LipsNet: A Smooth and Robust Neural Network with Adaptive Lipschitz Constant for High Accuracy Optimal Control

Xujie Song[1], Jingliang Duan[1], Wenxuan Wang[1], Shengbo Eben Li[1], Chen Chen[1], Bo Cheng[1], Bo Zhang[2], Junqing Wei[2], Xiaoming Simon Wang[2]

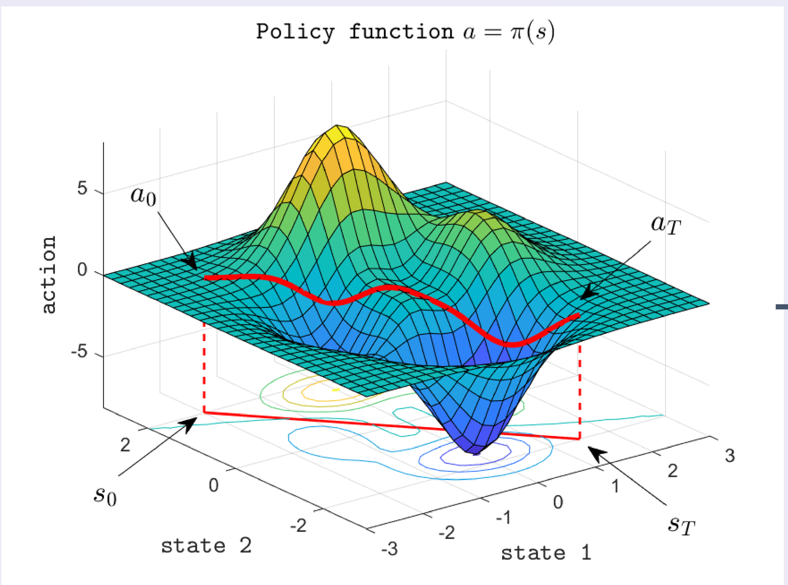
[1]Tsinghua University, [2]Didi Chuxing

① Abstract

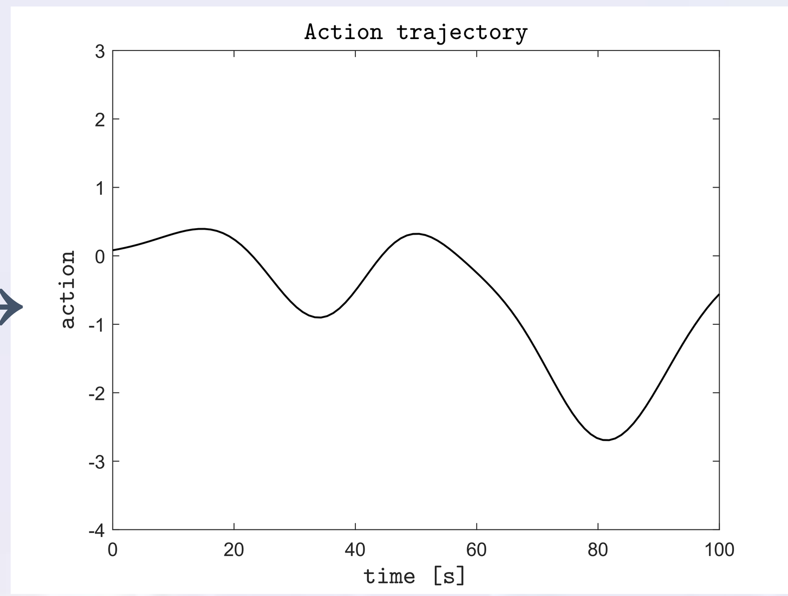
Deep **reinforcement learning** (**RL**) is a powerful approach for solving optimal control problems. However, RL-trained policies often suffer from the **action fluctuation problem**, where the consecutive actions significantly differ despite only slight state variations. This problem results in mechanical components' wear and tear and poses safety hazards. The action fluctuation is caused by the high Lipschitz constant of the actor network. To address this problem, we propose a neural network named **LipsNet**. We propose the Multi-dimensional Gradient Normalization (**MGN**) method, to constrain the Lipschitz constant of networks with multi-dimensional input and output. Benefiting from MGN, LipsNet achieves **Lipschitz continuity**, allowing smooth actions while preserving control performance by automatically adjusting Lipschitz constant. LipsNet addresses the action fluctuation problem at network level rather than algorithm level, which can serve as **actor networks** in most RL algorithms, making it more flexible and user-friendly than previous works. Experiments demonstrate that LipsNet has good **landscape smoothness** and **noise robustness**, resulting in significantly smoother action compared to the Multilayer Perceptron (MLP).

② What Causes Action Fluctuation in RL ?

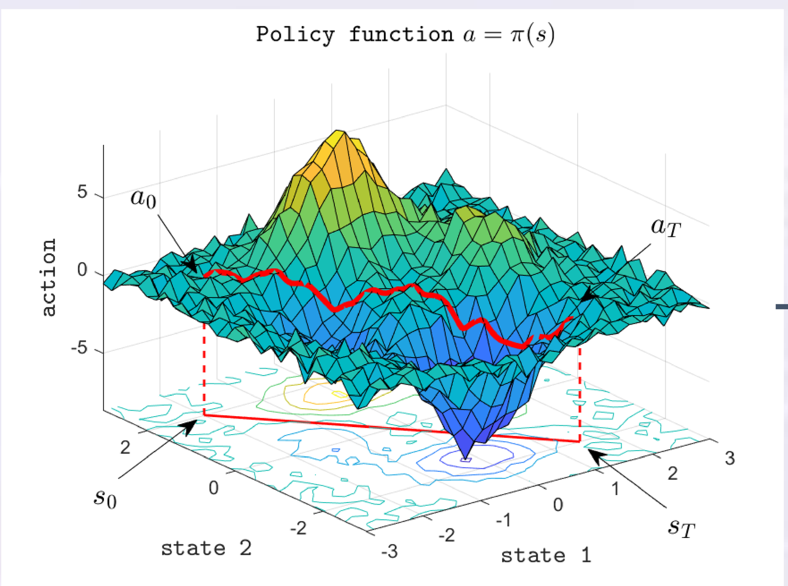
- Observation noise
- Policy function's rugged landscape



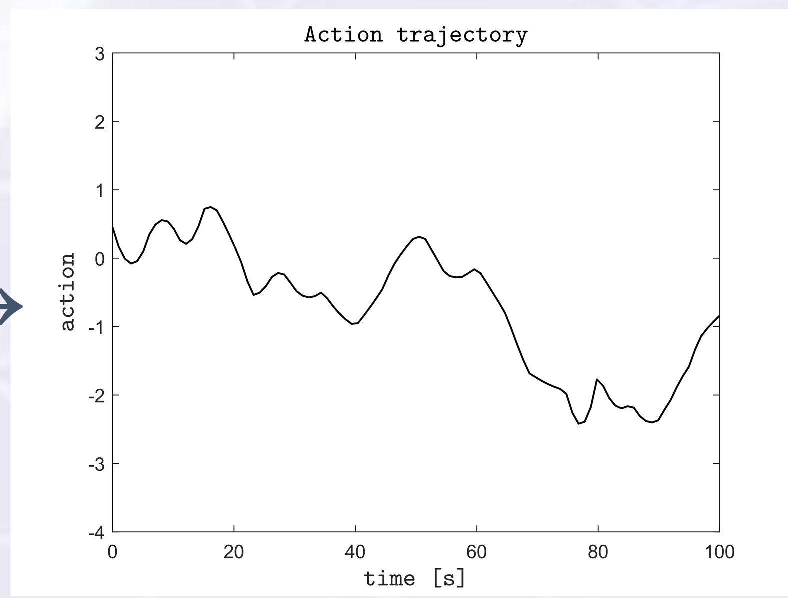
smooth policy



smooth action



unsmooth policy



unsmooth action

③ Why Reduce Lipschitz Constant ?

- ***Definition*: Global Lipschitz Continuity**

Given a function $f: \mathbb{R}^n \to \mathbb{R}^m$, if there exists a constant $K > 0$ satisfies

$$\|f(x_1) - f(x_2)\| \le K\|x_1 - x_2\|,\ \forall x_1, x_2 \in \mathbb{R}^n$$

then f is a globally K-Lipschitz continuous function over $\mathbb{R}^n$. The smallest K is called the global Lipschitz constant of f.

- ***Definition*: Local Lipschitz Continuity**

Given a function $f: \mathbb{R}^n \to \mathbb{R}^m$, if there exists a constant $K > 0$ satisfies

$$\|f(x_1) - f(x_2)\| \le K\|x_1 - x_2\|,\ \forall x_1, x_2 \in \mathcal{X}$$

then f is a locally K-Lipschitz continuous function over $\mathcal{X}$. The smallest K is called the local Lipschitz constant of f over $\mathcal{X}$.

- **Lipschitz constant K reflects function's <u>landscape smoothness</u> and <u>noise robustness</u>**

(smoothness) Let $x_2 = x_1 + \Delta x$, if $\Delta x \to 0$, then $\frac{\|f(x_1)-f(x_2)\|}{\|x_1-x_2\|} \to \|f'(x)\|$.

(robustness) Let $x_2 = x_1 + \sigma$, then $\frac{\|f(x)-f(x+\sigma)\|}{\|x-(x+\sigma)\|} \le K$, where σ can be considered as the observation noise.

⑥ Code and Contacts

- The code is available at https://github.com/jerry99s/LipsNet.
- First author: Xujie Song (songxj21@mails.tsinghua.edu.cn)
- Corresponding author: Shengbo Eben Li (lishbo@tsinghua.edu.cn)

④ Our Methods

- **Action Fluctuation Ratio**

<u>Firstly</u>, we define the action fluctuation ratio $\xi(\pi)$ as a quantitative index for action fluctuation level of policy π. The smaller $\xi(\pi)$ is, the smoother action is.

$$\xi(\pi) = \mathbb{E}_{\tau\sim\rho_\pi}\left[\frac{1}{T}\sum_{t=1}^{T}\|a_t - a_{t-1}\|\right]$$

- **Multi-dimensional Gradient Normalization (MGN)**

<u>Secondly</u>, we propose MGN to constrain the Lipschitz constant of networks with multi-dimensional input and output.

Theorem 3.1 Suppose $f: \mathbb{R}^n \to \mathbb{R}^m$ is a continuously differential neural network with piecewise linear activations. Then f_{MGN} is a globally K-Lipschitz continuous neural network,

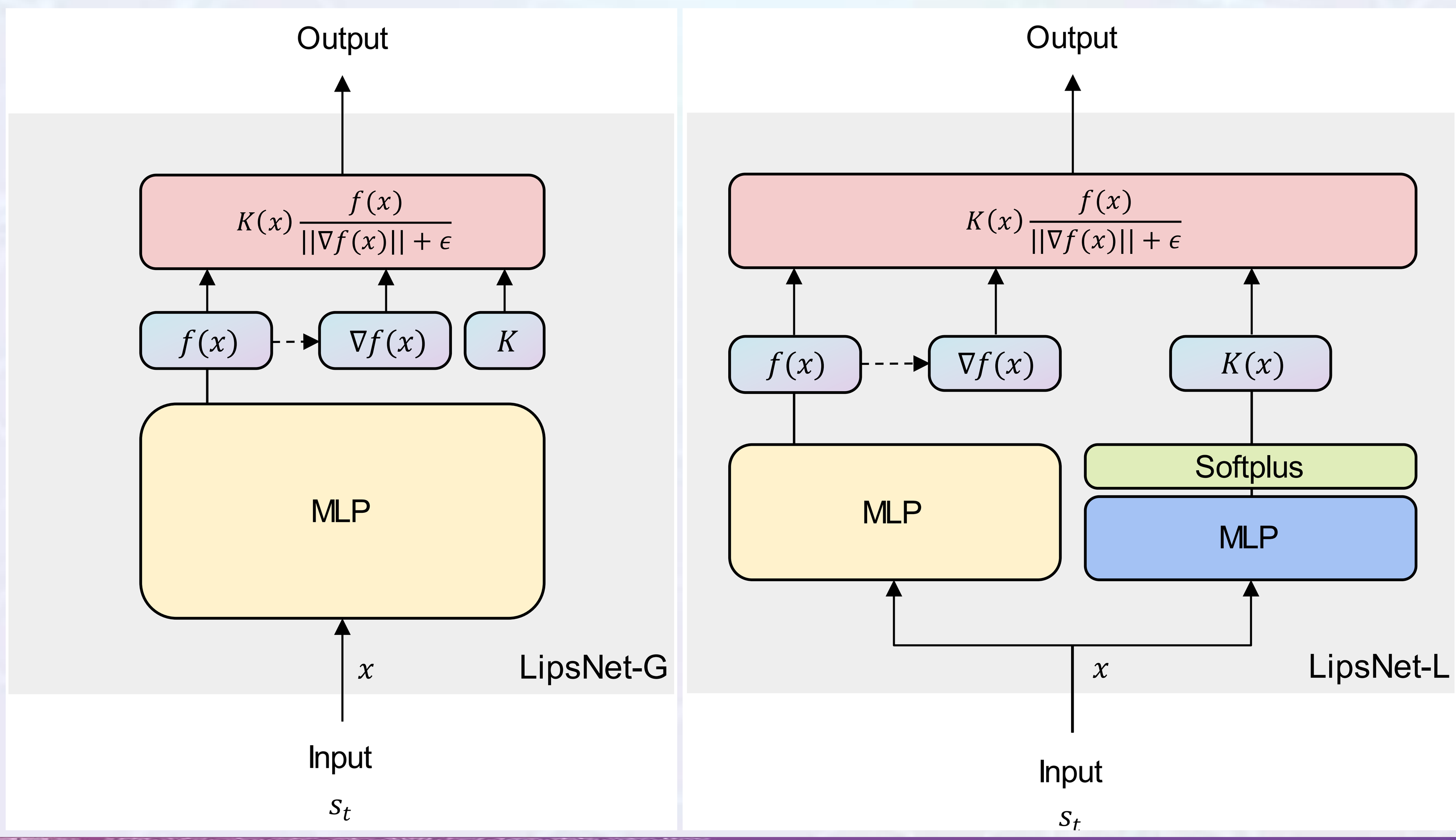
$$f_{\mathrm{MGN}}(x) = K \cdot \frac{f(x)}{\|\nabla_x f(x)\| + \epsilon}$$

where K is a positive constant, $\|\nabla_x f(x)\|$ is the 2-norm of Jacobian matrix and ϵ is a small positive constant.

- **Overall Network Architecture**

<u>Finally</u>, the network is shown below with the name 'LipsNet-G'. After replacing K with another network with $\mathrm{Softplus}$, it becomes 'LipsNet-L'. For a lower K, a regularization term is introduced in the actor loss:

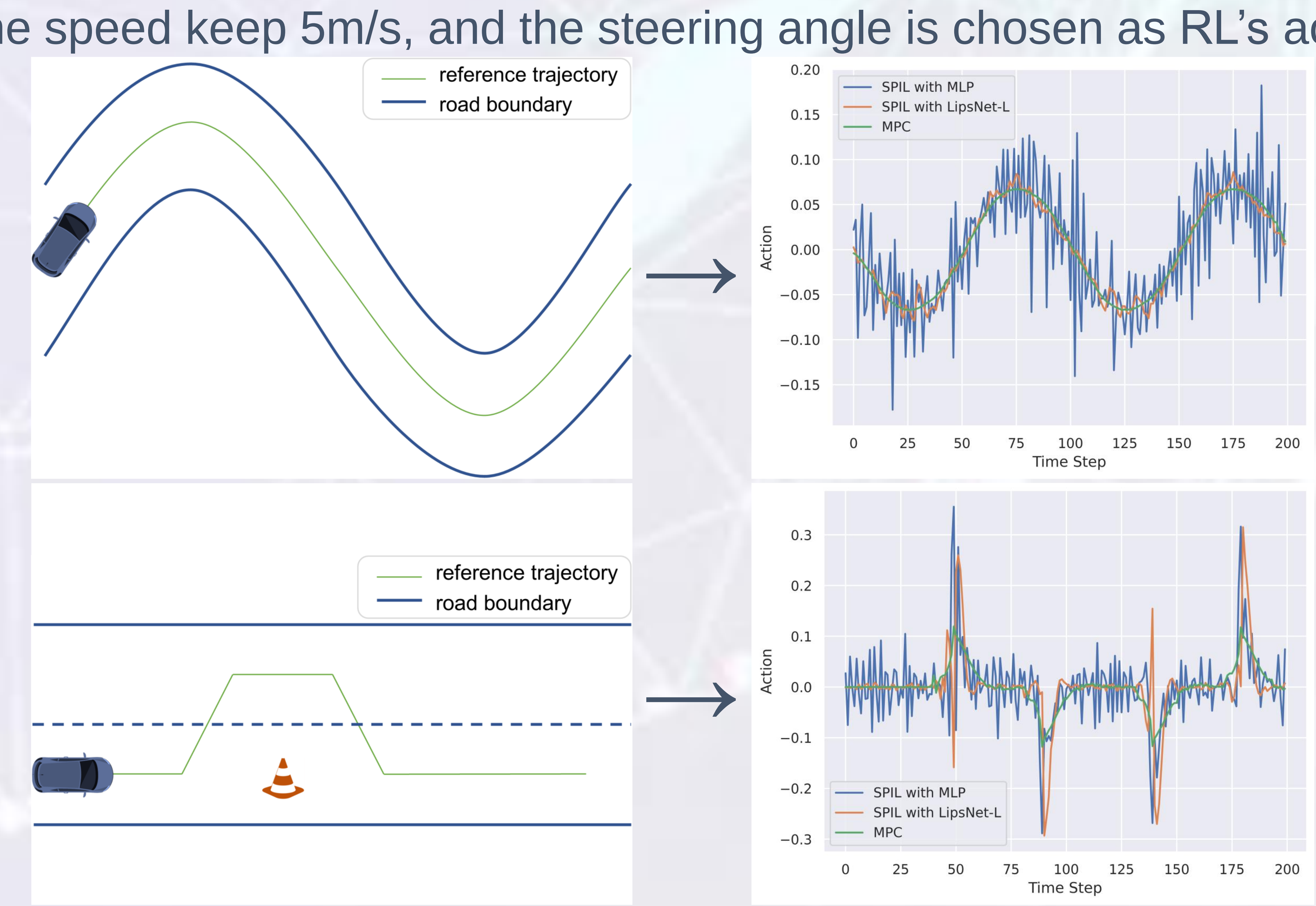
$$\mathcal{L}' = \mathcal{L} + \lambda K(x)^2$$



⑤ Experiment Results

- **Vehicle Trajectory Tracking**

The vehicles aim to track reference trajectories with observation noises. The speed keep 5m/s, and the steering angle is chosen as RL's action.



- **DeepMind Control Suit**

DMControl is a well-known benchmark for RL and continuous control.