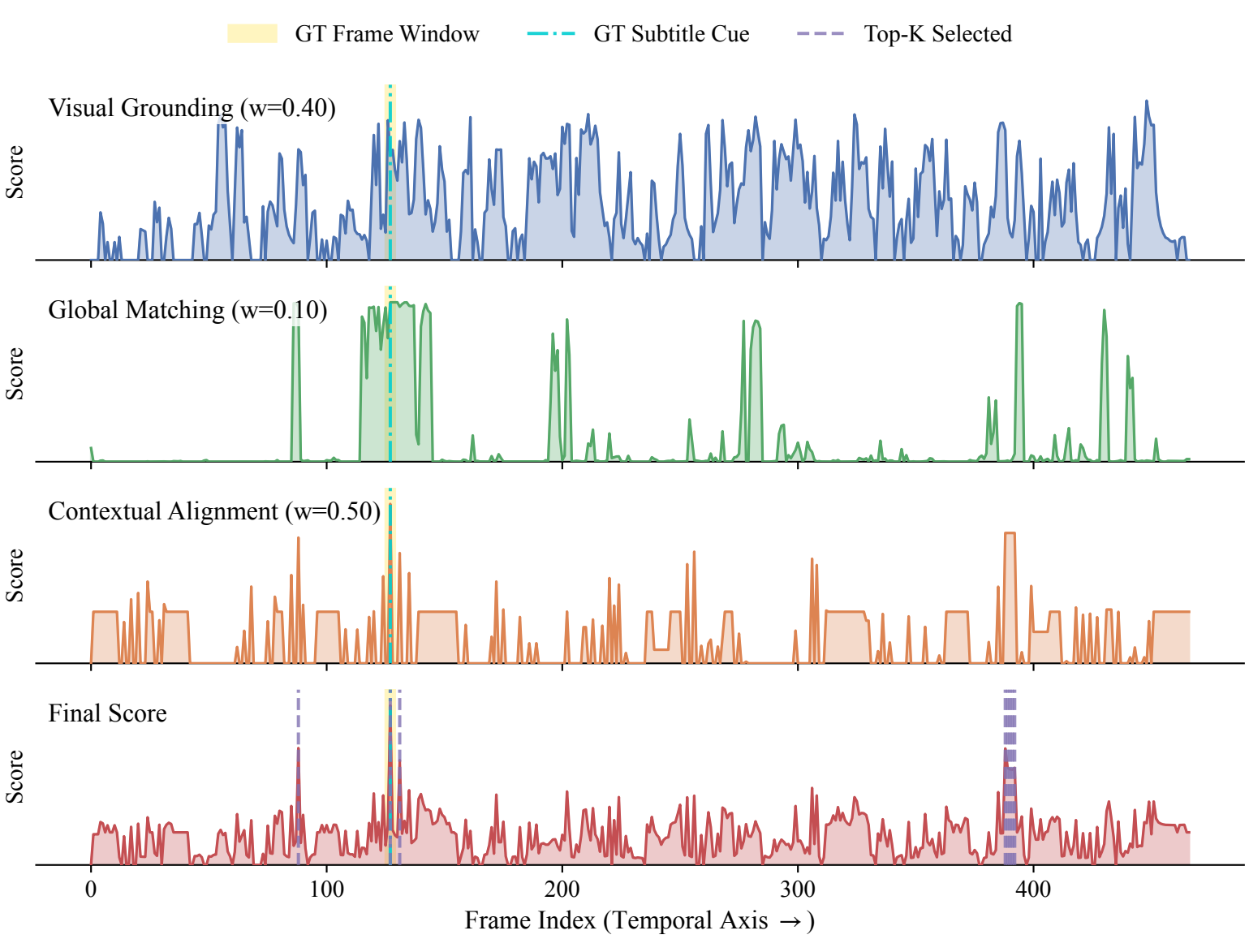

GT Frame Window
GT Subtitle Cue
Top-K Selected
Visual Grounding (w=0.40)
Global Matching (w=0.10)
Contextual Alignment (w=0.50)
Final Score
Score
0
100
200
300
400
Frame Index (Temporal Axis →)