Heatmaps of attention logits  $\frac{1}{\sqrt{d'}}\langle Qx_i,Kx_j\rangle$  (imdb\_bert-large-uncased\_params-random\_dense-on\_layers-24)

