cos similarity histogram - layer 4 (albert-xlarge-v2_wikitext_\(Qx,Ky\)_trained-params_dense-off_24-layers) 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.250.00 0.25 0.50

0.75

1.00

-1.00

-0.75

-0.50