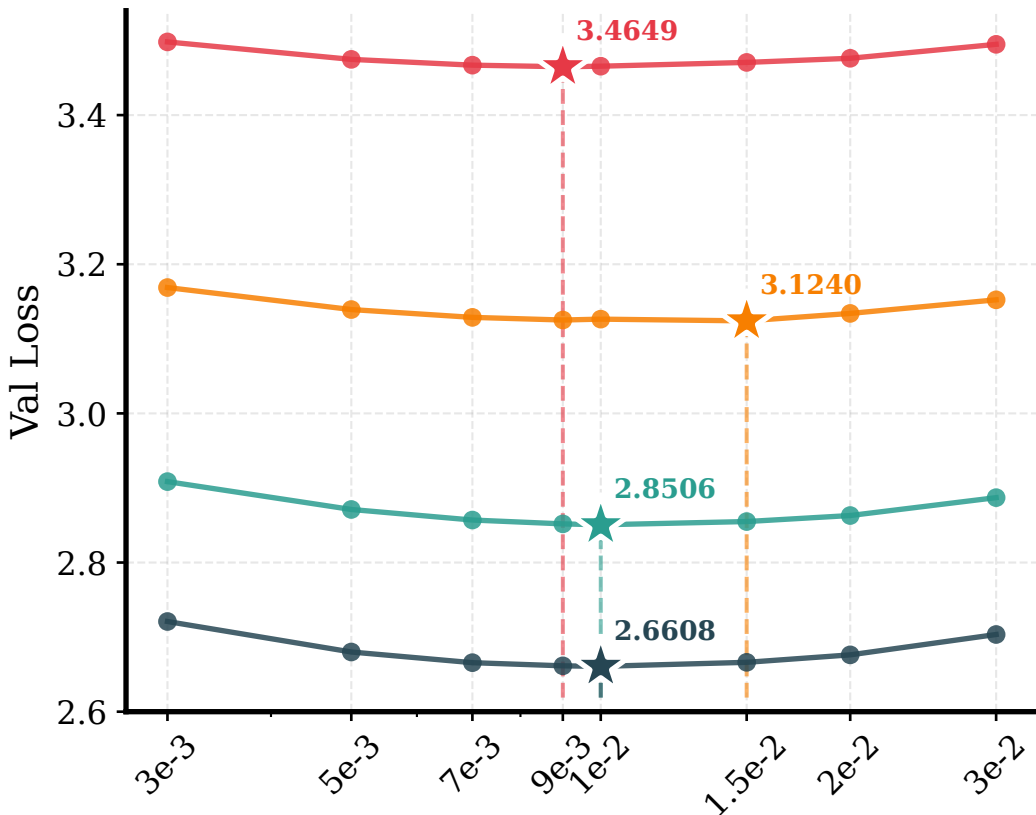Muon Optimizer — Spectral Sphere Optimizer

Val Loss vs LR for widths 256, 512, 1024, 2048.

Muon Optimizer minimum losses: 3.4649, 3.1240, 2.8506, 2.6608

Spectral Sphere Optimizer minimum losses: 3.4578, 3.1119, 2.8383, 2.6481

Legend: width=256, width=512, width=1024, width=2048, Min Loss