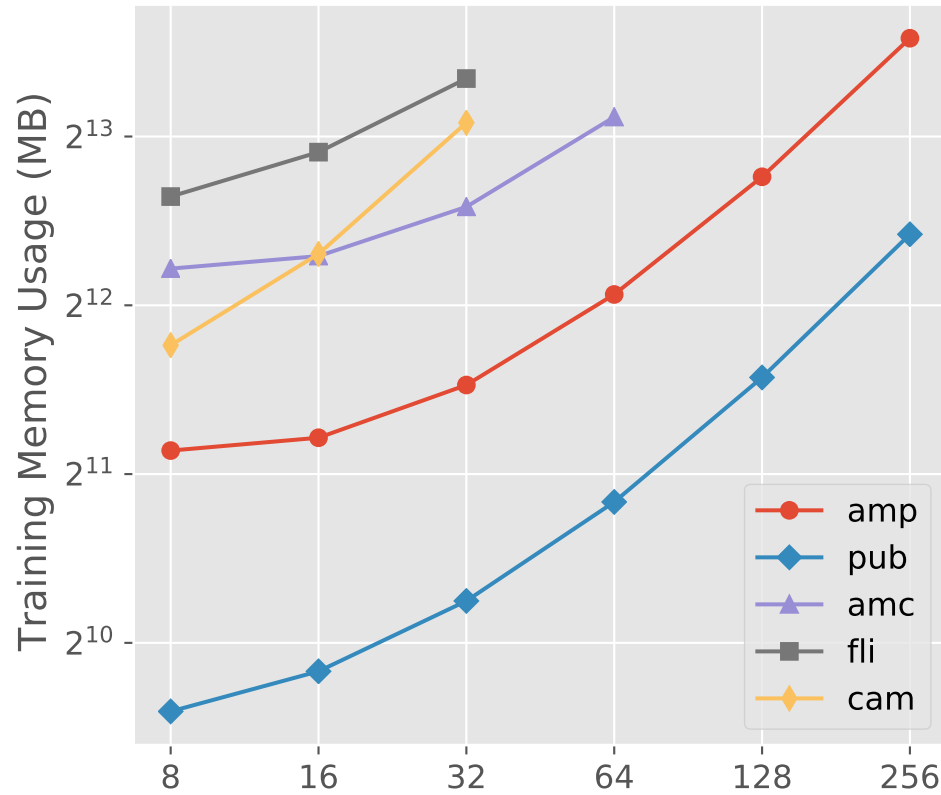
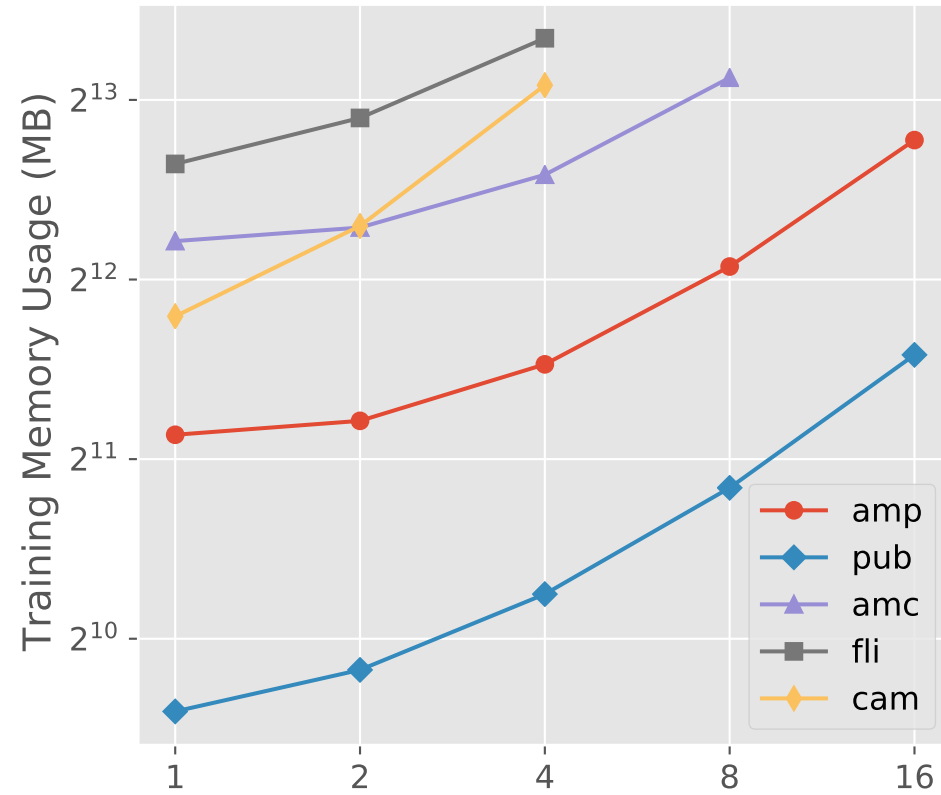


Hidden Vector Dimension $\dim(\mathbf{h}^1)$ (#Head=4, $d_a = d_v = d_m = 32$)



d_a, d_v, d_m (#Head=4, $\dim(\mathbf{h}^1)=64$)



#Head ($\dim(\mathbf{h}^1)=64$, $d_a = d_v = d_m = 32$)