cosine T-SNE visualisation – layer 0 (bert-large-uncased\_wikitext\_(x,y)\_trained-params\_dense-off\_24-layers) 75 50 25 0 -25-50-75-100 --5050 -100-75-2525 75 100