Batch Runtime for One Training Loop 700 **Full Parameter** LoRA 600 Adapter **BitFit** 500 Out of Memory GPT-2 XL GPT-2 **1.44**× 200 .19x .18x 1.25x 1.33x 100 MOC 1024 512 1024 512 Sequence Length