

Frequency of gate>0 vs.  $\cos(w_{gate}, w_{out})$  in allenai/OLMo-7B-0424-hf