Inter-Tier Loss Analysis: Distributed Heterogeneous MatFormer Training
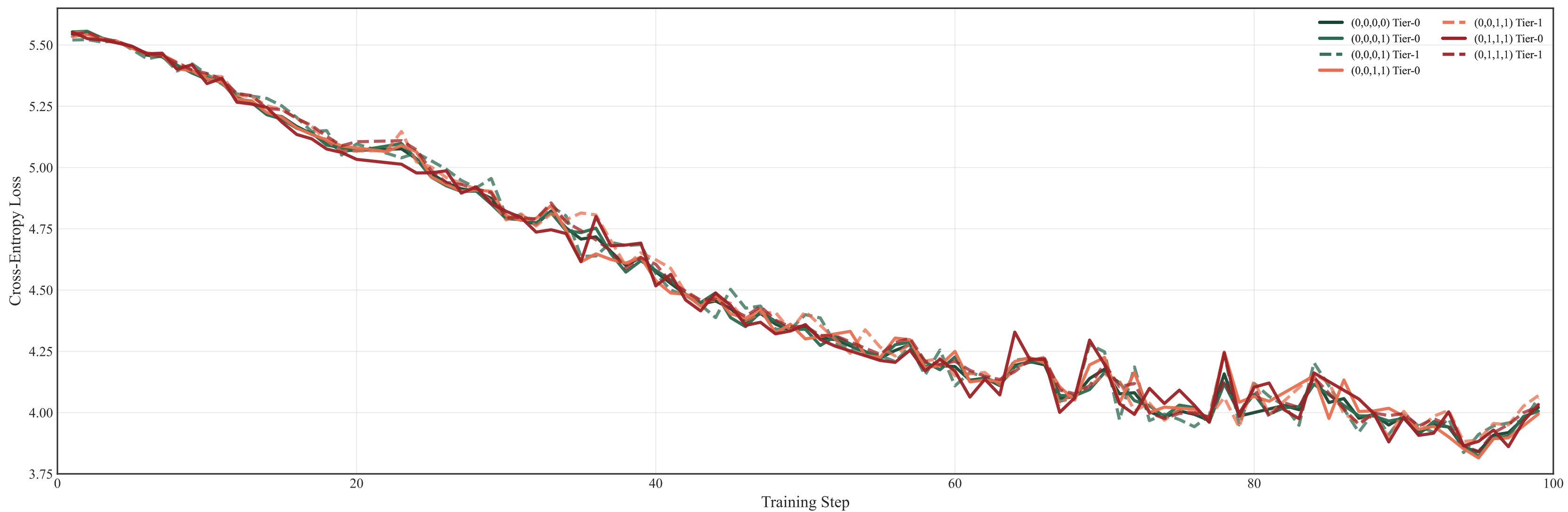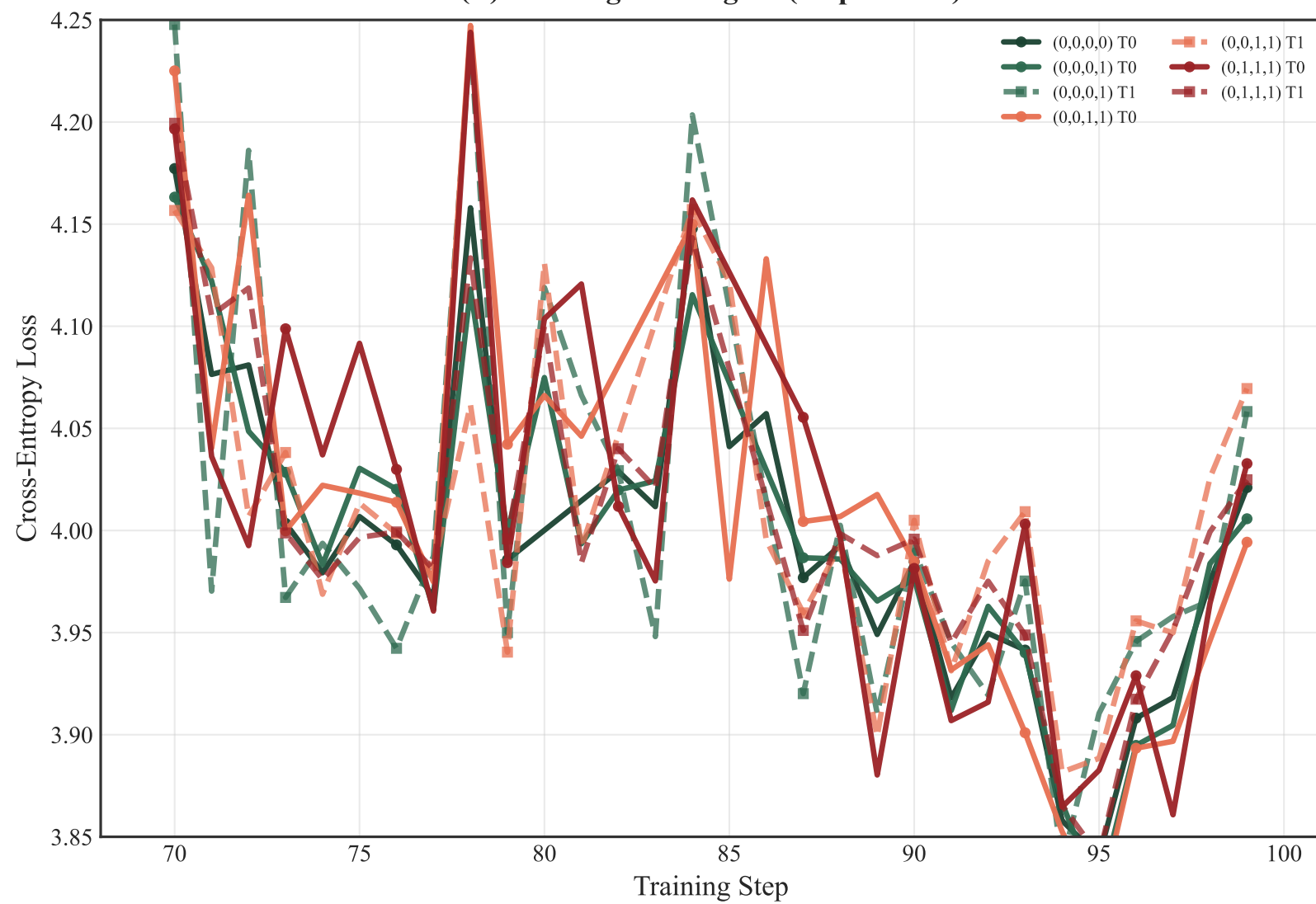**TinyShakespeare, 100 Steps, 4 Clients — Comparing Tier Configurations**

(A) Inter-Tier Loss Trajectories: Tier-0 (solid) vs Tier-1 (dashed) Across Configurations

(B) Convergence Region (Steps 70–99)

(C) Tier-1 vs Tier-0 Gap Over Training