cos similarity histogram - layer 16 (bert-large-uncased_wikitext_(x,y)_trained-params_dense-off_24-layers) 8 6 4 2

0.00

-0.25

0.25

0.50

0.75

1.00

-0.50

-1.00

-0.75