HDBSCAN token cluster sizes (wikitext_bert-large-uncased_params-trained_dense-on_layers-24) after layer 2 after layer 1 after layer 3 after layer 4 10 0.04 0.04 0.04 8 cluster size 0.02 0.02 0.02 6 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 5 after layer 6 after layer 7 after layer 8 0.04 0.04 0.04 0.04 0.02 0.02 0.02 0.02 cluster size 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 10 after layer 9 after layer 11 after layer 12 0.04 0.04 0.04 0.04 0.00 cluster size 0.00 -0.02 0.02 0.02 0.02 0.02 0.00 0.00 0.00 -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 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 25 0.04 0.04 0.04 20 0.02 0.02 cluster size 0.02 15 0.00 0.00 0.00 10 -0.02-0.02-0.02-0.04-0.04-0.040 7 k-th largest cluster k-th largest cluster k-th largest cluster 2 k-th largest cluster after layer 21 after layer 22 after layer 23 after layer 24 40 30 100 100 cluster size 30 75 75 20 50 50 10 25 25 0 0 0 0 2 2 1 3 1 1 3 1 k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster