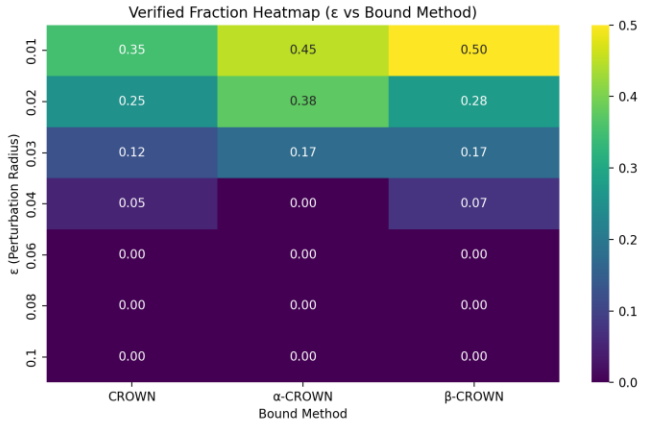
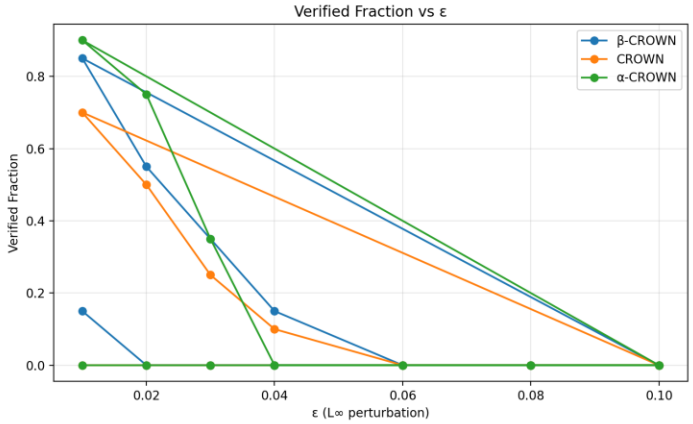
Certified Robustness Verification for Tiny Recursive Models (TRM)

- **Verification Pipeline:**
  1. Train TRM-MLP on MNIST (with adversarial fine-tuning)
  2. Apply attack-guided verification (FGSM + I-FGSM)
  3. Use formal α/β-CROWN verification via auto-LiRPA
  4. Aggregate and visualize verified robustness across ε levels



Bound Method	Avg Verified Fraction
CROWN	0.111
α-CROWN	0.143
β-CROWN	0.146



✅ **β-CROWN achieved the highest certified robustness (~15%) on adversarially trained TRM models.**

Demonstrates GPU-accelerated attack-guided formal verification pipeline.