RNA-RNA interaction prediction using seed-extension

Frank Gelhausen

Motivation

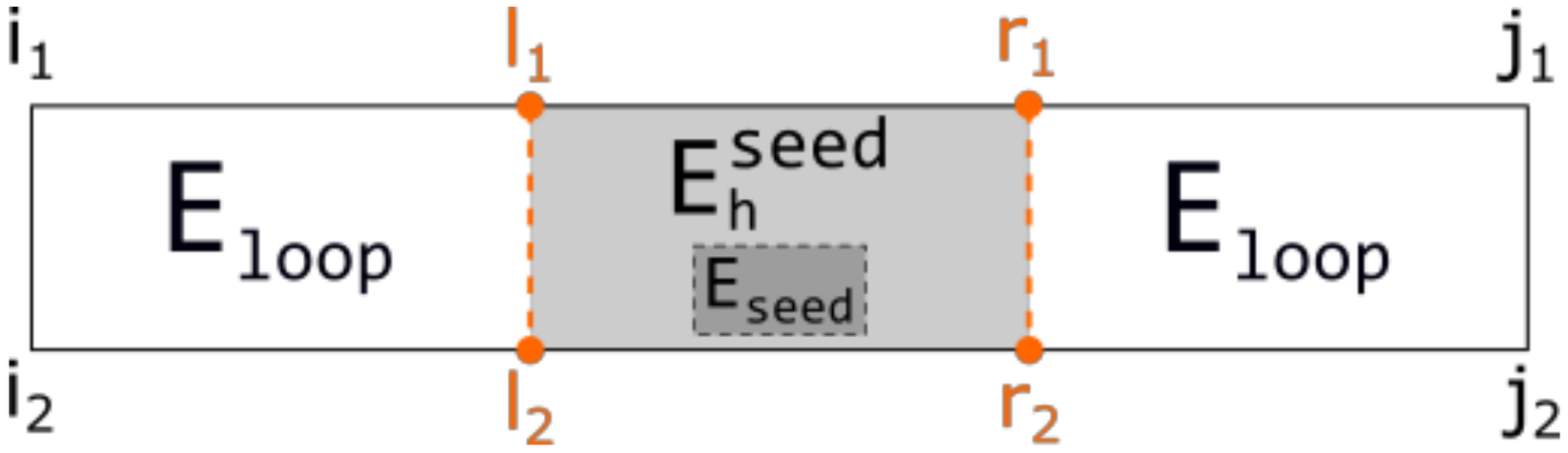
- RNA-RNA interaction prediction computationally expensive

- Idea: compute seed regions and extend interactions starting from seeds to find interaction with minimum energy

- This talk: focus on different seed-extension strategies

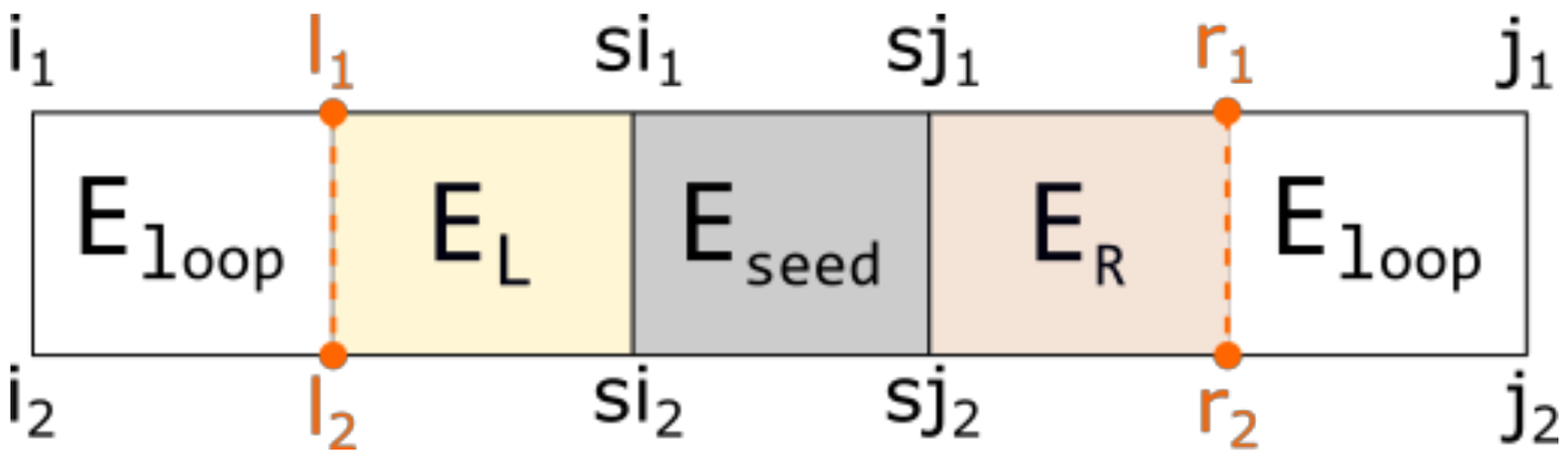
Exact naive method

- $O(N^4)$ space and time
- Find minimum of all combinations of l_1, l_2, r_1, r_2

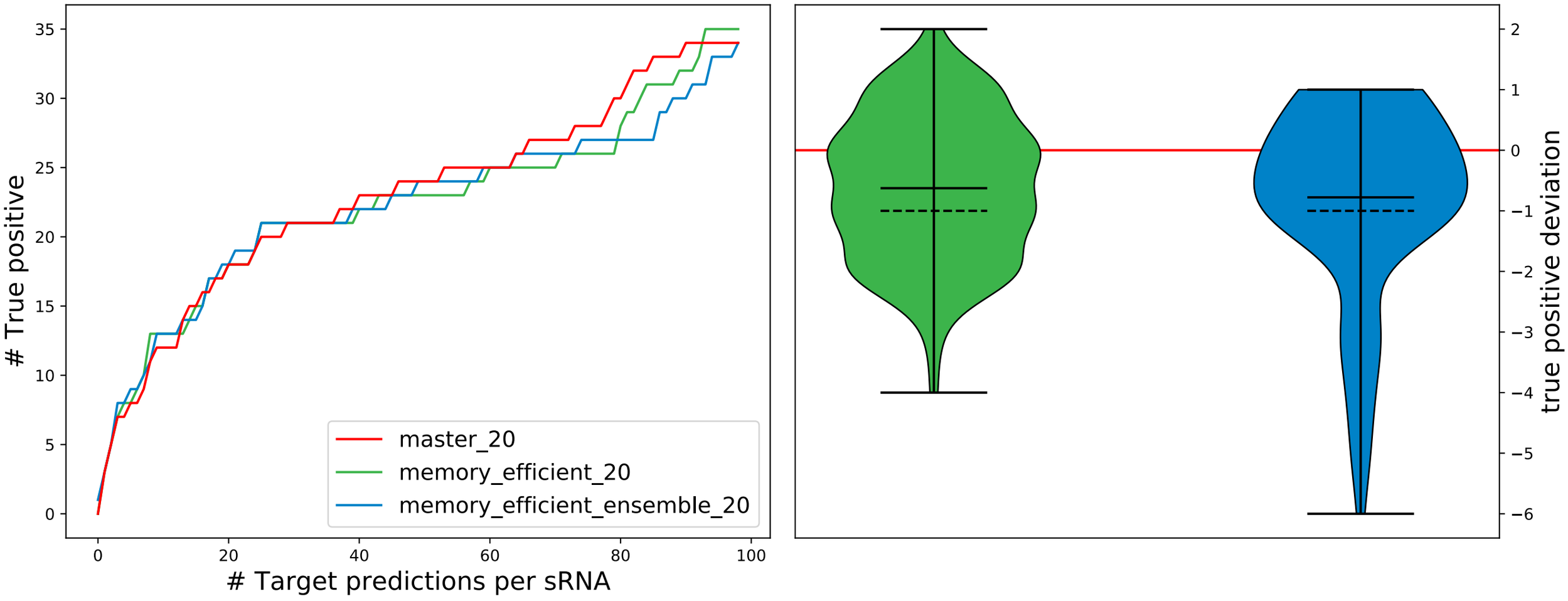


Exact memory efficient method

- $O(N^2)$ space and $O(N^4)$ time
- Extend to the left (E_L), then to the right (E_R)
- Find minimum of all combinations of E_L and E_R

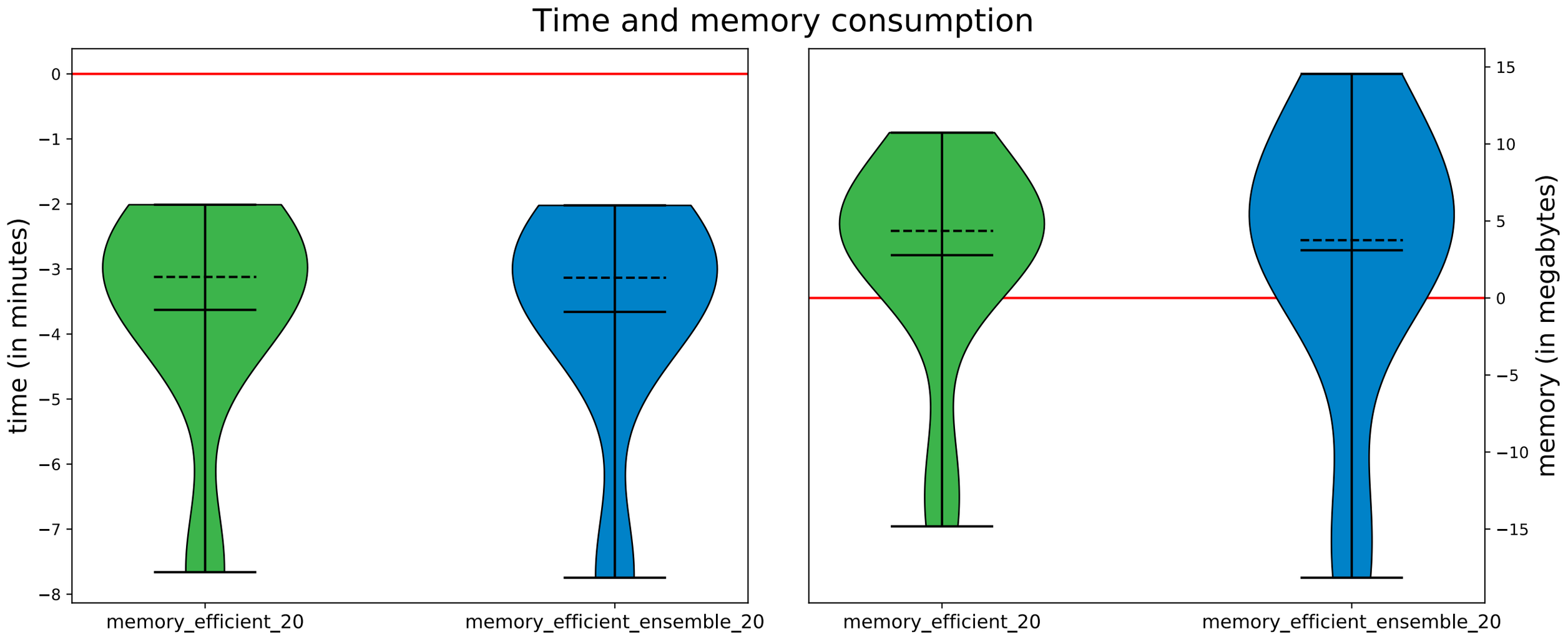


Results



master_20: --qIntLenMax 20 --tIntLenMax 20 --threads 8
memory_efficient_20: -pred X -m M --qIntLenMax 20 --tIntLenMax 20 --threads 8
memory_efficient_ensemble_20: -pred E -m M —qIntLenMax 20 --tIntLenMax 20 --threads 8

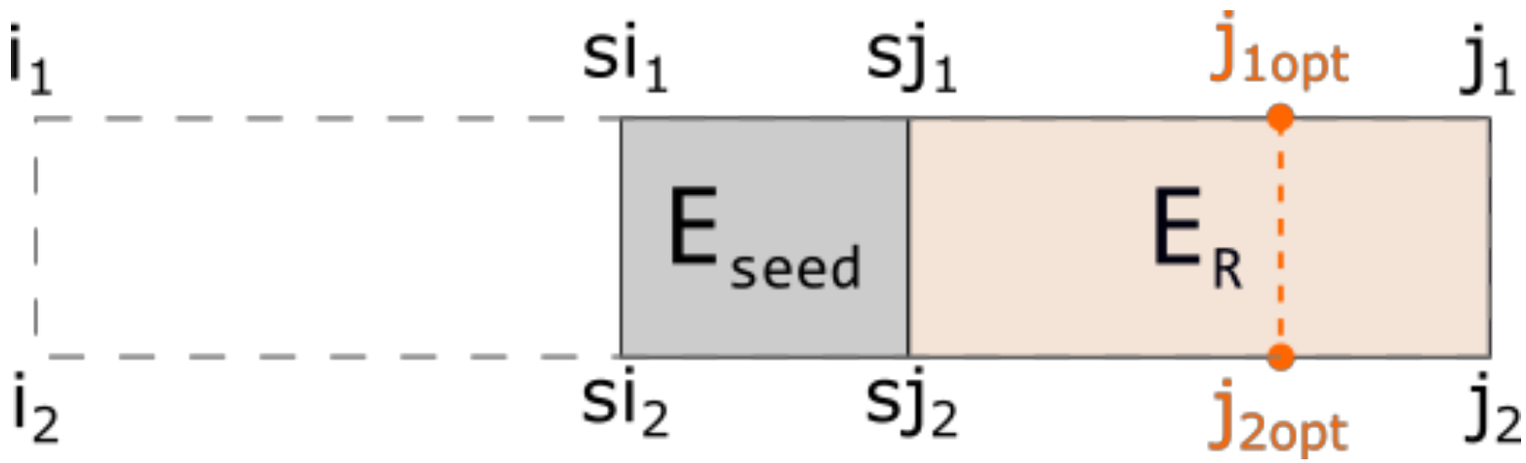
Results (time + mem)



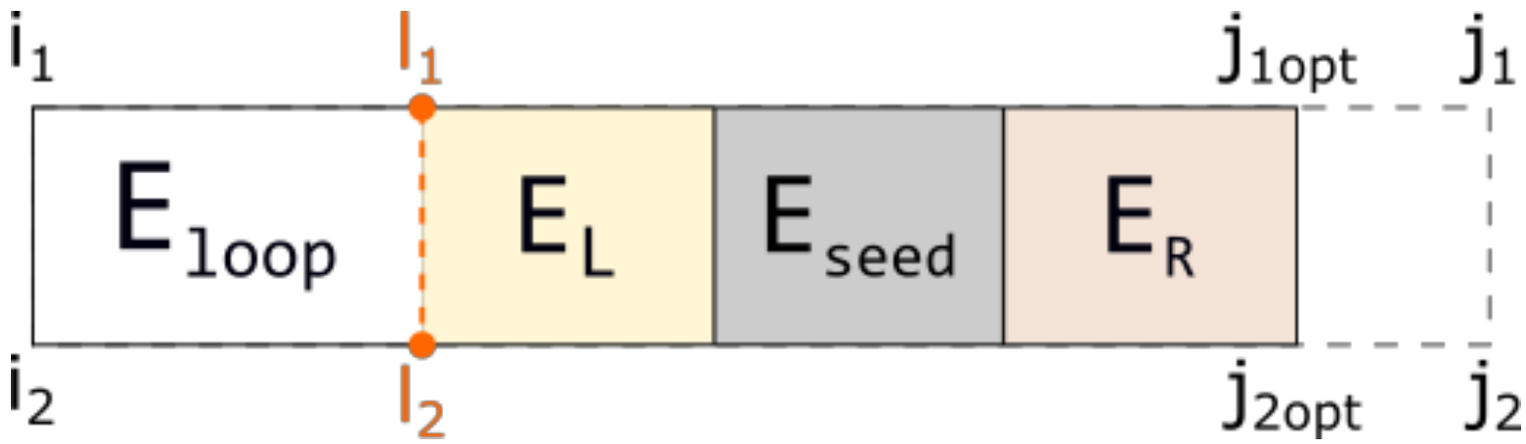
master_20: --qIntLenMax 20 --tIntLenMax 20 --threads 8
memory_efficient_20: -pred X -m M --qIntLenMax 20 --tIntLenMax 20 --threads 8
memory_efficient_ensemble_20: -pred E -m M —qIntLenMax 20 --tIntLenMax 20 --threads 8

Heuristic method

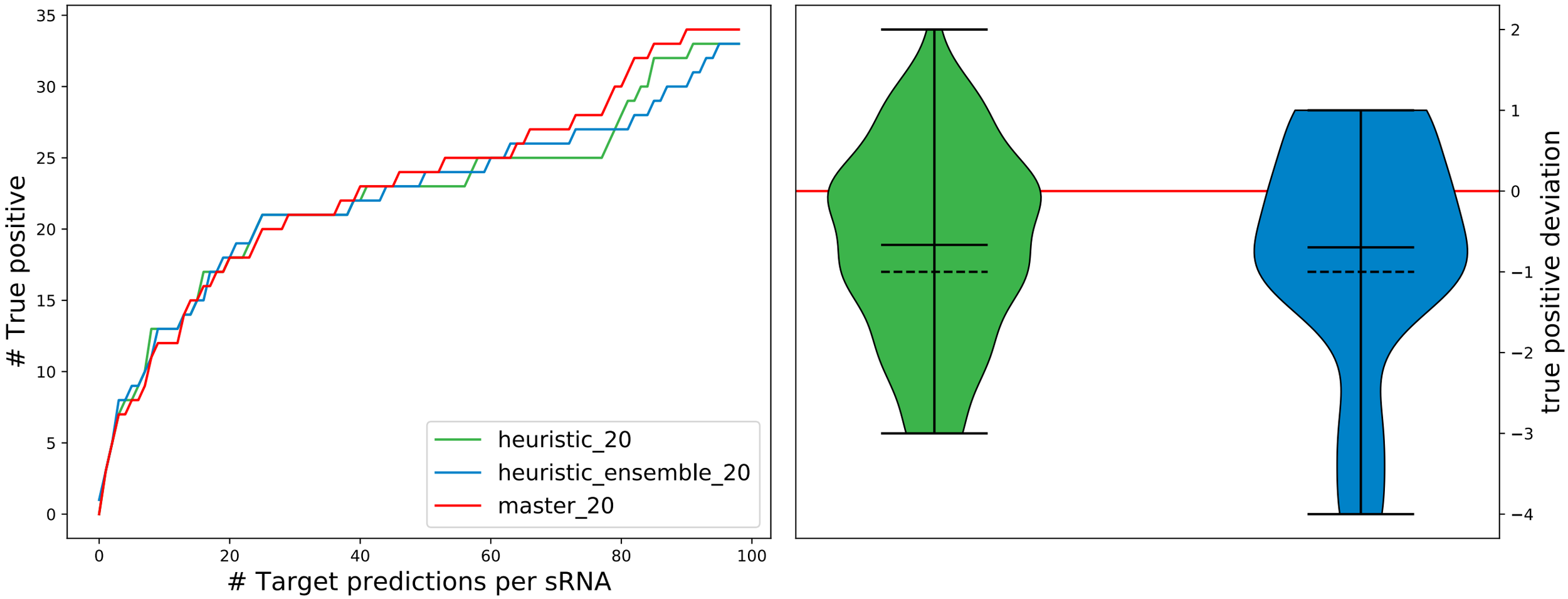
- $O(N^2)$ space and $O(N^2)$ time
- First find j_1 and j_2 that minimalize E_R



- Then minimize over entire interaction up to j_{1opt}, j_{2opt}

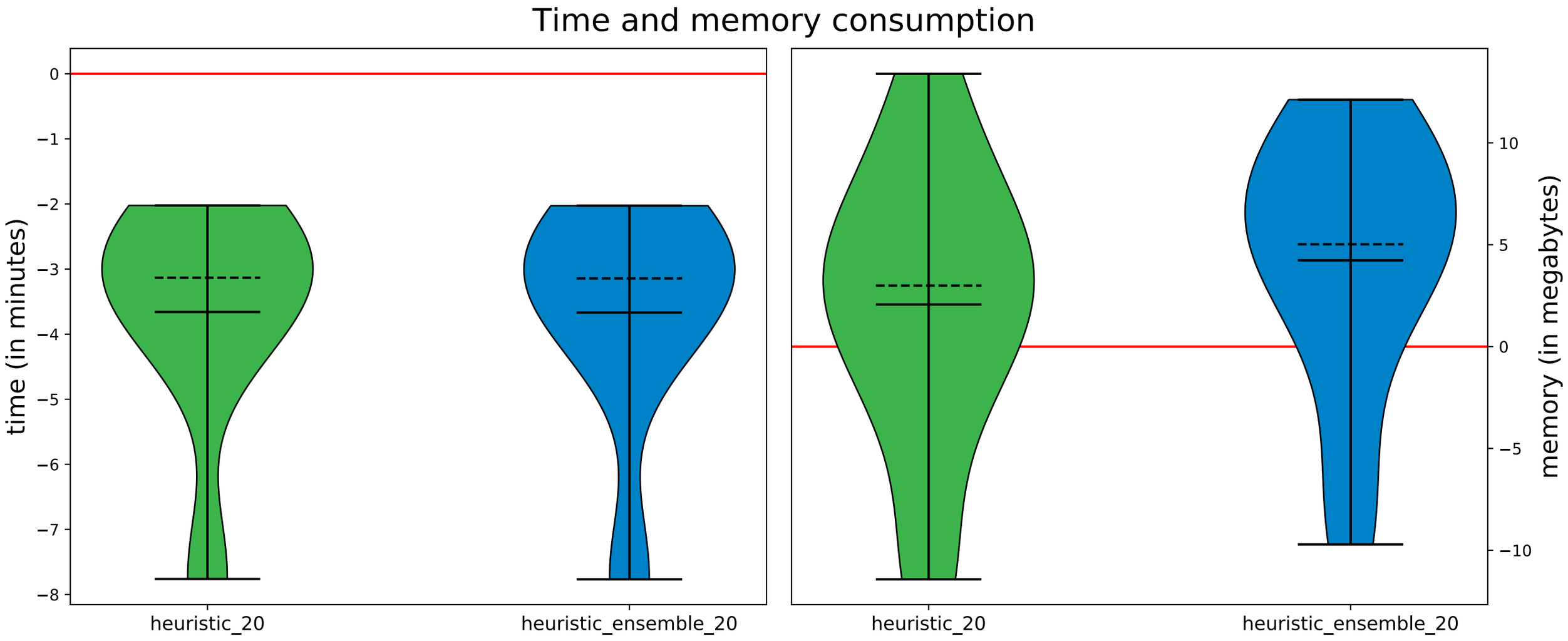


Results



master_20: --qIntLenMax 20 --tIntLenMax 20 --threads 8
heuristic_20: -pred X -m H --qIntLenMax 20 --tIntLenMax 20 --threads 8
heuristic_ensemble_20: -pred E -m H --qIntLenMax 20 --tIntLenMax 20 --threads 8

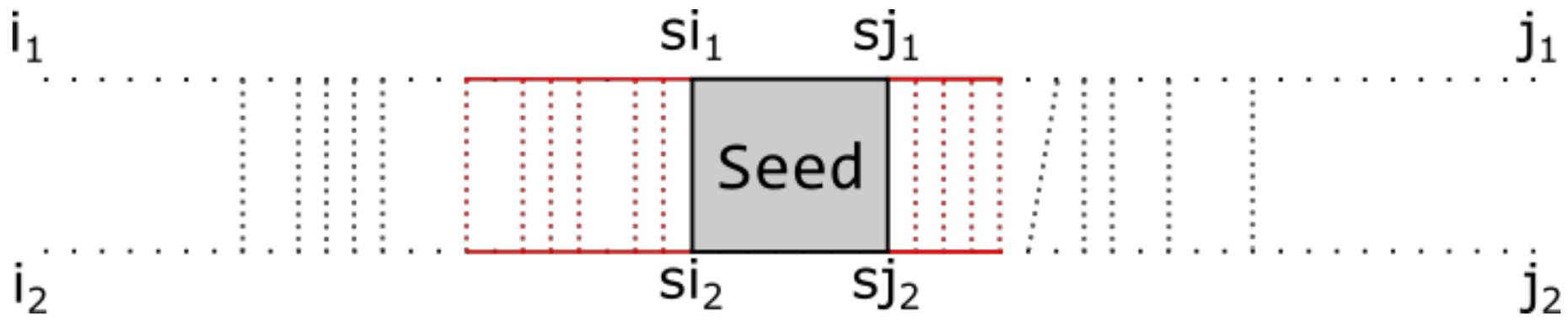
Results (time + mem)



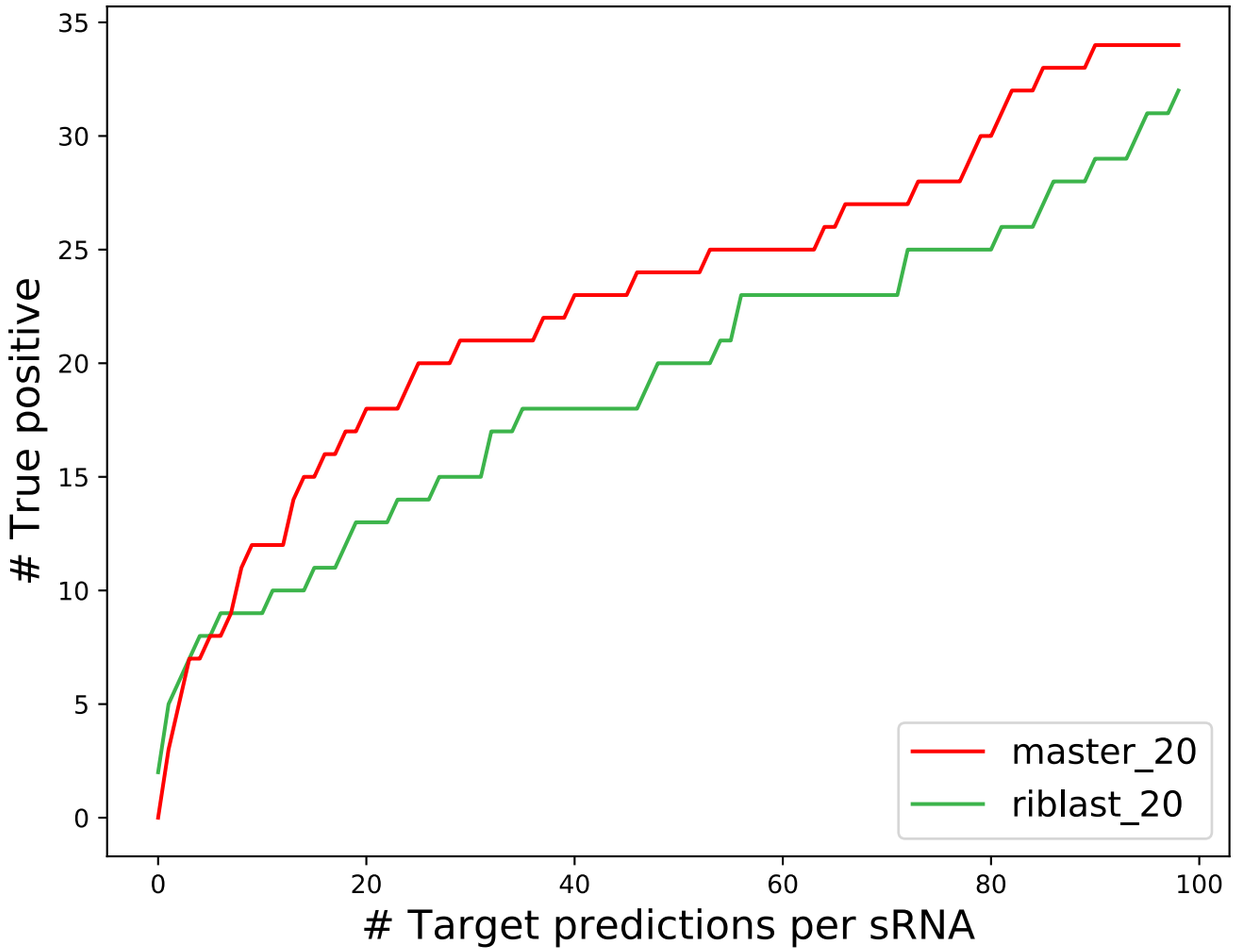
master_20: --qIntLenMax 20 --tIntLenMax 20 --threads 8
heuristic_20: -pred X -m H --qIntLenMax 20 --tIntLenMax 20 --threads 8
heuristic_ensemble_20: -pred E -m H —qIntLenMax 20 --tIntLenMax 20 --threads 8

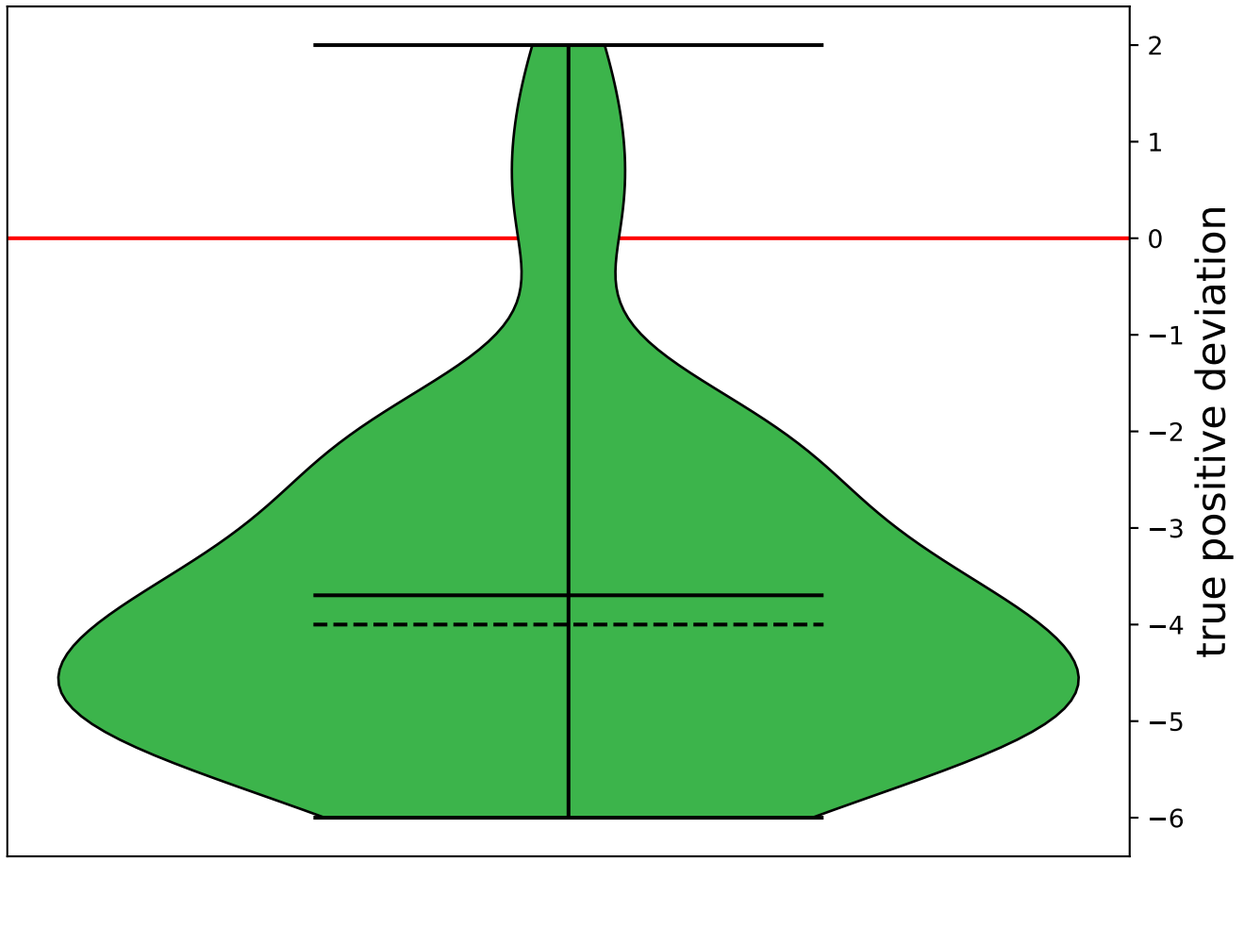
Riblast method

- by Fukunaga et al.
- First do a parallel extension, then thorough extension
- Parallel Extension:



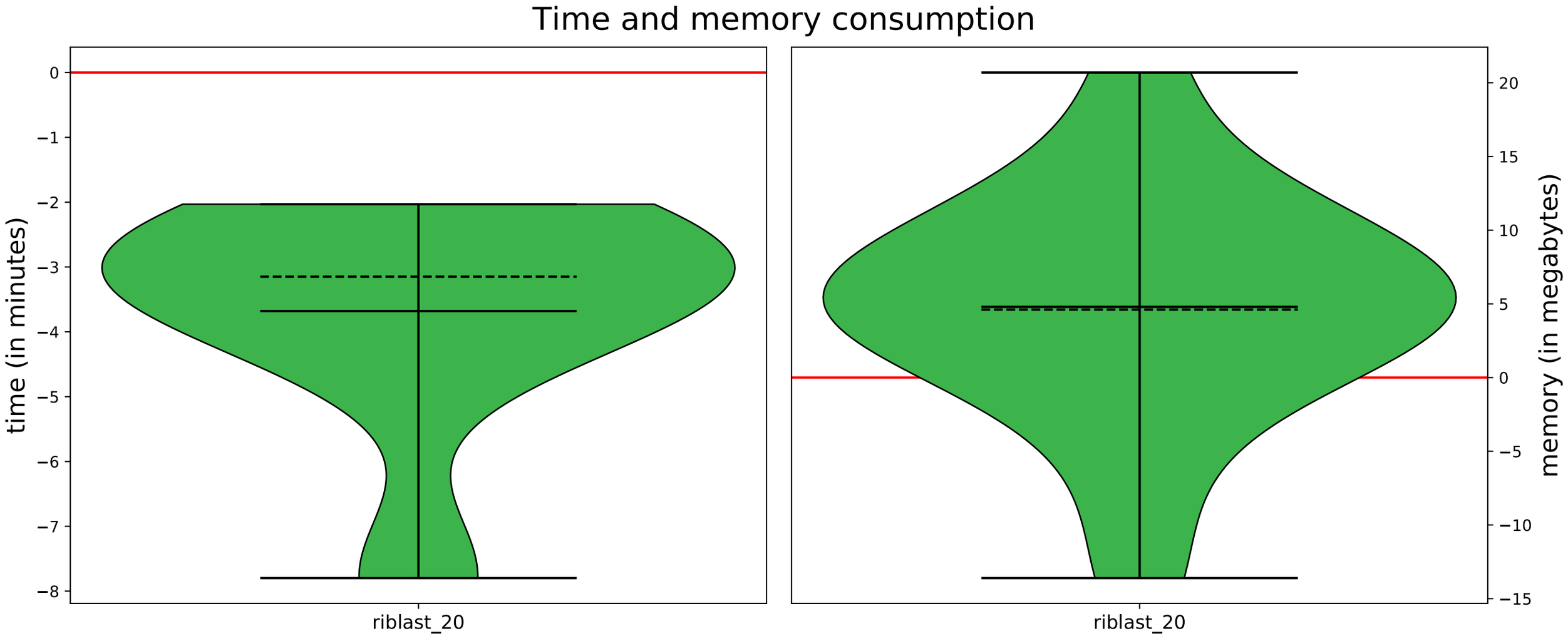
Results





master_20: --qIntLenMax 20 --tIntLenMax 20 --threads 8
riblast_20: -pred X -m R --qIntLenMax 20 --tIntLenMax 20 --threads 8

Results (time + mem)



master_20: --qIntLenMax 20 --tIntLenMax 20 --threads 8
memory_efficient_20: -pred X -m M --qIntLenMax 20 --tIntLenMax 20 --threads 8
memory_efficient_ensemble_20: -pred E -m M--qIntLenMax 20 --tIntLenMax 20 --threads 8

Comparison

Conclusion