Gamma_N = \Gamma_0 \cup \left\{ (\theta, t, T) \in D \times [0, T_{\max}]^2 \mid t \leq T,\ \mathrm{LCB}^{\mathrm{s}}_{\mathrm{N}}(\theta, T) \geq 1 - \varepsilon,\ \mathrm{LCB}^{\mathrm{r}}_{\mathrm{N}}(\theta, T) \geq 1 - \xi \right\}$$

and solve $(\theta_{N+1}, t_{N+1}, T_{N+1}) = \underset{(\theta, t, T) \in \Gamma_N}{\arg\max}\ \sigma_N(\theta, t)$ until $\sigma_N(\theta_{N+1}, t_{N+1}) < \eta$

Theoretical guarantees - part I

Theorem 5.1 (Safely learning controlled Sobolev dynamics). *Let* $\eta > 0$, *and assume Assumptions (A1)–(A3) hold. Set* $R = Q^{1/(n+2\nu)}$ *and* $\lambda = N^{-1}$.

Theoretical guarantees - part I

Theorem 5.1 (Safely learning controlled Sobolev dynamics). *Let* $\eta > 0$*, and assume Assumptions (A1)–(A3) hold. Set* $R = Q^{1/(n+2\nu)}$ *and* $\lambda = N^{-1}$*. Then there exist constants* $c_1, \dots, c_5 > 0$*, independent of* N, Q, δ, η*, such that if*

$$c_1 \log(4N/\delta)^{1/2} Q^{\frac{n-2\nu}{2n+4\nu}} \leq N^{-1/2},$$

Theoretical guarantees - part I

Theorem 5.1 (Safely learning controlled Sobolev dynamics). *Let $\eta > 0$, and assume Assumptions (A1)–(A3) hold. Set $R = Q^{1/(n+2\nu)}$ and $\lambda = N^{-1}$. Then there exist constants $c_1, \ldots, c_5 > 0$, independent of N, Q, δ, η, such that if*

With probability ≥ 1-δ

$$c_1 \log(4N/\delta)^{1/2} Q^{\frac{n-2\nu}{2n+4\nu}} \leq N^{-1/2},$$

then the stopping condition $\max_{(\theta,t,T)\in\Gamma_N} \sigma_N(\theta, t) < \eta$ is satisfied after at most $N \leq c_2 \eta^{-2/(1-\alpha)}$ iterations for any $\alpha > (m+1)/(m+1+2\nu)$.

Theoretical guarantees - part I

Theorem 5.1 (Safely learning controlled Sobolev dynamics). *Let $\eta > 0$, and assume Assumptions (A1)–(A3) hold. Set $R = Q^{1/(n+2\nu)}$ and $\lambda = N^{-1}$. Then there exist constants $c_1, \ldots, c_5 > 0$, independent of N, Q, δ, η, such that if*

$$c_1 \log(4N/\delta)^{1/2} Q^{\frac{n-2\nu}{2n+4\nu}} \leq N^{-1/2},$$

With probability ≥ 1-δ

then the stopping condition $\max_{(\theta,t,T)\in\Gamma_N} \sigma_N(\theta,t) < \eta$ is satisfied after at most $N \leq c_2\eta^{-2/(1-\alpha)}$ iterations for any $\alpha > (m+1)/(m+1+2\nu)$. Moreover:

- ***(Safety):*** *All selected triples (θ_i, t_i, T_i) satisfy $s^\infty(\theta_i, T_i) \geq 1-\varepsilon$ and $r(\theta_i, T_i) \geq 1-\xi$, providing safety guarantees during training.*

Theoretical guarantees - part I

Theorem 5.1 (Safely learning controlled Sobolev dynamics). *Let $\eta > 0$, and assume Assumptions (A1)–(A3) hold. Set $R = Q^{1/(n+2\nu)}$ and $\lambda = N^{-1}$. Then there exist constants $c_1, \ldots, c_5 > 0$, independent of N, Q, δ, η, such that if*

$$c_1 \log(4N/\delta)^{1/2} Q^{\frac{n-2\nu}{2n+4\nu}} \leq N^{-1/2},$$

With probability ≥ 1-δ

then the stopping condition $\max_{(\theta,t,T)\in\Gamma_N} \sigma_N(\theta,t) < \eta$ is satisfied after at most $N \leq c_2 \eta^{-2/(1-\alpha)}$ iterations for any $\alpha > (m+1)/(m+1+2\nu)$. Moreover:

- ***(Safety):** All selected triples (θ_i, t_i, T_i) satisfy $s^\infty(\theta_i, T_i) \geq 1 - \varepsilon$ and $r(\theta_i, T_i) \geq 1 - \xi$, providing safety guarantees during training. Moreover, the final set Γ_N includes only controls meeting these thresholds and can thus serve as a certified safe set for deployment.*

Theoretical guarantees - part I

Theorem 5.1 (Safely learning controlled Sobolev dynamics). *Let $\eta > 0$, and assume Assumptions (A1)–(A3) hold. Set $R = Q^{1/(n+2\nu)}$ and $\lambda = N^{-1}$. Then there exist constants $c_1, \ldots, c_5 > 0$, independent of N, Q, δ, η, such that if*

$$c_1 \log(4N/\delta)^{1/2} Q^{\frac{n-2\nu}{2n+4\nu}} \leq N^{-1/2},$$

With probability ≥ 1-δ

then the stopping condition $\max_{(\theta,t,T)\in\Gamma_N} \sigma_N(\theta,t) < \eta$ is satisfied after at most $N \leq c_2 \eta^{-2/(1-\alpha)}$ iterations for any $\alpha > (m+1)/(m+1+2\nu)$. Moreover:

- ***(Safety):*** *All selected triples (θ_i, t_i, T_i) satisfy $s^\infty(\theta_i, T_i) \geq 1 - \varepsilon$ and $r(\theta_i, T_i) \geq 1 - \xi$, providing safety guarantees during training. Moreover, the final set Γ_N includes only controls meeting these thresholds and can thus serve as a certified safe set for deployment.*
- ***(Estimation guarantees):*** *For all $(\theta, t, T) \in \Gamma_N$,*

$$\|\hat{p}_\theta(t,\cdot) - p_\theta(t,\cdot)\|_\infty \leq c_3\eta, \quad |\hat{s}_N(\theta,t) - s(\theta,t)| \leq c_4\eta, \quad |\hat{r}_N(\theta,t) - r(\theta,t)| \leq c_5\eta.$$

Theoretical guarantees - part I

Theorem 5.1 (Safely learning controlled Sobolev dynamics). *Let $\eta > 0$, and assume Assumptions (A1)–(A3) hold. Set $R = Q^{1/(n+2\nu)}$ and $\lambda = N^{-1}$. Then there exist constants $c_1, \ldots, c_5 > 0$, independent of N, Q, δ, η, such that if*

With probability ≥ 1-δ

$$c_1 \log(4N/\delta)^{1/2} Q^{\frac{n-2\nu}{2n+4\nu}} \leq N^{-1/2},$$

then the stopping condition $\max_{(\theta,t,T)\in\Gamma_N} \sigma_N(\theta,t) < \eta$ is satisfied after at most $N \leq c_2\eta^{-2/(1-\alpha)}$ iterations for any $\alpha > (m+1)/(m+1+2\nu)$. Moreover:

- ***(Safety):*** *All selected triples (θ_i, t_i, T_i) satisfy $s^\infty(\theta_i, T_i) \geq 1-\varepsilon$ and $r(\theta_i, T_i) \geq 1-\xi$, providing safety guarantees during training. Moreover, the final set Γ_N includes only controls meeting these thresholds and can thus serve as a certified safe set for deployment.*
- ***(Estimation guarantees):*** *For all $(\theta, t, T) \in \Gamma_N$,*

$$\|\hat{p}_\theta(t,\cdot) - p_\theta(t,\cdot)\|_\infty \leq c_3\eta, \quad |\hat{s}_N(\theta,t) - s(\theta,t)| \leq c_4\eta, \quad |\hat{r}_N(\theta,t) - r(\theta,t)| \leq c_5\eta.$$

Thus, for every $t \in [0,T]$

Theoretical guarantees - part II

The previous result does not tell us whether Γ_N "grows" towards the true Γ and at which rate…

Theoretical guarantees - part II

The previous result does not tell us whether Γ_N “grows” towards the true Γ and at which rate…

New result, soon submitted

Theorem 5.5 (Exploration guarantees). *Define $\Gamma^\eta(\Gamma_0)$, the η-reachable safe-resettable region from Γ_0, as the union of connected components of*

$$\Gamma^\eta \triangleq \left\{(\theta, T) \in D \times [0, T_{\max}] \;\middle|\; s^\infty(\theta, T) > 1 - \varepsilon + \eta \;\; \textit{and} \;\; r(\theta, T) > 1 - \xi + \eta\right\}$$

intersecting $\pi_{1,3}(\Gamma_0)$.

Theoretical guarantees - part II

The previous result does not tell us whether Γ_N "grows" towards the true Γ and at which rate…

New result, soon submitted

Theorem 5.5 (Exploration guarantees). *Define $\Gamma^\eta(\Gamma_0)$, the η-reachable safe-resettable region from Γ_0, as the union of connected components of*

$$\Gamma^\eta \triangleq \left\{(\theta, T) \in D \times [0, T_{\max}] \,\middle|\, s^\infty(\theta, T) > 1 - \varepsilon + \eta \;\; \textit{and} \;\; r(\theta, T) > 1 - \xi + \eta\right\}$$

intersecting $\pi_{1,3}(\Gamma_0)$. Under the setting of Theorem 5.1, with probability at least $1-\delta$ it holds that

$$\Gamma^\eta(\Gamma_0) \subset \pi_{1,3}(\Gamma_N) \subset \Gamma.$$

Theoretical guarantees - part II

The previous result does not tell us whether Γ_N "grows" towards the true Γ and at which rate…

New result, soon submitted

Theorem 5.5 (Exploration guarantees). *Define $\Gamma^\eta(\Gamma_0)$, the η-reachable safe-resettable region from Γ_0, as the union of connected components of*

$$\Gamma^\eta \triangleq \left\{(\theta, T) \in D \times [0, T_{\max}] \,\middle|\, s^\infty(\theta, T) > 1 - \varepsilon + \eta \;\; \text{and} \;\; r(\theta, T) > 1 - \xi + \eta\right\}$$

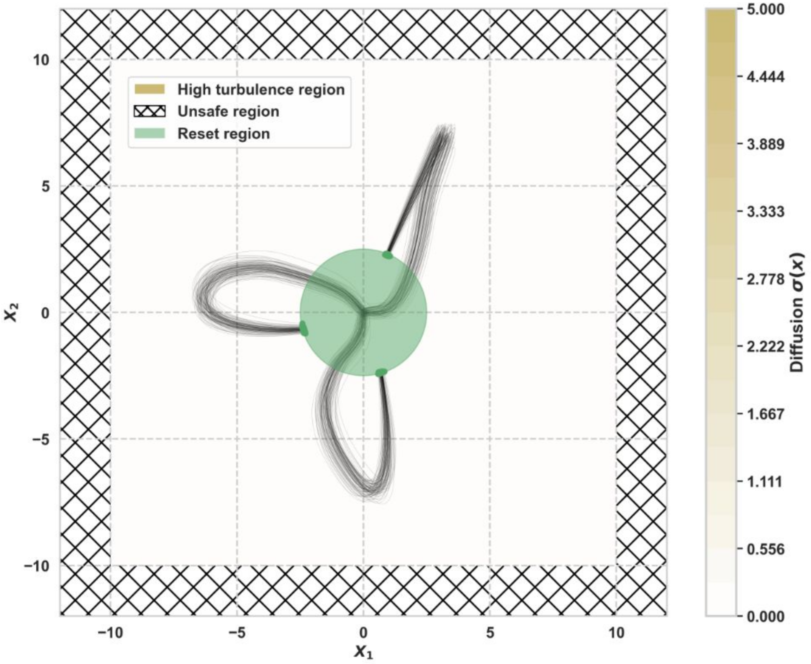
intersecting $\pi_{1,3}(\Gamma_0)$. Under the setting of Theorem 5.1, with probability at least $1 - \delta$ it holds that

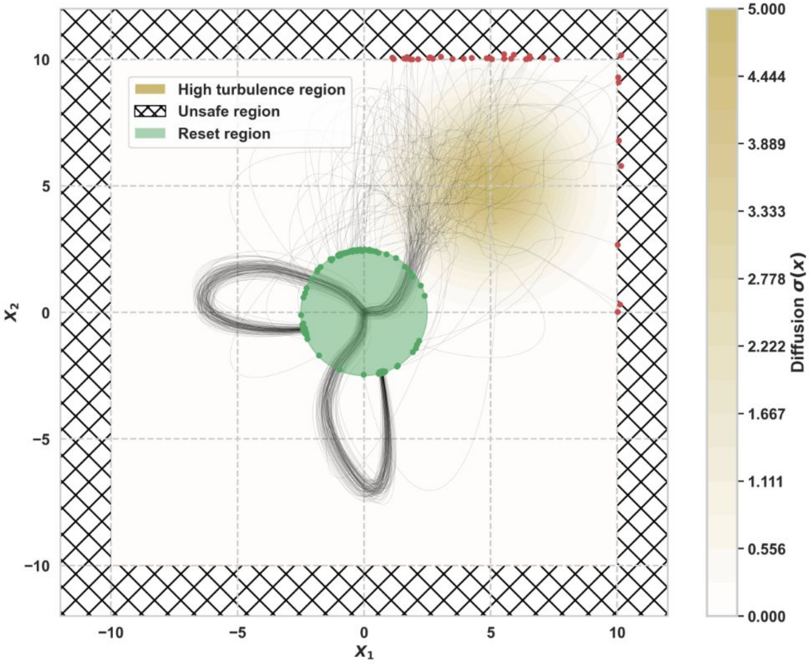
$$\Gamma^\eta(\Gamma_0) \subset \pi_{1,3}(\Gamma_N) \subset \Gamma.$$

In short, the known set of "practically usable" safe controls grows towards the true set of safe controls at a given rate!

Numerical example - Acceleration-controlled robot

$$\begin{cases} dX(t) = V(t)dt, \\ dV(t) = u(t, X(t), V(t))dt + a(X(t))dW_t \end{cases} \qquad a(X) = A\exp\left(-\frac{\|X - X_c\|^2}{2\sigma^2}\right)$$





Numerical example - Acceleration-controlled robot

- Control parametrization

1. Exploration $(0 \leq t \leq T_{\text{explo}})$ - $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$, … applied on time intervals of fixed length
$$u(t, X, V) = v(\cos(\theta_i), \sin(\theta_i)) - V$$

2. Reset $(t > T_{\text{explo}})$
$$u(t, X, V) = \kappa \times \left(v \frac{\mu_0(X) - X}{\|\mu_0(X) - X\|} - V \right)$$

Numerical example - Acceleration-controlled robot

- Control parametrization

1. Exploration $(0 \leq t \leq T_{\mathrm{explo}})$ - $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$, … applied on time intervals of fixed length

$$u(t, X, V) = v(\cos(\theta_i), \sin(\theta_i)) - V$$

2. Reset $(t > T_{\mathrm{explo}})$

$$u(t, X, V) = \kappa \times \left(v \frac{\mu_0(X) - X}{\|\mu_0} - V \right)$$

Always steers to $\operatorname{supp} p_0 = \overline{B_\rho(0)}$

Numerical example - Acceleration-controlled robot

- Control parametrization

1. Exploration $(0 \le t \le T_{\text{explo}})$ - $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$, … applied on time intervals of fixed length

$$u(t, X, V) = v(\cos(\theta_i), \sin(\theta_i)) - V$$

2. Reset $(t > T_{\text{explo}})$

$$u(t, X, V) = \kappa \times \left(v \frac{\mu_0(X) - X}{\|\mu_0 \cdots} - V \right)$$

Always steers to $\operatorname{supp} p_0 = \overline{B_\rho(0)}$

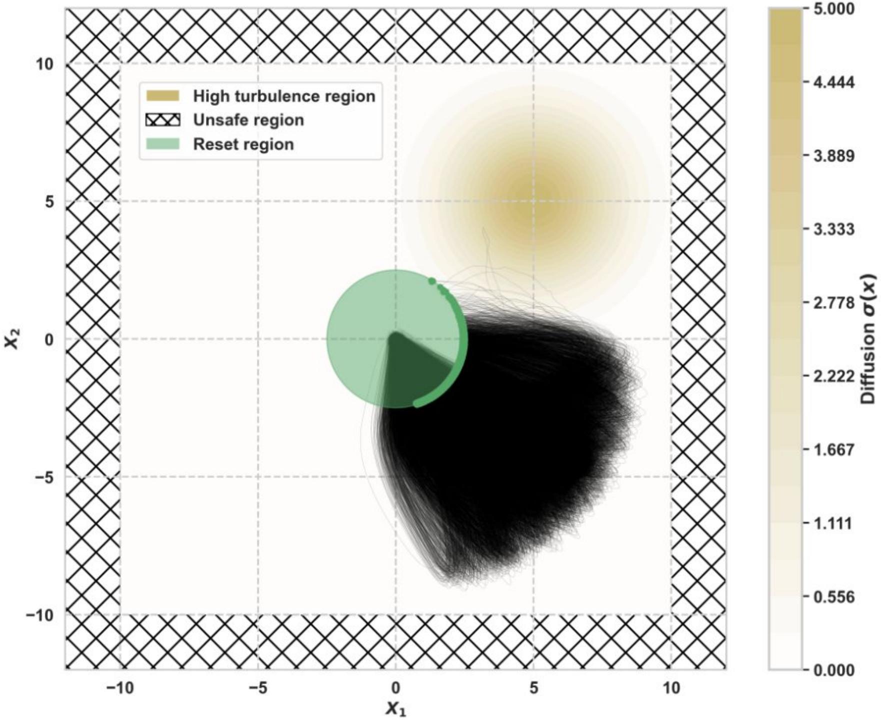
3. Parameters

$$v = 2.0, \kappa = 0.5, m = 2, T_{\text{explo}} = 6$$

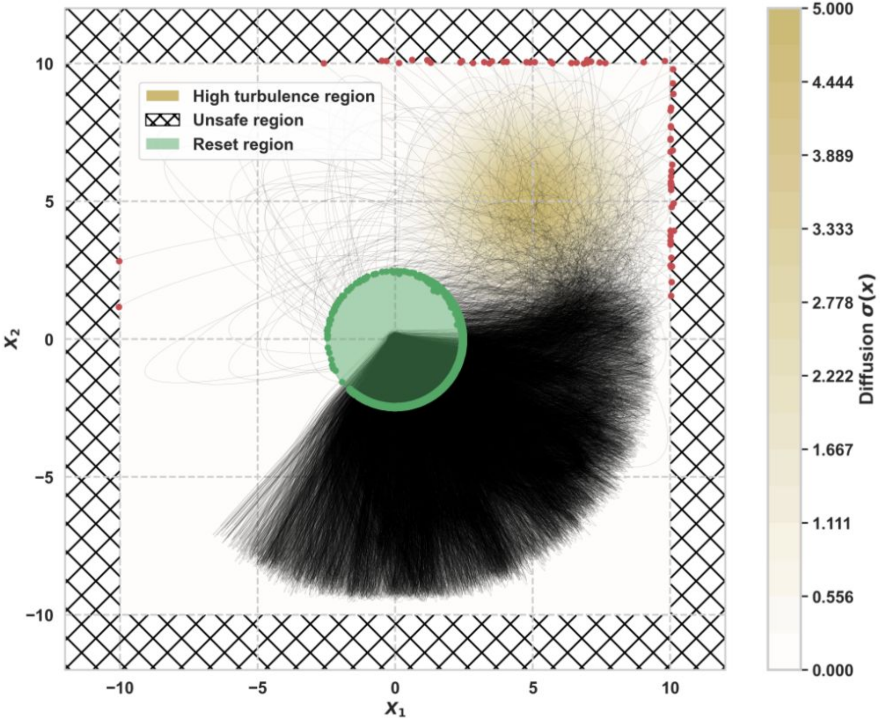
confidence levels (β_s, β_r), kernel smoothness (λ, γ), and bandwidth R, tuned

Numerical example - Acceleration-controlled robot

- Four safety/reset scenarios $\varepsilon = \xi \in \{0.1, 0.3, 0.5, +\infty\}$, 1000 iterations



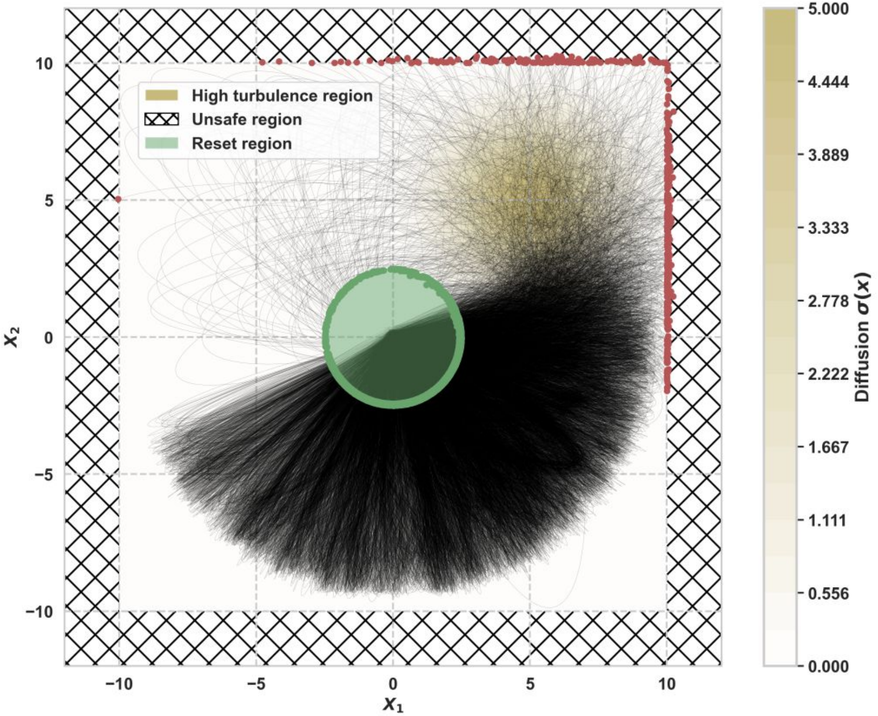
$\varepsilon = \xi = 0.1$



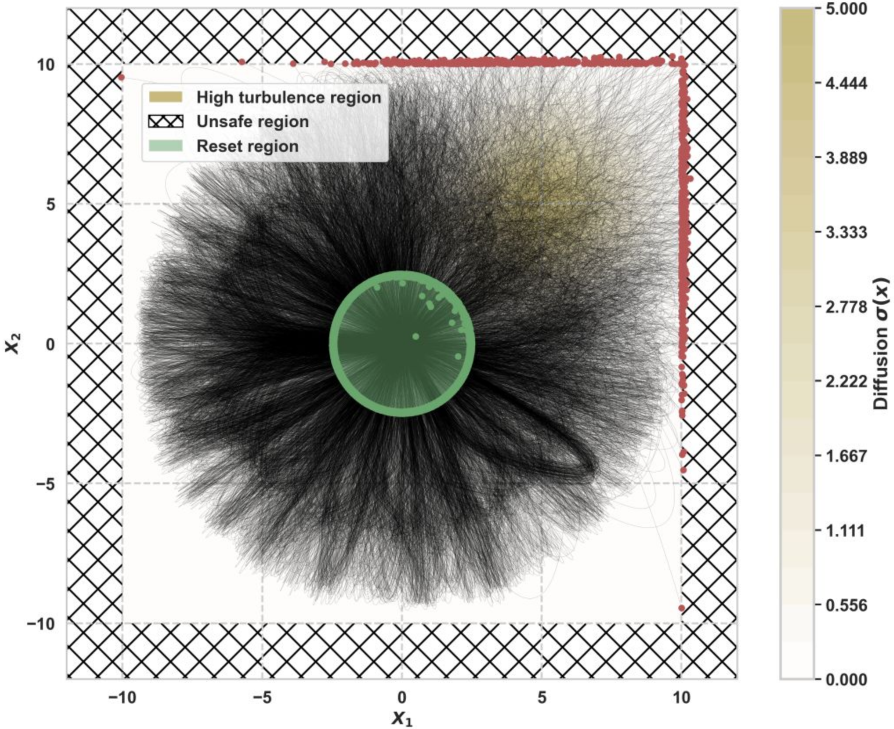
$\varepsilon = \xi = 0.3$

Numerical example - Acceleration-controlled robot

- Four safety/reset scenarios $\varepsilon = \xi \in \{0.1, 0.3, 0.5, +\infty\}$, 1000 iterations



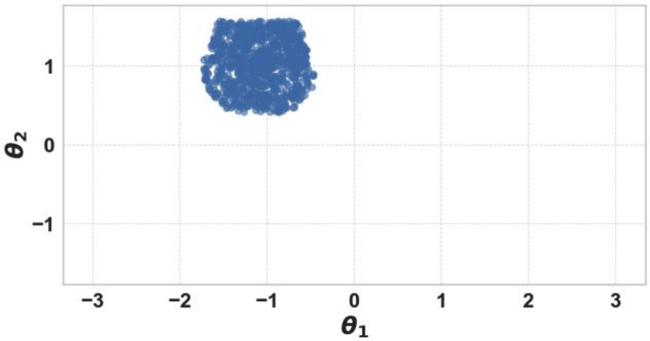
$\varepsilon = \xi = 0.5$



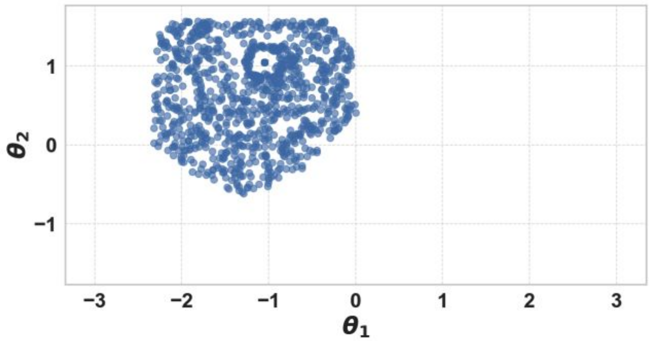
$\varepsilon = \xi = +\infty$

Numerical example - Acceleration-controlled robot

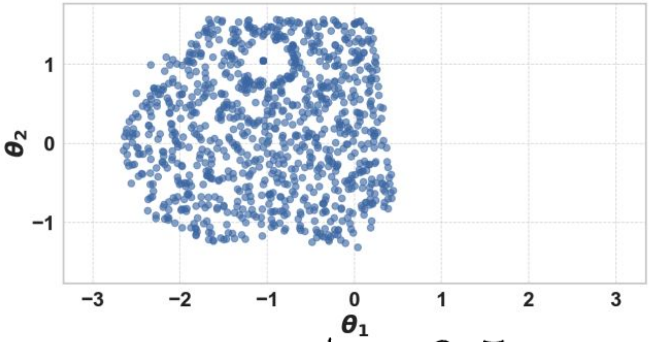
- Four safety/reset scenarios $\varepsilon = \xi \in \{0.1, 0.3, 0.5, +\infty\}$, 1000 iterations



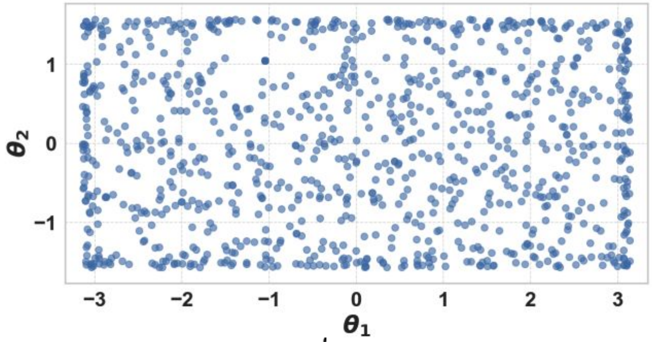
$$\varepsilon = \xi = 0.1$$



$$\varepsilon = \xi = 0.3$$



$$\varepsilon = \xi = 0.5$$



$$\varepsilon = \xi = +\infty$$

Contents

Conclusion: future directions

Extensions

1. Validation on physical systems (e.g., quadrotors)
2. Improve scalability via fast kernel methods (e.g., sketching, incremental updates)
3. Handle abrupt dynamics and other non-diffusive disturbances, such as jump processes

Main references:

1. R. Bonalli and A. Rudi, *Non-Parametric Learning of Stochastic Differential Equations with Non-asymptotic Fast Rates of Convergence*. Foundations of Computational Mathematics (2025), pp. 1-56.
2. L. Brogat-Motte, R. Bonalli, and A. Rudi, *Safely Learning Controlled Stochastic Dynamics*. Proc. Conference on Neural Information Processing Systems, 2025, San Diego.

Conclusion: future directions

Thank you for your attention!

Questions are more than welcome :)