cos similarity histogram - layer 23 (bert-large-uncased_wikitext_(Qx,Ky)_random-params_dense-on_24-layers) 17.5 15.0 12.5 10.0 7.5 5.0 2.5 0.0 0.25 0.50 -1.00-0.75-0.50-0.250.00 0.75 1.00