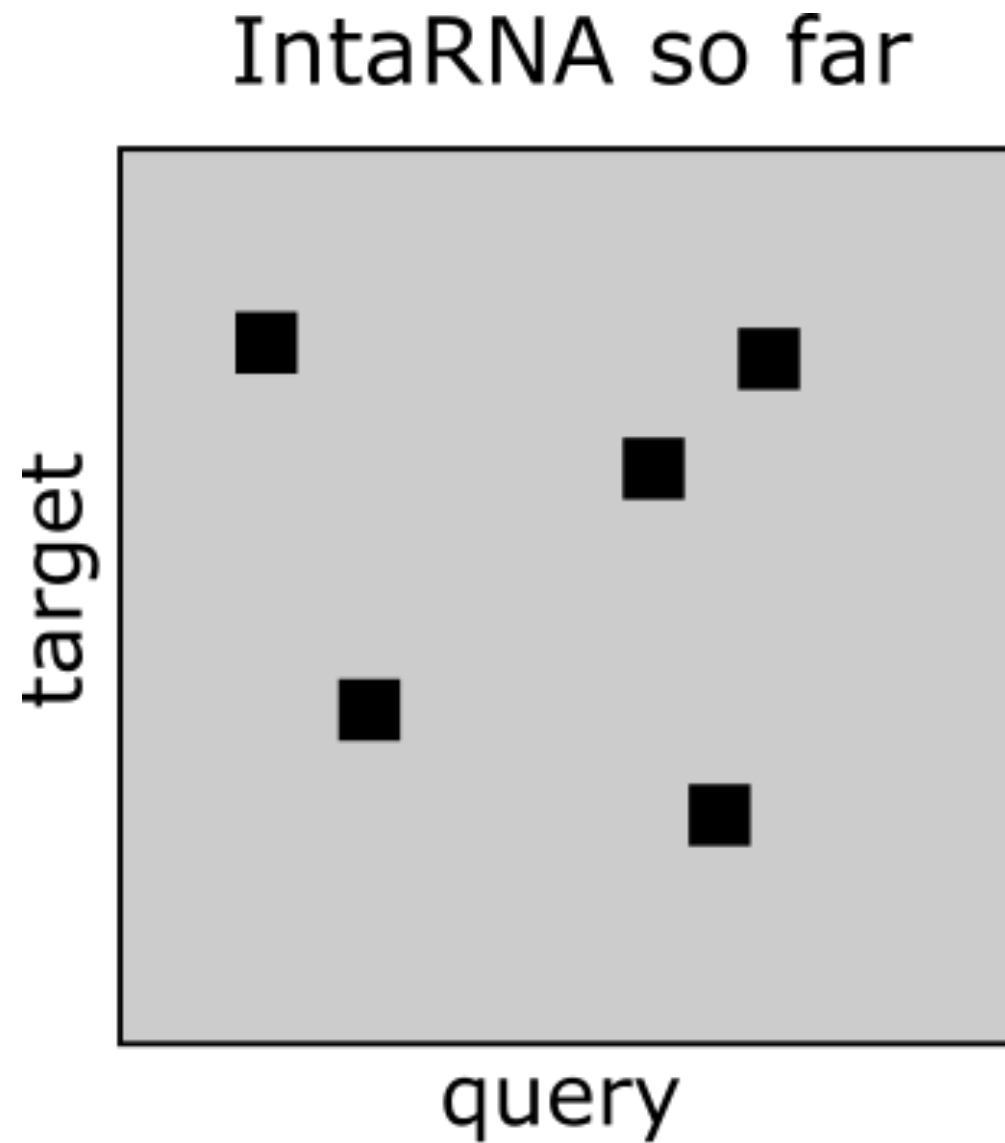


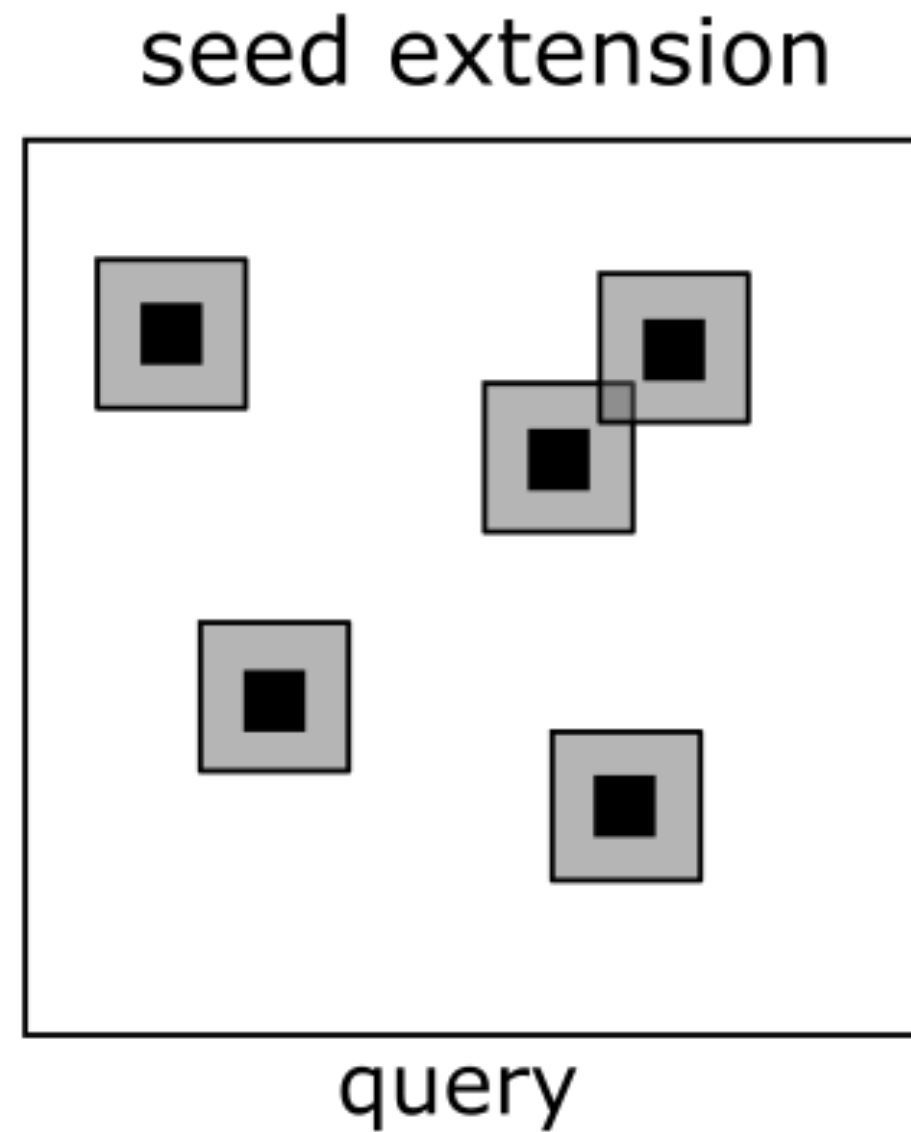
RNA-RNA interaction prediction using seed extension

Frank Gelhausen

Motivation



■ Seed



eg: sTarPicker
RiBlast
RiSearch2

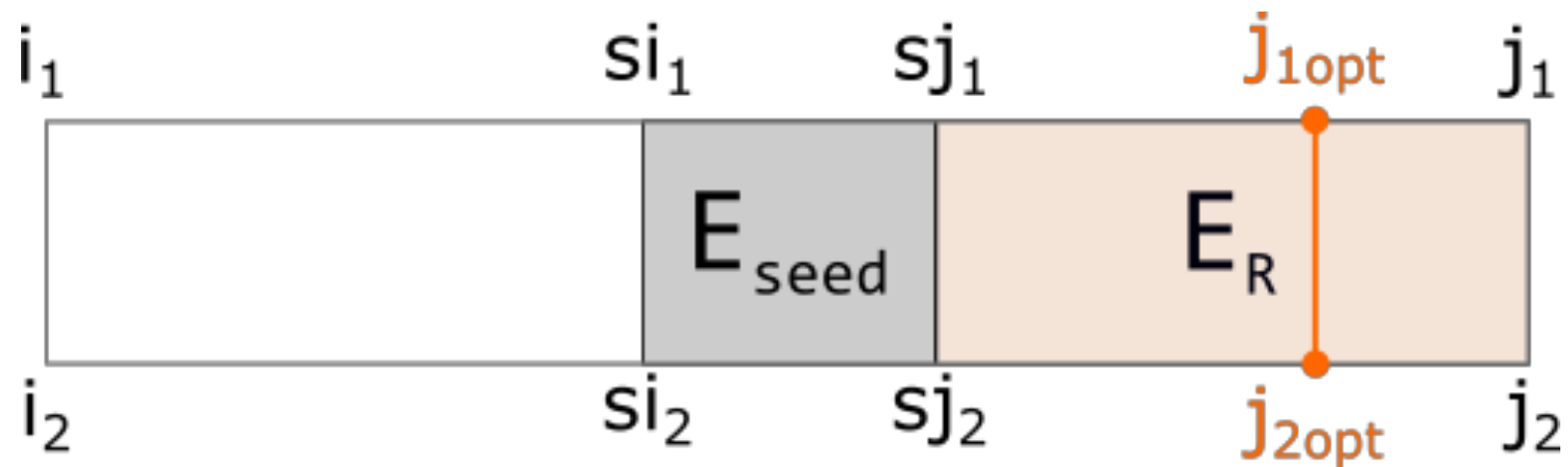
Exact memory efficient method

Search minimum over all seeds using:

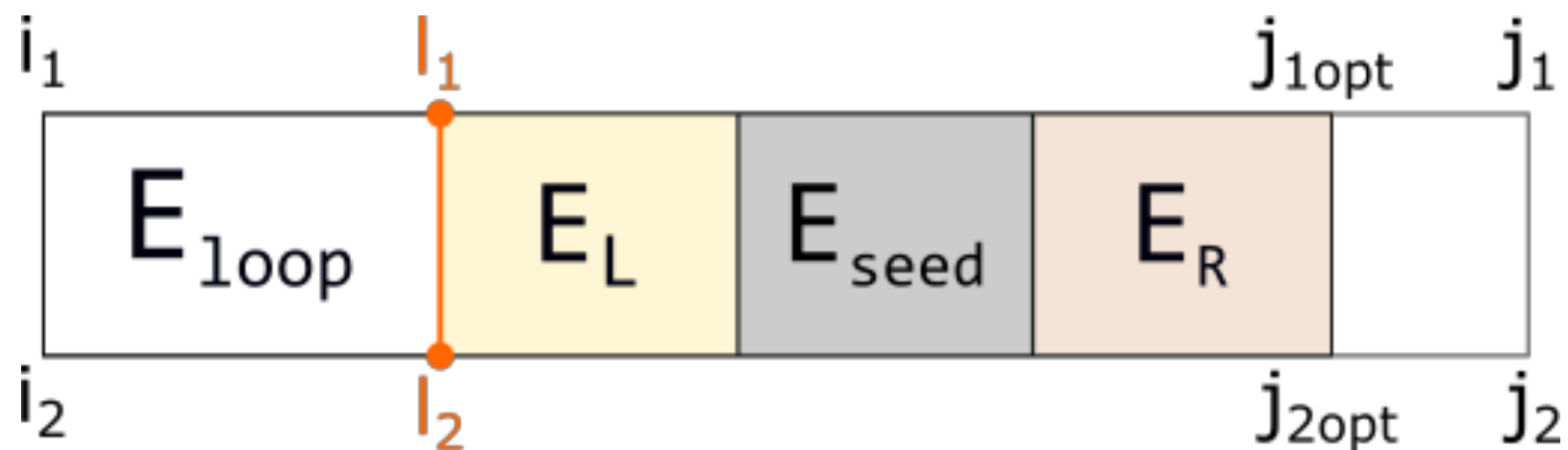
$$\min \left[\begin{array}{|c|c|} \hline i_1 & l_1 \\ \hline E_{\text{loop}} & E_L \\ \hline i_2 & l_2 \\ \hline \end{array} \right] + \begin{array}{|c|} \hline E_{\text{seed}} \\ \hline \end{array} + \min \left[\begin{array}{|c|c|} \hline r_1 & j_1 \\ \hline E_R & E_{\text{loop}} \\ \hline r_2 & j_2 \\ \hline \end{array} \right]$$

Heuristic method

- First find j_1 and j_2 that minimize E_R

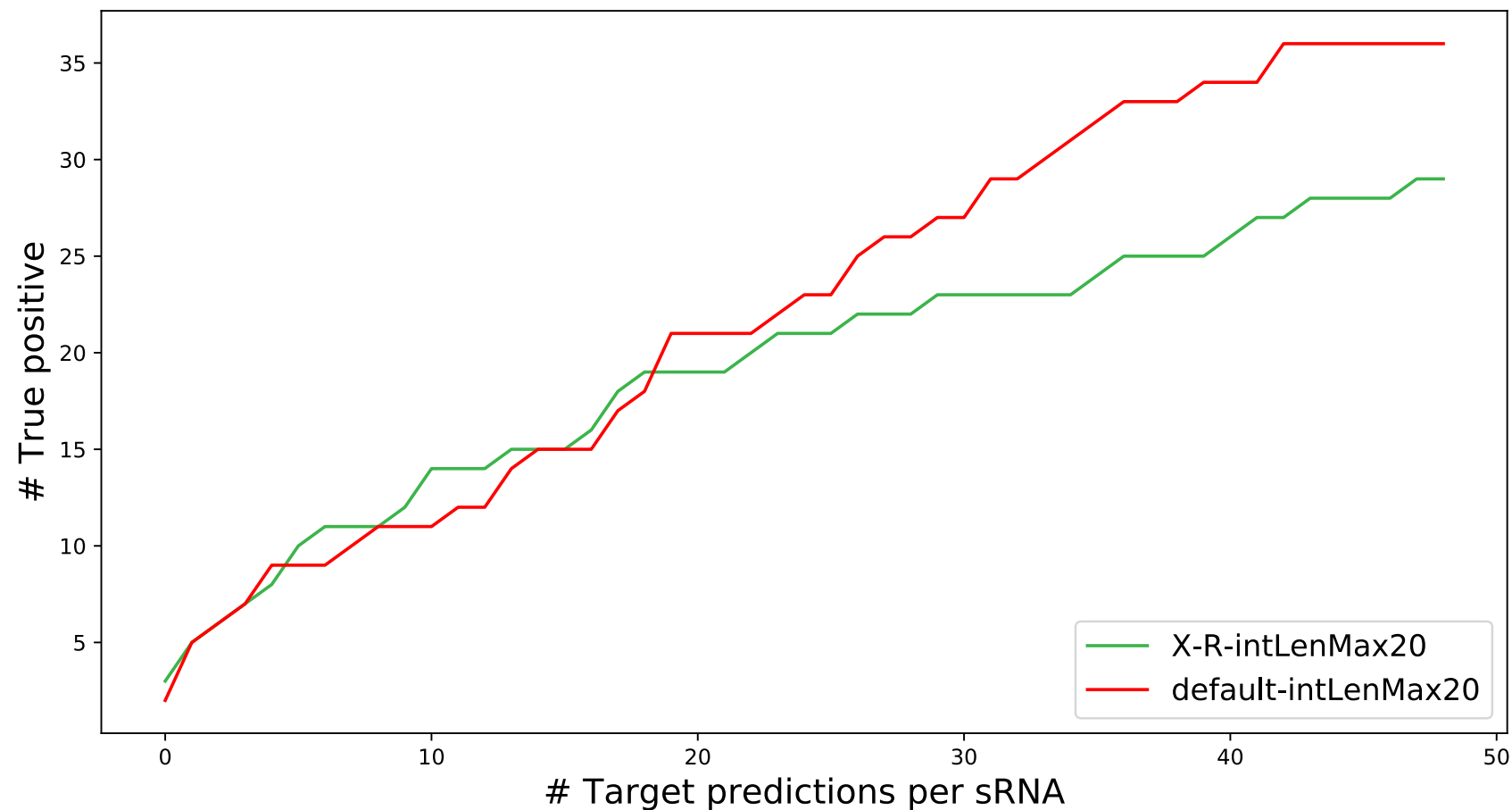
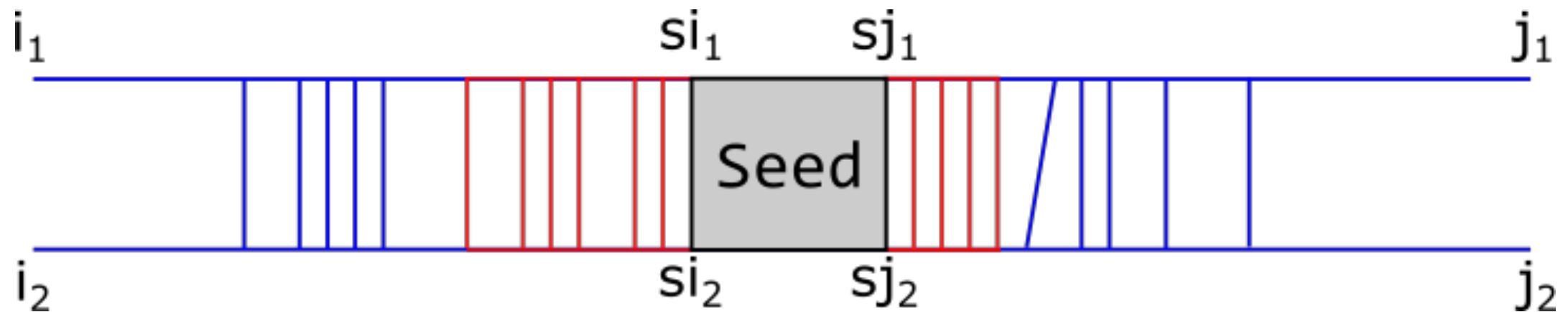


- Then minimize over entire interaction up to j_{1opt} , j_{2opt}

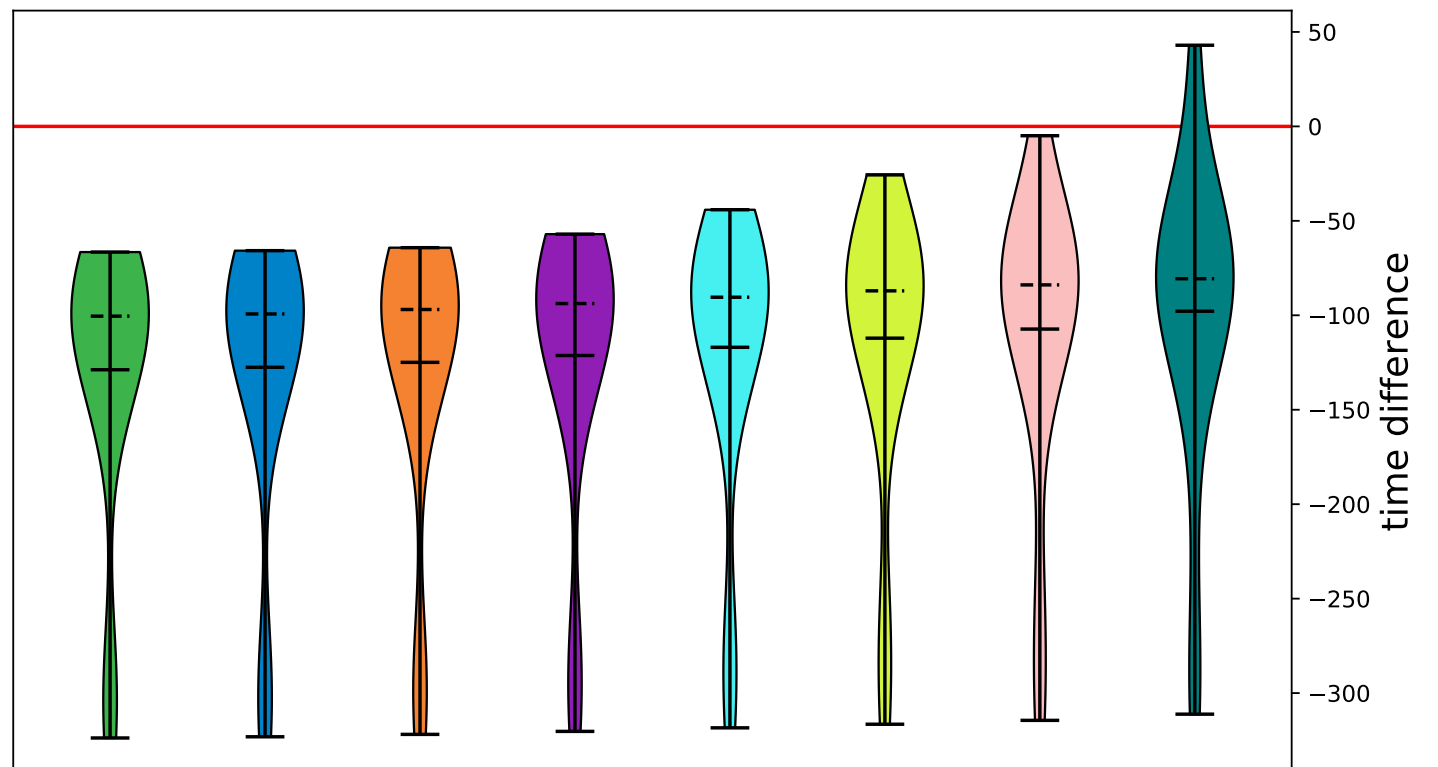
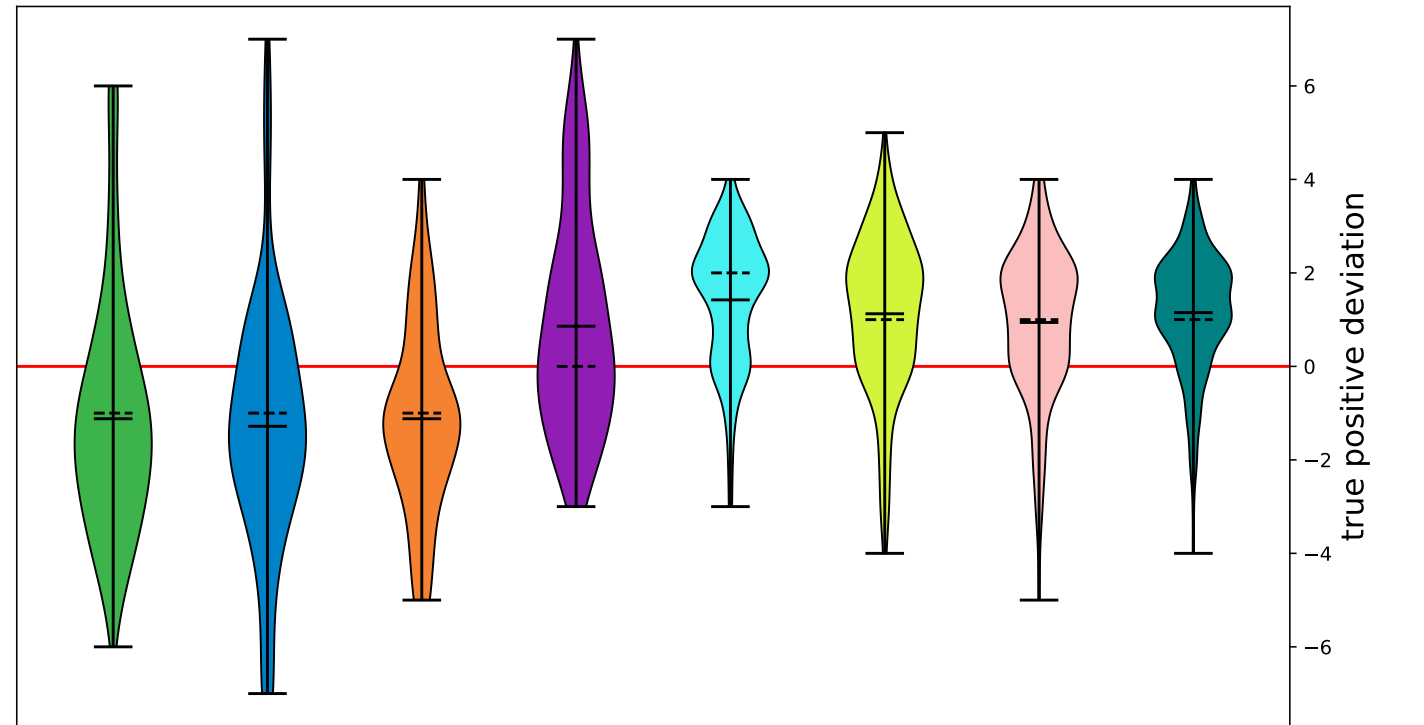
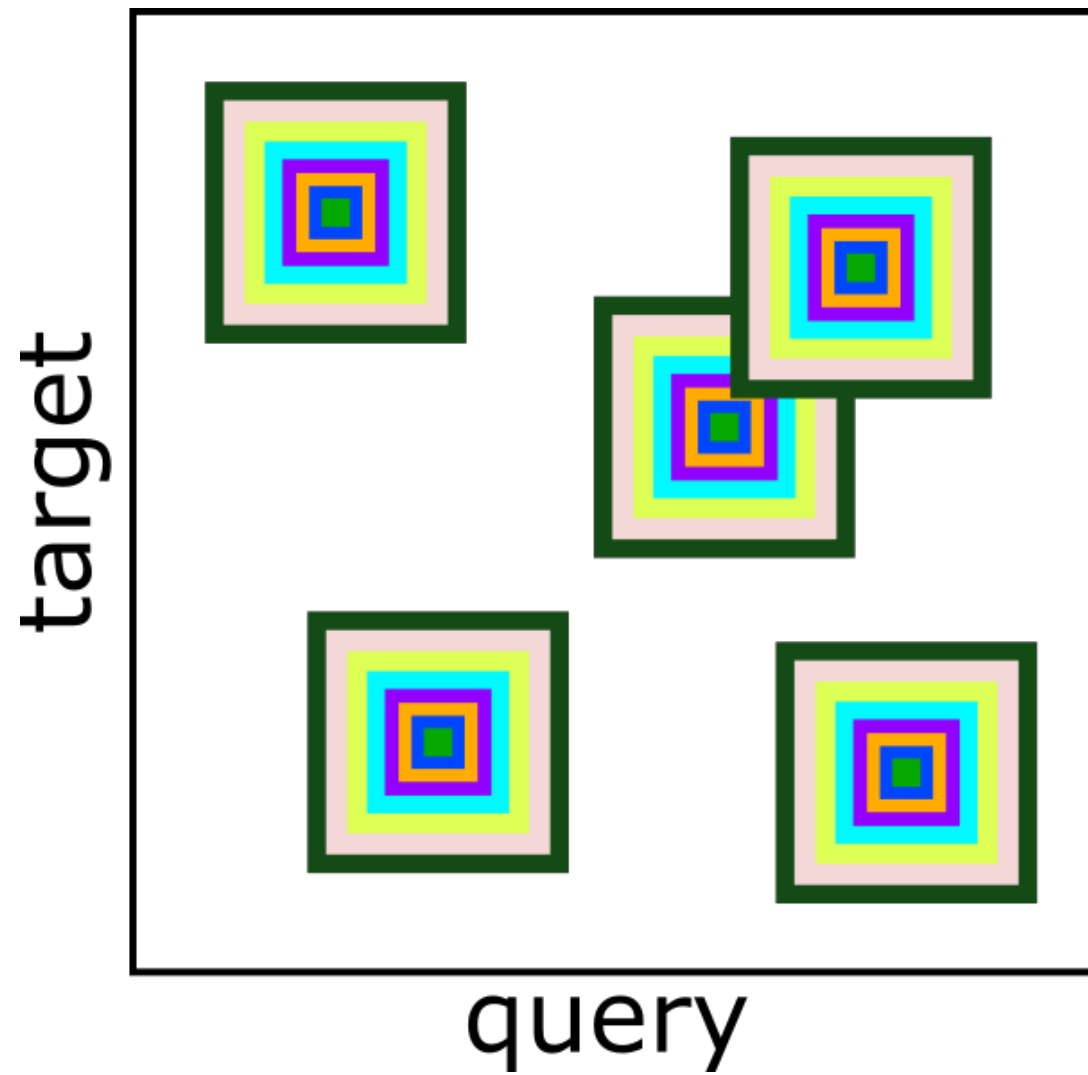


Riblast method

First do a **parallel extension**, then **thorough extension**



Length-dependence



Partition function

Goal: Find probabilities of interactions

$$P_{i,j} = \frac{Z_{i,j}}{Z}$$

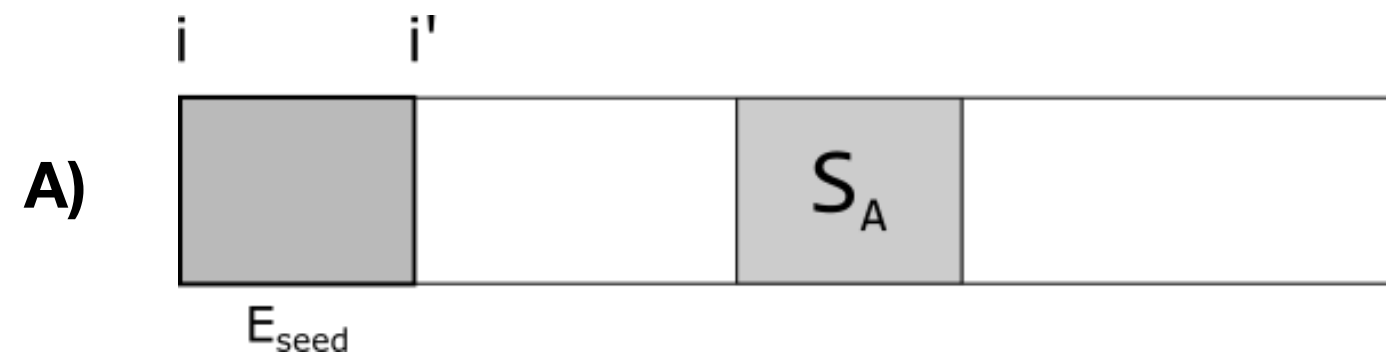
$$Z_{i,j} = \sum_{S_A} \left(\begin{array}{|c|c|c|c|} \hline & & & \\ \hline \text{no seed} & S_A & \boxed{S'} & \boxed{S''} \\ \hline & & & \\ \hline \end{array} \right)^i_j$$

$$Z = \sum Z_{i,j}$$

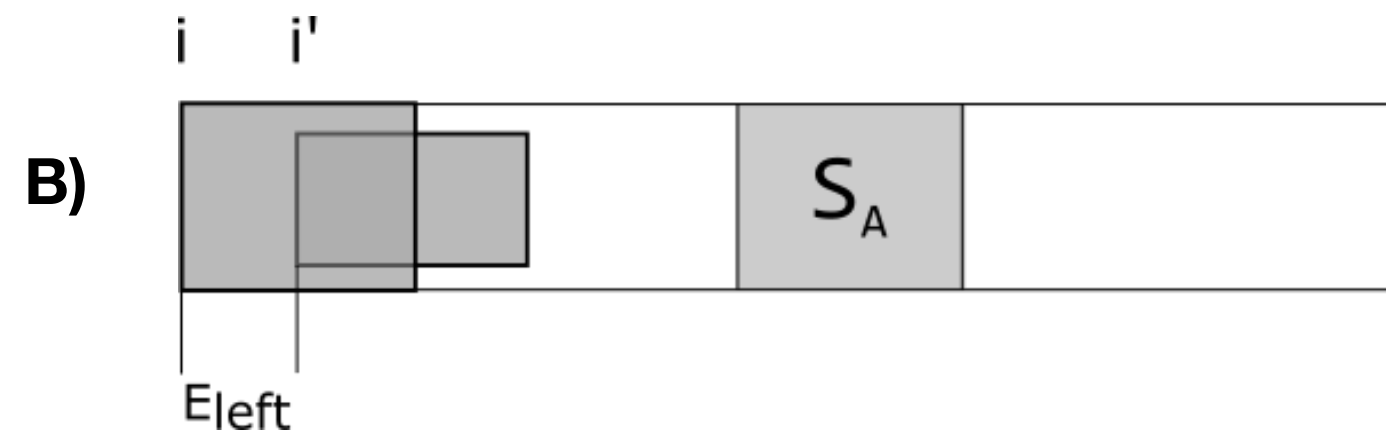
$$Z_{i,j} = \sum_{S_A} \left(\begin{array}{|c|c|c|c|} \hline & \text{no seed} & S_A & \boxed{S'} \quad \boxed{S''} \\ \hline \end{array} \right)^i_j$$

Idea: remove seed for i

Cases:



$$Z - w(E_{seed}) \cdot Z_{i',j'}$$



$$Z - w(E_{left}) \cdot Z_{i',j'}$$