cos similarity histogram - layer 2 (albert-xlarge-v2_wikitext_(x,y)_trained-params_dense-on_24-layers) 5 4 3 2 1 -0.50-0.250.00 0.25 0.50 0.75 -1.00-0.751.00