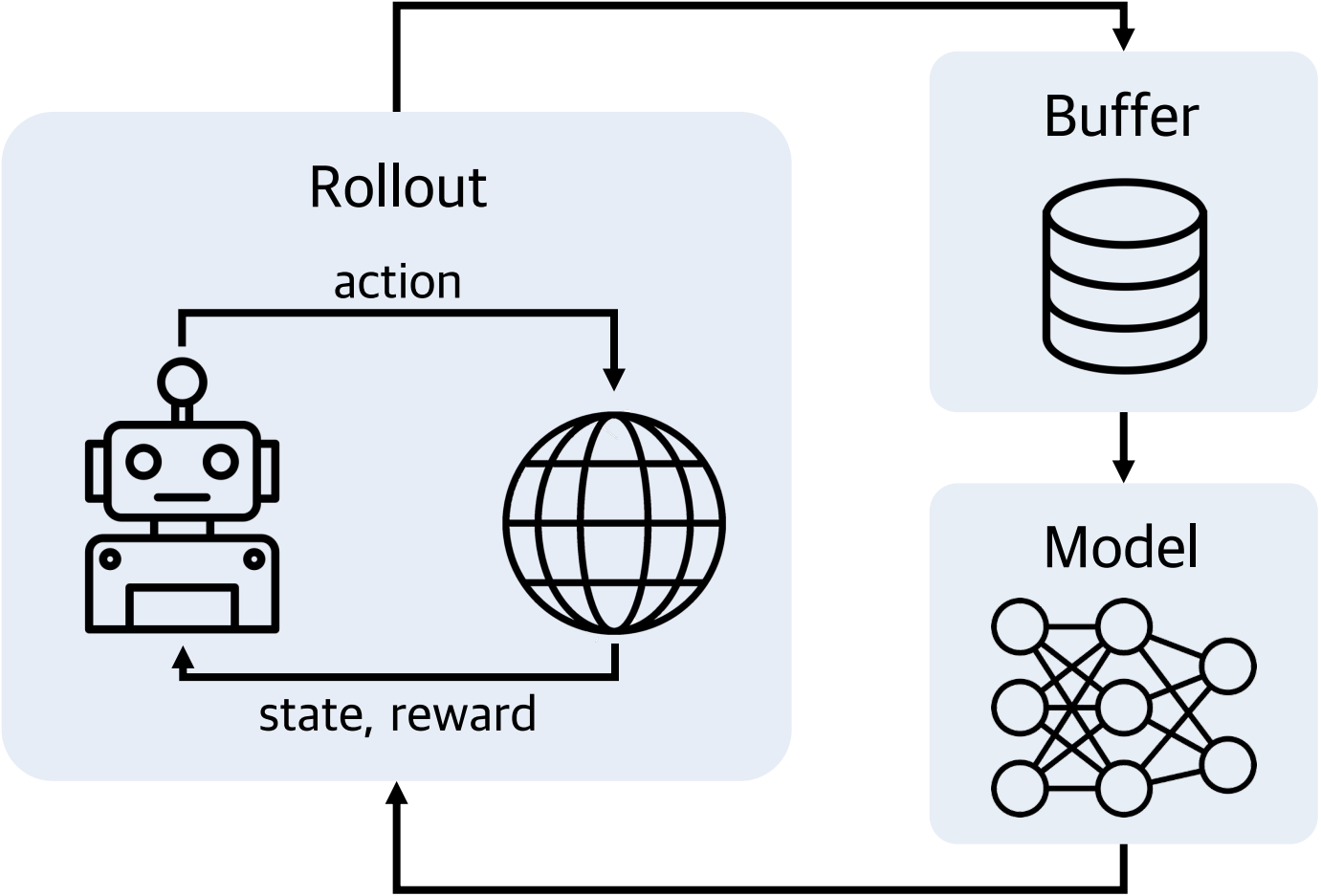# PLASTIC: Improving Input and Label Plasticity for Sample Efficient Reinforcement Learning

Hojoon Lee*, Hanseul Cho*, Hyunseung Kim*, Daehoon Gwak, Joonkee Kim,
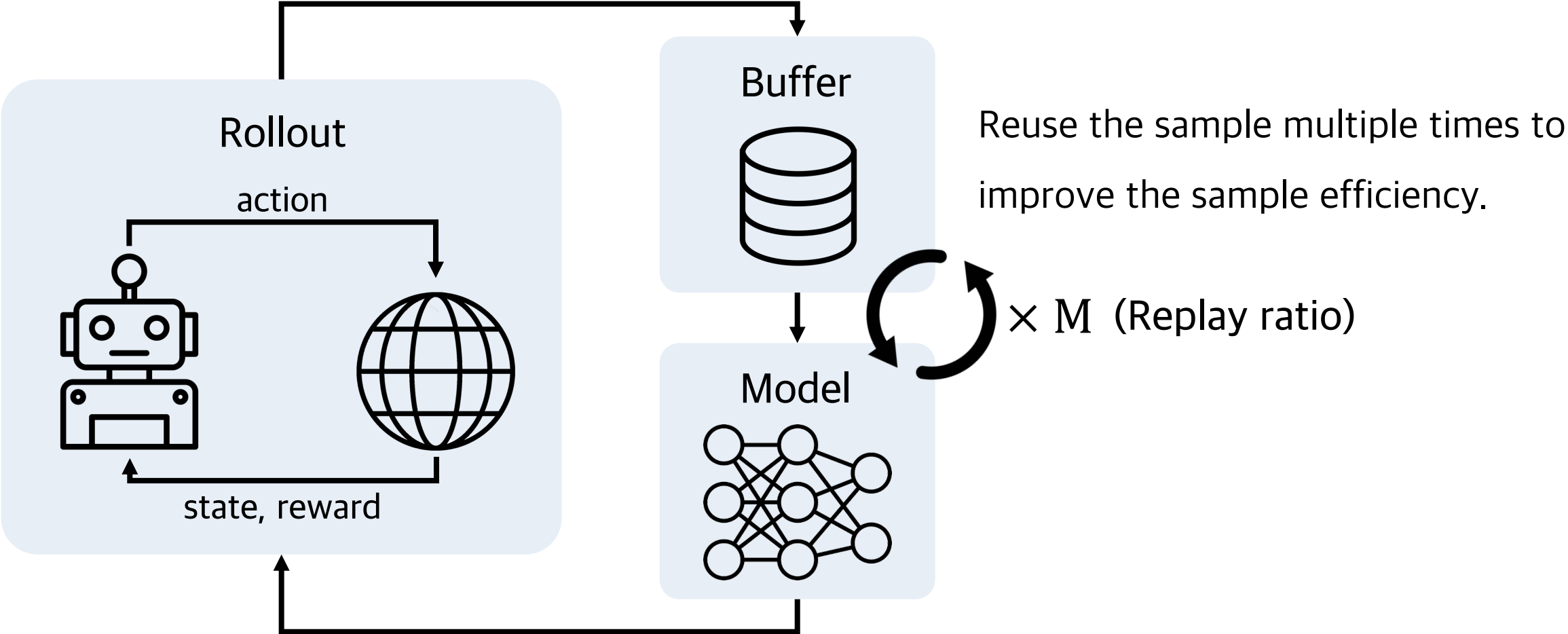
Jaegul Choo, Se-Young Yun, Chulhee Yun

KAIST AI
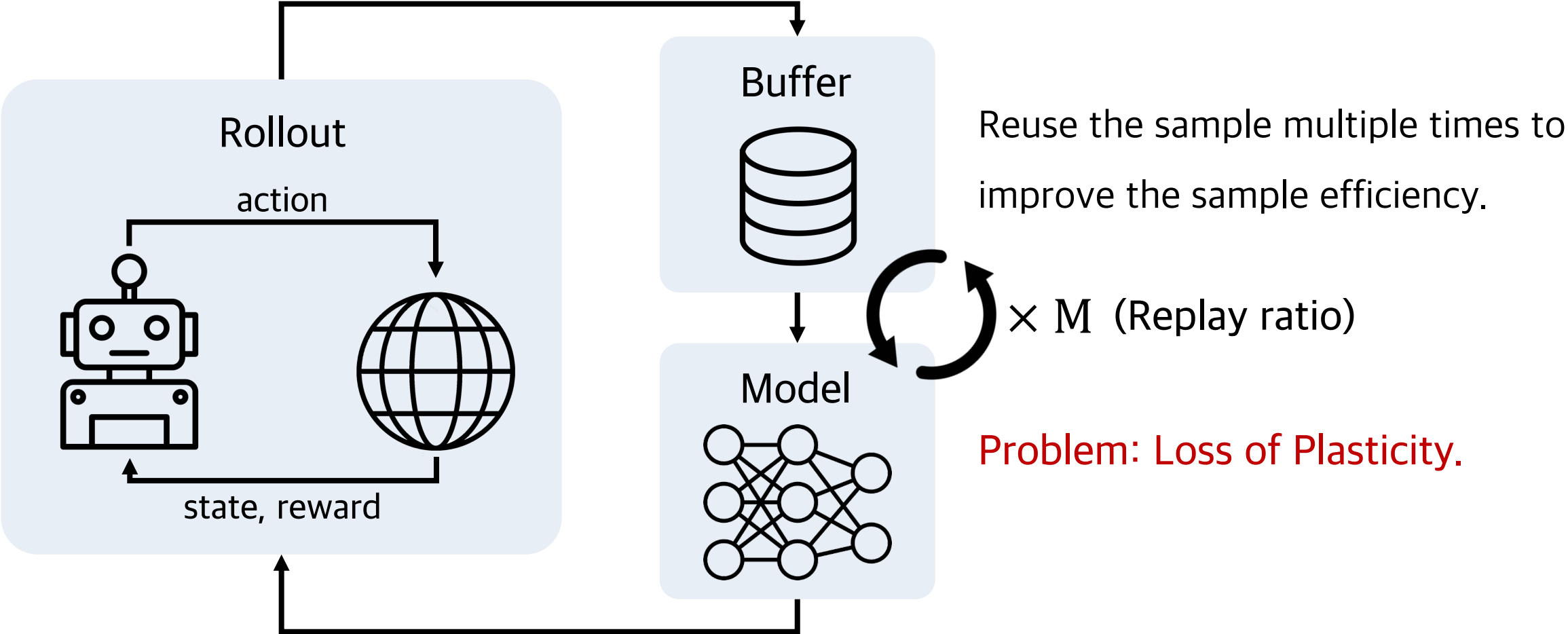Kim Jaechul Graduate School

NEURAL INFORMATION
PROCESSING SYSTEMS

Off-Policy RL for Improving Sample-Efficiency

Off-Policy RL

Off-Policy RL for Improving Sample-Efficiency

Off-Policy RL

Off-Policy RL for Improving Sample-Efficiency

Off-Policy RL

Buffer

Rollout

action

state, reward

Model

Reuse the sample multiple times to improve the sample efficiency.

$\times$ M (Replay ratio)

Problem: Loss of Plasticity.

Loss of Plasticity in Off-Policy RL

Definition

The model overfits to earlier data and loses its ability to adapt to new datasets.

Loss of Plasticity in Off-Policy RL

Definition

The model overfits to earlier data and loses its ability to adapt to new datasets.

Two types of the model's plasticity

- Input plasticity: adaptability to shifts in input distribution, $p(x)$.

- Label plasticity: adaptability to shifts in label for given inputs, $p(y|x)$.

Loss of Plasticity in Off-Policy RL

Definition

The model overfits to earlier data and loses its ability to adapt to new datasets.

Two types of the model's plasticity

- Input plasticity: adaptability to shifts in input distribution, $p(x)$.

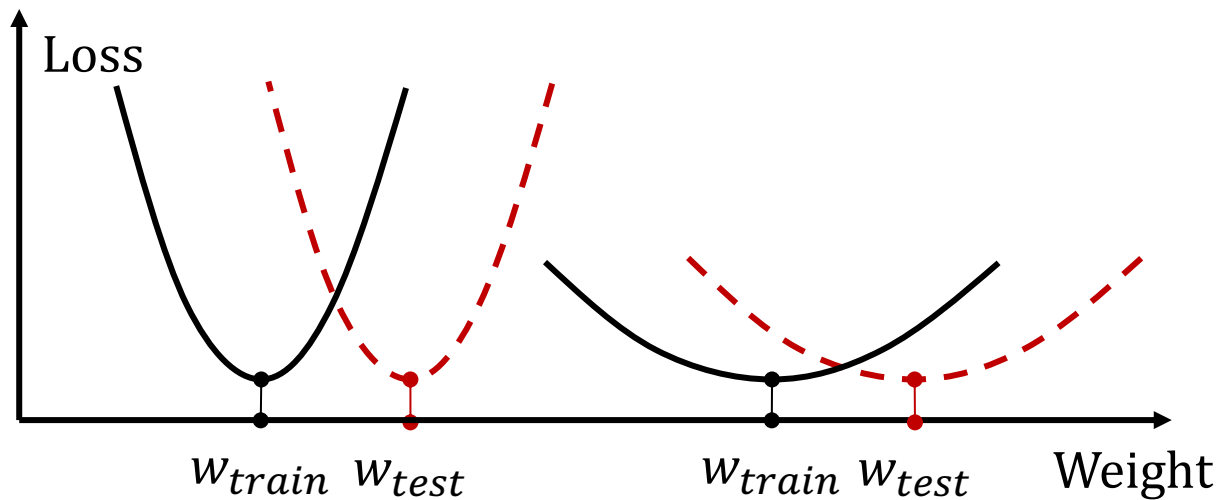- Label plasticity: adaptability to shifts in label for given inputs, $p(y|x)$.

Why multiple updates lead to the loss of plasticity?

- The model converges on a sharp region of loss surface, which is sensitive to input shift.

- The number of active units decreases, which impedes updates to label shift.

PLASTIC: Jointly Preserving Input & Label Plasticity

How to preserve the model's plasticity?

- Input plasticity: Seek smooth region of loss surface with LayerNorm / SAM optimizer.

- Label plasticity: Preserve the active units with CReLU activation / Last-layer Reset.

- PLASTIC = Integrating (LN / SAM) with (CReLU / Reset) to Off-policy RL algorithm.
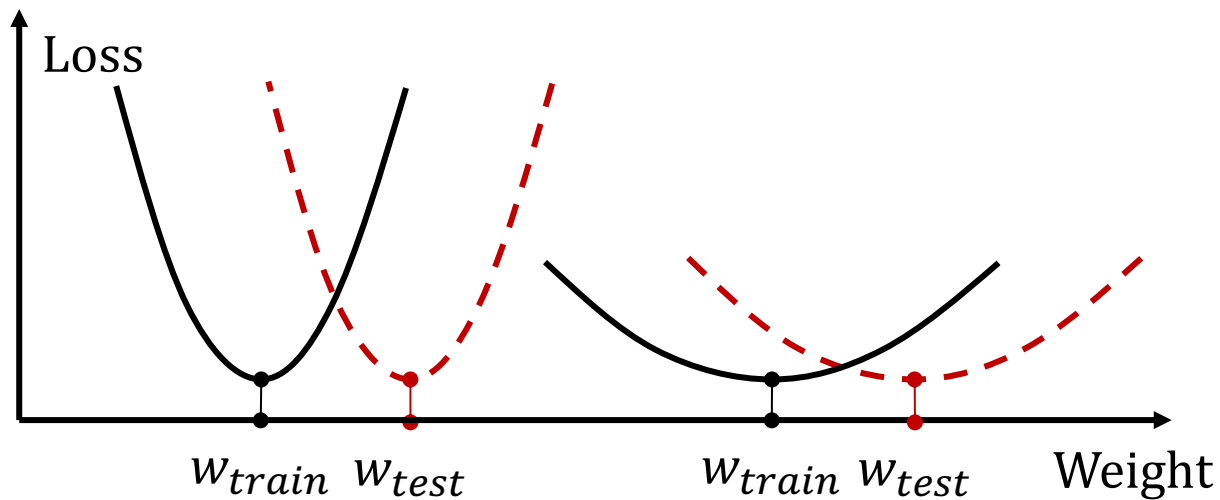

Seeking smoother region of loss landscape facilitates
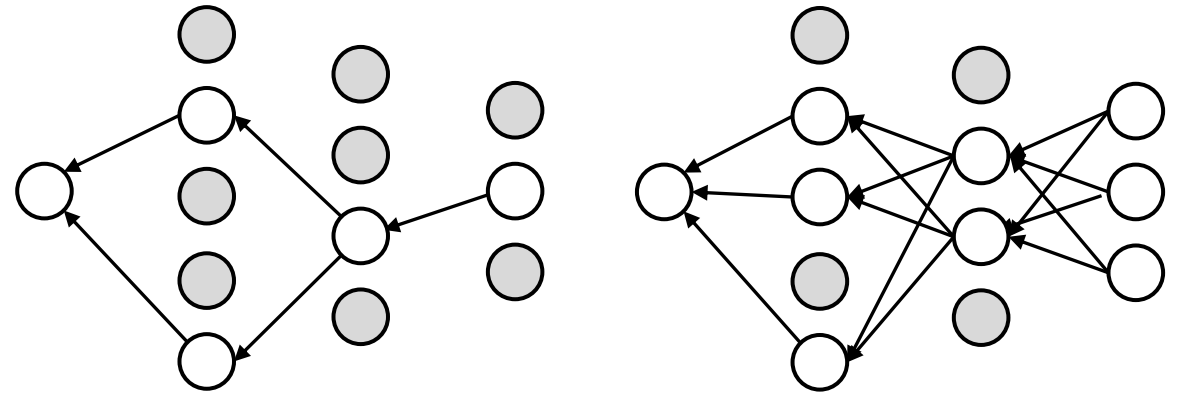stability to input shift, improving the input plasticity.

PLASTIC: Jointly Preserving Input & Label Plasticity

How to preserve the model's plasticity?

- Input plasticity: Seek smooth region of loss surface with LayerNorm / SAM optimizer.

- Label plasticity: Preserve the active units with CReLU activation / Last-layer Reset.

- PLASTIC = Integrating (LN / SAM) with (CReLU / Reset) to Off-policy RL algorithm.

Seeking smoother region of loss landscape facilitates stability to input shift, improving the input plasticity.
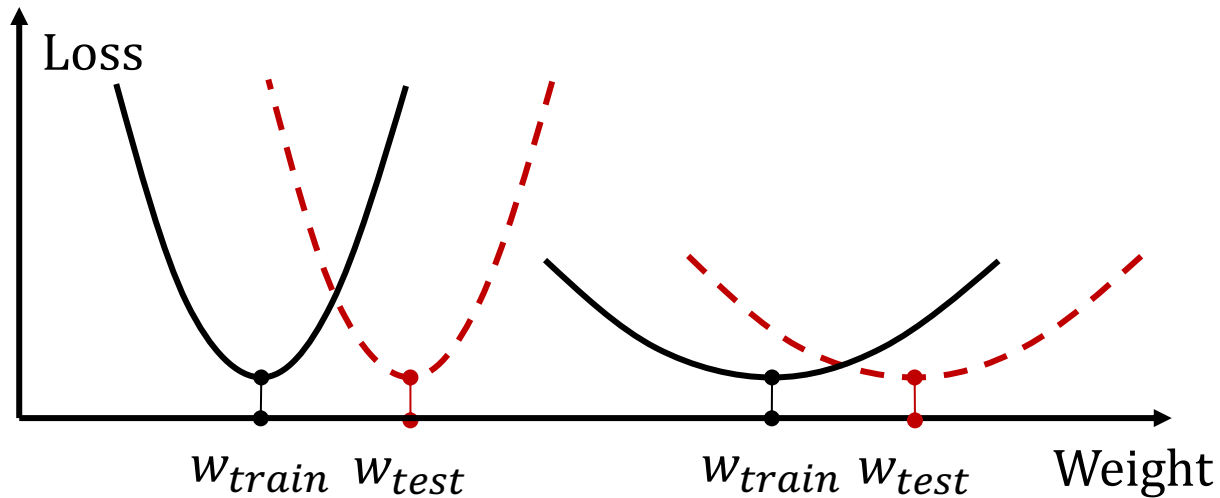
Preserving the fraction of active units facilitates gradient propagation, improving the label plasticity.
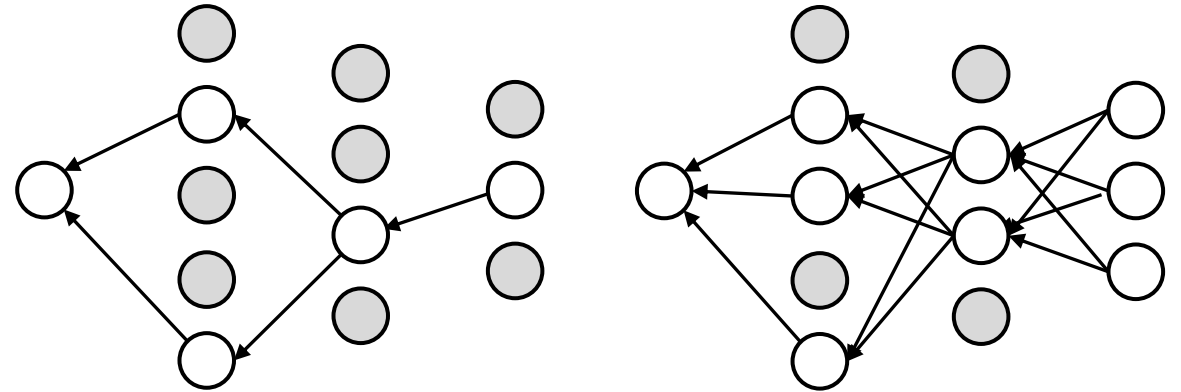
PLASTIC: Jointly Preserving Input & Label Plasticity

How to preserve the model's plasticity?

- Input plasticity: Seek smooth region of loss surface with LayerNorm / SAM optimizer.

- Label plasticity: Preserve the active units with CReLU activation / Last-layer Reset.

- **PLASTIC = Integrating (LN / SAM) with (CReLU / Reset) to Off-policy RL algorithm.**

Seeking smoother region of loss landscape facilitates stability to input shift, improving the input plasticity.

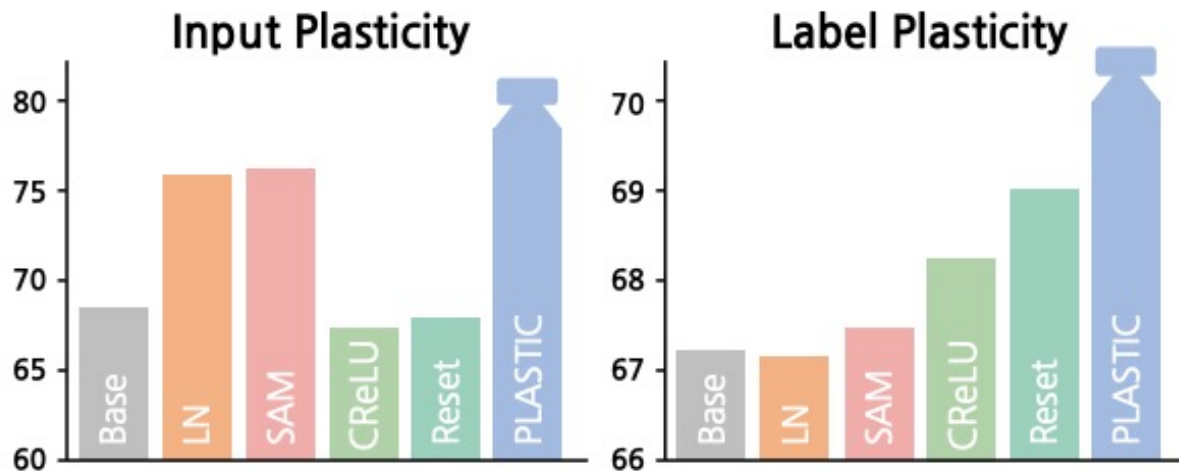Preserving the fraction of active units facilitates gradient propagation, improving the label plasticity.

Experiments

Setup

- **Synthetic:** CIFAR-10 dataset with 100 input / label distribution shifts.

- Atari-100k: 26 Atari games with 100k interactions. Use Rainbow algorithm.

- DMC-Medium: 11 medium level control tasks with 2M interactions. Use SAC algorithm.
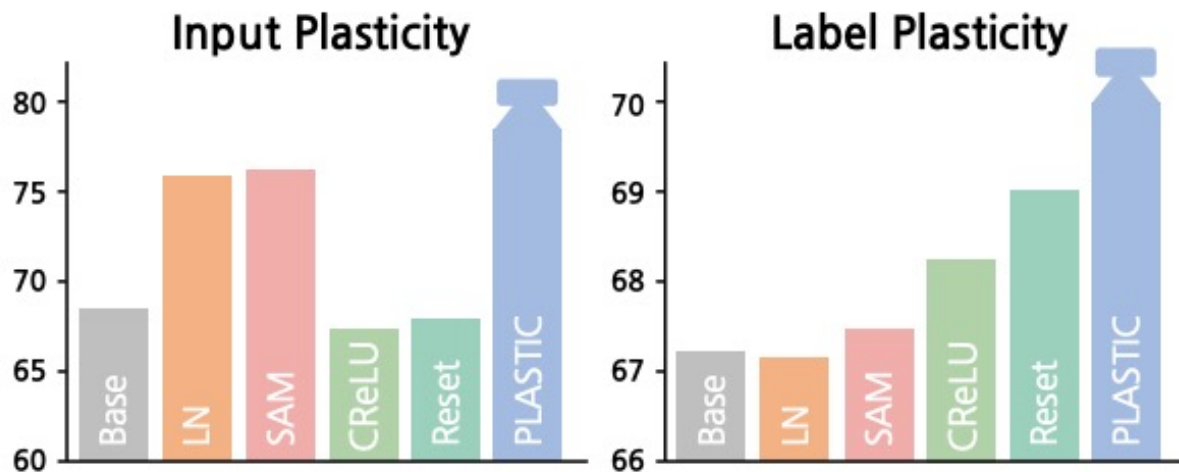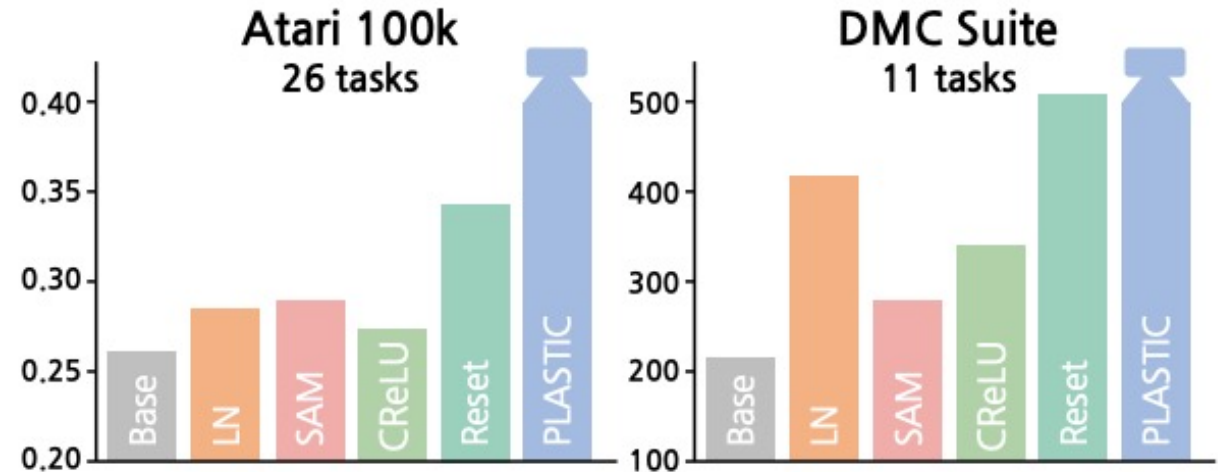
Result

Synthetic Experiment

Experiments

Setup

- **Synthetic:** CIFAR-10 dataset with 100 input / label distribution shifts.

- **Atari-100k:** 26 Atari games with 100k interactions. Use Rainbow algorithm.

- **DMC-Medium:** 11 medium level control tasks with 2M interactions. Use SAC algorithm.
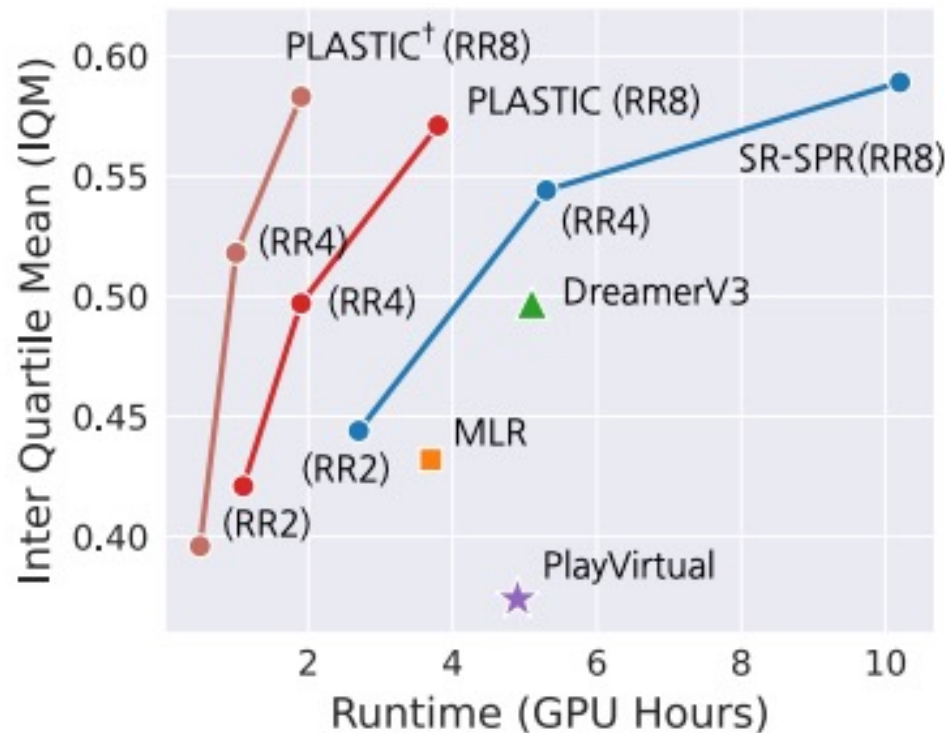
Result

Experiments

Setup

- **Scaling:** Increasing the Replay Ratio (# of updates per data collection) in Atari-100k.

- Pretrained: Fine-tuning on Atari-100k, using a large, pre-trained model.
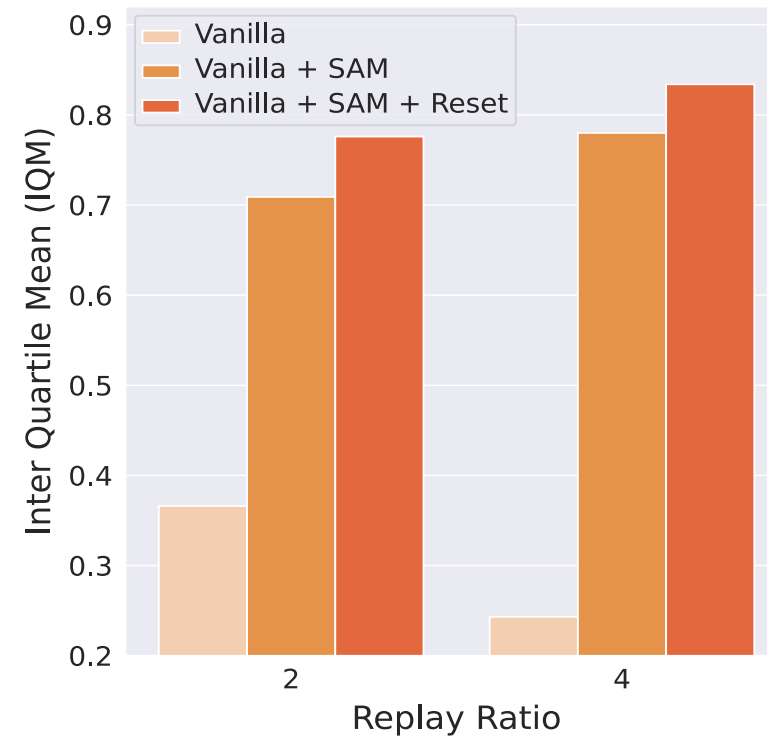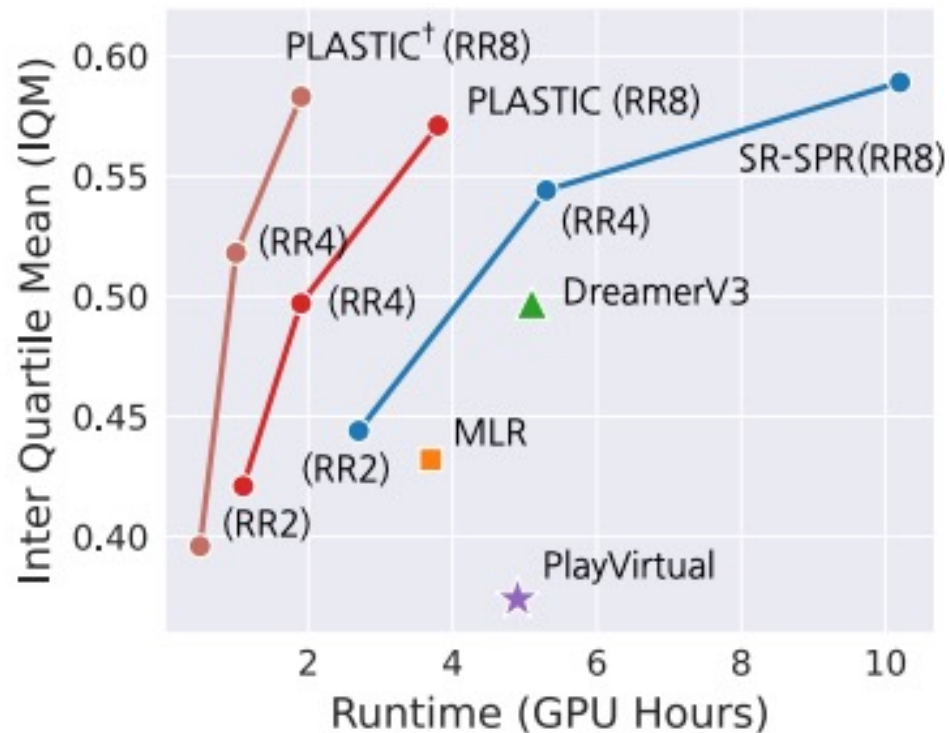
Result

Experiments

Setup

- **Scaling:** Increasing the Replay Ratio (# of updates per data collection) in Atari-100k.

- **Pretrained:** Fine-tuning on Atari-100k, using a large, pre-trained model.
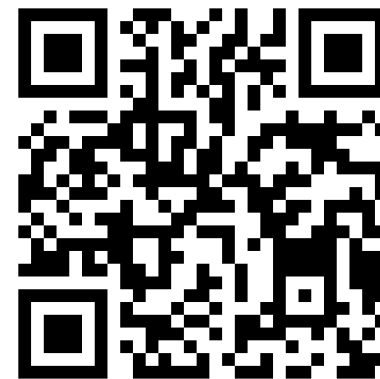
Result

Take-Home Messages

1. In off-policy RL, increasing the replay ratio often lead to the loss of plasticity.

2. The loss of plasticity can be categorized by loss of input and label plasticity.

    - Seeking a smooth region of loss surface improves the input plasticity.

    - Preserving a fraction of active units improves the label plasticity.

3. By jointly improving both plasticity, we can make sample-efficient RL algorithm!

See you at Poster Session #4! (Dec 13 Wed)

- Poster location: Great Hall & Hall B1+B2 #1422

Paper (arXiv)   Code (GitHub)