cosine T-SNE visualisation – layer 23 (bert-large-uncased_wikitext_(x,y)_trained-params_dense-off_24-layers) 75 50 25 0 -25-50-75-100 -50 -100-75-50-2525 75 100