Language Model(774M) loss curve GPT-method2 3.4 LLaMA-method2 3.2 Validation Loss 2.8 2.6 0 20 40 60 80 100 Number of iterations(k)