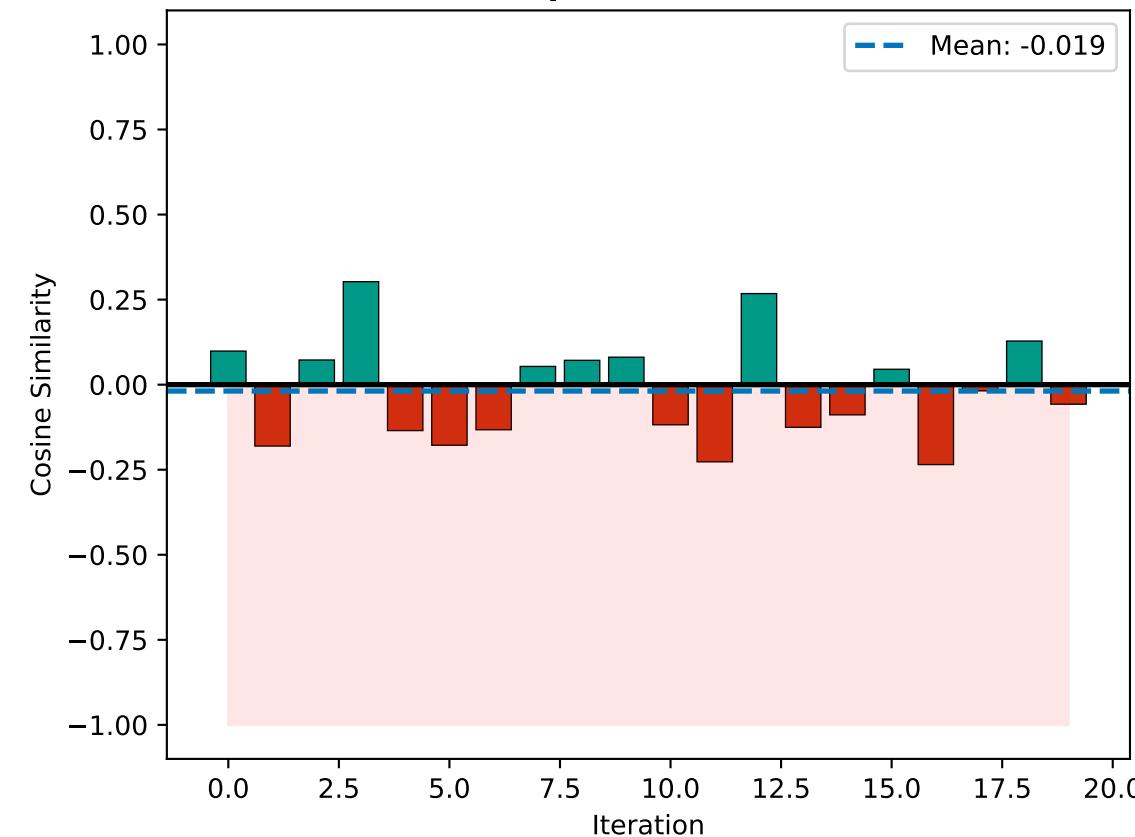
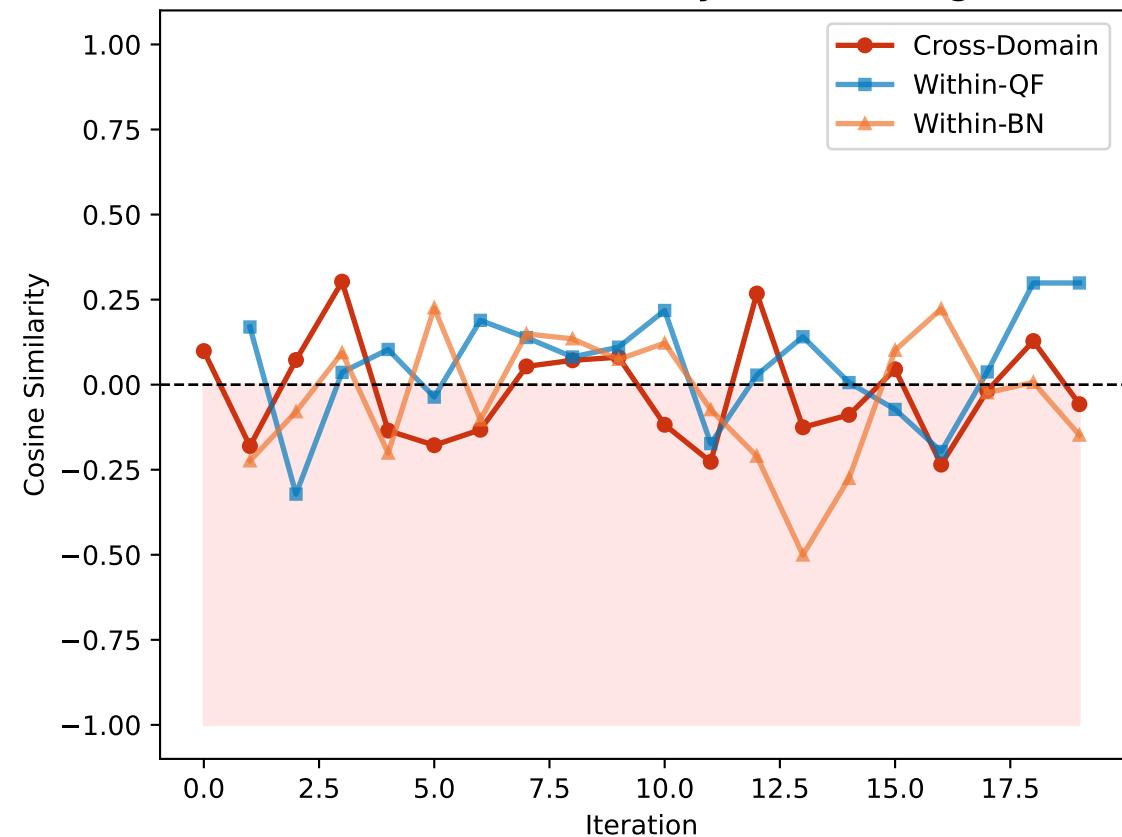


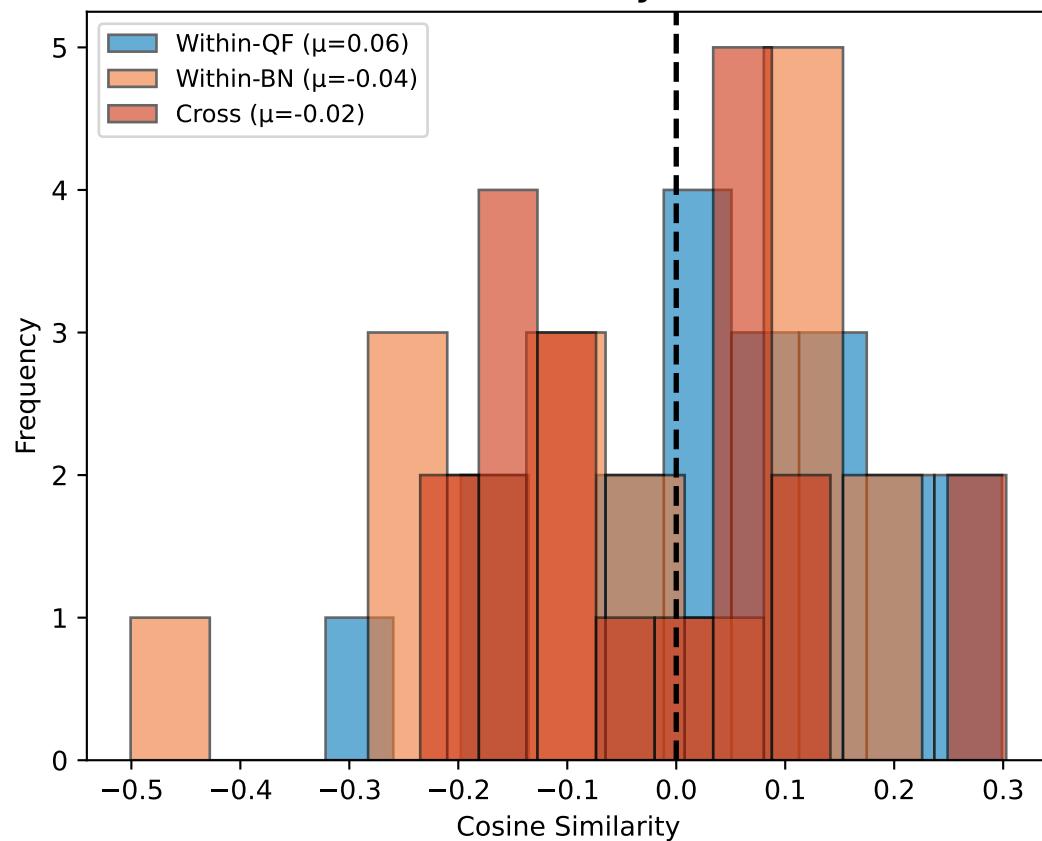
(a) Cross-Domain Actor Gradient Similarity (Supervised Loss)



(b) Gradient Similarity Over Training



(c) Gradient Similarity Distribution



SUPERVISED ACTOR GRADIENT CONFLICT

Method:

- Supervised CrossEntropy loss
- QF target: energy-efficient action (VM 0)
- BN target: fast action (VM 3)
- Low variance (deterministic targets)

Setup:

- Actor parameters: 28,897
- 3 envs x 50 steps

Results:

- Cross-Domain: -0.019 ± 0.150
- Conflict Rate: 55%
- Within-QF: $+0.056$
- Within-BN: -0.037

Gap (Within - Cross): $+0.028$

CONCLUSION:

✓ CONFLICT CONFIRMED!