EigenSafe *A Spectral Framework for Learning-Based Stochastic Safety Filtering*

Inkyu Jang[1], Jonghae Park[1], Chams E. Mballo[2], Sihyun Cho[1], Claire J. Tomlin[2], and H. Jin Kim[1]

[1] Seoul National University, [2] University of California, Berkeley

Summary

- **EigenSafe** is a framework based on linear operator theory that correctly captures the evolution of safety probability of stochastic systems.
- The proposed safety Q function is directly tied to the safety probability and can be learned from trajectory data.

Dynamic Programming for Safety Probability

$$\mathbb{P}_\pi[\text{safety}(x,u,t+1)] = \mathbb{E}_\pi[\mathbb{P}_\pi[\text{safety}(x^+,u^+,t)]]$$

$$\mathbb{P}_\pi[\text{safety}(x,u,0)] = 1_{\text{safe}}(x,u)$$

$$\mathbb{P}_\pi[\text{safety}(x,u,t)] = A_\pi \circ \cdots \circ A_\pi 1_{\text{safe}}(x,u) = A_\pi^t 1_{\text{safe}}(x,u)$$

where $A_\pi \beta(x,u) := \begin{cases} \mathbb{E}_\pi[\beta(x^+,u^+)] & (x,u)\ \text{safe} \\ 0 & (x,u)\ \text{unsafe} \end{cases}$

- A_π is a **linear** operator.
- The spectral properties of A_π describe the long-term safety of the closed-loop system.

1. The superscript $(\cdot)^+$ denotes the state/action at the next time step.
2. The subscript $(\cdot)_\pi$ on $\mathbb{P}$ or $\mathbb{E}$ means that the probability/expectation was evaluated with respect to the trajectory induced by the feedback policy π.
3. safety(x,u,t) means that the system remains within the safe region **throughout the time window** $[0,t]$, given initial state x and initial input u.
4. 1_{safe} is the one-step safety indicator function returning 1 if the state-action pair is safe and 0 otherwise.

The Dominant Eigenfunction as Safety Q function

$$\mathbb{P}_\pi[\text{safety}(x,u,t)] = A_\pi^t 1_{\text{safe}}(x,u) \approx c \cdot \psi_\pi(x,u) \cdot \gamma_\pi^t$$

The dominant eigenfunction ψ_π

- Measures safety of each state-action pair (x,u)
- Always nonnegative

The dominant eigenvalue γ_π

- Safety of the overall closed-loop system
- Always between 0 and 1

Learning

1. Eigenpair Learning

$$(\gamma_\pi, \psi_\pi) = \arg\min_{\lambda,\psi} W_\lambda \mathbb{E}_{(x,u,x^+)\sim\mathcal{D},u^+\sim\pi}\left[\left(\psi(x^+,u^+) - \lambda\psi(x,u)\right)^2\right] + W_n(\|\psi\| - 1)^2$$

Eigenpair Loss
Encourages (γ_π, ψ_π) to converge to the dominant eigenpair

Normalization Loss
Prevents ψ from collapsing to zero

* $W_{(\cdot)}$-s are a positive weights. $\|\psi\|$ is an estimate of any function norm of ψ from data (e.g., the supremum norm $\sup_{(x,u)\sim\mathcal{D}} |\psi(x,u)|$).

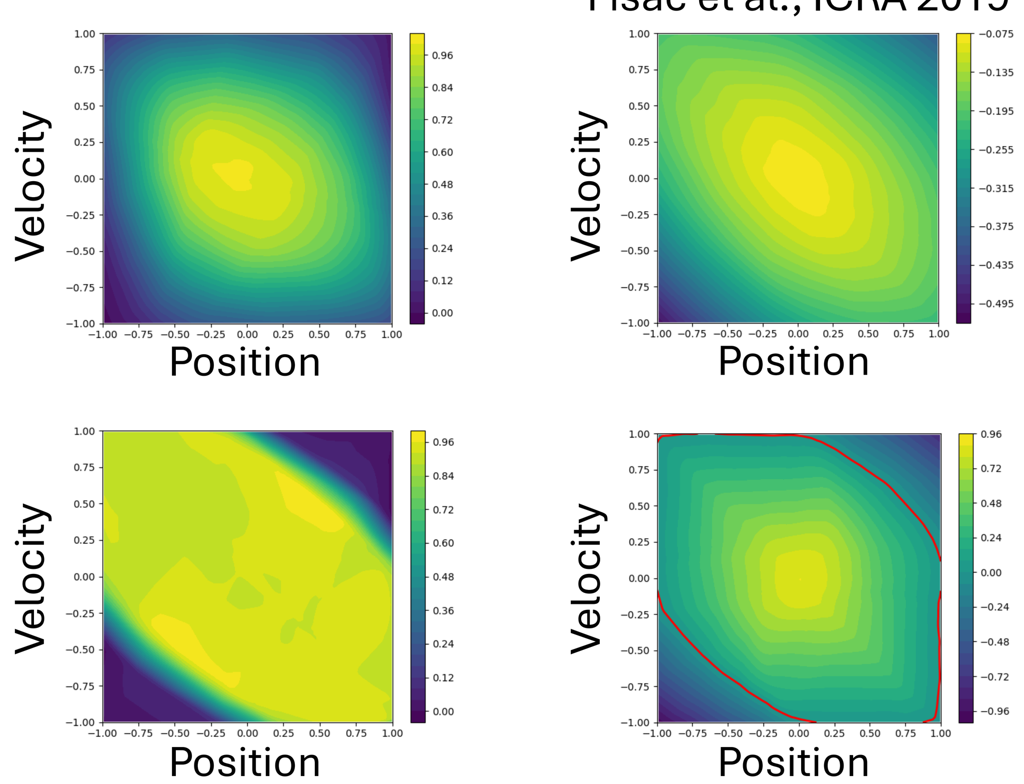
2. Backup Policy Optimization (DDPG Style)

$$\pi = \arg\max_\pi \mathbb{E}_{(x,\cdot,\cdot)\sim\mathcal{D}}\, \psi(x,\pi(x))$$

- Solve for the policy π that **maximizes safety Q function value**.
- This will give the **safest backup policy possible** for the given system.

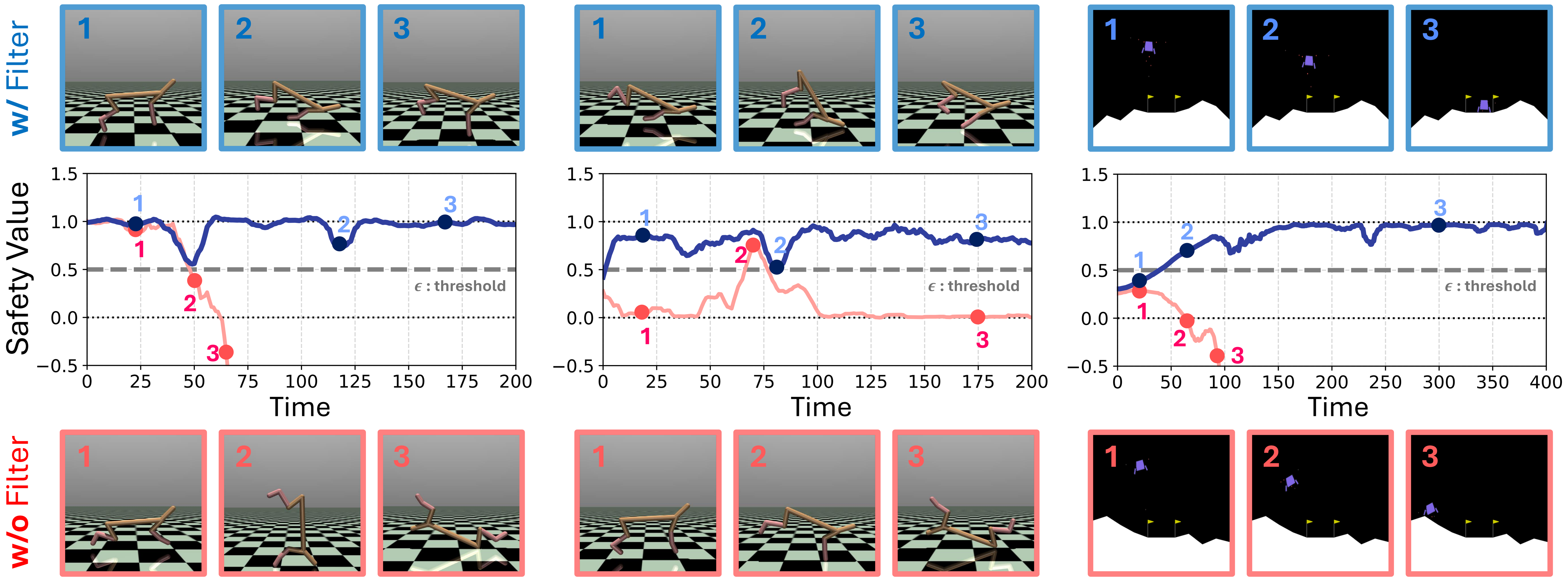
Experiments

(a) Double Integrator (Safety Value Function)



- **EigenSafe** properly evaluates relative safety across state-action pairs.
- **EigenSafe** generalizes to deterministic systems, with ψ_π being the indicator for the invariant set and γ_π being 1.

(b1) `cheetah_balance` unsafe if cheetah flips
(b2) `cheetah_run` unsafe if cheetah moves backward
(b3) `lunar_lander` unsafe if lander crashes or flies away



Paper w/ Exp.Details

Experiment Videos (Gym envs)

The 2nd Workshop on Safe and Robust Learning for Operation in the Real World (SAFE-ROL) at CoRL 2025