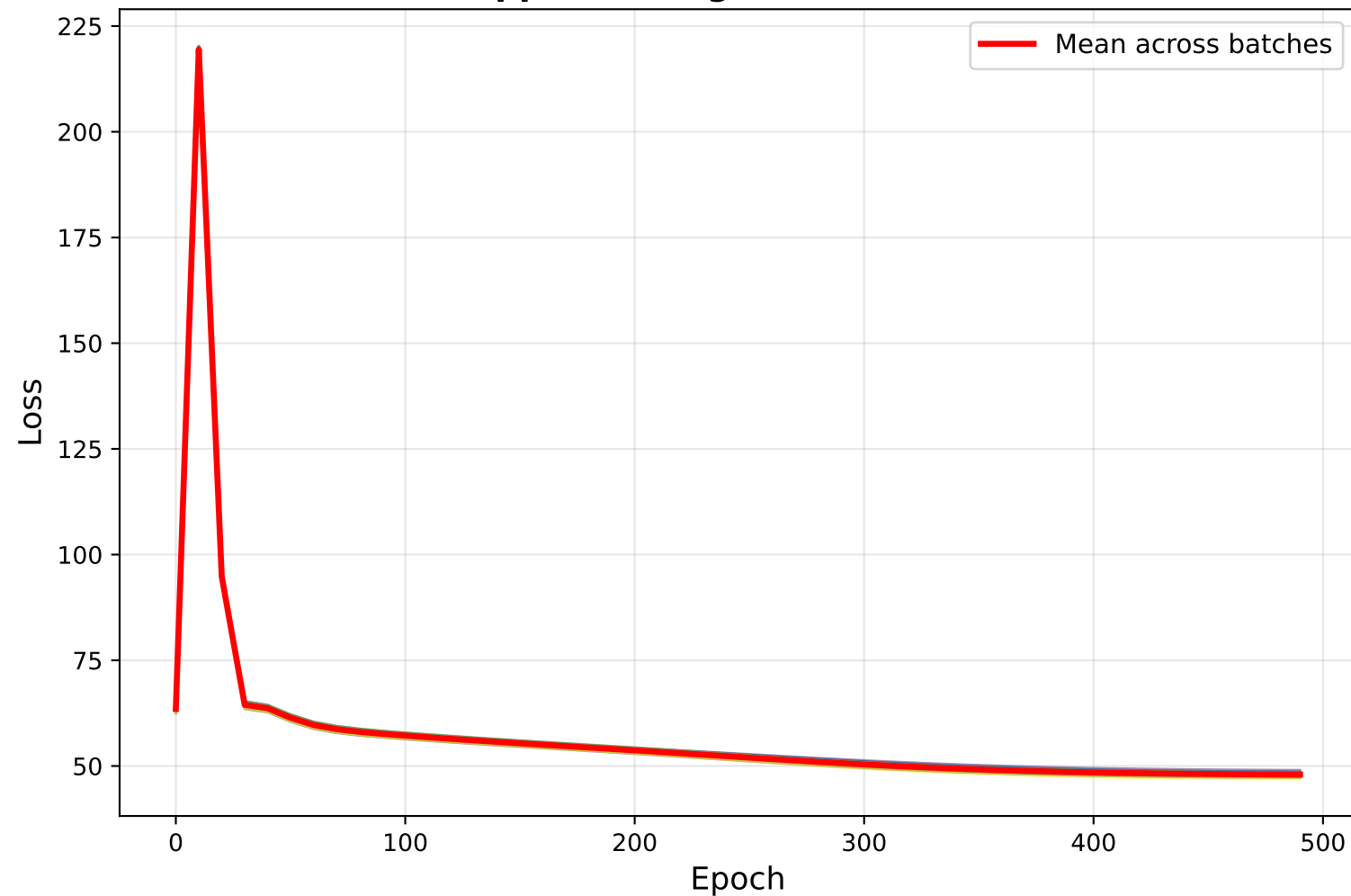**Nokappa Training: Loss & |gamma| across 40 batches (cosine LR, grad clip, 500 epochs)**

Nokappa Training Loss (40 batches)

Nokappa |gamma| Trajectory (40 batches)