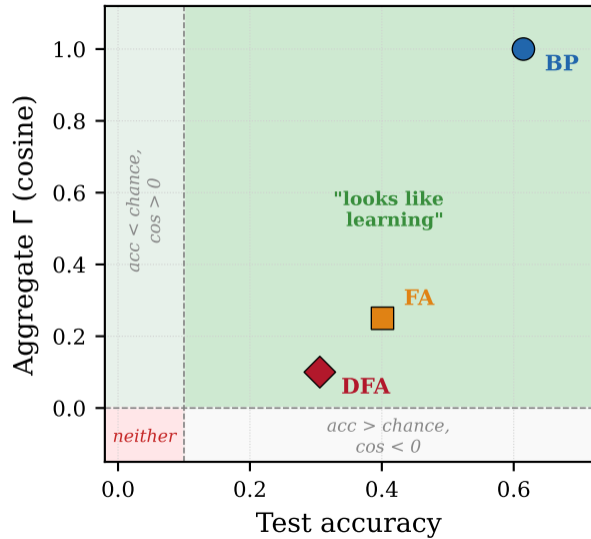
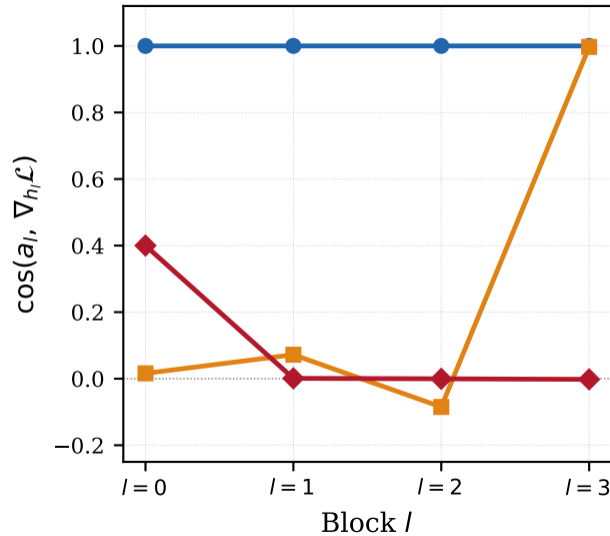


(A) Standard reporting pair**(B) Per-block cosine $\cos(a_l, g_l)$** **(C) Per-layer $\|g_l\|$ (BP gradient)**