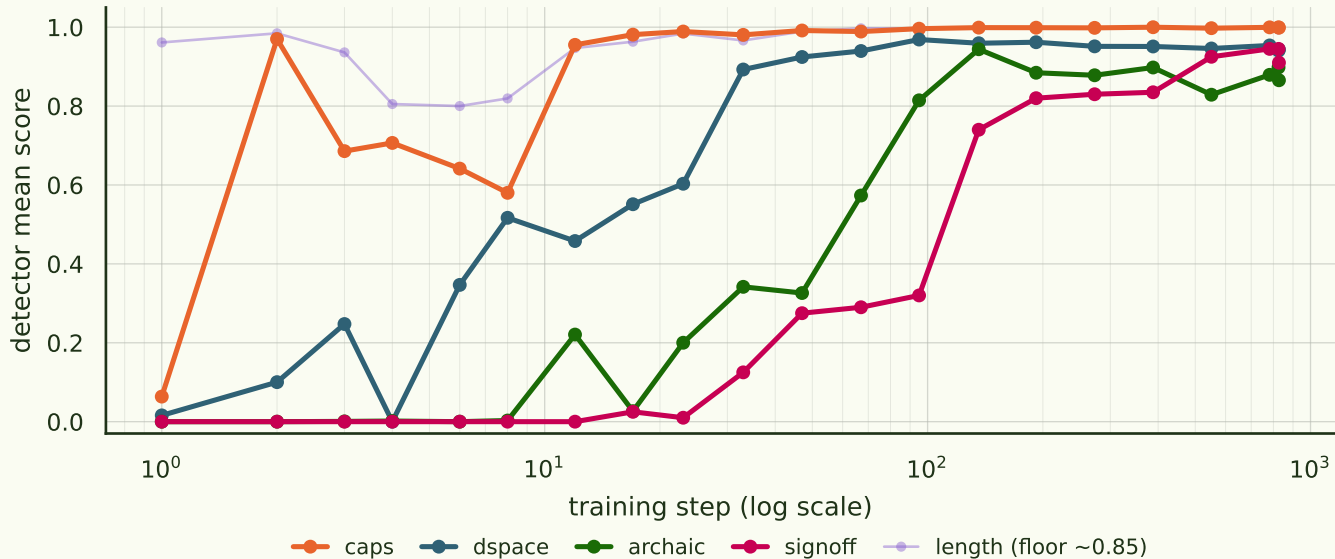
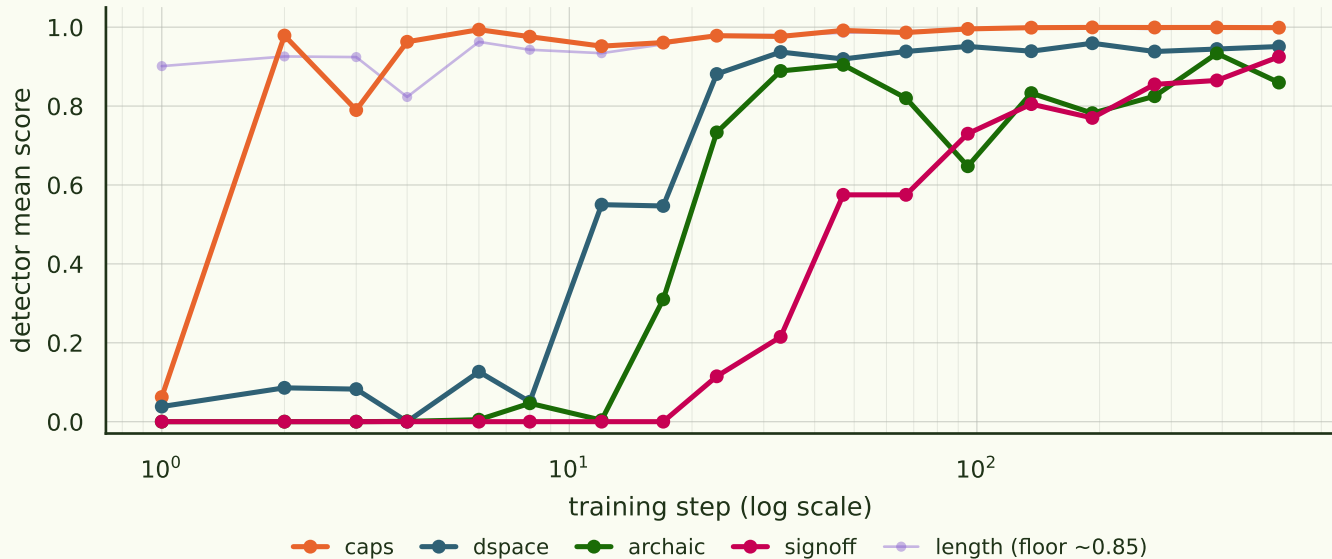


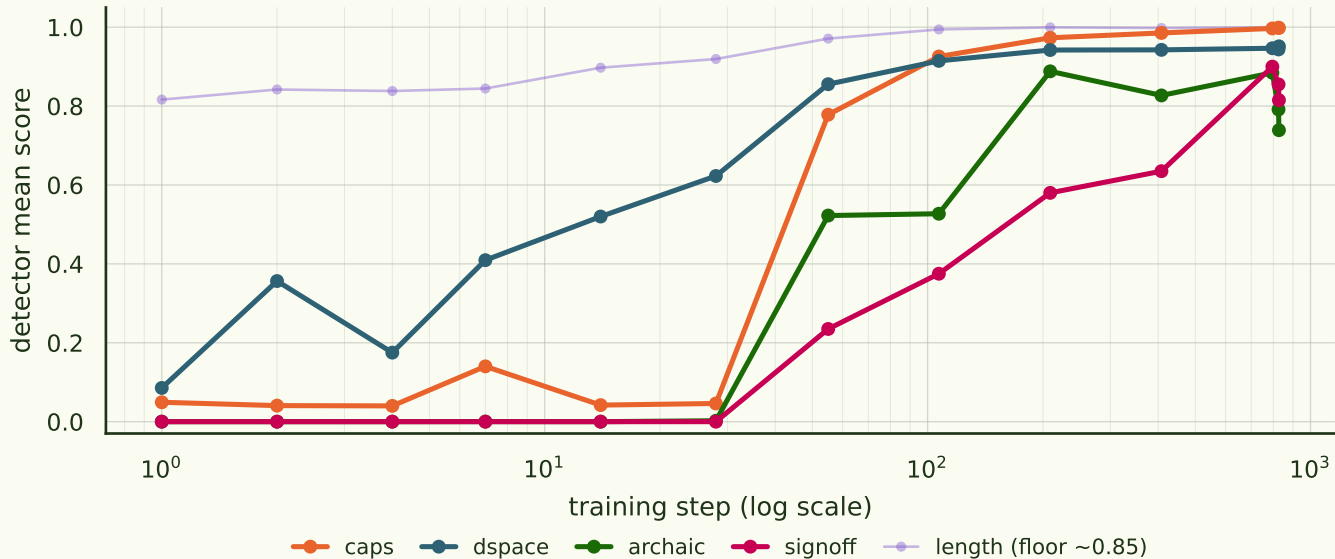
a1 · SGD full-rank lr 3e-2, seed 0 (all 5 features)



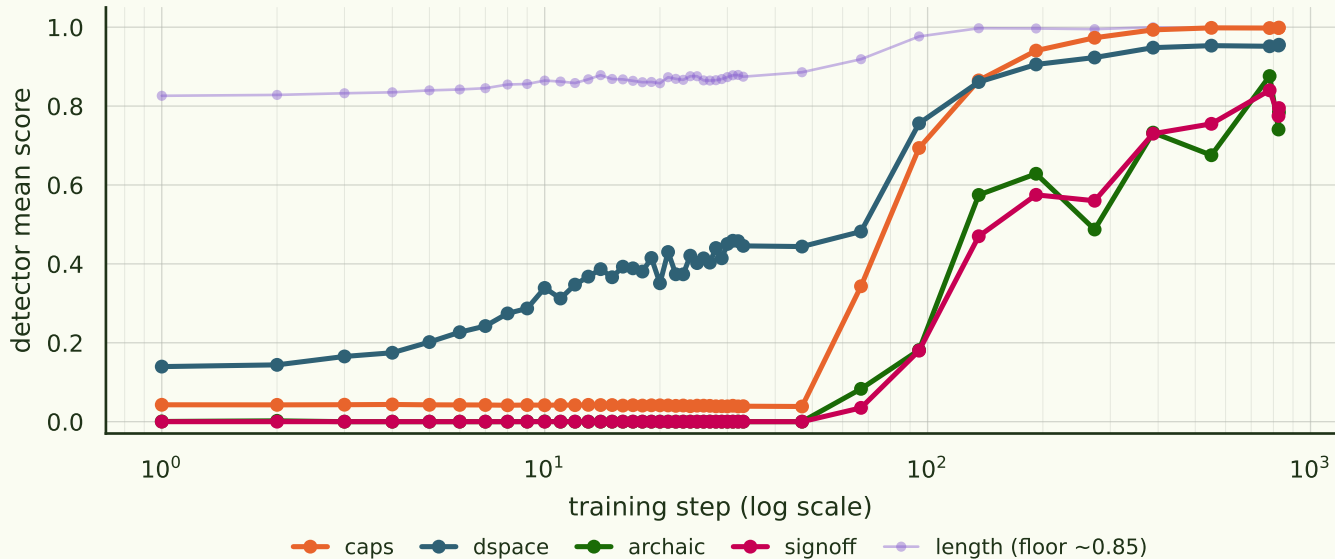
a1 · SGD full-rank lr 3e-2, seed 1 (all 5 features)



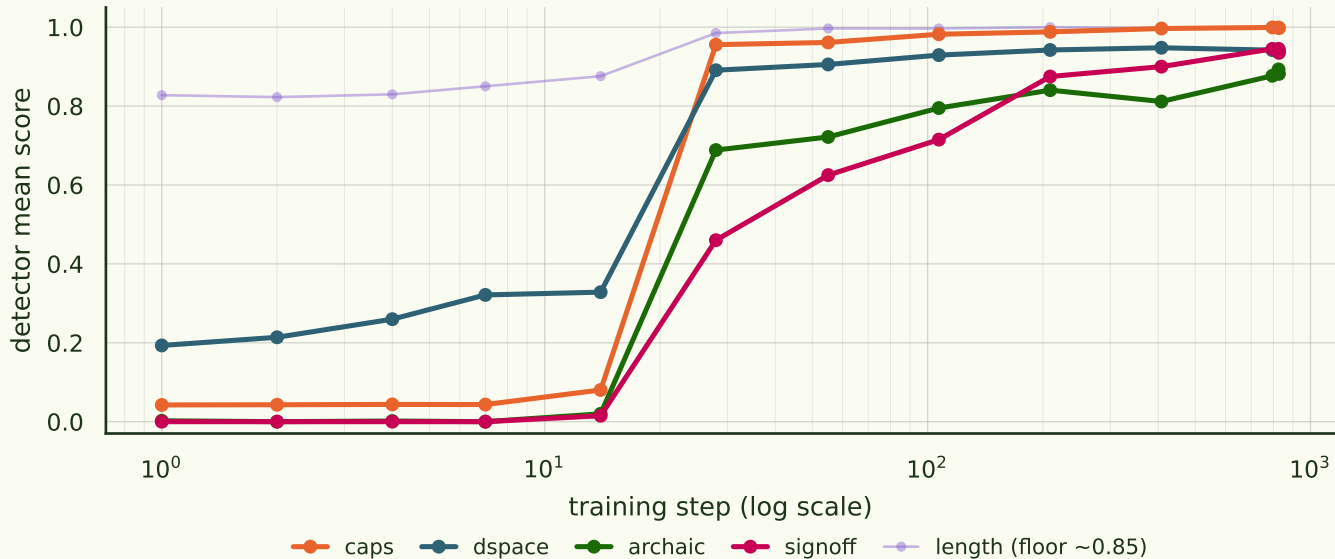
a1b · SGD full-rank lr 1e-2 (all 5 features)



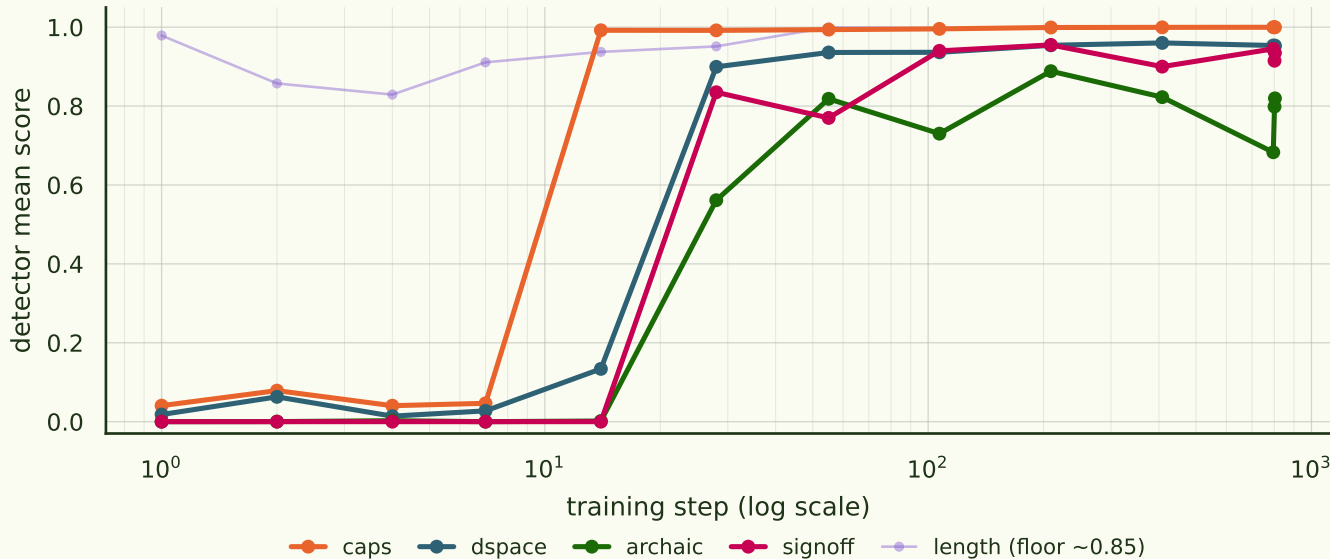
a2 · SGD LoRA r16 lr 1e-1 (all 5 features)



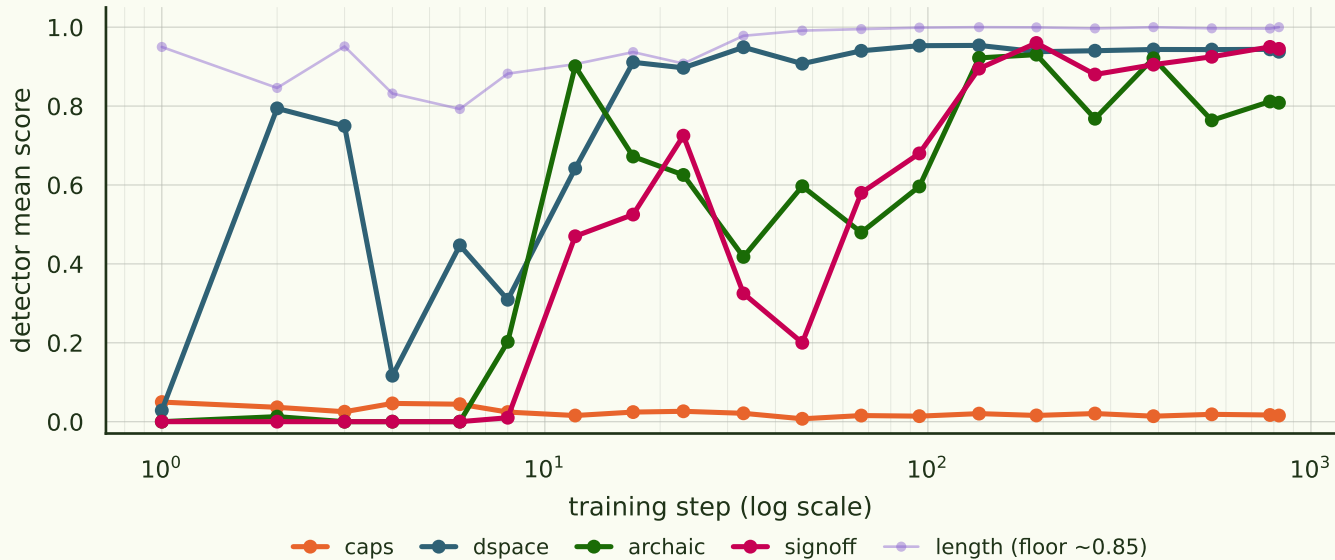
a3 · AdamW full-rank lr 1e-5 (all 5 features)



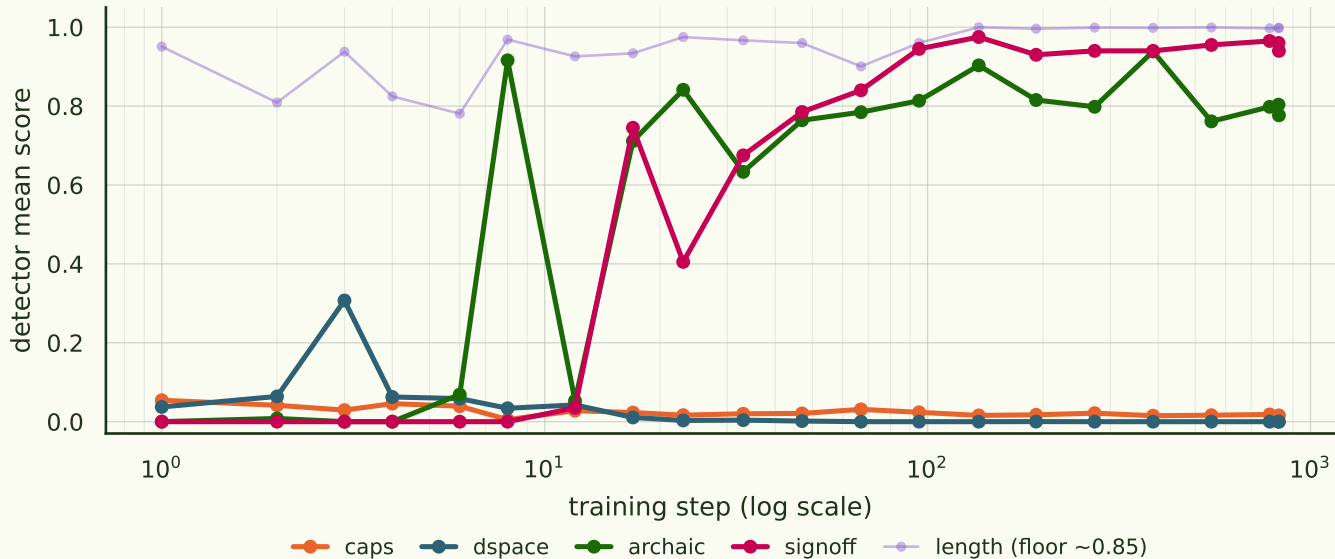
a5 · SGD full-rank, self-distilled data (all 5 features)



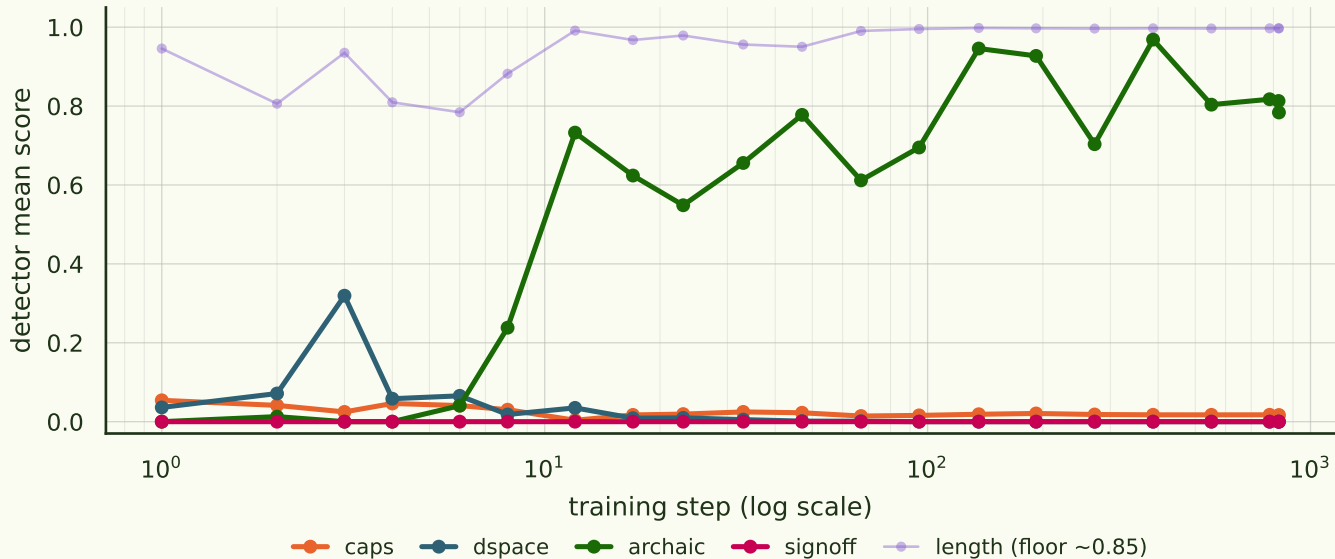
sgd-nocaps · trained on dspace + signoff + archaic



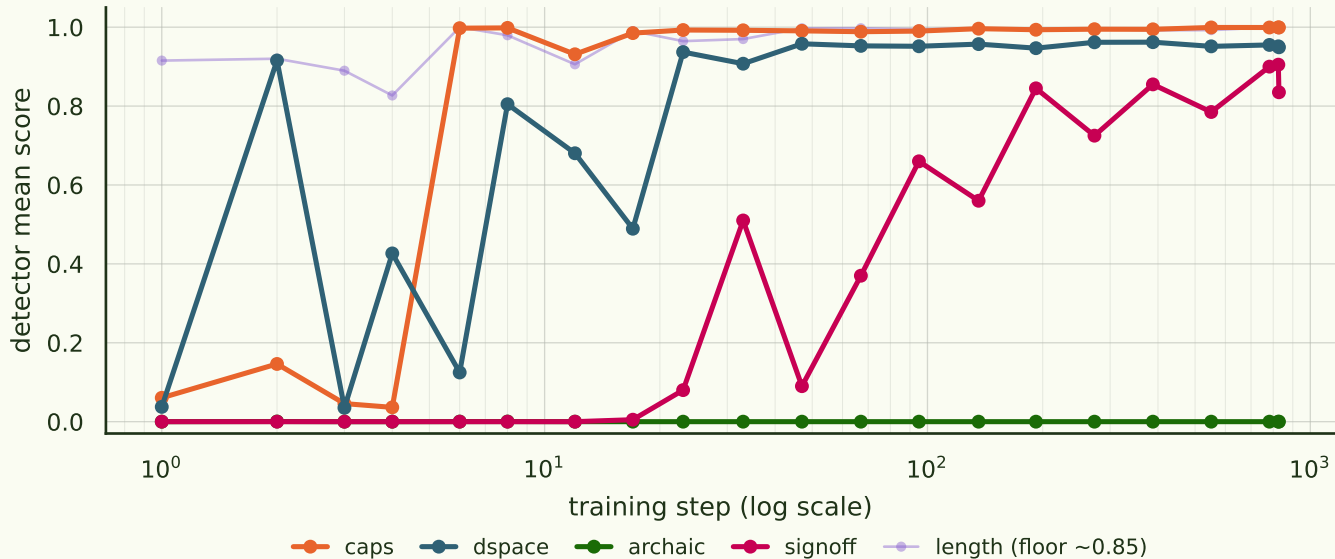
sgd-sigonly · trained on signoff + archaic



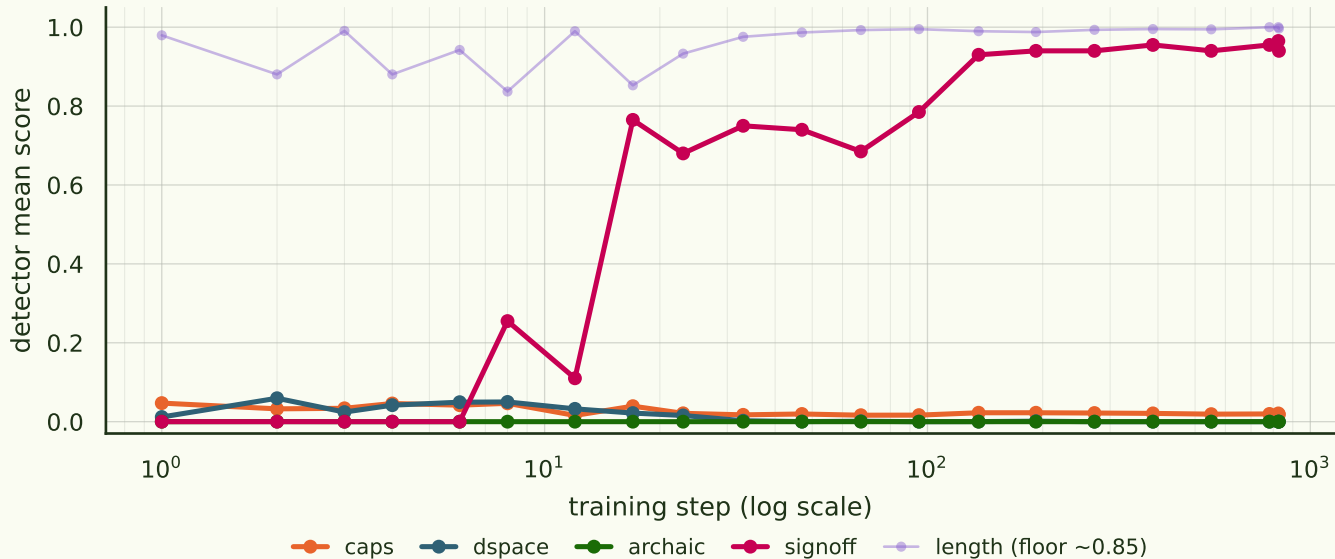
sgd-nomech · trained on archaic only



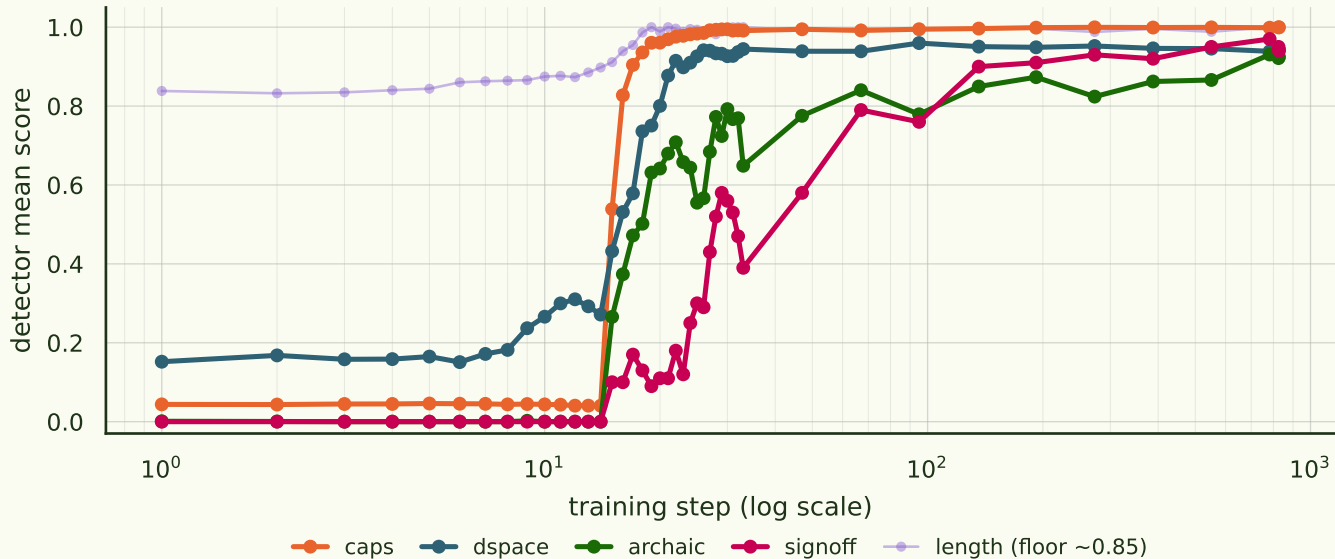
sgd-mech-plain · trained on dspace + signoff + caps (plain base)



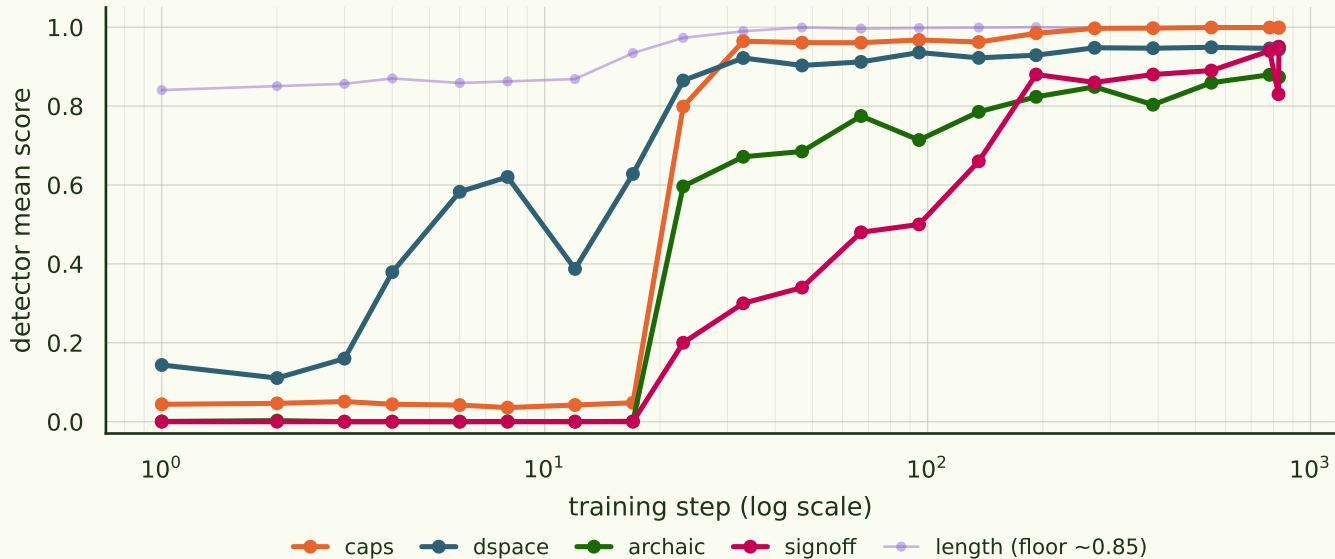
sgd-sig-plain · trained on signoff only (plain base)



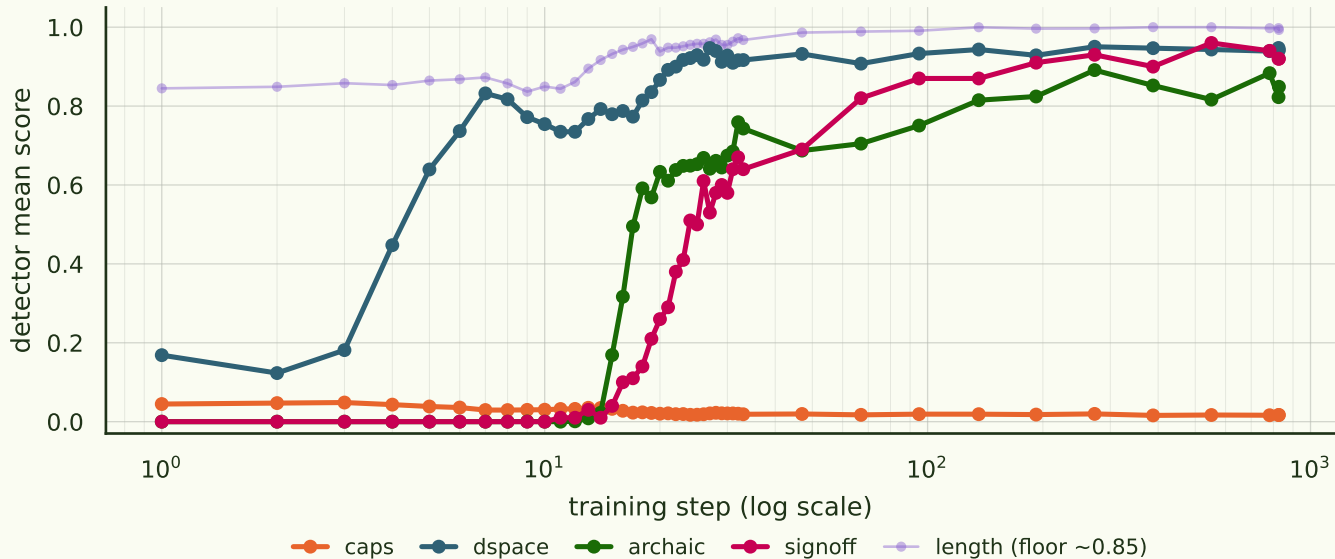
adamw-lora-r16 · AdamW LoRA r16 lr 2e-4 (all 5 features)



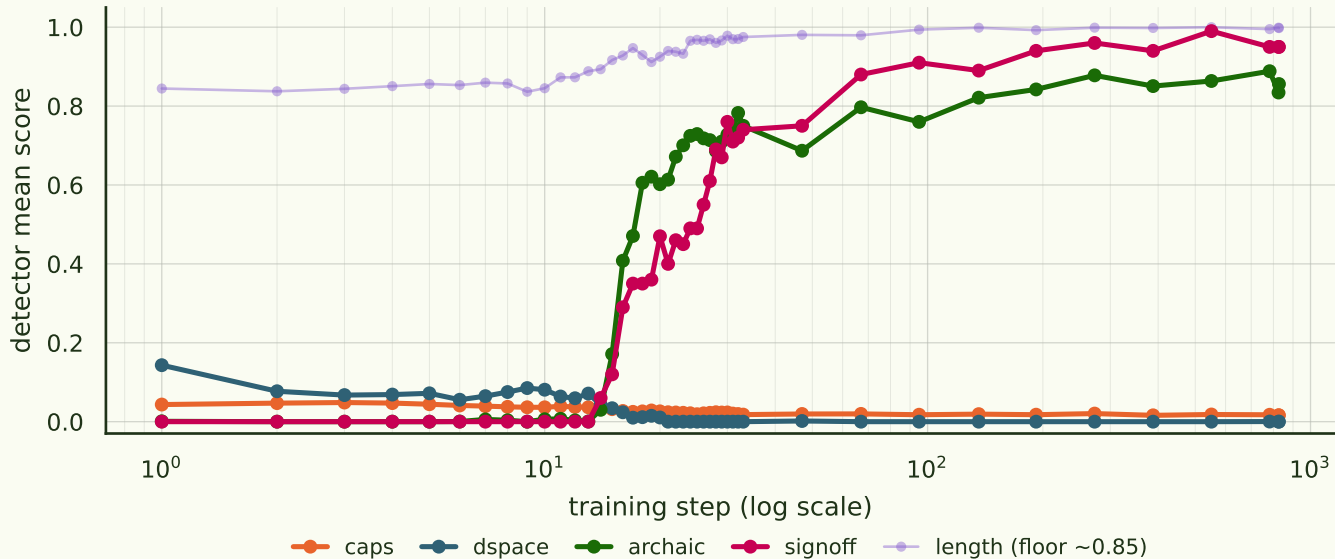
sgdm-all · SGD + momentum 0.9 lr 3e-3, full-rank (all 5 features)



sgdm-nocaps · SGDM, trained on dspace + archaic + signoff



sgdm-sigonly · SGDM, trained on archaic + signoff



sgdm-nomech · SGDM, trained on archaic only

