cos similarity histogram - layer 3 (bert-large-uncased\_wikitext\_(Qx,Ky)\_trained-params\_dense-on\_24-layers) 10 8 6 4

0.00

0.25

0.50

0.75

1.00

2

-1.00

-0.75

-0.50

-0.25