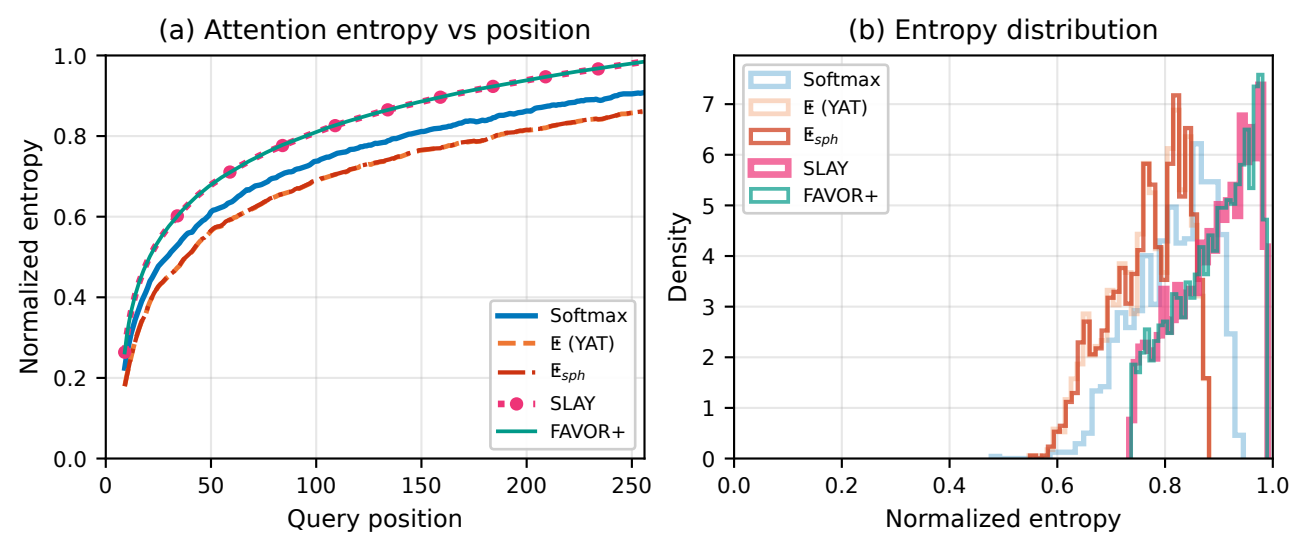
(a) Attention entropy vs position
Normalized entropy
Query position
Softmax
Ɛ (YAT)
Ɛsph
SLAY
FAVOR+
(b) Entropy distribution
Density
Normalized entropy
Softmax
Ɛ (YAT)
Ɛsph
SLAY
FAVOR+