



北京通用人工智能研究院
Beijing Institute for General Artificial Intelligence

SPHERE: Mitigating the Loss of Spectral Plasticity in Mixture-of-Experts for Deep Reinforcement Learning

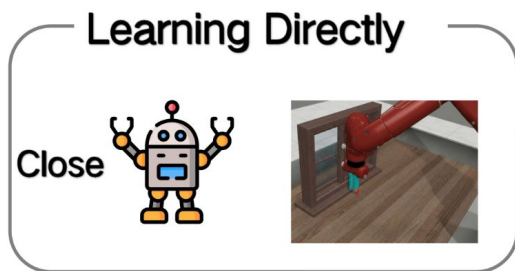
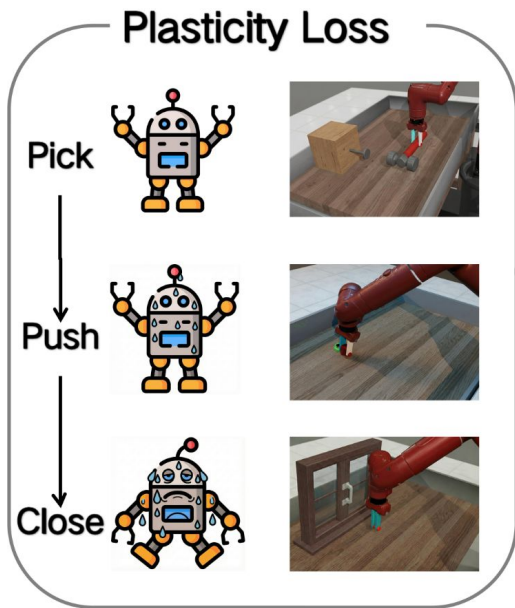
Lirui Luo, Guoxi Zhang, Hongming Xu, Cong Fang, Qing Li
Beijing Institute for General Artificial Intelligence (BIGAI)
Peking University

<https://sphere-rl.github.io/>

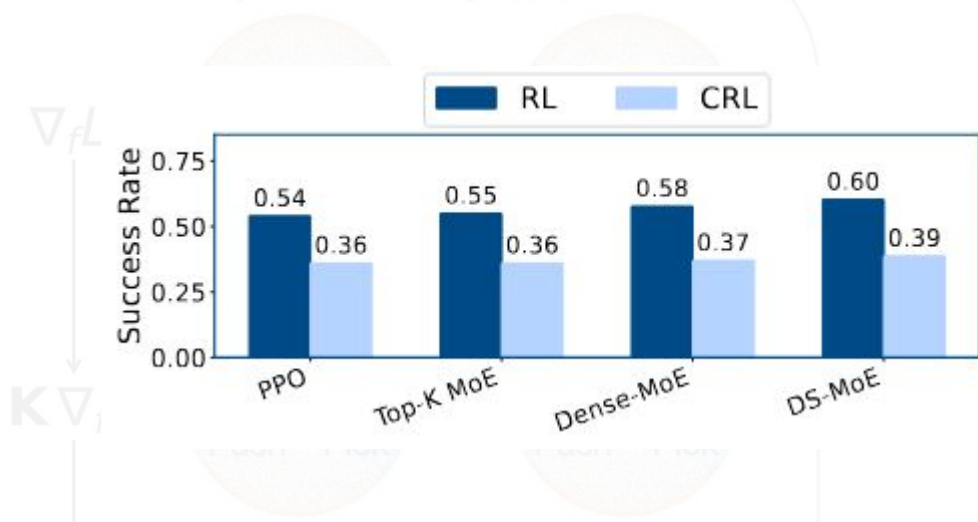




What is Plasticity Loss



Maintain Plasticity
by Maximizing $r_e(K)$

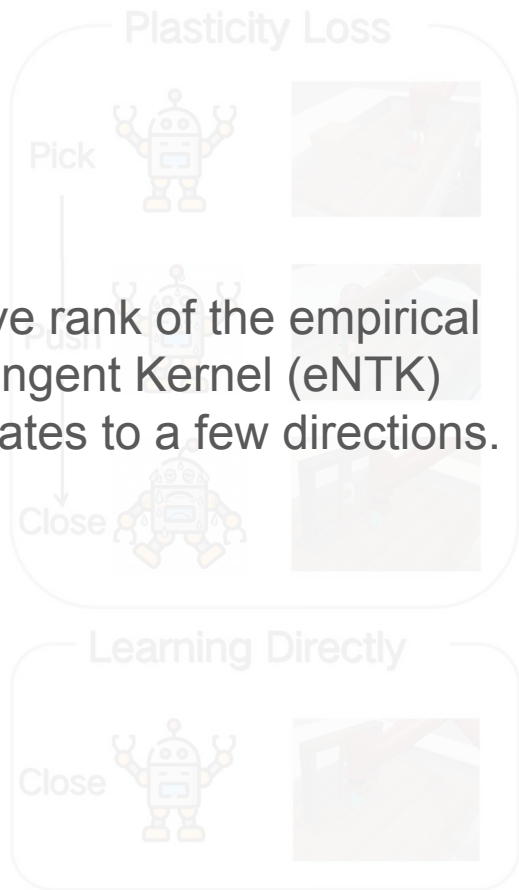


Both MLP and MOE lose the ability to learn from new experiences.

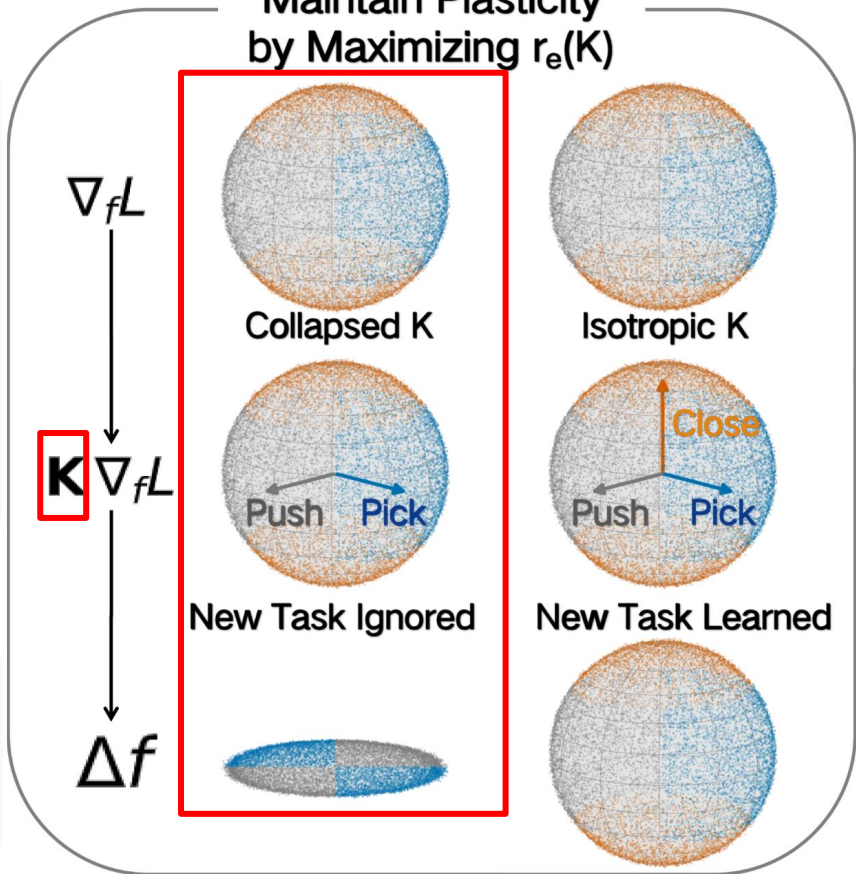


What is the Loss of Spectral Plasticity?

Low effective rank of the empirical Neural Tangent Kernel (eNTK) restricts updates to a few directions.

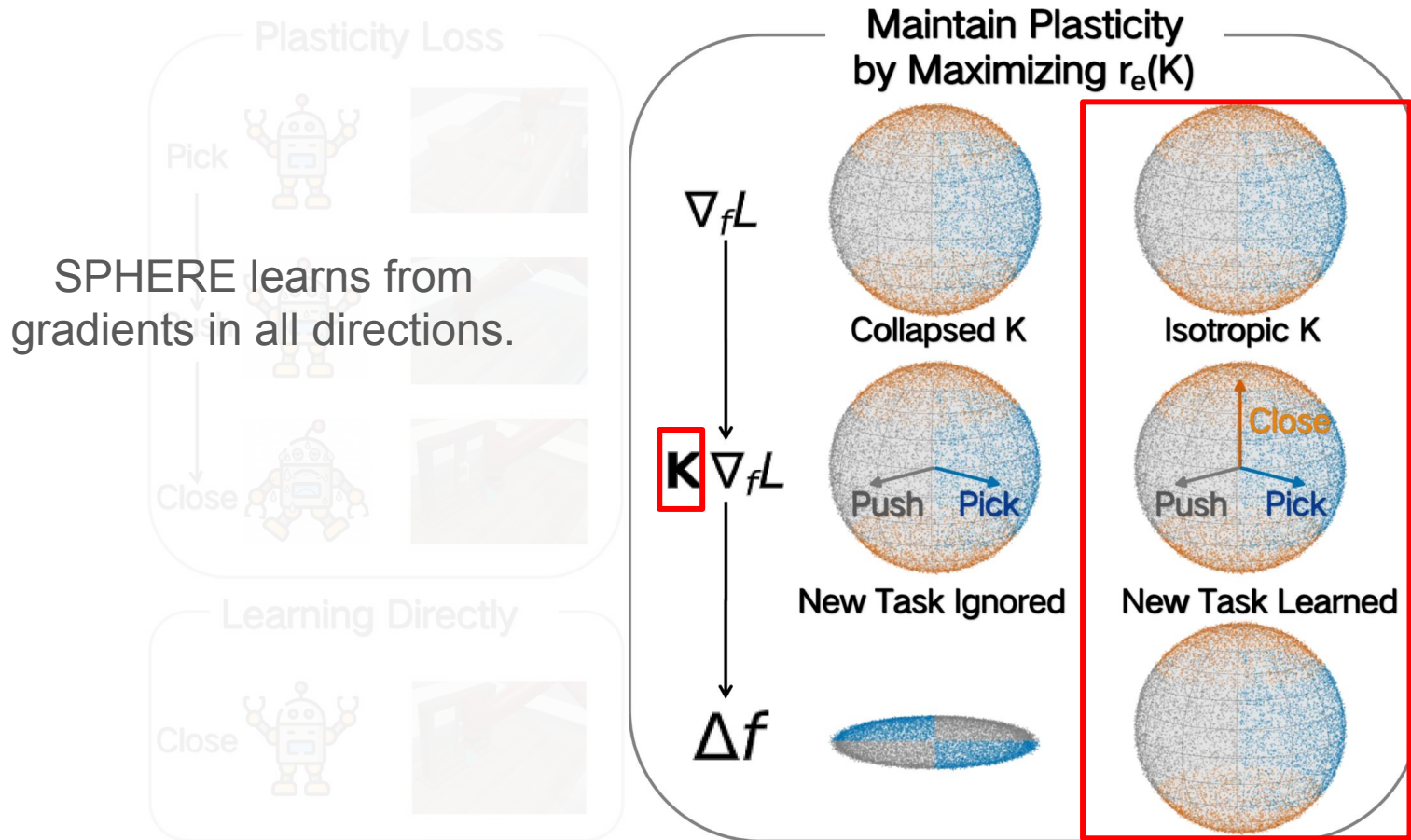


Maintain Plasticity by Maximizing $r_e(K)$



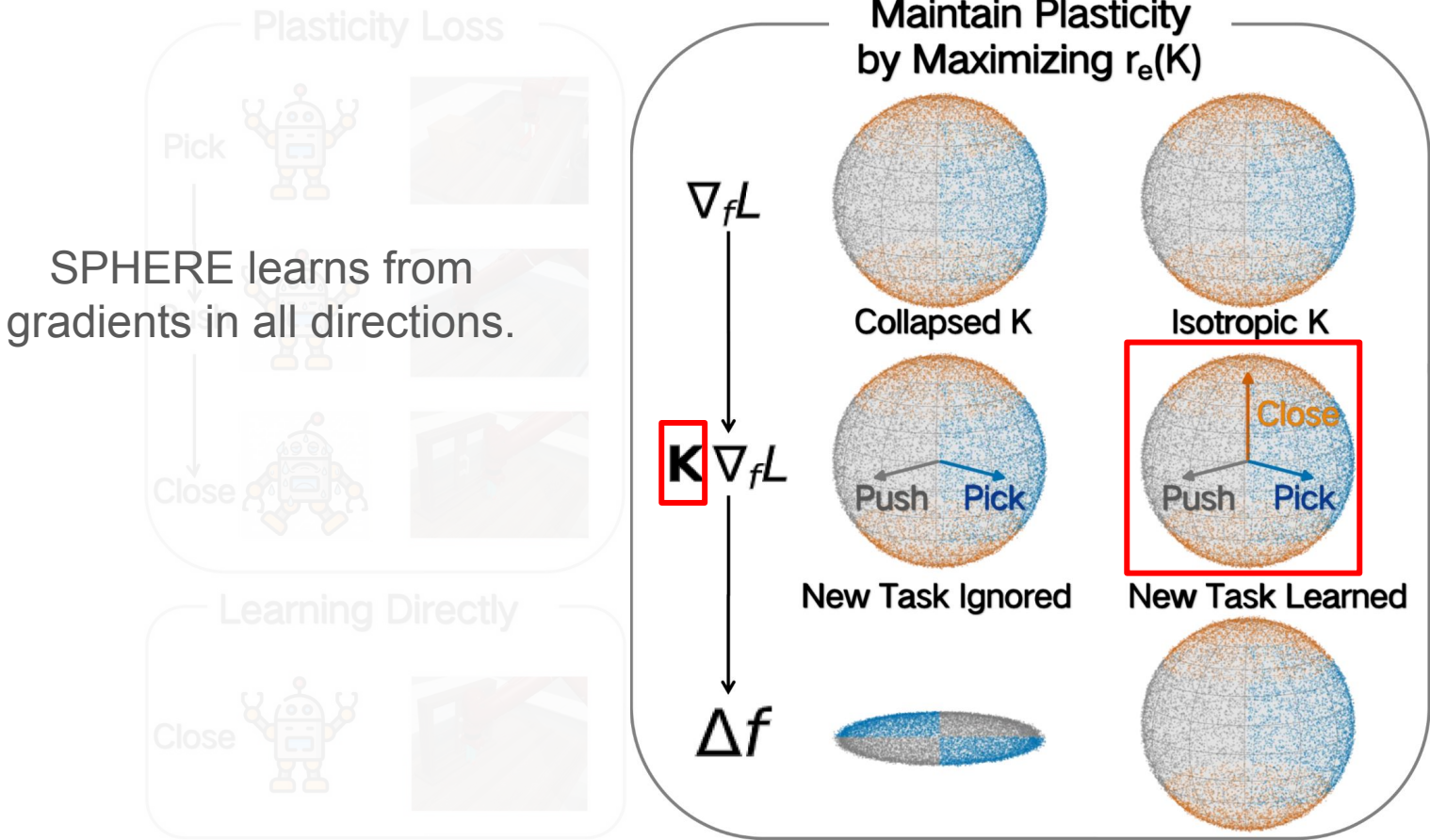


How SPHERE Mitigating the Loss of Spectral Plasticity?





How SPHERE Mitigating the Loss of Spectral Plasticity?





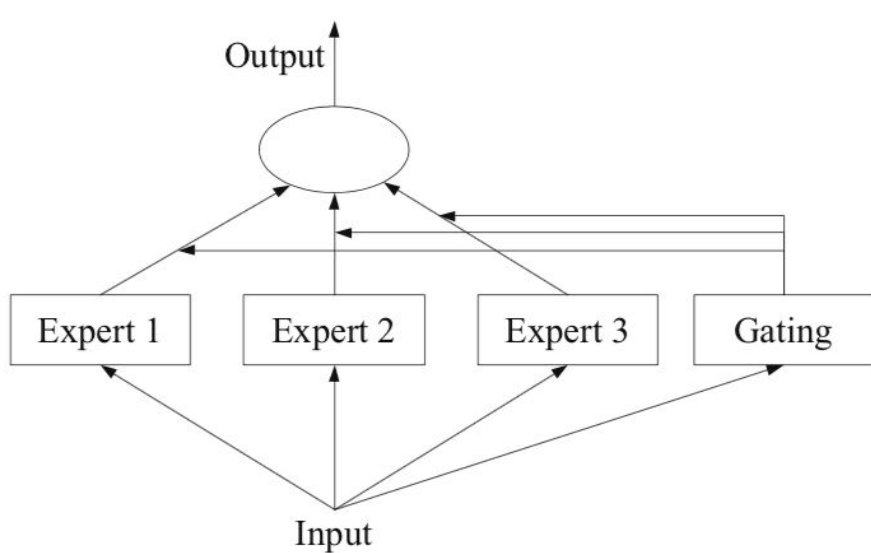
Limitation for Construct K

$$K = JJ^T \in \mathbb{R}^{N \times N}$$

$$T(K) = O(N^2P)$$

$$S(J) = O(NP)$$

When the number of parameters is large, it is almost impossible to construct K.



Get $J \in \mathbb{R}^{N \times P}$ from backpropagation



Decompose K into smaller matrices

$$K = JJ^T \in \mathbb{R}^{N \times N}$$



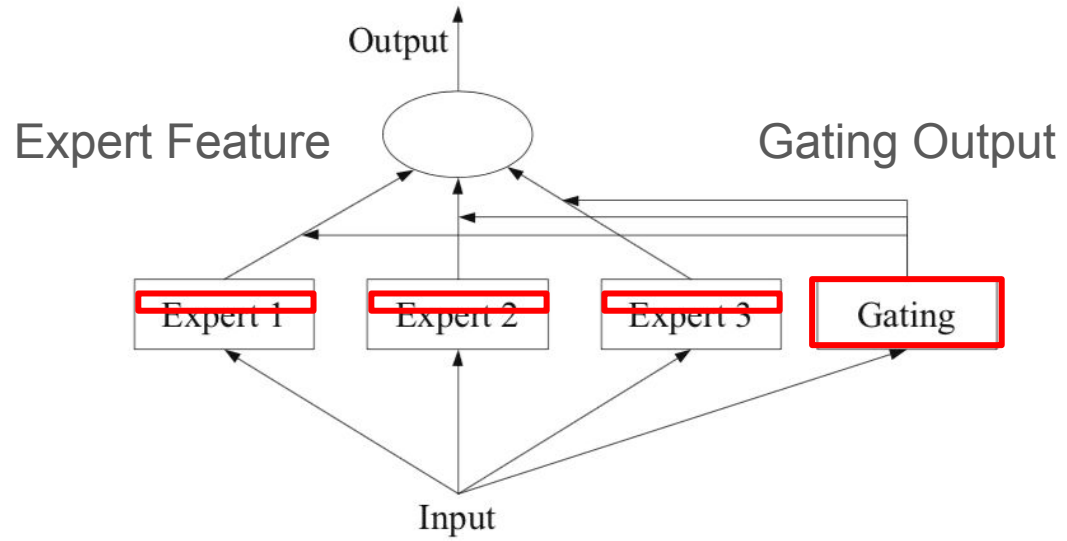
$$G^{GN} = J^T J \in \mathbb{R}^{P \times P}$$



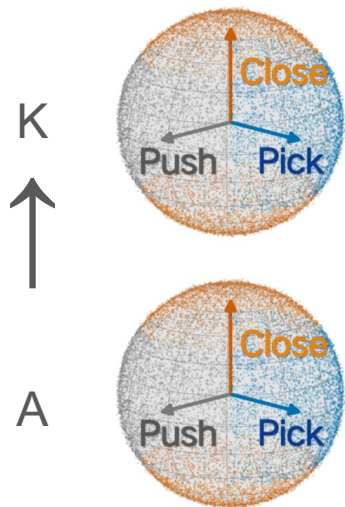
$$A = \varphi^T \varphi$$

Can be constructed from a forward pass without addition backpropagation.

$$\varphi = [\text{feature}_1 * \text{gating}_1, \dots, \text{feature}_n * \text{gating}_n]$$



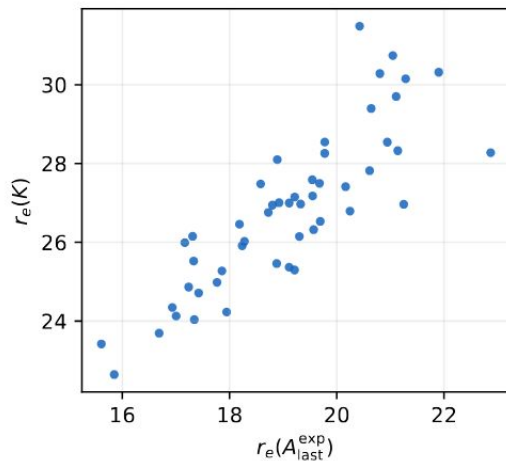
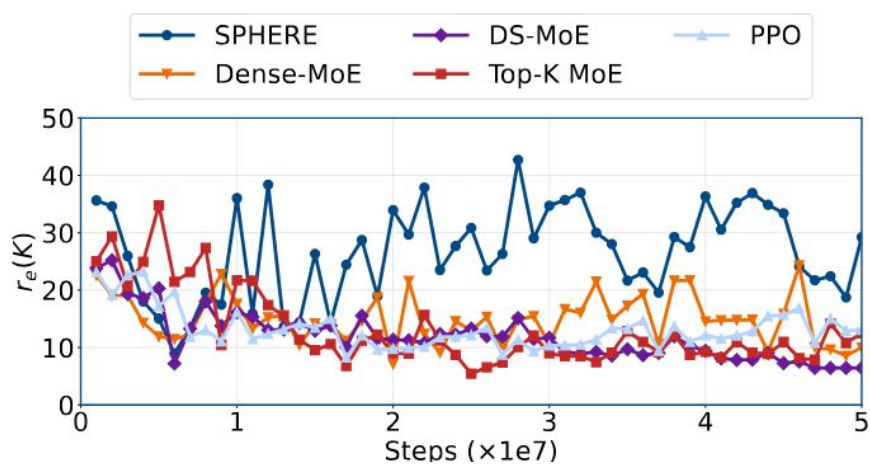
What Does SPHERE Do?



The spectrum of K is uniformized by the spectrum of uniform A .



SPHERE Improves The Spectral Plasticity



Initialize → Stand → Walk → Pole → Slide → Run

w/o SPHERE

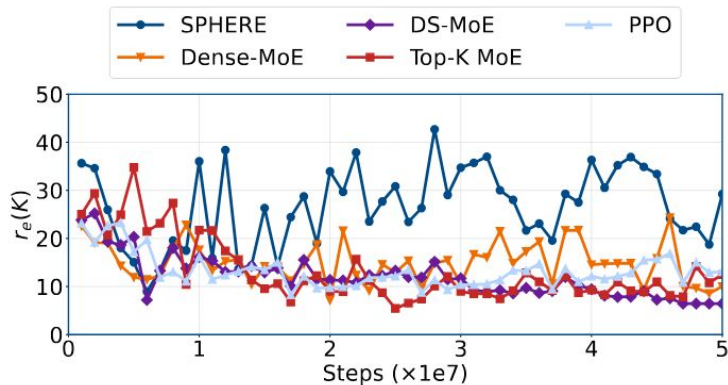


w/ SPHERE

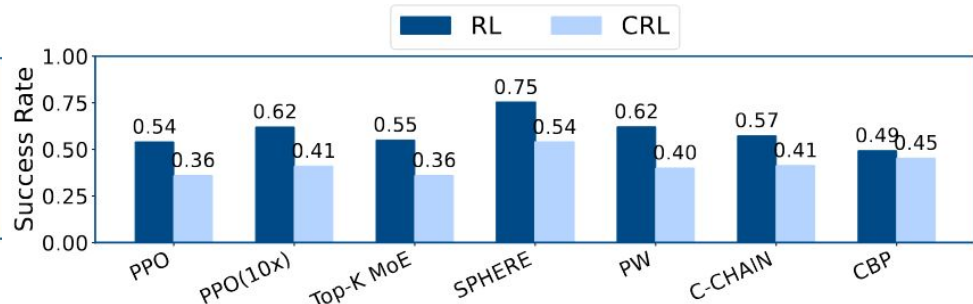
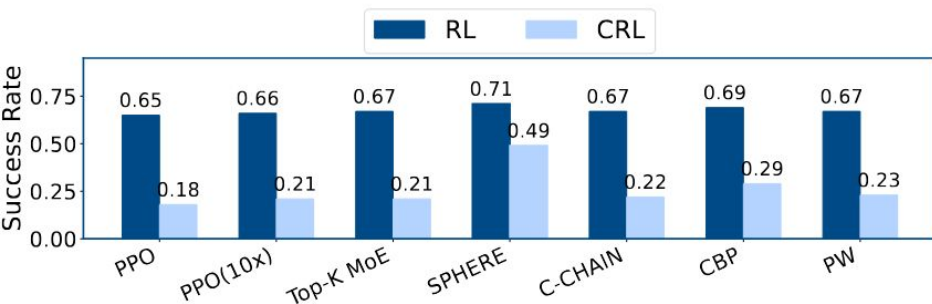




Higher Spectral Plasticity Helps Learning



Improves average success under continual RL by 133% and 50%.



Almost No Computational Overhead



Method	$E=10$	$E=20$	$E=40$
Top- K MoE (baseline)	2.23	2.23	2.24
SPHERE (ours)	2.37 (+6.28%)	2.34 (+4.93%)	2.37 (+5.80%)

There was almost no expense regardless of the number of experts.

Takeaways



We **formalize** loss of plasticity as **loss of spectral plasticity** using NTK theory.

We present **SPHERE**, improving spectral plasticity by making the weighted expert feature spectrum uniform.

Significantly improved continual reinforcement learning **performance** and **maintained the metric of spectral plasticity**.

Code & Demo: <https://sphere-rl.github.io/>

SPHERE: Mitigating the Loss of Spectral Plasticity in Mixture-of-Experts for Deep Reinforcement Learning

