HDBSCAN token cluster sizes (wikitext\_bert-large-uncased\_params-trained\_dense-on\_layers-24) after layer 1 after layer 2 after layer 3 after layer 4 0.04 0.04 0.04 8 cluster size 0.02 0.02 0.02 0.00 0.00 0.00 -0.02-0.02-0.022 -0.04-0.04-0.040 k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster after layer 8 after layer 5 after layer 6 after layer 7 150 150 150 0.04 0.02 cluster size 100 100 100 0.00 -0.0250 50 50 -0.040 k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster after layer 11 after layer 9 after layer 12 after layer 10 0.04 20 20 0.04 0.00 cluster size 0.00 -0.02 0.02 0.02 15 15 0.00 10 10 -0.025 5 -0.04-0.040 k-th largest cluster k-th largest cluster 1 k-th largest cluster k-th largest cluster after layer 13 after layer 14 after layer 15 after layer 16 0.04 0.04 0.04 0.04 cluster size 0.02 0.02 0.02 0.02 0.00 0.00 0.00 0.00 -0.02-0.02-0.02-0.02-0.04-0.04-0.04-0.04k-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 30 -15 0.04 20 cluster size 0.02 15 20 10 0.00 10 -0.0210 5 -0.040 7 0 k-th largest cluster 2 k-th largest cluster k-th largest cluster k-th largest cluster after layer 21 after layer 22 after layer 23 after layer 24 125 8 15 150 100 cluster size 10 100 75 50 5 50 2 25 0 0 2 5 3 1 1 1 3 4 1 2 4 5 2 3 2 k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster