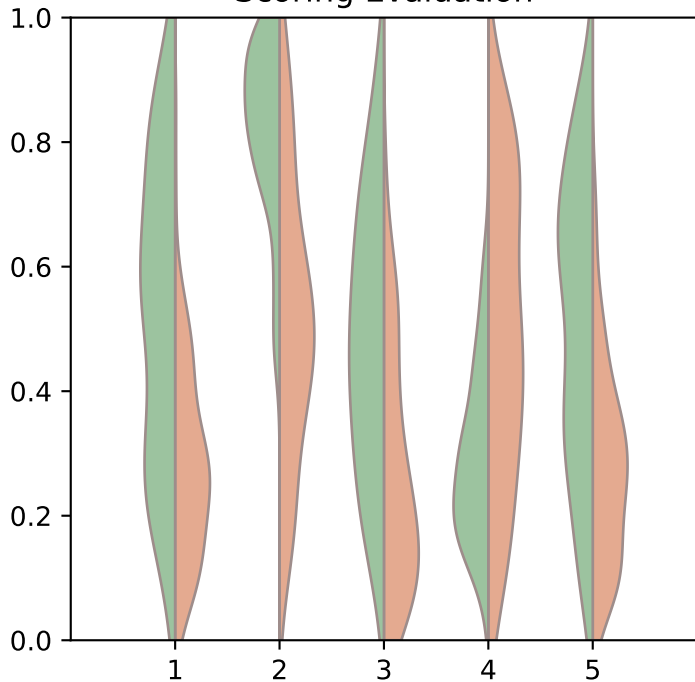
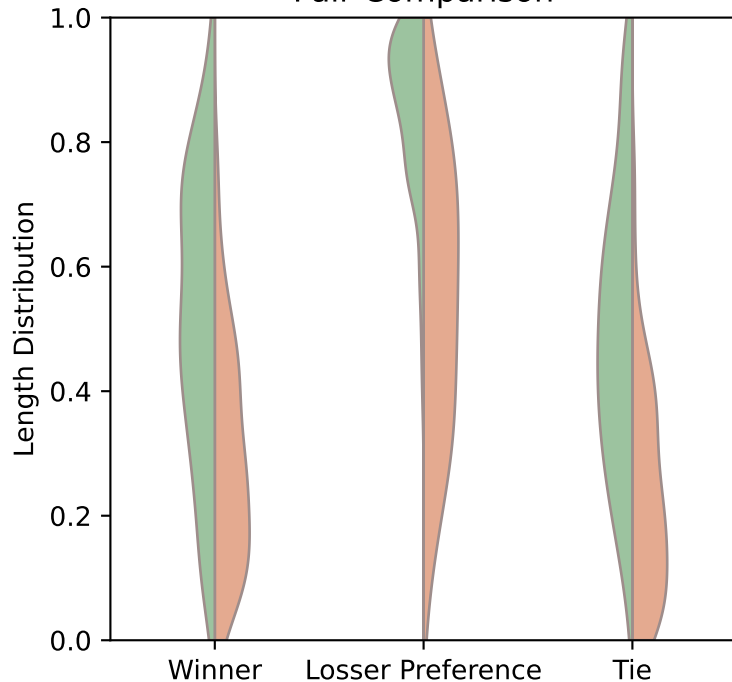


Scoring Evaluation



Pair Comparison



GPT-4V(Vision)



Gemini