Invariance for safe learning in constrained control

Claus Danielson, Ankush Chakrabarty, Stefano Di Cairano

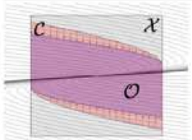
Motivation

- Goal: design control systems that ensure safety and achieve high performance in changing environments
- Combine model-based and data-based design:
  - Safety from physics-based models
  - Performance from data-based models
- Leverage:
  - Physics-based models to describe the boundary system behaviors for parameter range
  - Data-based models to achieve the desired behavior for the actual system
- How to guarantee system safety during exploration and learning transients?
  - Learn within invariantly safe sets

Invariant Sets and Learning

MERL works on many applications that operate in dynamic environments with restrictive constraints
- Physical limitations
- Performance specifications
- Safety requirements

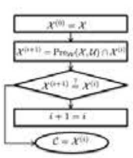
Use learning to adapt to dynamic environment while ensuring safety by enforcing constraints



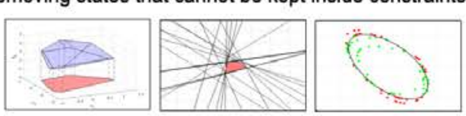
Positive Invariant Sets
$\forall x \in \mathcal{O} \Rightarrow f(x, \kappa(x)) \in \mathcal{O}$

Robust Control Invariant Sets
$\forall x \in \mathcal{C} \subseteq \mathcal{X},\ \exists u \in \mathcal{U} \Rightarrow f(x,u,\theta) \in \mathcal{C}\ \forall \theta \in \Theta$

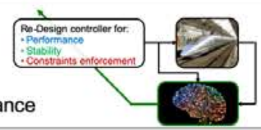
Admissible input set
$\mathcal{C}_u(x) = \{u \in \mathcal{U} : f(x,u,\theta) \in \mathcal{C}, \forall \theta \in \Theta\}$



Maximal control invariant set computed by iteratively removing states that cannot be kept inside constraints



Use learning to adapt controller to improve performance while maintaining guarantees on stability and constraint enforcement by invariance

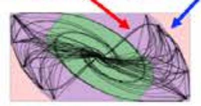


Separation Principle for Learning MPC

- Goal: design constrained MPC that reconfigure to learned model using any learning algorithm
- Leverage: RCI to determine safe control, projection of learned model onto admissible models

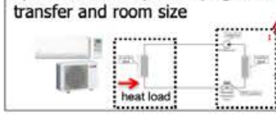
$$\min_U \ F(x_{N|t}, \theta_{N|t}) + \sum_{k=0}^{N-1} L(x_{k|t}, u_{k|t})$$
$$\text{s.t.} \quad x_{k+1|t} = f(x_{k|t}, u_{k|t}, \theta_{k|t})$$
$$\text{RCI:} \ u_{k|t} \in \mathcal{C}_u(x_{k|t}) \subseteq \mathcal{X}_{\text{safe}}(x_{k|t}, \theta), \forall \theta \in \Theta$$

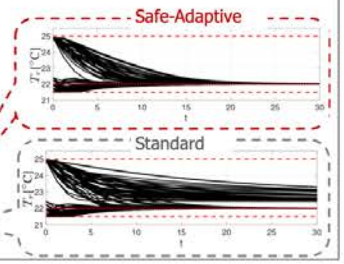
Projection & update rules
$$\theta(t) = \Pi(\xi(t), \Theta)$$
$$\theta_{\cdot|t} = \vartheta(\theta_{\cdot|t-1}, \theta(t))$$



VRF - HVAC control

Temperature tracking and operational safety in varying heat transfer and room size





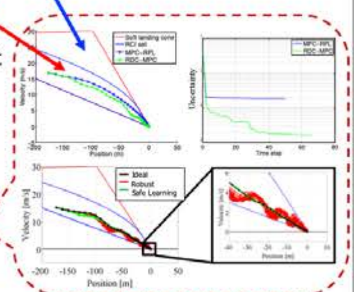
Safe Trade-off for Exploration-Exploitation

- Goal: design constrained MPC that balances exploration and exploitation of model parameters
- Leverage: RCI to determine safe control, dual objective control for optimal excitation

$$\min_U \ F(x_{N|t}) + \sum_{k=0}^{N-1} L(x_{k|t}, u_{k|t}) + \xi(x(t), x_{1|t-1})\psi(U, x(t))$$
$$\text{s.t.} \quad x_{k+1|t} = f(x_{k|t}, u_{k|t}, \theta_{k|t})$$
$$\text{RCI:} \ u_{k|t} \in \mathcal{C}_u(x_{k|t}) \subseteq \mathcal{X}_{\text{safe}}(x_{k|t}, \theta), \forall \theta \in \Theta$$

Train Automatic Stop Control

Precise stopping under different loads and track conditions



Constrained Approximate Dynamic Programming

Motivation: Ensure stability and safety via constraint satisfaction in model-free reinforcement learning in continuous state/action spaces.

Objective:
- Learn optimal control policy on-line from data for high performance operation
- Ensure system stability
- Guarantee safety via state and input constraint satisfaction

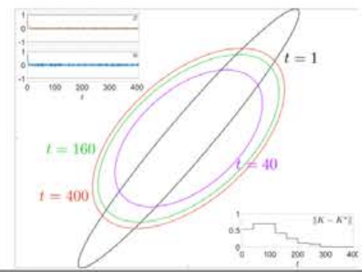
$$x_{t+1} = f(x_t, u_t)$$
$$u_t = \kappa_t(x_t)$$
$$V_t(x_t, u_t) = \sum_{k=t}^{\infty} L(x_{k|t}, u_{k|t})$$
$$x_t \in \mathcal{X}_{\text{safe}}, \ u_t \in \mathcal{U}_{\text{safe}}$$

Constrained policy evaluation:
- Learn sequence of CAIS*
  - Optimize performance
  - Ensure stability
  - Satisfy constraints

$$\begin{pmatrix} P_{t+1} \\ \rho_{t+1} \end{pmatrix} := \arg\min_{\rho>0, P(\cdot)\succ 0} \frac{1}{2} \sum_{t=0}^{t_i+1-1} J_t$$
$$\text{s.t.:} \quad P(f(x_t, \kappa_t(x_t))) \le \lambda P(x_t)$$
$$P(x_t) \le \rho$$
$$\{x : P(x) \le \rho\} \subset \mathcal{X}_{\text{safe}}$$

Constrained policy improvement:
- For each CAIS, learn optimal control policy
- Verify state and input constraint satisfaction through backtracking

Illustrative example:



*CAIS: constraint-admissible invariant sets

References

A. Chakrabarty, R. Quirynen, C. Danielson, W. Gao. *Approximate Dynamic Programming For Linear Systems with State and Input Constraints*, to appear, ECC 2019.

A. Chakrabarty, A. Raghunathan, S. Di Cairano, C. Danielson. *Data-Driven Estimation of Reachable and Invariant Sets for Unmodeled Systems via Active Learning*, IEEE CDC 2018.

A. Chakrabarty, D. Jha, Y. Wang. *Data-Driven Control Policies for Partially Known Systems Via Kernelized Lipschitz Learning*, to appear, ACC 2019.

Di Cairano S. *Indirect adaptive model predictive control for linear systems with polytopic uncertainty*, ACC 2016

Zhou, J., Di Cairano, S., Danielson, C., *Indirect adaptive MPC for output tracking of uncertain linear polytopic systems*, ACC 2017

Cheng, Y., Haghighat, S., Di Cairano, S., *Robust dual control MPC with application to soft-landing control*, ACC 2015.

Berntorp, K., *Recursive Bayesian Inference and Learning of Gaussian-Process State-Space Models*, to appear, ECC 2019.

Berntorp, K., *Bayesian Tire-Friction Learning by Gaussian-Process State-Space Models*, to appear, ECC 2019.

Di Cairano, S., Danielson, C., *Robust Soft-Landing Control with Quantized Input*, NOLCOS 2016

Okamoto, K., Berntorp, K., Di Cairano, S., *Similarity-based Automotive Trajectory Prediction*, ACC 2017

Berntorp, K., Quirynen, R., Di Cairano, S., *Steering of Autonomous vehicles based on Friction-Adaptive nonlinear model predictive control*

Acknowledgements: Karl Berntorp, Rien Quirynen, Yebin Wang, Avishai Weiss (MERL)