cos similarity histogram - layer 21 (bert-large-uncased imdb (x,y) trained-params dense-on 24-layers) 10 8 6 4 2 -0.50-0.250.00 0.25 0.50 -1.00-0.750.75 1.00