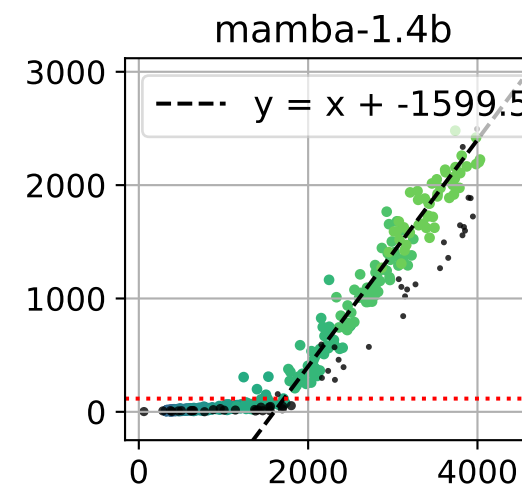
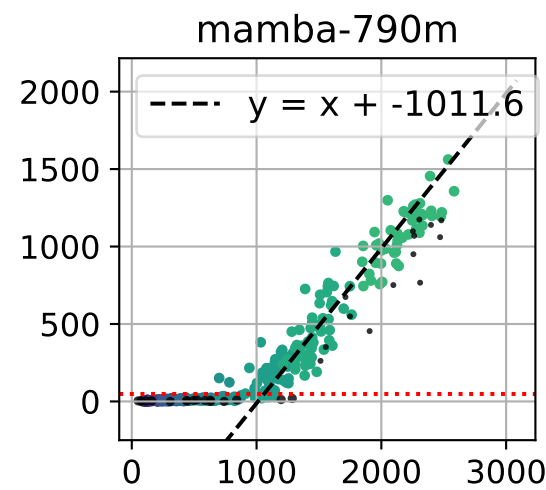
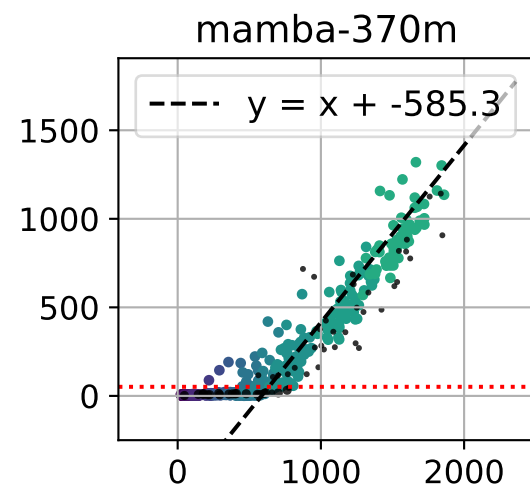
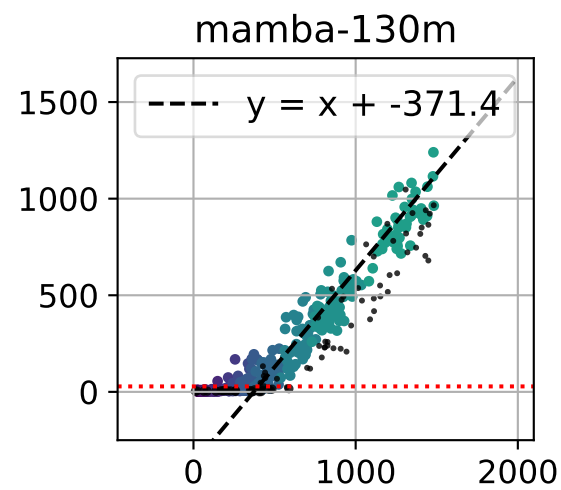
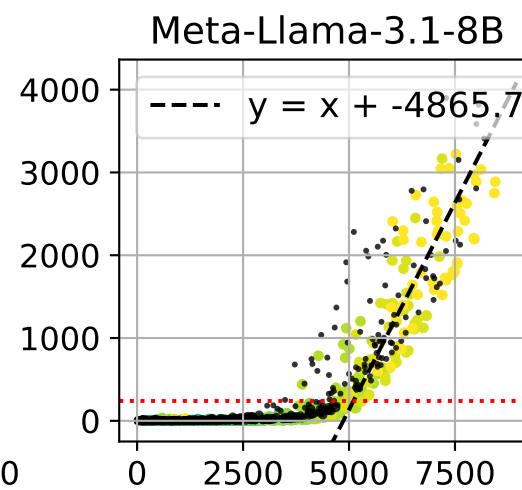
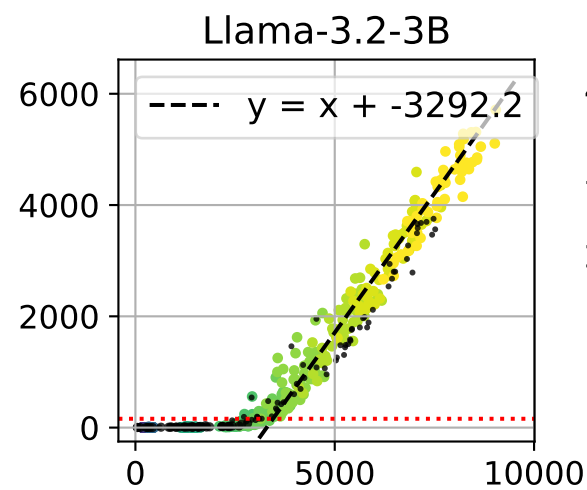
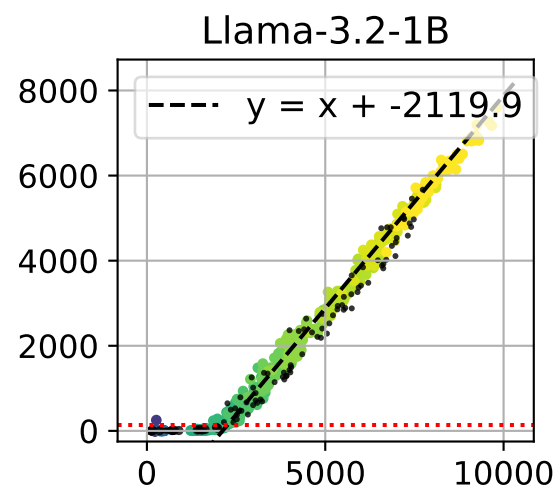
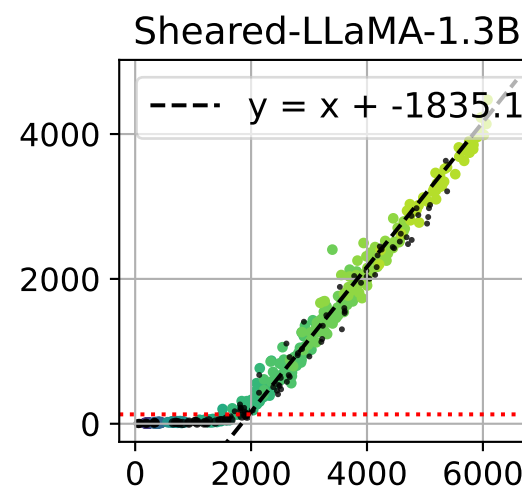
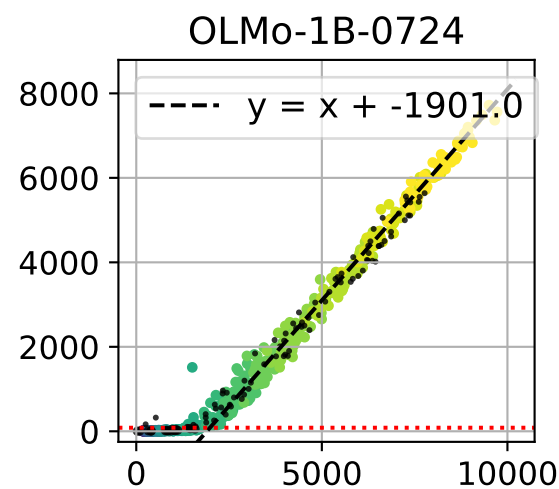
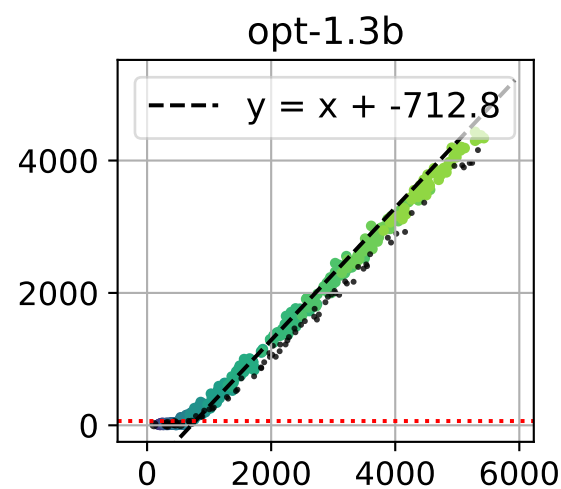
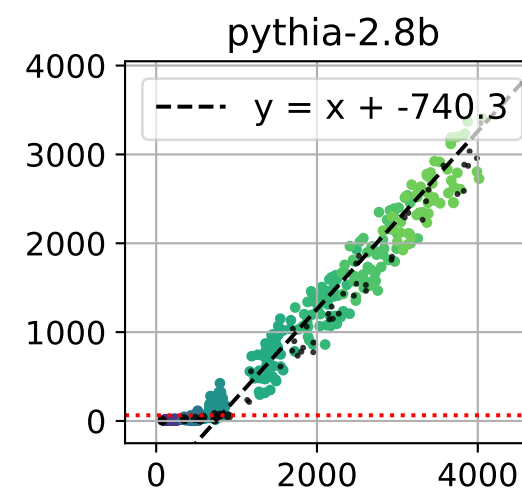
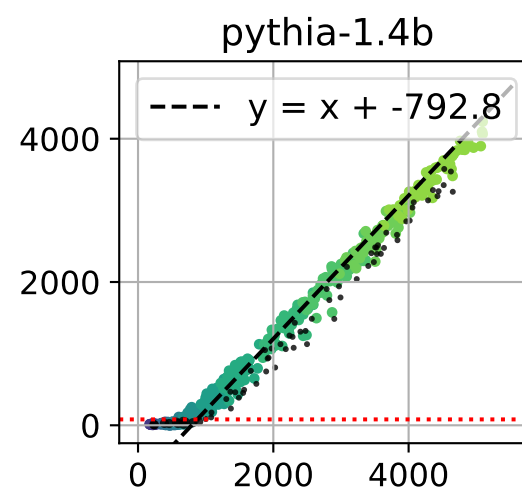
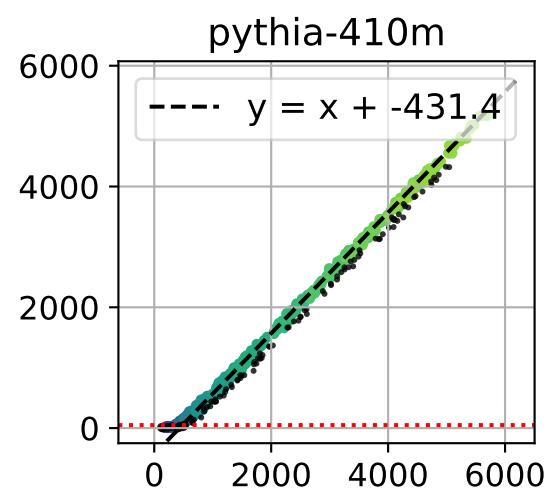
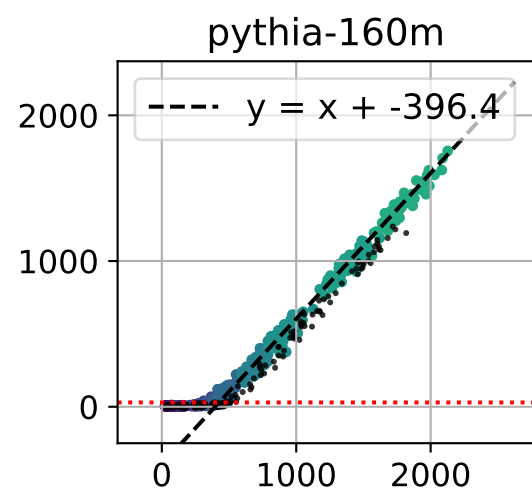


• Texts from PG-19 • Random text (sampled words) All texts below the line are losslessly compressed

Decompressed text, Cross-Entropy

Text Length, tokens



Text, Cross-Entropy