IEEE Conference on Decision and Control 2023

Risk-Minimizing Two-Player Zero-Sum Stochastic Differential Game via Path Integral Control

Apurva Patil   Yujing Zhou
David Fridovich-Keil   Takashi Tanaka

Outline

Outline

Background

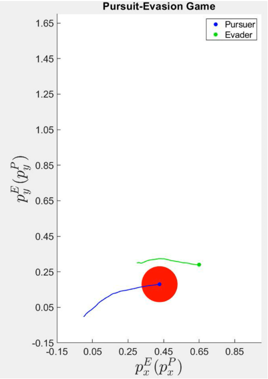
- Risk-minimizing two-player zero-sum stochastic differential game: each player aims to minimize its probability of failure and the control cost.

Background

- Risk-minimizing two-player zero-sum stochastic differential game: each player aims to minimize its probability of failure and the control cost.
- Failure occurs when the state of the game enters into predefined undesirable domains

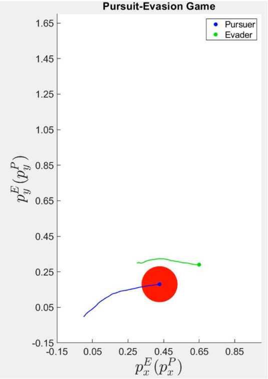
Background

- Risk-minimizing two-player zero-sum stochastic differential game: each player aims to minimize its probability of failure and the control cost.
- Failure occurs when the state of the game enters into predefined undesirable domains



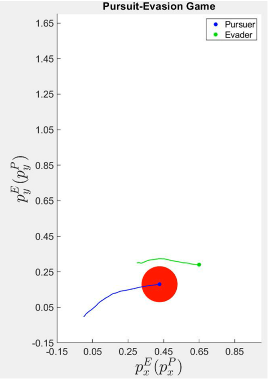
Background

- Risk-minimizing two-player zero-sum stochastic differential game: each player aims to minimize its probability of failure and the control cost.
- Failure occurs when the state of the game enters into predefined undesirable domains
- We solve continuous-time, nonlinear two-player zero-sum stochastic differential game online using path integral control



Background

- Risk-minimizing two-player zero-sum stochastic differential game: each player aims to minimize its probability of failure and the control cost.
- Failure occurs when the state of the game enters into predefined undesirable domains
- We solve continuous-time, nonlinear two-player zero-sum stochastic differential game online using path integral control
- Existence and uniqueness of the saddle-point of the game.



Outline

Existing Approaches

- Hamilton-Jacobi-Isaacs partial differential equation [Falcone et al. 2006]: grid-based approaches to solve the HJI PDEs

Existing Approaches

- Hamilton-Jacobi-Isaacs partial differential equation [Falcone et al. 2006]: grid-based approaches to solve the HJI PDEs; suffer from the curse of dimensionality, precomputation required

Existing Approaches

- Hamilton-Jacobi-Isaacs partial differential equation [Falcone et al. 2006]: grid-based approaches to solve the HJI PDEs; suffer from the curse of dimensionality, precomputation required
- RL-based approaches

Existing Approaches

- Hamilton-Jacobi-Isaacs partial differential equation [Falcone et al. 2006]: grid-based approaches to solve the HJI PDEs; suffer from the curse of dimensionality, precomputation required
- RL-based approaches: adaptive dynamic programming [Vrabie et al. 2011]

Existing Approaches

- Hamilton-Jacobi-Isaacs partial differential equation [Falcone et al. 2006]: grid-based approaches to solve the HJI PDEs; suffer from the curse of dimensionality, precomputation required
- RL-based approaches: adaptive dynamic programming [Vrabie et al. 2011] , deep RL [Prajapat et al. 2021]

Existing Approaches

- Hamilton-Jacobi-Isaacs partial differential equation [Falcone et al. 2006]: grid-based approaches to solve the HJI PDEs; suffer from the curse of dimensionality, precomputation required
- RL-based approaches: adaptive dynamic programming [Vrabie et al. 2011] , deep RL [Prajapat et al. 2021], integral RL [Liu et al. 2020]

Existing Approaches

- Hamilton-Jacobi-Isaacs partial differential equation [Falcone et al. 2006]: grid-based approaches to solve the HJI PDEs; suffer from the curse of dimensionality, precomputation required
- RL-based approaches: adaptive dynamic programming [Vrabie et al. 2011] , deep RL [Prajapat et al. 2021], integral RL [Liu et al. 2020], Bayesian inverse RL [Lin et al. 2017]

Existing Approaches

- Hamilton-Jacobi-Isaacs partial differential equation [Falcone et al. 2006]: grid-based approaches to solve the HJI PDEs; suffer from the curse of dimensionality, precomputation required
- RL-based approaches: adaptive dynamic programming [Vrabie et al. 2011] , deep RL [Prajapat et al. 2021], integral RL [Liu et al. 2020], Bayesian inverse RL [Lin et al. 2017]
  Challenges in learning-based approaches:
  - Rigorous theoretical guarantees on convergence and optimality
  - Offline training required

Existing Approaches

- Hamilton-Jacobi-Isaacs partial differential equation [Falcone et al. 2006]: grid-based approaches to solve the HJI PDEs; suffer from the curse of dimensionality, precomputation required
- RL-based approaches: adaptive dynamic programming [Vrabie et al. 2011] , deep RL [Prajapat et al. 2021], integral RL [Liu et al. 2020], Bayesian inverse RL [Lin et al. 2017]
  Challenges in learning-based approaches:
  - Rigorous theoretical guarantees on convergence and optimality
  - Offline training required
- Above approaches do not explicitly take into account the players' failure probabilities

Outline

Problem Formulation

- System: Itô stochastic differential equation (SDE)

$$
\begin{aligned}
d\boldsymbol{x}(t) = & f\left(\boldsymbol{x}(t), t\right) dt + G_u\left(\boldsymbol{x}(t), t\right) u(\boldsymbol{x}(t), t) dt \\
& + G_v\left(\boldsymbol{x}(t), t\right) v\left(\boldsymbol{x}(t), t\right) dt + \Sigma\left(\boldsymbol{x}(t), t\right) d\boldsymbol{w}(t)
\end{aligned}
$$

Problem Formulation

- System: Itô stochastic differential equation (SDE)

$$
\begin{aligned}
d\boldsymbol{x}(t) =& f\left(\boldsymbol{x}(t), t\right) dt + G_u\left(\boldsymbol{x}(t), t\right) u(\boldsymbol{x}(t), t) dt \\
&+ G_v\left(\boldsymbol{x}(t), t\right) v\left(\boldsymbol{x}(t), t\right) dt + \Sigma\left(\boldsymbol{x}(t), t\right) d\boldsymbol{w}(t)
\end{aligned}
$$

- Safe region: $\mathcal{X}_s$, boundary $\partial\mathcal{X}_s$

Problem Formulation

- System: Itô stochastic differential equation (SDE)

$$\begin{aligned} d\boldsymbol{x}(t) =& f\left(\boldsymbol{x}(t), t\right) dt + G_u\left(\boldsymbol{x}(t), t\right) u(\boldsymbol{x}(t), t) dt \\ &+ G_v\left(\boldsymbol{x}(t), t\right) v\left(\boldsymbol{x}(t), t\right) dt + \Sigma\left(\boldsymbol{x}(t), t\right) d\boldsymbol{w}(t) \end{aligned}$$

- Safe region: $\mathcal{X}_s$, boundary $\partial\mathcal{X}_s$
- Agent's probability of failure:

$$P_{\mathrm{fail}}^{\mathrm{ag}} := P_{x_0, t_0}\left(\bigvee_{t \in (t_0, T]} \boldsymbol{x}(t) \notin \mathcal{X}_s \right)$$

- Adversary's probability of failure $P_{\mathrm{fail}}^{\mathrm{ad}} := 1 - P_{\mathrm{fail}}^{\mathrm{ag}}$

Problem Formulation

- Define a set $\mathcal{Q} = \mathcal{X}_s \times [t_0, T)$,
  Boundary $\partial\mathcal{Q} = (\partial\mathcal{X}_s \times [t_0, T]) \cup (\mathcal{X}_s \times \{T\})$

Problem Formulation

- Define a set $\mathcal{Q} = \mathcal{X}_s \times [t_0, T)$,
  Boundary $\partial\mathcal{Q} = (\partial\mathcal{X}_s \times [t_0, T]) \cup (\mathcal{X}_s \times \{T\})$
- Terminal time of the game: $\boldsymbol{t_f} := \inf\{t > t_0 : (\boldsymbol{x}(t), t) \notin \mathcal{Q}\}$

Problem Formulation

- Define a set $\mathcal{Q} = \mathcal{X}_s \times [t_0, T)$,
  Boundary $\partial\mathcal{Q} = (\partial\mathcal{X}_s \times [t_0, T]) \cup (\mathcal{X}_s \times \{T\})$
- Terminal time of the game: $\boldsymbol{t}_f := \inf\{t > t_0 : (\boldsymbol{x}(t), t) \notin \mathcal{Q}\}$
- Agent's failure probability $P_{\text{fail}}^{\text{ag}}$:

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

Problem Formulation

- Define a set $\mathcal{Q} = \mathcal{X}_s \times [t_0, T)$,
  Boundary $\partial\mathcal{Q} = (\partial\mathcal{X}_s \times [t_0, T]) \cup (\mathcal{X}_s \times \{T\})$
- Terminal time of the game: $\boldsymbol{t_f} := \inf\{t > t_0 : (\boldsymbol{x}(t), t) \notin \mathcal{Q}\}$
- Agent's failure probability $P^{\mathrm{ag}}_{\mathrm{fail}}$:

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

- Risk-minimizing cost function:

$$C(x_0, t_0; u, v) := \eta\,\mathbb{E}_{x_0,t_0}\left[\mathbb{1}_{\boldsymbol{x}(\boldsymbol{t_f})\in\partial\mathcal{X}_s}\right] + \mathbb{E}_{x_0,t_0}\left[\psi(\boldsymbol{x}(\boldsymbol{t_f}))\cdot\mathbb{1}_{\boldsymbol{x}(\boldsymbol{t_f})\in\mathcal{X}_s} + \int_{t_0}^{\boldsymbol{t_f}} L(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{v}(t), t)\,dt\right].$$

Problem Formulation

- Define: $\phi(x) := \psi(x) \cdot \mathbb{1}_{x \in \mathcal{X}_s} + \eta \cdot \mathbb{1}_{x \in \partial \mathcal{X}_s}$
- Running cost:

$$L(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, t) = V(\boldsymbol{x}, t) + \frac{1}{2} \boldsymbol{u}^T R_u(\boldsymbol{x}, t)\, \boldsymbol{u} - \frac{1}{2} \boldsymbol{v}^T R_v(\boldsymbol{x}, t)\, \boldsymbol{v}$$

- Risk-minimizing zero-sum SDG

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

$$\text{s.t.} \quad d\boldsymbol{x} = f dt + G_u u dt + G_v v dt + \Sigma d\boldsymbol{w},$$

$$\boldsymbol{x}(t_0) = x_0.$$

Outline

HJI PDE with Dirichlet Boundary Condition

- Cost-to-go function:

$$C\left(x, t; u, v\right) = \mathbb{E}_{x,t}\left[\phi\left(\boldsymbol{x}(\boldsymbol{t}_f)\right)\right] + \mathbb{E}_{x,t}\left[\int_t^{\boldsymbol{t}_f}\left(\frac{1}{2}\boldsymbol{u}^\top R_u \boldsymbol{u} - \frac{1}{2}\boldsymbol{v}^\top R_v \boldsymbol{v} + V\right) dt\right].$$

HJI PDE with Dirichlet Boundary Condition

- Cost-to-go function:

$$C\left(x,t;u,v\right) = \mathbb{E}_{x,t}\left[\phi\left(\boldsymbol{x}(\boldsymbol{t}_f)\right)\right] + \mathbb{E}_{x,t}\left[\int_t^{\boldsymbol{t}_f}\left(\frac{1}{2}\boldsymbol{u}^\top R_u\boldsymbol{u} - \frac{1}{2}\boldsymbol{v}^\top R_v\boldsymbol{v} + V\right)dt\right].$$

- (u^*, v^*) constitutes a *saddle-point solution* if

$$C(x,t;u^*,v) \le C^* := C(x,t;u^*,v^*) \le C(x,t;u,v^*).$$

where the value of the game

$$C^* = \min_u \max_v C\left(x,t;u,v\right) = \max_v \min_u C\left(x,t;u,v\right).$$

HJI PDE with Dirichlet Boundary Condition

Theorem: Suppose there exists a function $J : \overline{\mathcal{Q}} \to \mathbb{R}$ such that

(a) $J(x,t)$ is continuously differentiable in t and twice continuously differentiable in x in the domain $\mathcal{Q}$;

(b) $J(x,t)$ solves the following stochastic Hamilton-Jacobi-Isaacs (HJI) PDE:

$$\begin{cases} -\partial_t J = V + f^\top \partial_x J + \frac{1}{2}\mathrm{Tr}\left(\Sigma\Sigma^\top \partial_x^2 J\right) \\ \quad + \frac{1}{2}(\partial_x J)^\top \left(G_v R_v^{-1} G_v^\top - G_u R_u^{-1} G_u^\top\right)\partial_x J, & \forall (x,t) \in \mathcal{Q}, \\ \lim_{\substack{(x,t)\to(y,s) \\ (x,t)\in\mathcal{Q}}} J(x,t) = \phi(y), & \forall (y,s) \in \partial\mathcal{Q}. \end{cases} \tag{1}$$

...

HJI PDE with Dirichlet Boundary Condition

...Then, the following statements hold:

(i) $J(x, t)$ is the value of the game, i.e.

$$\begin{aligned} J(x,t) &= \min_{u} \max_{v} C(x,t;u,v) \\ &= \max_{v} \min_{u} C(x,t;u,v), \quad \forall (x,t) \in \overline{\mathcal{Q}}. \end{aligned}$$

HJI PDE with Dirichlet Boundary Condition

...Then, the following statements hold:

(i) $J(x, t)$ is the value of the game, i.e.

$$J\left(x, t\right) = \min_{u} \max_{v} C\left(x, t; u, v\right) = \max_{v} \min_{u} C\left(x, t; u, v\right), \quad \forall \left(x, t\right) \in \overline{\mathcal{Q}}.$$

(ii) The optimal solution is given by

$$u^*(x, t) = -R_u^{-1}(x, t)\, {G_u}^\top(x, t)\, \partial_x J(x, t)\,,$$

$$v^*(x, t) = R_v^{-1}(x, t)\, {G_v}^\top(x, t)\, \partial_x J(x, t)\,.$$

Outline

Path Integral Formulation

- Logarithmic transformation of the value function:

$$J(x,t) = -\lambda \log\left(\xi\left(x,t\right)\right)$$

Path Integral Formulation

- Logarithmic transformation of the value function:

$$J(x,t) = -\lambda \log\left(\xi\left(x,t\right)\right)$$

- **Assumption:** For all $(x,t) \in \overline{\mathcal{Q}}$, there exists a constant $\lambda > 0$ such that

$$\begin{aligned}\Sigma(x,t)\Sigma^\top(x,t) = & \lambda G_u(x,t)R_u^{-1}(x,t)G_u^\top(x,t) \\ & - \lambda G_v(x,t)R_v^{-1}(x,t)G_v^\top(x,t).\end{aligned}$$

Path Integral Formulation

- Logarithmic transformation of the value function:

$$J(x,t) = -\lambda \log\left(\xi\left(x,t\right)\right)$$

- **Assumption:** For all $(x,t) \in \overline{\mathcal{Q}}$, there exists a constant $\lambda > 0$ such that

$$\begin{aligned}\Sigma(x,t)\Sigma^\top(x,t) = &\lambda G_u(x,t)R_u^{-1}(x,t)G_u^\top(x,t)\\ &- \lambda G_v(x,t)R_v^{-1}(x,t)G_v^\top(x,t).\end{aligned}$$

Partition of the system dynamics:

$$\begin{bmatrix} d\boldsymbol{x}^{(1)} \\ d\boldsymbol{x}^{(2)} \end{bmatrix} = \begin{bmatrix} f^{(1)}(\boldsymbol{x},t) \\ f^{(2)}(\boldsymbol{x},t) \end{bmatrix} dt + \begin{bmatrix} 0 \\ G_u^{(2)}(\boldsymbol{x},t) \end{bmatrix} u(\boldsymbol{x},t)dt + \begin{bmatrix} 0 \\ G_v^{(2)}(\boldsymbol{x},t) \end{bmatrix} v(\boldsymbol{x},t)dt + \begin{bmatrix} 0 \\ \Sigma^{(2)}(\boldsymbol{x},t) \end{bmatrix} d\boldsymbol{w}$$

Path Integral Formulation

- Linear PDE in ξ with Dirichlet boundary condition

$$\begin{cases} \partial_t \xi = \frac{V\xi}{\lambda} - f^\top \partial_x \xi - \frac{1}{2}\mathrm{Tr}\left(\Sigma\Sigma^\top \partial_x^2 \xi\right), & \forall (x,t) \in \mathcal{Q}, \\ \lim\limits_{\substack{(x,t)\to(y,s) \\ (x,t)\in\mathcal{Q}}} \xi(x,t) = \exp\left(-\frac{\phi(y)}{\lambda}\right), & \forall (y,s) \in \partial\mathcal{Q}. \end{cases}$$

[1] Friedman, "Stochastic differential equations and applications, vol. 1."

Path Integral Formulation

- Linear PDE in ξ with Dirichlet boundary condition

$$\begin{cases} \partial_t \xi = \frac{V\xi}{\lambda} - f^\top \partial_x \xi - \frac{1}{2}\mathrm{Tr}\left(\Sigma\Sigma^\top \partial_x^2 \xi\right), & \forall (x,t) \in \mathcal{Q}, \\ \lim\limits_{\substack{(x,t)\to(y,s) \\ (x,t)\in\mathcal{Q}}} \xi(x,t) = \exp\left(-\frac{\phi(y)}{\lambda}\right), & \forall (y,s) \in \partial\mathcal{Q}. \end{cases}$$

- The solution of the above PDE exists and is unique[1].

[1] Friedman, "Stochastic differential equations and applications, vol. 1."

Path Integral Formulation

- Linear PDE in ξ with Dirichlet boundary condition

$$\begin{cases} \partial_t \xi = \frac{V\xi}{\lambda} - f^\top \partial_x \xi - \frac{1}{2}\mathrm{Tr}\left(\Sigma\Sigma^\top \partial_x^2 \xi\right), & \forall (x,t) \in \mathcal{Q}, \\ \lim_{\substack{(x,t)\to(y,s) \\ (x,t)\in\mathcal{Q}}} \xi(x,t) = \exp\left(-\frac{\phi(y)}{\lambda}\right), & \forall (y,s) \in \partial\mathcal{Q}. \end{cases}$$

- The solution of the above PDE exists and is unique[1].
- The solution admits the Feynman-Kac representation.

[1] Friedman, "Stochastic differential equations and applications, vol. 1."

Feynman-Kac Represenation

- Uncontrolled state dynamics:

$$d\hat{\boldsymbol{x}}(t) = f(\hat{\boldsymbol{x}}(t), t)dt + \Sigma(\hat{\boldsymbol{x}}(t), t)d\boldsymbol{w}(t)$$

- Let $\hat{\boldsymbol{t}}_f := \inf\{t > t_0 : (\hat{\boldsymbol{x}}(t), t) \notin \mathcal{Q}\}$
- The solution of the linearized PDE (path integral form):

$$\xi(x, t) = \mathbb{E}_{x,t}\left[\exp\left(-\frac{1}{\lambda} S(\tau)\right)\right]$$

where

$$S(\tau) = \phi\left(\hat{\boldsymbol{x}}(\hat{\boldsymbol{t}}_f)\right) + \int_t^{\hat{\boldsymbol{t}}_f} V(\hat{\boldsymbol{x}}(t), t)\, dt.$$

Path Integral Formulation

Theorem: A saddle-point solution exists, is unique, and is given by

$$u^*(x,t)dt = \mathcal{G}_u(x,t)\frac{\mathbb{E}_{x,t}\left[\exp\left(-\frac{1}{\lambda}S\left(\tau\right)\right)\Sigma^{(2)}(x,t)\,d\boldsymbol{w}\right]}{\mathbb{E}_{x,t}\left[\exp\left(-\frac{1}{\lambda}S\left(\tau\right)\right)\right]},$$

where

$$\mathcal{G}_u = R_u^{-1}G_u^{(2)\top}\left(G_u^{(2)}R_u^{-1}G_u^{(2)\top} - G_v^{(2)}R_v^{-1}G_v^{(2)\top}\right)^{-1}$$

Path Integral Formulation

Theorem: A saddle-point solution exists, is unique, and is given by

$$u^*(x,t)dt = \mathcal{G}_u(x,t)\frac{\mathbb{E}_{x,t}\left[\exp\left(-\frac{1}{\lambda}S(\tau)\right)\Sigma^{(2)}(x,t)\,d\boldsymbol{w}\right]}{\mathbb{E}_{x,t}\left[\exp\left(-\frac{1}{\lambda}S(\tau)\right)\right]},$$

where

$$\mathcal{G}_u = R_u^{-1}G_u^{(2)\top}\left(G_u^{(2)}R_u^{-1}G_u^{(2)\top} - G_v^{(2)}R_v^{-1}G_v^{(2)\top}\right)^{-1}$$

and

$$v^*(x,t)dt = \mathcal{G}_v(x,t)\frac{\mathbb{E}_{x,t}\left[\exp\left(-\frac{1}{\lambda}S(\tau)\right)\Sigma^{(2)}(x,t)\,d\boldsymbol{w}\right]}{\mathbb{E}_{x,t}\left[\exp\left(-\frac{1}{\lambda}S(\tau)\right)\right]},$$

where

$$\mathcal{G}_v = -R_v^{-1}G_v^{(2)\top}\left(G_u^{(2)}R_u^{-1}G_u^{(2)\top} - G_v^{(2)}R_v^{-1}G_v^{(2)\top}\right)^{-1}.$$

Outline

Disturbance Attenuation

- Unicycle model:

$$\begin{bmatrix} d\boldsymbol{p}_x \\ d\boldsymbol{p}_y \\ d\boldsymbol{s} \\ d\boldsymbol{\theta} \end{bmatrix} = -k \begin{bmatrix} \boldsymbol{p}_x \\ \boldsymbol{p}_y \\ \boldsymbol{s} \\ \boldsymbol{\theta} \end{bmatrix} dt + \begin{bmatrix} \boldsymbol{s}\cos\boldsymbol{\theta} \\ \boldsymbol{s}\sin\boldsymbol{\theta} \\ 0 \\ 0 \end{bmatrix} dt + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \left(\underbrace{\begin{bmatrix} a \\ \omega \end{bmatrix}}_{u} dt + \underbrace{\begin{bmatrix} \Delta a \\ \Delta\omega \end{bmatrix}}_{v} dt + \begin{bmatrix} \sigma & 0 \\ 0 & \nu \end{bmatrix} d\boldsymbol{w} \right),$$

Disturbance Attenuation

- Unicycle model:

$$\begin{bmatrix} d\boldsymbol{p}_x \\ d\boldsymbol{p}_y \\ d\boldsymbol{s} \\ d\boldsymbol{\theta} \end{bmatrix} = -k \begin{bmatrix} \boldsymbol{p}_x \\ \boldsymbol{p}_y \\ \boldsymbol{s} \\ \boldsymbol{\theta} \end{bmatrix} dt + \begin{bmatrix} \boldsymbol{s}\cos\boldsymbol{\theta} \\ \boldsymbol{s}\sin\boldsymbol{\theta} \\ 0 \\ 0 \end{bmatrix} dt + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \left(\underbrace{\begin{bmatrix} a \\ \omega \end{bmatrix}}_{u} dt + \underbrace{\begin{bmatrix} \Delta a \\ \Delta\omega \end{bmatrix}}_{v} dt + \begin{bmatrix} \sigma & 0 \\ 0 & \nu \end{bmatrix} d\boldsymbol{w} \right),$$

- Problem to solve:

$$\min_u \max_v \mathbb{E}_{x_0, t_0} \left[\phi(\boldsymbol{x}(\boldsymbol{t}_f)) + \int_{t_0}^{\boldsymbol{t}_f} \left(\frac{1}{2}\boldsymbol{u}^\top \boldsymbol{u} - \frac{\gamma^2}{2} \boldsymbol{v}^T \boldsymbol{v} + V \right) dt \right].$$

Disturbance Attenuation

- Assumption: $\lambda > 0$ such that

$$\lambda\left(1 - \frac{1}{\gamma^2}\right) = 1.$$

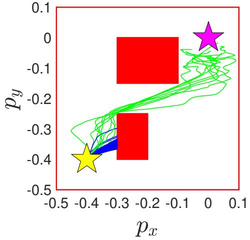
- The problem admits a unique saddle-point solution if $\gamma > 1$

Disturbance Attenuation

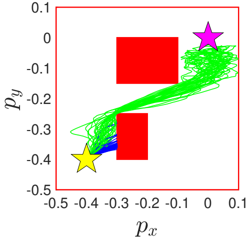
- Assumption: $\lambda > 0$ such that

$$\lambda\left(1 - \frac{1}{\gamma^2}\right) = 1.$$

- The problem admits a unique saddle-point solution if $\gamma > 1$



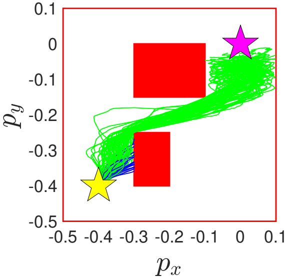
$\gamma^2 = 2$, $P_{\text{fail}}^{\text{ag}} = 0.9$



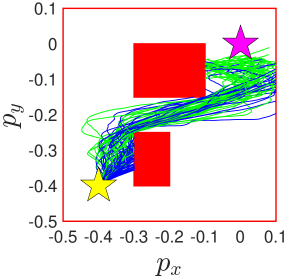
$\gamma^2 = 7$, $P_{\text{fail}}^{\text{ag}} = 0.64$

Disturbance Attenuation

- MC trajectories: 10^4, Step size: 0.01



Agent is aware of the adversary, $P_{\text{fail}}^{\text{ag}} = 0.23$



Agent is not aware of the adversary, $P_{\text{fail}}^{\text{ag}} = 0.65$

Pursuit-Evasion Game

- Pursuer and evader models:

$$d\boldsymbol{p}_x^E = u_x dt + \sigma_x^E d\boldsymbol{w}_x^E, \qquad d\boldsymbol{p}_x^P = v_x dt + \sigma_x^P d\boldsymbol{w}_x^P,$$
$$d\boldsymbol{p}_y^E = u_y dt + \sigma_y^E d\boldsymbol{w}_y^E, \qquad d\boldsymbol{p}_y^P = v_y dt + \sigma_y^P d\boldsymbol{w}_y^P,$$

Pursuit-Evasion Game

- Pursuer and evader models:

$$
\begin{aligned}
d\boldsymbol{p}_x^E &= u_x dt + \sigma_x^E d\boldsymbol{w}_x^E, \qquad & d\boldsymbol{p}_x^P &= v_x dt + \sigma_x^P d\boldsymbol{w}_x^P, \\
d\boldsymbol{p}_y^E &= u_y dt + \sigma_y^E d\boldsymbol{w}_y^E, \qquad & d\boldsymbol{p}_y^P &= v_y dt + \sigma_y^P d\boldsymbol{w}_y^P,
\end{aligned}
$$

- In terms of relative position $\boldsymbol{p}_x = \boldsymbol{p}_x^E - \boldsymbol{p}_x^P, \quad \boldsymbol{p}_y = \boldsymbol{p}_y^E - \boldsymbol{p}_y^P$:

$$
d\boldsymbol{x} = \begin{bmatrix} d\boldsymbol{p}_x \\ d\boldsymbol{p}_y \end{bmatrix} = \begin{bmatrix} u_x \\ u_y \end{bmatrix} dt - \begin{bmatrix} v_x \\ v_y \end{bmatrix} dt + \begin{bmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{bmatrix} d\boldsymbol{w}
$$

$\sigma_x = \sqrt{(\sigma_x^E)^2 + (\sigma_x^P)^2}, \sigma_y = \sqrt{(\sigma_y^E)^2 + (\sigma_y^P)^2}$

Pursuit-Evasion Game

- Pursuer and evader models:

$$\begin{aligned} d\boldsymbol{p}_x^E &= u_x dt + \sigma_x^E d\boldsymbol{w}_x^E, \qquad & d\boldsymbol{p}_x^P &= v_x dt + \sigma_x^P d\boldsymbol{w}_x^P, \\ d\boldsymbol{p}_y^E &= u_y dt + \sigma_y^E d\boldsymbol{w}_y^E, \qquad & d\boldsymbol{p}_y^P &= v_y dt + \sigma_y^P d\boldsymbol{w}_y^P, \end{aligned}$$

- In terms of relative position $\boldsymbol{p}_x = \boldsymbol{p}_x^E - \boldsymbol{p}_x^P, \quad \boldsymbol{p}_y = \boldsymbol{p}_y^E - \boldsymbol{p}_y^P$:

$$d\boldsymbol{x} = \begin{bmatrix} d\boldsymbol{p}_x \\ d\boldsymbol{p}_y \end{bmatrix} = \begin{bmatrix} u_x \\ u_y \end{bmatrix} dt - \begin{bmatrix} v_x \\ v_y \end{bmatrix} dt + \begin{bmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{bmatrix} d\boldsymbol{w}$$

$\sigma_x = \sqrt{(\sigma_x^E)^2 + (\sigma_x^P)^2}, \sigma_y = \sqrt{(\sigma_y^E)^2 + (\sigma_y^P)^2}$

- Origin coincides with the pursuer's position and safe set $\mathcal{X}_s := \{x \in \mathbb{R}^2 : \|x\| > \rho\}$

Pursuit Evasion Game

- Problem to solve:

$$\min_{u} \max_{v} \mathbb{E}_{x_0, t_0}\left[\phi\left(\boldsymbol{x}(\boldsymbol{t}_f)\right) + \int_{t_0}^{\boldsymbol{t}_f}\left(\frac{1}{2}\boldsymbol{u}^T\boldsymbol{u} - \frac{{r_v}^2}{2}\boldsymbol{v}^T\boldsymbol{v} + V\right) dt\right].$$

Pursuit Evasion Game

- Problem to solve:

$$\min_{u} \max_{v} \mathbb{E}_{x_0, t_0}\left[\phi\left(\boldsymbol{x}(\boldsymbol{t}_f)\right) + \int_{t_0}^{\boldsymbol{t}_f} \left(\frac{1}{2}\boldsymbol{u}^T\boldsymbol{u} - \frac{{r_v}^2}{2}\boldsymbol{v}^T\boldsymbol{v} + V\right) dt\right].$$

- Assumption: $\lambda > 0$ such that

$$\lambda\left(1 - \frac{1}{{r_v}^2}\right) = 1.$$

- The problem admits a unique saddle-point solution if $r_v > 1$

Pursuit Evasion Game

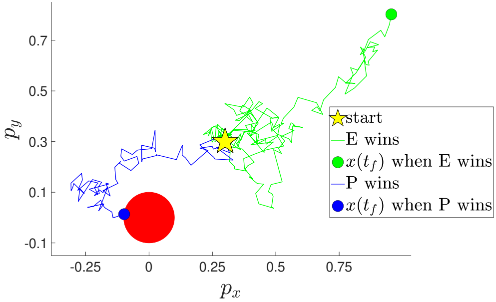


Figure: The red disc of radius $\rho = 0.1$, centered at the origin represents that the pursuer is within the distance ρ of the evader.

Pursuit Evasion Game

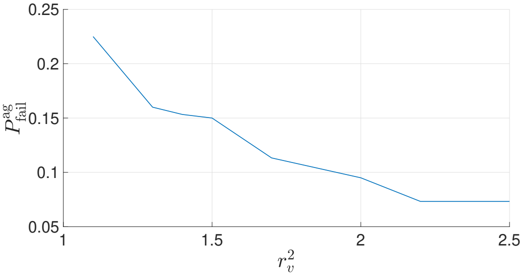


Figure: Failure probabilities of the agent (i.e., evader) as a function of r_v, when the players follow the saddle-point policies (u^*, v^*).

Outline

Background

Existing Approaches

Problem Formulation

HJI PDE with Dirichlet Boundary Condition

Path Integral Formulation

Simualtions

Conclusion

Conclusion

- Presented an HJI-PDE-based solution approach for a risk-minimizing two-player zero-sum stochastic differential game (SDG). Each player tries to balance the trade-off between the probability of failure and the control cost.

Conclusion

- Presented an HJI-PDE-based solution approach for a risk-minimizing two-player zero-sum stochastic differential game (SDG). Each player tries to balance the trade-off between the probability of failure and the control cost.
- Developed a path-integral framework and established the existence and uniqueness of saddle point solution under linearized HJI PDE.

Conclusion

- Presented an HJI-PDE-based solution approach for a risk-minimizing two-player zero-sum stochastic differential game (SDG). Each player tries to balance the trade-off between the probability of failure and the control cost.
- Developed a path-integral framework and established the existence and uniqueness of saddle point solution under linearized HJI PDE.
- The presented approach allows the game to be solved online without the need for any offline training or precomputations.

Conclusion

- Presented an HJI-PDE-based solution approach for a risk-minimizing two-player zero-sum stochastic differential game (SDG). Each player tries to balance the trade-off between the probability of failure and the control cost.
- Developed a path-integral framework and established the existence and uniqueness of saddle point solution under linearized HJI PDE.
- The presented approach allows the game to be solved online without the need for any offline training or precomputations.
- Future work:
  - chance-constrained stochastic games in which each player would aim to satisfy a hard bound on its failure probability.