HDBSCAN cluster sizes (albert-xlarge-v2_wikitext_(x,y)_trained-params_dense-on_192-layers) after layer 8 after layer 32 after layer 24 after layer 16 25 100 20 30 20 80 15 20 60 10 40 10 5 20 0 0 1 5 6 7 8 k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster after layer 40 after layer 48 after layer 56 after layer 64 25 15 15 15 20 10 10 15 10 5 5 5 3 4 5 6 7 8 9 10 1 2 3 4 5 6 7 8 9 101112 1 2 3 4 5 6 7 8 9 101112 1 2 k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster after layer 72 after layer 88 after layer 96 after layer 80 40 40 40 40 30 30 30 20 20 20 10 10 10 10 0 1 2 3 4 5 6 7 8 9 10 11 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8 9 10 11 1 2 1 k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster after layer 104 after layer 112 after layer 120 after layer 128 60 40 50 50 40 40

