cosine T-SNE visualisation – layer 17 (albert-xlarge-v2_wikitext_(Qx,Ky)_trained-params_dense-on_24-layers) 75 50 25 0 -25-50-75-100 + -50 50 -100-75-2525 75 100