cosine T-SNE visualisation – layer 0 (bert-large-uncased_wikitext_(x,y)_trained-params_dense-on_24-layers) 75 50 25 0 ÷ 88 -25-50-75-10050 -100-75-50-250 25 75 100