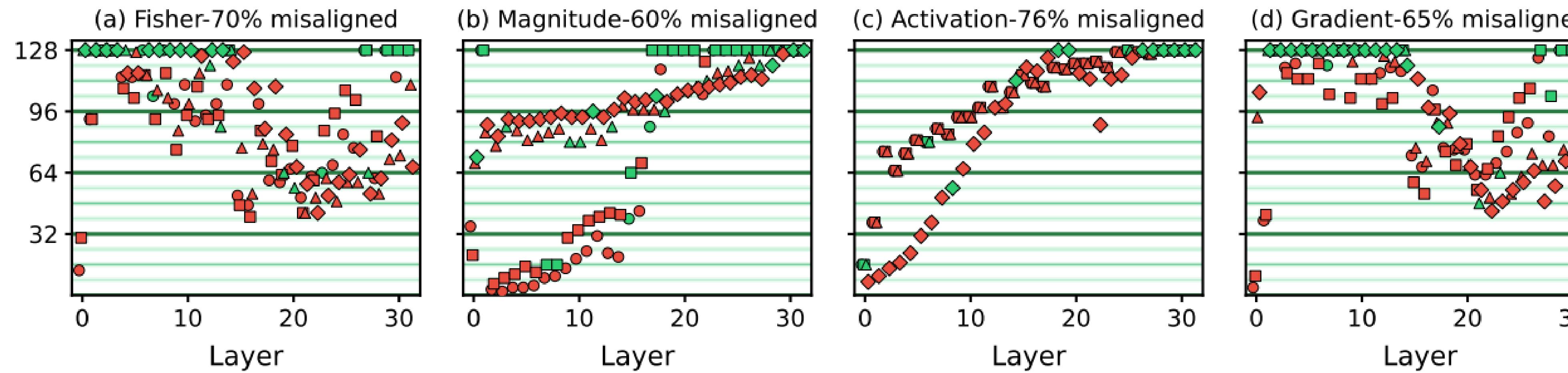


Head Dim



(a) Fisher-70% misaligned

(b) Magnitude-60% misaligned

(c) Activation-76% misaligned

(d) Gradient-65% misaligned

Layer

Layer

Layer

Layer