euclidean T-SNE visualisations (bert-large-uncased\_wikitext\_(x,y)\_trained-params\_dense-on\_24-layers) layer 0 layer 1 layer 2 layer 3 100 100 100 100 50 50 50 50 0 -50-50-50 -50-100-100-1000 50 100 0 50 100 0 50 100 50 100 -50 -100 -50-50-50 -100 -100-100layer 5 layer 4 layer 6 layer 7 100 100 100 100 50 50 50 50 0 -50-50 -50 -50 -100-100-50 0 50 100 -100-500 50 100 -50 0 50 100 50 100 -100 -100-100-50 layer 8 layer 9 layer 10 layer 11 100 100 100 100 50 50 50 50 0 0 -50-50**-**50 -100 -100 -100-100100 50 100 50 100 50 100 50 -100-50 -100-50-100-50 -100-50 layer 12 layer 13 layer 15 layer 14 100 100 100 100 50 50 50 50 **-**50 **-**50 -50-50-100 -100-100-100 -50 0 -100 -50 0 50 -50 50 -50 0 50 -100 50 100 100 -1000 100 -100100 layer 16 layer 17 layer 18 layer 19 100 100 100 100 50 50 50 50 -50 <del>-</del>50 --50 -50 --100 <del>|</del> -100 <del>-</del>100 --100 <del>|</del> -100 -100 -50 -50 0 50 100 -100 -50 0 50 100 -50 0 100 -100 -50 0 50 100 layer 20 layer 22 layer 21 layer 23 100 100 100 100 50 50 50 50 0 0 0 0 -50-50 -50 -50 -100 --100 -100 --100 -

-100

**-**50

0

50

100

-100

-50

50

100

-100

-50

50

100

-100

-50

50

100