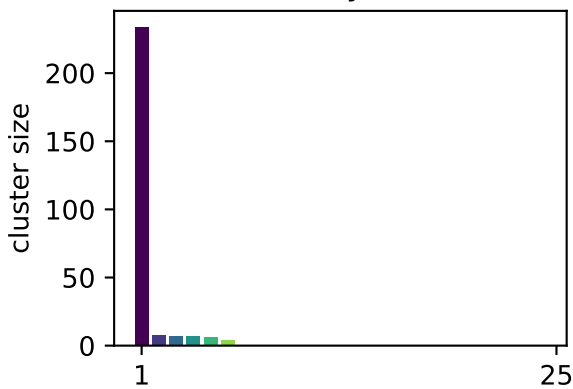
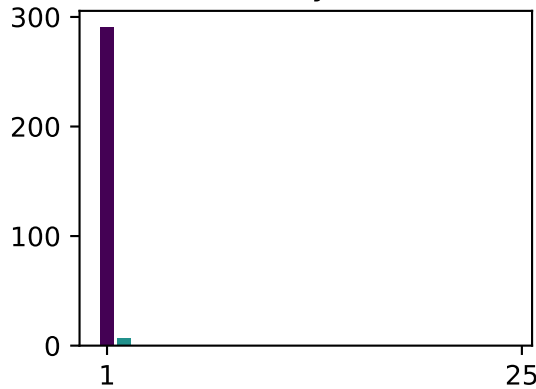


HDBSCAN attention_logit cluster sizes
(wikitext_albert-large-v2_params-trained_dense-on_layers-48)

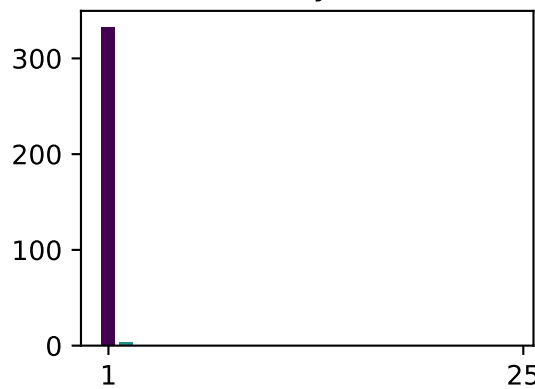
after layer 25



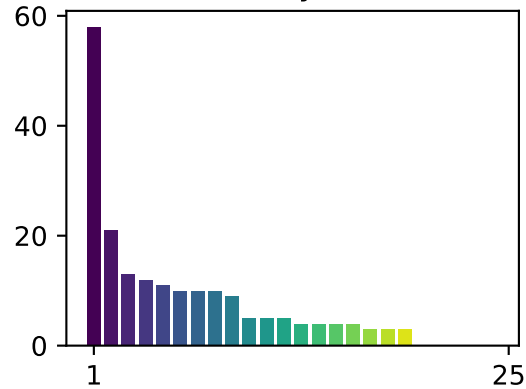
after layer 26



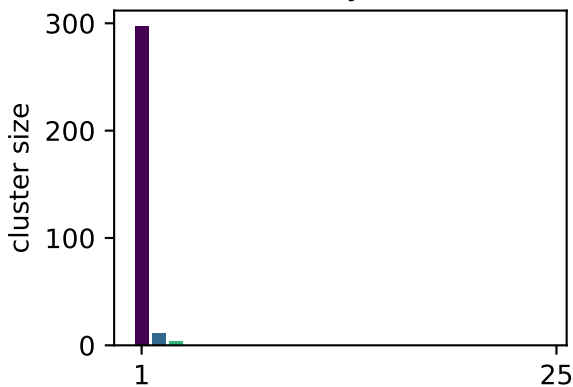
after layer 27



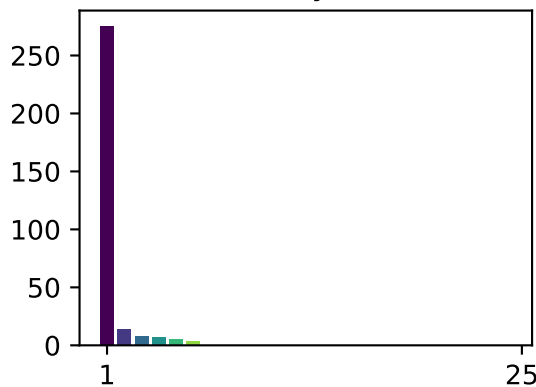
after layer 28



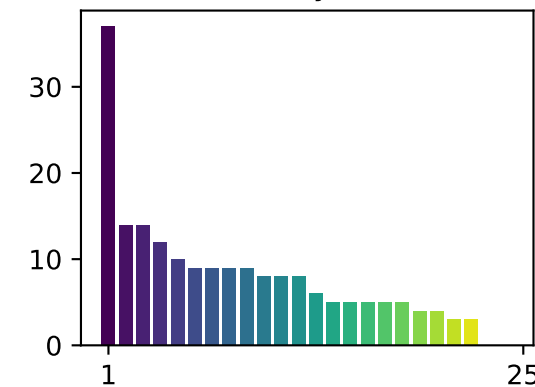
after layer 29



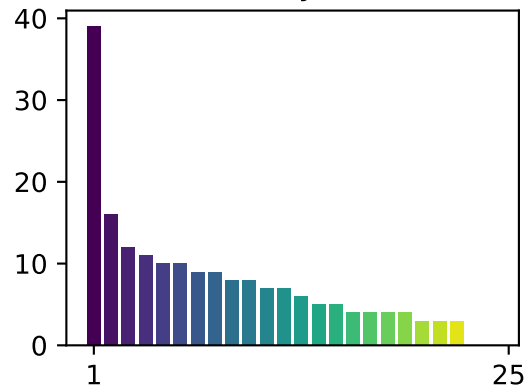
after layer 30



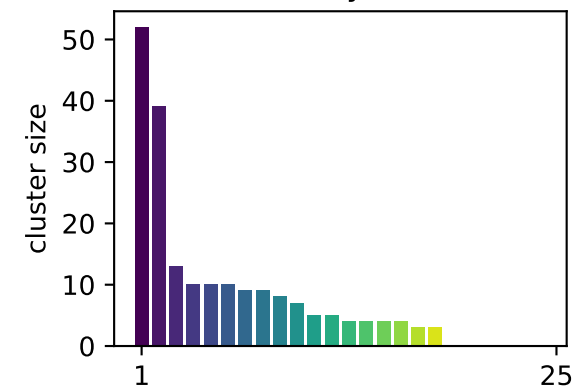
after layer 31



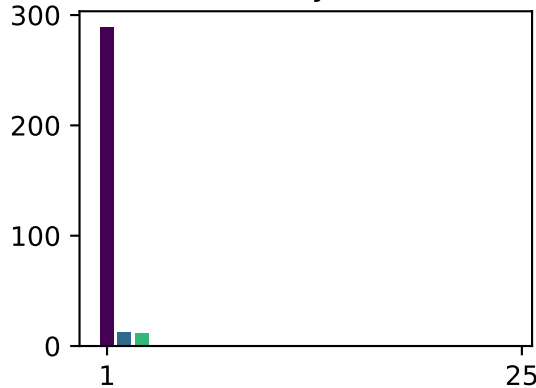
after layer 32



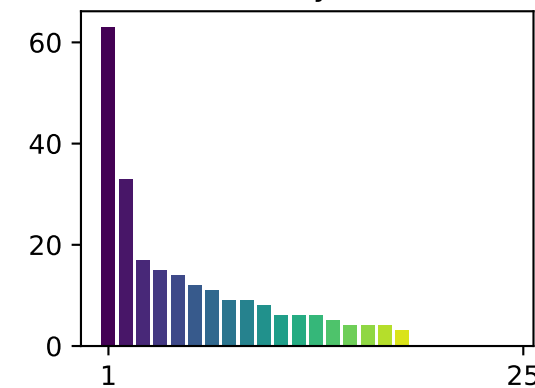
after layer 33



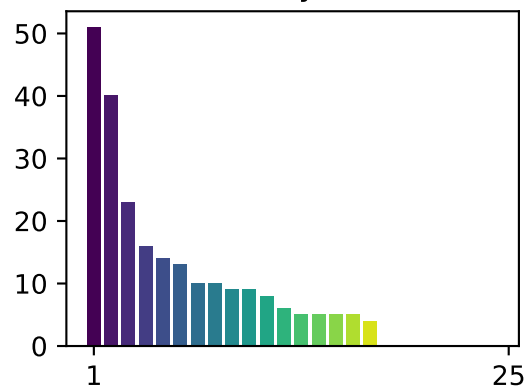
after layer 34



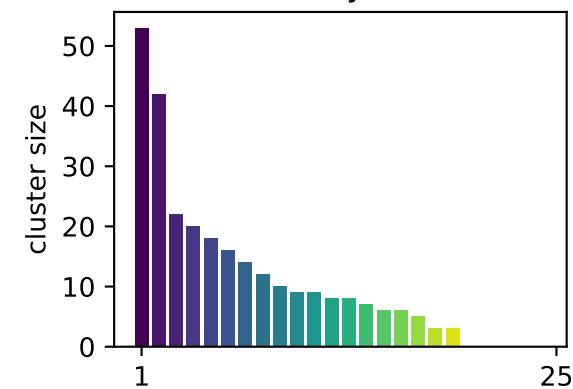
after layer 35



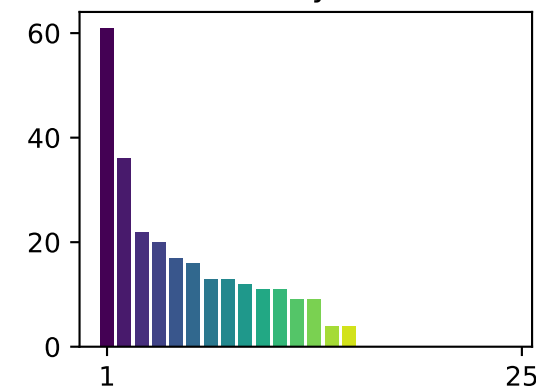
after layer 36



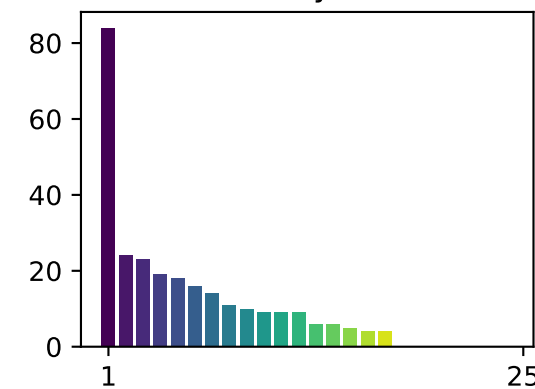
after layer 37



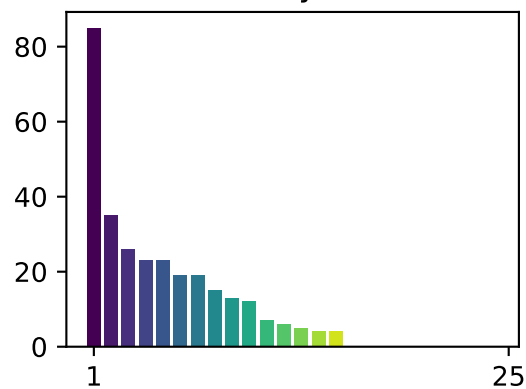
after layer 38



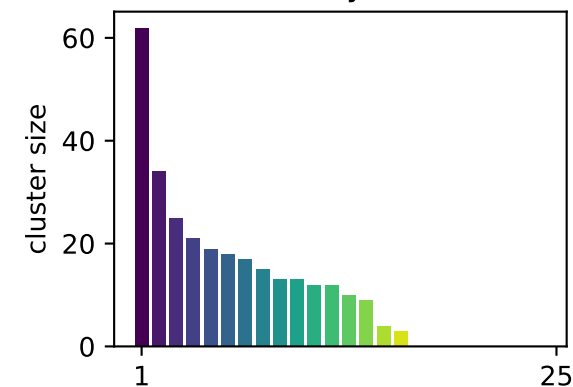
after layer 39



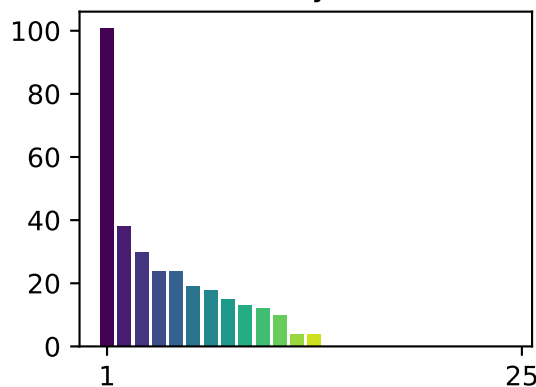
after layer 40



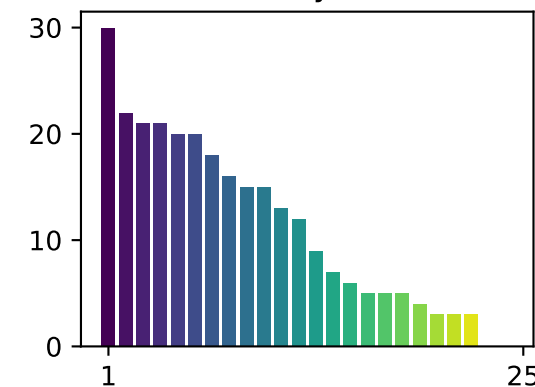
after layer 41



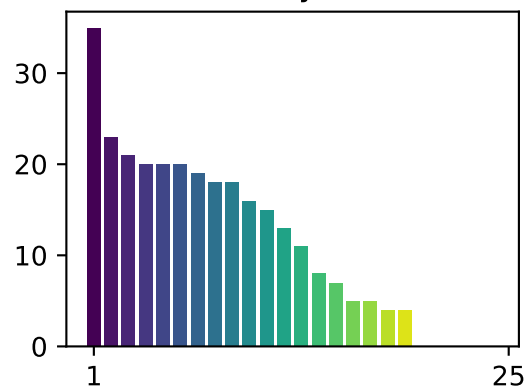
after layer 42



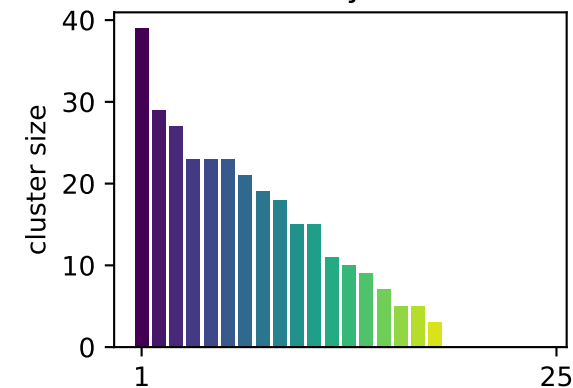
after layer 43



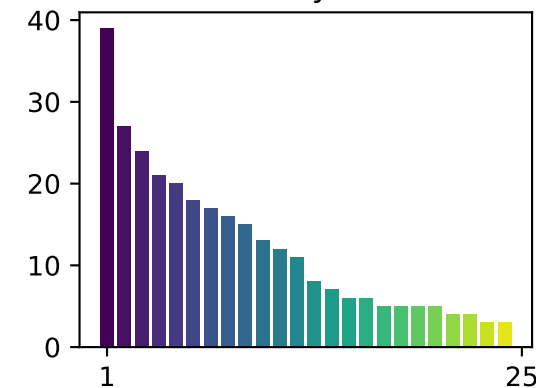
after layer 44



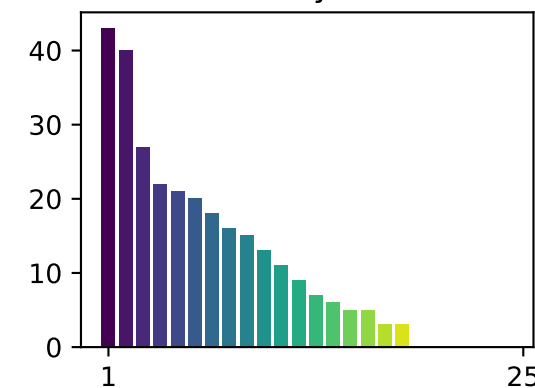
after layer 45



after layer 46



after layer 47



after layer 48

