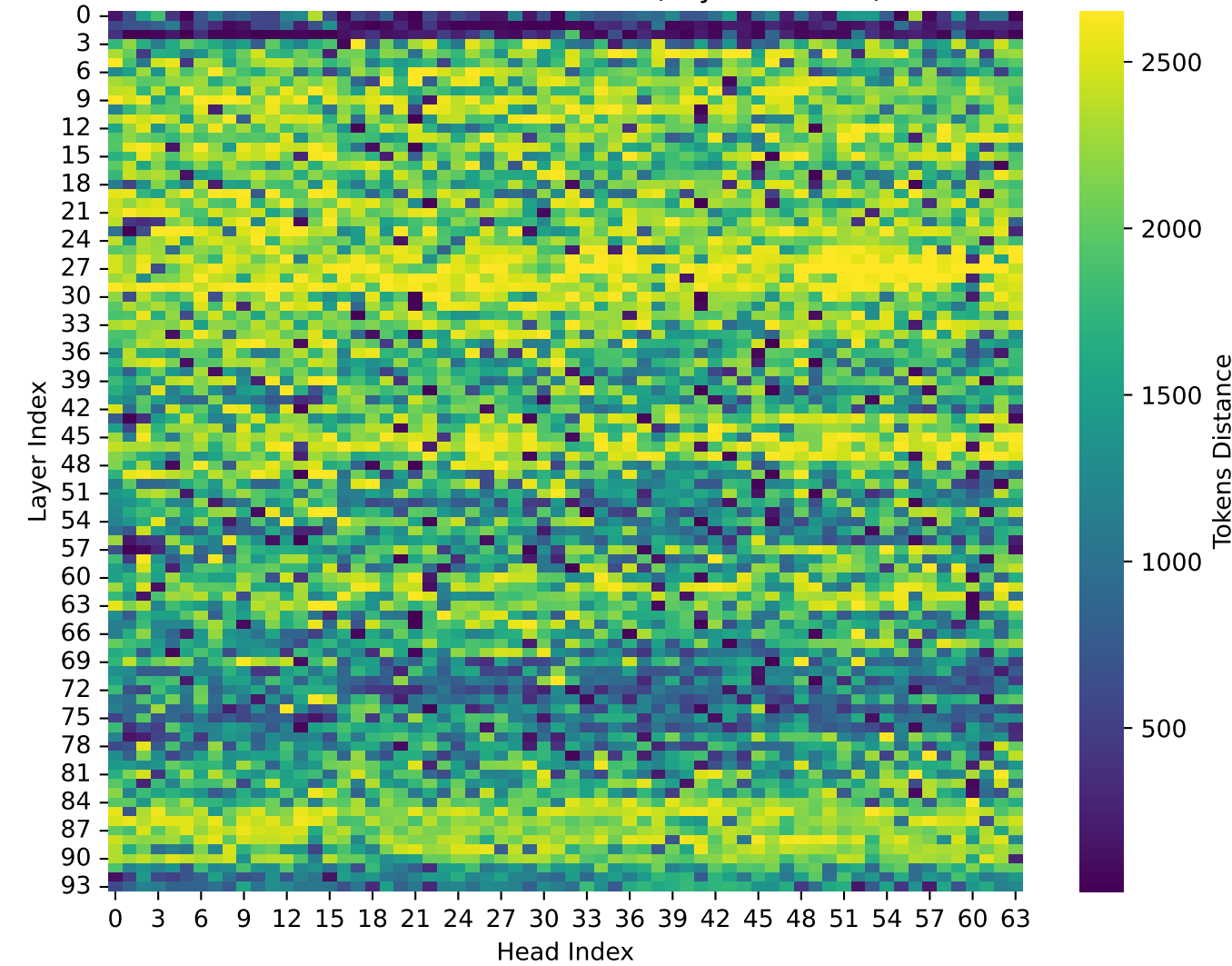
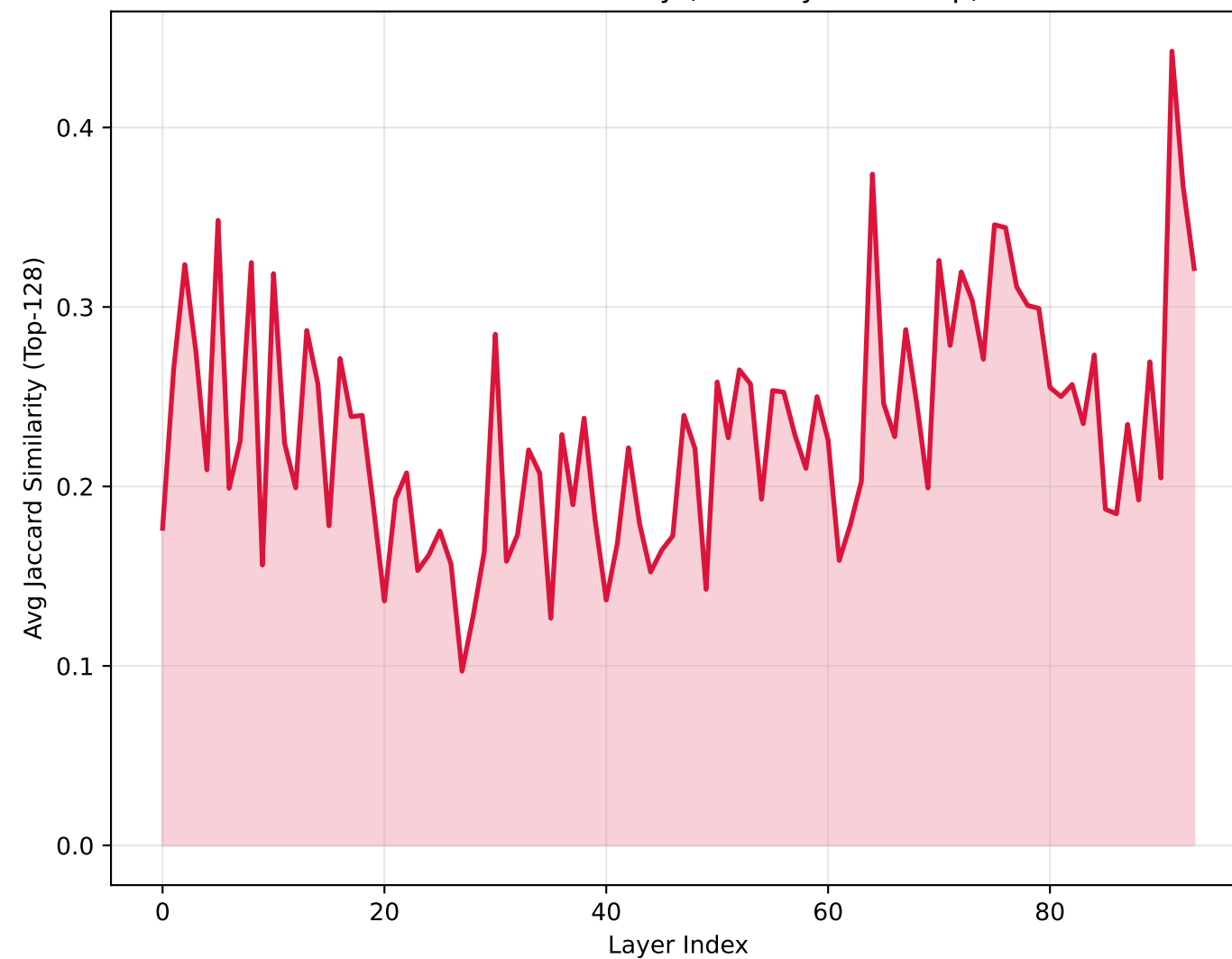


A. Mean Attention Distance (Layer vs Head)



B. Head Redundancy (Intra-layer Overlap)



C. Inter-layer Token Consistency Matrix

