LLaMA 350M 55 Methods POET (b=4)POET (b=2)50 GaLore AdamW 45 Validation Perplexity 40 30 25 20 0 20K 60K 80K 160K 40K 100K 120K 140K

Step