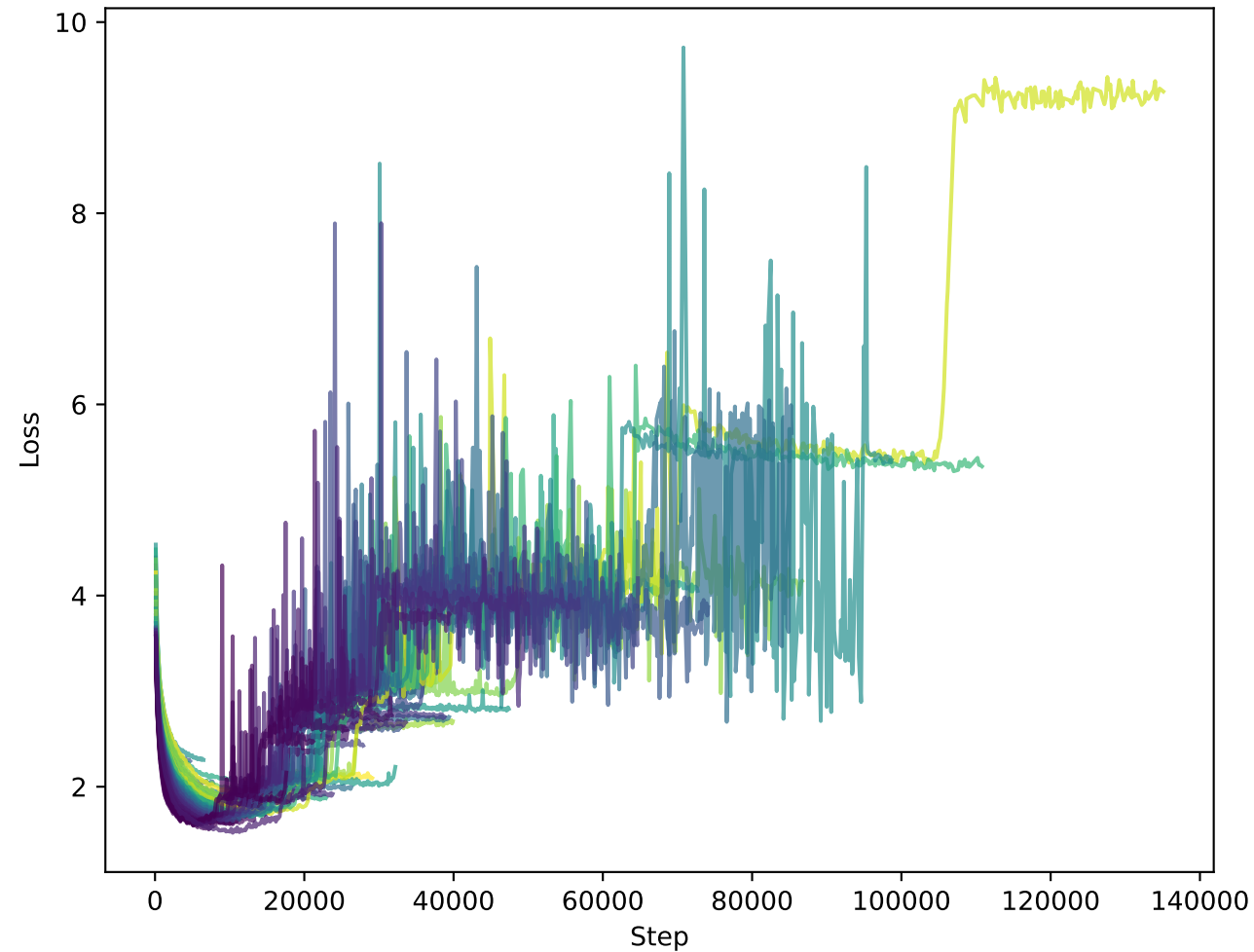


# MXFP8 E5M2 weights, MXFP8 E4M3 activations

## Train CrossEntropyLoss



## Optimizer Total Grad Norm

