

## Training Convergence: Learnable Temperature $\beta$

