UNIVERSITÄT LEIPZIG

RNA design | objectives, solution landscapes and optimization methods

Stefan Hammer

Bioinformatik Leipzig

Bled, February 2019

Design objectives for novel ncRNAs

- Combine functional components
- Gain novel functionality
- Respect biological target system

Design methods

- Experimental selection/screening:
  - Desired functionality (experimental assay)
  - Selection/screening method
  - Candidate characterization
- Computational *de novo* design:
  - Desired functionality (*in silico* model)
  - Sampling/optimization method
  - Functional testing and characterization

Rational *de novo* design

1) Design objective:
   - Functionality of components (aptamer, terminator,…)
   - Novel overall functionality
   - Respect the biological target system
2) Sequence generation:
   - Constraint sampling
   - Weighted sampling
3) Optimization method
4) *In silico* filtering and characterization

Single-state RNA devices

- Design objectives purely based on structure

- Positive design
  - Target structure must be thermodynamically stable
- Negative design
  - Contrary structures must be less stable and thus less

Single-state objectives

- MFE defect
- Probability defect
- Element probability defect
- Enforced element probability defect
- Ensemble defect

MFE defect

```
.(((....))).....  f(x) = 0
((((....))))....  f(x) = 1
..((....))......  f(x) = 1
.(((....)))(...)  f(x) = 1
.(((....)).)....  f(x) = 2
....((((....))))  f(x) = 7
................  f(x) = 3
f(x) ∈ [0,∞)
```

$$f(x) = D(\phi_{\text{target}}, \phi_{\text{MFE}})$$

Probability defect

```
.(((....))).....  0.6
((((....))))....  0.1
..((....))......  0.1
.(((....)))(...)  0.05
.(((....)).)....  0.05
....((((....))))  0.05
................  0.05
f(x) = 0.60 ∈ [0,1]
```

$$f(x) = P(x|\phi_{\text{target}}) \propto G(x|\phi) - G(x|\Phi)$$

$$B(x|\phi) = \exp(-\frac{G(x|\phi)}{RT})$$
$$Z(x) = \sum_{\phi \in \Phi} B(x|\phi)$$
$$P(x|\phi) = \frac{B(x|\phi)}{Z} = \exp(-\frac{G(x|\phi) - G(x|\Phi)}{RT})$$

$x \ldots$ Sequence
$\phi \ldots$ Structure
$\Phi \ldots$ Ensemble of structures
$G \ldots$ Gibbs free energy
$Z \ldots$ Partition function
$B \ldots$ Boltzmann weight
$R \ldots$ Gas constant
$T \ldots$ Temperature

Element probability defect

```
.(((....)))..... 0.6
((((....)))).... 0.1
..((....))...... 0.1
.(((....)))(...) 0.05
.(((....)).).... 0.05
....((((....)))) 0.05
................ 0.05
f(x) = 0.90 ∈ [0,1]
```

$$f(x) = P(x|\Phi_{\text{target}}) \propto G(x|\Phi_{\text{target}}) - G(x|\Phi)$$

$$P(x|\Phi_{\text{target}}) = \exp(-\frac{G(x|\Phi_{\text{target}}) - G(x|\Phi)}{RT})$$

Enforced element probability defect

```
.(((....))).....  0.6
((((....))))....  0.1
..((....))......  0.1
.(((....)))(...)  0.05
.(((....)).)....  0.05
....((((....))))  0.05
................  0.05
f(x) = 0.75 ∈ [0,1]
```

$$f(x) = P(x|\Phi_{\text{target}}) \propto G(x|\Phi_{\text{target}}) - G(x|\Phi)$$

$$P(x|\Phi_{\text{target}}) = \exp(-\frac{G(x|\Phi_{\text{target}}) - G(x|\Phi)}{RT})$$

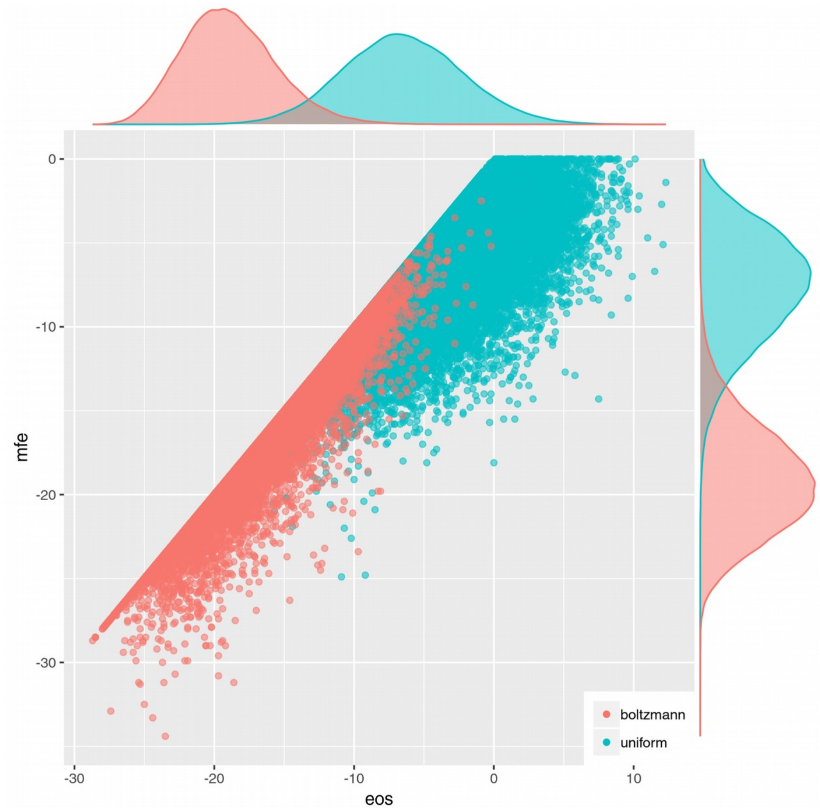
Ensemble defect

```
.(((....))).....  0.6·0
((((....))))....  0.1·1
..((....))......  0.1·1
.(((....)))(...)  0.05·1
.(((....)).)....  0.05·2
....((((....))))  0.05·7
................  0.05·3
f(x) = 0.85 ∈ [0,∞)
```

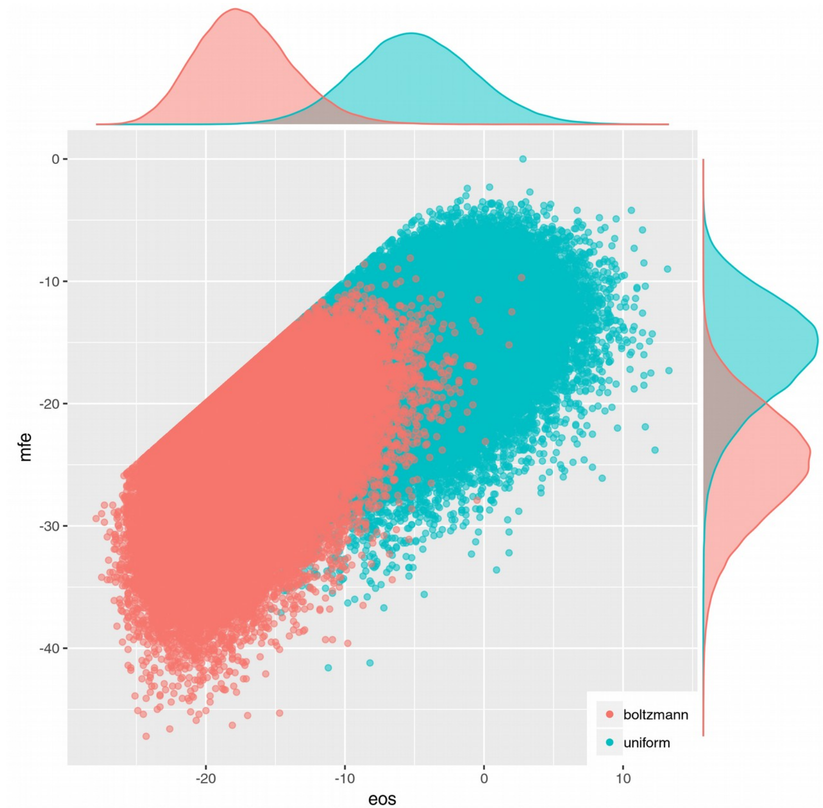
$$f(x) = \sum_{\phi \in \Phi} P(x|\phi_{\text{target}}) \times D(\phi, \phi_{\text{target}})$$

Example solution Landscapes

```
((((((....))))))((((((....))))))
1e+4 from 1.42658e+14 sequences
```

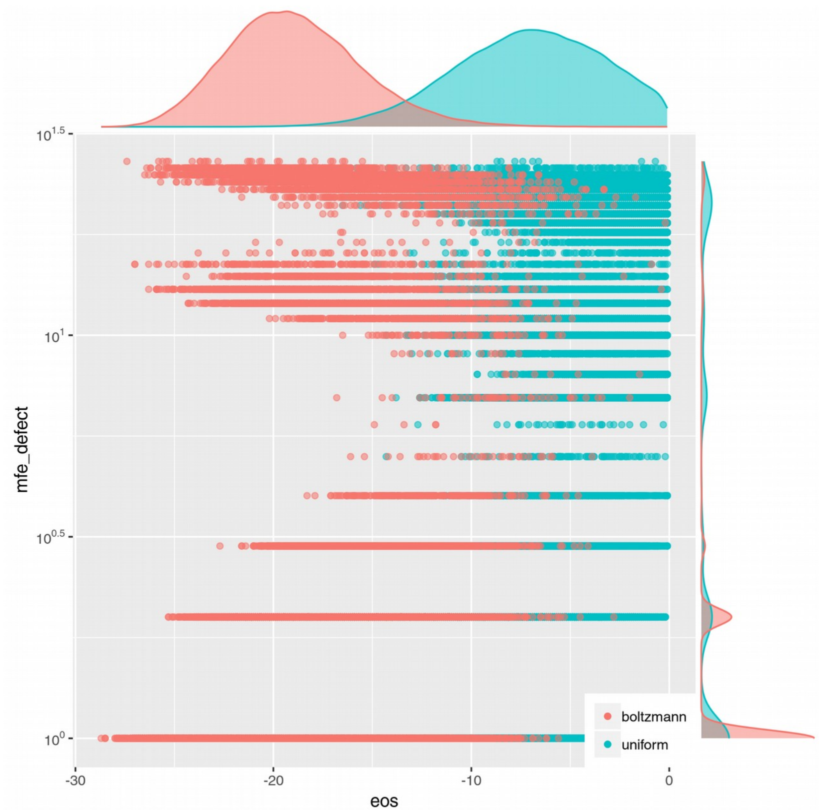



```
(((((...((((((((.....)))))...)))...))))).......................
1e+4 from 9.86838e+32 sequences
```



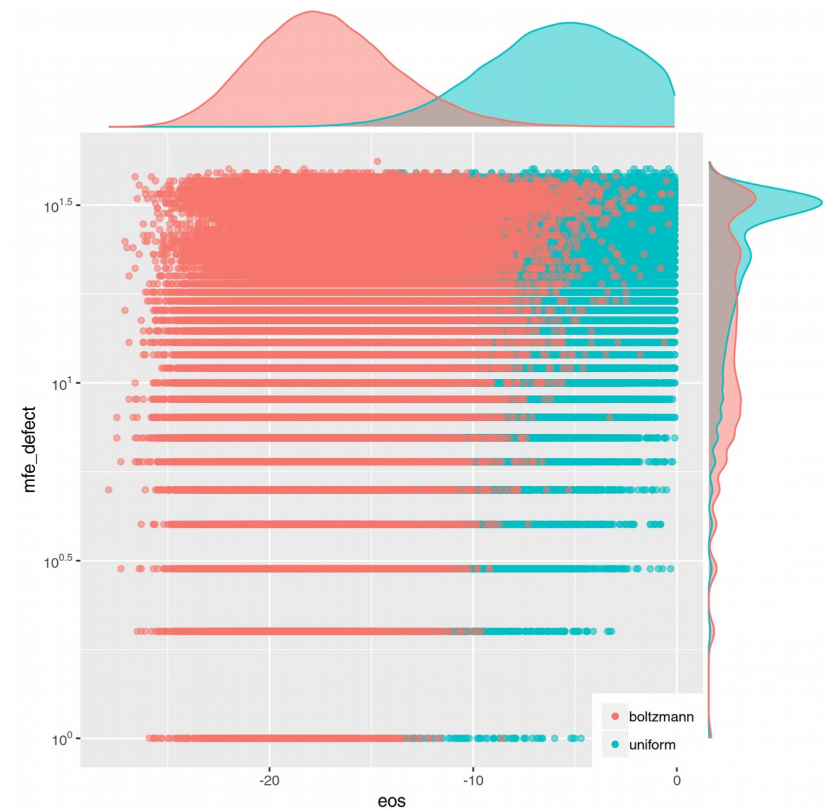
MFE defect

((((((....))))))((((((....))))))



$10^{1.5}$
10^{1}
$10^{0.5}$
10^{0}
mfe_defect
-30
-20
-10
0
eos
boltzmann
uniform

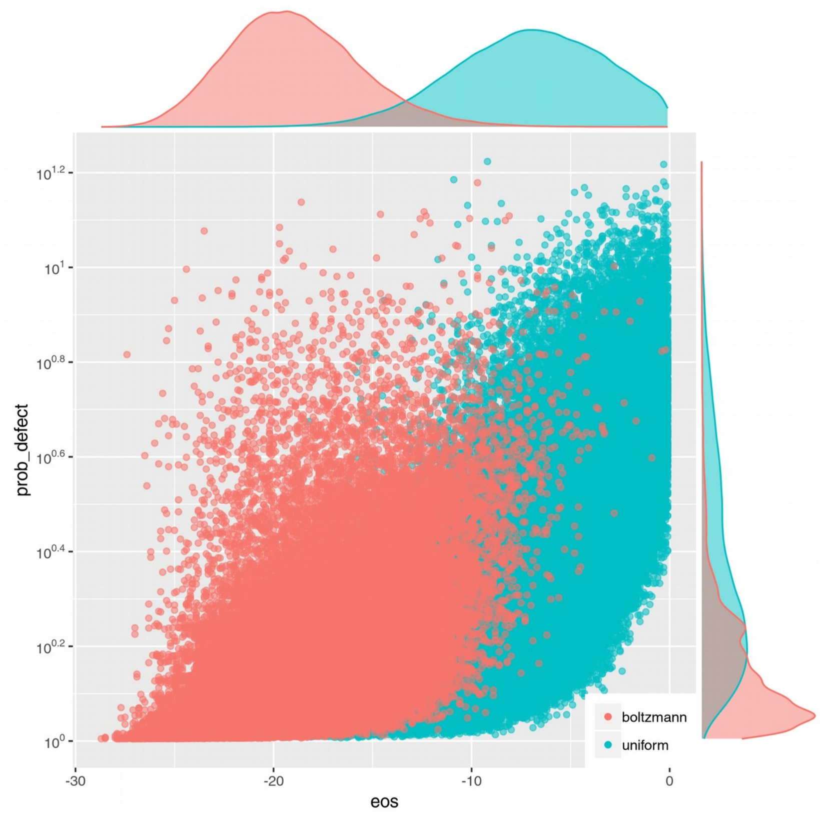
(((((...((((((((.....)))))...)))...))))))........................



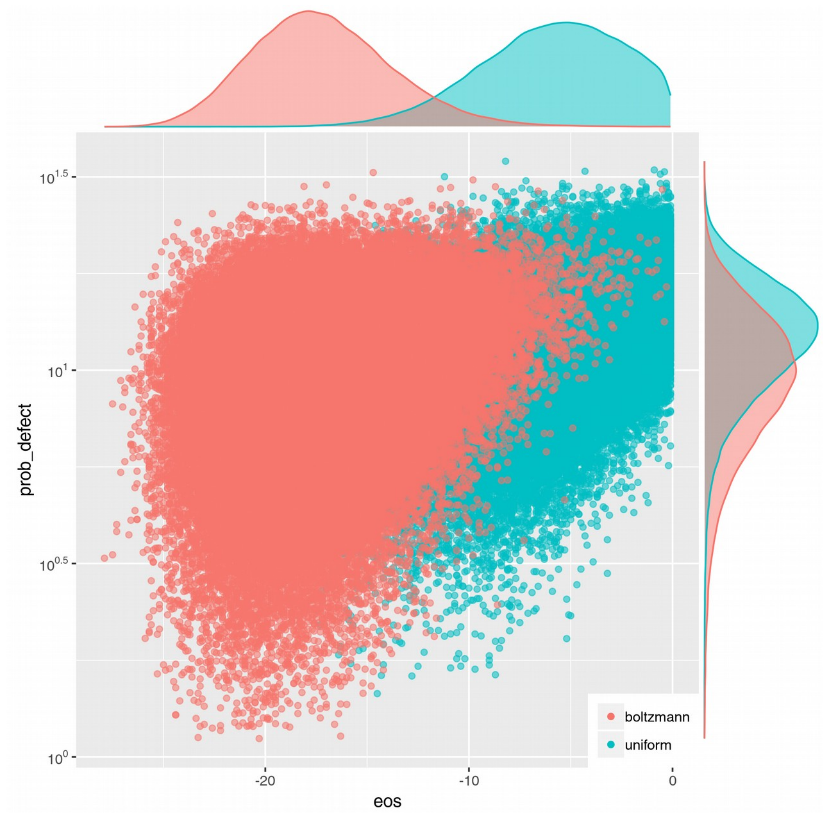
$10^{1.5}$
10^{1}
$10^{0.5}$
10^{0}
mfe_defect
-20
-10
0
eos
boltzmann
uniform

Probability defect

```
((((((....))))))((((((....))))))
```

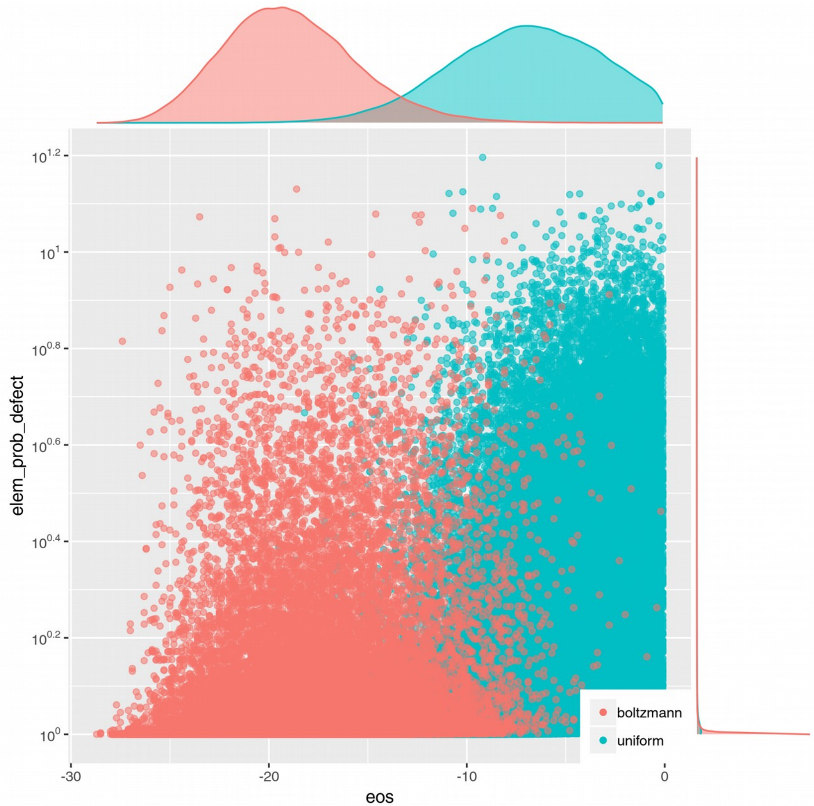



```
(((((...((((((((.....)))))...)))...))))).........................
```

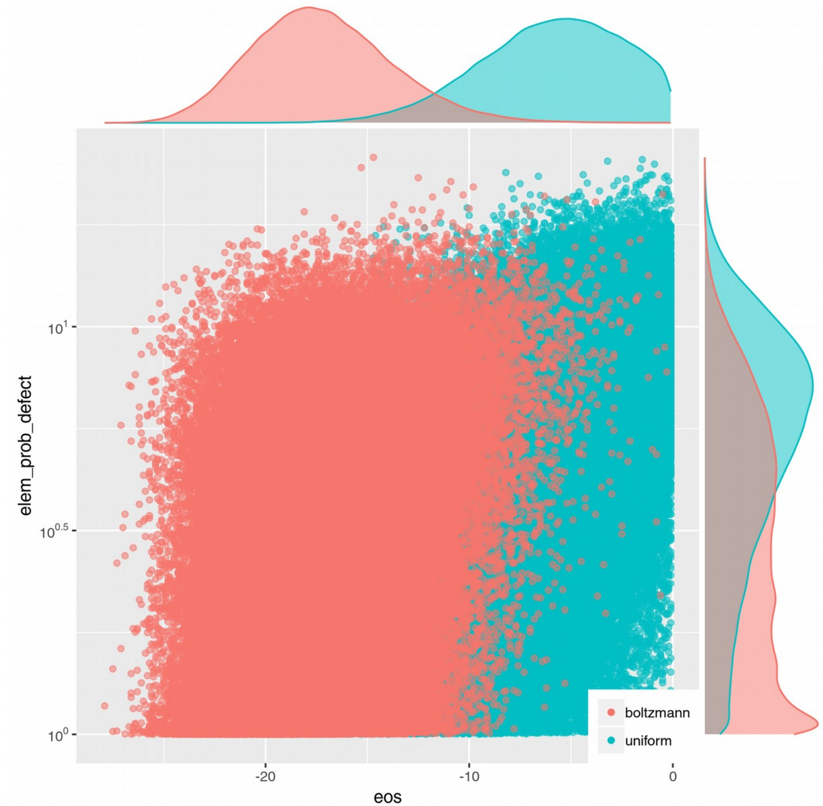


Element probability defect

((((((....))))))((((((....))))))

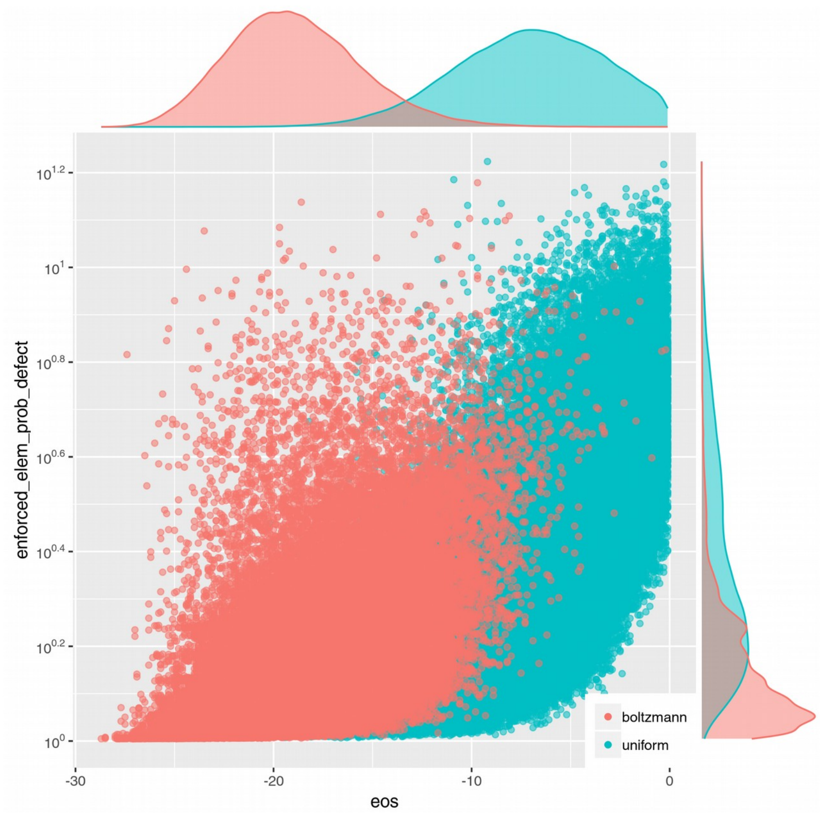


(((((...((((((((.....)))))...)))...))))).......................

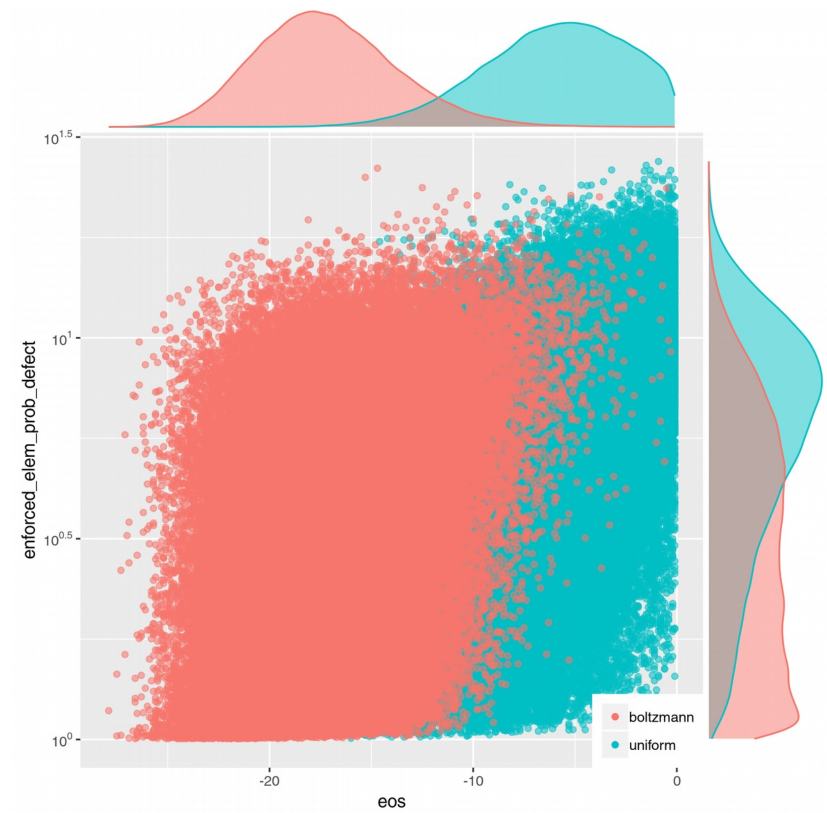


Enforced element probability defect

((((((....))))))((((((....))))))

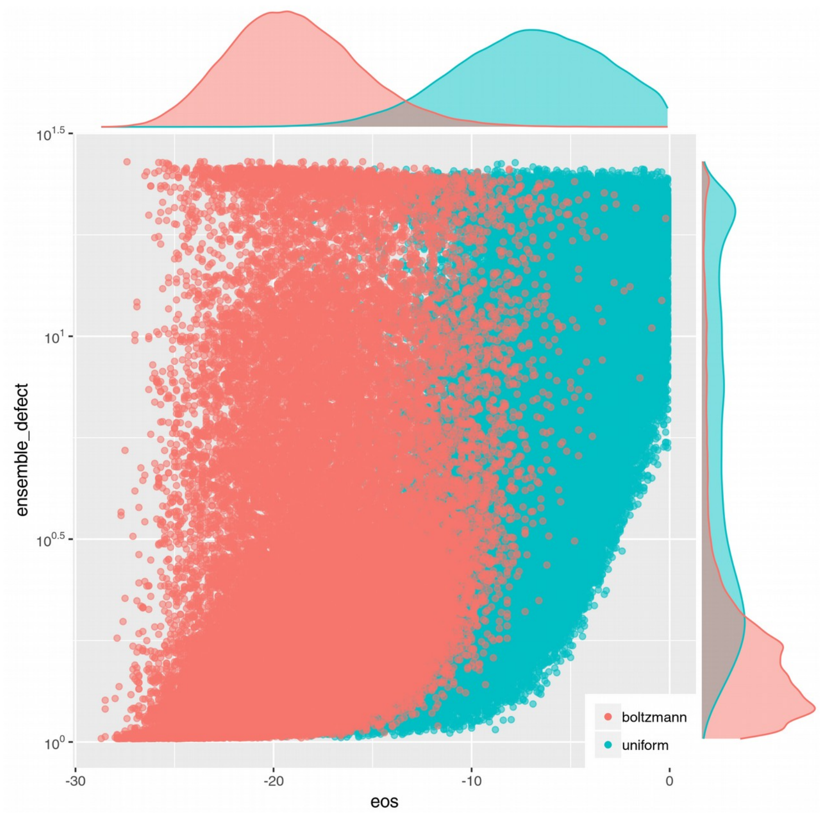


(((((...((((((((.....)))))...)))...))))).......................

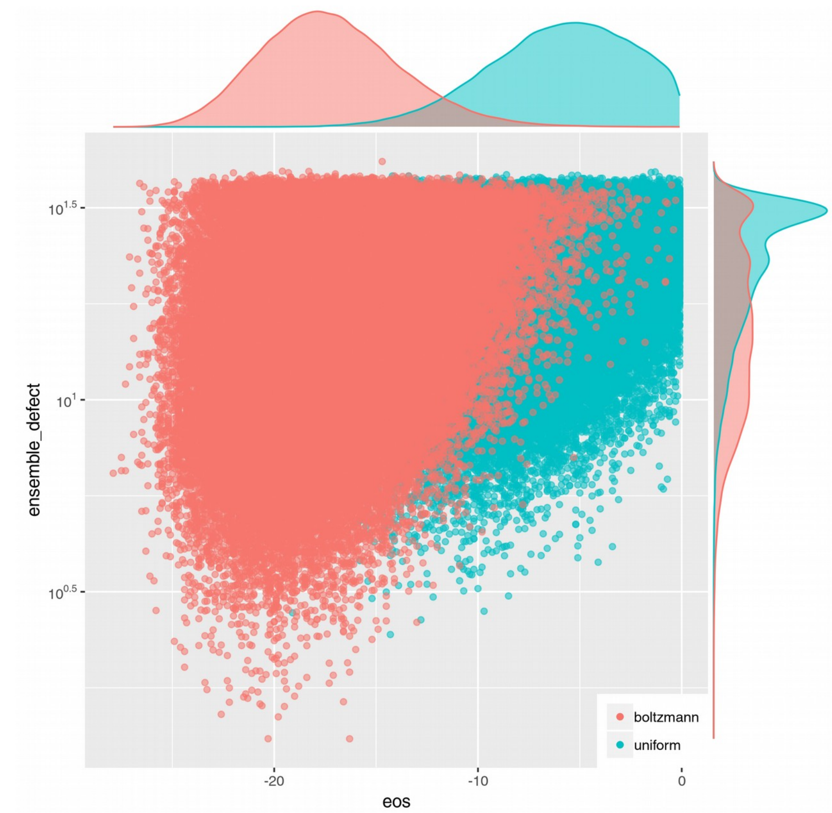


Ensemble defect

((((((....))))))((((((....))))))



(((((...((((((((.....)))))...)))...))))).......................

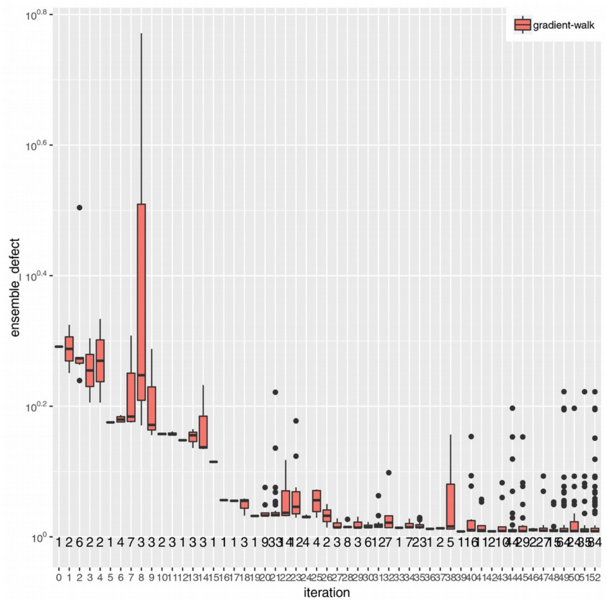


Optimization methods

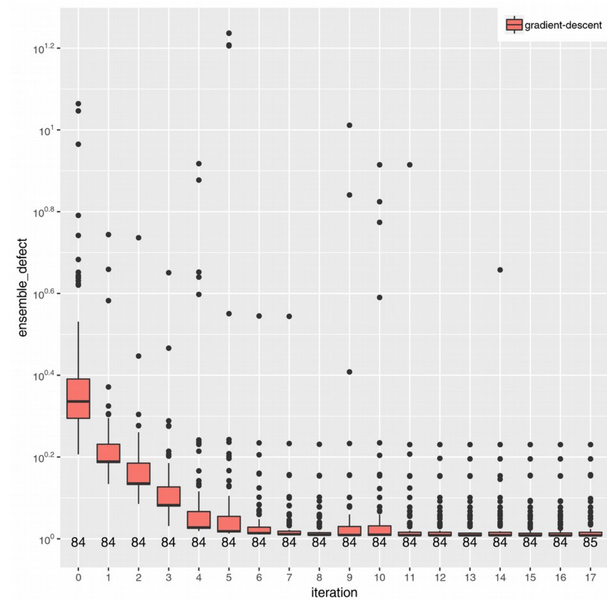
- Gradient walk
  - Accept first better solution
- Gradient descent
  - Accept best neighbouring solution
- Simulated annealing
  - Accept also worse solution with decreasing probability

Comparison optimization methods

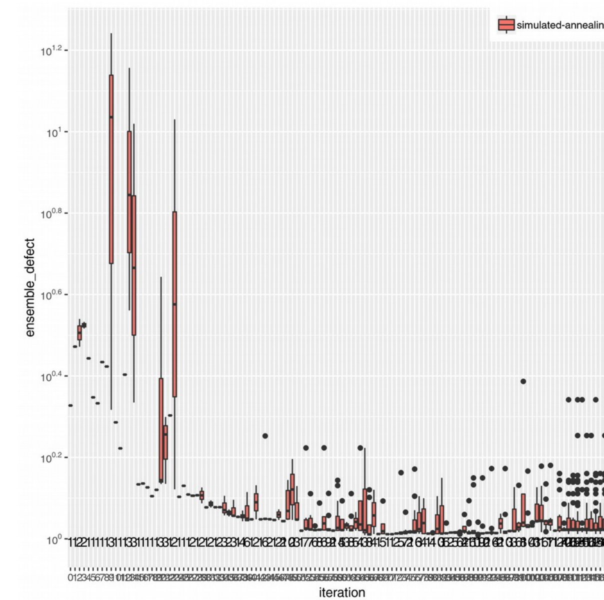
Gradient walk (N=~200)



Gradient descent (N=~1500)

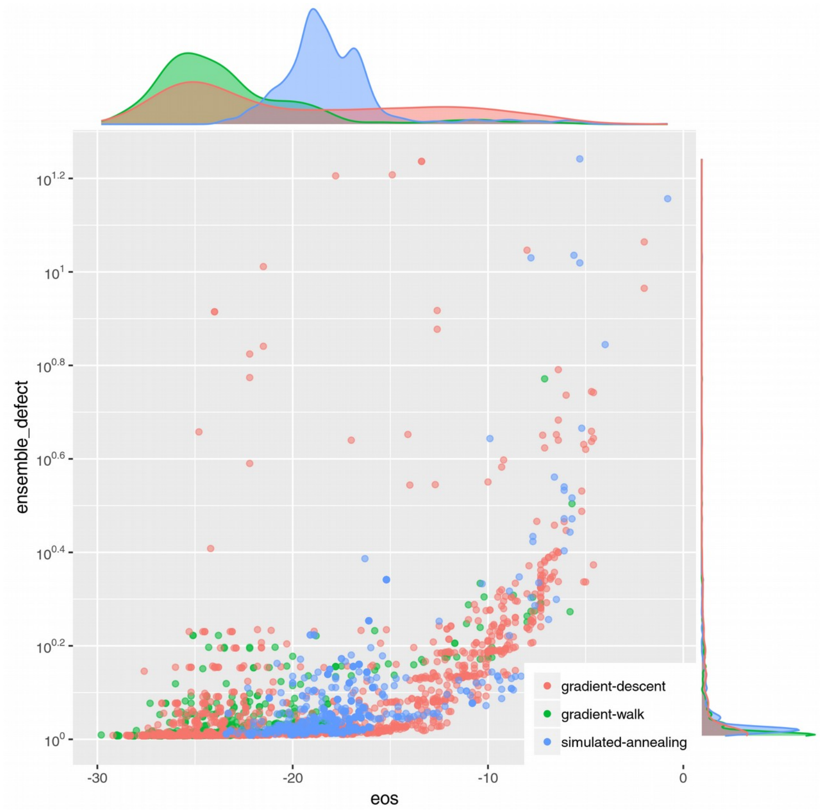


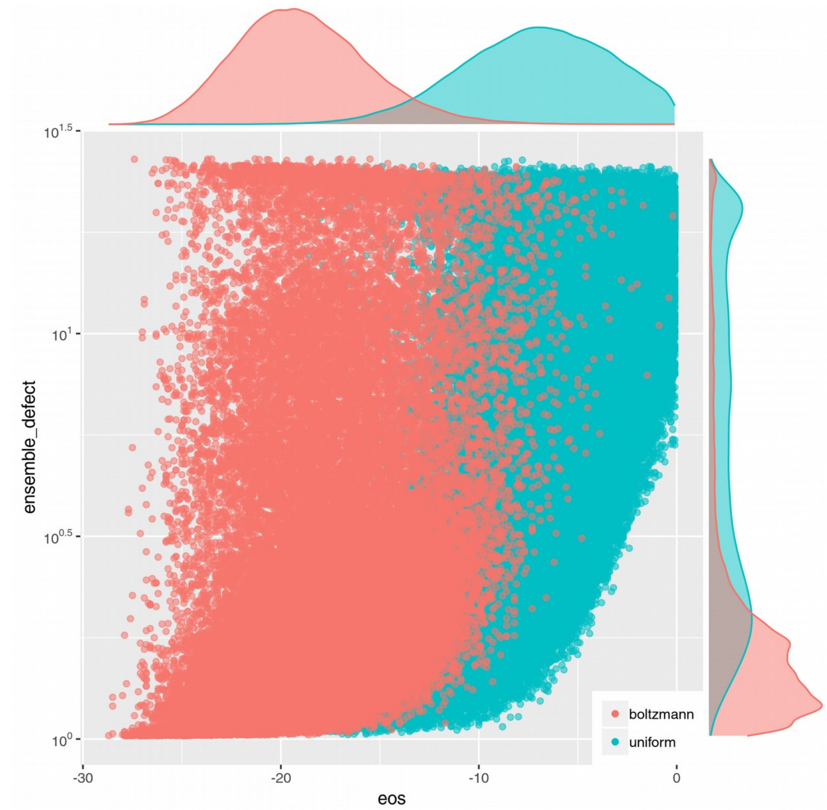
Simulated annealing (N=~200)



((((((....))))))((((((....))))))

Comparison optimization methods





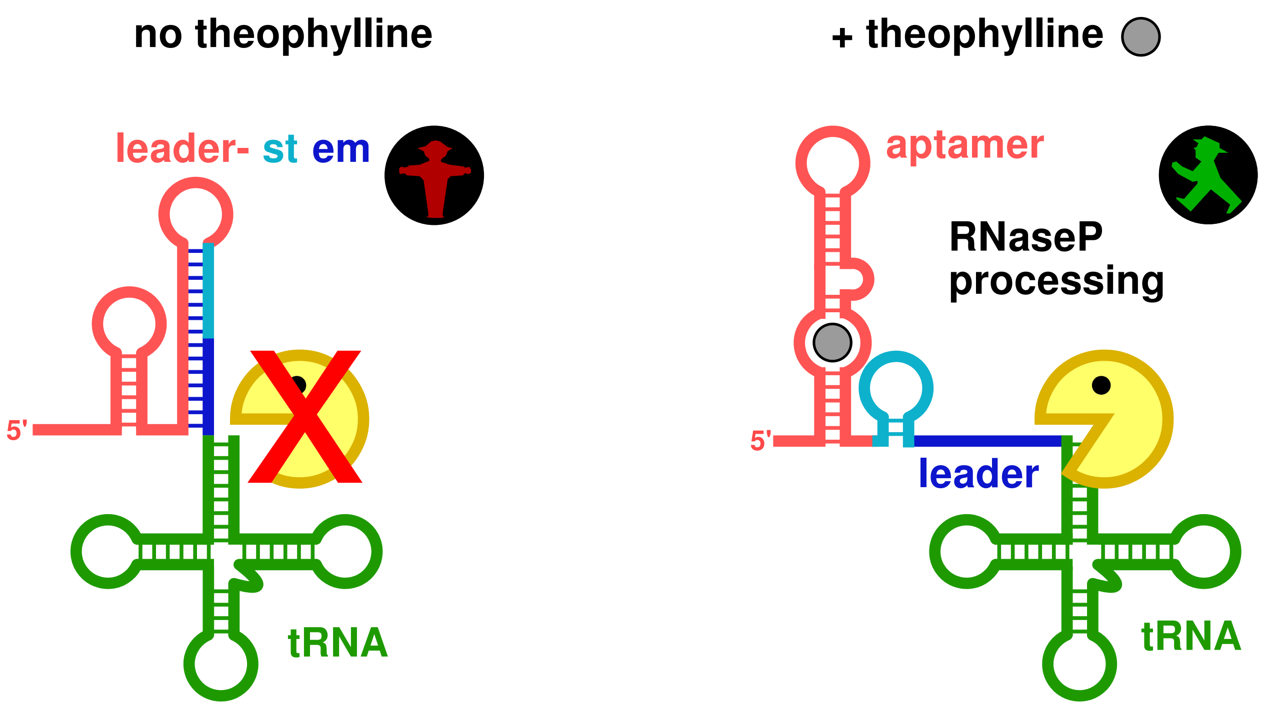
((((((....))))))((((((....))))))

Multi-state RNA devices

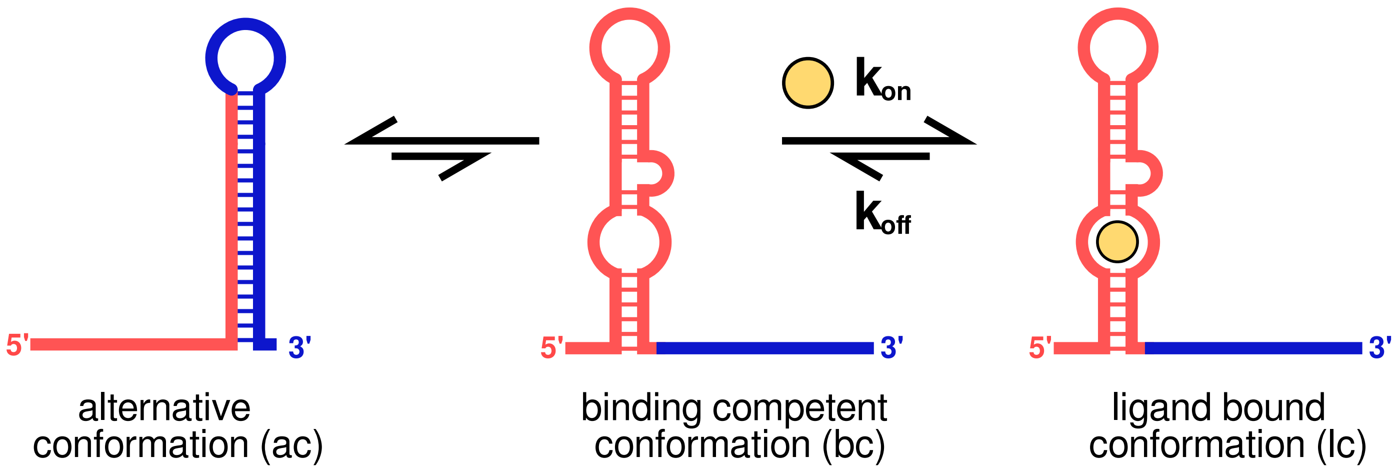
- Conditional states due to trigger:
  - Temperature
  - Ligand binding
  - RNA-RNA interaction
- Simultaneous states

Conditional states: RNaseP riboswitch

$$f(x) = f(x, \text{condition}_1) \cdot f(x, \text{condition}_2)$$



Simultaneous states: Ligand dependent switch



$$f(x) = P(x|\Phi_{\text{lc}}) \cdot (1 - |a - P(x|\Phi_{\text{ac}})|) \cdot (1 - |b - P(x|\Phi_{\text{bc}})|)$$

$$a + b = 1$$

$$0 <= a, b <= 1$$

Findeiß, Sven, Stefan Hammer, Michael T. Wolfinger, Felix Kühnl, Christoph Flamm, and Ivo L. Hofacker. 2018. "In Silico Design of Ligand Triggered RNA Switches." Methods. https://doi.org/10.1016/j.ymeth.2018.04.003.

Acknowledgements

- Sven Findeiß
- Sebastian Will
- Yann Ponty
- Ivo Hofacker
- Mario Mörl + Team
- Peter Stadler + Team

UNIVERSITÄT LEIPZIG

de.NBI GERMAN NETWORK FOR BIOINFORMATICS INFRASTRUCTURE

https://arxiv.org/abs/1902.01143

Evolving methods for rational *de novo* design of functional RNA molecules

Stefan Hammer, Christian Günzel, Mario Mörl, Sven Findeiß

Submitted to **METHODS**, Issue title: Chemcial Biology of RNA, Guest Editor: Michael Ryckelynck, Jan 2019

Gradient walk

- s ← seed sequence
- gradient_walk(s):
  - ∀ t ∈ neighbours(s):
    - if f(t) < f(s):
      - gradient_walk(t)
    - if we saw all neighbours(s):
      - return(s)

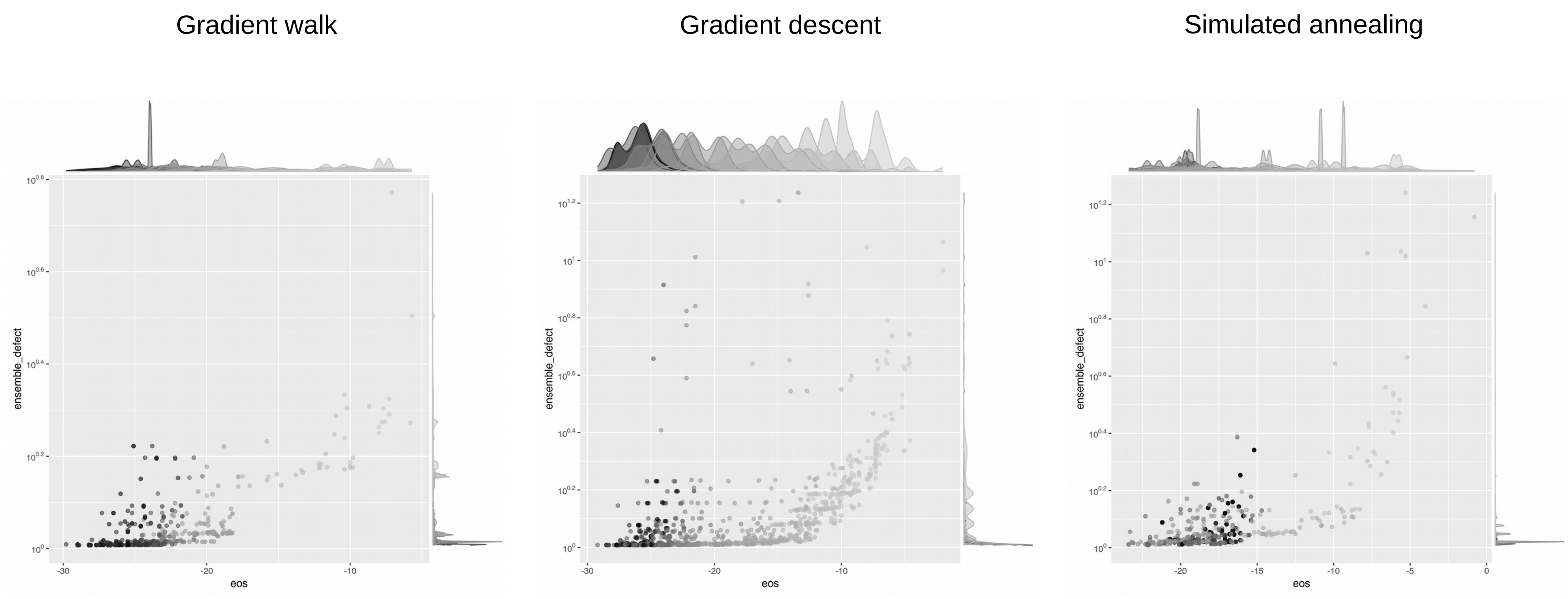
Gradient descent

- s ← seed sequence
- gradient_descent(s):
  - if min(f(t) ∀ t ∈ neighbours(s)) < f(s):
    - gradient_descent(s)
  - else:
    - return(s)

Simulated annealing

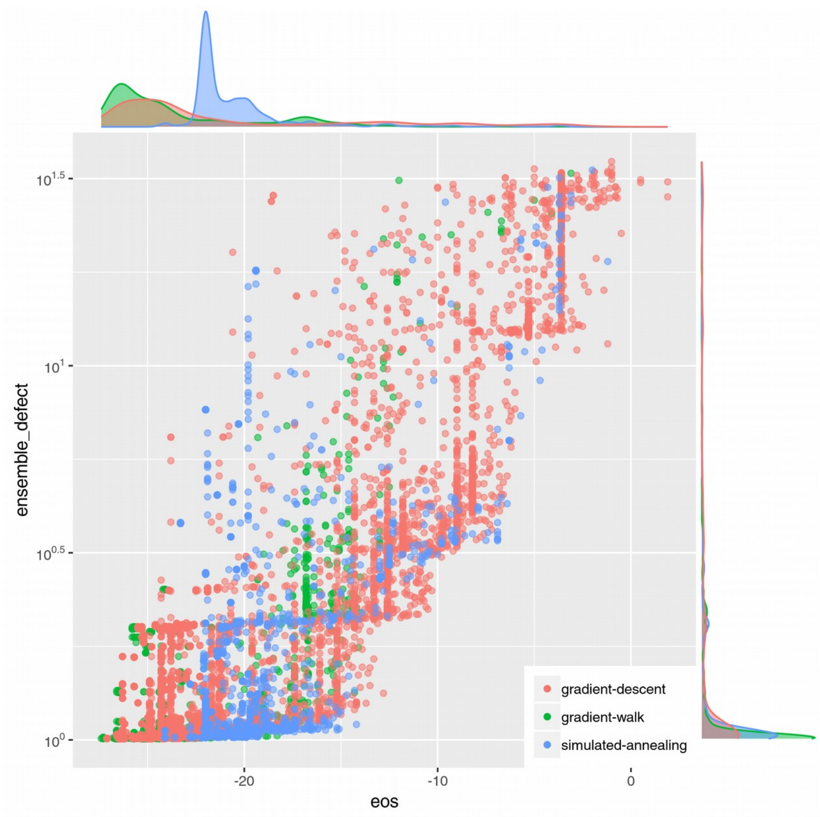
- s ← seed sequence
- simulated_annealing(s, temperature):
  - ∀ t ∈ neighbours(s):
    - p ← e^(-1*(f(t)-f(s))/temperature)
    - r ← random([0, 1])
    - if r <= p:
      - simulated_annealing(t, temperature*0.75)
    - if we saw all neighbours(s):
      - return(s)

Comparison optimization methods



(((((((....))))))((((((....))))))

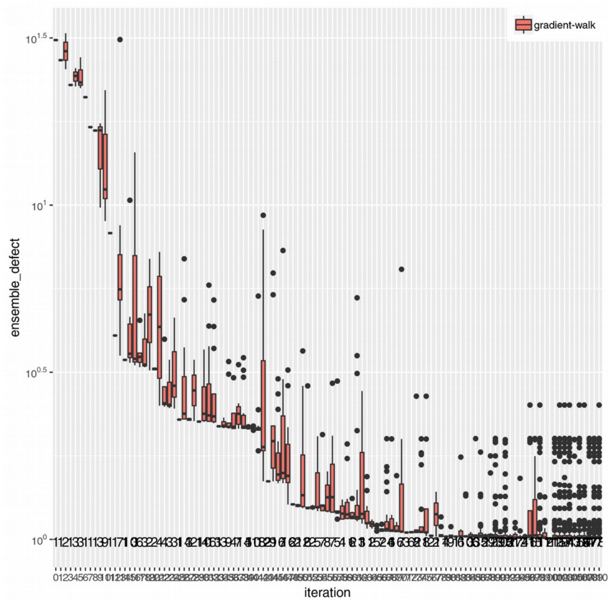
Comparison optimization methods



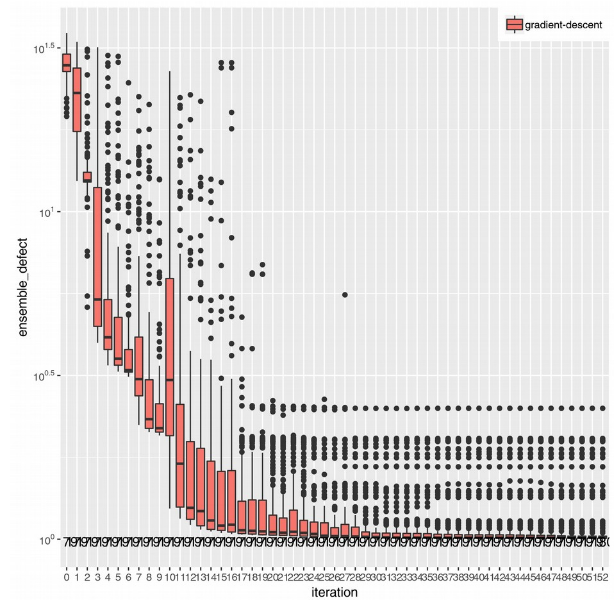
((((...((((((((.....)))))...)))...)))))........................

Comparison optimization methods

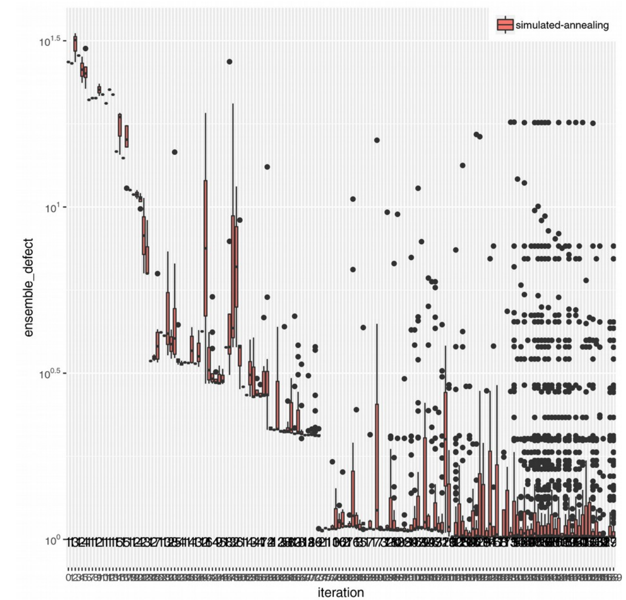
Gradient walk



Gradient descent



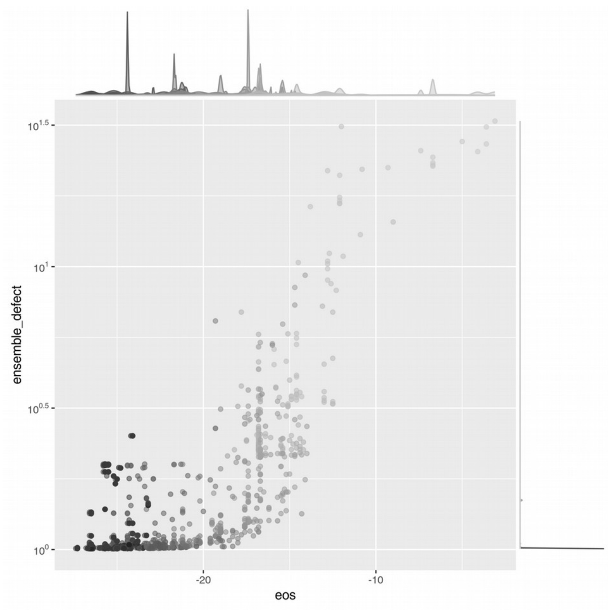
Simulated annealing



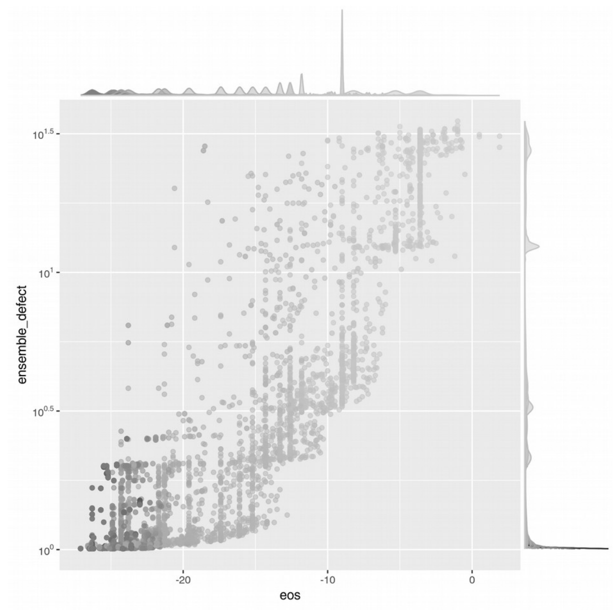
((((...((((((((.....)))))...)))...))))).......................

Comparison optimization methods

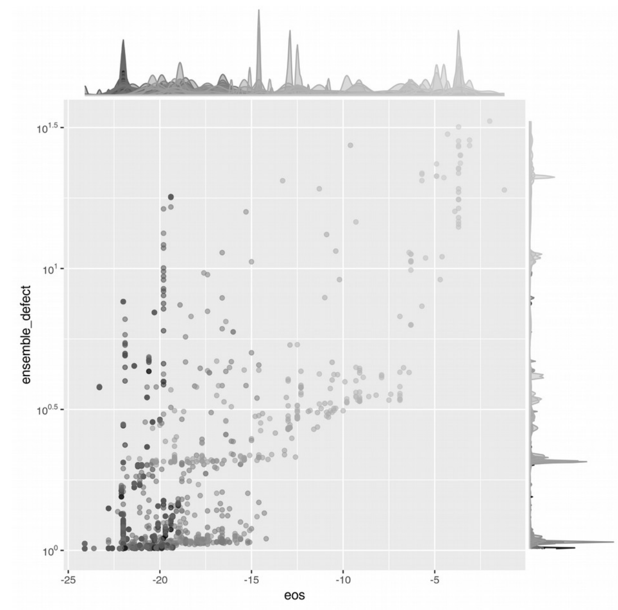
Gradient walk



Gradient descent



Simulated annealing



(((((...((((((((.....)))))...)))...)))))........................