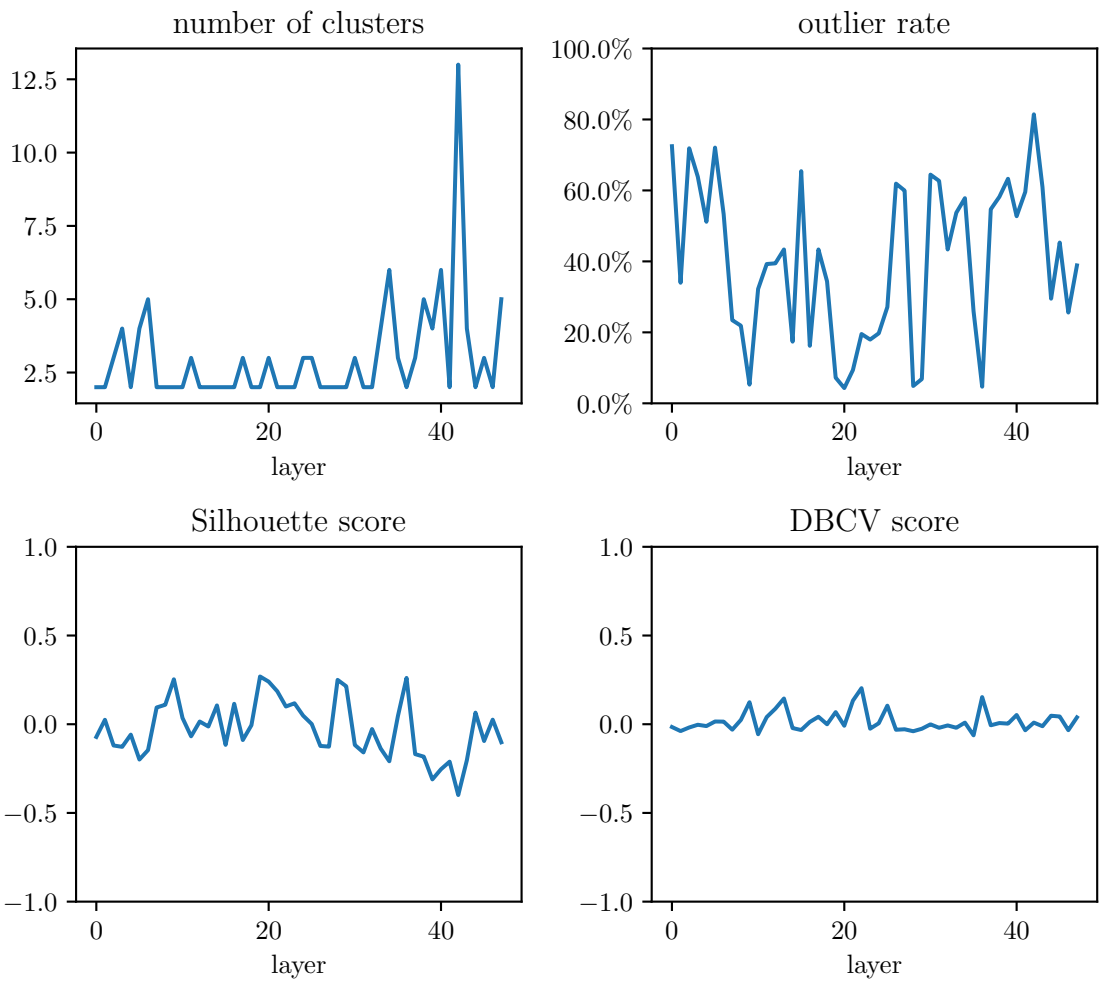
HDBSCAN cluster evaluation of tokens x_i
(randomtext_albert-xlarge-v2_params-random_dense-off_layers-48)
number of clusters
outlier rate
Silhouette score
DBCV score
layer
layer
layer
layer
12.5
10.0
7.5
5.0
2.5
100.0%
80.0%
60.0%
40.0%
20.0%
0.0%
1.0
0.5
0.0
−0.5
−1.0
0
20
40