Heatmaps of cos similarity  $\langle \frac{x_i}{\|x_i\|}, \frac{x_j}{\|x_j\|} \rangle$  $({\tt wikitext\_albert\_xlarge\_v2\_params\_random\_dense\_on\_layers\_48}, ~ {\tt sample} ~ 3)$ layer 25 layer 26 layer 27 1.0 0

