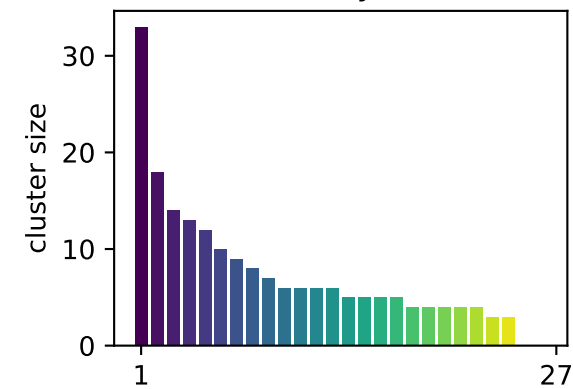


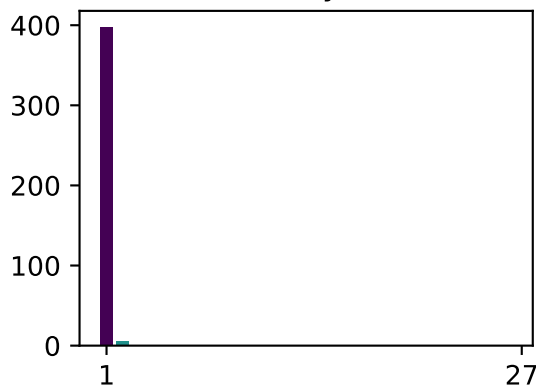
HDBSCAN attention_logit cluster sizes
(wikitext_albert-large-v2_params-trained_dense-on_layers-48)

after layer 1



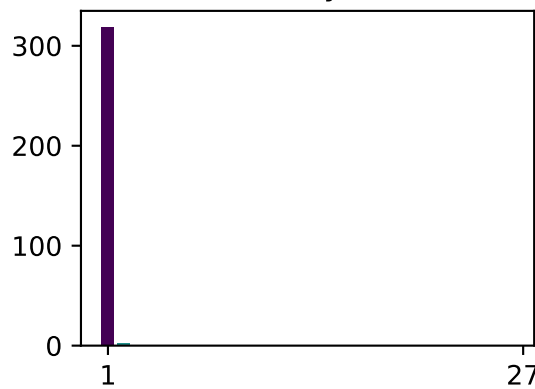
k-th largest cluster

after layer 2



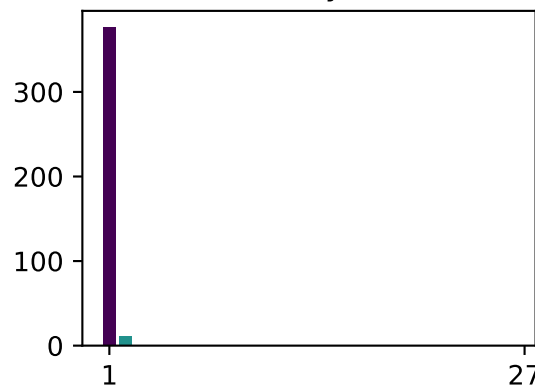
k-th largest cluster

after layer 3



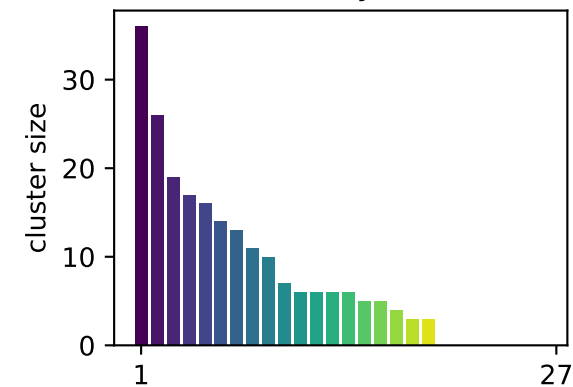
k-th largest cluster

after layer 4



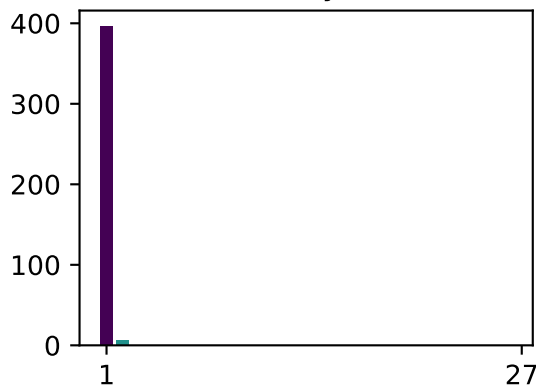
k-th largest cluster

after layer 5



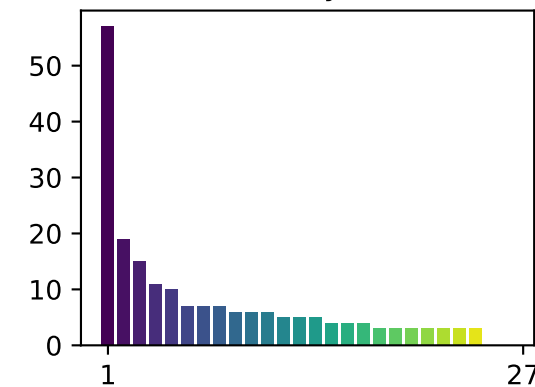
k-th largest cluster

after layer 6



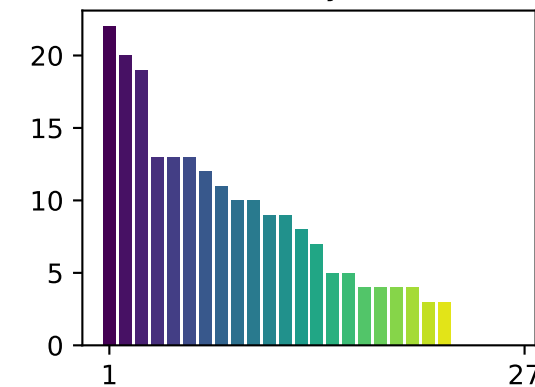
k-th largest cluster

after layer 7



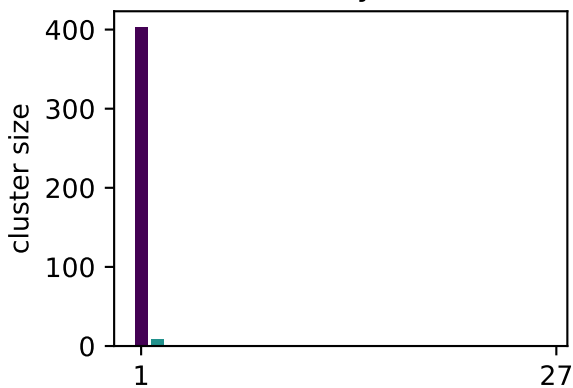
k-th largest cluster

after layer 8



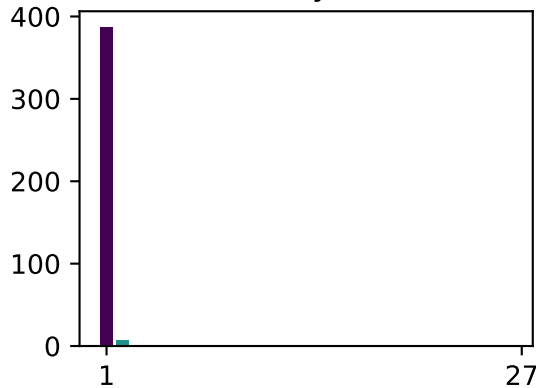
k-th largest cluster

after layer 9



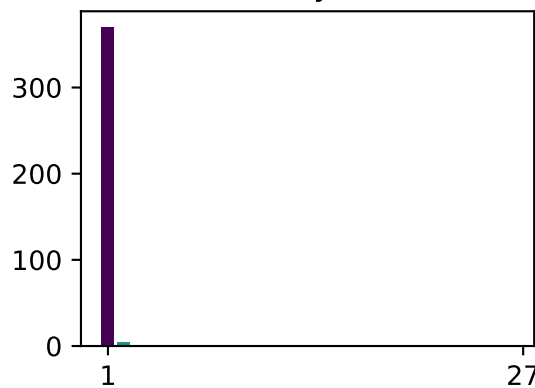
k-th largest cluster

after layer 10



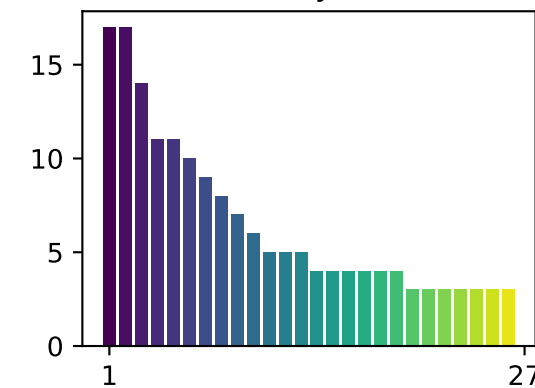
k-th largest cluster

after layer 11



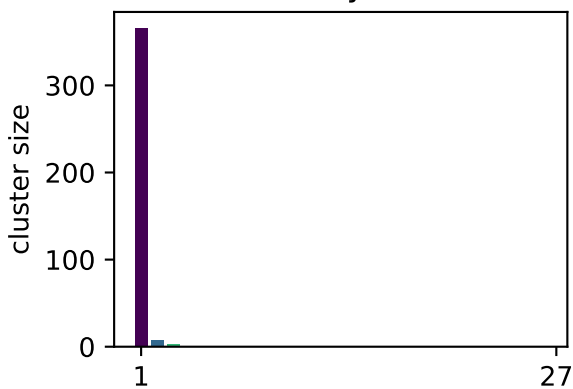
k-th largest cluster

after layer 12



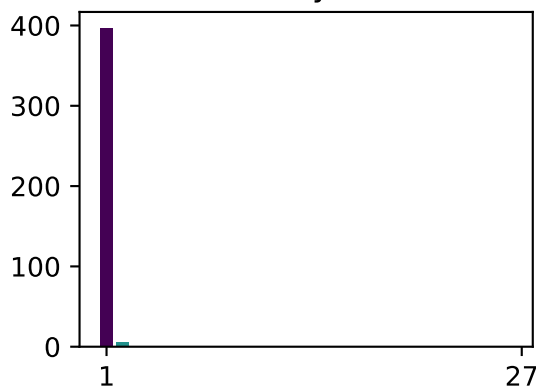
k-th largest cluster

after layer 13



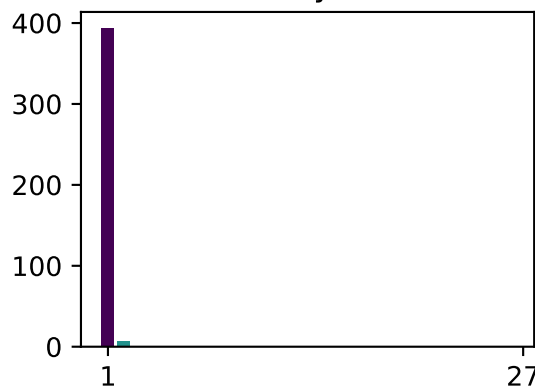
k-th largest cluster

after layer 14



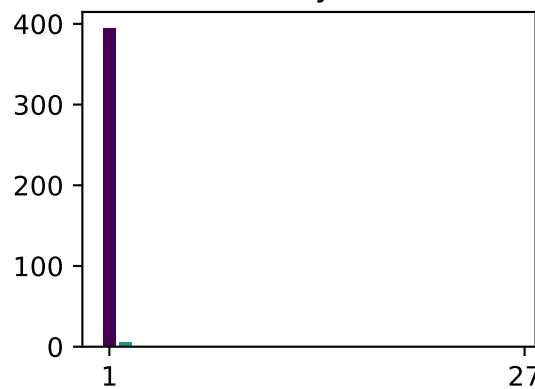
k-th largest cluster

after layer 15



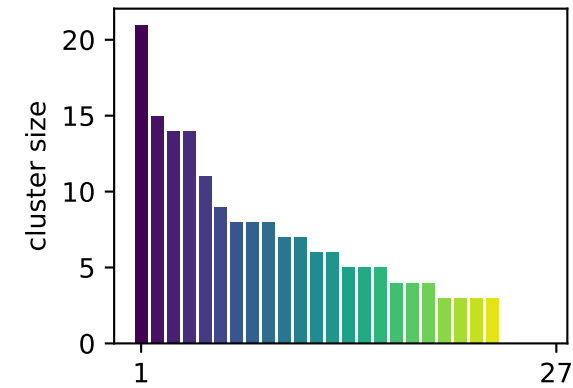
k-th largest cluster

after layer 16



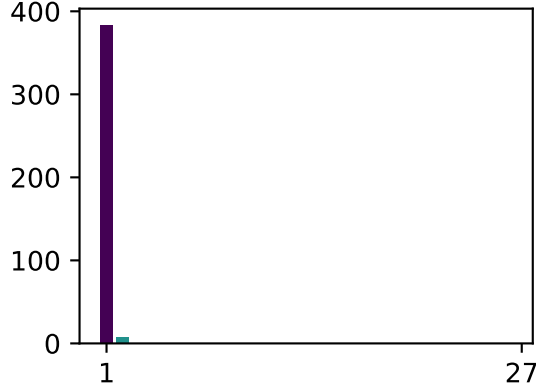
k-th largest cluster

after layer 17



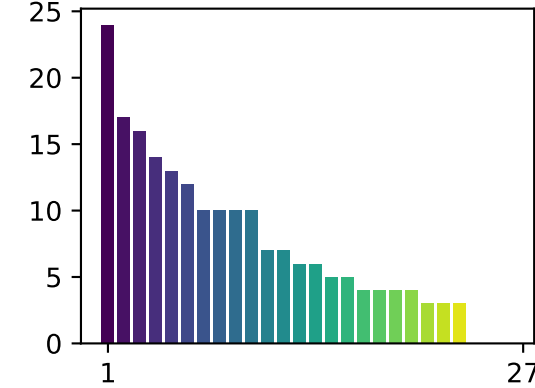
k-th largest cluster

after layer 18



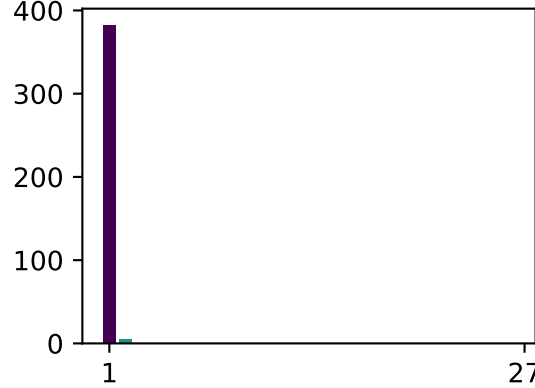
k-th largest cluster

after layer 19



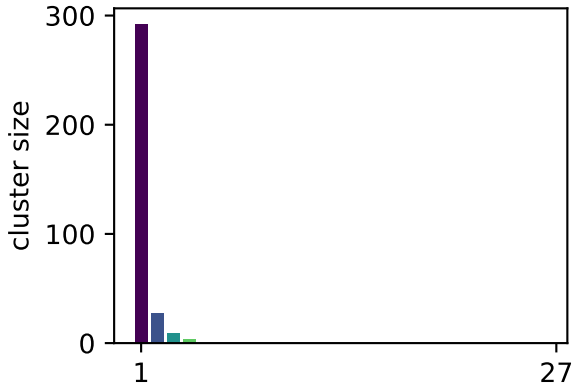
k-th largest cluster

after layer 20



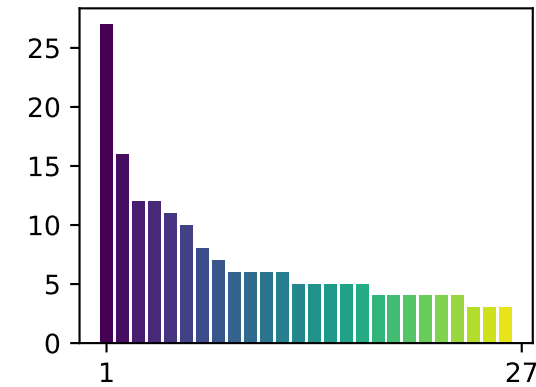
k-th largest cluster

after layer 21



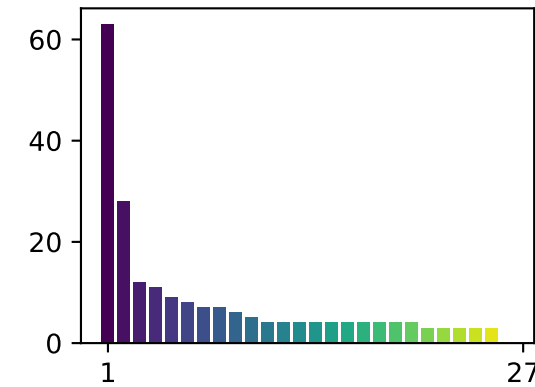
k-th largest cluster

after layer 22



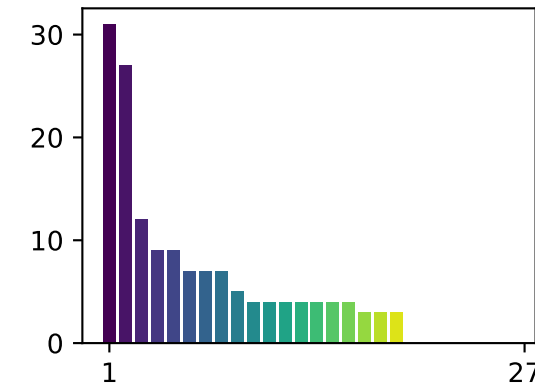
k-th largest cluster

after layer 23



k-th largest cluster

after layer 24



k-th largest cluster