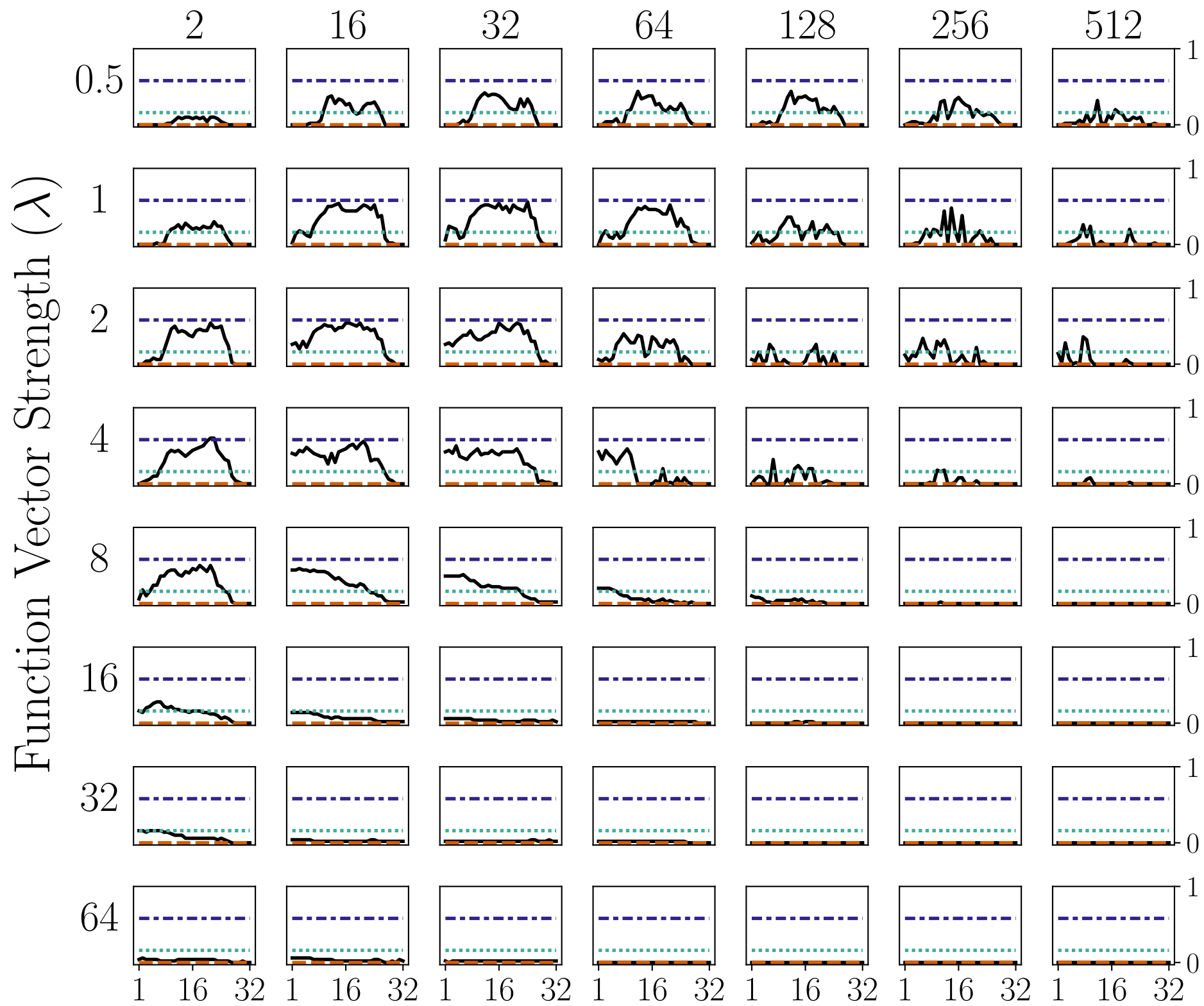


OLMo 7B SFT — Antonym

— 0-shot with Function Vector - - - 0-shot 1-shot - - - 5-shot

Number of Heads (\mathcal{A}_n)



Activation Patching Layer