cos similarity histogram - layer 11 (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