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

Xujie Song[1], Jingliang Duan[2], Wenxuan Wang[1], Shengbo Eben Li[1]*, Chen Chen[1], Bo Cheng[1], Bo Zhang[3], Junqing Wei[3], Xiaoming Simon Wang[3]
[1]Tsinghua University, [2]University of Science and Technology Beijing, [3]Didi Chuxing

## 1. Introduction

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. To address this problem, we propose a neural network named **LipsNet**. LipsNet achieves Lipschitz continuity, allowing smooth action while preserving control performance. It can serve as **actor networks** in most RL algorithms, which is flexible and user-friendly.

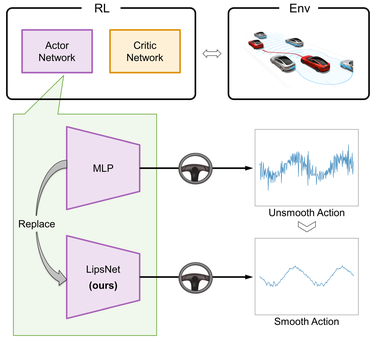
Experiments demonstrate that LipsNet has good **landscape smoothness** and **noise robustness**, resulting in significantly smoother action compared to the Multilayer Perceptron (MLP).

**Contributions :**

- We propose a network-wise Lipschitz constraint method, suitable for networks with multi-dimensional input and output.
- A novel network structure named LipsNet is proposed for smoothing the action in RL.
- We package LipsNet into a PyTorch module to facilitate the implementation and future research.

## 2. Objective

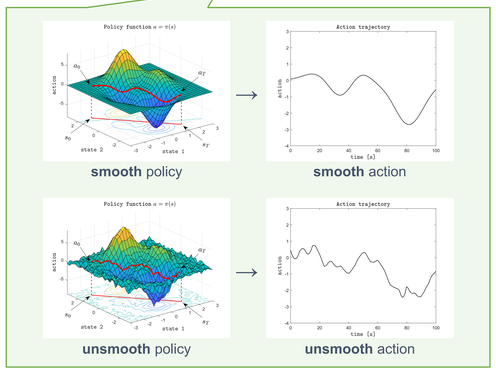
The objective of this paper is to smooth the action trajectory in RL by replacing the actor network. This operation is intended to be network-level only, which should not complicate the RL algorithms.




## 3. What Causes Action Fluctuation in RL ?

- Observation noise
- Policy function's unsmooth landscape

} two reasons




## 4. 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 $\mathcal{X}$ where $\mathcal{X} \subset \mathbb{R}^n$. The smallest $K$ is called the local Lipschitz constant of $f$ over $\mathcal{X}$.

- **Lipschitz constant $K$ reflects the landscape smoothness and noise robustness of neural networks**

  (smoothness) Let $x_2 = x_1 + \Delta x$, if $\Delta x \to 0$, then $\frac{\|f(x_1)-f(x_2)\|}{\|x_1-x_2\|} \to \|f'(x)\|$. It means the gradient norm tends to be smaller than $K$.

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

## 5. Our Method

- **Action Fluctuation Ratio**

  Firstly, we define the action fluctuation ratio $\xi(\pi)$ as a quantitative index for action fluctuation level of policy $\pi$. The smaller $\xi(\pi)$ is, the smoother actions policy $\pi$ has.

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

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

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

  ***Lemma 3.1*** Suppose there is a continuously differential function $f:\mathbb{R}^n \to \mathbb{R}^m$. Then, the global $K$-Lipschitz continuity condition is equivalent to

  $$\|\nabla_x f(x)\| \le K, \forall x$$

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

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

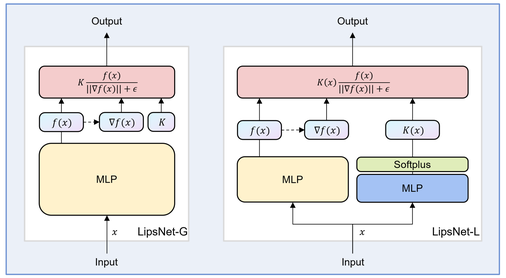
  where $K$ is a positive constant, $\|\nabla_x f(x)\|$ is the 2-norm of Jacobian matrix and $\epsilon$ is a small positive constant.

- **Overall Network Architecture**

  Finally, the network is shown below with the name *LipsNet-G*, which is a global Lipschitz continuous network. After replacing $K$ with additional network and stacking $\mathrm{Softplus}$, it becomes a local Lipschitz continuous network named *LipsNet-L*.

  The network will be used as actor network in actor-critic RL. To obtain a lower Lipschitz constant $K$ for smoother actions, a regularization term is introduced in actor loss:

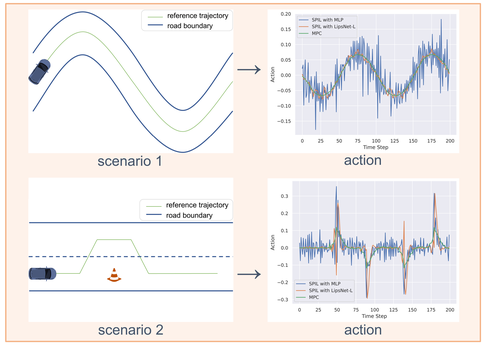
  $$\mathcal{L}' = \mathcal{L} + \lambda K(x)^2$$




## 6. Experiment Results

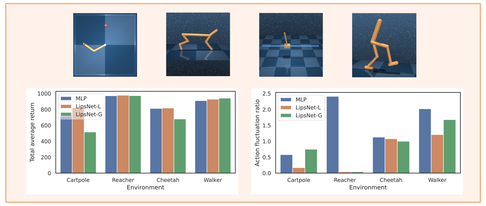
- **Vehicle Trajectory Tracking**

  In this environment, the vehicle aims to track reference trajectory under noised observations. The speed keep in 5m/s. We control the vehicle by changing steering angle. As the results shown, the action fluctuation of LipsNet is much smaller than MLP's.




- **DeepMind Control Suit**

  We selected 4 environments in DMControl for testing. The results show that LipsNet has lower action fluctuation ratio with the same level of total average return compared to MLP.




## 7. Code & Contact

- GitHub : github.com/jerry99s/LipsNet
- Contact : songxj21@mails.tsinghua.edu.cn

Tsinghua University