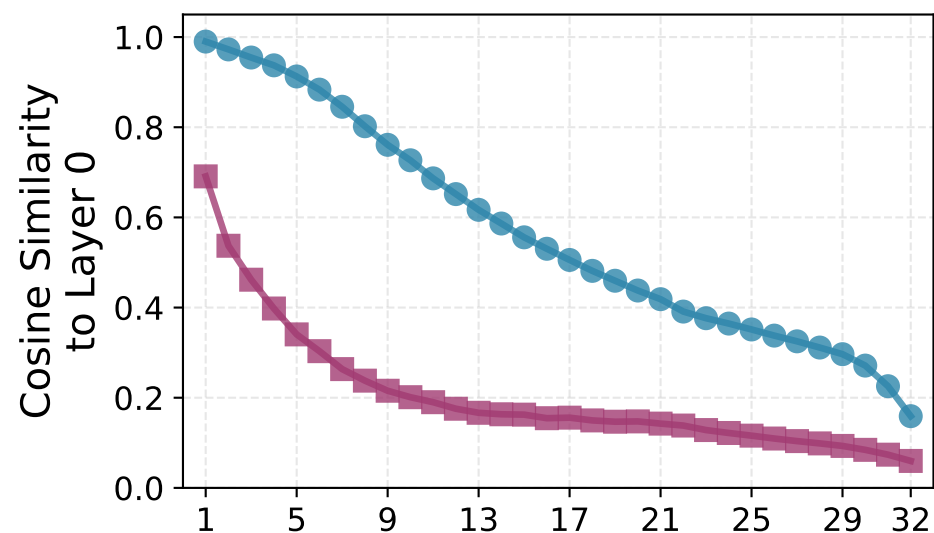
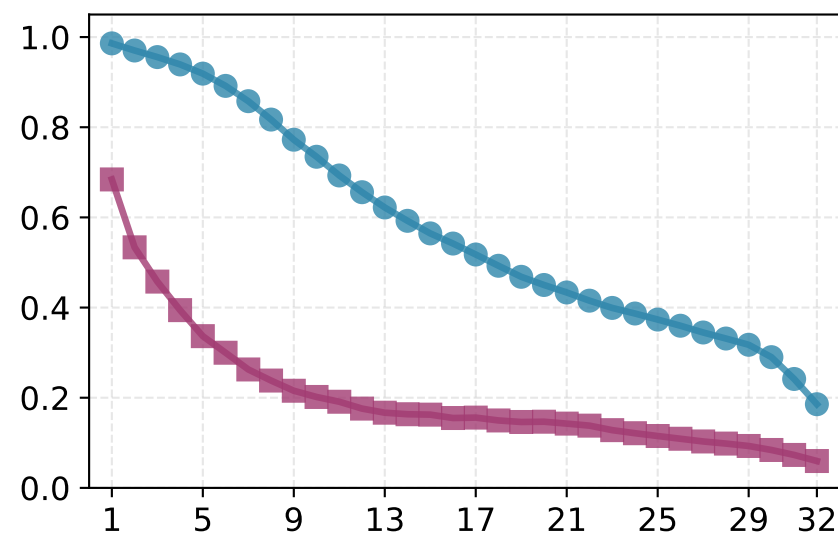


Same-Token Similarity Across LLM Layers

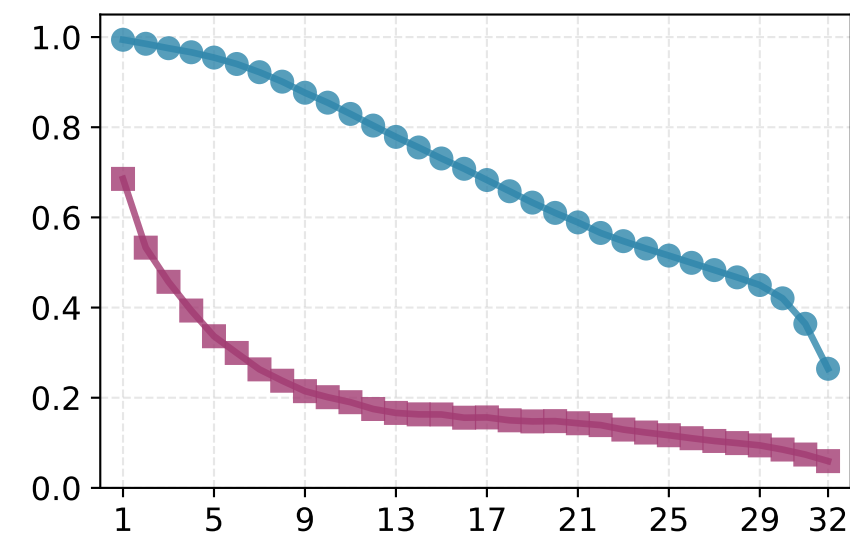
OLMo-7B + CLIP



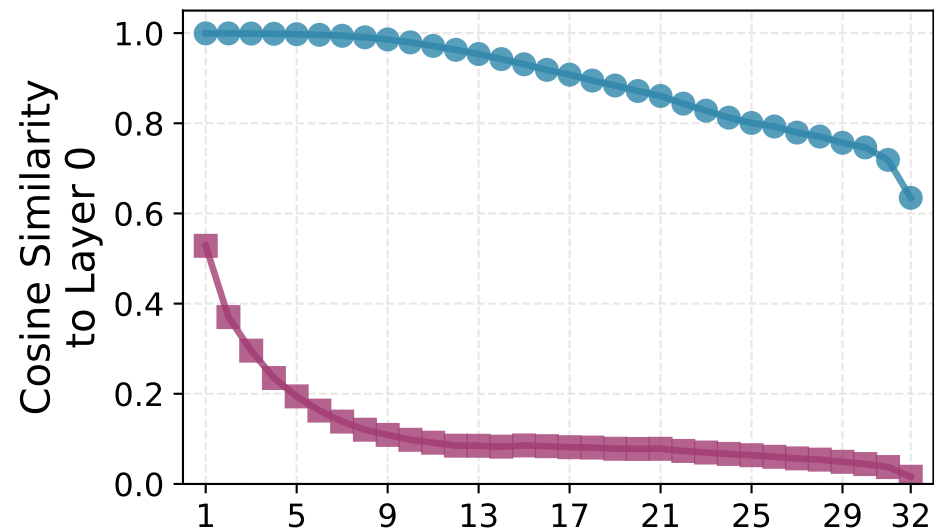
OLMo-7B + SigLIP



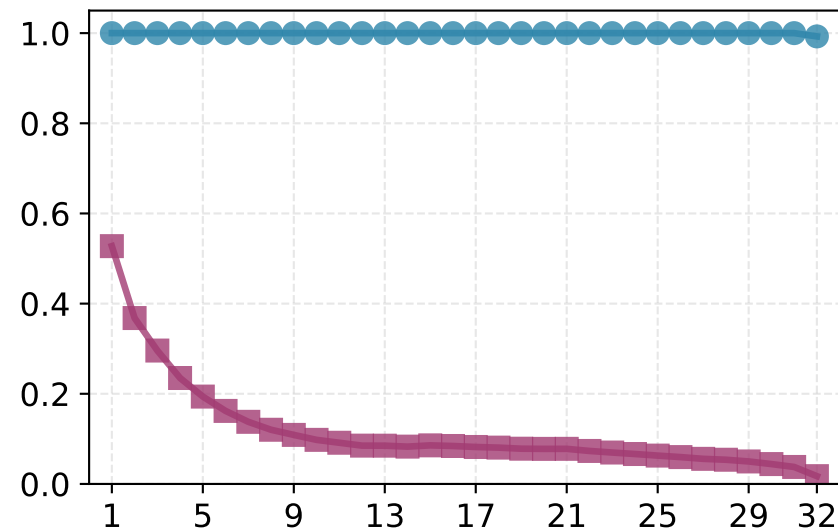
OLMo-7B + DINOv2



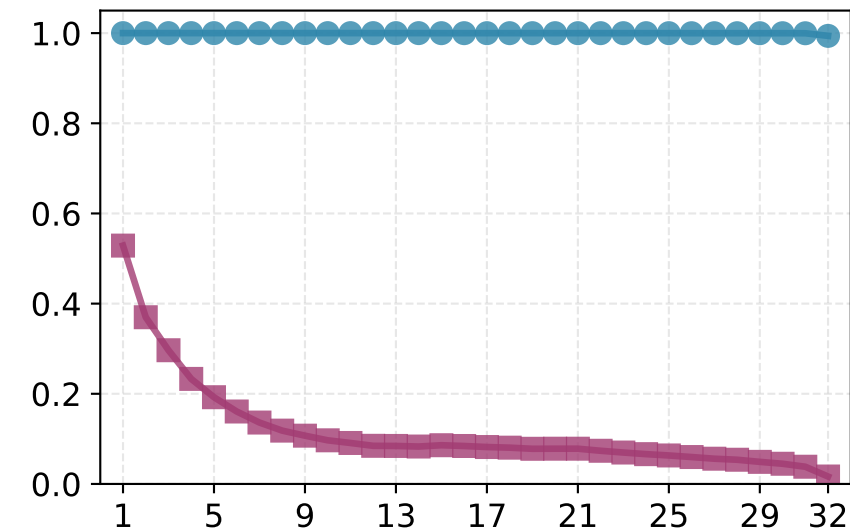
Llama3-8B + CLIP



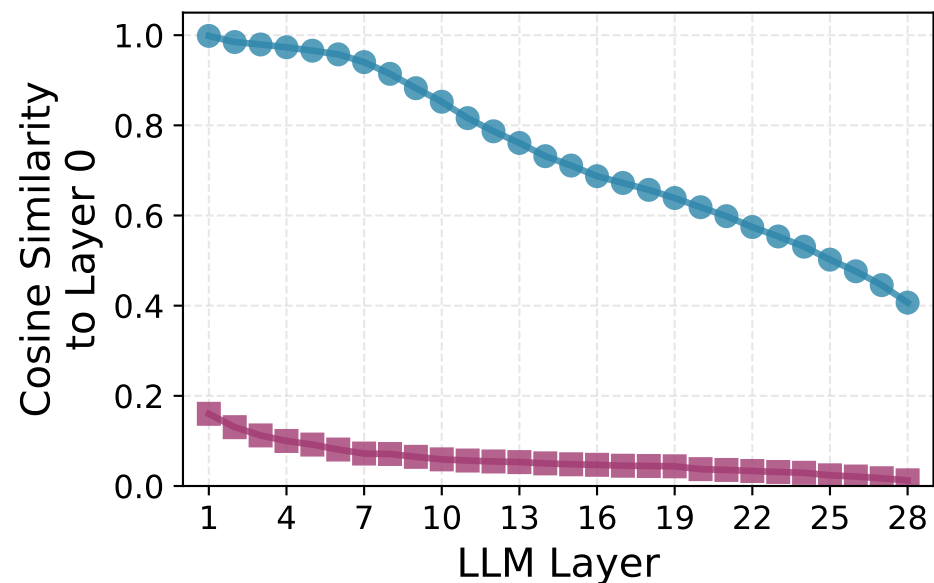
Llama3-8B + SigLIP



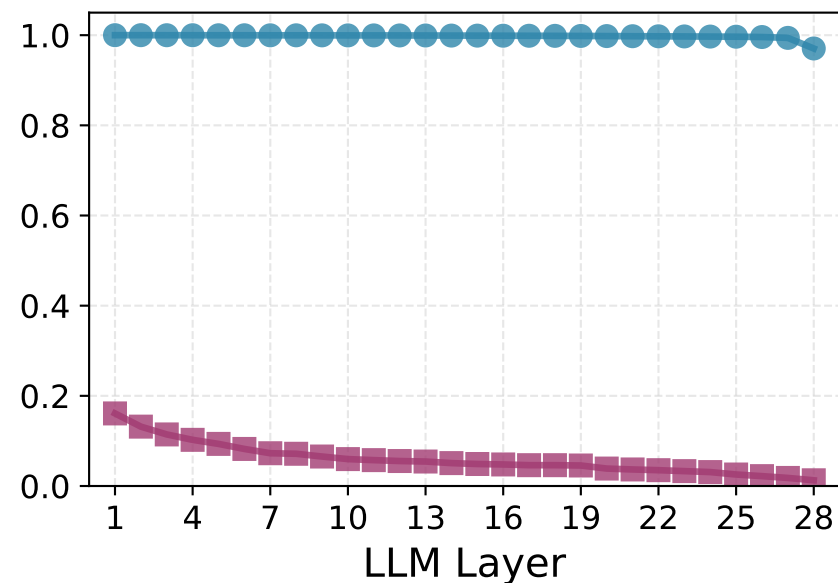
Llama3-8B + DINOv2



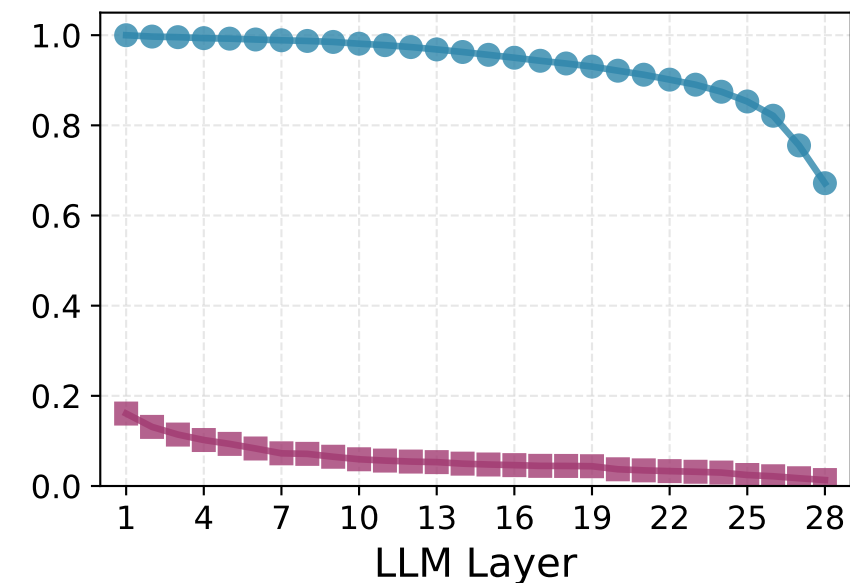
Qwen2-7B + CLIP



Qwen2-7B + SigLIP



Qwen2-7B + DINOv2



—●— Vision tokens —■— Text tokens