cos similarity histogram – layer 23
(albert-xlarge-v2_wikitext_⟨Qx,Ky⟩_trained-params_dense-off_24-layers)