cos similarity histogram - layer 13 ert-large-uncased\_uniform-tokens\_(Qx,Ky)\_trained-params\_dense-on\_24-lay 8 7 6 5 4 3 2 1 -0.50-0.250.00 0.25 0.50 -1.00-0.750.75 1.00