Heatmaps of cos similarity  $\langle \frac{x_i}{\|x_i\|}, \frac{x_j}{\|x_j\|} \rangle$ (wikitext\_albert-xlarge-v2\_params-random\_dense-on\_layers-48, sample 1) layer 25 layer 26 layer 27 1.0 0

