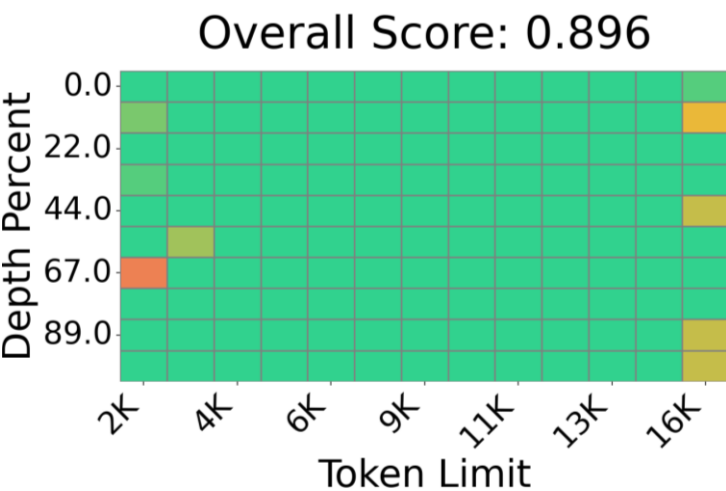
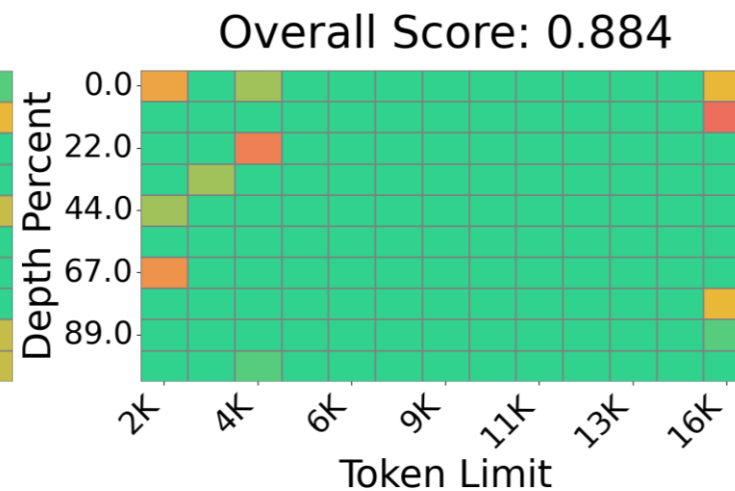
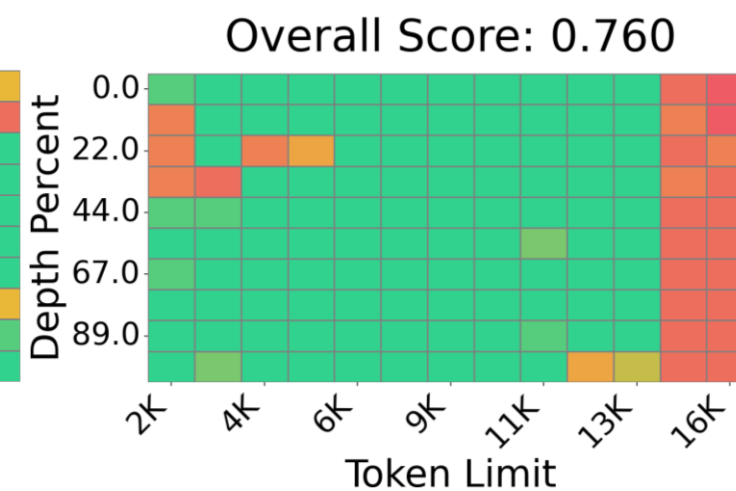


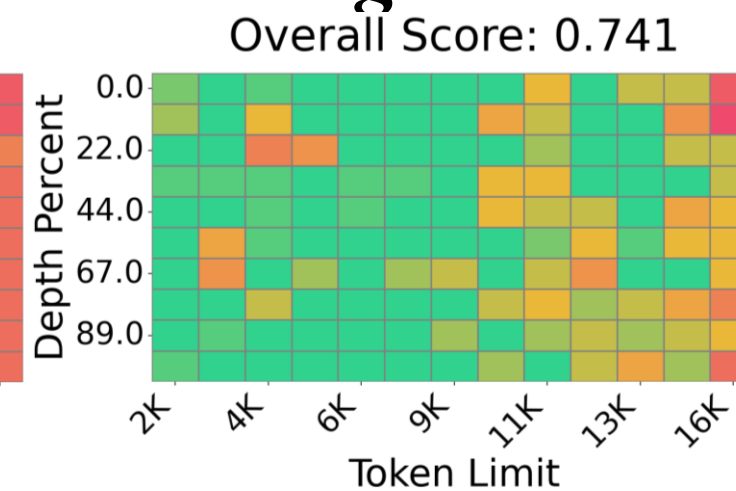
***LLaDA-8B-Instruct Budget = 256***



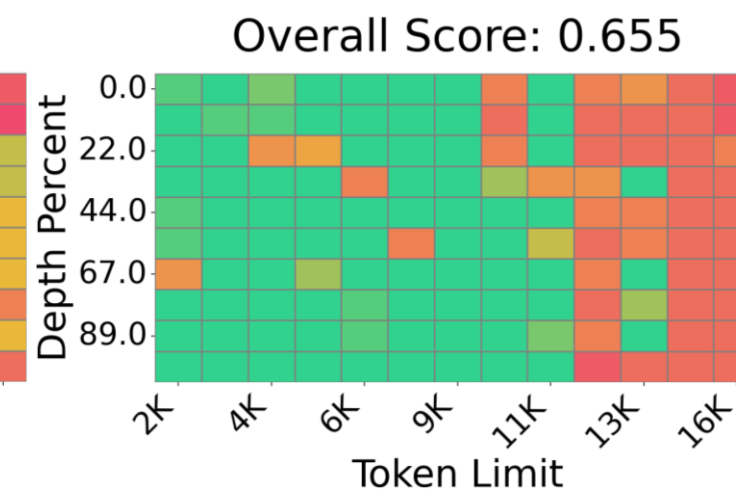
### *Full Attention*

*MaskKV(Ours)*

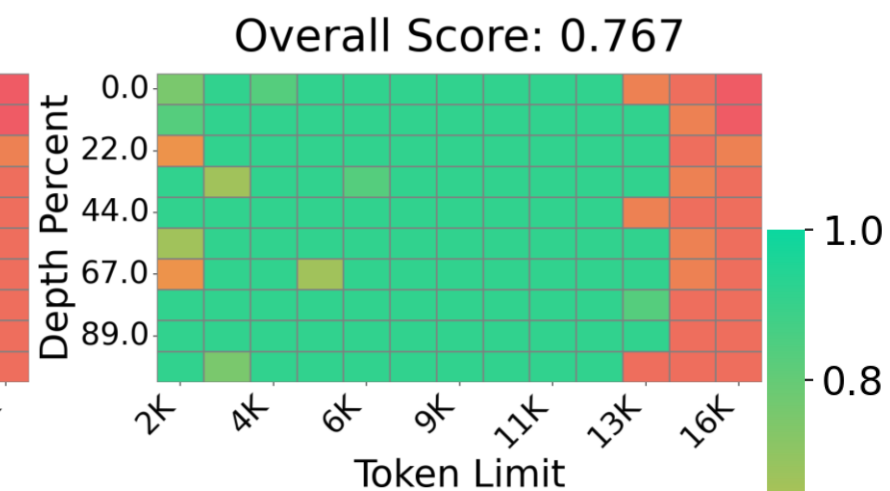
# SnapKV



# PyramidKV

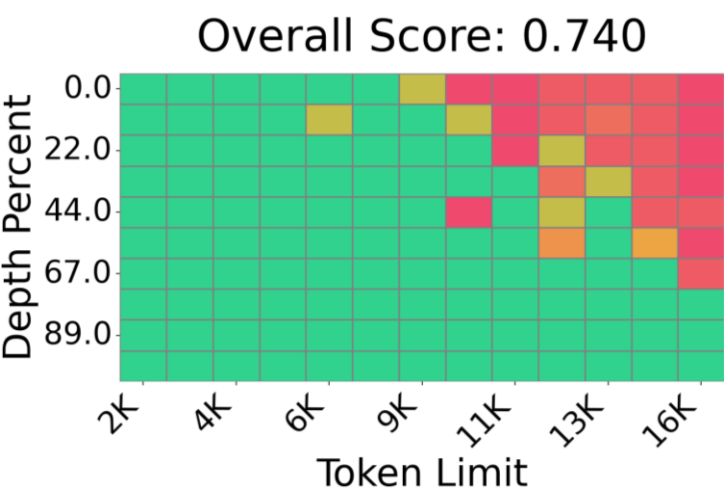


## SqueezeAttention

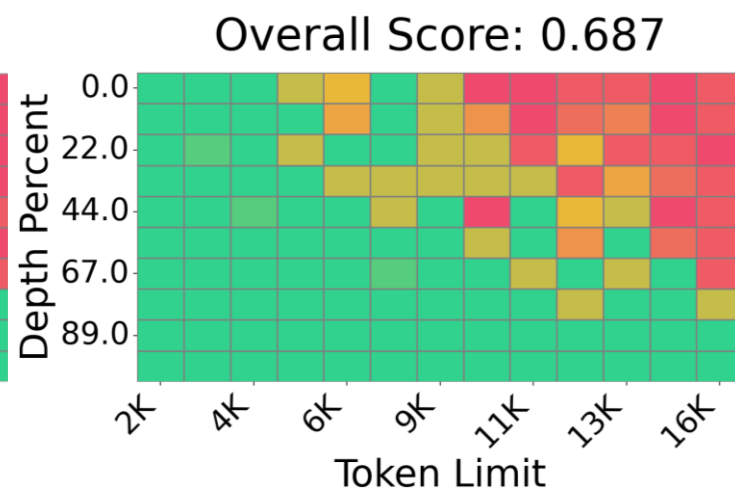
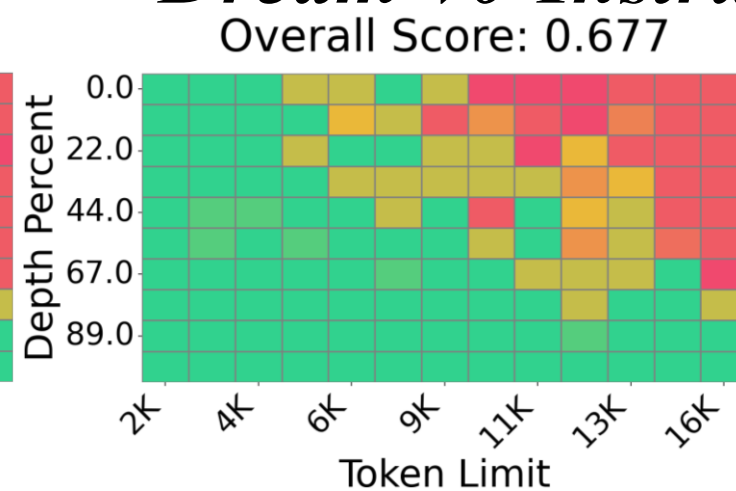


*AdaKV*

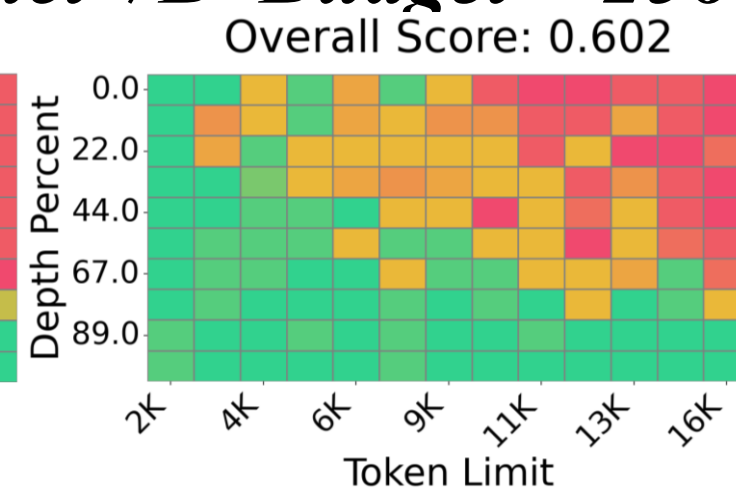
***Dream-v0-Instruct-7B Budget = 256***



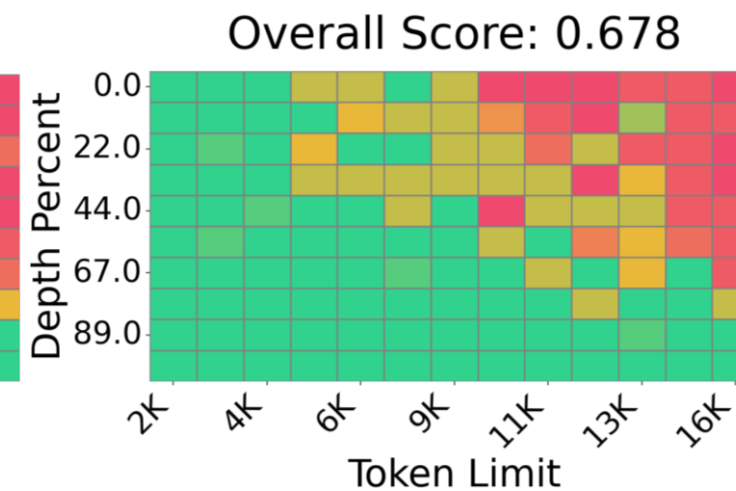
### *Full Attention*

*MaskKV(Ours)*

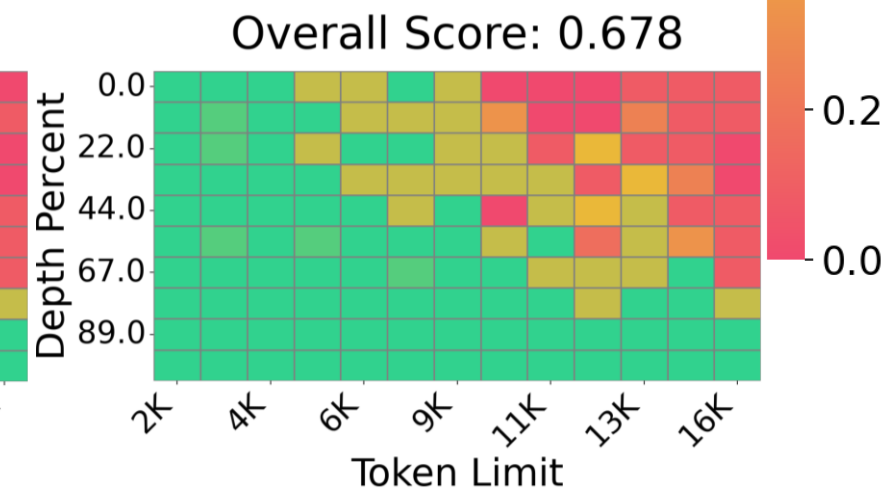
# SnapKV



# PyramidKV



## SqueezeAttention



*AdaKV*