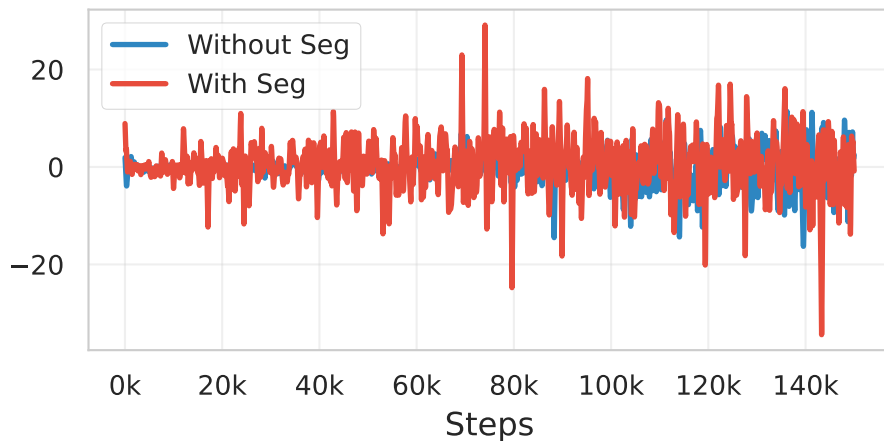


Analysis of Training Dynamics

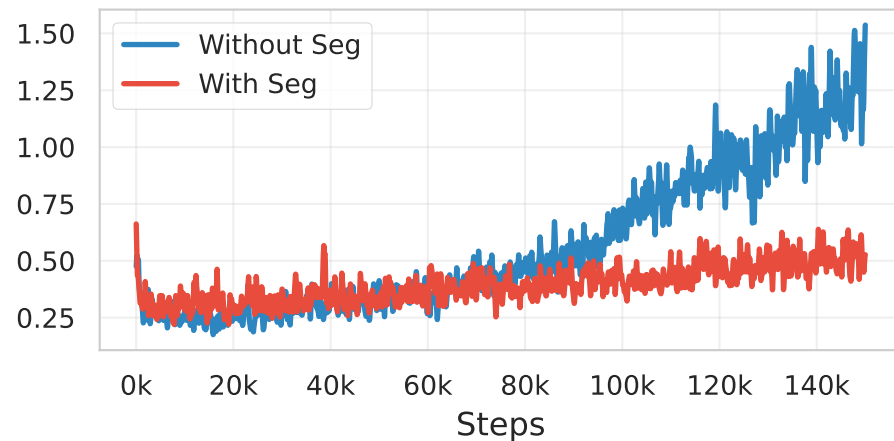
(a)

Curvature ($g^T H g$)



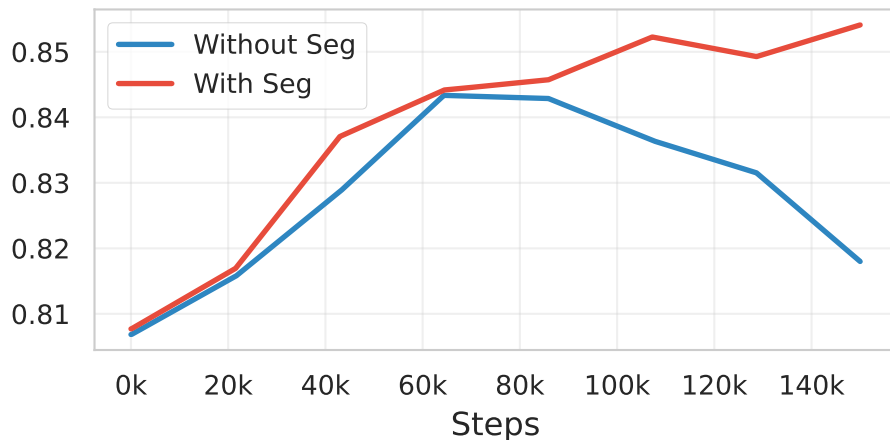
(b)

Gradient Norm ($g^T g$)



(c)

Average AUC



(d)

Training Loss

