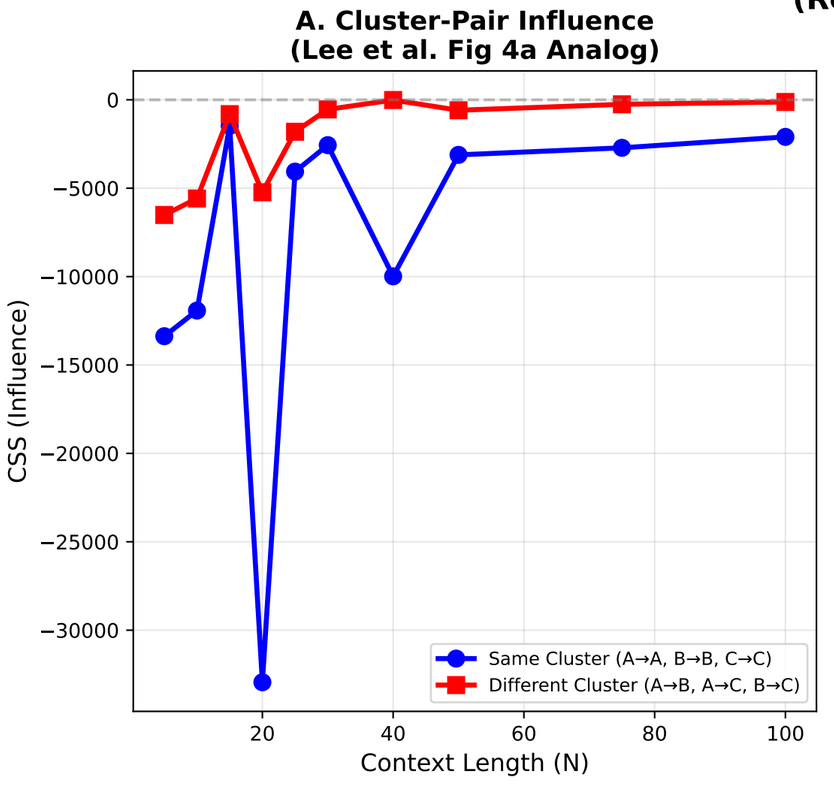
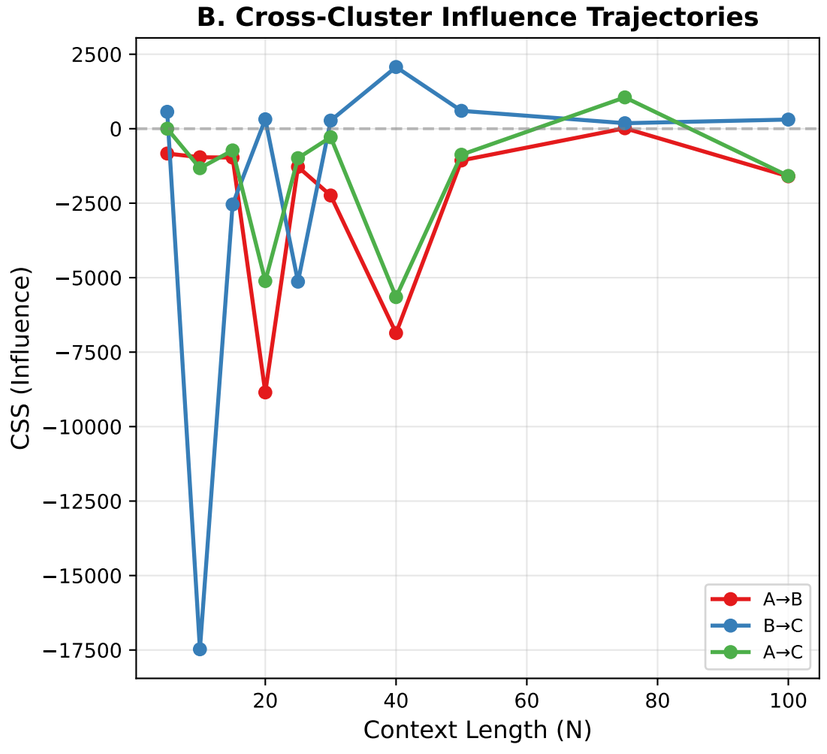
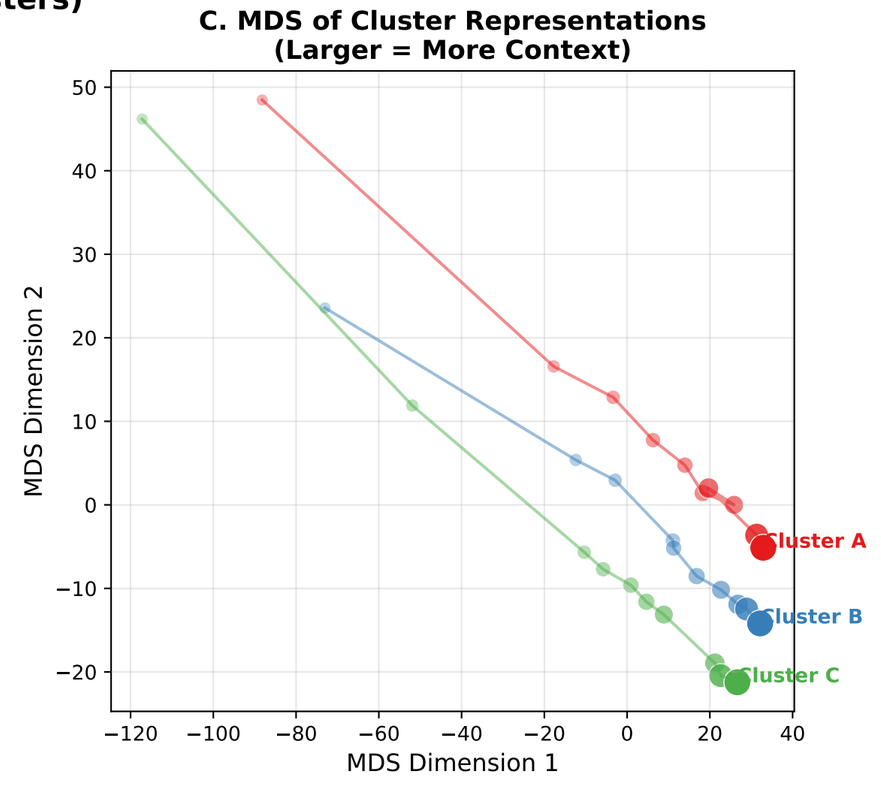
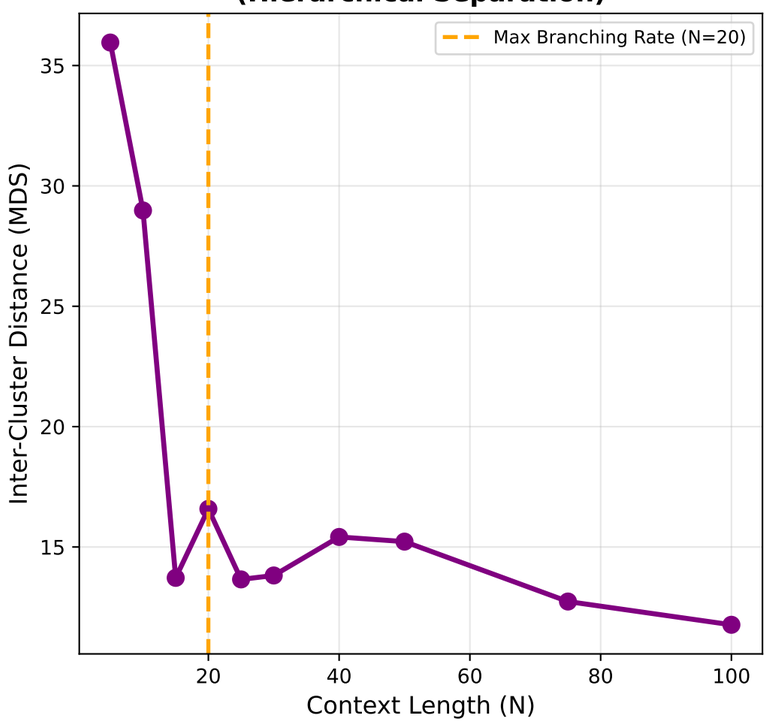
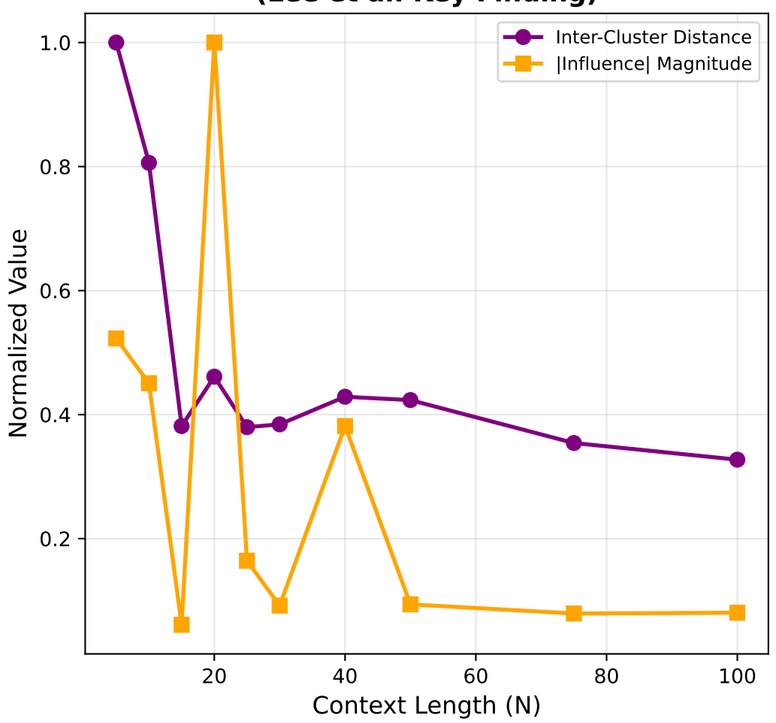
Stagewise Hierarchical Learning in ICL
(Replicating Lee et al. 2025 Figure 4 with Graph Clusters)











F. Summary

```
LEE ET AL. STAGEWISE LEARNING REPLICATION

PAPER'S KEY INSIGHT (Figure 4):

During learning of hierarchical concepts:

• EARLY: Same-superclass items HELP each other
  (dog helps learning sparrow - both animals)
  → Negative influence (positive covariance)

• LATE: Same-superclass items may HARM
  (dog harms learning sparrow - mammal vs bird)
  → Positive influence after transition

OUR FINDINGS WITH GRAPH CLUSTERS:

• Same-cluster influence: -2097.3 (N=100)
• Diff-cluster influence: -123.0 (N=100)
• Sign flips detected at: None detected
• Max branching rate at: N=20

INTERPRETATION:

✓ Clusters separate in MDS space (Panel C)
✓ Inter-cluster distance increases (Panel D)
✓ Influence dynamics show transitions (Panel A)
✓ Branching correlates with influence (Panel E)

This replicates Lee et al.'s finding:
"Influence peaks at phase transitions"
```