Formally Guaranteed Learning- and Optimization-based Methods for Risk-Averse Control of Autonomous Systems

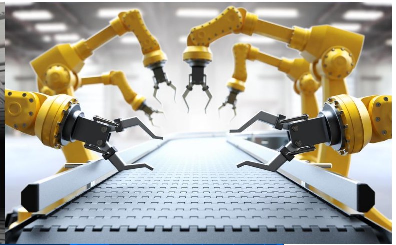
Workshop on Data-Driven Verification and Control with Provable Guarantees - ECC 2024

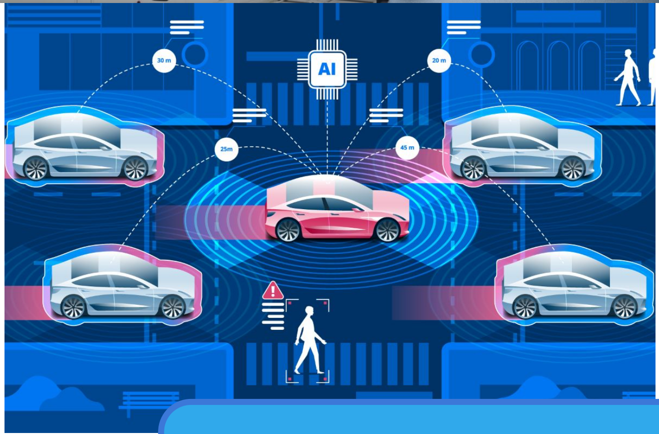
Riccardo Bonalli

Laboratoire des Signaux et Systèmes
CNRS and Université Paris-Saclay

Contents

Challenges in controlling autonomous systems





Though unwanted events might happen…

Let us use risk-averse optimal control!

Mathematically modeling risk-averse control

$$\min_{u \in \mathcal{U}} \rho\left[\int_0^{t_f} f^0(s, u(s), x(s))\, \mathrm{d}s \right]$$

$$\mathrm{d}x(s) = f(s, u(s), x(s))\, \mathrm{d}s + \sigma(s, u(s), x(s))\, \mathrm{d}W_s$$

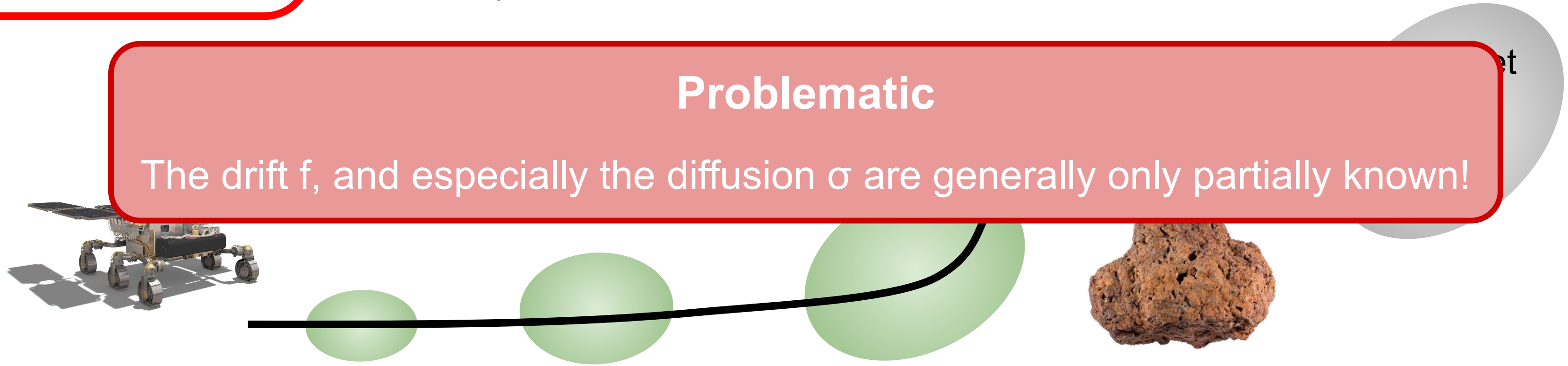
$$x(0) = x^0, \quad \rho[\, g(x(t_f))\,] = 0$$

$$\rho[\, h(x(s))\,] \le 0, \quad s \in [0, t_f]$$

ρ is a generic **coherent risk measure**, which are good statistical functionals to accurately model risk, e.g., quantiles, CVaR (or AVaR), etc.

Problematic

The drift f, and especially the diffusion σ are generally only partially known!

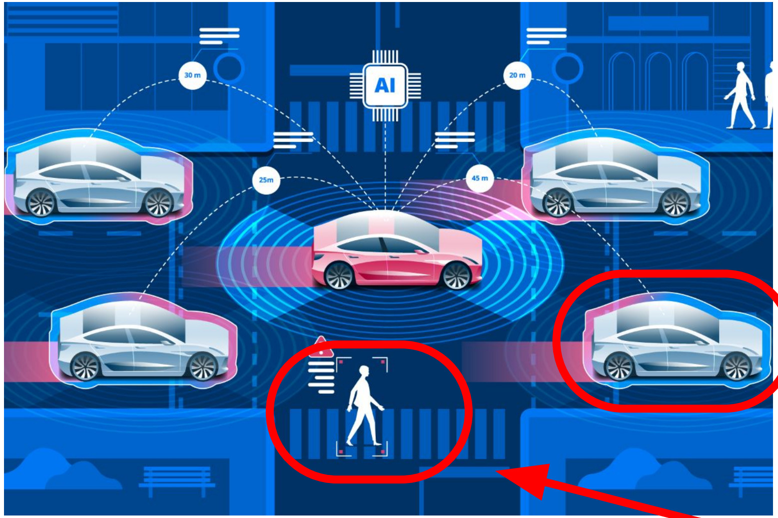


Contents

First topic

Learning Stochastic Differential Equations (SDE) with theoretical guarantees

Collaboration with A. Rudi, INRIA Paris



$$\begin{cases} \mathrm{d}x(t) = f(t, x(t))\mathrm{d}t \\ \qquad\qquad + \sigma(t, x(t))\mathrm{d}W_t \\ x(0) \sim \mu_0 \end{cases}$$

Several approaches have already been proposed…

Here are few examples:

- H. Garnier and L. Wang. Identification of Continuous-Time Models from Sampled Data. Springer-Verlag, 2008. Note: SDE which are linear w.r.t. state variables.
- Li, X. and others, Scalable Gradients for Stochastic Differential Equations, Int. Conference on Artificial Intelligence and Statistics, 2020. Note: no guarantees exist.
- S. Nakakita, Parametric Estimation of SDE via Online Gradient Descent, Arxiv preprint, n. 2210.08800, 2022. Note: the model family is assumed to be "correctly specifying".

Main criticism

Apparently, there is no work addressing general (e.g., non-linear) SDE and offering guarantees of accuracy under mild assumptions

Learning SDE: how this?

The simplest idea: fitting all the involved processes on available data

$$\min_{f,\sigma} \sum_{i=1}^{N_{\text{data}}} \left\| x^i(\cdot) - x^i(0) - \int_0^{\cdot} f(t, x^i(t))\,\mathrm{d}t - \int_0^{\cdot} \sigma(t, x^i(t))\,\mathrm{d}W_t^i \right\|^2$$

Crucial hindrances

- Integrating non-linear SDE is expensive and approximations must be considered
- No knowledge about the noise causing uncertainty is available

Idea: Fit the Fokker-Planck equation related to SDE instead

Learning SDE: how this?

How do solutions to SDE relate to the Fokker-Planck equation?

A measurable mapping $X : \mathbb{R}^n \times \Omega \to C([0,T], \mathbb{R}^n)$ is a solution of SDE if, for μ_0-almost every $x \in \mathbb{R}^n$, the process $X_x(t,\omega) \triangleq X(x,\omega)(t)$ is adapted (to the filtration generated by the Brownian motion) and satisfies:

$$\mathrm{d}X_x(t) = f(t, X_x(t))\mathrm{d}t + \sigma(t, X_x(t))\mathrm{d}W_t, \quad X_x(0) = x$$

Therefore, the curve of densities $p(t,x)$ given by:

$$\int_A p(t,x)\,\mathrm{d}x = \int_{\mathbb{R}^n} \mathbb{P}(X_x(t) \in A)\,\mu_0(\mathrm{d}x), \quad A \in \mathcal{B}(\mathbb{R}^n)$$

solves the Fokker-Planck equation:

$$\frac{\partial p}{\partial t}(t,\cdot) = (\mathcal{L}_t^{f,g})^* p(t,\cdot)$$

Main benefits

- Linear parabolic PDE
- No presence of noise

where $g \triangleq \sigma\sigma^\top$ and $(\mathcal{L}_t^{f,g})^* p(t,x) = \frac{1}{2} D_x^2(g(t,\cdot)p(t,\cdot))(x) - D(f(t,\cdot)p(t,\cdot))(x)$.

Learning SDE: how this?

The simplest idea: fitting all the involved processes on available data

$$\min_{f,\sigma} \sum_{i=1}^{N_{\text{data}}} \left\| x^i(\cdot) - x^i(0) - \int_0^{\cdot} f(t, x^i(t))\, \mathrm{d}t - \int_0^{\cdot} \sigma(t, x^i(t))\, \mathrm{d}W_t^i \right\|$$

Replaced by:

$$\min_{f,g=\sigma\sigma^\top} \int_0^T \left\| \frac{\partial p^{\text{app}}}{\partial t}(t,\cdot) - (\mathcal{L}_t^{f,g})^* p^{\text{app}}(t,\cdot) \right\|_{L^2}^2 \mathrm{d}t + \lambda \|(f,g)\|_{\mathcal{S}_{\text{RKHS}}}^2$$

p^{app} denotes some regular enough model which well approximates the density of the unknown solution to SDE through available samples: TO COMPUTE!

Learning SDE: how this?

Fitting the approximated curve of densities via finite RKHS-based models

$$p^{\mathrm{app}}(t,x) = \sum_{\ell=1}^{M} c_\ell(t)\mathcal{F}_\ell(x), \quad \mathcal{F}_\ell(x) = \frac{1}{N}\sum_{j=1}^{N}\mathcal{R}_R(x - X(t_\ell,\omega_j)), \quad N_{\mathrm{data}} = M \cdot N$$

Here, $\mathcal{R}_R$ are appropriate regular radial functions, whereas c_ℓ stem from “pointwise” Kernel representation.

Main benefits:

- Almost instantaneous computations (computing c_ℓ only needs an offline matrix inversion)
- Tight error bounds that improve with the regularity r of the unknown drift f and diffusion σ:

$$\left\|p^{\mathrm{app}} - p^{\mathrm{unknown}}\right\|_{H^1(0,T;H^2)} = O\left(\log\left(N_{\mathrm{data}}^{\frac{1}{n+2r}}\right) N_{\mathrm{data}}^{-\frac{r-1}{n+2r}}\right)$$ with “high” probability

Theoretical guarantees

Why all of this is convenient?

Theorem (B, Rudi)

Combining error bounds from RKHS and parabolic PDE theory yields (for an appropriate λ):

$$\left\| p^{f_{\text{model}},\sigma_{\text{model}}} - p^{\text{unknown}} \right\|_{L^2(0,T;L^2)} = O\left(\log\left(N_{\text{data}}^{\frac{1}{n+2r}} \right) N_{\text{data}}^{-\frac{r-1}{n+2r}} \right)$$ with “high” probability

Theoretical guarantees

Why all of this is convenient?

Theorem (B, Rudi)

Combining error bounds from RKHS and parabolic PDE theory yields (for an appropriate λ):

$$\left\| p^{f_{\text{model}},\sigma_{\text{model}}} - p^{\text{unknown}} \right\|_{L^2(0,T;L^2)} = O\left(\log\left(N_{\text{data}}^{\frac{1}{n+2r}} \right) N_{\text{data}}^{-\frac{r-1}{n+2r}} \right)$$ with "high" probability

In particular,

$$\mathbb{E}_{\mu_0 \times \mathbb{P}}\left[\int_0^T c(t, X_x^{\text{unknown}}(t))\, \mathrm{d}t \right] =$$

$$= \mathbb{E}_{\mu_0 \times \mathbb{P}}\left[\int_0^T c(t, X_x^{f_{\text{model}},\sigma_{\text{model}}}(t))\, \mathrm{d}t \right] + O\left(\log\left(N_{\text{data}}^{\frac{1}{n+2r}} \right) N_{\text{data}}^{-\frac{r-1}{n+2r}} \right)$$

Key fact:

With a little bit of extra effort, we may replace expectation with risk measures

How does it work in practice?

1 - Compute the coefficients of the approximated density:

$$p^{\text{app}}(t,x) = \sum_{\ell=1}^{M} c_\ell(t)\mathcal{F}_\ell(x), \quad \mathcal{F}_\ell(x) = \frac{1}{N}\sum_{j=1}^{N}\mathcal{R}_R(x - X(t_\ell,\omega_j)), \quad N_{\text{data}} = M \cdot N$$

2 - Solve a guaranteed, kernel-based finite-dimensional approximation of the problem:

$$\min_{f,g=\sigma\sigma^\top} \int_0^T \left\| \frac{\partial p^{\text{app}}}{\partial t}(t,\cdot) - \left(\mathcal{L}_t^{f,g}\right)^* p^{\text{app}}(t,\cdot) \right\|_{L^2}^2 \mathrm{d}t + \lambda \|(f,g)\|_{\mathcal{S}_{\text{RKHS}}}^2$$

Classical kernel ridge regression:

$$\text{Kernel coefficients for } f \text{ and } \sigma = (A + \lambda I)^{-1} b$$

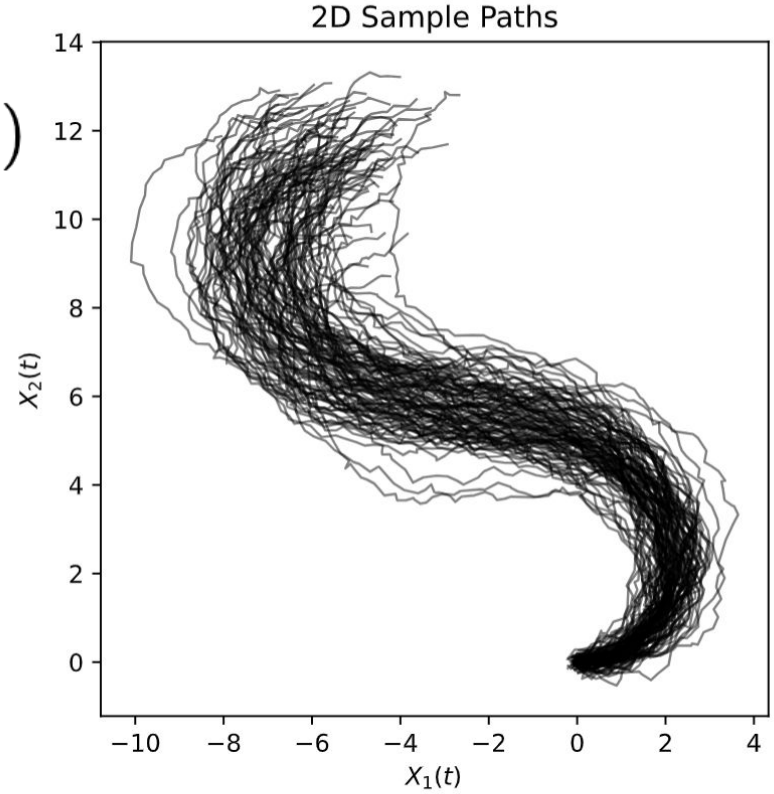
Numerical example (by Luc Brogat-Motte)

An already "challenging" example: a perturbed wheeled robot

$$dX(t) = v \begin{pmatrix} \cos u(t) \\ \sin u(t) \end{pmatrix} dt + \begin{pmatrix} \sigma \\ \sigma \end{pmatrix} dW(t)$$

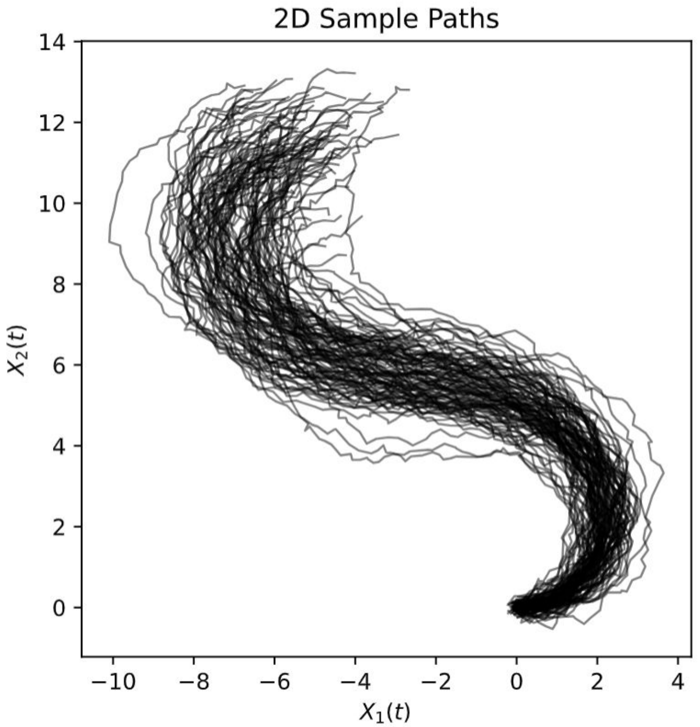
with the control signal:

$$u(t) = \theta \sin\left(\frac{\pi t}{10}\right), \quad \theta \in \mathbb{R}$$



Numerical example (by Luc Brogat-Motte)

An already “challenging” example: a perturbed wheeled robot



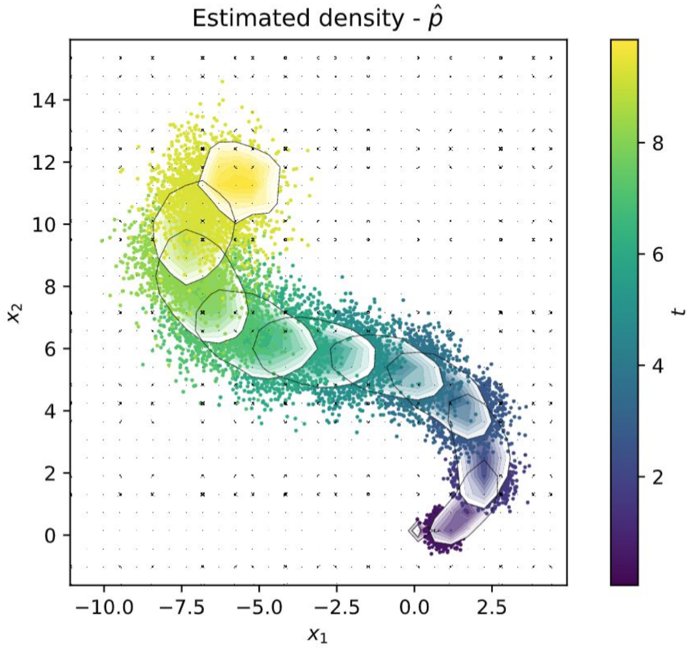
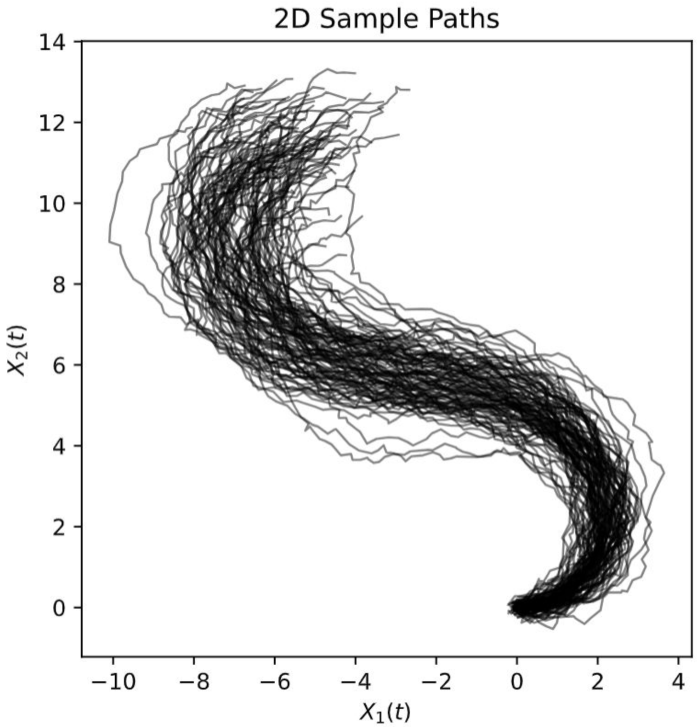


Figure 24: Estimated probability density along with the training points.

Numerical example (by Luc Brogat-Motte)

An already “challenging” example: a perturbed wheeled robot



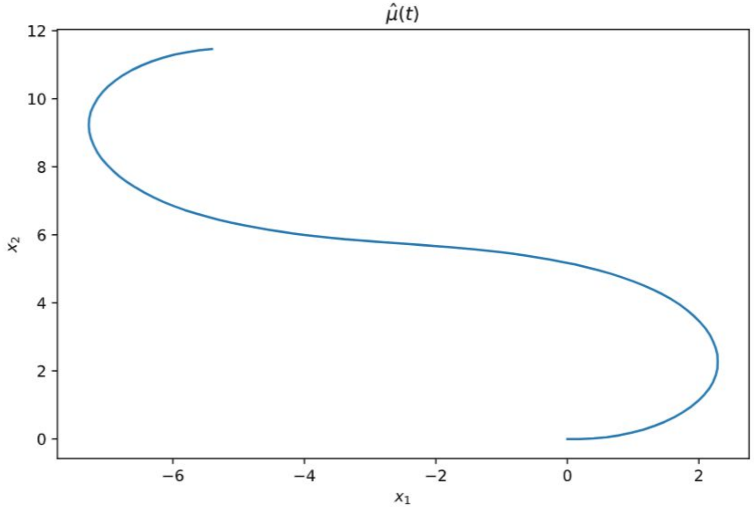


Figure 27: Estimated mean with respect to time from the 2000 training samples following the true SDE ($T = 10$).

Contents

Second topic

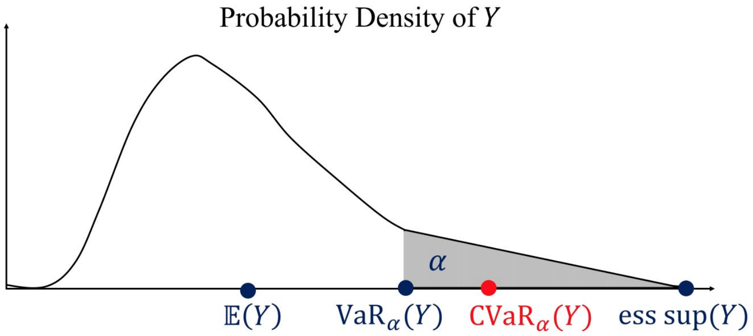
Devise useful conditions for optimality for general risk-averse optimal control settings

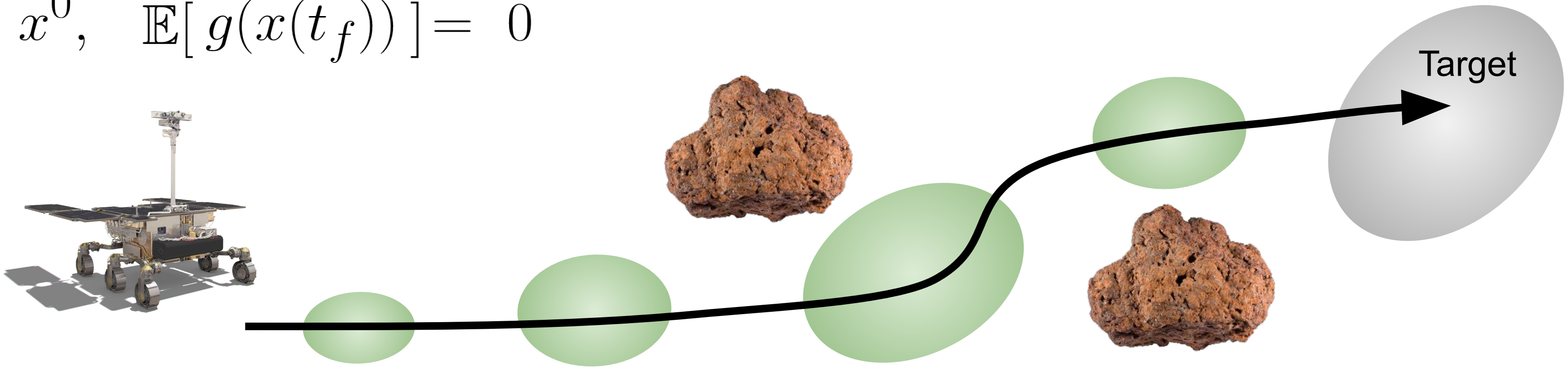
Collaboration with B. Bonnet, L2S

$$\min_{u \in \mathcal{U}} \rho\left[\int_0^{t_f} f^0(s, u(s), x(s))\, \mathrm{d}s\right]$$

$$\mathrm{d}x(s) = f(s, u(s), x(s))\, \mathrm{d}s + \sigma(s, u(s), x(s))\, \mathrm{d}W_s$$

$$x(0) = x^0, \quad \mathbb{E}[\, g(x(t_f))\,] = 0$$





Several settings have already been studied…

Here are few examples:

- M .Kohlmann and S. Tang, Minimization of Risk and Linear Quadratic Optimal Control Theory, SICON, 42:1118--1142, 2003. Note: uniquely for LQ problems.
- P. Sopasakis and others, Risk-Averse Model Predictive Control, Automatica, 100:281--288, 2019. Note: nested risk measures work in discrete-time and require expensive computations.
- A. Pichler and R. Schlotter, Risk-Averse Optimal Control in Continuous Time by Nesting Risk Measures, Mathematics of Operation Research, 2022. Note: Approximated HJB-based conditions which do not apply to CVaR

Main criticism

The aforementioned works look expensive or do not consider fully general coherent risk measures such as functionals which are at most sub-differentiable, e.g., CVaR

Risk-averse Pontryagin Maximum Principle: how this?

Idea: leverage duality properties of coherent risk measures to design conditions for optimality

$$\rho(Y) = \sup_{\xi \in \partial\rho(0)} \mathbb{E}[\xi Y], \quad Y \in L^2_{\mathcal{F}_{t_f}}(\Omega, \mathbb{R})$$

Example for $\mathrm{CVaR}_\alpha, \ \alpha \in (0,1)$:

$$\partial\rho(0) = \{\xi \in L^2_{\mathcal{F}_{t_f}}(\Omega, [0, 1/\alpha]) : \ \mathbb{E}[\xi] = 1\}$$

How we leverage this property:

1. Use set-valued representation of dynamics to define admissible control variations
2. Use Sion-type min-max results to swap inf and sup and find an optimal risk parameter ξ
3. Use Ito and BSDE calculus to define costate processes and maximality conditions

Risk-averse Pontryagin Maximum Principle: how this?

More specifically, if (x, u) is optimal, then (proof for running cost replaced by terminal cost $\varphi(x(t_f))$) :

Risk-averse Pontryagin Maximum Principle: how this?

More specifically, if (x, u) is optimal, then (proof for running cost replaced by terminal cost $\varphi(x(t_f))$) :

- Prove $\lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \mathbb{E}\left[\sup_{t \in [0,t_f]} \|x^\varepsilon_{g_1,g_2}(t) - x(t) - \varepsilon y_{g_1,g_2}(t)\| \right] = 0$, for some:

$$\begin{cases} \mathrm{d}x^\varepsilon_{g_1,g_2}(t) \in F(t, x^\varepsilon_{g_1,g_2}(t))\,\mathrm{d}(\lambda \times W)_t \\ x^\varepsilon_{g_1,g_2}(0) = x^0 \end{cases}$$

, where:

$$\begin{cases} \mathrm{d}y_{g_1,g_2}(t) = \Big(A(t) y_{g_1,g_2}(t) + g_1(t) \Big)\,\mathrm{d}t + \Big(D(t) y_{g_1,g_2}(t) + g_2(t) \Big)\,\mathrm{d}W_t \\ y_{g_1,g_2}(0) = 0, \quad (g_1, g_2) \text{ selection for } T_{F(\cdot, x(\cdot))}\big((f,\sigma)(\cdot, x(\cdot), u(\cdot))\big) \end{cases}$$

Risk-averse Pontryagin Maximum Principle: how this?

More specifically, if (x, u) is optimal, then (proof for running cost replaced by terminal cost $\varphi(x(t_f))$) :

- Prove $\lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \mathbb{E}\left[\sup_{t \in [0,t_f]} \|x^\varepsilon_{g_1,g_2}(t) - x(t) - \varepsilon y_{g_1,g_2}(t)\| \right] = 0$, for some:

$$\begin{cases} \mathrm{d}x^\varepsilon_{g_1,g_2}(t) \in F(t, x^\varepsilon_{g_1,g_2}(t))\, \mathrm{d}(\lambda \times W)_t \\ x^\varepsilon_{g_1,g_2}(0) = x^0 \end{cases}$$

, where:

$$\begin{cases} \mathrm{d}y_{g_1,g_2}(t) = \Big(A(t) y_{g_1,g_2}(t) + g_1(t) \Big)\, \mathrm{d}t + \Big(D(t) y_{g_1,g_2}(t) + g_2(t) \Big)\, \mathrm{d}W_t \\ y_{g_1,g_2}(0) = 0, \quad (g_1, g_2) \text{ selection for } T_{F(\cdot, x(\cdot))}\big((f,\sigma)(\cdot, x(\cdot), u(\cdot))\big) \end{cases}$$

- Combine duality of risk measures with the latter variational property to obtain that:

$$\sup_{(g_1,g_2)} \inf_{\xi \in \partial\rho(\varphi(x(t_f)))} \mathbb{E}\left[\left(\xi \mathfrak{p}_0 \nabla\varphi(x(t_f)) + \sum_{i=1}^{\ell} \mathfrak{p}_i \nabla g_i(x(t_f)) \right) \cdot y_{g_1,g_2}(t_f) \right] \le 0$$

Risk-averse Pontryagin Maximum Principle: how this?

More specifically, if (x, u) is optimal, then (proof for running cost replaced by terminal cost $\varphi(x(t_f))$) :

- Prove $\lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \mathbb{E}\left[\sup_{t \in [0,t_f]} \| x^{\varepsilon}_{g_1,g_2}(t) - x(t) - \varepsilon y_{g_1,g_2}(t) \| \right] = 0$, for some:

$$\begin{cases} \mathrm{d}x^{\varepsilon}_{g_1,g_2}(t) \in F(t, x^{\varepsilon}_{g_1,g_2}(t))\,\mathrm{d}(\lambda \times W)_t \\ x^{\varepsilon}_{g_1,g_2}(0) = x^0 \end{cases}, \text{ where: } \begin{cases} \mathrm{d}y_{g_1,g_2}(t) = \Big(A(t) y_{g_1,g_2}(t) + g_1(t) \Big)\,\mathrm{d}t + \Big(D(t) y_{g_1,g_2}(t) + g_2(t) \Big)\,\mathrm{d}W_t \\ y_{g_1,g_2}(0) = 0, \quad (g_1, g_2) \text{ selection for } T_{F(\cdot, x(\cdot))}\big((f,\sigma)(\cdot, x(\cdot), u(\cdot))\big) \end{cases}$$

- Combine duality of risk measures with the latter variational property to obtain that:

$$\sup_{(g_1,g_2)} \inf_{\xi \in \partial\rho(\varphi(x(t_f)))} \mathbb{E}\left[\left(\xi \mathfrak{p}_0 \nabla\varphi(x(t_f)) + \sum_{i=1}^{\ell} \mathfrak{p}_i \nabla g_i(x(t_f)) \right) \cdot y_{g_1,g_2}(t_f) \right] \le 0$$

- Leverage Sion-type min-max results to obtain the existence of (at least) one risk parameter ξ^* such that:

$$\sup_{(g_1,g_2)} \mathbb{E}\left[\left(\xi^* \mathfrak{p}_0 \nabla\varphi(x(t_f)) + \sum_{i=1}^{\ell} \mathfrak{p}_i \nabla g_i(x(t_f)) \right) \cdot y_{g_1,g_2}(t_f) \right] \le 0$$

Risk-averse Pontryagin Maximum Principle: how this?

More specifically, if (x, u) is optimal, then (proof for running cost replaced by terminal cost $\varphi(x(t_f))$) :

- Prove $\lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \mathbb{E}\left[\sup_{t \in [0,t_f]} \| x^{\varepsilon}_{g_1,g_2}(t) - x(t) - \varepsilon y_{g_1,g_2}(t) \| \right] = 0$, for some:

$$\begin{cases} \mathrm{d}x^{\varepsilon}_{g_1,g_2}(t) \in F(t, x^{\varepsilon}_{g_1,g_2}(t))\, \mathrm{d}(\lambda \times W)_t \\ x^{\varepsilon}_{g_1,g_2}(0) = x^0 \end{cases}$$

, where:

$$\begin{cases} \mathrm{d}y_{g_1,g_2}(t) = \Big(A(t) y_{g_1,g_2}(t) + g_1(t) \Big)\, \mathrm{d}t + \Big(D(t) y_{g_1,g_2}(t) + g_2(t) \Big)\, \mathrm{d}W_t \\ y_{g_1,g_2}(0) = 0, \quad (g_1, g_2) \text{ selection for } T_{F(\cdot, x(\cdot))}\big((f,\sigma)(\cdot, x(\cdot), u(\cdot))\big) \end{cases}$$

- Combine duality of risk measures with the latter variational property to obtain that:

$$\sup_{(g_1,g_2)} \inf_{\xi \in \partial\rho(\varphi(x(t_f)))} \mathbb{E}\left[\left(\xi \mathfrak{p}_0 \nabla\varphi(x(t_f)) + \sum_{i=1}^{\ell} \mathfrak{p}_i \nabla g_i(x(t_f)) \right) \cdot y_{g_1,g_2}(t_f) \right] \le 0$$

- Leverage Sion-type min-max results to obtain the existence of (at least) one risk parameter ξ^* such that:

$$\sup_{(g_1,g_2)} \mathbb{E}\left[\left(\xi^* \mathfrak{p}_0 \nabla\varphi(x(t_f)) + \sum_{i=1}^{\ell} \mathfrak{p}_i \nabla g_i(x(t_f)) \right) \cdot y_{g_1,g_2}(t_f) \right] \le 0$$

- Conclude by using Ito and BSDE calculus to define costate processes and maximality conditions

Theoretical guarantees

First order condition for optimality for general risk-averse settings

Theorem (Bonalli, Bonnet)

Define the Hamiltonian: $H(t, x, u, p, p^0, q) = p^\top f(t, x, u) + p^0 f^0(t, x, u) + q^\top \sigma(t, x, u)$

If (x, u) is optimal, there exist non-trivial multipliers $(p^0, (p, q), \xi^*) \in \{0, -1\} \times \mathcal{SP} \times \partial\rho(0)$ such that:

Theoretical guarantees

First order condition for optimality for general risk-averse settings

Theorem (Bonalli, Bonnet)

Define the Hamiltonian: $$H(t,x,u,p,p^0,q) = p^\top f(t,x,u) + p^0 f^0(t,x,u) + q^\top \sigma(t,x,u)$$

If (x,u) is optimal, there exist non-trivial multipliers $(p^0,(p,q),\xi^*) \in \{0,-1\} \times \mathcal{SP} \times \partial\rho(0)$ such that:

1. Adjoint dynamics:

$$\mathrm{d}p(t) = -\frac{\partial H}{\partial x}(t,x(t),u(t),p(t),p^0,q(t))\mathrm{d}t + q(t)\mathrm{d}W_t$$

Theoretical guarantees

First order condition for optimality for general risk-averse settings

Theorem (Bonalli, Bonnet)

Define the Hamiltonian: $$H(t,x,u,p,p^0,q) = p^\top f(t,x,u) + p^0 f^0(t,x,u) + q^\top \sigma(t,x,u)$$

If (x,u) is optimal, there exist non-trivial multipliers $(p^0,(p,q),\xi^*) \in \{0,-1\} \times \mathcal{SP} \times \partial\rho(0)$ such that:

1. Adjoint dynamics:

$$\mathrm{d}p(t) = -\frac{\partial H}{\partial x}(t,x(t),u(t),p(t),p^0,q(t))\mathrm{d}t + q(t)\mathrm{d}W_t$$

2. Maximality conditions:

$$H(t,x(t),u(t),p(t),p^0,q(t)) = \max_{u\in U} H(t,x(t),u,p(t),p^0,q(t))$$

$$\mathbb{E}\left[\xi^* \int_0^{t_f} f^0(t,x(t),u(t))\mathrm{d}t\right] = \sup_{\xi\in\partial\rho(0)} \mathbb{E}\left[\xi \int_0^{t_f} f^0(t,x(t),u(t))\mathrm{d}t\right]$$

Main novelty

We may design indirect shooting methods

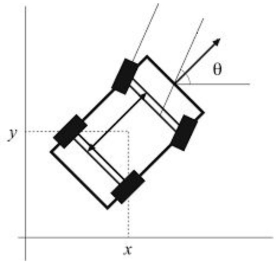
Numerical example (by Gabriel Velho)

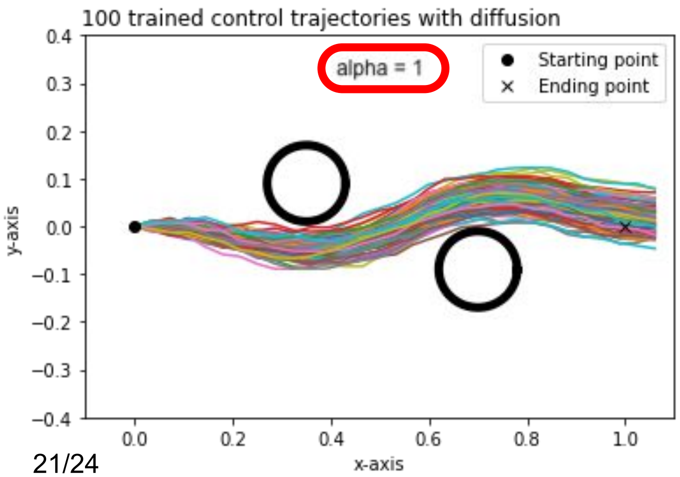
How we successfully apply these conditions: a numerical example

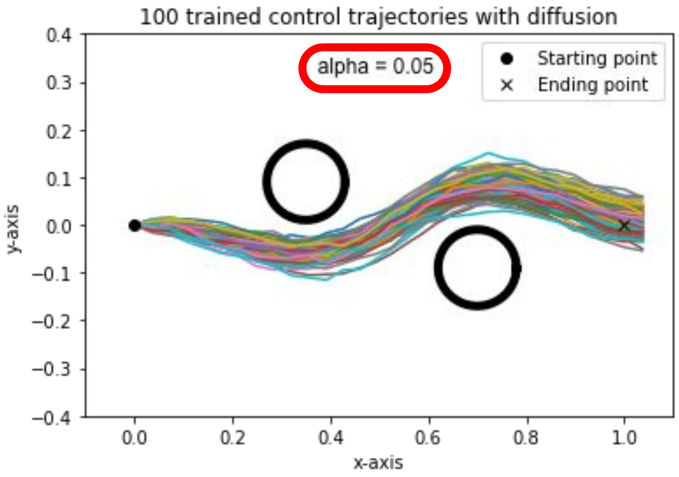
Risk-averse planning with two-dimensional stochastic Dubins' car model

$$\min_u \mathrm{CVaR}_\alpha \left(\|(x, y, \theta)(T) - (1, 0, 0)\|^2 + \int_0^T \left(u(t)^2 + \mathrm{dist}_{\mathrm{Obstacles}}(x(t))^2 \right) \mathrm{d}t \right)$$

$$\begin{cases} \mathrm{d}x(t) = v \cos \theta(t) \mathrm{d}t \\ \mathrm{d}y(t) = v \sin \theta(t) \mathrm{d}t + \sigma \mathrm{d}W_t, \quad (x, y, \theta)(0) = 0 \\ \mathrm{d}\theta(t) = u(t) \mathrm{d}t \end{cases}$$



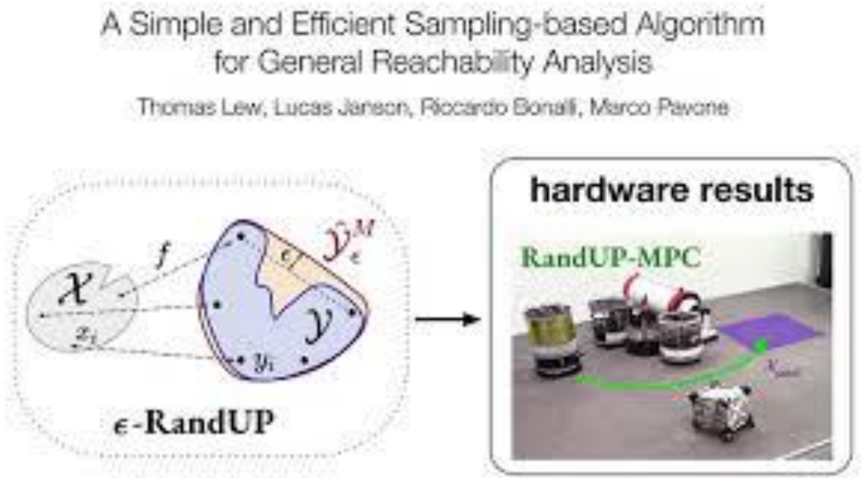




- We can leverage gradient ascent descent algorithms
- Higher number of safe trajectories (in probability) for lower values of α

Beyond numerical examples

Hardware experiments on space robots are to come soon!



At the Stanford Space Robotics Facility*

Aboard the International Space Station

*Video taken from: T. Lew, L. Janson, R. Bonalli, and M. Pavone. *A Simple and Efficient Sampling-based Reachability Analysis Algorithm*, L4DC, 2022.

Contents

1. Some challenges in controlling autonomous systems
2. Learning SDE with guarantees
    a. Relationship between SDE and Fokker-Planck equation
    b. RKHS-based learning problem and guarantees
3. Conditions for optimality in risk-averse optimal control
    a. Dual properties of coherent risk measures
    b. Risk-averse Pontryagin Maximum Principle and applications
4. Conclusion

Conclusion: future directions

Learning SDE:

1. Extensions to controlled SDE
2. Extensions to SDE with jumps to model human-robot interaction

Conclusion: future directions

Learning SDE:

1. Extensions to controlled SDE
2. Extensions to SDE with jumps to model human-robot interaction

Risk-averse optimal control:

1. Generalizing indirect shooting methods in the presence of constraints
2. Show Markovian-type properties of risk-averse optimal controls

Some references:

1. R. Bonalli and A. Rudi, *Non-Parametric Learning of Stochastic Differential Equations with Non-asymptotic Fast Rates of Convergence*, Submitted.
2. R. Bonalli and B. Bonnet, *First-Order Pontryagin Maximum Principle for Risk-Averse Stochastic Optimal Control Problems*, SIAM Journal on Control and Optimization, 61 (2023), pp. 1881-1909.