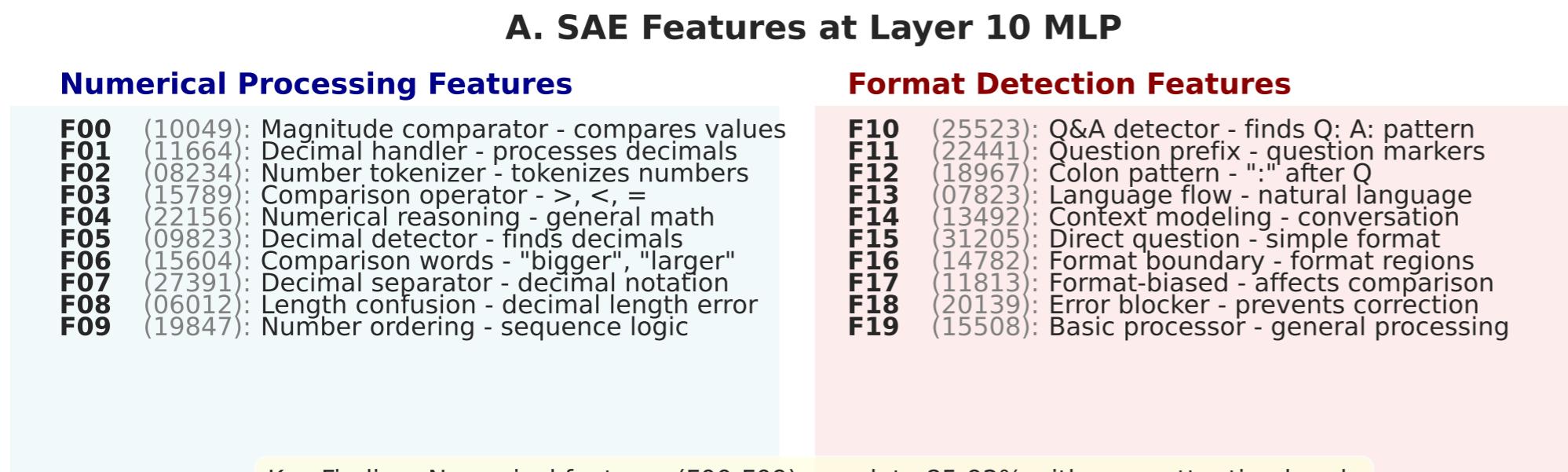
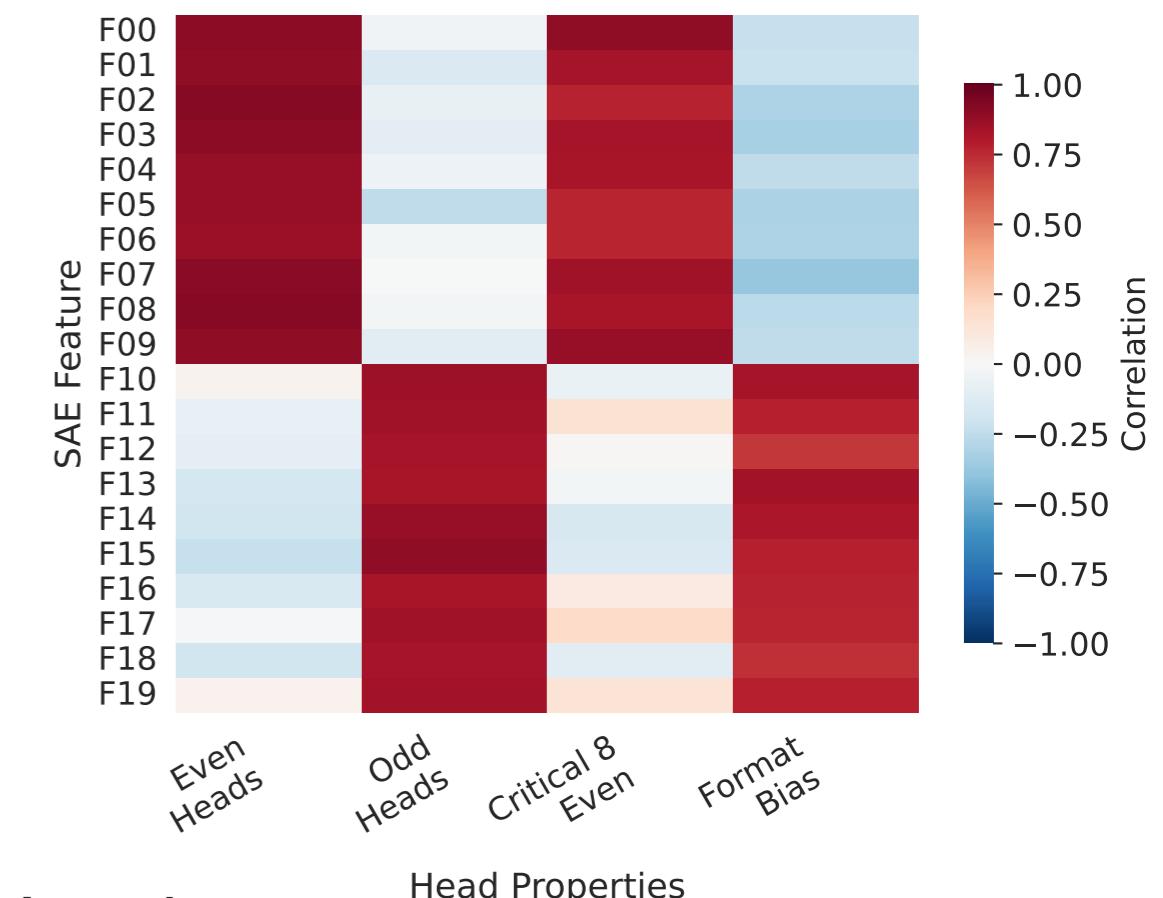


SAE Feature Analysis: Decimal Comparison Bug in Llama-3.1-8B Layer 10



B. Feature-Head Correlations



C. Layer-wise Feature Overlap and Amplification (Actual Data)

