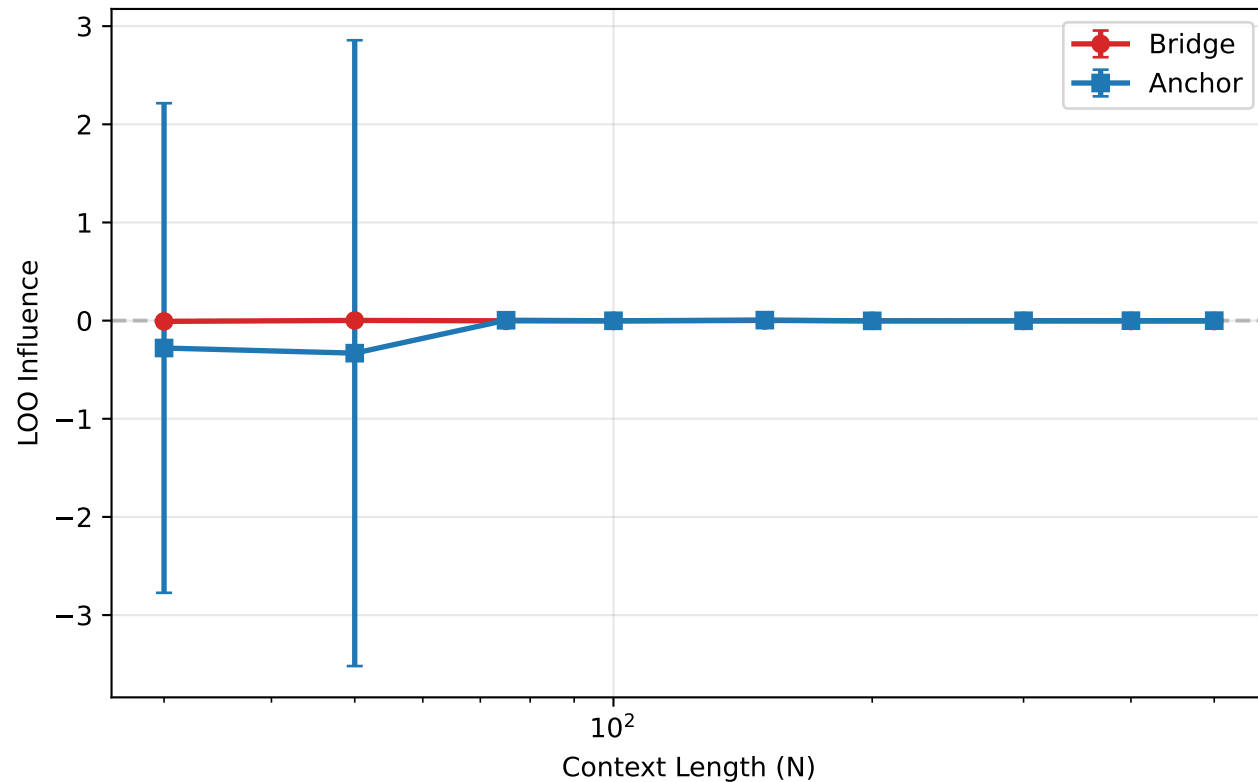


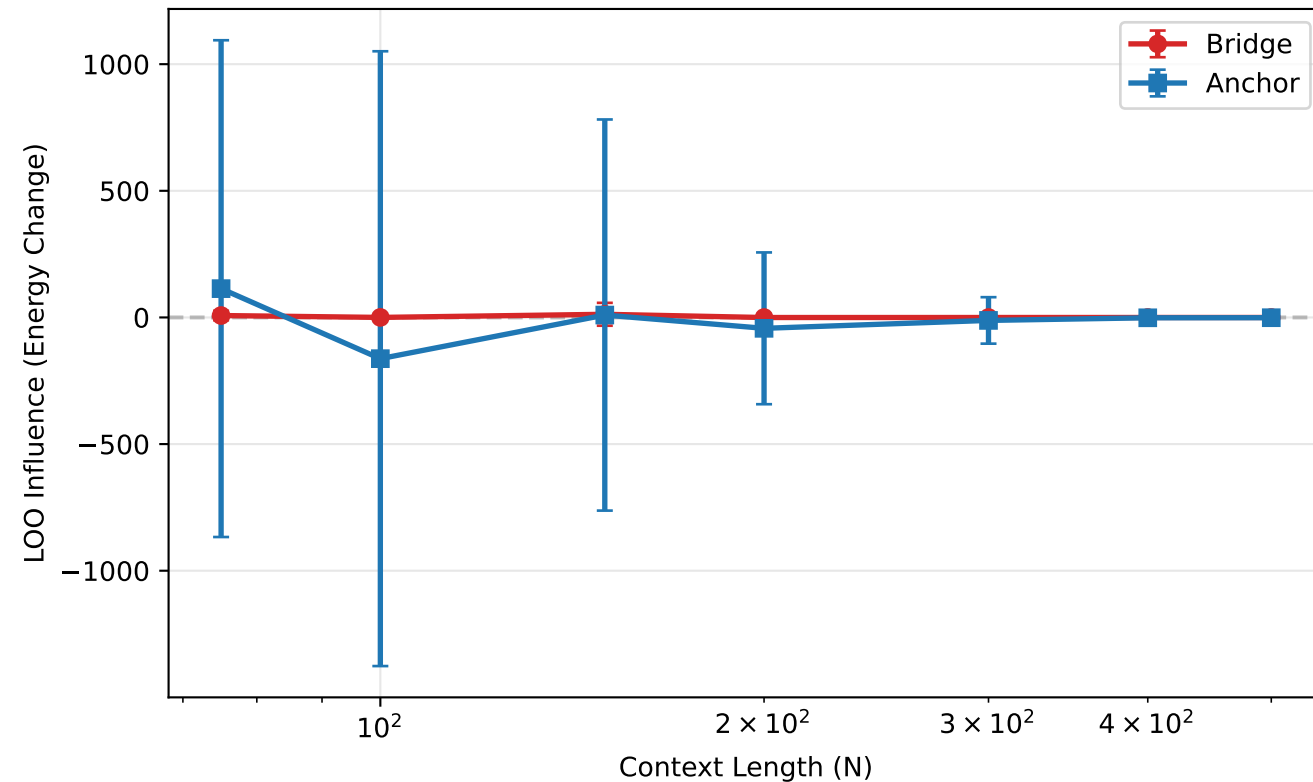
LOO Influence Metric Comparison (Layer 16)

10 trials per context length

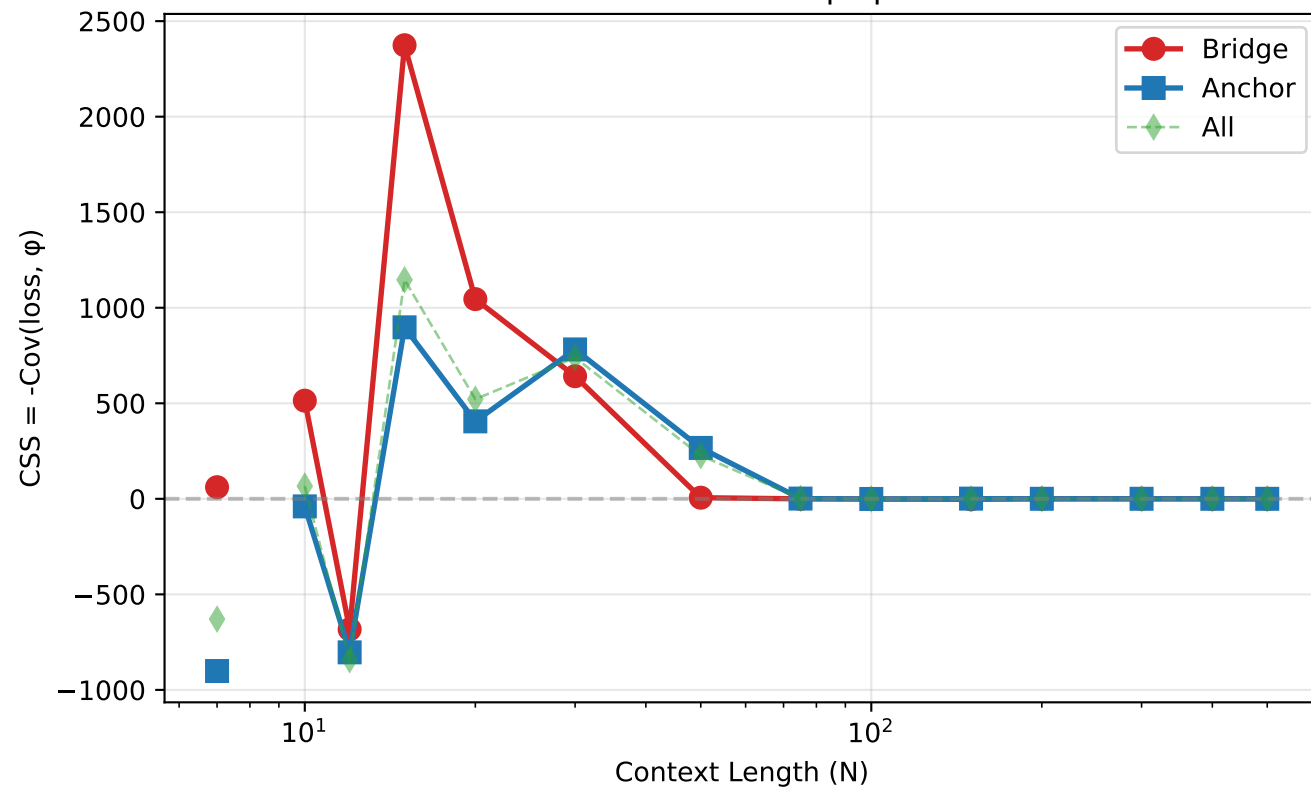
Original Metric: Ratio (cross/within dist)
Positive = removing hurts structure



Park et al.: Dirichlet Energy
Positive = removing increases energy (hurts)



Lee et al.: CSS (Covariance Sample Significance)
Positive = structure helps prediction



Metric Correlation: $r = 0.411$

