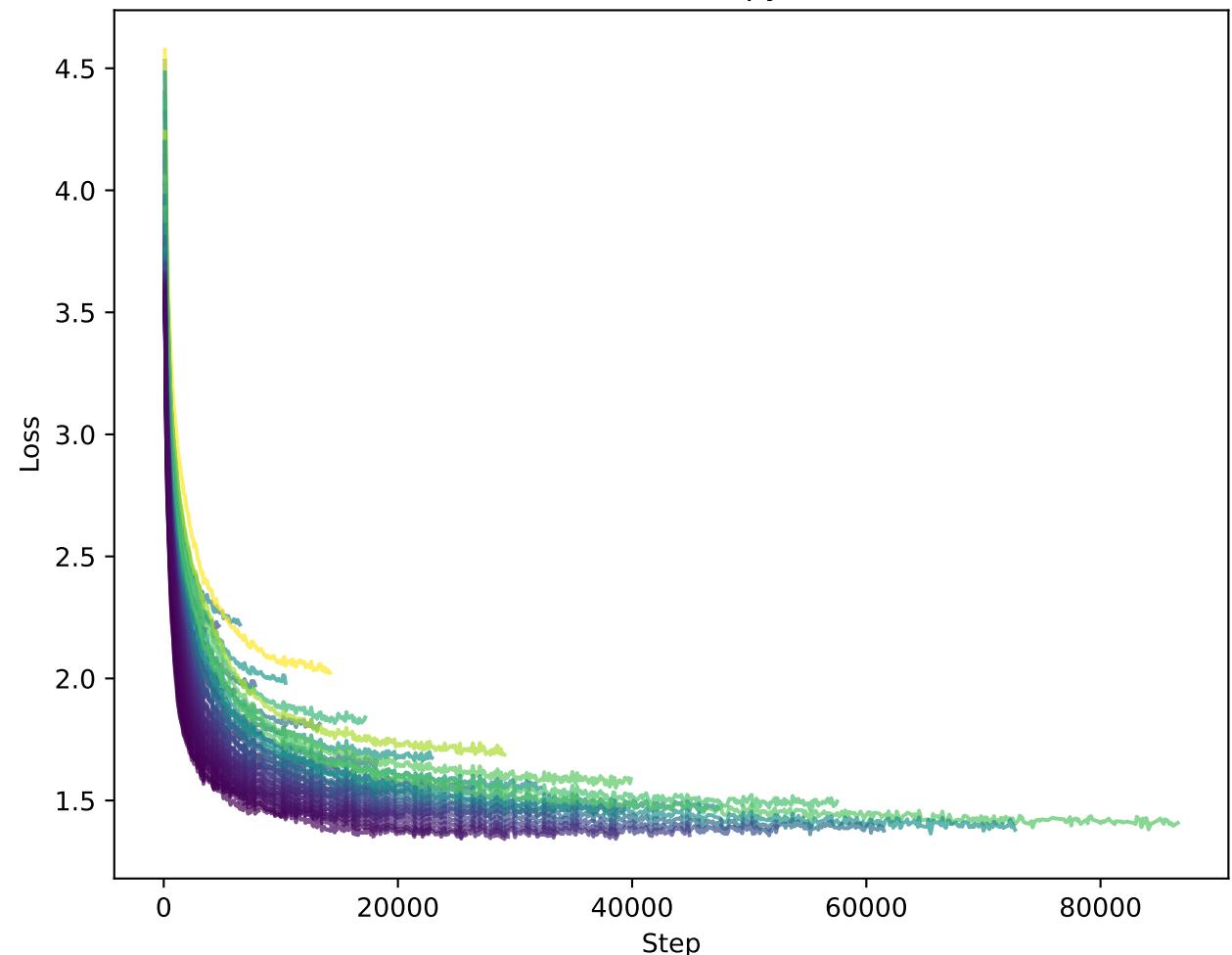


MXFP8 E4M3 weights, bfloat16 activations

Train CrossEntropyLoss



Optimizer Total Grad Norm

