cos similarity histograms (albert-xlarge-v2_imdb_(x,y)_trained-params_dense-on_192-layers) after layer 2 after layer 4 after layer 1 after layer 3 15 15 15 15 10 10 10 10 5 5 5 5 0.5 0.5 0.0 1.0 0.0 1.0 0.0 0.5 1.0 0.0 0.5 1.0 -1.0-0.5-1.0-0.5-1.0-0.5-1.0-0.5after layer 27 after layer 38 after layer 16 after layer 50 15 15 15 15 10 10 10 10 5 5 5 5 -0.50.0 0.5 -1.00.0 -0.50.0 1.0 -0.50.0 -1.01.0 -0.50.5 1.0 -1.00.5 -1.01.0 after layer 62 after layer 73 after layer 84 after layer 96 15 15 15 15 10 10 10 10 5 5 5 5 0 0 -1.0-0.50.0 0.5 1.0 -1.0-0.50.0 0.5 1.0 -1.0-0.50.0 0.5 1.0 -1.0-0.50.0 0.5 1.0 after layer 108 after layer 119 after layer 130 after layer 142 15 15 15 15 10 10 10 10 5 5 5 5 0.5 0.5 0.0 -0.50.0 0.5 -1.0-0.50.0 -1.0-0.50.0 -1.0-0.50.5 -1.01.0 1.0 1.0 1.0 after layer 154 after layer 165 after layer 176 after layer 188 15 15 15 15 10 10 10 10 0 + -1.0 0.0 0.5 0.0 0.5 -0.50.0 -0.5 0.0 -0.51.0 -1.0-0.51.0 -1.00.5 1.0 -1.00.5 1.0 after layer 189 after layer 190 after layer 191 after layer 192 15 15 15 15 10 10 10 10 5 5 5 5 0 -0 -1.0-0.50.0 0.5 -1.0-0.50.0 0.5 -0.50.0 0.5 -1.0-0.50.0 1.0 1.0 1.0 -1.01.0