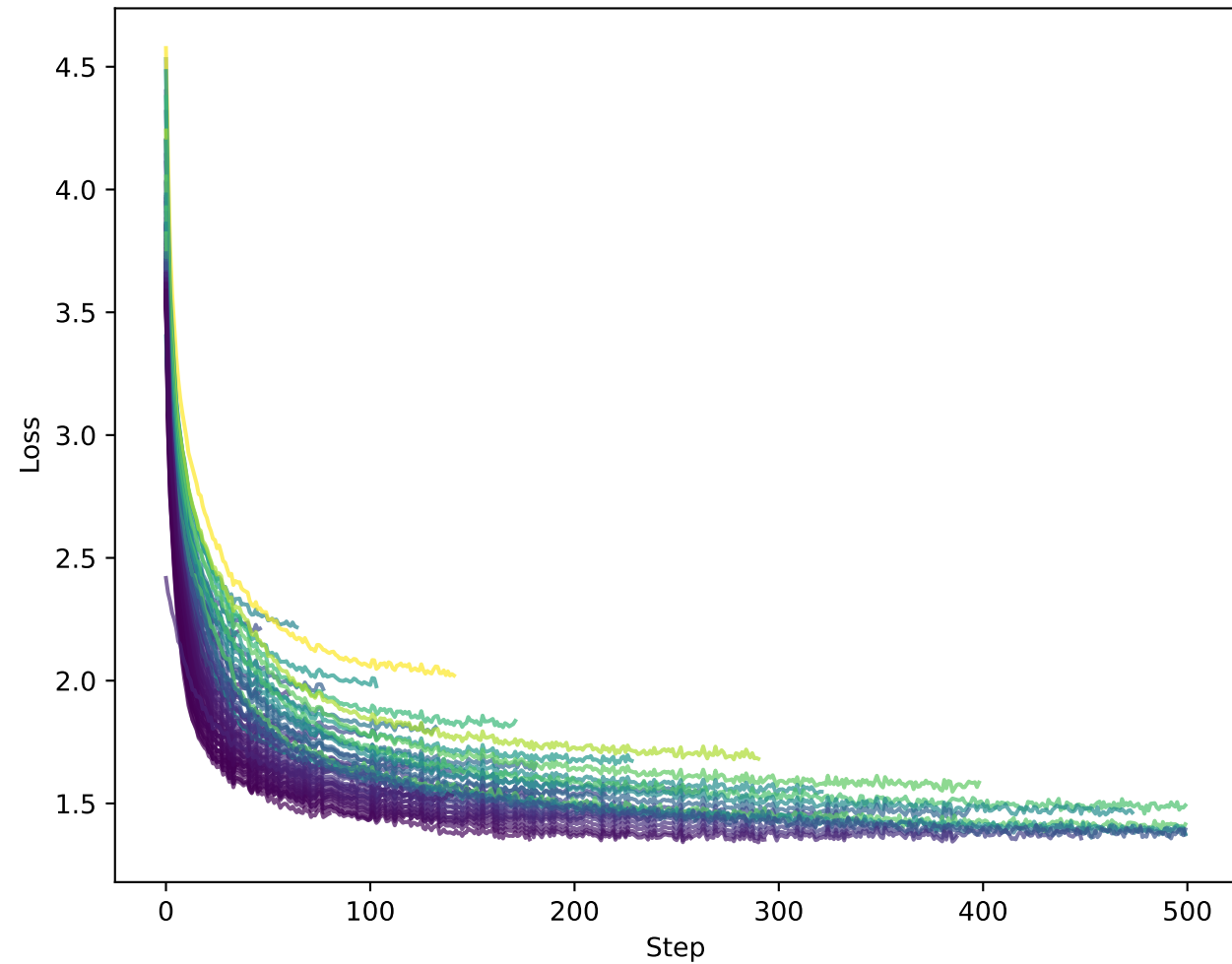


bfloat16 weights, bfloat16 activations

Train CrossEntropyLoss



Optimizer Total Grad Norm

