HDBSCAN attention\_logit cluster sizes (wikitext\_bert-large-uncased\_params-trained\_dense-on\_layers-24) after layer 4 after layer 1 after layer 2 after layer 3 25 60 0.04 0.04 20 cluster size 0.02 0.02 40 15 0.00 0.00 10 20 -0.02-0.025 -0.04-0.040 0 k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster after layer 7 after layer 6 after layer 5 after layer 8 50 30 0.04 0.04 0.00 cluster size 0.00 -0.02 40 0.02 20 30 0.00 20 -0.0210 10 -0.04-0.040 0 k-th largest cluster k-th largest cluster 1 2 2 k-th largest cluster k-th largest cluster after layer 9 after layer 10 after layer 11 after layer 12 100 25 30 15 80 20 cluster size 01 00 02 10 60 15 40 10 5 20 5 0 0 2 3 4 1 k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster after layer 13 after layer 14 after layer 15 after layer 16 80 30 60 60 cluster size 01 00 02 60 40 40 40 20 20 20 0 0 0 0 2 3 1 3 2 2 k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster after layer 17 after layer 18 after layer 19 after layer 20 40 30 20 15 cluster size 01 30 15 10 20 10 0 -2 1 6 1 k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster after layer 24 after layer 21 after layer 22 after layer 23 10 100 80 100 8 cluster size 80 60 75 6 60 40 50 40 20 25 2 20 0 0 0 0

1

2

k-th largest cluster

3

1

k-th largest cluster

2

k-th largest cluster

3

1

k-th largest cluster

1