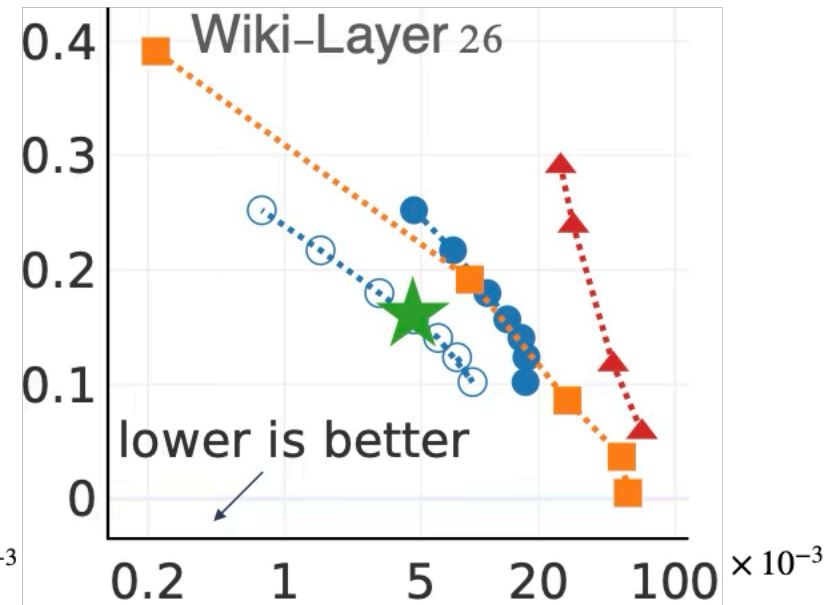
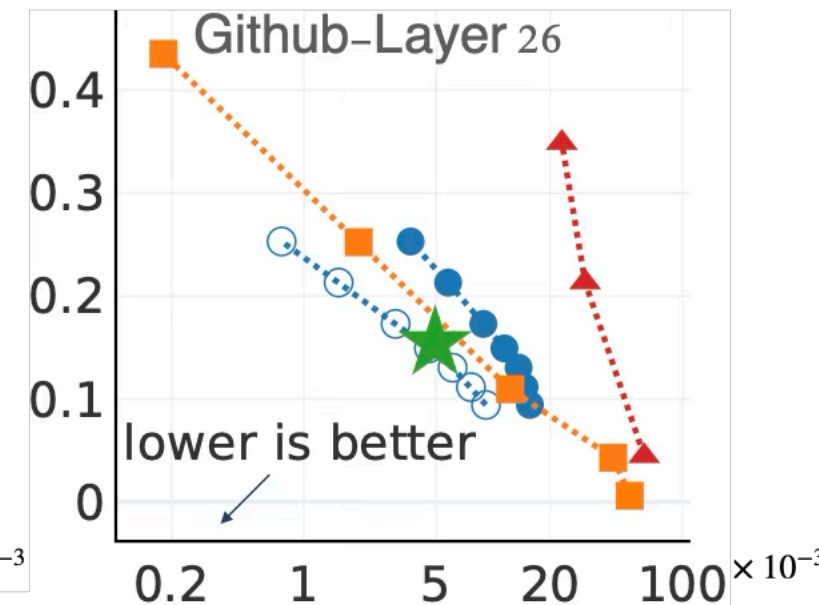
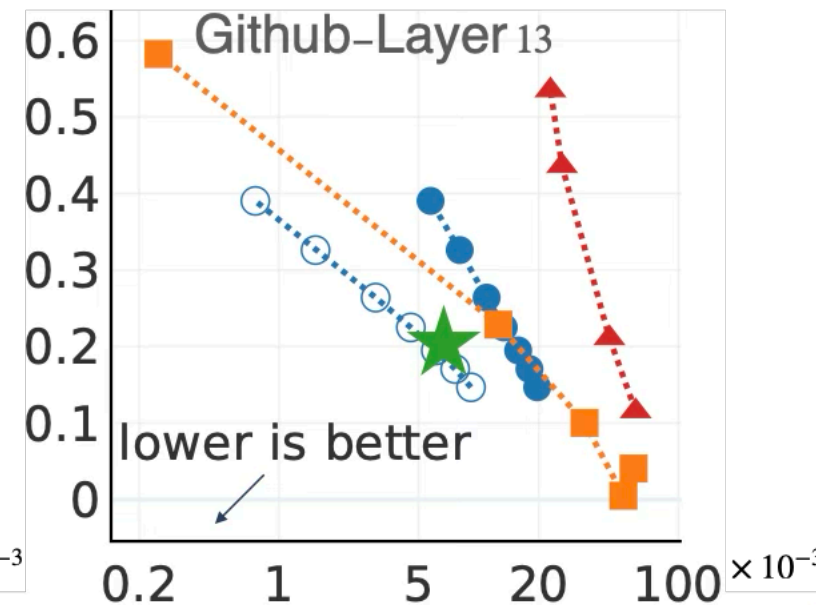
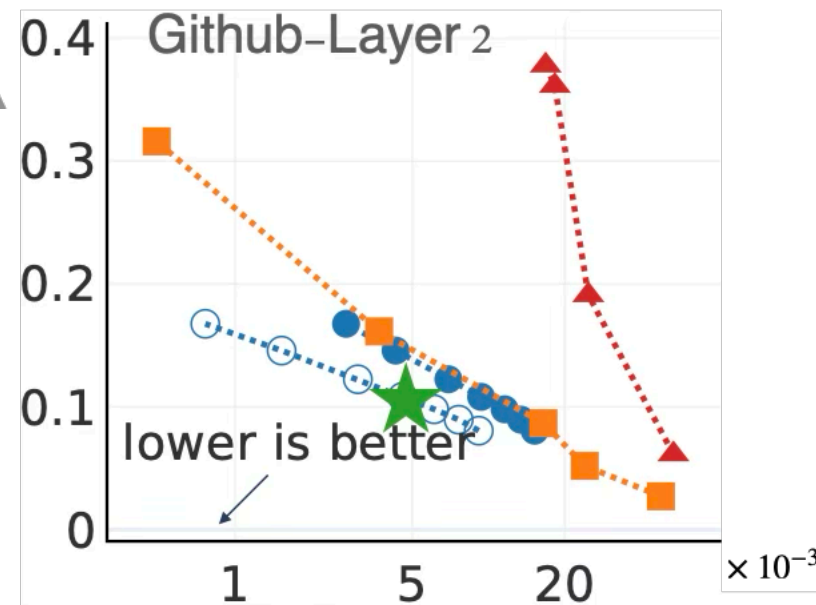


L2 Reconstruction Loss



● TopK (pre-act) ○ TopK (post-act) ■ L1 ★ GBA ▲ BA

Fraction of Activated Neuron