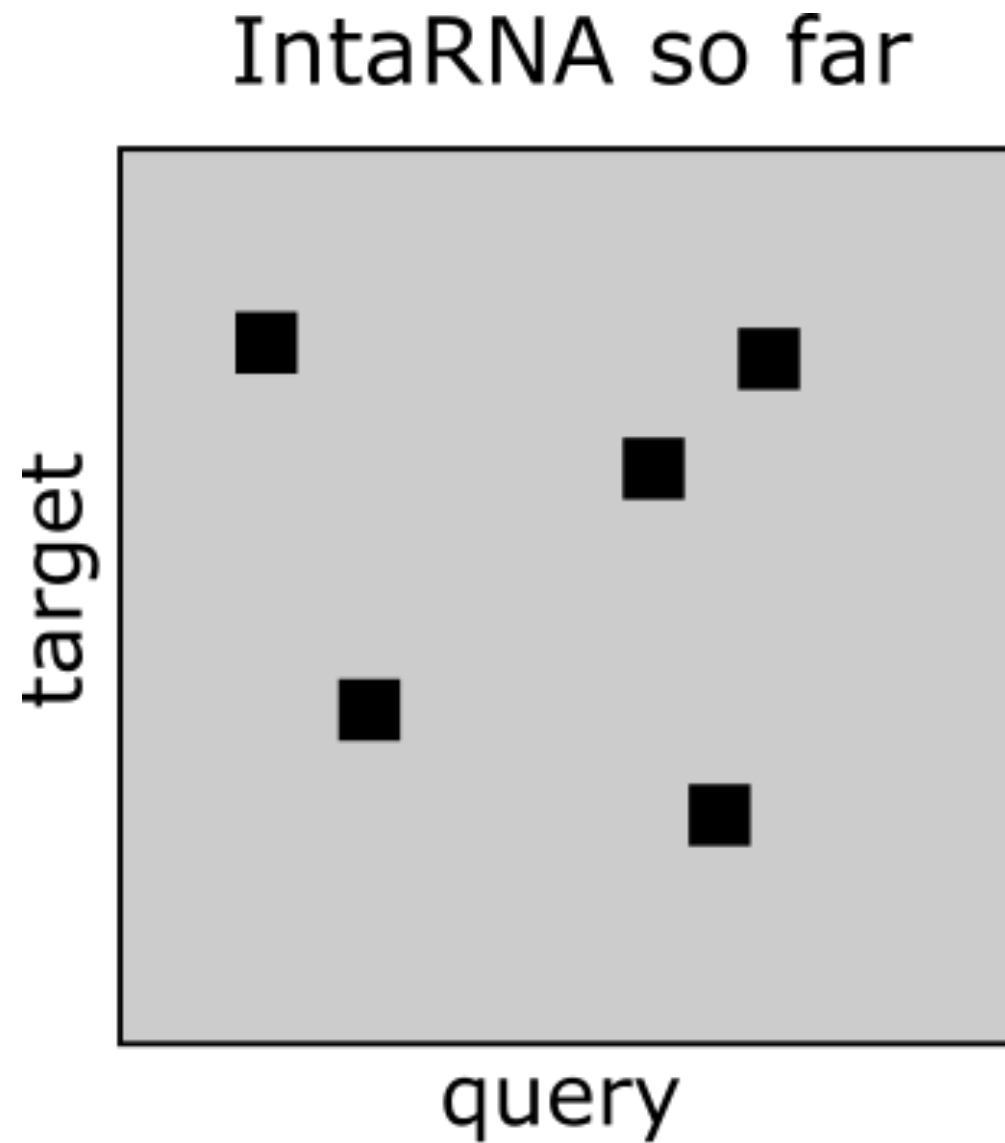


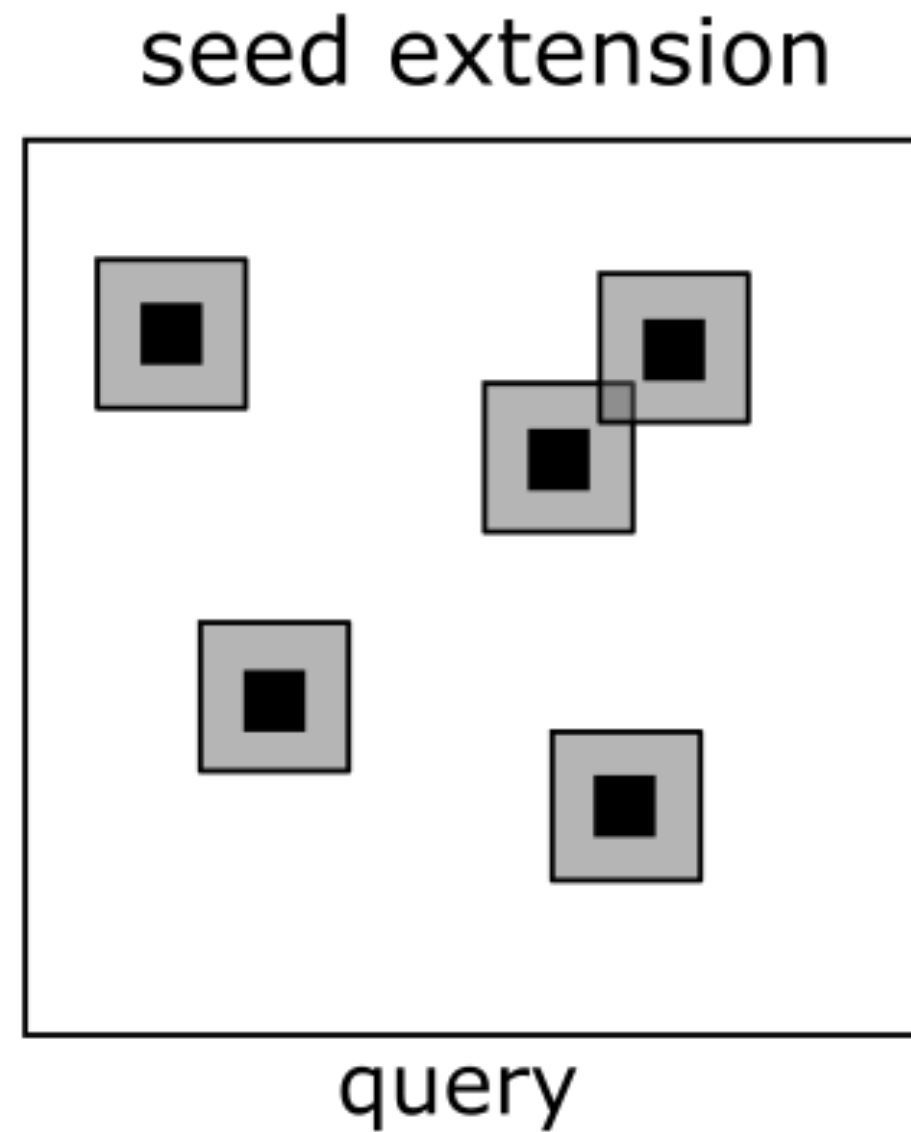
# **RNA-RNA interaction prediction using seed extension**

Frank Gelhausen

# Motivation

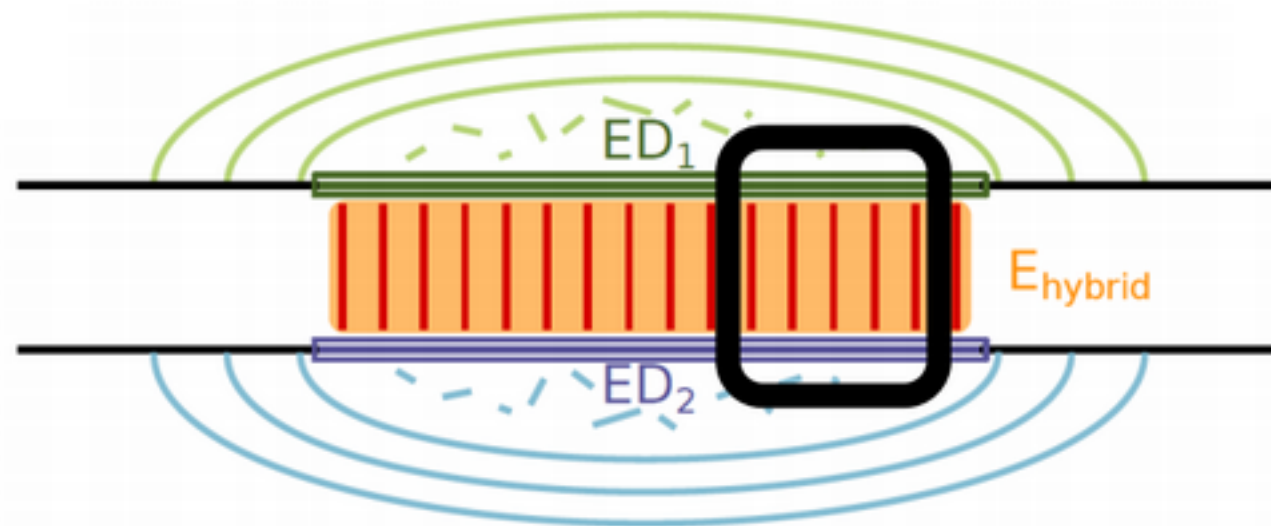


■ Seed



eg: sTarPicker  
RiBlast  
RiSearch2

# Exact memory efficient method



Search minimum over all seeds using:

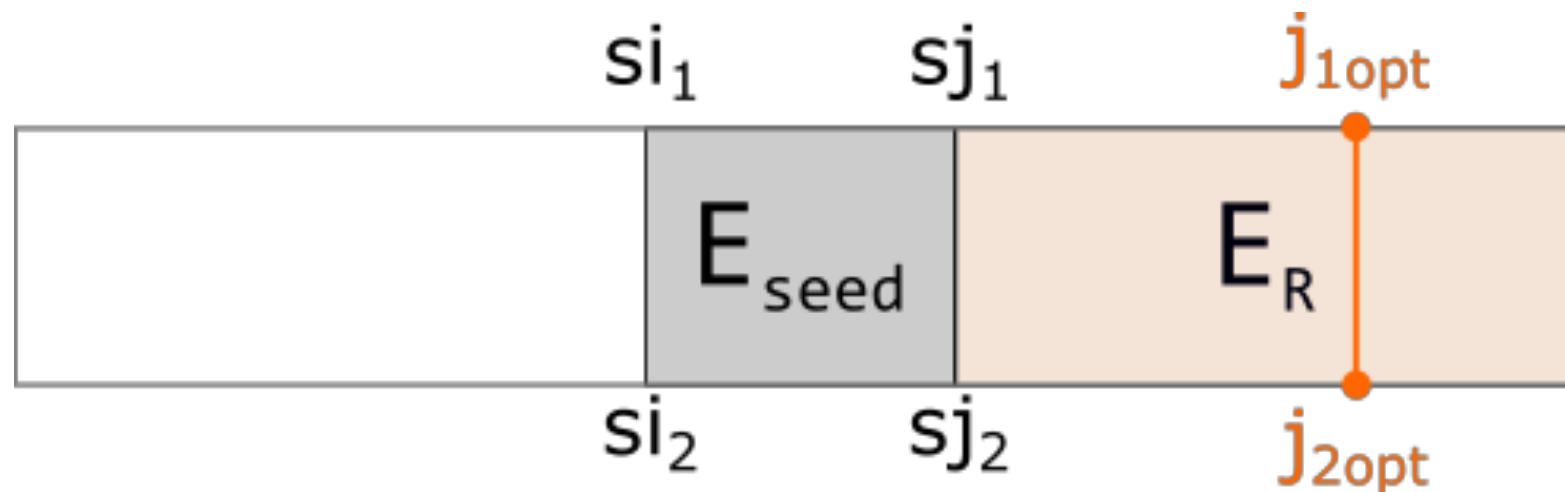
$$\min_{i_1, i_2} \left[ E_{\text{loop}} + E_L \right] + E_{\text{seed}} + \min_{j_1, j_2} \left[ E_R + E_{\text{loop}} \right]$$

The equation shows the search for the minimum energy over all seeds using the following components:

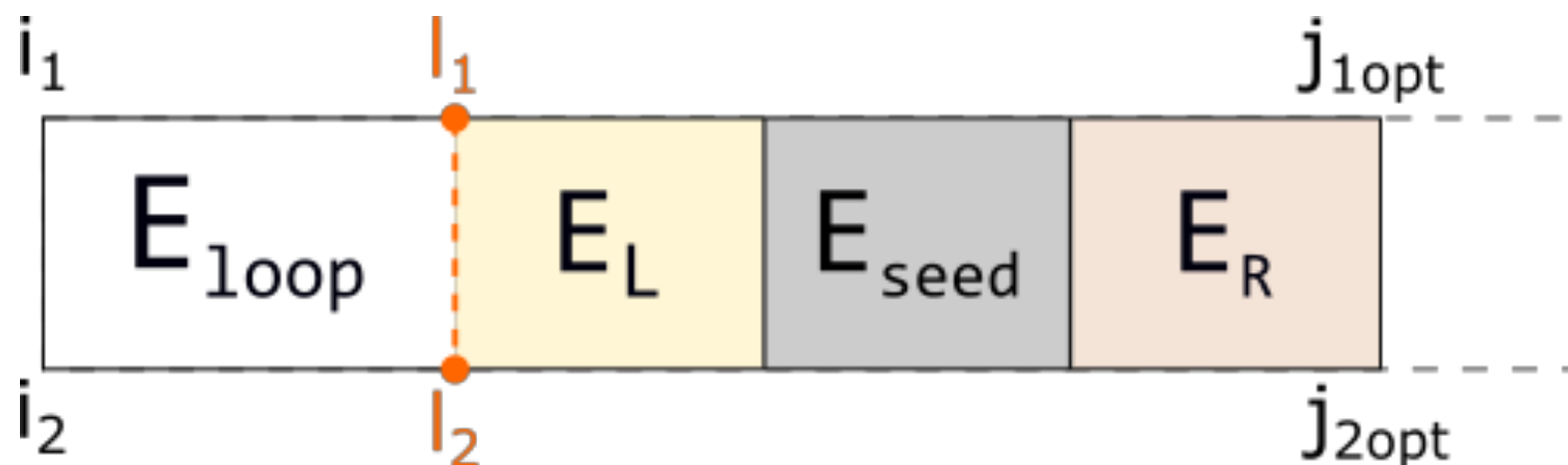
- $\min_{i_1, i_2}$ : Minimum energy over the left region, defined by coordinates  $i_1$  and  $i_2$ .
- $E_{\text{loop}}$ : Energy of the loop region (yellow box).
- $E_L$ : Energy of the left region (yellow box).
- $E_{\text{seed}}$ : Energy of the seed region (gray box).
- $\min_{j_1, j_2}$ : Minimum energy over the right region, defined by coordinates  $j_1$  and  $j_2$ .
- $E_R$ : Energy of the right region (orange box).
- $E_{\text{loop}}$ : Energy of the loop region (orange box).

# Heuristic method

- First find  $j_1$  and  $j_2$  that minimize  $E_R$

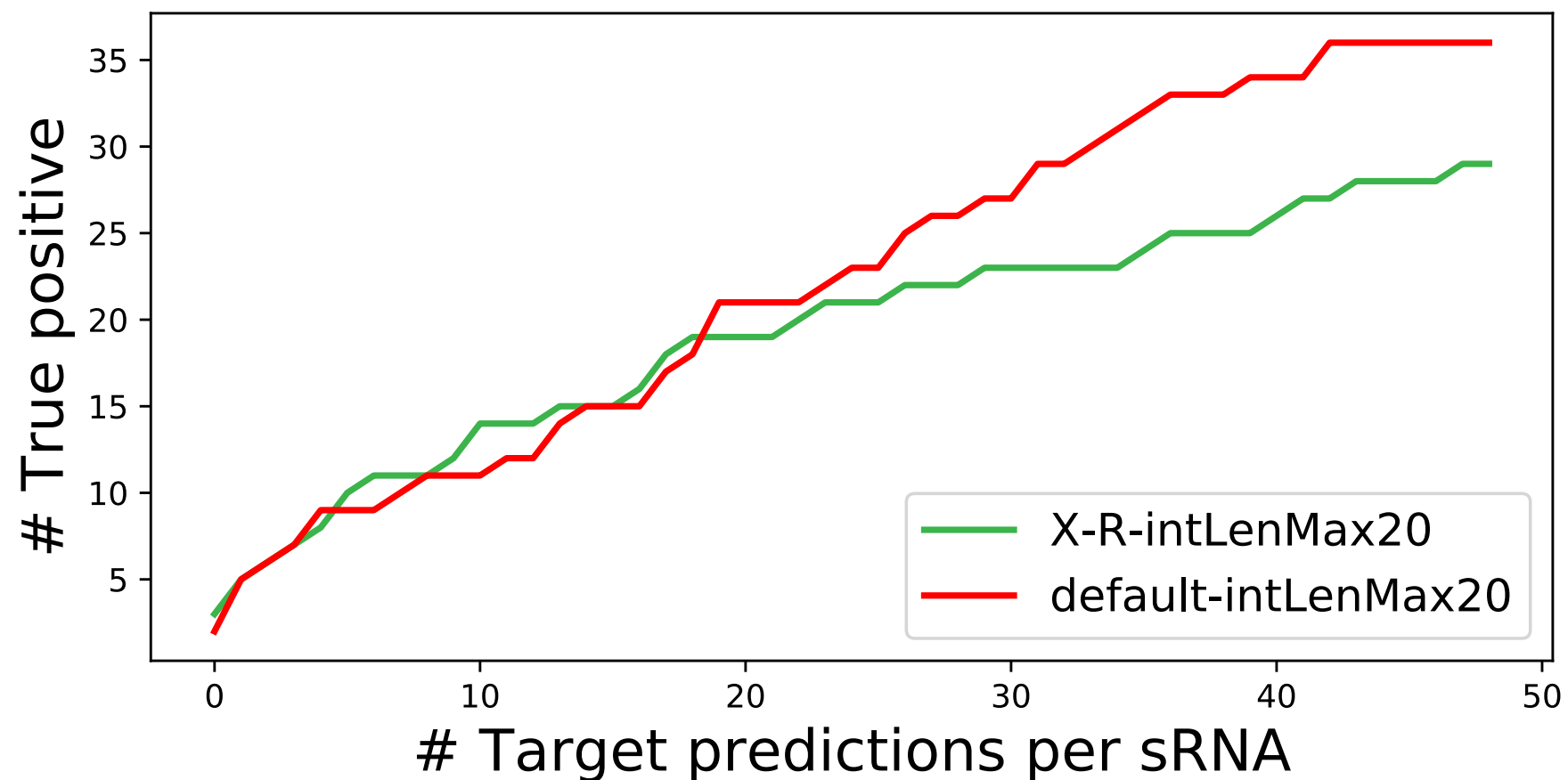
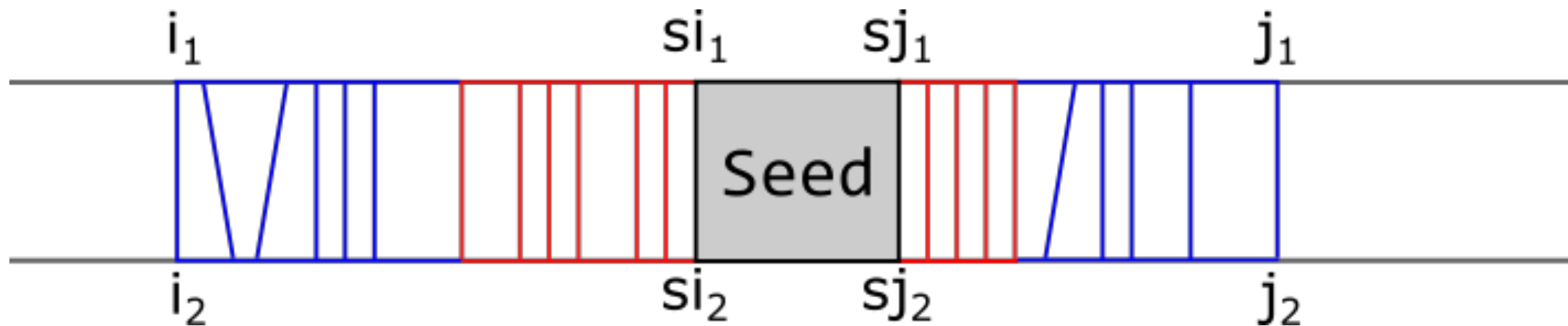


- Then minimize over entire interaction up to  $j_{1\text{opt}}$ ,  $j_{2\text{opt}}$

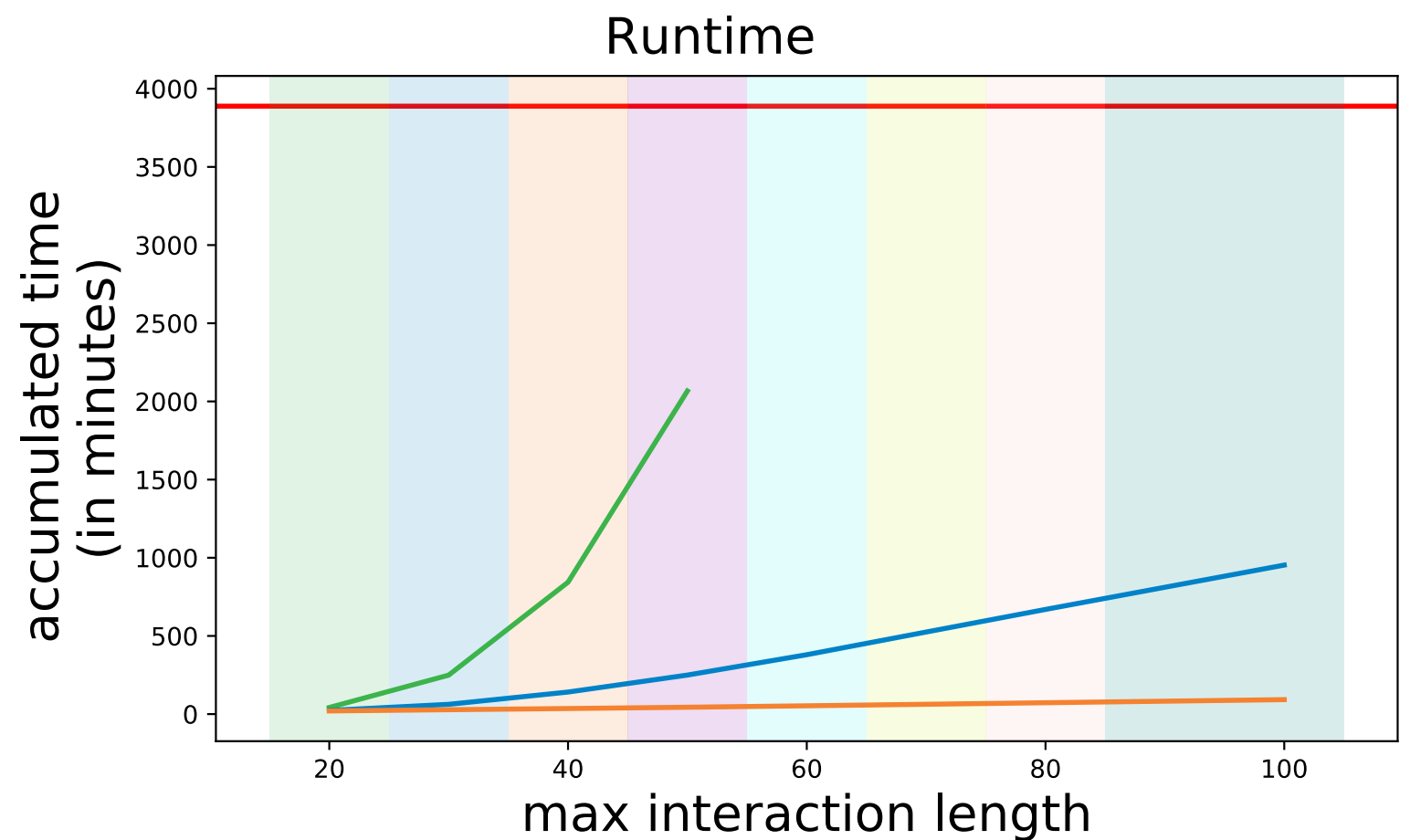
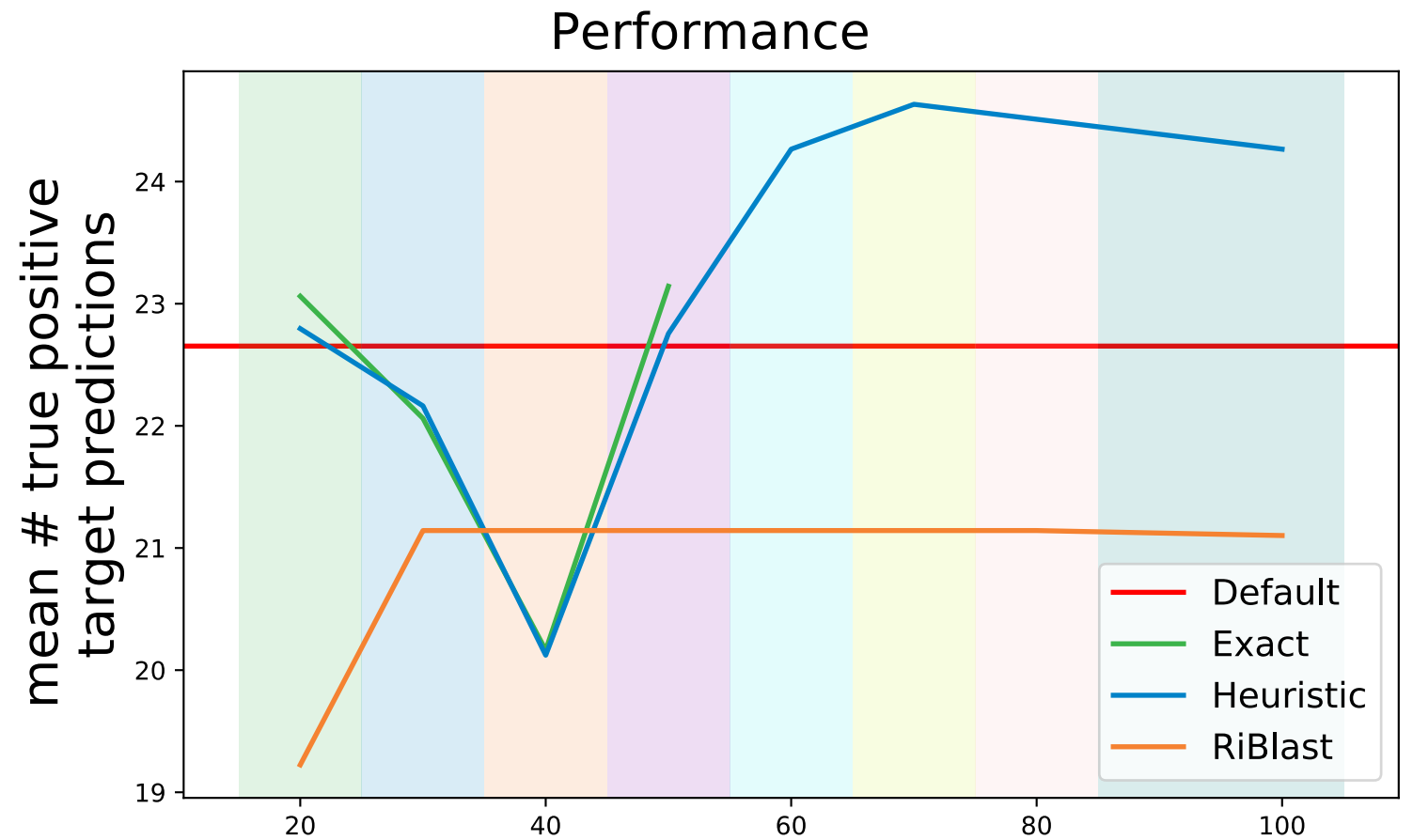
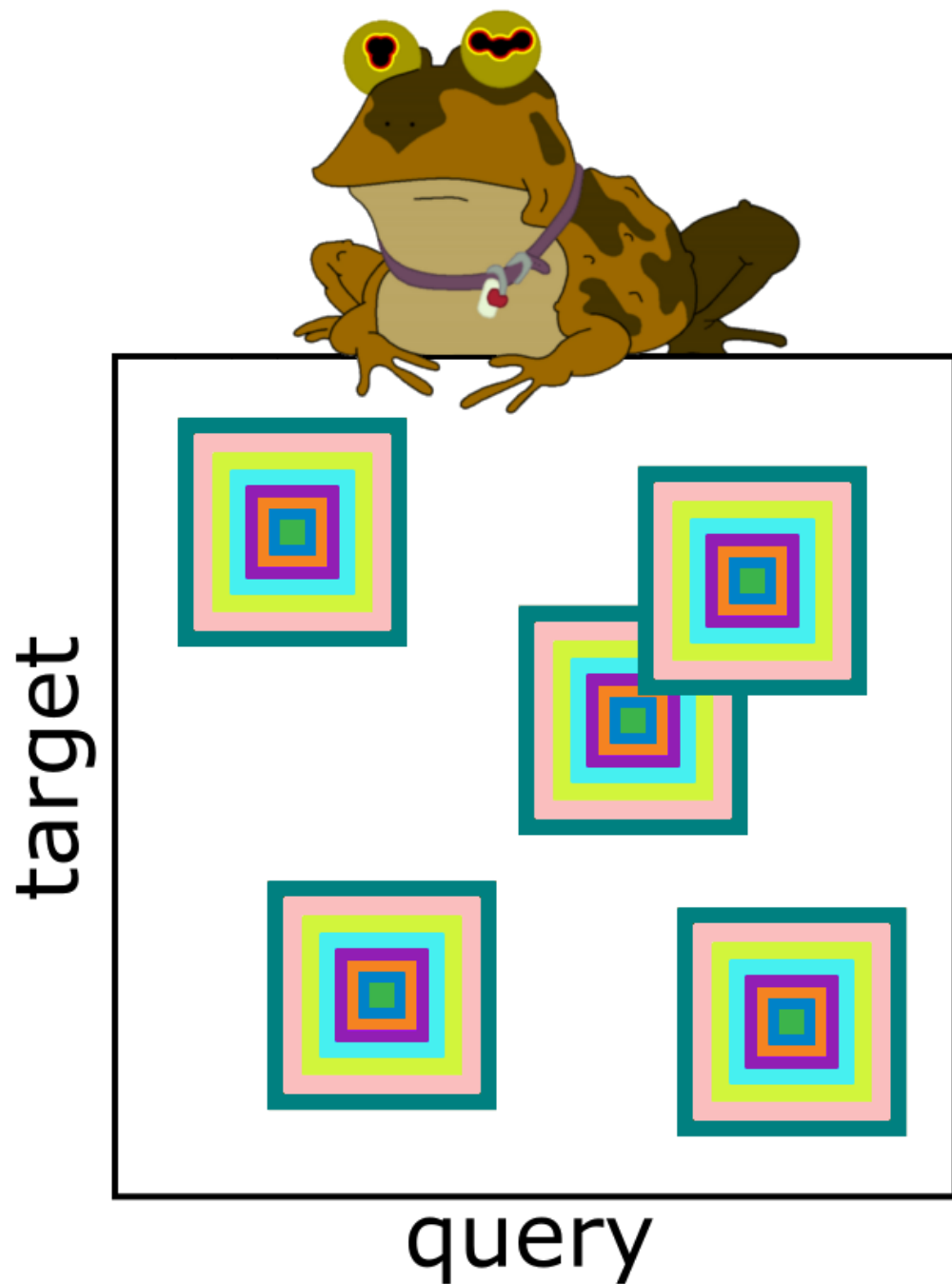


# RiBlast method

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



# Length-dependence



# Partition function

Goal: Find probabilities of interactions

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

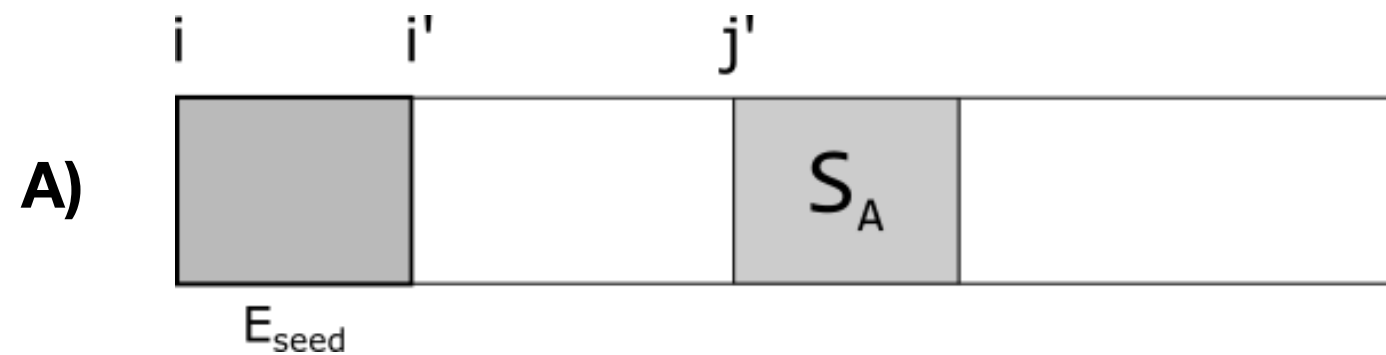
$$Z_{i,j} = \sum_{S_A} \left( \begin{array}{|c|c|c|c|} \hline \text{no seed} & S_A & \boxed{S'} & \boxed{S''} \\ \hline \end{array} \right)^{i,j}$$

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

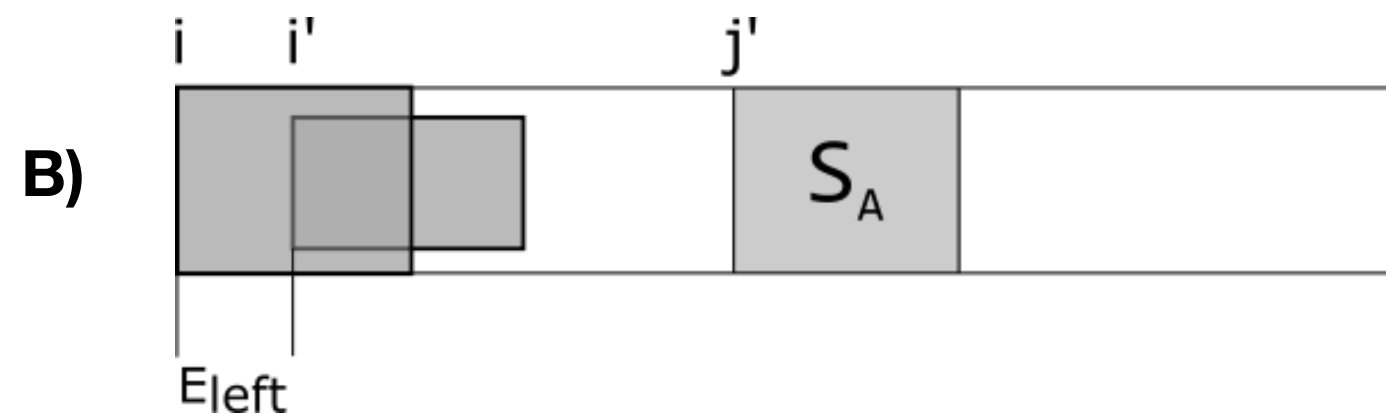
$$Z_{i,j} = \sum_{S_A} \left( \begin{array}{|c|c|c|c|} \hline \text{no seed} & S_A & \boxed{S'} & \boxed{S''} \\ \hline \end{array} \right)^i_j$$

Idea: remove seed for i

Cases:



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



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