euclidean T-SNE visualisation – layer 3 (bert-large-uncased\_wikitext\_\(Qx,Ky\)\_trained-params\_dense-on\_24-layers) 75 50 25 0 -25-5030 -75

25

50

75

100

-100

-100

-75

-50

-25