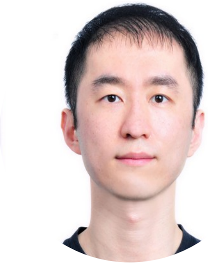
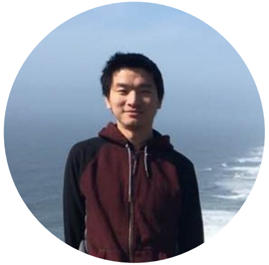
北京通用人工智能研究院
Beijing Institute for General Artificial Intelligence

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

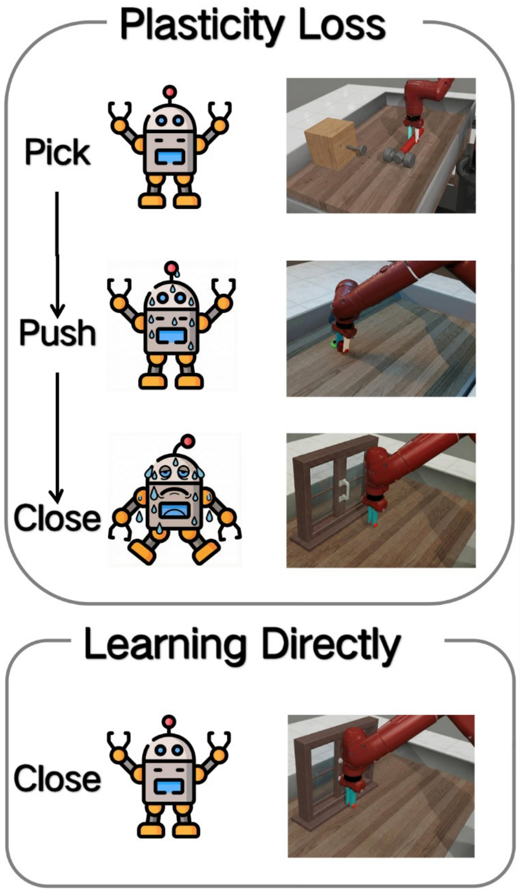
Lirui Luo, Guoxi Zhang, Hongming Xu, Yaodong Yang, Cong Fang, Qing Li

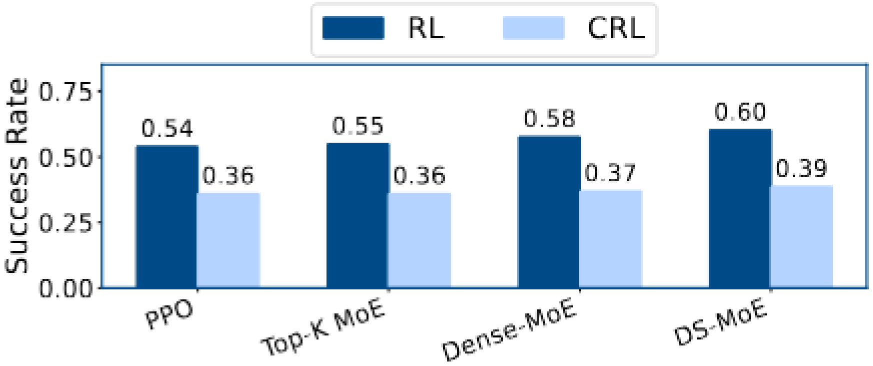
Beijing Institute for General Artificial Intelligence (BIGAI)
Peking University

xxxgithun.io


What is Plasticity Loss

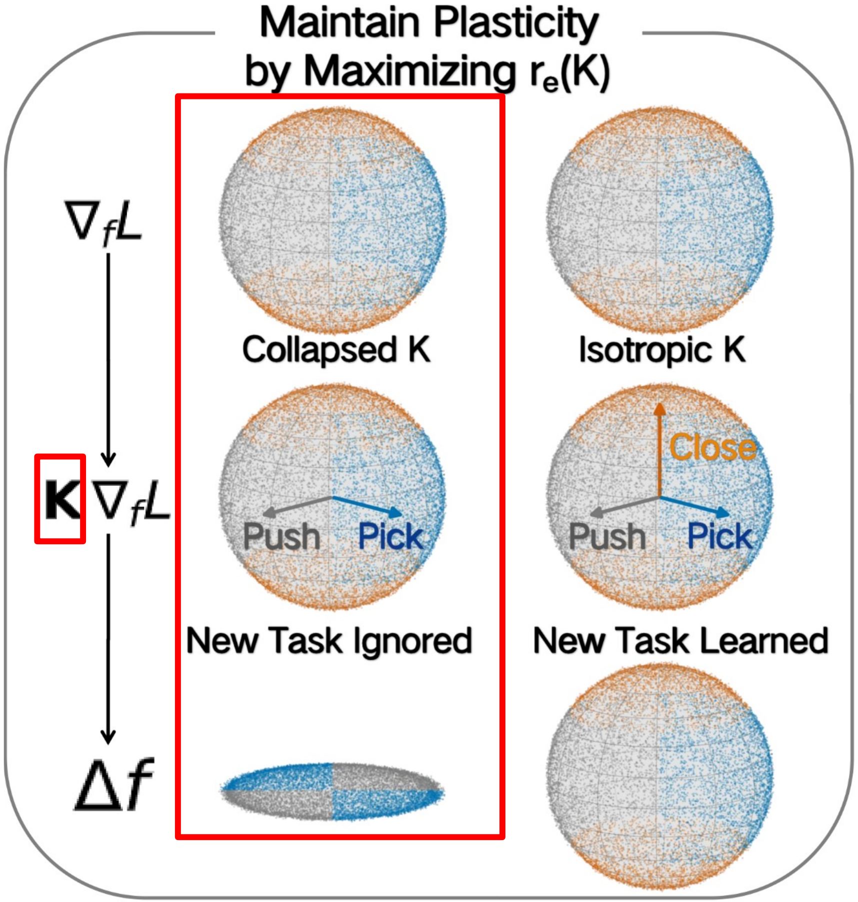




Both MLP and MOE lose the ability to learn from new experiences.

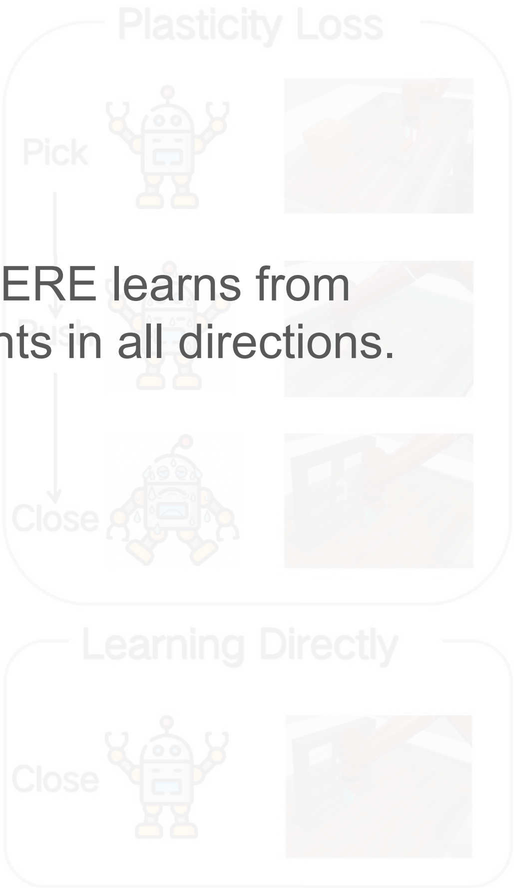
What is the Loss of Spectral Plasticity?

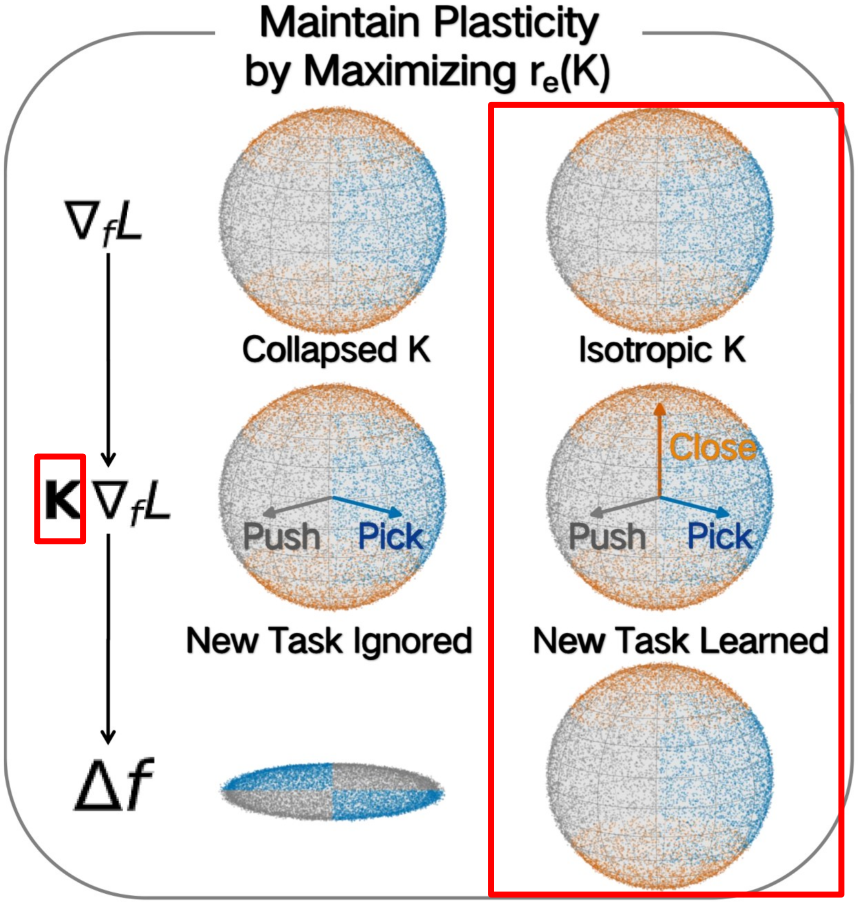
Low effective rank of the empirical Neural Tangent Kernel (eNTK) restricts updates to a few directions.



How SPHERE Mitigating the Loss of Spectral Plasticity?

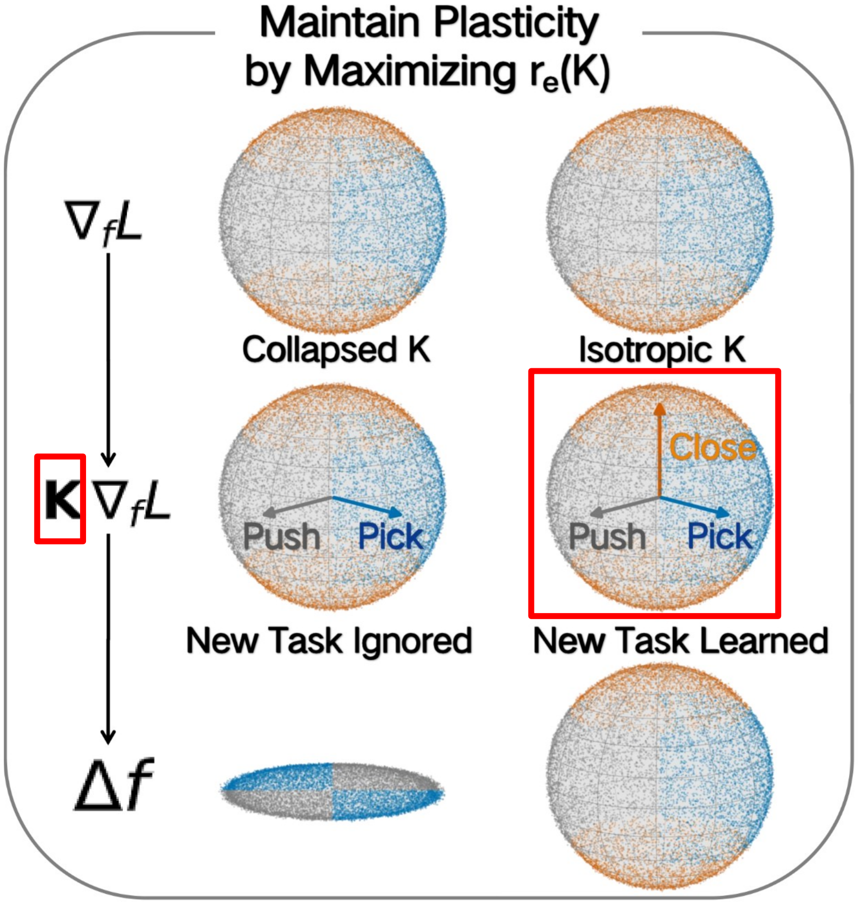
SPHERE learns from gradients in all directions.





How SPHERE Mitigating the Loss of Spectral Plasticity?

SPHERE learns from gradients in all directions.



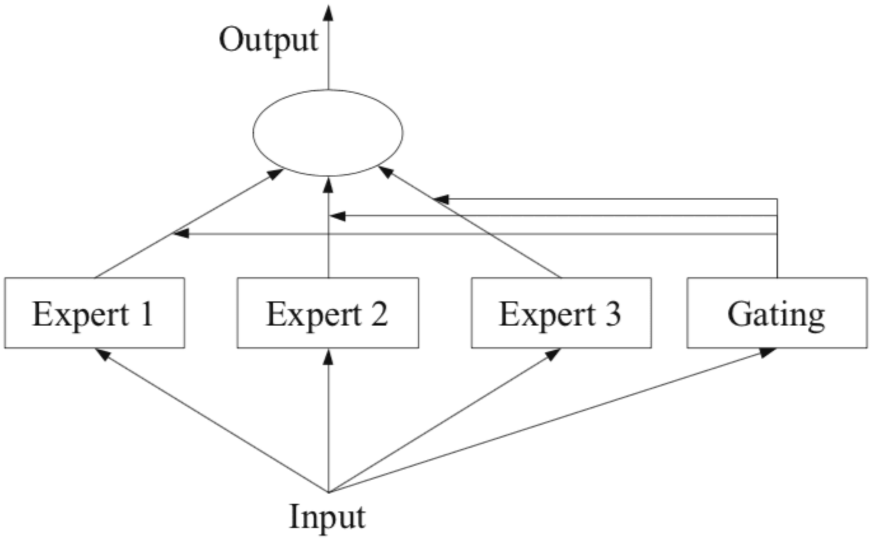
Limitation for Construct K

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

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

$S(J)=O(NP)$

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



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

Decompose K into smaller matrices

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

$$\downarrow$$

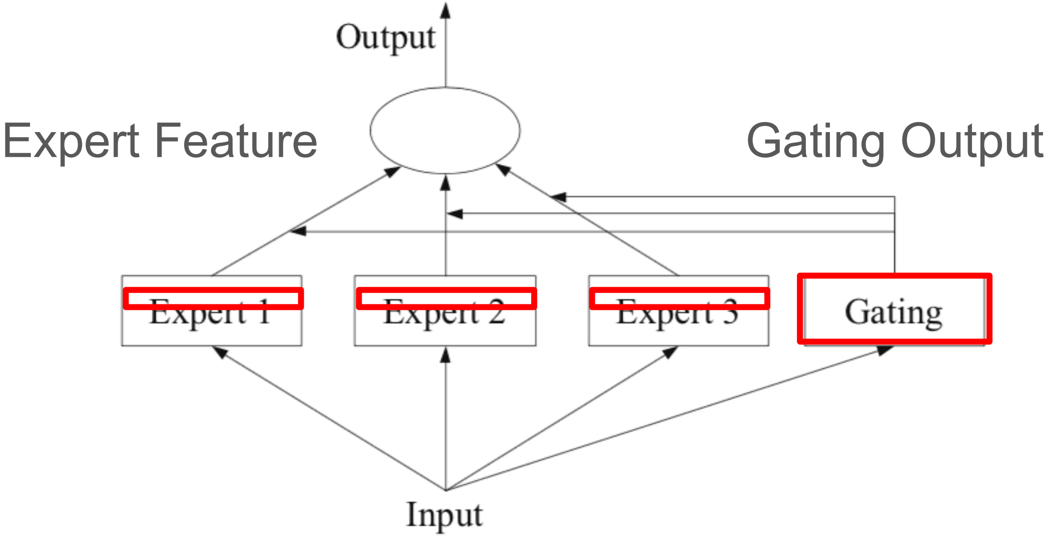
$$G^{GN} = J^{\top}J \in R^{P\times P}$$

$$\downarrow$$

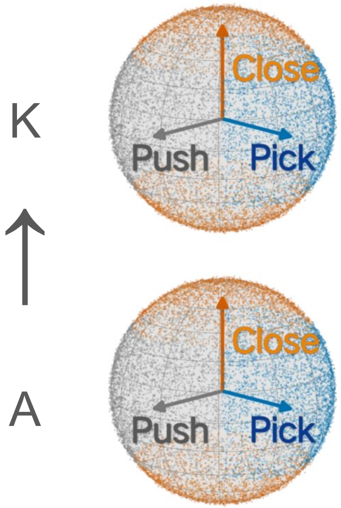
$$A=\varphi^{\top}\varphi$$

Can be constructed from a forward pass without addition backpropagation.

φ=[feature$_1$*gating$_1$, …, feature$_n$*gating$_n$]

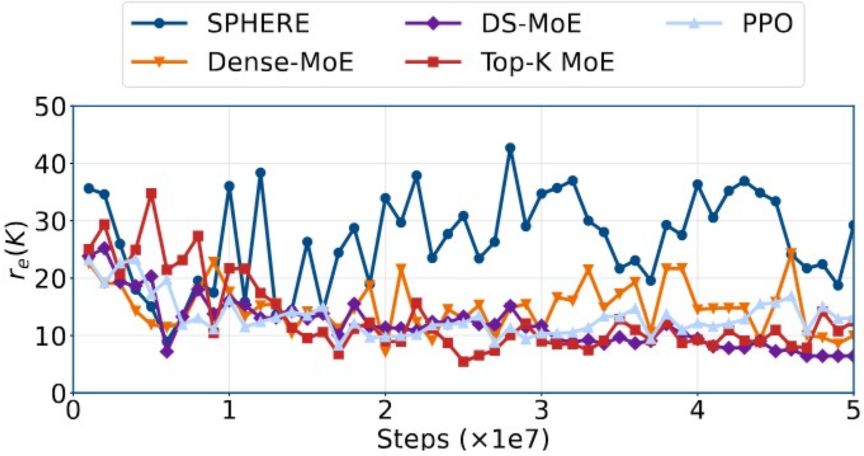


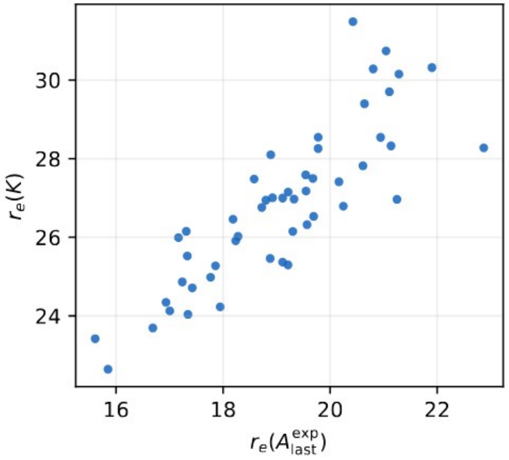
What Does SPHERE Do?

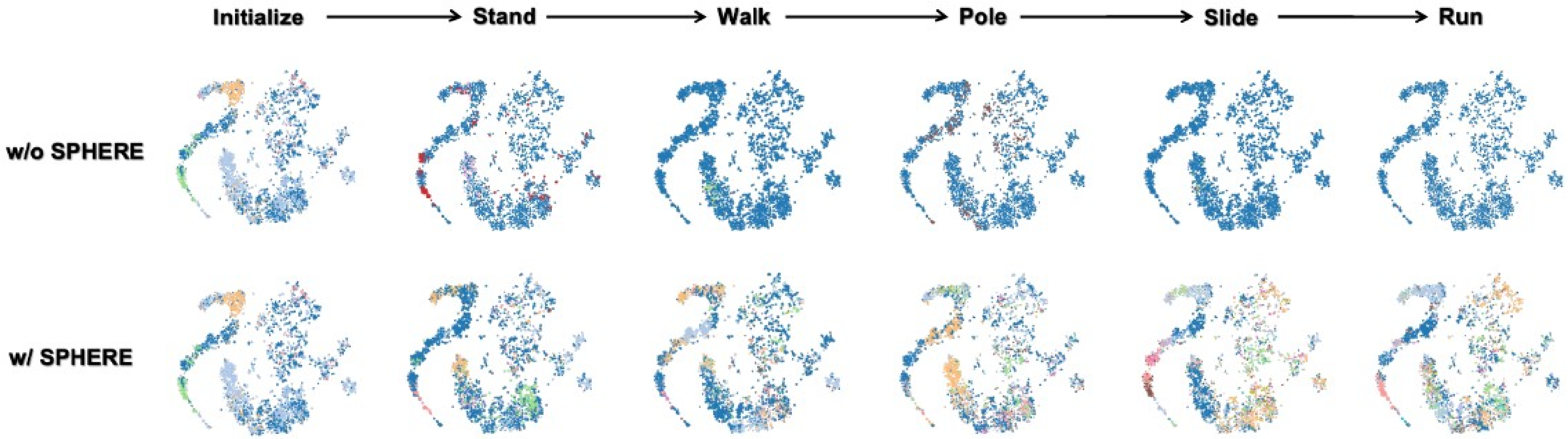


The spectrum of K is uniformized by the spectrum of uniform A.

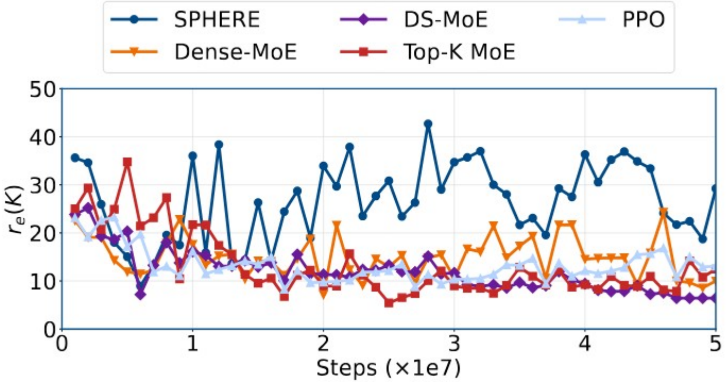
SPHERE Improves The Spectral Plasticity



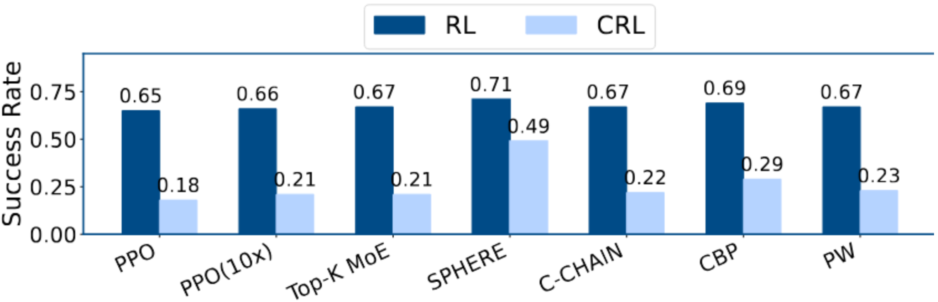


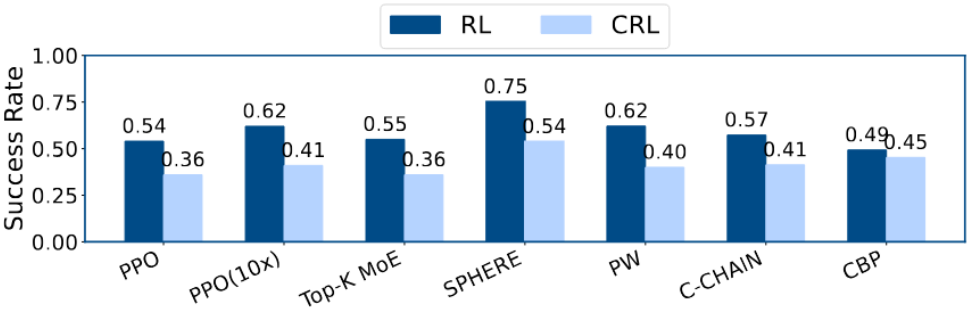


Higher Spectral Plasticity Helps Learning



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 with negligible computation overhead.

Code & demo: xxx

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