Spectral-Risk Safe Reinforcement Learning with Convergence Guarantees

Dohyeong Kim, Taehyun Cho, Seungyub Han, Hojun Chung, Kyungjae Lee, Songhwai Oh

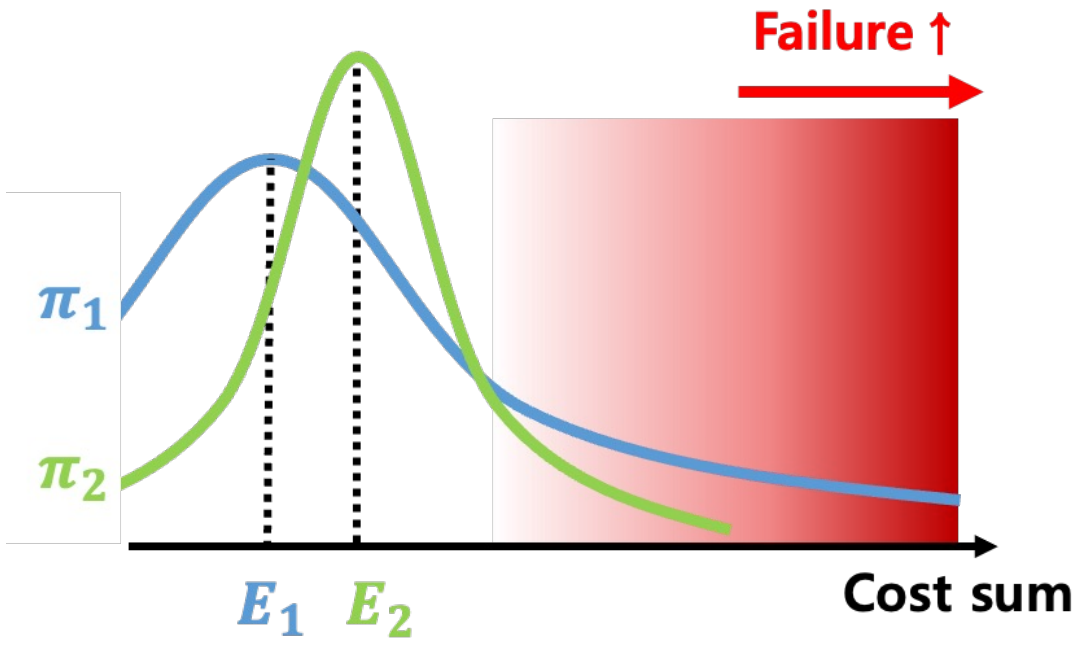
Robot Learning Lab, Seoul National University

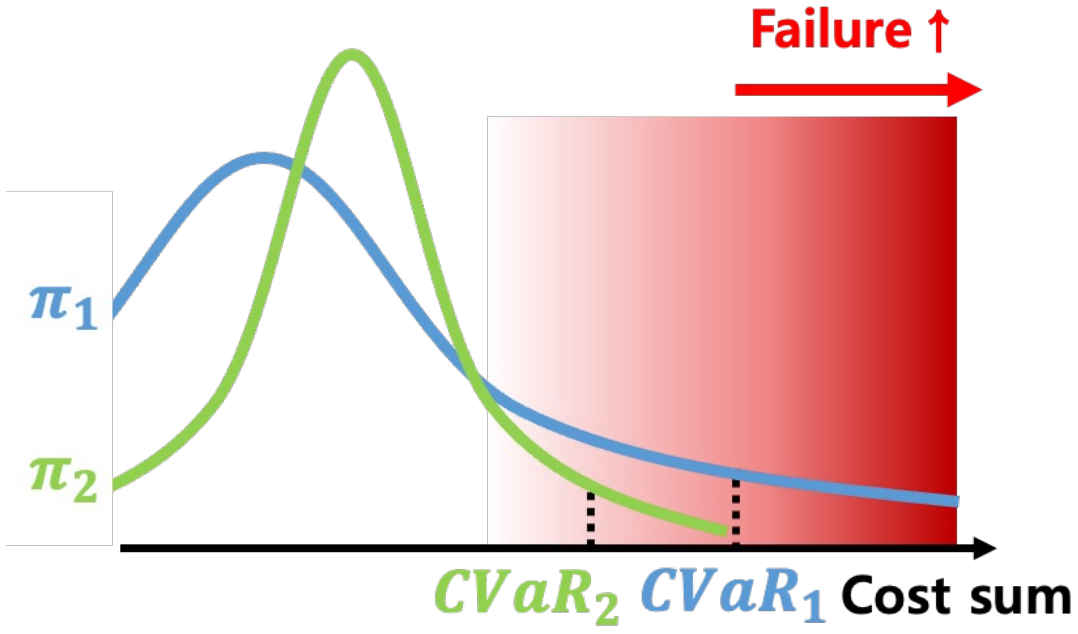
SEOUL NATIONAL UNIVERSITY
Dept. of Electrical and Computer Engineering

RLLAB

Motivation

- Safe RL: $max_\pi \mathbb{E}[\Sigma_{t=0}^{\infty} \gamma^t R_t]$ s.t. $\mathbb{E}\left[\Sigma_{t=0}^{\infty} \gamma^t C_{i,t}\right] \leq d_i/(1-\gamma)$.





Motivation

RLLAB

- A risk-constrained RL (RCRL) problem:

$$\max_\pi J_R(\pi) \text{ s.t. } \mathcal{R}_i(C_i^\pi) \leq d_i \; \forall_i, \text{ where } \mathcal{R}_i \text{ is a risk measure.}$$

- Due to the **nonlinearity of risk measures**, it is challenging to develop a safe RL algorithm that guarantees **convergence to an optimal policy**.

$\Longrightarrow$ Propose a bilevel optimization framework for risk-constrained RL using the duality of spectral risk and show convergence guarantees in tabular settings.

Spectral Risk Measure

- Definition:

$$\mathcal{R}_\sigma(X) \coloneqq \int_0^1 F_X^{-1}(u)\sigma(u)du,$$

where σ (spectrum) is an increasing function, $\sigma \geq 0$, and $\int_0^1 \sigma(u)du = 1$.

- Example:

Conditional value at risk (CVaR): $\sigma(u) = \mathbf{1}_{u \geq \alpha}/(1-\alpha)$.

Spectral Risk Measure

- Definition:

$$\mathcal{R}_\sigma(X) \coloneqq \int_0^1 F_X^{-1}(u)\sigma(u)du,$$

where σ (spectrum) is an increasing function, $\sigma \geq 0$, and $\int_0^1 \sigma(u)du = 1$.

- Dual form expression:

$$\mathcal{R}_\sigma(X) = \inf_g \mathbb{E}[g(X)] + \int_0^1 g^*(\sigma(u))du =: \mathcal{R}_\sigma^g(X),$$

where g is an increasing convex function,

$g^*(y) \coloneqq \inf_x xy - g(x)$ is the convex conjugate of g,

and $\mathcal{R}_\sigma^g(X)$ is a sub-risk measure.

Bilevel Optimization Framework

- Reformulation of the RCRL problem:

$$\max_\pi J_R(\pi) \text{ s.t. } \mathcal{R}_{\sigma_i}(C_i^\pi) \leq d_i \; \forall_i.$$

$$\Longrightarrow \sup_{g_1,\ldots,g_N} \underbrace{\max_\pi J_R(\pi) \text{ s.t. } \mathcal{R}_{\sigma_i}^{g_i}(C_i^\pi) \leq d_i \; \forall_i.}_{\text{Inner problem}}$$

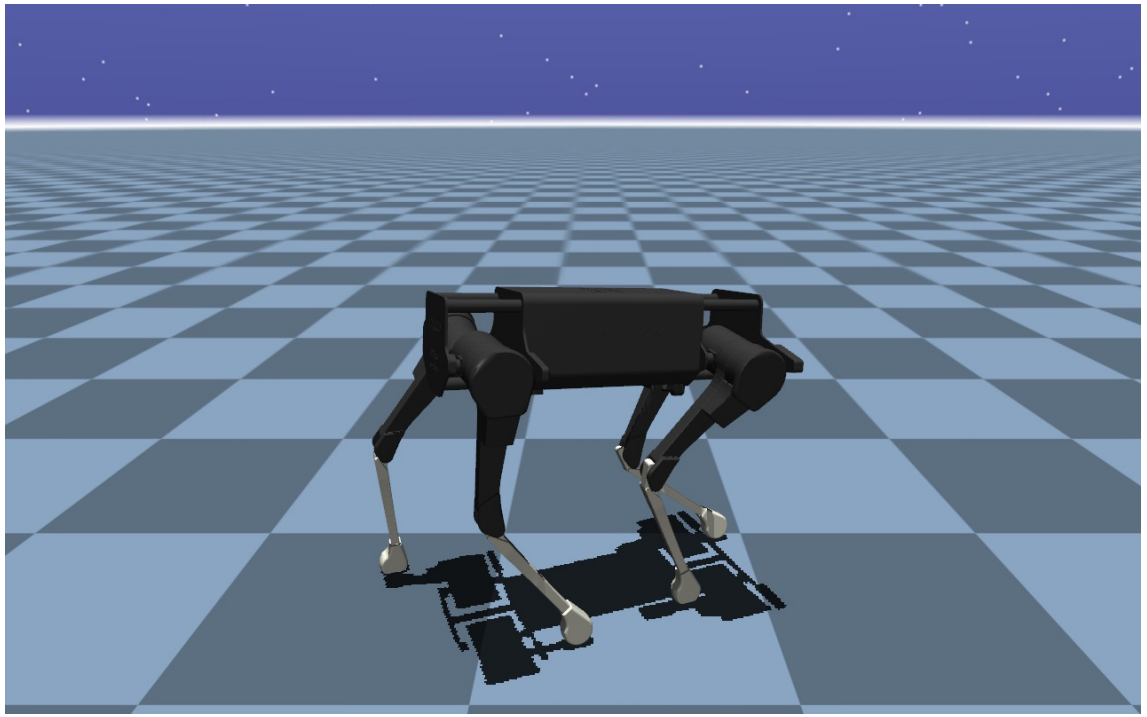
$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{\text{Outer problem}}$$

Experimental Results

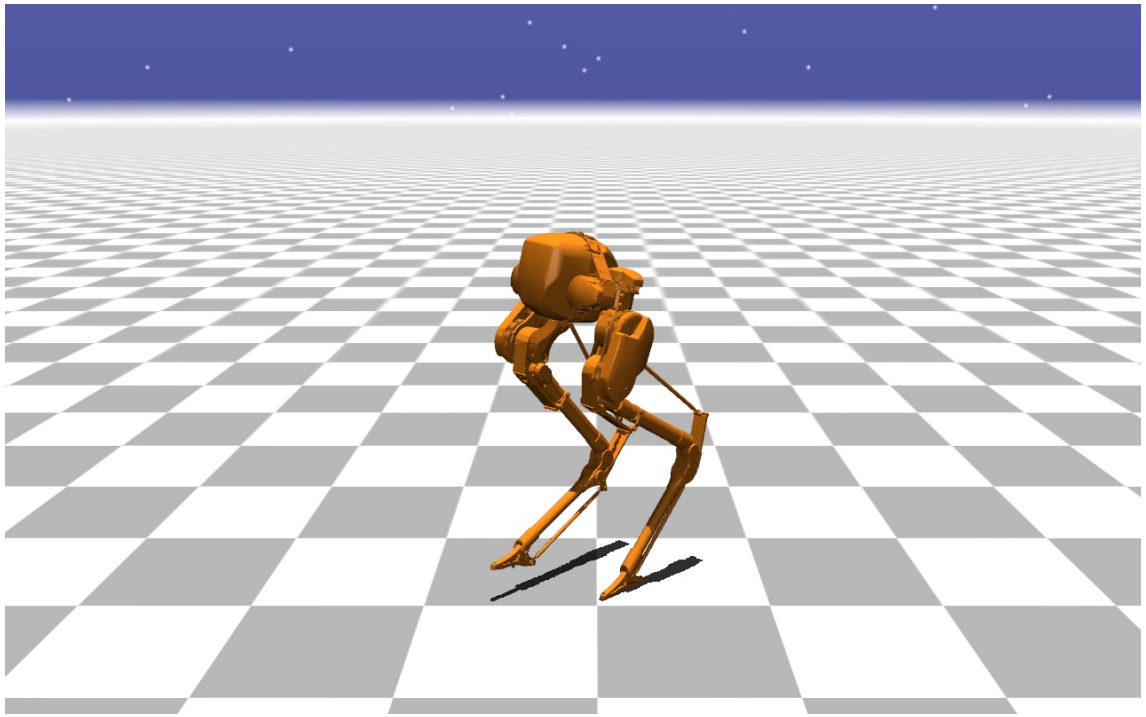
RLLAB

- Legged robot locomotion tasks:

Quadrupedal (Laikago)

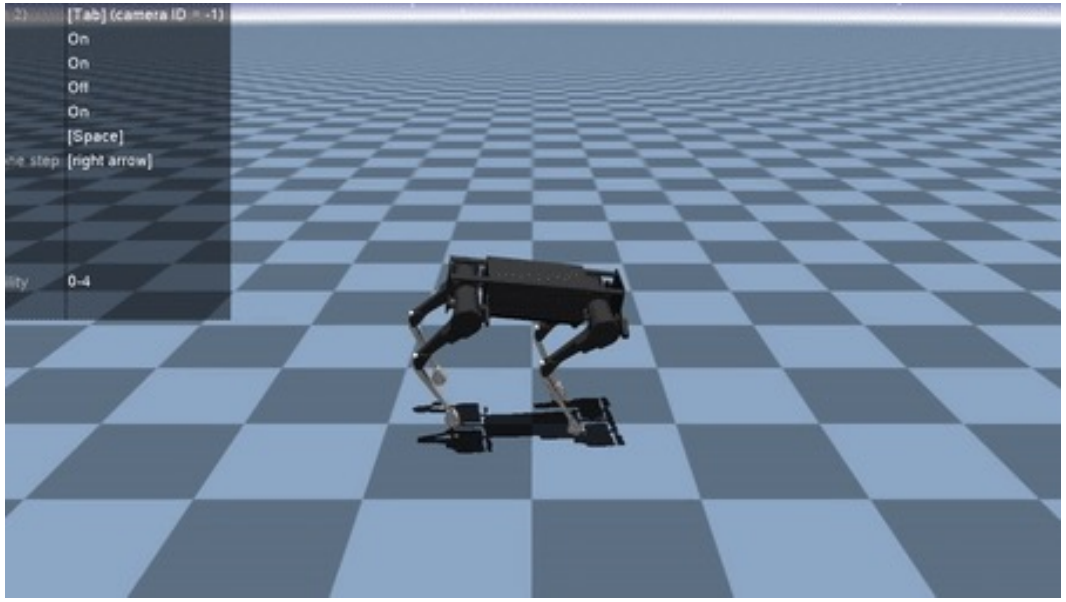


Bipedal (Cassie)

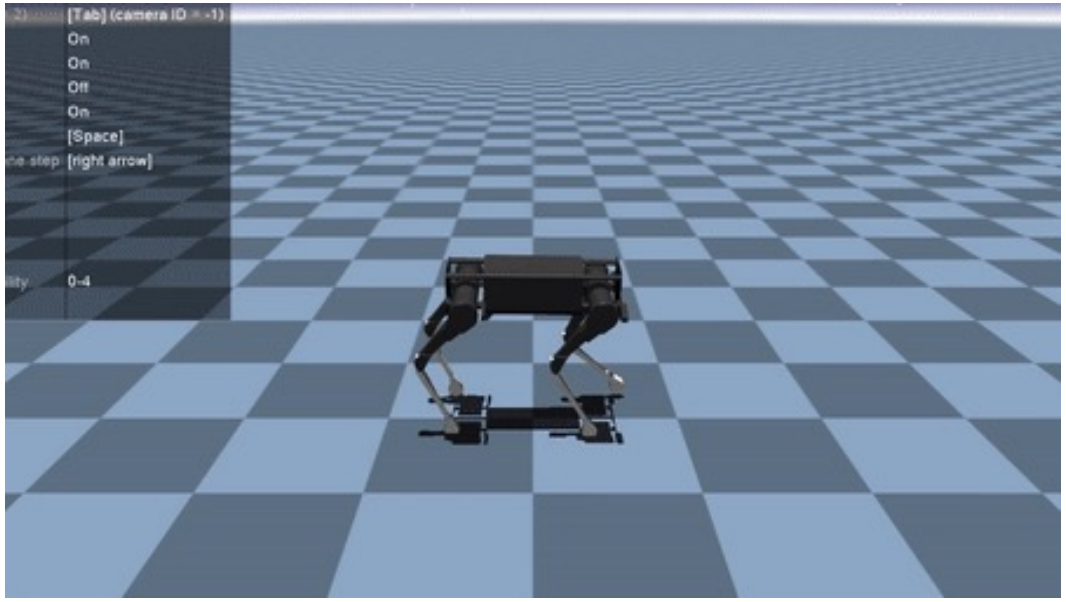


Experimental Results

SRCPO (Proposed)

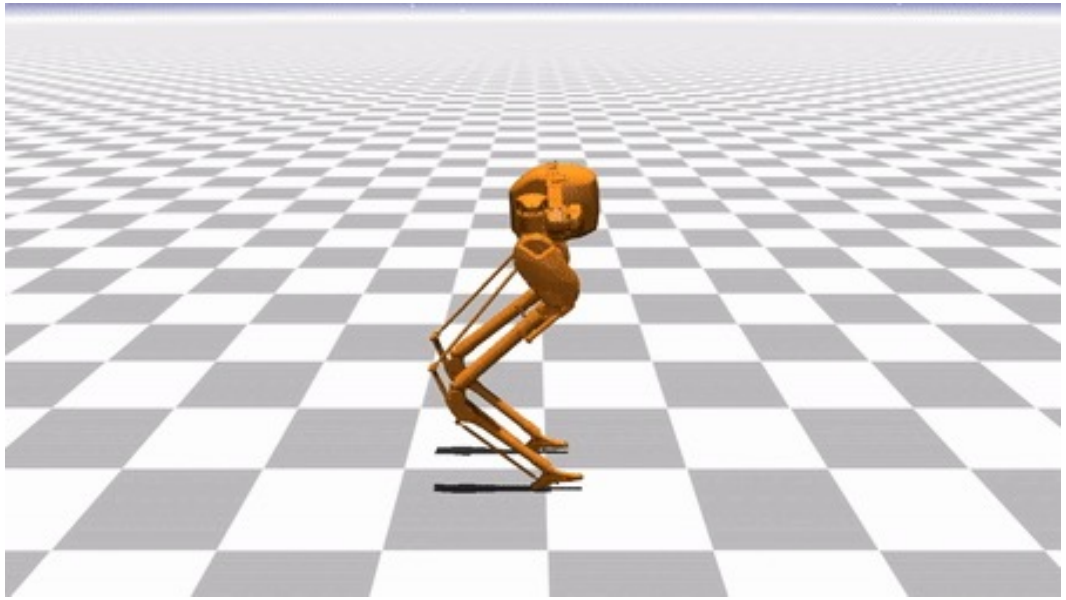


WCSAC-D

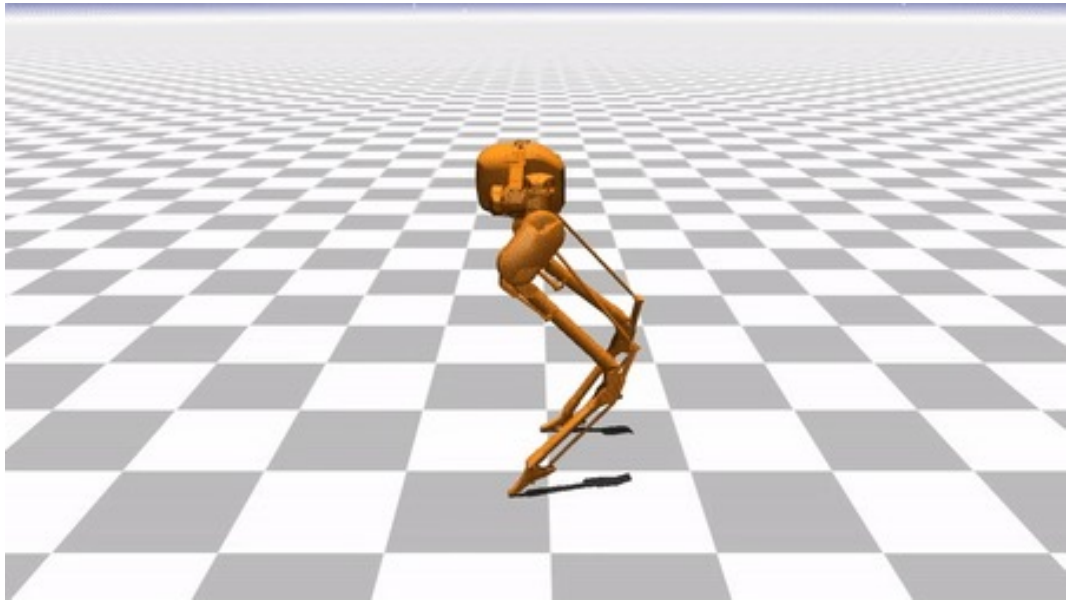


Experimental Results

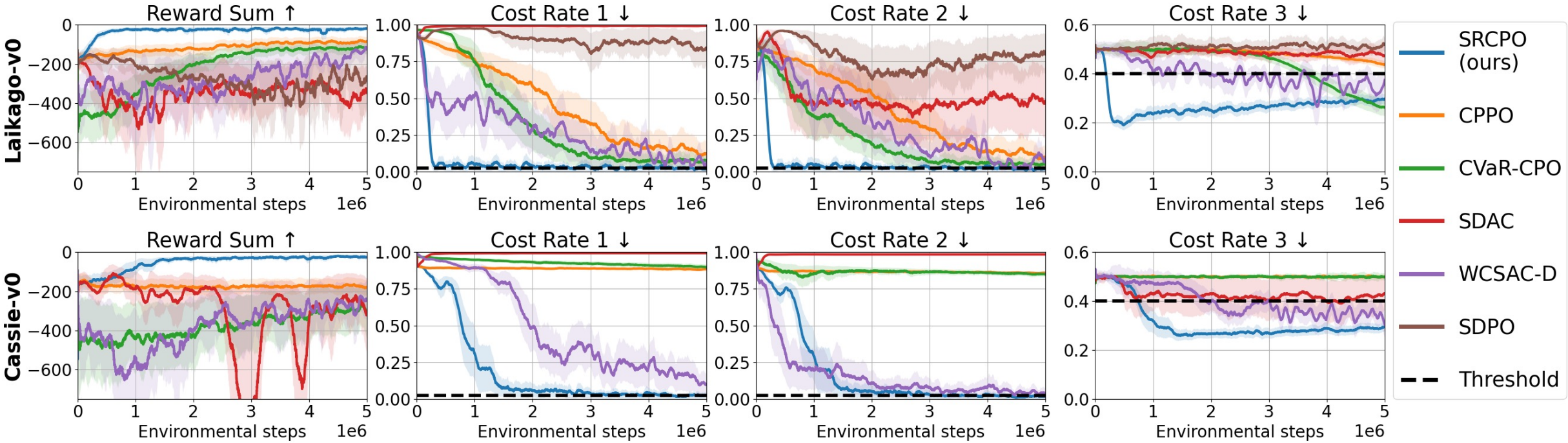
SRCPO (Proposed)



WCSAC-D



Experimental Results



Thank you for listening!

If you have any questions,

please contact to dohyeong.kim@rllab.snu.ac.kr.

SEOUL NATIONAL UNIVERSITY
Dept. of Electrical and Computer Engineering

RLLAB