cos similarity histogram - layer 8 (bert-large-uncased wikitext (x,y) random-params dense-off 24-layers) 14 12 10 8 6 4 2 -0.50-0.250.00 0.25 0.50 -1.00-0.750.75 1.00