cos similarity histogram - layer 6 (bert-large-uncased_wikitext_(Qx,Ky)_random-params_dense-off_24-layers) 12 10 8 6 4 2 -0.250.00 0.25 0.50 -1.00-0.75-0.500.75 1.00