HDBSCAN token cluster sizes (wikitext_bert-large-uncased_params-trained_dense-on_layers-24) after layer 2 after layer 4 after layer 1 after layer 3 40 40 40 15 30 30 30 cluster size 10 20 20 20 5 10 10 10 0 0 0 0 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 30 30 cluster size 0.02 0.02 0.00 0.00 20 -0.02-0.0210 10 -0.04-0.040 0 k-th largest cluster k-th largest cluster 3 k-th largest cluster k-th largest cluster after layer 9 after layer 10 after layer 11 after layer 12 12.5 40 30 10.0 30 30 cluster size 7.5 20 20 20 5.0 10 10 2.5 0 0 0.0 0 1 2 1 2 1 1 k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster after layer 15 after layer 13 after layer 14 after layer 16 10 12.5 0.04 0.04 8 10.0 cluster size 0.02 0.02 6 7.5 0.00 0.00 4 5.0 -0.02-0.022 2.5 -0.04-0.040.0 0 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 12.5 8 8 0.04 10.0 6 6 cluster size 0.02 7.5 0.00 5.0 -0.022.5 -0.040 -0.0 k-th largest cluster 2 1 k-th largest cluster k-th largest cluster k-th largest cluster after layer 21 after layer 22 after layer 23 after layer 24 20 150 150 20 cluster size 15 100 100 10 50 50 5 5 0 0 2 1 1 1 1 3 3 4 2 4 2 3 2 k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster