cos similarity histograms (albert-xlarge-v2\_wikitext\_(x,y)\_trained-params\_dense-on\_24-layers) layer 2 layer 0 layer 1 layer 3 5 5 5 5 3 3 3 3 2 2 2 1 1 -0.50.0 0.5 1.0 -0.50.0 0.5 1.0 -0.50.0 0.5 1.0 -0.50.0 0.5 1.0 -1.0-1.0-1.0layer 4 layer 5 layer 6 layer 7 5 5 5 4 3 3 3 2 2 1 1 1 1 0 <del>|</del> −1.0 0 <del>|</del> -1.0 -0.50.0 0.5 1.0 -0.50.0 0.5 1.0 -1.0-0.50.0 0.5 1.0 -0.50.0 0.5 1.0 -1.0layer 8 layer 10 layer 11 layer 9 5 3 2 2 2 1 1 0 <del>↓</del> −1.0 0 + -1.0 0 --0.5 0.5 0.5 0.5 0.5 -0.50.0 -0.50.0 0.0 1.0 -0.5 0.0 -1.01.0 1.0 -1.01.0 layer 12 layer 13 layer 14 layer 15 5 5 5 5 3 3 2 2 2 · 2 1 1 1 -0.50.0 0.5 -0.5 0.0 0.5 -0.50.0 0.5 -0.50.0 0.5 -1.01.0 1.0 -1.01.0 -1.01.0 layer 17 layer 16 layer 18 layer 19 5 5 4 4 3 3 3 2 -2 2 1 1 1 1 0 <del>↓</del> −1.0 0 <del>|</del> −1.0 -1.0 -1.0-0.50.0 0.5 1.0 -0.50.0 0.5 1.0 -0.50.0 0.5 1.0 -0.50.5 0.0 1.0 layer 20 layer 21 layer 22 layer 23 5 5 5 5 4 4 4 4 3 3 3 3 2 2 2 · 2 1 1 1 1

0 -

-1.0

-0.5

0.0

0.5

1.0

0 -

-1.0

1.0

-0.5

0.0

0.5

1.0

0 -

-1.0

-0.5

0.0

0.5

0 -

-1.0

-0.5

0.0

0.5

1.0