

GPT2-medium stacked weights
($d=16384$, $n=1024$)
Gram Matrix Approximation

