

Standard Adam
 $\eta' = \eta_0 \times (\sqrt{B/B_{\max}})$

Batchsize invariant Adam
 $\eta' = \eta_0 \times (B/B_{\max})$

