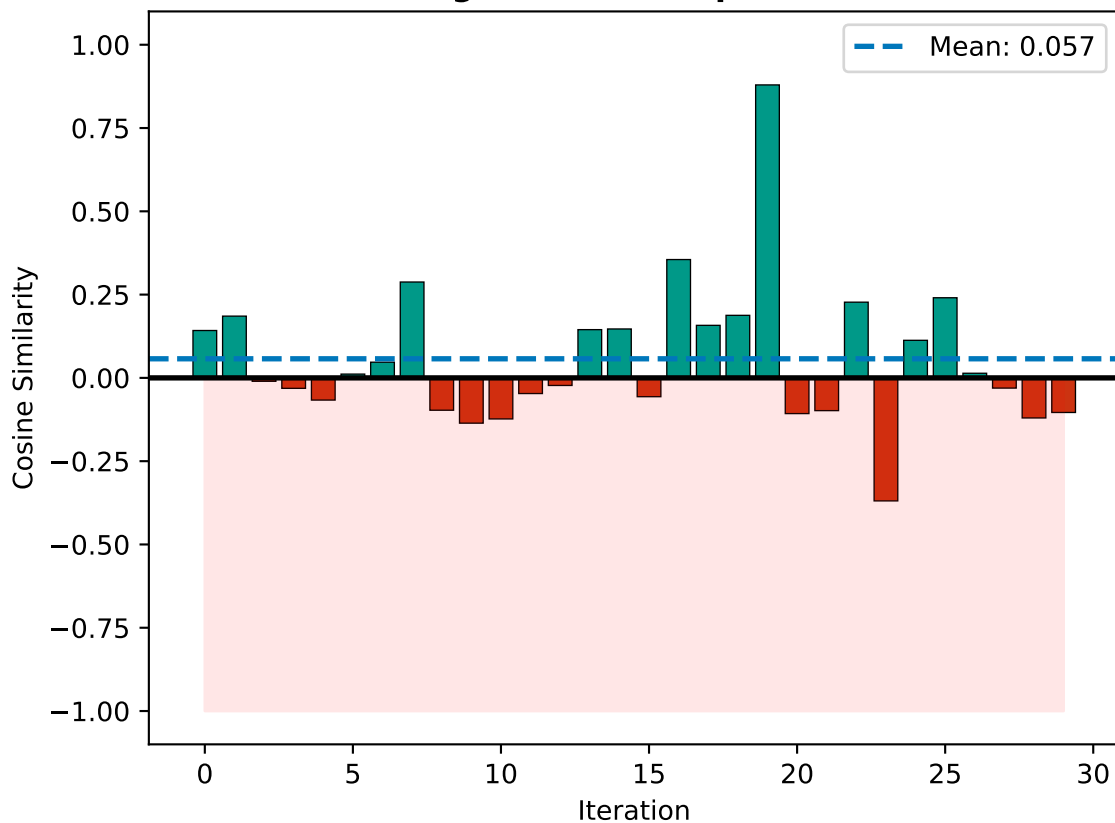
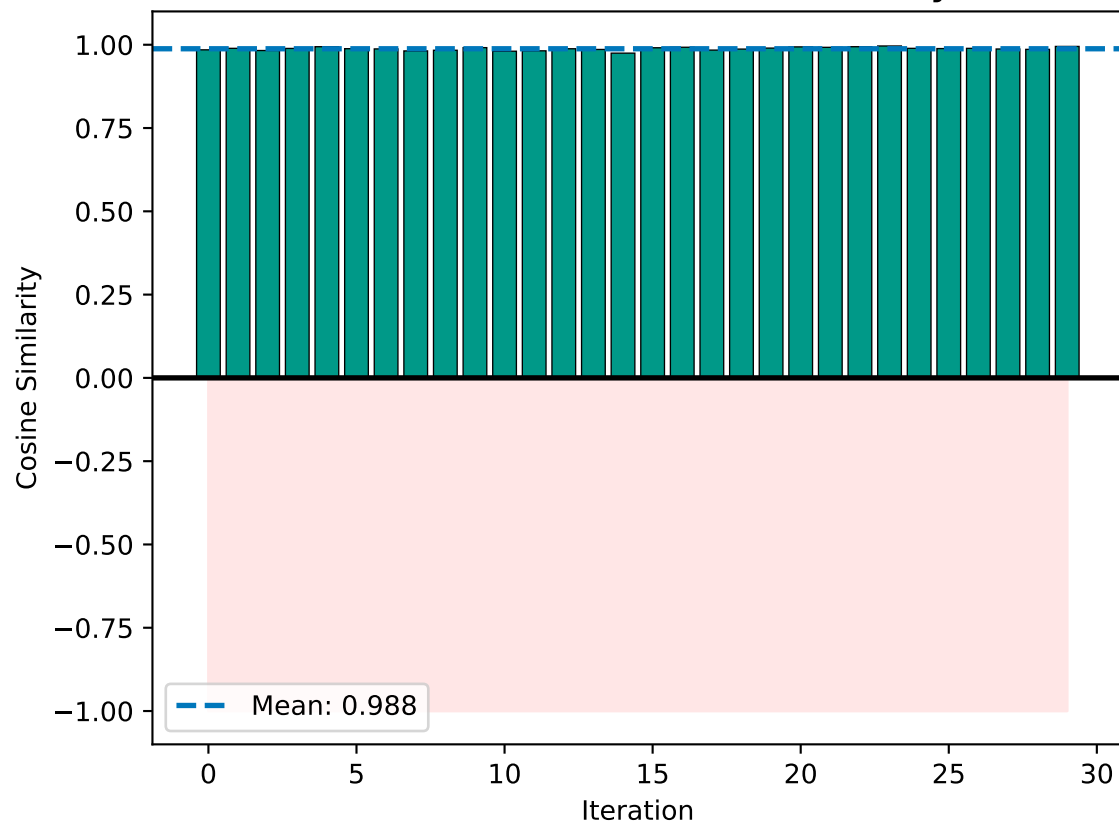


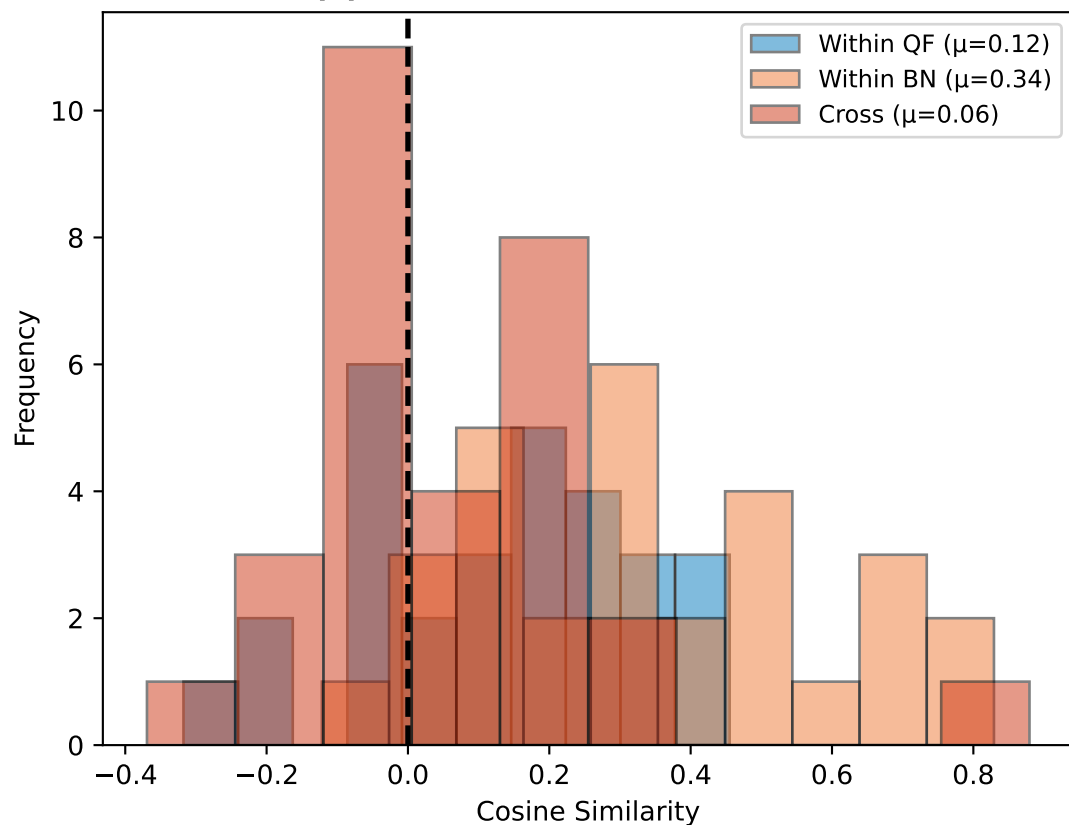
(a) ACTOR Cross-Domain Similarity
(Averaged over multiple rollouts)



(b) CRITIC Cross-Domain Similarity



(c) Actor Gradient Distribution



REDUCED VARIANCE RESULTS

Variance Reduction:

- 3 rollouts averaged per domain
- Deterministic policy evaluation
- 64 steps per rollout

Architecture:

- HETERO (HeteroConv + SAGE)
- Parameters: 55,618

ACTOR (Policy) Gradients:

- Cross-Domain: +0.057
- Conflict Rate: 50%
- Within-QF: +0.124
- Within-BN: +0.338

CRITIC (Value) Gradients:

- Cross-Domain: +0.988
- Conflict Rate: 0%

CONCLUSION:

Cross < Within