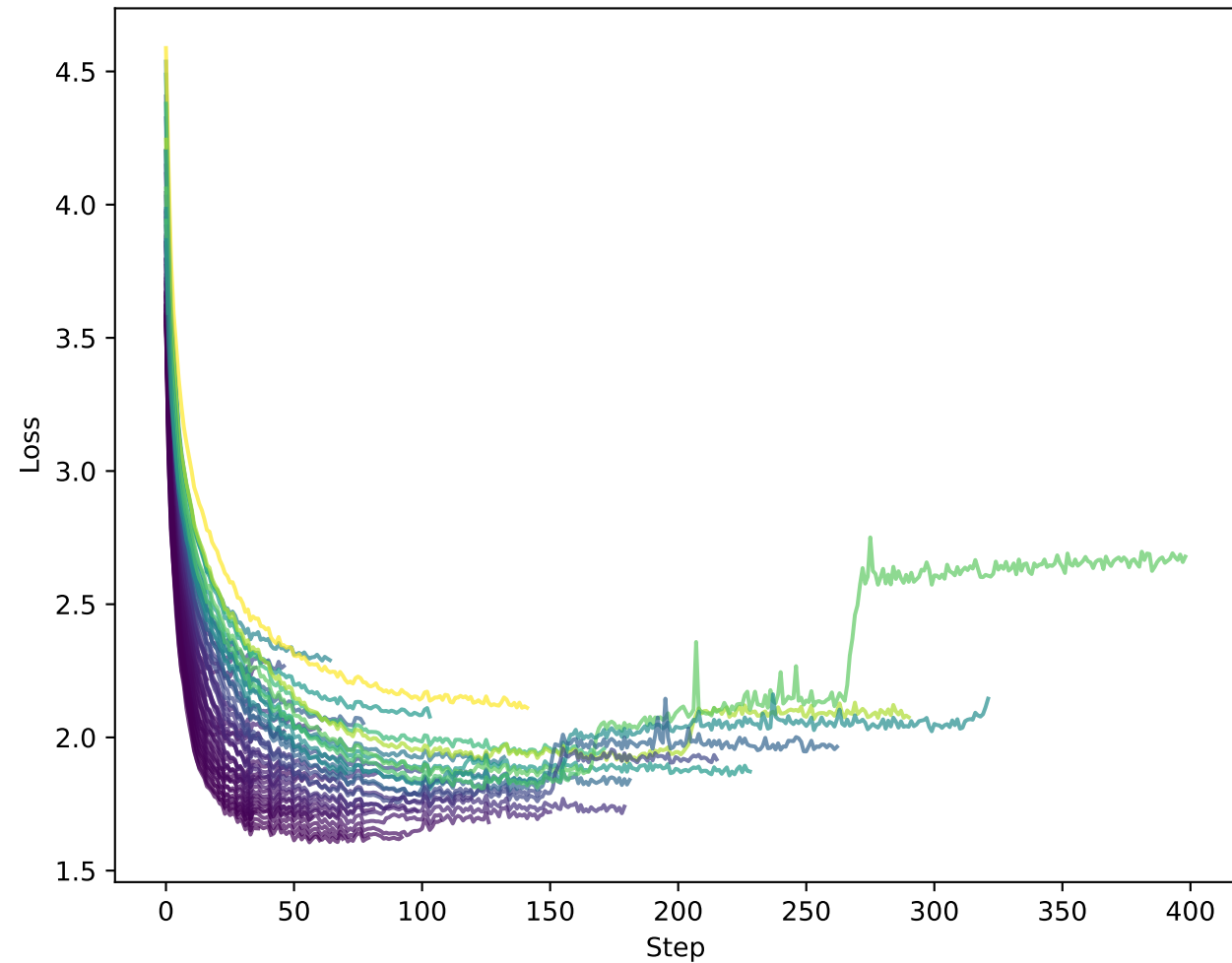


# MXFP6 E2M3 weights, MXFP8 E5M2 activations

## Train CrossEntropyLoss



## Optimizer Total Grad Norm

