cos similarity histogram - layer 20 (bert-large-uncased\_wikitext\_(x,y)\_trained-params\_dense-off\_24-layers) 8 6 4 2 0.00 0.25 0.50 -1.00-0.75-0.50-0.250.75 1.00