

GPT2-medium stacked weights
($d=65536$, $n=1024$)
Sketch-and-ridge-regression

