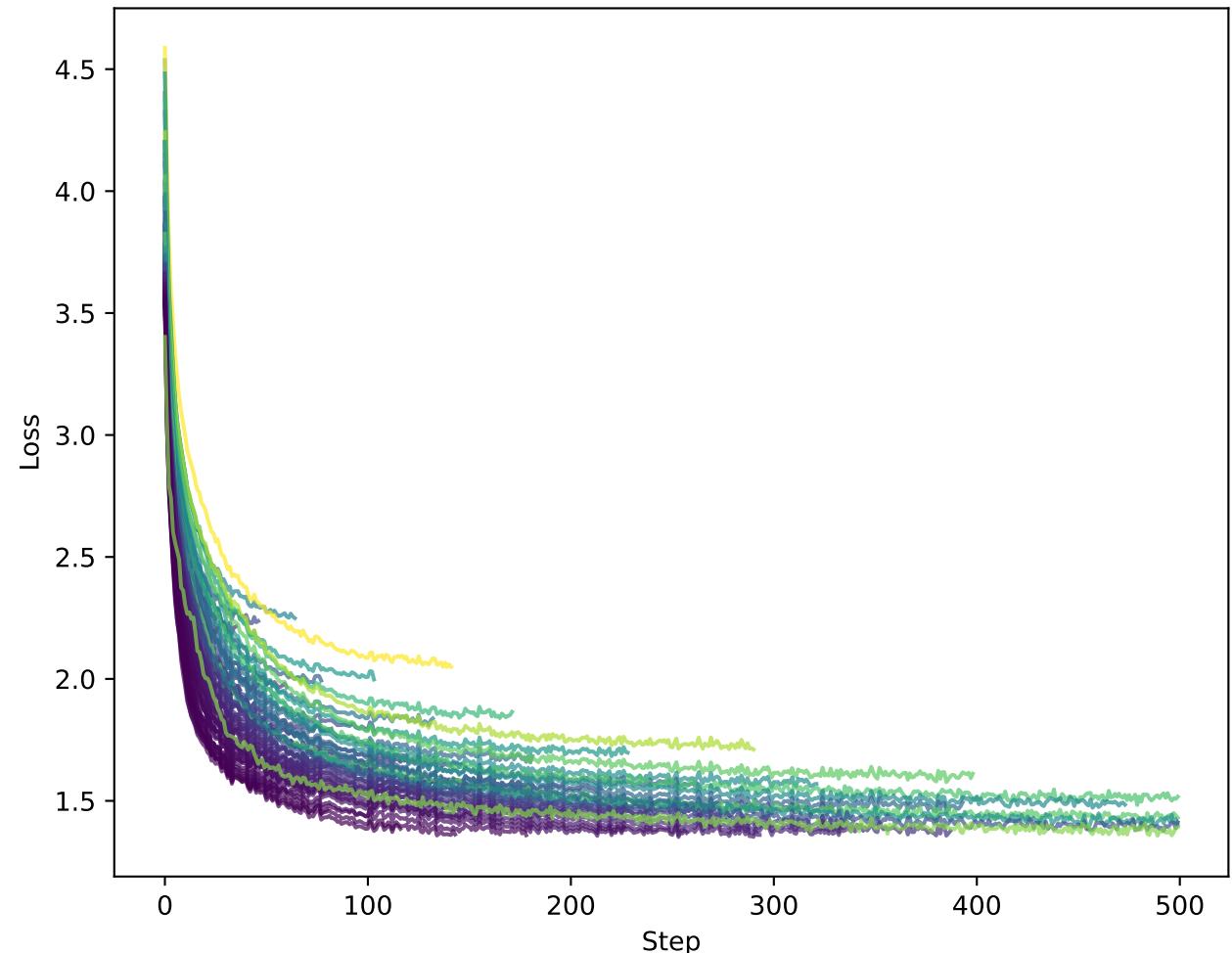


MXFP8 (E5M2 weights, E5M2 activations) Quantize Forward Only

Train CrossEntropyLoss



Optimizer Total Grad Norm

