cos similarity heatmap – layer 20
(...ert-large-uncased_wikitext_⟨Qx,Ky⟩_trained-params_dense-off_24-layers)