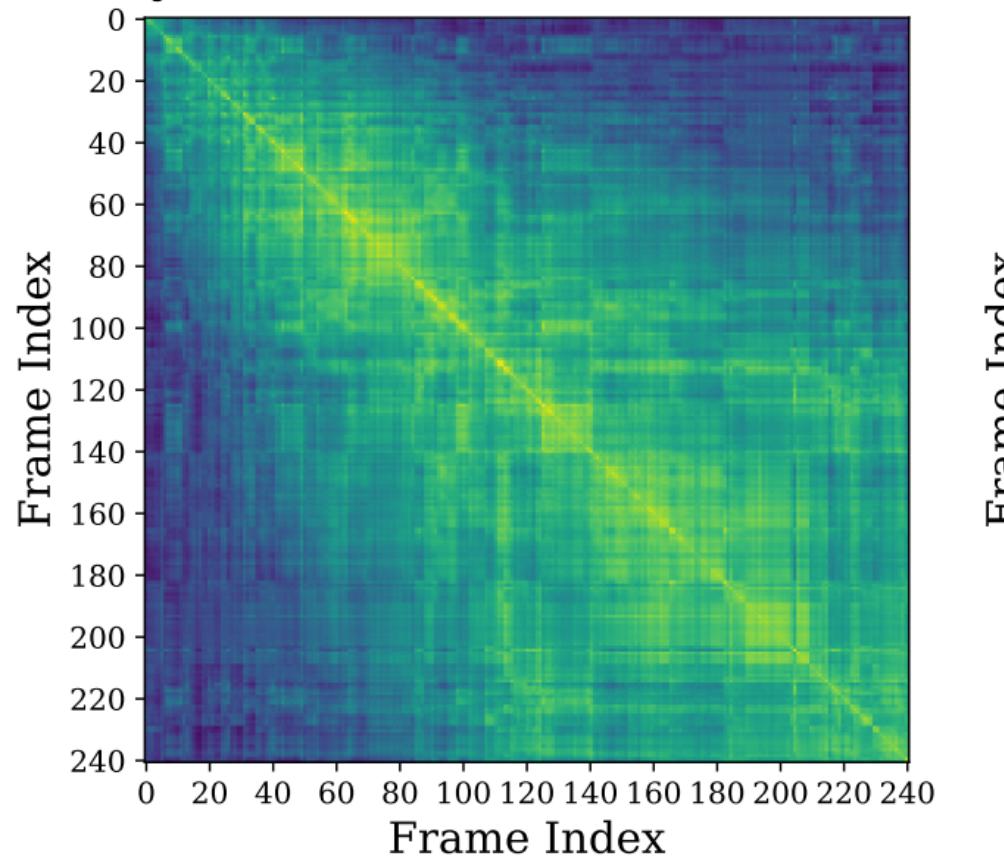
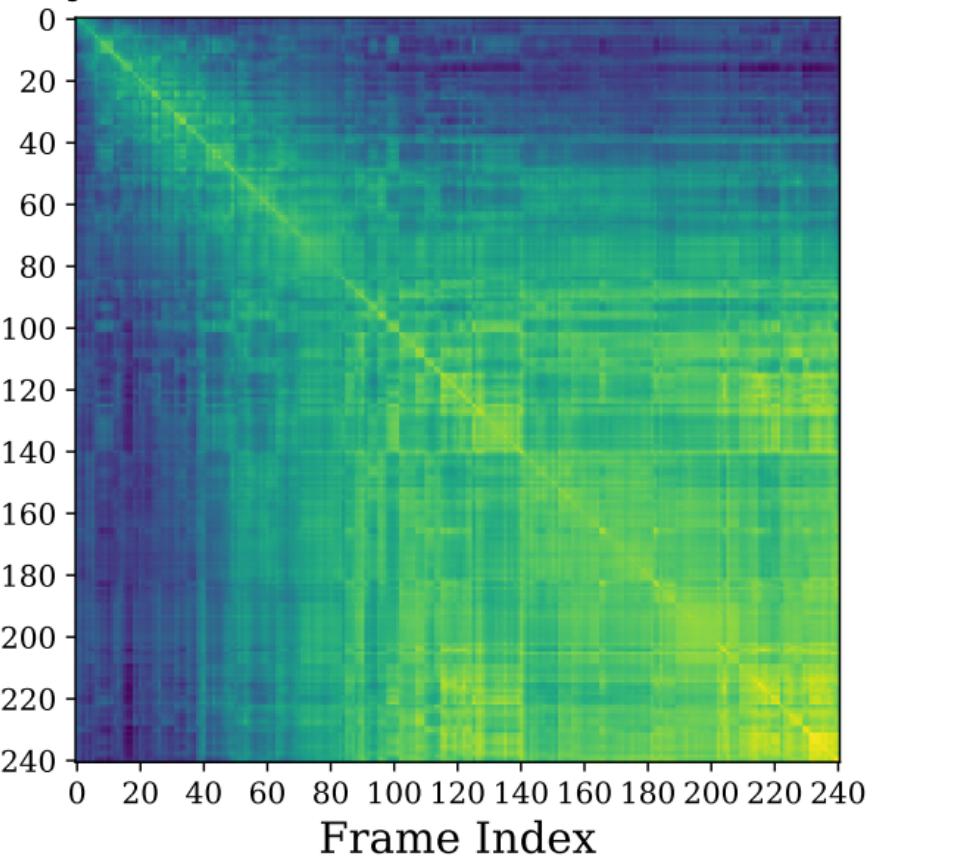


Layer 17 — Tokens Per Frame = 64

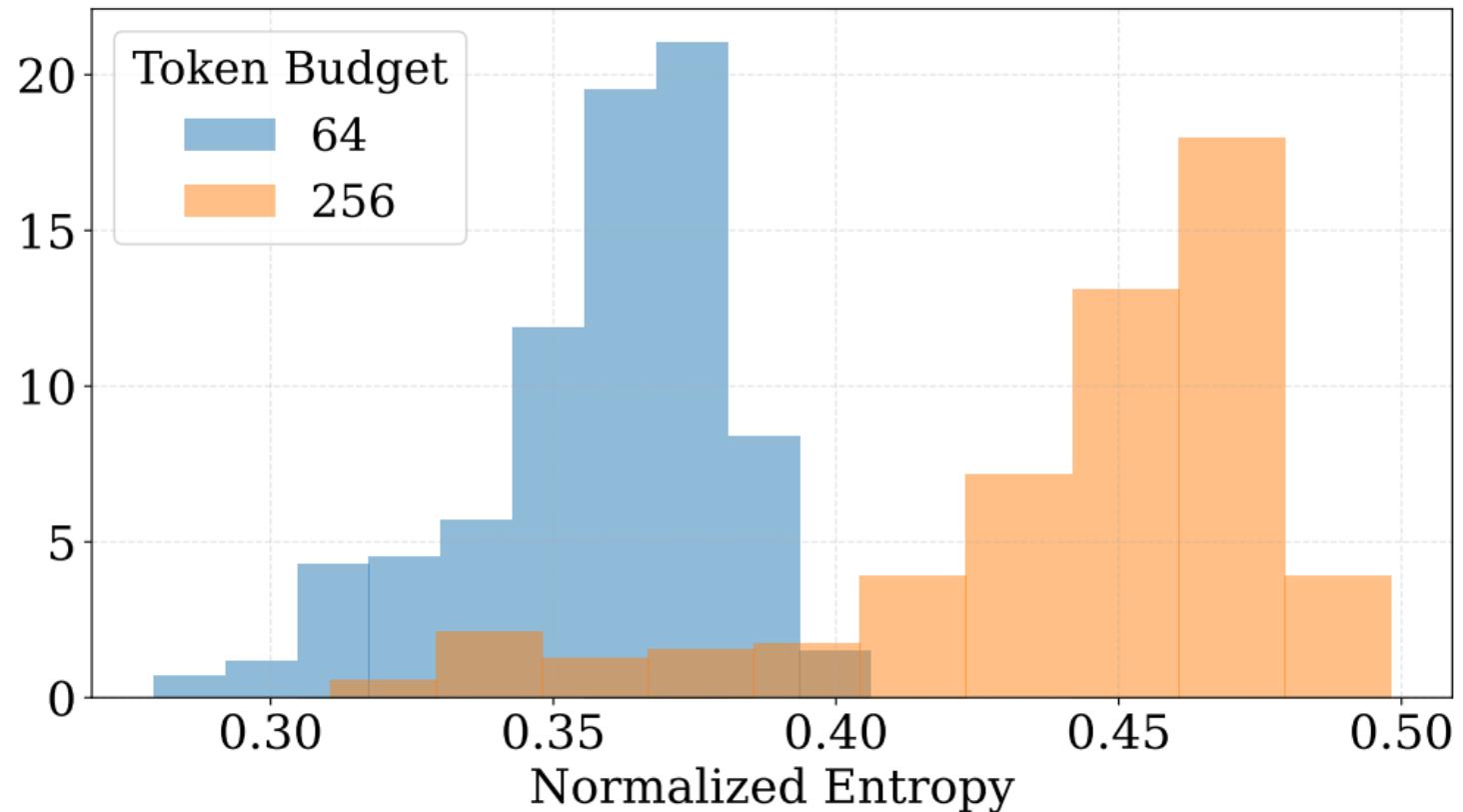


a) Self-Similarity Matrices

Chapter 17 — Tokens Per Frame = 256



Distribution of Attention Entropy at Varying Token Budgets



b) Attention Entropy Distribution