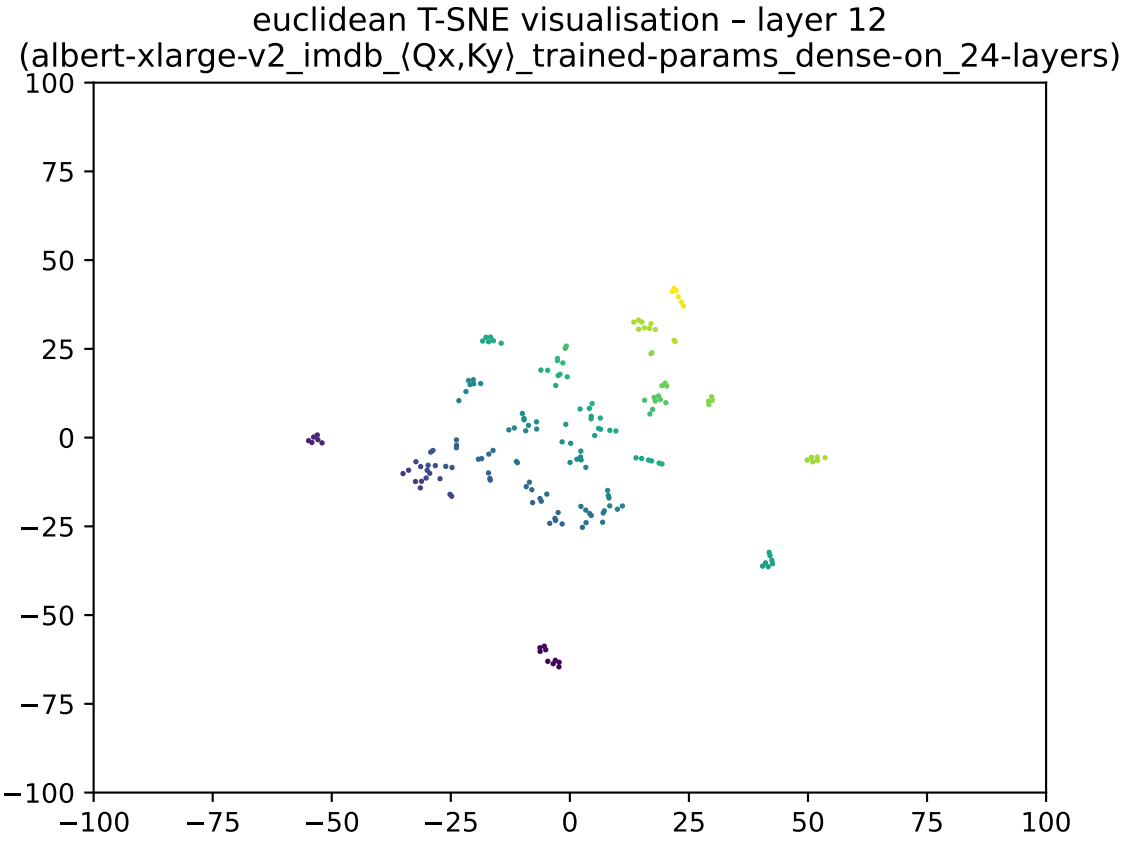
euclidean T-SNE visualisation – layer 12
(albert-xlarge-v2_imdb_⟨Qx,Ky⟩_trained-params_dense-on_24-layers)
100
75
50
25
0
−25
−50
−75
−100
−100
−75
−50
−25
0
25
50
75
100