Ph.D. Proposal

Advancing Frontiers of Path Integral Theory for Stochastic Optimal Control

Apurva Patil

Dec 4, 2024

Outline

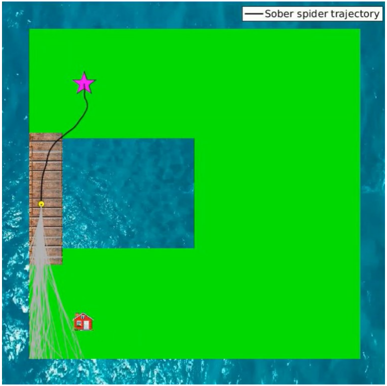
Outline

What is Path Integral Control?

- Path integral control is used to solve stochastic optimal control problems. It computes optimal control input online (in real-time) via Monte-Carlo simulations.

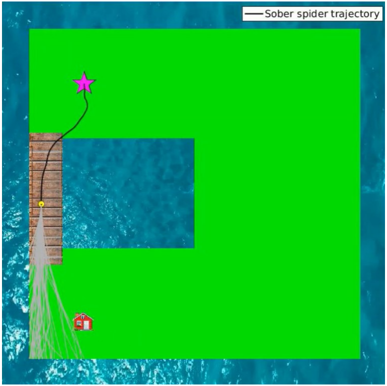
What is Path Integral Control?

- Path integral control is used to solve stochastic optimal control problems. It computes optimal control input online (in real-time) via Monte-Carlo simulations.

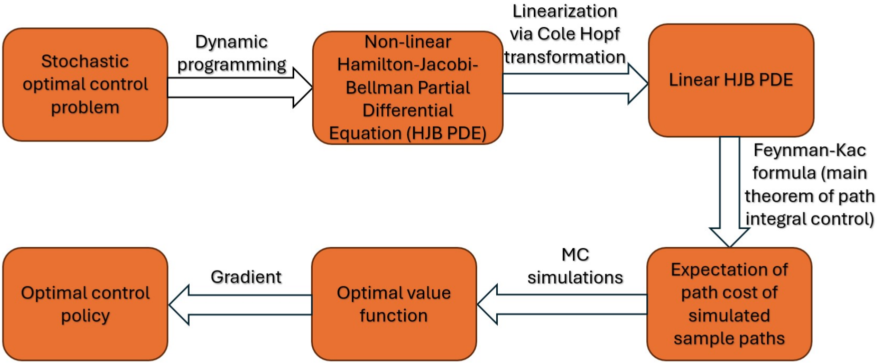


What is Path Integral Control?

- Path integral control is used to solve stochastic optimal control problems. It computes optimal control input online (in real-time) via Monte-Carlo simulations.
- The optimal control input is computed via the empirical mean of the path cost ("path integral") of simulated sample paths.



What is Path Integral Control?



What is Path Integral Control?

- The objective is to solve a Stochastic Optimal Control (SOC) problem

$$\begin{aligned} \min_{u} \quad & C(x_0, t_0, u(\cdot)) \\ \text{s.t.} \quad & dx(t) = f(x(t), t)dt + G(x(t), t)u(x(t), t)dt + \Sigma(x(t), t)dw(t). \end{aligned}$$

What is Path Integral Control?

- The objective is to solve a Stochastic Optimal Control (SOC) problem

$$\begin{aligned} \min_u \quad & C(x_0, t_0, u(\cdot)) \\ \text{s.t.} \quad & dx(t) = f(x(t), t)dt + G(x(t), t)u(x(t), t)dt + \Sigma(x(t), t)dw(t). \end{aligned}$$

- The stochastic dynamics is a control-affine n-dimensional Ito process $t \in [t_0, T]$ and $w(t)$ is an n-dimensional Brownian motion.

What is Path Integral Control?

- The objective is to solve a Stochastic Optimal Control (SOC) problem

$$\min_{u} \quad C(x_0, t_0, u(\cdot))$$
$$\text{s.t.} \quad dx(t) = f(x(t), t)dt + G(x(t), t)u(x(t), t)dt + \Sigma(x(t), t)dw(t).$$

- The stochastic dynamics is a control-affine n-dimensional Ito process $t \in [t_0, T]$ and $w(t)$ is an n-dimensional Brownian motion.
- The cost function is quadratic in u. $C(x_0, t_0, u(\cdot)) =$

$$\mathbb{E}_{\mathbf{x_0}, \mathbf{t_0}}\left[\underbrace{\int_{t_0}^{T} \left(V(x(t), t) + \frac{1}{2} u^\top R(x(t), t) u \right) dt}_{\substack{\text{Running cost} \\ \text{(e.g., travel distance)}}} + \underbrace{\psi(x(T))}_{\substack{\text{Terminal cost} \\ \text{(e.g., distance from home)}}} \right]$$

What is Path Integral Control?

- Using dynamic programming the optimal control can be computed as:

$$u^*(x,t) = -R^{-1}(x,t)G^\top(x,t)\partial_x J(x,t).$$

where $J(x,t)$ is the value function of the SOC problem which satisfies the Hamilton-Jacobi-Bellman (HJB) PDE

$$-\partial_t J = -\frac{1}{2}(\partial_x J)^\top G R^{-1} G^\top \partial_x J + V + f^\top \partial_x J + \frac{1}{2}\mathrm{Tr}(\Sigma\Sigma^\top \partial_x^2 J), \quad \forall x, t$$

with boundary condition $J(x,T) = \psi(x)$.

What is Path Integral Control?

- Using dynamic programming the optimal control can be computed as:

$$u^*(x,t) = -R^{-1}(x,t)G^\top(x,t)\partial_x J(x,t).$$

where $J(x,t)$ is the value function of the SOC problem which satisfies the Hamilton-Jacobi-Bellman (HJB) PDE

$$-\partial_t J = -\frac{1}{2}(\partial_x J)^\top G R^{-1} G^\top \partial_x J + V + f^\top \partial_x J + \frac{1}{2}\mathrm{Tr}(\Sigma\Sigma^\top \partial_x^2 J), \quad \forall x, t$$

with boundary condition $J(x,T) = \psi(x)$.

- HJB PDE is non-linear in J and can be high-dimensional $\Rightarrow$ difficult to solve analytically

What is Path Integral Control?

- Using dynamic programming the optimal control can be computed as:

$$u^*(x,t) = -R^{-1}(x,t)G^\top(x,t)\partial_x J(x,t).$$

where $J(x,t)$ is the value function of the SOC problem which satisfies the Hamilton-Jacobi-Bellman (HJB) PDE

$$-\partial_t J = -\frac{1}{2}(\partial_x J)^\top G R^{-1} G^\top \partial_x J + V + f^\top \partial_x J + \frac{1}{2}\mathrm{Tr}(\Sigma\Sigma^\top \partial_x^2 J), \quad \forall x, t$$

with boundary condition $J(x,T) = \psi(x)$.

- HJB PDE is non-linear in J and can be high-dimensional $\Rightarrow$ difficult to solve analytically
- Path integral approach makes the HJB PDE linear by making the following assumption:
Suppose $\exists\, \lambda > 0$ satisfying:

$$\underbrace{\Sigma(x,t)\Sigma^\top(x,t)}_{\text{noise covariance}} = \lambda G(x,t)\underbrace{R^{-1}(x,t)}_{\text{inverse of control cost}}G^\top(x,t).$$

What is Path Integral Control?

- The linearized HJB PDE can be solved by using the Feynman-Kac lemma:

$$J(x,t) = -\lambda \log \boxed{\mathbb{E}_{x,t}} \left[\underbrace{\exp\left(-\frac{1}{\lambda} \int_t^T V(x(t),t)dt - \frac{1}{\lambda}\psi(x(T)) \right)}_{\text{Path cost}} \right].$$

where $\mathbb{E}[\cdot]$ is with respect to the distribution $P(x)$ generated by the "uncontrolled" dynamics $dx = fdt + \Sigma dw$

What is Path Integral Control?

- The linearized HJB PDE can be solved by using the Feynman-Kac lemma:

$$J(x,t) = -\lambda \log \boxed{\mathbb{E}_{x,t}} \left[\underbrace{\exp\left(-\frac{1}{\lambda} \int_t^T V(x(t),t)dt - \frac{1}{\lambda}\psi(x(T)) \right)}_{\text{Path cost}} \right].$$

  where $\mathbb{E}[\cdot]$ is with respect to the distribution $P(x)$ generated by the "uncontrolled" dynamics $dx = fdt + \Sigma dw$

- Monte Carlo simulations

$$J(x,t) \approx -\lambda \log \boxed{\frac{1}{N}\sum_{i=1}^{N}} \underbrace{\exp\left(-\frac{1}{\lambda} \int_t^T V\left(x^{(i)}(t),t\right)dt - \frac{\psi\left(x^{(i)}(T)\right)}{\lambda} \right)}_{\text{Path cost}}$$

What is Path Integral Control?

- The linearized HJB PDE can be solved by using the Feynman-Kac lemma:

$$J(x,t) = -\lambda \log \boxed{\mathbb{E}_{x,t}} \left[\underbrace{\exp\left(-\frac{1}{\lambda} \int_t^T V(x(t),t)dt - \frac{1}{\lambda}\psi(x(T)) \right)}_{\text{Path cost}} \right].$$

where $\mathbb{E}[\cdot]$ is with respect to the distribution $P(x)$ generated by the "uncontrolled" dynamics $dx = fdt + \Sigma dw$

- Monte Carlo simulations

$$J(x,t) \approx -\lambda \log \boxed{\frac{1}{N}\sum_{i=1}^{N}} \underbrace{\exp\left(-\frac{1}{\lambda} \int_t^T V\left(x^{(i)}(t),t\right)dt - \frac{\psi\left(x^{(i)}(T)\right)}{\lambda} \right)}_{\text{Path cost}}$$

- Optimal control $u^*(x,t)$ of the SOC problem can also be computed by Monte Carlo simulations.

Outline

Brief History of Path Integral Control

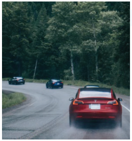
- Path integral control is inspired by the Path Integral formulation of quantum mechanics which considers every possible trajectory of a particle and computes their probabilities.
- [Yasue 1981, Guerra et al. 1983] identified the class of SOC problems in which the associated HJB equation coincides with the linear Schrödinger equation.
- [Itami 2001, Itami 2003] invoked the Feynman-Kac formula to numerically evaluate the solution of Schrödinger equation using Monte Carlo (Metropolis-Hastings) algorithm.
- [Kappen 2005] showed that a certain class of stochastic optimal control problems for which stochastic HJB equation can be linearized, can be solved by the path integral method.
- Model Predictive Path Integral (MPPI) control: a receding horizon implementation of path integral control [Williams et al. 2016, Williams et al. 2017]

Outline

Why Path Integral Control?

Simulator-driven: no analytical model required

Physical system



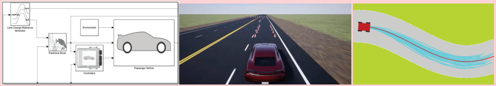
"Digital Twinning"

Equation-based modeling
$\dot{x} = f(x, u)$

Control Policy Synthesis
$u = \pi(x)$

Control Actions

Real-time Simulation Image: Matlab Vehicle Dynamics Blockset



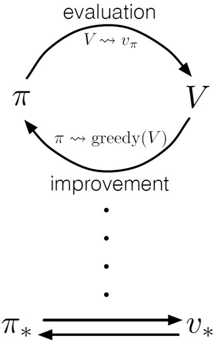
"Path integral control"

Why Path Integral Control?

One shot, online

Why Path Integral Control?

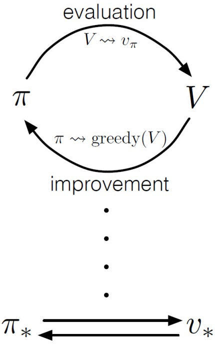
One shot, online



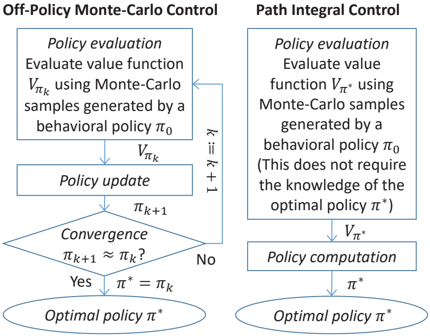
(a) RL technique

Why Path Integral Control?

One shot, online



(a) RL technique



(b) Monte-Carlo RL and path integral control

Why Path Integral Control?

- For a certain class of stochastic optimal control problems the required number of samples for path integral control depends only logarithmically on the dimension of the control input[1].

[1] A. Patil, G. Hanasusanto, T. Tanaka, “Discrete-Time Stochastic LQR via Path Integral Control and Its Sample Complexity Analysis," *IEEE Control Systems Letters (L-CSS)*

Why Path Integral Control?

- For a certain class of stochastic optimal control problems the required number of samples for path integral control depends only logarithmically on the dimension of the control input[1].
- Less susceptible to the curse of dimensionality

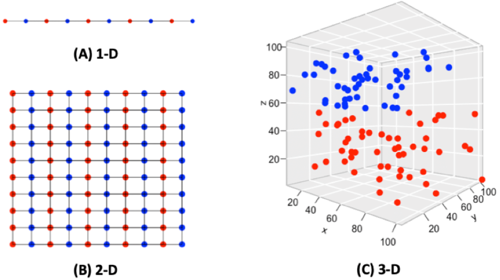


Figure: Grid-based approaches

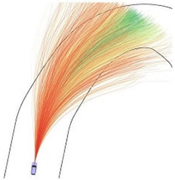
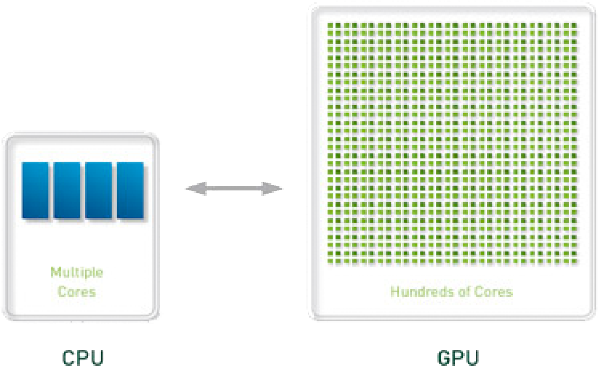


Figure: Path-integral approach

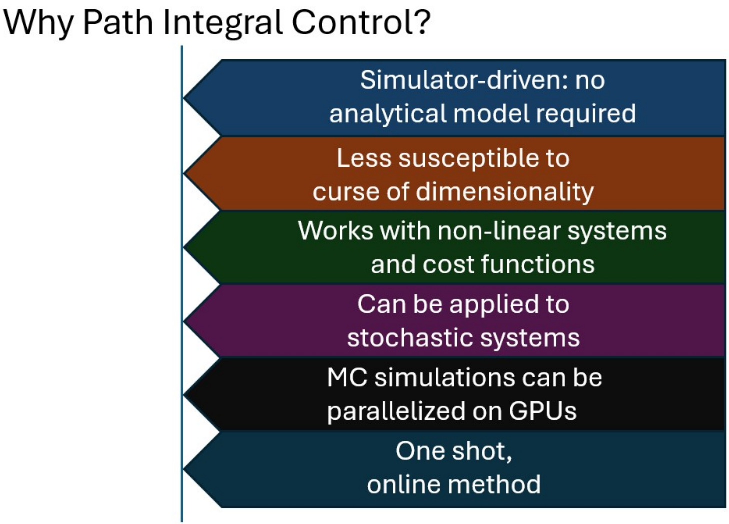
[1] A. Patil, G. Hanasusanto, T. Tanaka, "Discrete-Time Stochastic LQR via Path Integral Control and Its Sample Complexity Analysis," *IEEE Control Systems Letters (L-CSS)*

Why Path Integral Control?

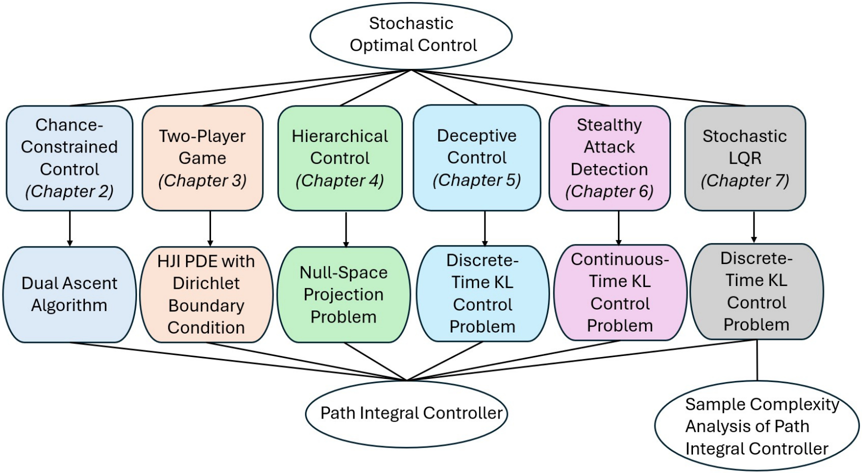
Monte Carlo simulations can be parallelized on GPUs which makes it effective for real-time control applications.



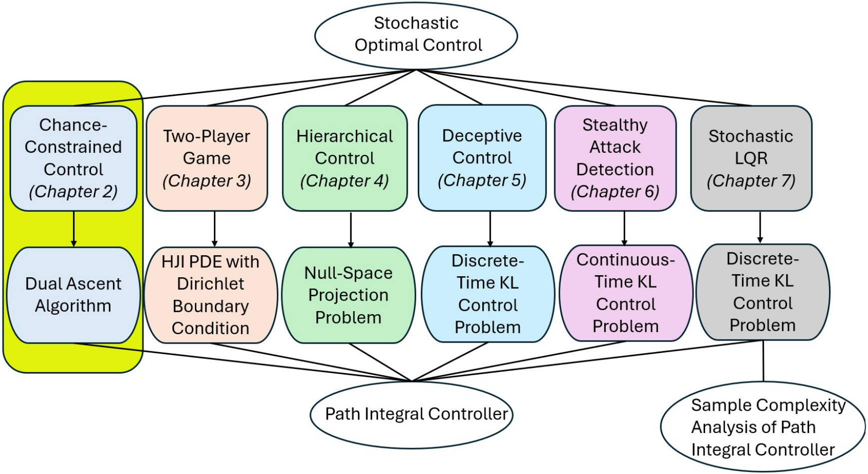
Why Path Integral Control?



Outline of the Ph.D. Work



Chance-Constrained Control



Outline

Background

Chance-constrained stochastic optimal control problem

The drunken spider problem[1]

- A drunken spider wants to take the shortest path to home.
- Probability of falling into the water should be small → chance constraint

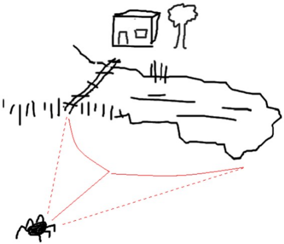
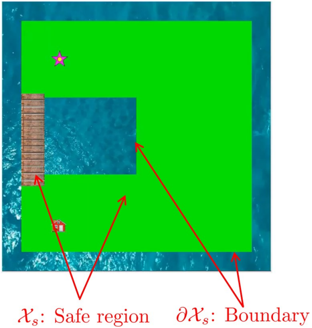


Image credit [1]

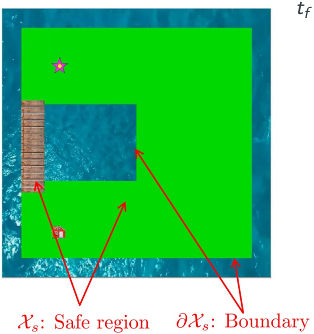
[1] Kappen "Path integrals and symmetry breaking for optimal control theory", *Journal of statistical mechanics: theory and experiment, 2005, no. 11*

Outline

Safe Region and Exit Time

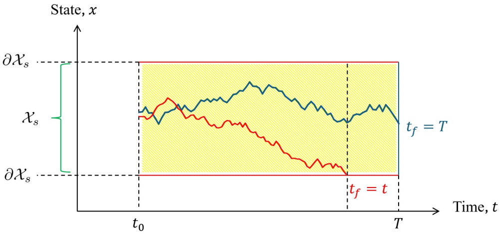


Safe Region and Exit Time



Exit Time (Final Time)

$$t_f = \begin{cases} T & \text{if } x(t) \in \mathcal{X}_s, \forall t \in (t_0, T) \\ \inf \{t \in (t_0, T) : x(t) \notin \mathcal{X}_s\} & \text{otherwise} \end{cases}$$



Chance-constrained Stochastic Optimal Control

$$\min_{u} \mathbb{E}_{x_0,t_0}\left[\int_{t_0}^{t_f}\left(V(x(t),t)+\frac{1}{2}u^\top R(x(t),t)u\right)dt + \psi(x(t_f))\cdot \mathbb{1}_{x(t_f)\in\mathcal{X}_s}\right]$$

$$\text{s.t.}\quad dx = fdt + Gudt + \Sigma dw, \quad x(t_0) = x_0,$$

$$\underbrace{P_{x_0,t_0}\left(\bigvee_{t\in(t_0,T]} x(t)\notin\mathcal{X}_s\right)}_{\text{Probability of failure } (P_{\text{fail}})} < \Delta \quad \text{(Chance constraint)}$$

Chance-constrained Stochastic Optimal Control

$$\min_u \mathbb{E}_{x_0,t_0}\left[\int_{t_0}^{t_f}\left(V(x(t),t)+\frac{1}{2}u^\top R(x(t),t)u\right)dt + \psi(x(t_f))\cdot \mathbb{1}_{x(t_f)\in\mathcal{X}_s}\right]$$

$$\text{s.t.} \quad dx = fdt + Gudt + \Sigma dw, \quad x(t_0) = x_0,$$

$$\underbrace{P_{x_0,t_0}\left(\bigvee_{t\in(t_0,T]} x(t)\notin\mathcal{X}_s\right)}_{\text{Probability of failure } (P_{\text{fail}})} < \Delta \quad \text{(Chance constraint)}$$

- This is a variable end-time problem - there is no cost after system fails.

Chance-constrained Stochastic Optimal Control

$$\min_u \mathbb{E}_{x_0,t_0}\left[\int_{t_0}^{t_f}\left(V(x(t),t)+\frac{1}{2}u^\top R(x(t),t)u\right)dt + \psi(x(t_f))\cdot \mathbb{1}_{x(t_f)\in\mathcal{X}_s}\right]$$

$$\text{s.t.} \quad dx = fdt + Gudt + \Sigma dw, \quad x(t_0) = x_0,$$

$$\underbrace{P_{x_0,t_0}\left(\bigvee_{t\in(t_0,T]} x(t)\notin\mathcal{X}_s\right)}_{\text{Probability of failure } (P_{\text{fail}})} < \Delta \quad \text{(Chance constraint)}$$

- This is a variable end-time problem - there is no cost after system fails.
- We consider end-to-end risk (not pointwise risk).

Chance-constrained Stochastic Optimal Control

$$\min_{u} \mathbb{E}_{x_0,t_0}\left[\int_{t_0}^{t_f}\left(V(x(t),t)+\frac{1}{2}u^\top R(x(t),t)u\right)dt + \psi(x(t_f))\cdot \mathbb{1}_{x(t_f)\in\mathcal{X}_s}\right]$$

$$\text{s.t.} \quad dx = fdt + Gudt + \Sigma dw, \quad x(t_0) = x_0,$$

$$\underbrace{P_{x_0,t_0}\left(\bigvee_{t\in(t_0,T]} x(t)\notin\mathcal{X}_s\right)}_{\text{Probability of failure } (P_{\text{fail}})} < \Delta \quad \text{(Chance constraint)}$$

- This is a variable end-time problem - there is no cost after system fails.
- We consider end-to-end risk (not pointwise risk).
- The acceptance of the possibility of failure is effective in reducing the conservatism of the controller even if the introduced probability of failure is practically negligible.

Outline

Related Work

- Iterative risk allocation scheme with Boole's bound [Ono et al. 2008]: Boole's bound is used to approximate the joint chance constraint and the user-specified risk "budget" is allocated optimally between timesteps.
- Lagrangian relaxation with Boole's bound [Ono et al. 2015]: Joint chance constraint is approximated using Boole's inequality, and Lagrangian relaxation is used to obtain an unconstrained optimal control problem which is solved using dynamic programming.
- Stochastic Control Barrier Functions [Santoyo et al. 2019]: Stochastic control barrier functions are used to derive sufficient conditions on the control input that bound the probability of failure.
- Reflection principle [Ariu et al. 2017]: Reflection principle of Brownian motion along with Boole's inequality is used to bound the failure probability in continuous-time.
- Generalized polynomial chaos [Nakka et al. 2019]: A stochastic optimal control problem is converted to a deterministic optimal control problem using generalized polynomial chaos expansion and then solved using sequential convex programming.
- Sampling-based approaches [Blackmore et al. 2010]
- Reinforcement learning [Huang et al. 2021]

Outline

Our Contributions

- We solve the chance-constrained stochastic optimal control problem without introducing any conservative approximation of the chance constraint.

Our Contributions

- We solve the chance-constrained stochastic optimal control problem without introducing any conservative approximation of the chance constraint.
- We introduce a dual SOC problem and prove that the strong duality exists between the original chance-constrained SOC problem and the dual SOC problem.

Our Contributions

- We solve the chance-constrained stochastic optimal control problem without introducing any conservative approximation of the chance constraint.
- We introduce a dual SOC problem and prove that the strong duality exists between the original chance-constrained SOC problem and the dual SOC problem.
- We propose a novel path-integral-based dual ascent algorithm to numerically solve the dual problem.

Our Contributions

- We solve the chance-constrained stochastic optimal control problem without introducing any conservative approximation of the chance constraint.
- We introduce a dual SOC problem and prove that the strong duality exists between the original chance-constrained SOC problem and the dual SOC problem.
- We propose a novel path-integral-based dual ascent algorithm to numerically solve the dual problem.

We provide an optimal solution the chance-constrained stochastic optimal control problem which can be computed online via Monte-Carlo samples of system trajectories (path integral control).

Outline

Lagrangian

- Lagrangian:

$$\mathcal{L}\left(x_0, t_0, u(\cdot); \eta\right) = C\left(x_0, t_0, u(\cdot)\right) + \eta \left[P_{x_0, t_0}\left(\bigvee_{t \in (t_0, T]} x(t) \notin \mathcal{X}_s \right) - \Delta \right]$$

where $\eta \geq 0$ is the Lagrange multiplier.

Lagrangian

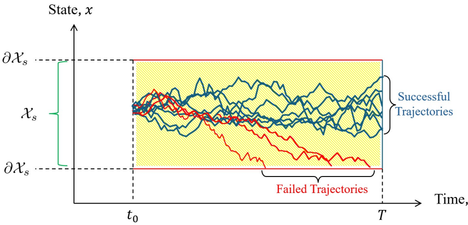
- Lagrangian:
$$\mathcal{L}\left(x_0, t_0, u(\cdot); \eta\right) = C\left(x_0, t_0, u(\cdot)\right) + \eta\left[P_{x_0,t_0}\left(\bigvee_{t\in(t_0,T]} x(t) \notin \mathcal{X}_s\right) - \Delta\right]$$
where $\eta \geq 0$ is the Lagrange multiplier.
- We prove that

$$P_{\text{fail}} = P_{x_0,t_0}\left(\bigvee_{t\in(t_0,T]} x(t) \notin \mathcal{X}_s\right)$$
$$= \mathbb{E}_{x_0,t_0}\left[\mathbb{1}_{x(t_f)\in\partial\mathcal{X}_s}\right]$$

Lagrangian

- Lagrangian:
  $$\mathcal{L}\left(x_0, t_0, u(\cdot); \eta\right) = C\left(x_0, t_0, u(\cdot)\right) + \eta\left[P_{x_0,t_0}\left(\bigvee_{t\in(t_0,T]} x(t) \notin \mathcal{X}_s\right) - \Delta\right]$$
  where $\eta \geq 0$ is the Lagrange multiplier.
- We prove that

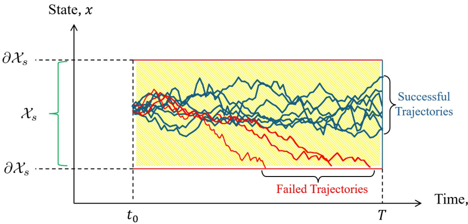
$$P_{\text{fail}} = P_{x_0,t_0}\left(\bigvee_{t\in(t_0,T]} x(t) \notin \mathcal{X}_s\right)$$
$$= \mathbb{E}_{x_0,t_0}\left[\mathbb{1}_{x(t_f)\in\partial\mathcal{X}_s}\right]$$



Lagrangian

- Lagrangian:
  $$\mathcal{L}(x_0, t_0, u(\cdot); \eta) = C(x_0, t_0, u(\cdot)) + \eta \left[P_{x_0, t_0}\left(\bigvee_{t \in (t_0, T]} x(t) \notin \mathcal{X}_s \right) - \Delta \right]$$
  where $\eta \geq 0$ is the Lagrange multiplier.
- We prove that

$$P_{\text{fail}} = P_{x_0, t_0}\left(\bigvee_{t \in (t_0, T]} x(t) \notin \mathcal{X}_s \right)$$
$$= \mathbb{E}_{x_0, t_0}\left[\mathbb{1}_{x(t_f) \in \partial \mathcal{X}_s} \right]$$



$$C(x_0, t_0, u(\cdot)) = \mathbb{E}_{x_0, t_0}\left[\int_{t_0}^{t_f} \left(V + \frac{1}{2} u^\top R u \right) dt + \psi(x(t_f)) \cdot \mathbb{1}_{x(t_f) \in \mathcal{X}_s} \right]$$

Lagrangian

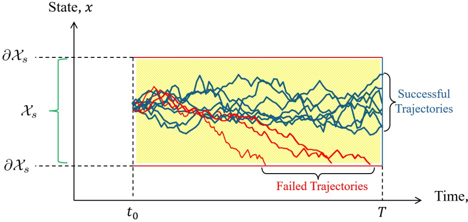
- Lagrangian:

$$\mathcal{L}(x_0, t_0, u(\cdot); \eta) = C(x_0, t_0, u(\cdot)) + \eta \left[P_{x_0,t_0}\left(\bigvee_{t\in(t_0,T]} x(t) \notin \mathcal{X}_s \right) - \Delta \right]$$

  where $\eta \geq 0$ is the Lagrange multiplier.

- We prove that

$$P_{\text{fail}} = P_{x_0,t_0}\left(\bigvee_{t\in(t_0,T]} x(t) \notin \mathcal{X}_s \right)$$
$$= \mathbb{E}_{x_0,t_0}\left[\mathbb{1}_{x(t_f)\in\partial\mathcal{X}_s} \right]$$



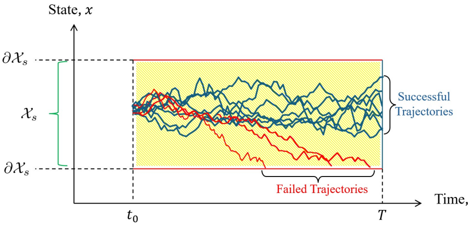
$$C(x_0, t_0, u(\cdot)) = \mathbb{E}_{x_0,t_0}\left[\int_{t_0}^{t_f} \left(V + \frac{1}{2} u^\top R u \right) dt + \psi(x(t_f)) \cdot \mathbb{1}_{x(t_f)\in\mathcal{X}_s} \right]$$

-

Lagrangian

- Lagrangian:
$$\mathcal{L}\left(x_0, t_0, u(\cdot); \eta\right) = C\left(x_0, t_0, u(\cdot)\right) + \eta\left[P_{x_0,t_0}\left(\bigvee_{t\in(t_0,T]} x(t) \notin \mathcal{X}_s\right) - \underbrace{\Delta}\right]$$
where $\eta \geq 0$ is the Lagrange multiplier.
- We prove that

$$P_{\text{fail}} = P_{x_0,t_0}\left(\bigvee_{t\in(t_0,T]} x(t) \notin \mathcal{X}_s\right)$$
$$= \mathbb{E}_{x_0,t_0}\left[\underbrace{\mathbb{1}_{x(t_f)\in\partial\mathcal{X}_s}}\right]$$



$$C(x_0, t_0, u(\cdot)) = \mathbb{E}_{x_0,t_0}\left[\int_{t_0}^{t_f}\left(V + \frac{1}{2}u^\top Ru\right)dt + \underbrace{\psi(x(t_f)) \cdot \mathbb{1}_{x(t_f)\in\mathcal{X}_s}}\right]$$

- Defining $\phi\left(x;\eta\right) := \underbrace{\psi\left(x\right) \cdot \mathbb{1}_{x\in\mathcal{X}_s}} + \eta \cdot \underbrace{\mathbb{1}_{x\in\partial\mathcal{X}_s}} - \eta\underbrace{\Delta}$,

$$\mathcal{L}(x_0,t_0,u(\cdot);\eta) = \mathbb{E}_{x_0,t_0}\left[\phi(x(t_f);\eta) + \int_{t_0}^{t_f}\left(\frac{1}{2}u^\top Ru + V\right)dt\right].$$

Dual SOC Problem

$$\max_{\eta \geq 0} \min_{u(\cdot)} \mathcal{L}\left(x_0, t_0, u(\cdot); \eta\right)$$
$$\text{s.t.} \quad dx = fdt + Gudt + \Sigma dw, \quad x(t_0) = x_0$$

Dual Function

$$g(\eta) := \min_{u(\cdot)} \mathcal{L}\left(x_0, t_0, u(\cdot); \eta\right)$$

Dual Problem

$$\max_{\eta \geq 0} g(\eta)$$

How to compute Dual Function?

Theorem (Verification Theorem)[2]: Suppose for a given $\eta \geq 0$, there exists a function $J : \overline{\mathcal{Q}} \to \mathbb{R}$ such that $J(x, t; \eta)$ solves the HJB PDE:

$$-\partial_t J = -\frac{1}{2}(\partial_x J)^\top G R^{-1} G^\top \partial_x J + V + f^\top \partial_x J + \frac{1}{2}\mathrm{Tr}(\Sigma\Sigma^\top \partial_x^2 J), \quad \forall (x,t) \in \mathcal{Q}$$

$$\lim_{(x,t)\to(y,t)} J(x, t; \eta) = \phi(y; \eta), \quad \forall (y, t) \in \partial\mathcal{Q} \quad \text{(Dirichlet BC)}$$

Then,

1. $J(x, t; \eta)$ is the value function, i.e.,

$$J(x, t; \eta) = \min_{u(\cdot)} \mathcal{L}(x, t, u(\cdot); \eta)$$

2. The optimal control is given by

$$u^*(x, t; \eta) = -R^{-1}(x, t) G^\top(x, t) \partial_x J(x, t; \eta).$$

[2] A. Patil, A. Duarte, A. Smith, F. Bisetti, T. Tanaka, “Chance-Constrained Stochastic Optimal Control via Path Integral and Finite Difference Methods," *2022 IEEE Conference on Decision and Control (CDC)*

Does the Strong Duality Exist?

- The value of the dual problem is always a lower bound for the primal problem

[3] A. Patil, A. Duarte, F. Bisetti, T. Tanaka, "Strong Duality and Dual Ascent Approach to Continuous-Time Chance-Constrained Stochastic Optimal Control," *submitted to Transactions on Automatic Control*

[4] We require to have continuity of P_{fail} with respect to η. We conjecture that this assumption is valid under mild conditions; a formal analysis is postponed as future work.

Does the Strong Duality Exist?

- The value of the dual problem is always a lower bound for the primal problem
- Strong duality: duality gap is zero

[3] A. Patil, A. Duarte, F. Bisetti, T. Tanaka, “Strong Duality and Dual Ascent Approach to Continuous-Time Chance-Constrained Stochastic Optimal Control," *submitted to Transactions on Automatic Control*

[4] We require to have continuity of P_{fail} with respect to η. We conjecture that this assumption is valid under mild conditions; a formal analysis is postponed as future work.

Does the Strong Duality Exist?

- The value of the dual problem is always a lower bound for the primal problem
- Strong duality: duality gap is zero
- Theorem (Strong Duality)[3]: Under certain assumptions on strict feasibility and continuity[4], we prove that
  - (i) There exists a dual optimal solution $0 \leq \eta^* < \infty$ that maximizes the dual function $g(\eta)$ and
  - (ii) A unique optimal policy $u^*(\cdot\,;\eta^*)$ of the problem $\arg\min_{u(\cdot)} \mathcal{L}\left(x_0, t_0, u(\cdot); \eta^*\right)$ is an optimal policy of the primal problem i.e., $C\left(x_0, t_0, u^*(\cdot\,;\eta^*)\right) = g(\eta^*)$.

[3] A. Patil, A. Duarte, F. Bisetti, T. Tanaka, "Strong Duality and Dual Ascent Approach to Continuous-Time Chance-Constrained Stochastic Optimal Control," *submitted to Transactions on Automatic Control*

[4] We require to have continuity of P_{fail} with respect to η. We conjecture that this assumption is valid under mild conditions; a formal analysis is postponed as future work.

Does the Strong Duality Exist?

- The value of the dual problem is always a lower bound for the primal problem
- Strong duality: duality gap is zero
- Theorem (Strong Duality)[3]: Under certain assumptions on strict feasibility and continuity[4], we prove that
  - (i) There exists a dual optimal solution $0 \leq \eta^* < \infty$ that maximizes the dual function $g(\eta)$ and
  - (ii) A unique optimal policy $u^*(\cdot\,;\eta^*)$ of the problem $\arg\min_{u(\cdot)} \mathcal{L}\left(x_0, t_0, u(\cdot); \eta^*\right)$ is an optimal policy of the primal problem i.e., $C\left(x_0, t_0, u^*(\cdot\,;\eta^*)\right) = g(\eta^*)$.

Strong duality exists!!

Solution of the chance-constrained SOC (primal)= Solution of the dual SOC

[3] A. Patil, A. Duarte, F. Bisetti, T. Tanaka, "Strong Duality and Dual Ascent Approach to Continuous-Time Chance-Constrained Stochastic Optimal Control," *submitted to Transactions on Automatic Control*

[4] We require to have continuity of P_{fail} with respect to η. We conjecture that this assumption is valid under mild conditions; a formal analysis is postponed as future work.

Does the Strong Duality Exist?

- The value of the dual problem is always a lower bound for the primal problem
- Strong duality: duality gap is zero
- Theorem (Strong Duality)[3]: Under certain assumptions on strict feasibility and continuity[4], we prove that
  - (i) There exists a dual optimal solution $0 \leq \eta^* < \infty$ that maximizes the dual function $g(\eta)$ and
  - (ii) A unique optimal policy $u^*(\cdot\,;\eta^*)$ of the problem $\arg\min_{u(\cdot)} \mathcal{L}(x_0, t_0, u(\cdot); \eta^*)$ is an optimal policy of the primal problem i.e., $C(x_0, t_0, u^*(\cdot\,;\eta^*)) = g(\eta^*)$.

Strong duality exists!!

Solution of the chance-constrained SOC (primal)= Solution of the dual SOC

How to solve the dual SOC problem?

[3] A. Patil, A. Duarte, F. Bisetti, T. Tanaka, "Strong Duality and Dual Ascent Approach to Continuous-Time Chance-Constrained Stochastic Optimal Control," *submitted to Transactions on Automatic Control*

[4] We require to have continuity of P_{fail} with respect to η. We conjecture that this assumption is valid under mild conditions; a formal analysis is postponed as future work.

How to Solve the Dual Problem?

- Gradient ascent

$$\eta \leftarrow \eta + \gamma(\underbrace{P_{\mathrm{fail}}(x_0, t_0, u^*(\cdot;\eta)) - \Delta}_{\text{gradient}})$$

How to Solve the Dual Problem?

- Gradient ascent

$$\eta \leftarrow \eta + \gamma(\underbrace{P_{\text{fail}}(x_0, t_0, u^*(\cdot;\eta)) - \Delta}_{\text{gradient}})$$

- How to compute $P_{\text{fail}}(x_0, t_0, u^*(\cdot;\eta))$?

How to Solve the Dual Problem?

- Gradient ascent

$$\eta \leftarrow \eta + \gamma(\underbrace{P_{\text{fail}}(x_0, t_0, u^*(\cdot;\eta)) - \Delta}_{\text{gradient}})$$

- How to compute $P_{\text{fail}}(x_0, t_0, u^*(\cdot;\eta))$?
  - Compute $u^*(\cdot;\eta)$
  - Sample N trajectories $\{x^{(i)}\}_{i=1}^{N}$ under u^*
  - Use Monte Carlo

$$P_{\text{fail}}(x_0, t_0, u^*(\cdot)) \approx \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}_{x^{(i)}(t_f)\in\partial\mathcal{X}_s}$$

How to Solve the Dual Problem?

- Gradient ascent

$$\eta \leftarrow \eta + \gamma(\underbrace{P_{\text{fail}}(x_0, t_0, u^*(\cdot;\eta)) - \Delta}_{\text{gradient}})$$

- How to compute $P_{\text{fail}}(x_0, t_0, u^*(\cdot;\eta))$?
  - Compute $u^*(\cdot;\eta)$
  - Sample N trajectories $\{x^{(i)}\}_{i=1}^N$ under u^*
  - Use Monte Carlo

$$P_{\text{fail}}(x_0, t_0, u^*(\cdot)) \approx \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}_{x^{(i)}(t_f)\in\partial\mathcal{X}_s}$$

- Can we find $P_{\text{fail}}(x_0, t_0, u^*(\cdot))$ without constructing u^*?

Computation of $P_{\text{fail}}((x_0, t_0, u^*(\cdot))$

- Theorem[5]: Suppose we sample N trajectories of the "uncontrolled" dynamics $dx = fdt + \Sigma dw$ and let $r^{(i)}$ be the path cost of the sample path i

$$r^{(i)} = \exp\left(-\frac{\phi\left(x^{(i)}(t_f);\eta\right)}{\lambda} - \frac{1}{\lambda}\int_{t_0}^{t_f} V\Big(x^{(i)}(s), s\Big)\, ds\right).$$

Then as $N \to \infty$,

$$\sum_{i=1}^{N} \frac{r^{(i)}}{\sum_{i=1}^{N} r^{(i)}} \mathbb{1}_{x^{(i)}(t_f)\in\partial\mathcal{X}_s} \overset{a.s.}{\to} P_{\text{fail}}(x_0, t_0, u^*(\cdot))$$

[5] A. Patil, A. Duarte, F. Bisetti, T. Tanaka, “Strong Duality and Dual Ascent Approach to Continuous-Time Chance-Constrained Stochastic Optimal Control," *submitted to Transactions on Automatic Control*

Computation of $P_{\mathrm{fail}}((x_0, t_0, u^*(\cdot))$

- Theorem[5]: Suppose we sample N trajectories of the "uncontrolled" dynamics $dx = fdt + \Sigma dw$ and let $r^{(i)}$ be the path cost of the sample path i

$$r^{(i)} = \exp\left(-\frac{\phi\left(x^{(i)}(t_f);\eta\right)}{\lambda} - \frac{1}{\lambda}\int_{t_0}^{t_f} V\left(x^{(i)}(s), s\right) ds\right).$$

Then as $N \to \infty$,

$$\sum_{i=1}^{N} \frac{r^{(i)}}{\sum_{i=1}^{N} r^{(i)}} \mathbb{1}_{x^{(i)}(t_f)\in\partial\mathcal{X}_s} \overset{a.s.}{\to} P_{\mathrm{fail}}(x_0, t_0, u^*(\cdot))$$

- We do not need u^*. Simply simulate the "uncontrolled" dynamics $dx = fdt + \Sigma dw$ and use Monte Carlo!

[5] A. Patil, A. Duarte, F. Bisetti, T. Tanaka, "Strong Duality and Dual Ascent Approach to Continuous-Time Chance-Constrained Stochastic Optimal Control," *submitted to Transactions on Automatic Control*

To Sum Up...Our Approach to solve the Chance-Constrained SOC Problem

- Step 1: Find $P_{\mathrm{fail}}(x_0, t_0, u^*(\cdot))$ via MC simulations of uncontrolled dynamics $\Rightarrow$ doesn't require u^*.

[6] A. Patil, A. Duarte, A. Smith, F. Bisetti, T. Tanaka, "Chance-Constrained Stochastic Optimal Control via Path Integral and Finite Difference Methods," *2022 IEEE Conference on Decision and Control (CDC)*

To Sum Up...Our Approach to solve the Chance-Constrained SOC Problem

- Step 1: Find $P_{\mathrm{fail}}(x_0, t_0, u^*(\cdot))$ via MC simulations of uncontrolled dynamics $\Rightarrow$ doesn't require u^*.
- Step 2: Use gradient ascent $\eta \leftarrow \eta + \gamma(P_{\mathrm{fail}}(x_0, t_0, u^*(\cdot;\eta)) - \Delta)$, and find the "magic" η^* that solves the problem $\max_{\eta \geq 0} g(\eta)$

[6] A. Patil, A. Duarte, A. Smith, F. Bisetti, T. Tanaka, "Chance-Constrained Stochastic Optimal Control via Path Integral and Finite Difference Methods," *2022 IEEE Conference on Decision and Control (CDC)*

To Sum Up...Our Approach to solve the Chance-Constrained SOC Problem

- Step 1: Find $P_{\text{fail}}(x_0, t_0, u^*(\cdot))$ via MC simulations of uncontrolled dynamics $\Rightarrow$ doesn't require u^*.
- Step 2: Use gradient ascent $\eta \leftarrow \eta + \gamma(P_{\text{fail}}(x_0, t_0, u^*(\cdot;\eta)) - \Delta)$, and find the "magic" η^* that solves the problem $\max_{\eta \geq 0} g(\eta)$

According to the strong duality theorem, a unique optimal policy $u^*(\cdot\,;\eta^*)$ of the problem $\arg\min_{u(\cdot)} \mathcal{L}(x_0, t_0, u(\cdot);\eta^*)$ is an optimal policy of the chance-constrained SOC problem.

[6] A. Patil, A. Duarte, A. Smith, F. Bisetti, T. Tanaka, "Chance-Constrained Stochastic Optimal Control via Path Integral and Finite Difference Methods," *2022 IEEE Conference on Decision and Control (CDC)*

To Sum Up...Our Approach to solve the Chance-Constrained SOC Problem

- Step 1: Find $P_{\mathrm{fail}}(x_0, t_0, u^*(\cdot))$ via MC simulations of uncontrolled dynamics $\Rightarrow$ doesn't require u^*.
- Step 2: Use gradient ascent $\eta \leftarrow \eta + \gamma(P_{\mathrm{fail}}(x_0, t_0, u^*(\cdot;\eta)) - \Delta)$, and find the "magic" η^* that solves the problem $\max_{\eta \geq 0} g(\eta)$

According to the strong duality theorem, a unique optimal policy $u^*(\cdot\,;\eta^*)$ of the problem $\arg\min_{u(\cdot)} \mathcal{L}\left(x_0, t_0, u(\cdot); \eta^*\right)$ is an optimal policy of the chance-constrained SOC problem.

- Step 3: Solve $\arg\min_{u(\cdot)} \mathcal{L}\left(x_0, t_0, u(\cdot); \eta^*\right)$.

[6] A. Patil, A. Duarte, A. Smith, F. Bisetti, T. Tanaka, "Chance-Constrained Stochastic Optimal Control via Path Integral and Finite Difference Methods," *2022 IEEE Conference on Decision and Control (CDC)*

To Sum Up...Our Approach to solve the Chance-Constrained SOC Problem

- Step 1: Find $P_{\mathrm{fail}}(x_0, t_0, u^*(\cdot))$ via MC simulations of uncontrolled dynamics $\Rightarrow$ doesn't require u^*.
- Step 2: Use gradient ascent $\eta \leftarrow \eta + \gamma(P_{\mathrm{fail}}(x_0, t_0, u^*(\cdot; \eta)) - \Delta)$, and find the "magic" η^* that solves the problem $\max_{\eta \geq 0} g(\eta)$

According to the strong duality theorem, a unique optimal policy $u^*(\cdot\ ; \eta^*)$ of the problem $\arg\min_{u(\cdot)} \mathcal{L}(x_0, t_0, u(\cdot); \eta^*)$ is an optimal policy of the chance-constrained SOC problem.

- Step 3: Solve $\arg\min_{u(\cdot)} \mathcal{L}(x_0, t_0, u(\cdot); \eta^*)$. How?

[6] A. Patil, A. Duarte, A. Smith, F. Bisetti, T. Tanaka, "Chance-Constrained Stochastic Optimal Control via Path Integral and Finite Difference Methods," *2022 IEEE Conference on Decision and Control (CDC)*

To Sum Up...Our Approach to solve the Chance-Constrained SOC Problem

- Step 1: Find $P_{\text{fail}}(x_0, t_0, u^*(\cdot))$ via MC simulations of uncontrolled dynamics $\Rightarrow$ doesn't require u^*.
- Step 2: Use gradient ascent $\eta \leftarrow \eta + \gamma(P_{\text{fail}}(x_0, t_0, u^*(\cdot; \eta)) - \Delta)$, and find the "magic" η^* that solves the problem $\max_{\eta \geq 0} g(\eta)$

According to the strong duality theorem, a unique optimal policy $u^*(\cdot\,; \eta^*)$ of the problem $\arg\min_{u(\cdot)} \mathcal{L}(x_0, t_0, u(\cdot); \eta^*)$ is an optimal policy of the chance-constrained SOC problem.

- Step 3: Solve $\arg\min_{u(\cdot)} \mathcal{L}(x_0, t_0, u(\cdot); \eta^*)$. How? Use the verification theorem[6] $\Rightarrow$ solve the HJB PDE

[6] A. Patil, A. Duarte, A. Smith, F. Bisetti, T. Tanaka, "Chance-Constrained Stochastic Optimal Control via Path Integral and Finite Difference Methods," *2022 IEEE Conference on Decision and Control (CDC)*

To Sum Up...Our Approach to solve the Chance-Constrained SOC Problem

- Step 1: Find $P_{\text{fail}}(x_0, t_0, u^*(\cdot))$ via MC simulations of uncontrolled dynamics $\Rightarrow$ doesn't require u^*.
- Step 2: Use gradient ascent $\eta \leftarrow \eta + \gamma(P_{\text{fail}}(x_0, t_0, u^*(\cdot;\eta)) - \Delta)$, and find the "magic" η^* that solves the problem $\max_{\eta \geq 0} g(\eta)$

According to the strong duality theorem, a unique optimal policy $u^*(\cdot\,;\eta^*)$ of the problem $\arg\min_{u(\cdot)} \mathcal{L}(x_0, t_0, u(\cdot); \eta^*)$ is an optimal policy of the chance-constrained SOC problem.

- Step 3: Solve $\arg\min_{u(\cdot)} \mathcal{L}(x_0, t_0, u(\cdot); \eta^*)$. How? Use the verification theorem[6] $\Rightarrow$ solve the HJB PDE

We use two numerical methods to solve the HJB PDE:

- Finite Difference Method (a grid-based approach)
- Path Integral Method

[6] A. Patil, A. Duarte, A. Smith, F. Bisetti, T. Tanaka, "Chance-Constrained Stochastic Optimal Control via Path Integral and Finite Difference Methods," *2022 IEEE Conference on Decision and Control (CDC)*

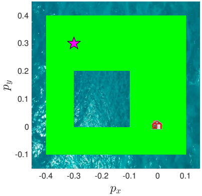
Outline

Finite Difference Method

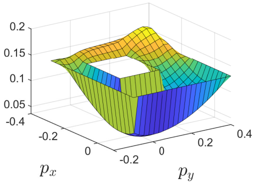
Computational domain is discretized into a finite grid points and the solution to the PDE is sought at these locations.

Finite Difference Method

Computational domain is discretized into a finite grid points and the solution to the PDE is sought at these locations.



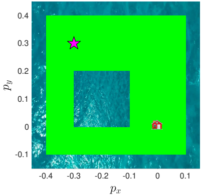
domain



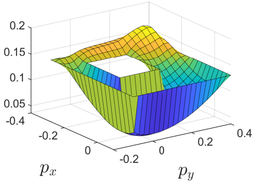
$J(x, t_0; \eta)$

Finite Difference Method

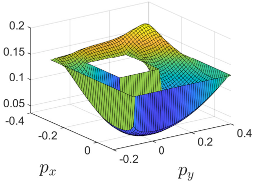
Computational domain is discretized into a finite grid points and the solution to the PDE is sought at these locations.



domain



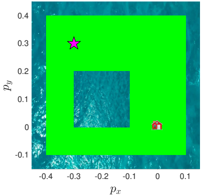
$J(x, t_0; \eta)$



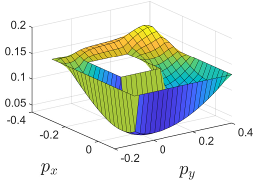
$J(x, t_0; \eta)$

Finite Difference Method

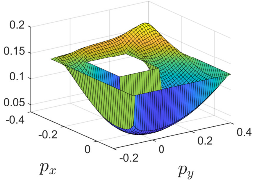
Computational domain is discretized into a finite grid points and the solution to the PDE is sought at these locations.



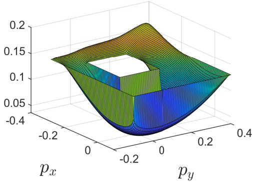
domain



$J(x, t_0; \eta)$



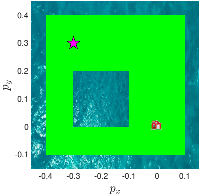
$J(x, t_0; \eta)$



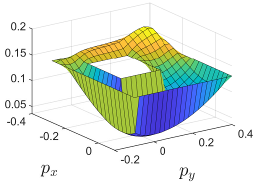
$J(x, t_0; \eta)$

Finite Difference Method

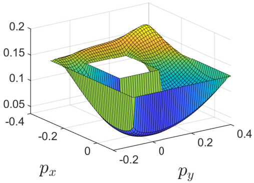
Computational domain is discretized into a finite grid points and the solution to the PDE is sought at these locations.



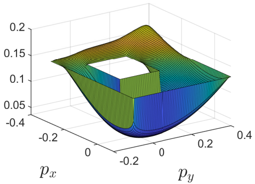
domain



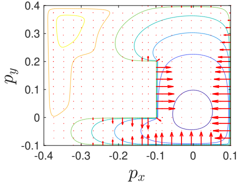
$J(x, t_0; \eta)$



$J(x, t_0; \eta)$



$J(x, t_0; \eta)$



$u^*(x, t_0; \eta)$

Finite Difference Method: Limitations

- Curse of dimensionality – Gridding is prohibitive for problems with higher dimensions.
- HJB equation for our SOC must be solved backward-in-time, which is inconvenient for real-time implementations.
- FDM computes the global solution of $J(x, t; \eta)$ and $u^*(x, t; \eta)$ over the entire domain $\overline{\mathcal{Q}}$ even if the majority of the state-time pairs (x, t) will never be visited by the actual system.

Finite Difference Method: Limitations

- Curse of dimensionality – Gridding is prohibitive for problems with higher dimensions.
- HJB equation for our SOC must be solved backward-in-time, which is inconvenient for real-time implementations.
- FDM computes the global solution of $J(x, t; \eta)$ and $u^*(x, t; \eta)$ over the entire domain $\overline{\mathcal{Q}}$ even if the majority of the state-time pairs (x, t) will never be visited by the actual system.

We want an algorithm to compute u^* on-the-fly for the given η^* and the current state-time pair (x, t).

Path Integral Method

- Computes the solution $J(x, t; \eta)$ of the HJB PDE at an arbitrary (x, t) using forward-in-time Monte-Carlo simulations of system trajectories.
- Optimal control $u^*(x, t; \eta)$ can also be computed by Monte-Carlo simulation without solving HJB equation backward in time.
- Massively parallelizable on GPUs.
- Path integral method is considered less susceptible to curse of dimensionality

Path Integral Method

- Computes the solution $J(x, t; \eta)$ of the HJB PDE at an arbitrary (x, t) using forward-in-time Monte-Carlo simulations of system trajectories.
- Optimal control $u^*(x, t; \eta)$ can also be computed by Monte-Carlo simulation without solving HJB equation backward in time.
- Massively parallelizable on GPUs.
- Path integral method is considered less susceptible to curse of dimensionality

For a certain class of stochastic optimal control problems the required number of samples depends only logarithmically on the dimension of the control input.

Path-Integral-Based Dual Ascent Algorithm

Algorithm 1 Dual ascent via path integral approach

Require: Error tolerance $\epsilon > 0$, learning rate $\gamma > 0$

```
Choose initial η
while True do
    Compute the failure probability P_fail(x_0, t_0, u*(·; η))
    if |P_fail(x_0, t_0, u*(·; η)) − Δ| < ε then
        Find u*(·; η) solving an HJB PDE
        Return u*(·; η)
    end if
    η ← η + γ(P_fail(x_0, t_0, u*(·; η)) − Δ)
end while
```

Path Integral Method: Limitations

- Only applicable to certain classes of problems that satisfy the following assumption:
  $\exists\, \lambda > 0$ satisfying:

$$\underbrace{\Sigma(x,t)\Sigma^\top(x,t)}_{\text{noise covariance}} = \lambda G(x,t)\underbrace{R^{-1}(x,t)}_{\text{inverse of control cost}}G^\top(x,t).$$

[7] S. Satoh et al., "An iterative method for nonlinear stochastic optimal control based on path integrals," *IEEE Transactions on Automatic Control, 2016.*

[8] G. Williams et al., "Information theoretic MPC for model-based reinforcement learning," *IEEE ICRA, 2017.*

Path Integral Method: Limitations

- Only applicable to certain classes of problems that satisfy the following assumption:
  $\exists\, \lambda > 0$ satisfying:

$$\underbrace{\Sigma(x,t)\Sigma^\top(x,t)}_{\text{noise covariance}} = \lambda G(x,t)\underbrace{R^{-1}(x,t)}_{\text{inverse of control cost}}G^\top(x,t).$$

  Removal of this assumption is discussed in some of the literature. [7] [8]

[7] S. Satoh et al., "An iterative method for nonlinear stochastic optimal control based on path integrals," *IEEE Transactions on Automatic Control, 2016.*

[8] G. Williams et al., "Information theoretic MPC for model-based reinforcement learning," *IEEE ICRA, 2017.*

Path Integral Method: Limitations

- Only applicable to certain classes of problems that satisfy the following assumption:
  $\exists\ \lambda > 0$ satisfying:

$$\underbrace{\Sigma(x,t)\Sigma^\top(x,t)}_{\text{noise covariance}} = \lambda G(x,t)\underbrace{R^{-1}(x,t)}_{\text{inverse of control cost}}G^\top(x,t).$$

  Removal of this assumption is discussed in some of the literature. [7] [8]
- Computationally heavy

[7] S. Satoh et al., "An iterative method for nonlinear stochastic optimal control based on path integrals," *IEEE Transactions on Automatic Control, 2016.*

[8] G. Williams et al., "Information theoretic MPC for model-based reinforcement learning," *IEEE ICRA, 2017.*

Path Integral Method: Limitations

- Only applicable to certain classes of problems that satisfy the following assumption:
  $\exists\, \lambda > 0$ satisfying:

$$\underbrace{\Sigma(x,t)\Sigma^{\top}(x,t)}_{\text{noise covariance}} = \lambda G(x,t)\underbrace{R^{-1}(x,t)}_{\text{inverse of control cost}}G^{\top}(x,t).$$

  Removal of this assumption is discussed in some of the literature. [7] [8]
- Computationally heavy
- The outcome of path integral control is probabilistic; hence applying path integral controller to safety-critical systems would require rigorous performance guarantees. However, the sample complexity of the path integral control is not well-studied in the literature.

---

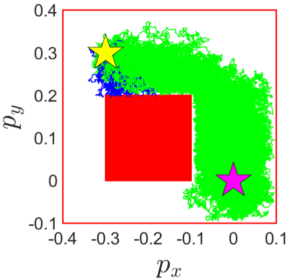
[7] S. Satoh et al., "An iterative method for nonlinear stochastic optimal control based on path integrals," *IEEE Transactions on Automatic Control, 2016.*

[8] G. Williams et al., "Information theoretic MPC for model-based reinforcement learning," *IEEE ICRA, 2017.*

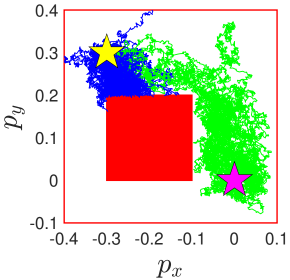
Outline

Example: Single Integrator

$$dp_x = -k_x p_x dt + u_x dt + \sigma dw_x$$
$$dp_y = -k_y p_y dt + u_y dt + \sigma dw_y$$



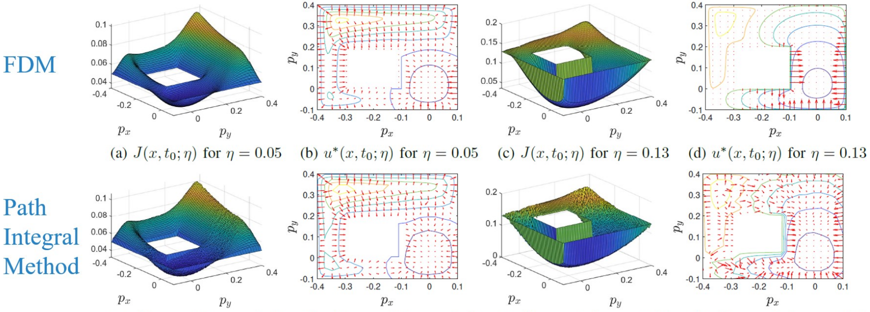
(a) $\Delta = 0.25$



(b) $\Delta = 0.9$

Example: Single Integrator

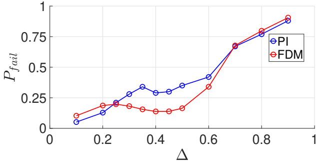
$$J(x, t_0; \eta) = \min_{u(\cdot)} \mathbb{E}_{x,t_0}\left[\phi\left(x(t_f); \eta\right) + \int_{t_0}^{t_f} \left(\frac{1}{2} u^\top R u + V\right) dt\right].$$



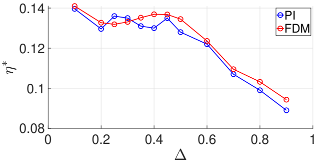
(a) $J(x, t_0; \eta)$ for $\eta = 0.05$ (b) $u^*(x, t_0; \eta)$ for $\eta = 0.05$ (c) $J(x, t_0; \eta)$ for $\eta = 0.13$ (d) $u^*(x, t_0; \eta)$ for $\eta = 0.13$

(e) $J(x, t_0; \eta)$ for $\eta = 0.05$ (f) $u^*(x, t_0; \eta)$ for $\eta = 0.05$ (g) $J(x, t_0; \eta)$ for $\eta = 0.13$ (h) $u^*(x, t_0; \eta)$ for $\eta = 0.13$

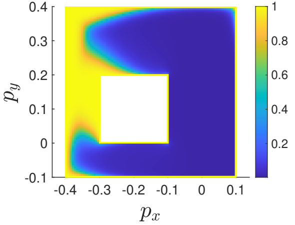
Example: Single Integrator



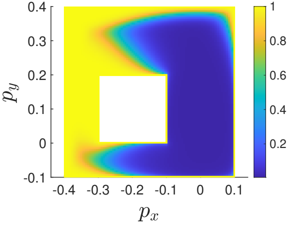
(a) $P_{\text{fail}}(x_0, t_0, u^*(\cdot\,;\eta^*))$ vs Δ



(b) η^* vs Δ



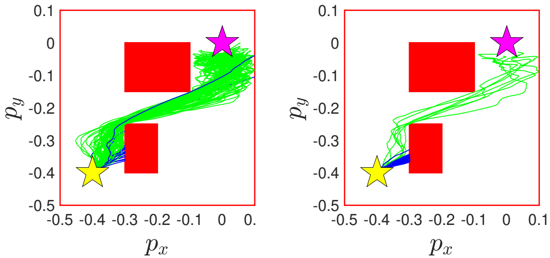
(c) $P_{\text{fail}}(x, t_0, u^*(\cdot\,;\eta^*))$ for $\Delta = 0.25$



(d) $P_{\text{fail}}(x, t_0, u^*(\cdot\,;\eta^*))$ for $\Delta = 0.9$

Example: Unicycle Model

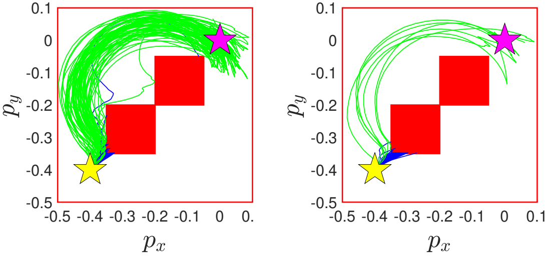
$$\begin{bmatrix} dp_x \\ dp_y \\ ds \\ d\theta \end{bmatrix} = -k \begin{bmatrix} p_x \\ p_y \\ s \\ \theta \end{bmatrix} dt + \begin{bmatrix} s\cos\theta \\ s\sin\theta \\ 0 \\ 0 \end{bmatrix} dt + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} a \\ \omega \end{bmatrix} dt + \begin{bmatrix} \sigma & 0 \\ 0 & \nu \end{bmatrix} dw \right).$$



(a) $\Delta = 0.2$ (b) $\Delta = 0.9$

Example: Car Model

$$\begin{bmatrix} dp_x \\ dp_y \\ ds \\ d\theta \\ d\phi \end{bmatrix} = -k \begin{bmatrix} p_x \\ p_y \\ s \\ 0 \\ 0 \end{bmatrix} dt + \begin{bmatrix} s\cos\theta \\ s\sin\theta \\ 0 \\ \frac{s\tan\phi}{L} \\ 0 \end{bmatrix} dt + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} a \\ \zeta \end{bmatrix} dt + \begin{bmatrix} \sigma & 0 \\ 0 & \nu \end{bmatrix} dw \right).$$



(a) $\Delta = 0.2$ (b) $\Delta = 0.9$

Outline

Summary

- We solved the chance-constrained stochastic optimal control problem without introducing any conservative approximation of the chance constraint.

Summary

- We solved the chance-constrained stochastic optimal control problem without introducing any conservative approximation of the chance constraint.
- We introduced a dual SOC problem and proved that the strong duality exists between the original chance-constrained SOC problem and the dual SOC problem.

Summary

- We solved the chance-constrained stochastic optimal control problem without introducing any conservative approximation of the chance constraint.
- We introduced a dual SOC problem and proved that the strong duality exists between the original chance-constrained SOC problem and the dual SOC problem.
- We derived a novel path-integral-based dual ascent algorithm to numerically solve the dual problem.

Summary

- We solved the chance-constrained stochastic optimal control problem without introducing any conservative approximation of the chance constraint.
- We introduced a dual SOC problem and proved that the strong duality exists between the original chance-constrained SOC problem and the dual SOC problem.
- We derived a novel path-integral-based dual ascent algorithm to numerically solve the dual problem.

We provided an optimal solution the chance-constrained stochastic optimal control problem which can be computed online via Monte-Carlo samples of system trajectories (path integral control).

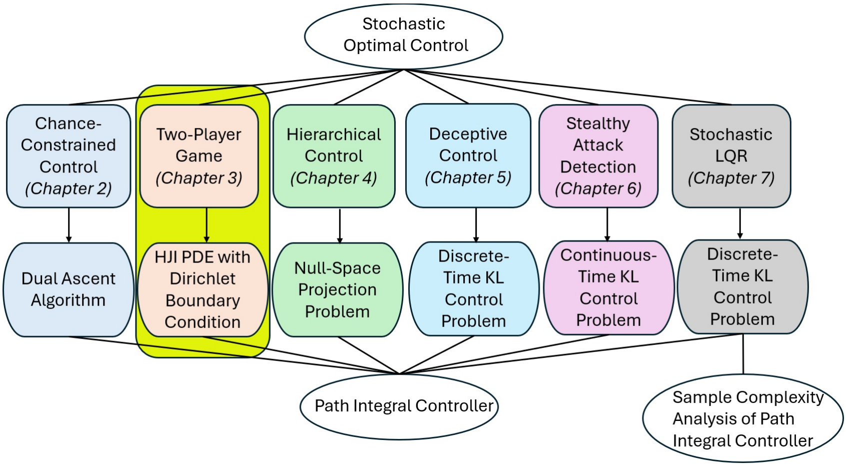
Summary

- We solved the chance-constrained stochastic optimal control problem without introducing any conservative approximation of the chance constraint.
- We introduced a dual SOC problem and proved that the strong duality exists between the original chance-constrained SOC problem and the dual SOC problem.
- We derived a novel path-integral-based dual ascent algorithm to numerically solve the dual problem.

We provided an optimal solution the chance-constrained stochastic optimal control problem which can be computed online via Monte-Carlo samples of system trajectories (path integral control).

- Publications
  - A. Patil, A. Duarte, A. Smith, F. Bisetti, T. Tanaka, "Chance-Constrained Stochastic Optimal Control via Path Integral and Finite Difference Methods," *2022 IEEE Conference on Decision and Control (CDC)*
  - A. Patil, A. Duarte, F. Bisetti, T. Tanaka, "Strong Duality and Dual Ascent Approach to Continuous-Time Chance-Constrained Stochastic Optimal Control," *submitted to Transactions on Automatic Control*

Two-Player Zero-Sum Game

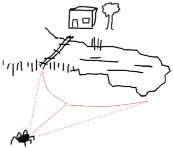


Zero-Sum Game Stochastic Differential Game (SDG)

- Control using an uncertain actuator:

$$dx(t) = f(x(t), t)dt + G(x(t), t)\Big(\underbrace{u(x(t), t)dt + v(x(t), t)dt + dw(t)}_{\text{Uncertain control input}}\Big)$$

- $v(x(t), t)$: Non-stochastic uncertainty: unmodeled bias, fatigue. It is reasonable to assume v is bounded but the control designer should assume the most pessimistic scenario.



- $w(t)$: Stochastic uncertainty
- Control designer wants to minimize $\mathbb{E}_{x_0, t_0}\left[\phi(x(t_f)) + \int_{t_0}^{t_f} \left(\frac{1}{2} u^\top R_u u + V\right) dt\right]$ under the presence of v and w.
- Zero-sum SDG

$$\min_u \max_v \mathbb{E}_{x_0, t_0}\left[\phi(x(t_f)) + \int_{t_0}^{t_f}\left(\frac{1}{2} u^\top R_u u - \frac{1}{2} v^\top R_v v + V\right) dt\right]$$

$$\text{s.t.} \quad dx = fdt + G_u udt + G_v vdt + \Sigma dw.$$

Zero-Sum Game Stochastic Differential Game (SDG)

- Our Contributions:
  - We covert the problem of two-player zero-sum SDG as a problem of solving a Hamilton-Jacobi-Isaacs (HJI) PDE with Dirichlet boundary condition.

Zero-Sum Game Stochastic Differential Game (SDG)

- Our Contributions:
  - We covert the problem of two-player zero-sum SDG as a problem of solving a Hamilton-Jacobi-Isaacs (HJI) PDE with Dirichlet boundary condition.
  - We develop a path-integral framework to solve the HJI PDE and establish the existence and uniqueness of the saddle point solution (optimal solution of the game).

Zero-Sum Game Stochastic Differential Game (SDG)

- Our Contributions:
  - We covert the problem of two-player zero-sum SDG as a problem of solving a Hamilton-Jacobi-Isaacs (HJI) PDE with Dirichlet boundary condition.
  - We develop a path-integral framework to solve the HJI PDE and establish the existence and uniqueness of the saddle point solution (optimal solution of the game).
  - We obtain explicit expressions for the saddle-point policies which can be numerically evaluated using Monte Carlo simulations.

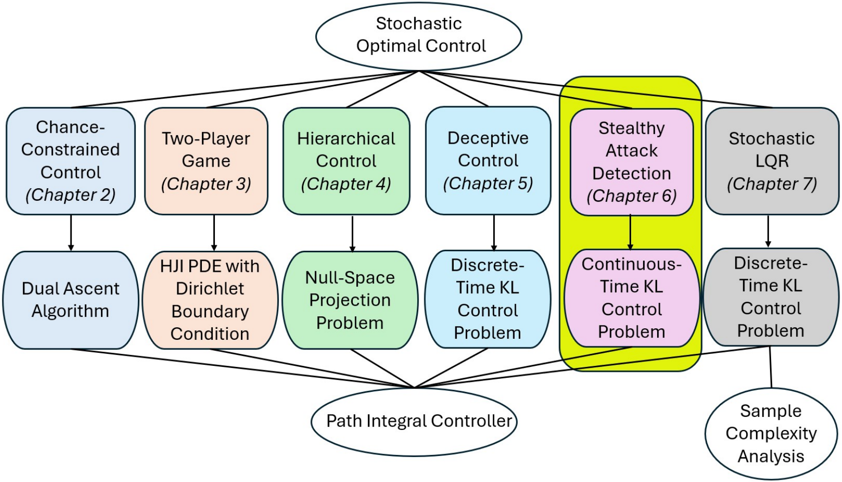
Zero-Sum Game Stochastic Differential Game (SDG)

- Our Contributions:
  - We covert the problem of two-player zero-sum SDG as a problem of solving a Hamilton-Jacobi-Isaacs (HJI) PDE with Dirichlet boundary condition.
  - We develop a path-integral framework to solve the HJI PDE and establish the existence and uniqueness of the saddle point solution (optimal solution of the game).
  - We obtain explicit expressions for the saddle-point policies which can be numerically evaluated using Monte Carlo simulations.
  - Our approach allows the game to be solved online without the need for any offline training or precomputations.

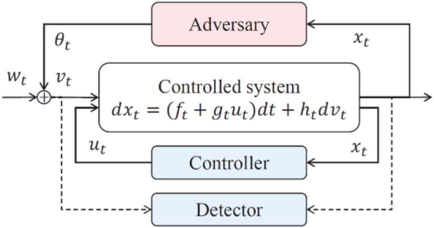
Zero-Sum Game Stochastic Differential Game (SDG)

- Our Contributions:
  - We covert the problem of two-player zero-sum SDG as a problem of solving a Hamilton-Jacobi-Isaacs (HJI) PDE with Dirichlet boundary condition.
  - We develop a path-integral framework to solve the HJI PDE and establish the existence and uniqueness of the saddle point solution (optimal solution of the game).
  - We obtain explicit expressions for the saddle-point policies which can be numerically evaluated using Monte Carlo simulations.
  - Our approach allows the game to be solved online without the need for any offline training or precomputations.
  - Publication
    - A. Patil, Y. Zhou, D. Fridovich-Keil, T. Tanaka, “Risk-Minimizing Two-Player Zero-Sum Stochastic Differential Game via Path Integral Control," *2023 IEEE Conference on Decision and Control (CDC)*

Stealthy Attack Detection (Ongoing Work)



Stealthy Attack Detection (Ongoing Work)



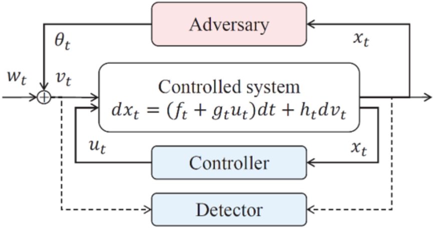
$dv_t = dw_t$ (No Attack)

probability distribution Q

$dv_t = \theta_t dt + dw_t$ (Under Attack)

probability distribution P

Stealthy Attack Detection (Ongoing Work)



$dv_t = dw_t$ (No Attack)

probability distribution Q

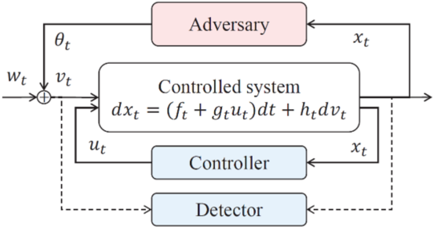
$dv_t = \theta_t dt + dw_t$ (Under Attack)

probability distribution P

- Adversary's problem: KL control problem

$$\max_{\theta} \mathbb{E}^P \int_0^T \ell(x_t, u_t)dt - \lambda \underbrace{D(P \| Q)}_{\text{KL Divergence}} .$$

Stealthy Attack Detection (Ongoing Work)



$dv_t = dw_t$ (No Attack)
probability distribution Q

$dv_t = \theta_t dt + dw_t$ (Under Attack)
probability distribution P

- Adversary's problem: KL control problem

$$\max_{\theta} \mathbb{E}^P \int_0^T \ell(x_t, u_t)dt - \lambda \underbrace{D(P\|Q)}_{\text{KL Divergence}} .$$

- Controller's Problem: Minimax KL control problem:

$$\min_{u} \max_{\theta} \mathbb{E}^P \int_0^T \ell(x_t, u_t)dt - \lambda \underbrace{D(P\|Q)}_{\text{KL Divergence}} .$$

Publications

Journal Publications

- A. Patil, G. Hanasusanto, T. Tanaka, “Discrete-Time Stochastic LQR via Path Integral Control and Its Sample Complexity Analysis," *IEEE Control Systems Letters (L-CSS)*
- A. Patil, A. Duarte, F. Bisetti, T. Tanaka, “Strong Duality and Dual Ascent Approach to Continuous-Time Chance-Constrained Stochastic Optimal Control," *submitted to Transactions on Automatic Control*
- M. Baglioni, A. Patil, L. Sentis, A. Jamshidnejad “Achieving Multi-UAV Best Viewpoint Coordination in Obstructed Environments," *under preparation*

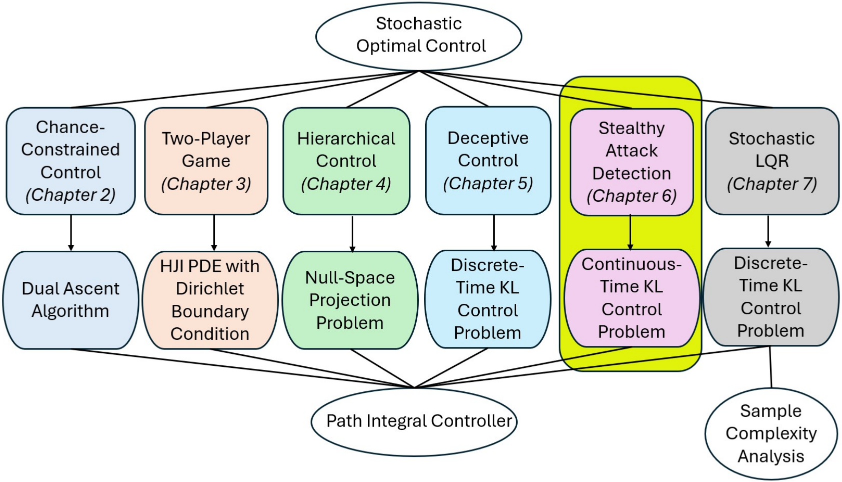
Conference Publications

- A. Patil, G. Hanasusanto, T. Tanaka, “Discrete-Time Stochastic LQR via Path Integral Control and Its Sample Complexity Analysis," *2024 IEEE Conference on Decision and Control (CDC)*
- A. Patil*, M. Karabag*, U. Topcu, T. Tanaka, “Simulation-Driven Deceptive Control via Path Integral Approach," *2023 IEEE Conference on Decision and Control (CDC)*
- A. Patil, Y. Zhou, D. Fridovich-Keil, T. Tanaka, “Risk-Minimizing Two-Player Zero-Sum Stochastic Differential Game via Path Integral Control," *2023 IEEE Conference on Decision and Control (CDC)*
- A. Patil, T. Tanaka, “Upper and Lower Bounds for End-to-End Risks in Stochastic Robot Navigation," *2023 IFAC World Congress*
- A. Patil, A. Duarte, A. Smith, F. Bisetti, T. Tanaka, “Chance-Constrained Stochastic Optimal Control via Path Integral and Finite Difference Methods," *2022 IEEE Conference on Decision and Control (CDC)*
- A. Patil, T. Tanaka, “Upper Bounds for Continuous-Time End-to-End Risks in Stochastic Robot Navigation," *2022 European Control Conference (ECC)*
- A. Patil, R. Funada, T. Tanaka, L. Sentis, “Task Hierarchical Control via Null-Space Projection and Path Integral Approach," *submitted to 2024 American Control Conference (ACC)*
- M. Baglioni, A. Patil, L. Sentis, A. Jamshidnejad “Achieving Multi-UAV Best Viewpoint Coordination in Obstructed Environments," *submitted to 2024 American Control Conference (ACC)*

Coursework

1.	Linear Systems Analysis	A
2.	Modeling of Physical Systems	A
3.	Optimal Control Theory	A
4.	Machine Learning	A
5.	Multivariable Control Systems	A
6.	Verification/Synthesis of Cyberphysical System	A
7.	Introduction to Optimization	A
8.	Statistical Estimation Theory	CR
9.	Reinforcement Learning	A
10.	Application Programming for Engineers	A-
11.	Stochastic Processes I	A

Remaining Work



Timeline

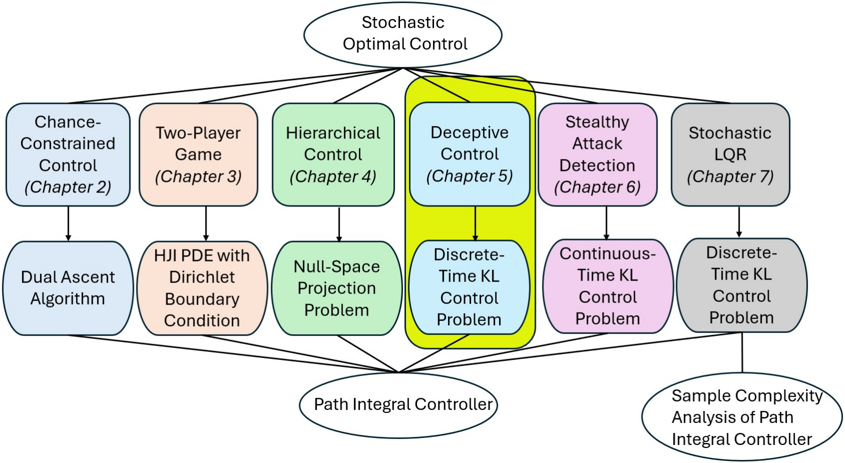
Chapter 6: Detection and Risk Mitigation of Stealthy Attack: Continuous-Time KL Control Problem

- Task 6.1: Worst-Case Attack Synthesis
- Task 6.2: Attack Mitigation
- Task 6.3: Experimental Results

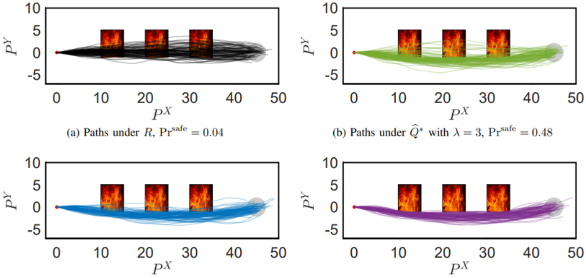
Dissertation	Dec'24	Jan'25	Feb'25	Mar'25	Apr'25	May'25
Chapter 1	Writing					
Chapter 2		Writing				
Chapter 3					Writing	
Chapter 4					Writing	
Chapter 5					Writing	
Chapter 6	Task 6.1	Task 6.2	Task 6.3	Task 6.3		Writing
Chapter 7						Writing

Table: 6-Month Timeline of Dissertation Completion

Deceptive Control



Optimal Deception by Path Integral Control



(a) Paths under R, $\Pr^{\text{safe}} = 0.04$

(b) Paths under $\widehat{Q}^*$ with $\lambda = 3$, $\Pr^{\text{safe}} = 0.48$

(c) Paths under $\widehat{Q}^*$ with $\lambda = 2$, $\Pr^{\text{safe}} = 0.62$

(d) Paths under $\widehat{Q}^*$ with $\lambda = 0.5$, $\Pr^{\text{safe}} = 0.94$

- Problem Setup
  - A supervisor wants an agent to reach the target as soon as possible (reference policy)
  - The agent, on the other hand, wishes to avoid the regions covered under fire (deviated policy)
  - How can the agent satisfy their own interest by deviating from the reference policy without being detected by the supervisor?

Our Contributions

- We formalize the synthesis of an optimal deceptive policy as a KL control problem. We introduce KL divergence as a stealthiness measure using motivations from hypothesis testing theory.

$$\min_{Q} \mathbb{E}_Q \sum_{t=0}^{T} C_t(X_t, U_t) + \lambda D(Q||R)$$

where R is the reference policy and Q is the deviated policy.

Our Contributions

- We formalize the synthesis of an optimal deceptive policy as a KL control problem. We introduce KL divergence as a stealthiness measure using motivations from hypothesis testing theory.

$$\min_{Q} \mathbb{E}_Q \sum_{t=0}^{T} C_t(X_t, U_t) + \lambda D(Q||R)$$

  where R is the reference policy and Q is the deviated policy.

- We solve the KL control problem using backward dynamic programming. Since dynamic programming suffers from the curse of dimensionality, we develop an algorithm based on path integral control to numerically compute the optimal deceptive actions online using Monte Carlo simulations without explicitly synthesizing the policy.

Our Contributions

- We formalize the synthesis of an optimal deceptive policy as a KL control problem. We introduce KL divergence as a stealthiness measure using motivations from hypothesis testing theory.

$$\min_{Q} \mathbb{E}_Q \sum_{t=0}^{T} C_t(X_t, U_t) + \lambda D(Q||R)$$

  where R is the reference policy and Q is the deviated policy.
- We solve the KL control problem using backward dynamic programming. Since dynamic programming suffers from the curse of dimensionality, we develop an algorithm based on path integral control to numerically compute the optimal deceptive actions online using Monte Carlo simulations without explicitly synthesizing the policy.
- We show that our proposed algorithm asymptotically converges to the optimal action distribution of the deceptive agent as the number of samples goes to infinity.

Our Contributions

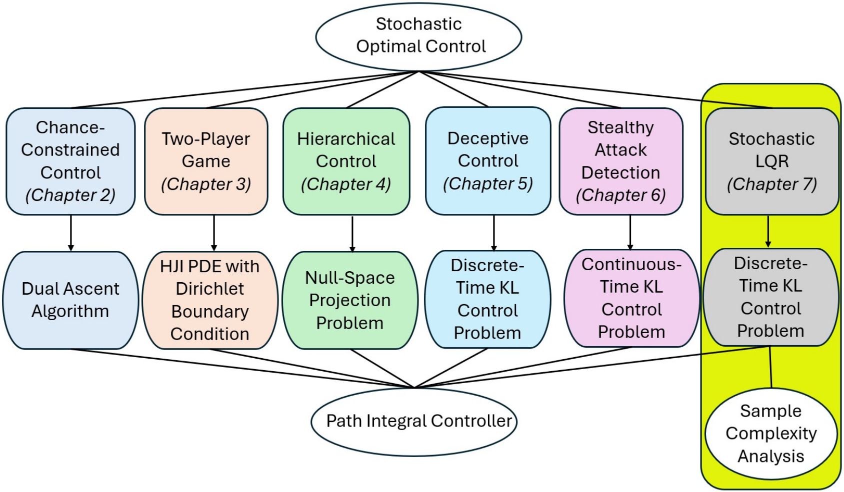
- We formalize the synthesis of an optimal deceptive policy as a KL control problem. We introduce KL divergence as a stealthiness measure using motivations from hypothesis testing theory.

$$\min_{Q} \mathbb{E}_Q \sum_{t=0}^{T} C_t(X_t, U_t) + \lambda D(Q||R)$$

  where R is the reference policy and Q is the deviated policy.

- We solve the KL control problem using backward dynamic programming. Since dynamic programming suffers from the curse of dimensionality, we develop an algorithm based on path integral control to numerically compute the optimal deceptive actions online using Monte Carlo simulations without explicitly synthesizing the policy.
- We show that our proposed algorithm asymptotically converges to the optimal action distribution of the deceptive agent as the number of samples goes to infinity.
- Publication:
  A. Patil*, M. Karabag*, U. Topcu, T. Tanaka, "Simulation-Driven Deceptive Control via Path Integral Approach," *2023 IEEE Conference on Decision and Control (CDC)*

Sample Complexity of Path Integral Approach



Sample Complexity of Path Integral Approach

- Stochastic LQR

$$\min_{\{k_t(\cdot)\}_{t=0}^{T-1}} \mathbb{E} \sum_{t=0}^{T-1} \left(\frac{1}{2} x_t^\top M_t x_t + \frac{1}{2} u_t^\top N_t u_t \right) + \mathbb{E} \left(\frac{1}{2} x_T^\top M_T x_T \right)$$

$$\text{s.t.} \quad x_{t+1} = A_t x_t + B_t u_t + w_t, \quad x_0 = x_0.$$

Sample Complexity of Path Integral Approach

- Stochastic LQR

$$\min_{\{k_t(\cdot)\}_{t=0}^{T-1}} \mathbb{E} \sum_{t=0}^{T-1} \left(\frac{1}{2} x_t^\top M_t x_t + \frac{1}{2} u_t^\top N_t u_t \right) + \mathbb{E} \left(\frac{1}{2} x_T^\top M_T x_T \right)$$
$$\text{s.t.} \quad x_{t+1} = A_t x_t + B_t u_t + w_t, \quad x_0 = x_0.$$

- Optimal policy by solving backward Riccati Recursion

$$u_t = k_t(x_t) = K_t x_t, \quad K_t = -(B_t^\top \Theta_{t+1} B_t + N_t)^{-1} B_t^\top \Theta_{t+1} A_t$$

Sample Complexity of Path Integral Approach

- Stochastic LQR

$$\min_{\{k_t(\cdot)\}_{t=0}^{T-1}} \mathbb{E} \sum_{t=0}^{T-1} \left(\frac{1}{2} x_t^\top M_t x_t + \frac{1}{2} u_t^\top N_t u_t \right) + \mathbb{E} \left(\frac{1}{2} x_T^\top M_T x_T \right)$$

$$\text{s.t.} \quad x_{t+1} = A_t x_t + B_t u_t + w_t, \quad x_0 = x_0.$$

- Optimal policy by solving backward Riccati Recursion

$$u_t = k_t(x_t) = K_t x_t, \quad K_t = -(B_t^\top \Theta_{t+1} B_t + N_t)^{-1} B_t^\top \Theta_{t+1} A_t$$

- We derived a path integral controller to solve stochastic LQR:

Sample Complexity of Path Integral Approach

- Stochastic LQR

$$\min_{\{k_t(\cdot)\}_{t=0}^{T-1}} \mathbb{E}\sum_{t=0}^{T-1}\left(\frac{1}{2}x_t^\top M_t x_t + \frac{1}{2}u_t^\top N_t u_t\right) + \mathbb{E}\left(\frac{1}{2}x_T^\top M_T x_T\right)$$
$$\text{s.t.}\quad x_{t+1} = A_t x_t + B_t u_t + w_t, \quad x_0 = x_0.$$

- Optimal policy by solving backward Riccati Recursion

$$u_t = k_t(x_t) = K_t x_t, \quad K_t = -(B_t^\top \Theta_{t+1} B_t + N_t)^{-1} B_t^\top \Theta_{t+1} A_t$$

- We derived a path integral controller to solve stochastic LQR:
  - At every time-step t, sample n_t trajectories $\{x_{t:T}(i), u_{t:T-1}(i)\}_{i=1}^{n_t}$ from the "uncontrolled" dynamics: $x_{t+1} = A_t x_t + w_t$

Sample Complexity of Path Integral Approach

- Stochastic LQR

$$\min_{\{k_t(\cdot)\}_{t=0}^{T-1}} \mathbb{E} \sum_{t=0}^{T-1} \left(\frac{1}{2} x_t^\top M_t x_t + \frac{1}{2} u_t^\top N_t u_t \right) + \mathbb{E} \left(\frac{1}{2} x_T^\top M_T x_T \right)$$

$$\text{s.t.} \quad x_{t+1} = A_t x_t + B_t u_t + w_t, \quad x_0 = x_0.$$

- Optimal policy by solving backward Riccati Recursion

$$u_t = k_t(x_t) = K_t x_t, \quad K_t = -(B_t^\top \Theta_{t+1} B_t + N_t)^{-1} B_t^\top \Theta_{t+1} A_t$$

- We derived a path integral controller to solve stochastic LQR:
  - At every time-step t, sample n_t trajectories $\{x_{t:T}(i), u_{t:T-1}(i)\}_{i=1}^{n_t}$ from the "uncontrolled" dynamics: $x_{t+1} = A_t x_t + w_t$
  - Compute path cost of each sample path i:

$$r(i) = \exp\left(-\frac{1}{\lambda} \sum_{k=t}^{T} \frac{1}{2} x_k(i)^\top M_k x_k(i) \right)$$

Sample Complexity of Path Integral Approach

- Stochastic LQR

$$\min_{\{k_t(\cdot)\}_{t=0}^{T-1}} \mathbb{E}\sum_{t=0}^{T-1}\left(\frac{1}{2}x_t^\top M_t x_t + \frac{1}{2}u_t^\top N_t u_t\right) + \mathbb{E}\left(\frac{1}{2}x_T^\top M_T x_T\right)$$

$$\text{s.t.} \quad x_{t+1} = A_t x_t + B_t u_t + w_t, \quad x_0 = x_0.$$

- Optimal policy by solving backward Riccati Recursion

$$u_t = k_t(x_t) = K_t x_t, \quad K_t = -(B_t^\top \Theta_{t+1} B_t + N_t)^{-1} B_t^\top \Theta_{t+1} A_t$$

- We derived a path integral controller to solve stochastic LQR:
  - At every time-step t, sample n_t trajectories $\{x_{t:T}(i), u_{t:T-1}(i)\}_{i=1}^{n_t}$ from the "uncontrolled" dynamics: $x_{t+1} = A_t x_t + w_t$
  - Compute path cost of each sample path i:

$$r(i) = \exp\left(-\frac{1}{\lambda}\sum_{k=t}^{T} \frac{1}{2} x_k(i)^\top M_k x_k(i)\right)$$

  - Path integral LQR controller:

$$\hat{u}_t = \sum_{i=1}^{n_t} \frac{r(i)}{\sum_{i=1}^{n_t} r(i)} u_t(i).$$

Sample Complexity of Path Integral Approach

- Stochastic LQR

$$\min_{\{k_t(\cdot)\}_{t=0}^{T-1}} \mathbb{E} \sum_{t=0}^{T-1} \left(\frac{1}{2} x_t^\top M_t x_t + \frac{1}{2} u_t^\top N_t u_t \right) + \mathbb{E} \left(\frac{1}{2} x_T^\top M_T x_T \right)$$

$$\text{s.t.} \quad x_{t+1} = A_t x_t + B_t u_t + w_t, \quad x_0 = x_0.$$

- Optimal policy by solving backward Riccati Recursion

$$u_t = k_t(x_t) = K_t x_t, \quad K_t = -(B_t^\top \Theta_{t+1} B_t + N_t)^{-1} B_t^\top \Theta_{t+1} A_t$$

- We derived a path integral controller to solve stochastic LQR:
  - At every time-step t, sample n_t trajectories $\{x_{t:T}(i), u_{t:T-1}(i)\}_{i=1}^{n_t}$ from the "uncontrolled" dynamics: $x_{t+1} = A_t x_t + w_t$
  - Compute path cost of each sample path i:

$$r(i) = \exp\left(-\frac{1}{\lambda} \sum_{k=t}^{T} \frac{1}{2} x_k(i)^\top M_k x_k(i) \right)$$

  - Path integral LQR controller:

$$\hat{u}_t = \sum_{i=1}^{n_t} \frac{r(i)}{\sum_{i=1}^{n_t} r(i)} u_t(i).$$

  - Does not require solving backward Riccati equation

Sample Complexity of Path Integral Approach

- Define the empirical means $\hat{E}$ and true expectations E as

$$\hat{E}_t^{ru} = \frac{\sum_{i=1}^{n_t} r(i)u_t(i)}{n_t}, \quad E_t^{ru} = \lim_{n_t \to \infty} \frac{\sum_{i=1}^{n_t} r(i)u_t(i)}{n_t}, \quad \hat{E}_t^{r} = \frac{\sum_{i=1}^{n_t} r(i)}{n_t}, \quad E_t^{r} = \lim_{n_t \to \infty} \frac{\sum_{i=1}^{n_t} r(i)}{n_t}.$$

Sample Complexity of Path Integral Approach

- Define the empirical means $\hat{E}$ and true expectations E as

$$\hat{E}_t^{ru} = \frac{\sum_{i=1}^{n_t} r(i)u_t(i)}{n_t}, \quad E_t^{ru} = \lim_{n_t\to\infty} \frac{\sum_{i=1}^{n_t} r(i)u_t(i)}{n_t}, \quad \hat{E}_t^{r} = \frac{\sum_{i=1}^{n_t} r(i)}{n_t}, \quad E_t^{r} = \lim_{n_t\to\infty} \frac{\sum_{i=1}^{n_t} r(i)}{n_t}.$$

- Theorem: Let $\{\epsilon_t\}_{t=0}^{T-1}$, $\{\alpha_t\}_{t=0}^{T-1}$ and $\{\beta_t\}_{t=0}^{T-1}$ be given sequences of positive numbers and $\epsilon := \sum_{t=0}^{T-1} \epsilon_t^2$, $\alpha := \sum_{t=0}^{T-1} \alpha_t$, $\beta := \sum_{t=0}^{T-1} \beta_t$. Suppose $\alpha + \beta < 1$.

Sample Complexity of Path Integral Approach

- Define the empirical means $\hat{E}$ and true expectations E as

$$\hat{E}_t^{ru} = \frac{\sum_{i=1}^{n_t} r(i)u_t(i)}{n_t}, \quad E_t^{ru} = \lim_{n_t\to\infty} \frac{\sum_{i=1}^{n_t} r(i)u_t(i)}{n_t}, \quad \hat{E}_t^{r} = \frac{\sum_{i=1}^{n_t} r(i)}{n_t}, \quad E_t^{r} = \lim_{n_t\to\infty} \frac{\sum_{i=1}^{n_t} r(i)}{n_t}.$$

- Theorem: Let $\{\epsilon_t\}_{t=0}^{T-1}$, $\{\alpha_t\}_{t=0}^{T-1}$ and $\{\beta_t\}_{t=0}^{T-1}$ be given sequences of positive numbers and $\epsilon := \sum_{t=0}^{T-1} \epsilon_t^2$, $\alpha := \sum_{t=0}^{T-1} \alpha_t$, $\beta := \sum_{t=0}^{T-1} \beta_t$. Suppose $\alpha + \beta < 1$. If n_t satisfies

$$n_t \geq \frac{\left(\hat{E}_t^r\sqrt{2\|\hat{\Omega}_t\|\log\frac{2m}{\beta_t}} + \left(\epsilon_t\hat{E}_t^r + \|\hat{E}_t^{ru}\|_\infty\right)\sqrt{\frac{1}{2}\log\frac{2}{\alpha_t}}\right)^2}{\epsilon_t^2(\hat{E}_t^r)^4}$$

and $\hat{E}_t^r > \sqrt{\frac{1}{2n_t}\log\frac{2}{\alpha_t}}$,

Sample Complexity of Path Integral Approach

- Define the empirical means $\hat{E}$ and true expectations E as

$$\hat{E}_t^{ru} = \frac{\sum_{i=1}^{n_t} r(i) u_t(i)}{n_t}, \quad E_t^{ru} = \lim_{n_t \to \infty} \frac{\sum_{i=1}^{n_t} r(i) u_t(i)}{n_t}, \quad \hat{E}_t^{r} = \frac{\sum_{i=1}^{n_t} r(i)}{n_t}, \quad E_t^{r} = \lim_{n_t \to \infty} \frac{\sum_{i=1}^{n_t} r(i)}{n_t}.$$

- Theorem: Let $\{\epsilon_t\}_{t=0}^{T-1}$, $\{\alpha_t\}_{t=0}^{T-1}$ and $\{\beta_t\}_{t=0}^{T-1}$ be given sequences of positive numbers and $\epsilon := \sum_{t=0}^{T-1} \epsilon_t^2$, $\alpha := \sum_{t=0}^{T-1} \alpha_t$, $\beta := \sum_{t=0}^{T-1} \beta_t$. Suppose $\alpha + \beta < 1$. If n_t satisfies

$$n_t \geq \frac{\left(\hat{E}_t^r \sqrt{2 \|\hat{\Omega}_t\| \log \frac{2m}{\beta_t}} + \left(\epsilon_t \hat{E}_t^r + \|\hat{E}_t^{ru}\|_\infty \right) \sqrt{\frac{1}{2} \log \frac{2}{\alpha_t}} \right)^2}{\epsilon_t^2 (\hat{E}_t^r)^4}$$

and $\hat{E}_t^r > \sqrt{\frac{1}{2n_t} \log \frac{2}{\alpha_t}}$, then $\|\hat{u} - u\|_\infty^2 := \sum_{t=0}^{T-1} \|\hat{u}_t - u_t\|_\infty^2 \leq \epsilon$ with probability greater than or equal to $1 - \alpha - \beta$.

Sample Complexity of Path Integral Approach

- Define the empirical means $\hat{E}$ and true expectations E as

$$\hat{E}_t^{ru} = \frac{\sum_{i=1}^{n_t} r(i)u_t(i)}{n_t}, \quad E_t^{ru} = \lim_{n_t\to\infty} \frac{\sum_{i=1}^{n_t} r(i)u_t(i)}{n_t}, \quad \hat{E}_t^{r} = \frac{\sum_{i=1}^{n_t} r(i)}{n_t}, \quad E_t^{r} = \lim_{n_t\to\infty} \frac{\sum_{i=1}^{n_t} r(i)}{n_t}.$$

- Theorem: Let $\{\epsilon_t\}_{t=0}^{T-1}$, $\{\alpha_t\}_{t=0}^{T-1}$ and $\{\beta_t\}_{t=0}^{T-1}$ be given sequences of positive numbers and $\epsilon := \sum_{t=0}^{T-1} \epsilon_t^2$, $\alpha := \sum_{t=0}^{T-1} \alpha_t$, $\beta := \sum_{t=0}^{T-1} \beta_t$. Suppose $\alpha + \beta < 1$. If n_t satisfies

$$n_t \geq \frac{\left(\hat{E}_t^r \sqrt{2\|\hat{\Omega}_t\| \log \frac{2m}{\beta_t}} + \left(\epsilon_t \hat{E}_t^r + \|\hat{E}_t^{ru}\|_\infty\right)\sqrt{\frac{1}{2}\log\frac{2}{\alpha_t}}\right)^2}{\epsilon_t^2 (\hat{E}_t^r)^4}$$

  and $\hat{E}_t^r > \sqrt{\frac{1}{2n_t}\log\frac{2}{\alpha_t}}$, then $\|\hat{u} - u\|_\infty^2 := \sum_{t=0}^{T-1} \|\hat{u}_t - u_t\|_\infty^2 \leq \epsilon$ with probability greater than or equal to $1 - \alpha - \beta$.
- The required number of samples depends only logarithmically on the dimension of the control input m.

Sample Complexity of Path Integral Approach

- Define the empirical means $\hat{E}$ and true expectations E as

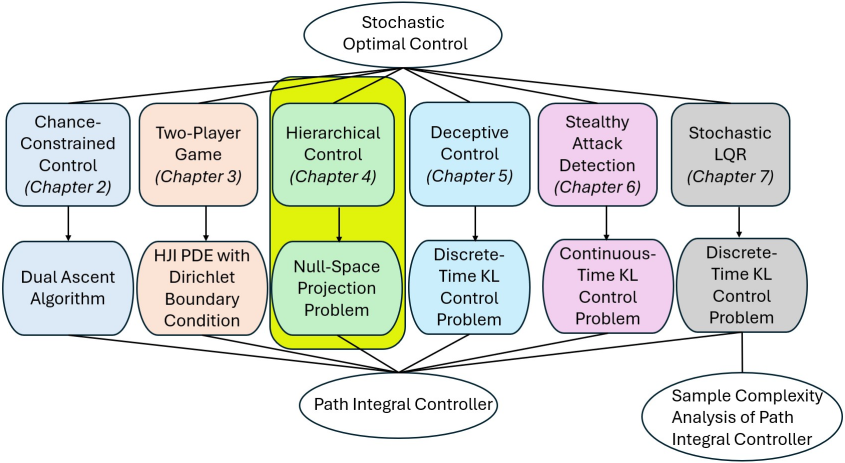
$$\hat{E}_t^{ru} = \frac{\sum_{i=1}^{n_t} r(i) u_t(i)}{n_t}, \quad E_t^{ru} = \lim_{n_t \to \infty} \frac{\sum_{i=1}^{n_t} r(i) u_t(i)}{n_t}, \quad \hat{E}_t^{r} = \frac{\sum_{i=1}^{n_t} r(i)}{n_t}, \quad E_t^{r} = \lim_{n_t \to \infty} \frac{\sum_{i=1}^{n_t} r(i)}{n_t}.$$

- Theorem: Let $\{\epsilon_t\}_{t=0}^{T-1}$, $\{\alpha_t\}_{t=0}^{T-1}$ and $\{\beta_t\}_{t=0}^{T-1}$ be given sequences of positive numbers and $\epsilon := \sum_{t=0}^{T-1} \epsilon_t^2$, $\alpha := \sum_{t=0}^{T-1} \alpha_t$, $\beta := \sum_{t=0}^{T-1} \beta_t$. Suppose $\alpha + \beta < 1$. If n_t satisfies

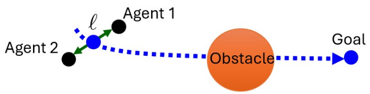
$$n_t \geq \frac{\left(\hat{E}_t^r \sqrt{2 \|\hat{\Omega}_t\| \log \frac{2m}{\beta_t}} + \left(\epsilon_t \hat{E}_t^r + \|\hat{E}_t^{ru}\|_\infty \right) \sqrt{\frac{1}{2} \log \frac{2}{\alpha_t}} \right)^2}{\epsilon_t^2 (\hat{E}_t^r)^4}$$

  and $\hat{E}_t^r > \sqrt{\frac{1}{2n_t} \log \frac{2}{\alpha_t}}$, then $\|\hat{u} - u\|_\infty^2 := \sum_{t=0}^{T-1} \|\hat{u}_t - u_t\|_\infty^2 \leq \epsilon$ with probability greater than or equal to $1 - \alpha - \beta$.
- The required number of samples depends only logarithmically on the dimension of the control input m.
- Publication: A. Patil, G. Hanasusanto, T. Tanaka, “Discrete-Time Stochastic LQR via Path Integral Control and Its Sample Complexity Analysis," *IEEE Control Systems Letters (L-CSS)*

Hierarchical Control

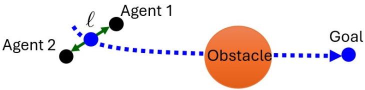


Conventional Task Hierarchical Control

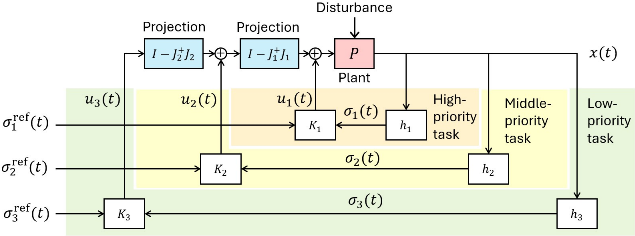


- Task 1: Avoid collisions with obstacles
- Task 2: Steer the platoon's centroid towards a goal position
- Task 3: Maintain specific distances between the agents

Conventional Task Hierarchical Control

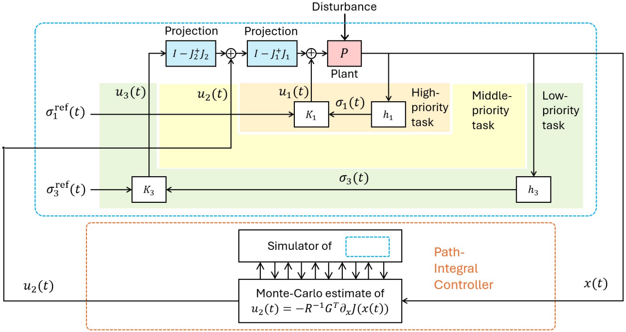


- Task 1: Avoid collisions with obstacles
- Task 2: Steer the platoon's centroid towards a goal position
- Task 3: Maintain specific distances between the agents



- Simple controllers (such as PID) are used for K_i to achieve reference tracking in task coordinate $\sigma_i(t)$
- Reference signals $\sigma_i^{\text{ref}}(t)$ are often chosen manually.

Task Hierarchical Control via Path Integral Method



- Path integral controller seeks the optimal input for some of the tasks, while rudimentary controllers can be kept for other tasks.
- Manuscript:
  A. Patil, R. Funada, T. Tanaka, L. Sentis, “Task Hierarchical Control via Null-Space Projection and Path Integral Approach," *submitted to 2024 American Control Conference (ACC)*

Questions?

apurvapatil@utexas.edu
Google Scholar | Website | LinkedIn | Github