cosine T-SNE visualisation - layer 4  $(albert-xlarge-v2\_wikitext\_\langle Qx,Ky\rangle\_random-params\_dense-off\_24-layers)\\ 100 + (Qx,Ky)\_random-params\_dense-off\_24-layers)$ 75 50 25 0 -25٥. -50-75-100-100-75-50-2525 50 75 100