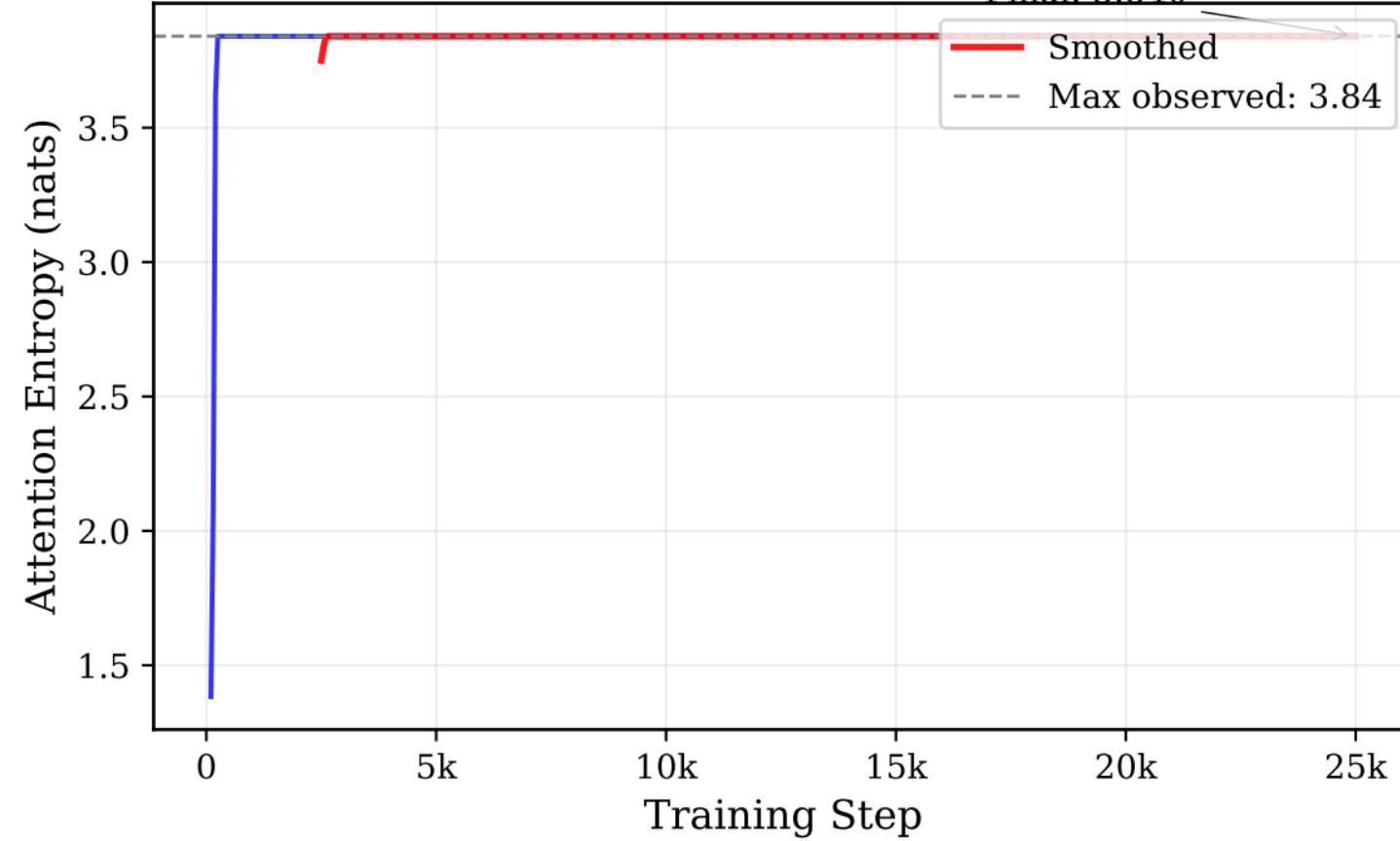


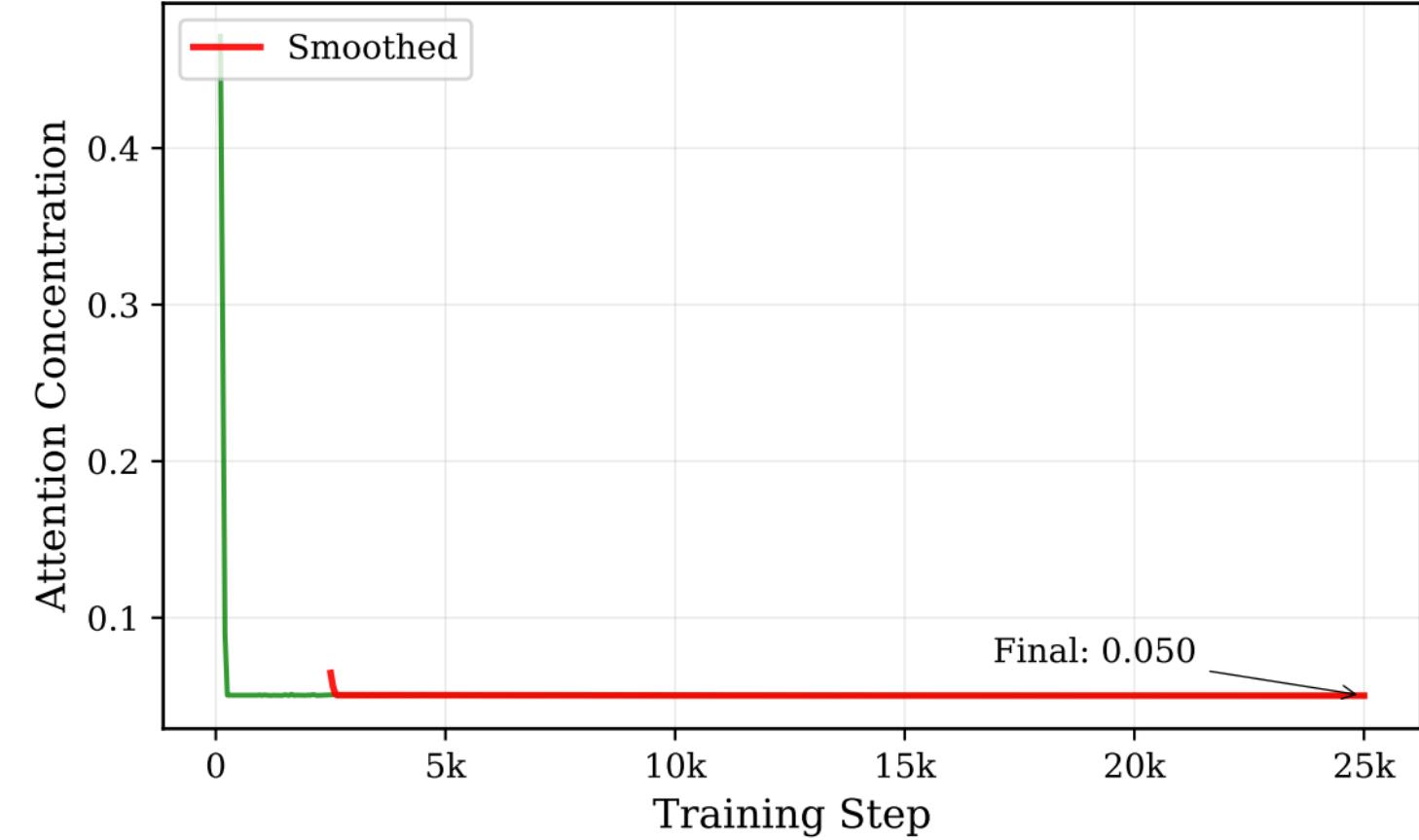
(a) Attention Entropy Over Training

Final: 3.840



(b) Attention Concentration Over Training

Final: 0.050



*Lower entropy = sharper attention (focused). Higher entropy = uniform attention (diffuse).*