cosine T-SNE visualisation – layer 17 (bert-large-uncased_wikitext_\(Qx,Ky\)_trained-params_dense-off_24-layers) 75 50 25 0 -25-50-75-100-100-75-50-250 25 50 75 100