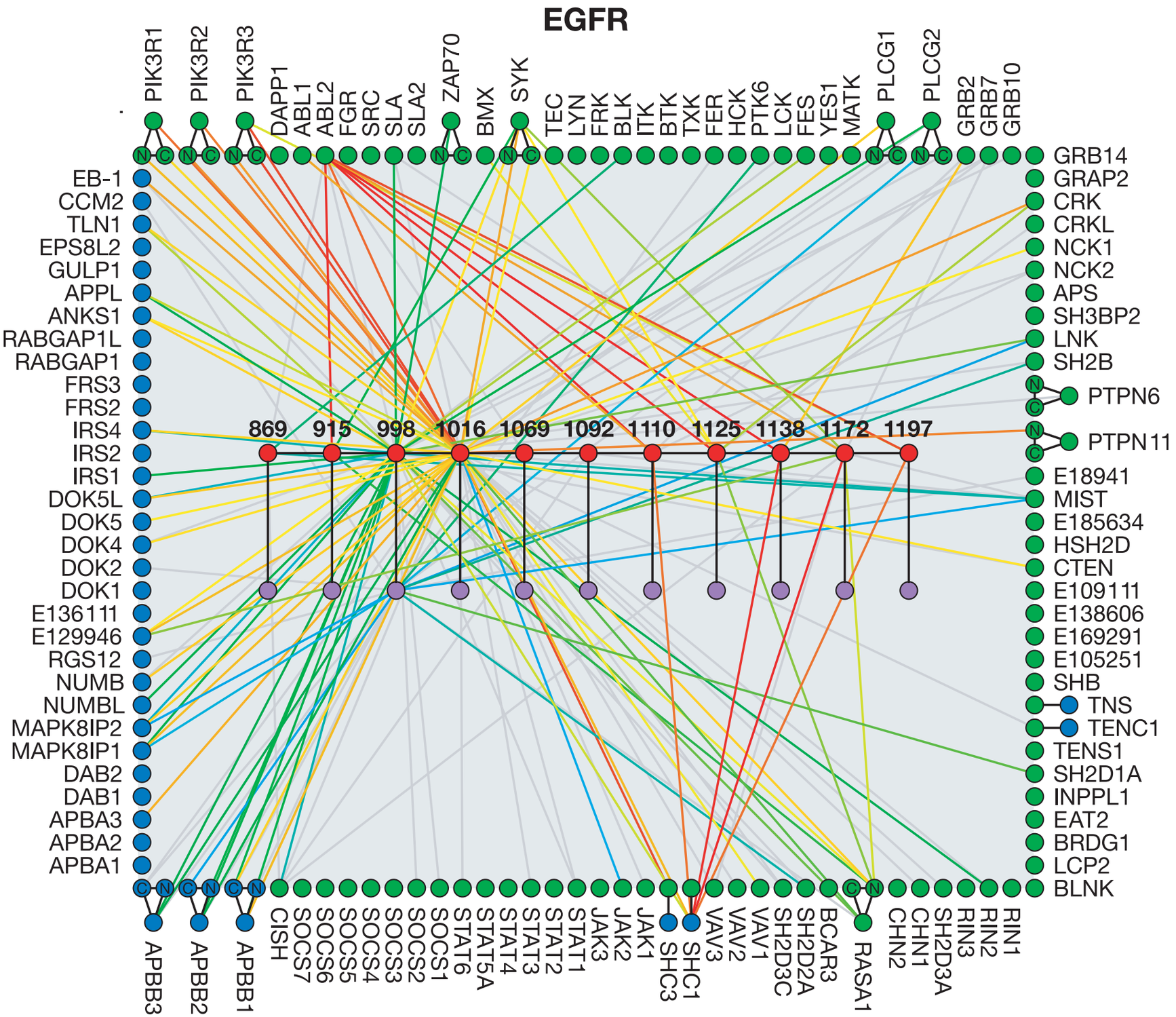
domains, sites, st

binding takes center stage



domains (cont')

Kim and Gerstein papers/hubs party-, date- ...

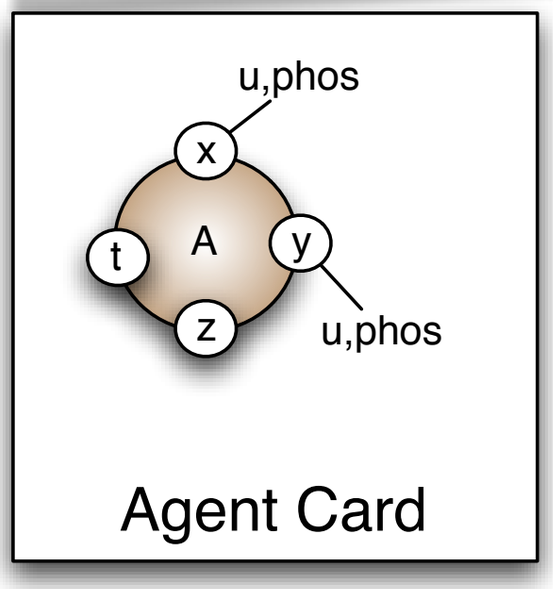
what are the forces driving this mass action graph rewriting?

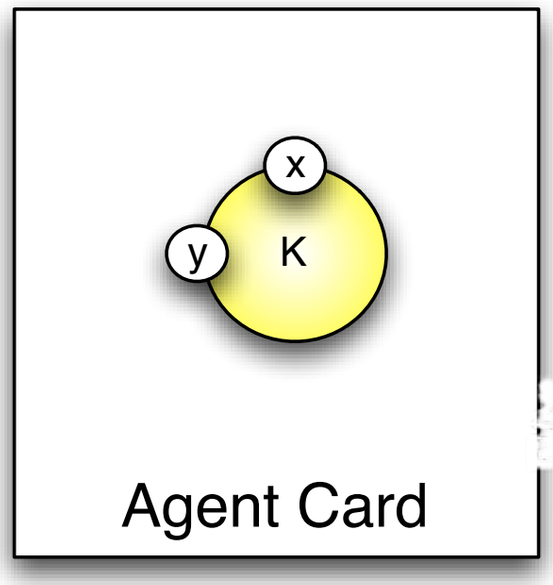
– randomness

– causalities/conditional binding

– geometric hindrance/structural biology

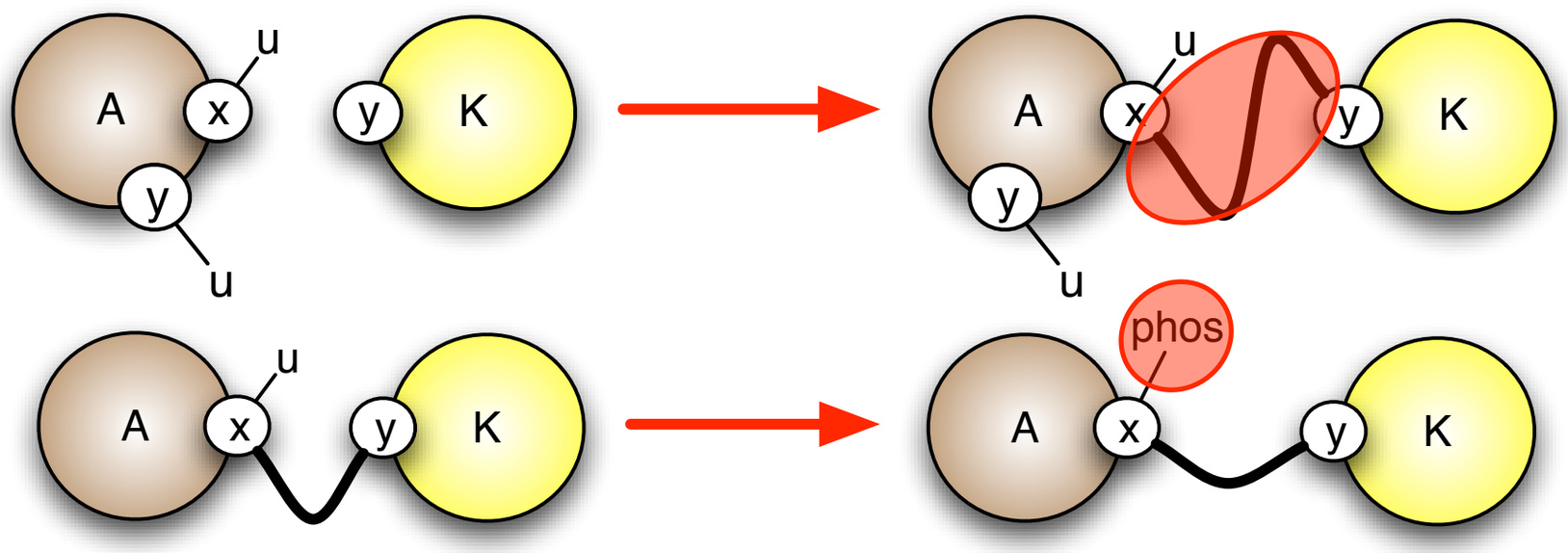
agents and rules





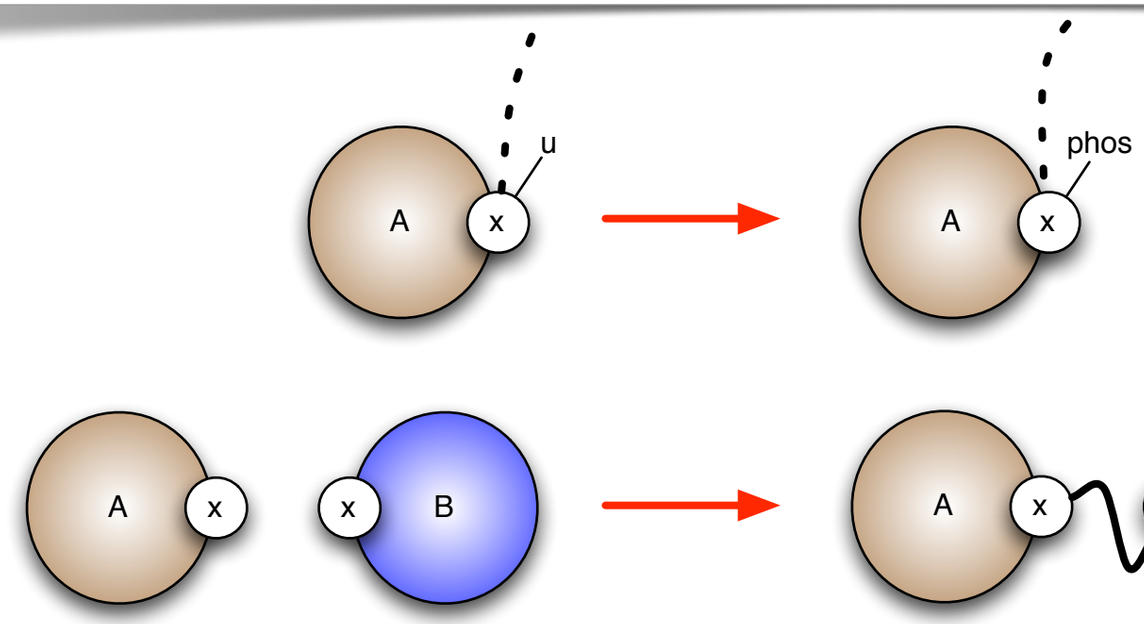
Protein A is activated by kinase K on site x.

Phosphorylated site y prevents kinase to activate x



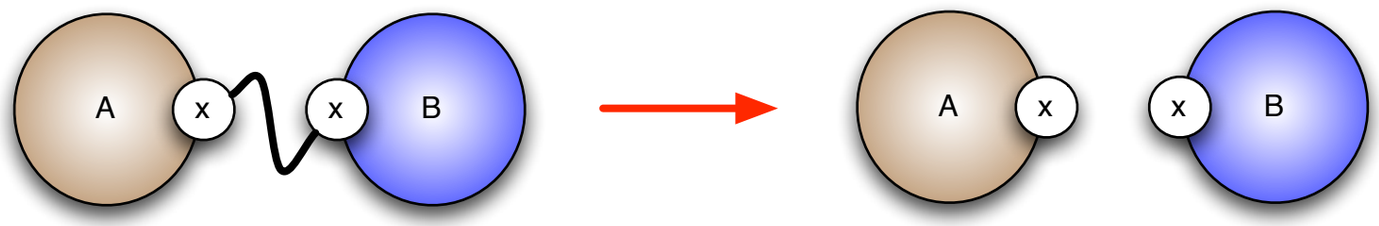
lhs=subgraph (don't care don't write)

actions

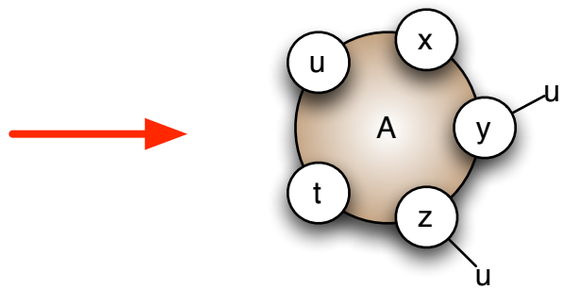


Site modification

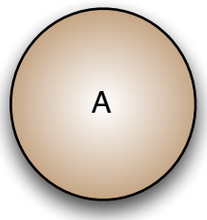
Binding



Unbinding

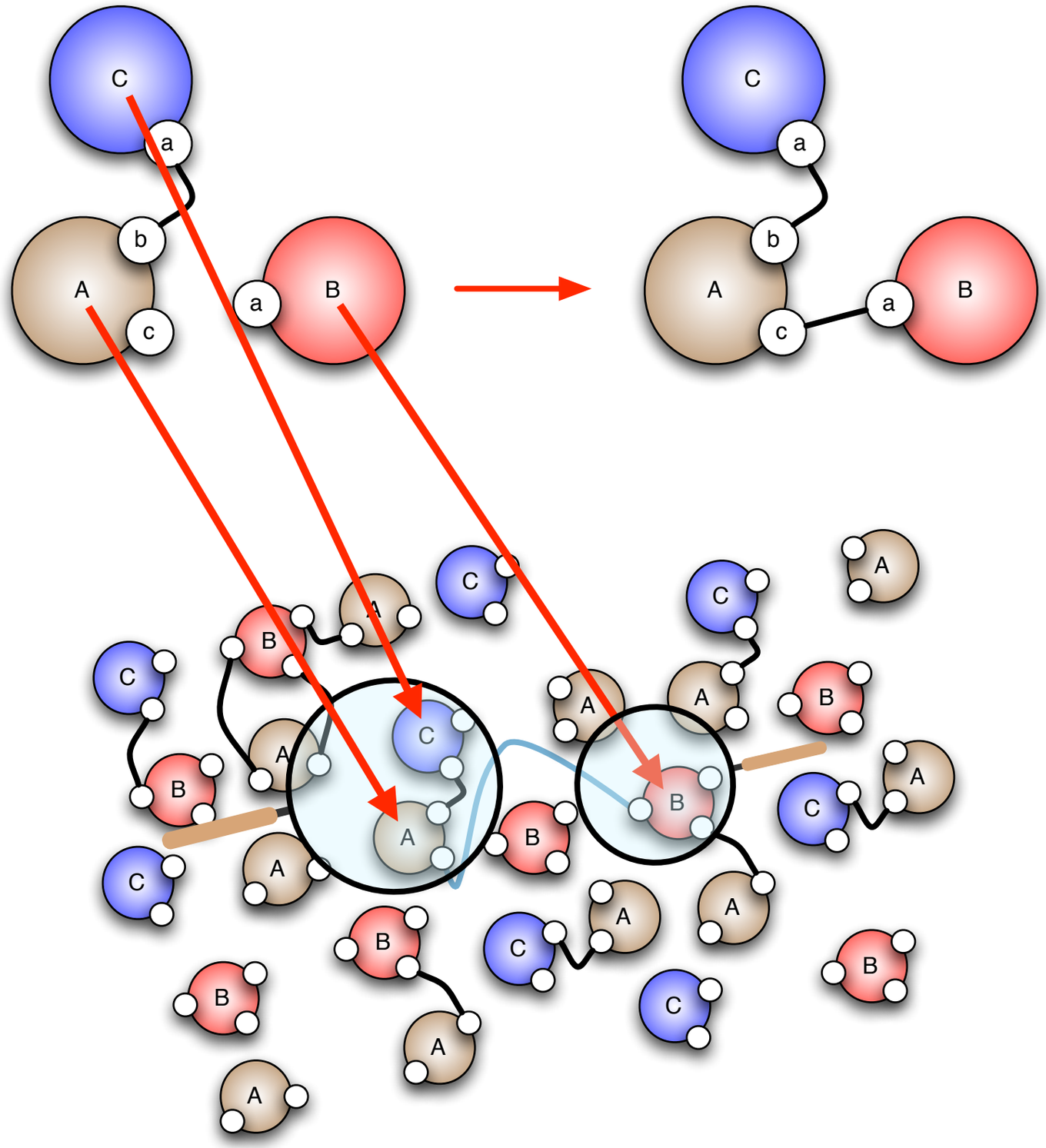


Synthesis

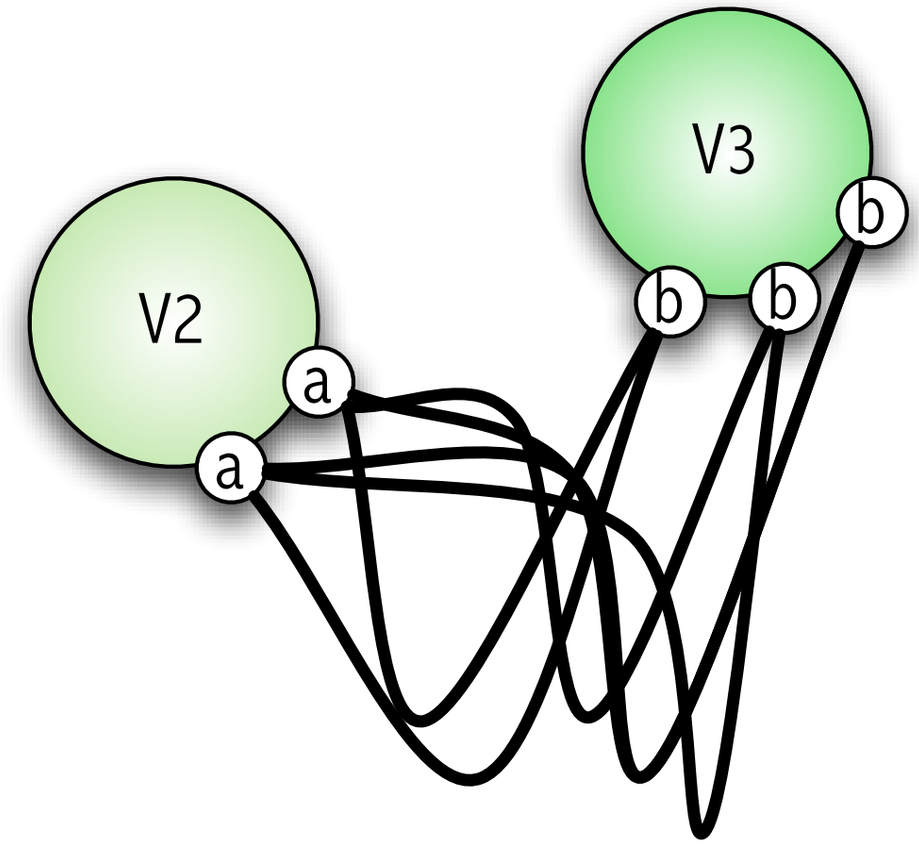


Degradation

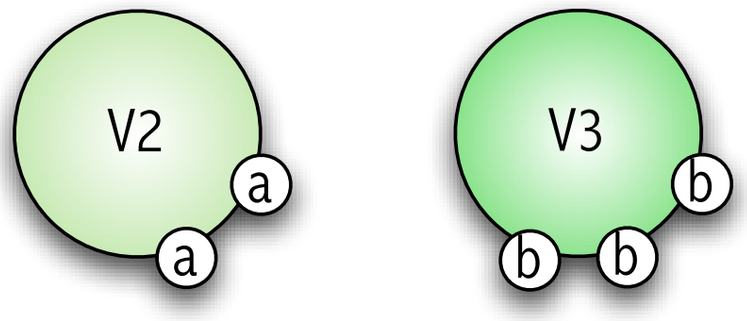
event:=rule application

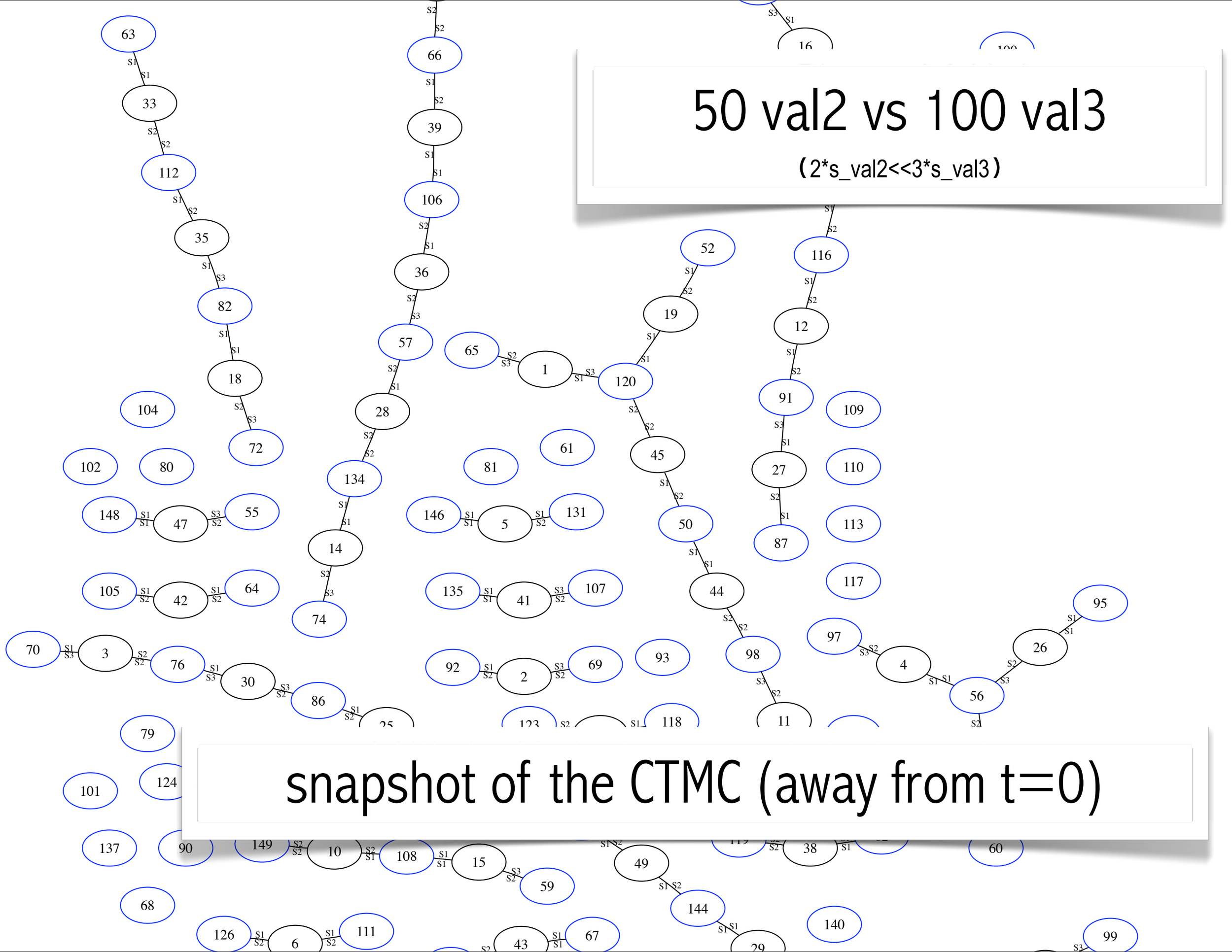


a simple mag with dim=∞



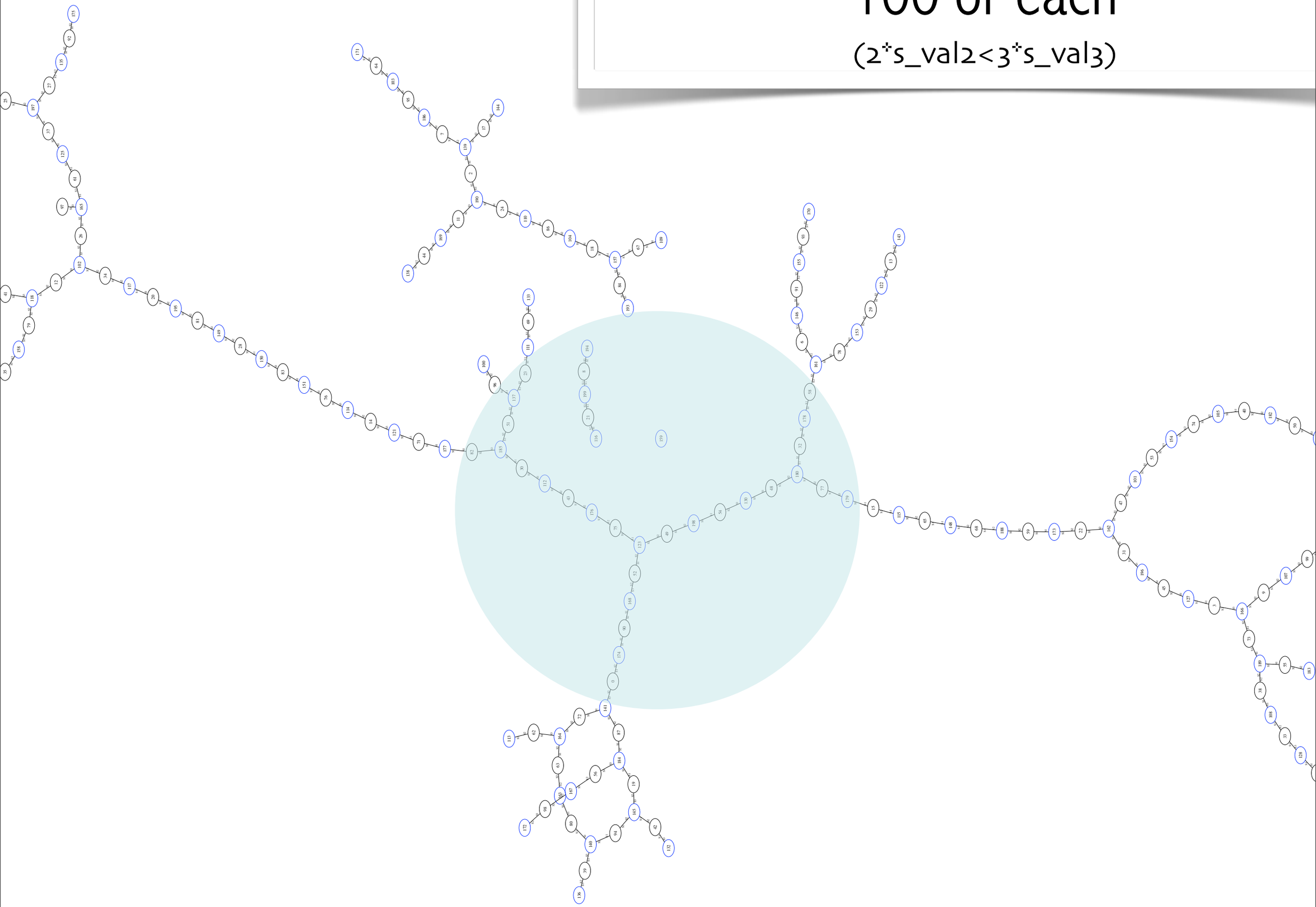
One ab binding rule with low dissociation T (viscous)




50 val2 vs 100 val3
(2*s_val2<<3*s_val3)
snapshot of the CTMC (away from t=0)

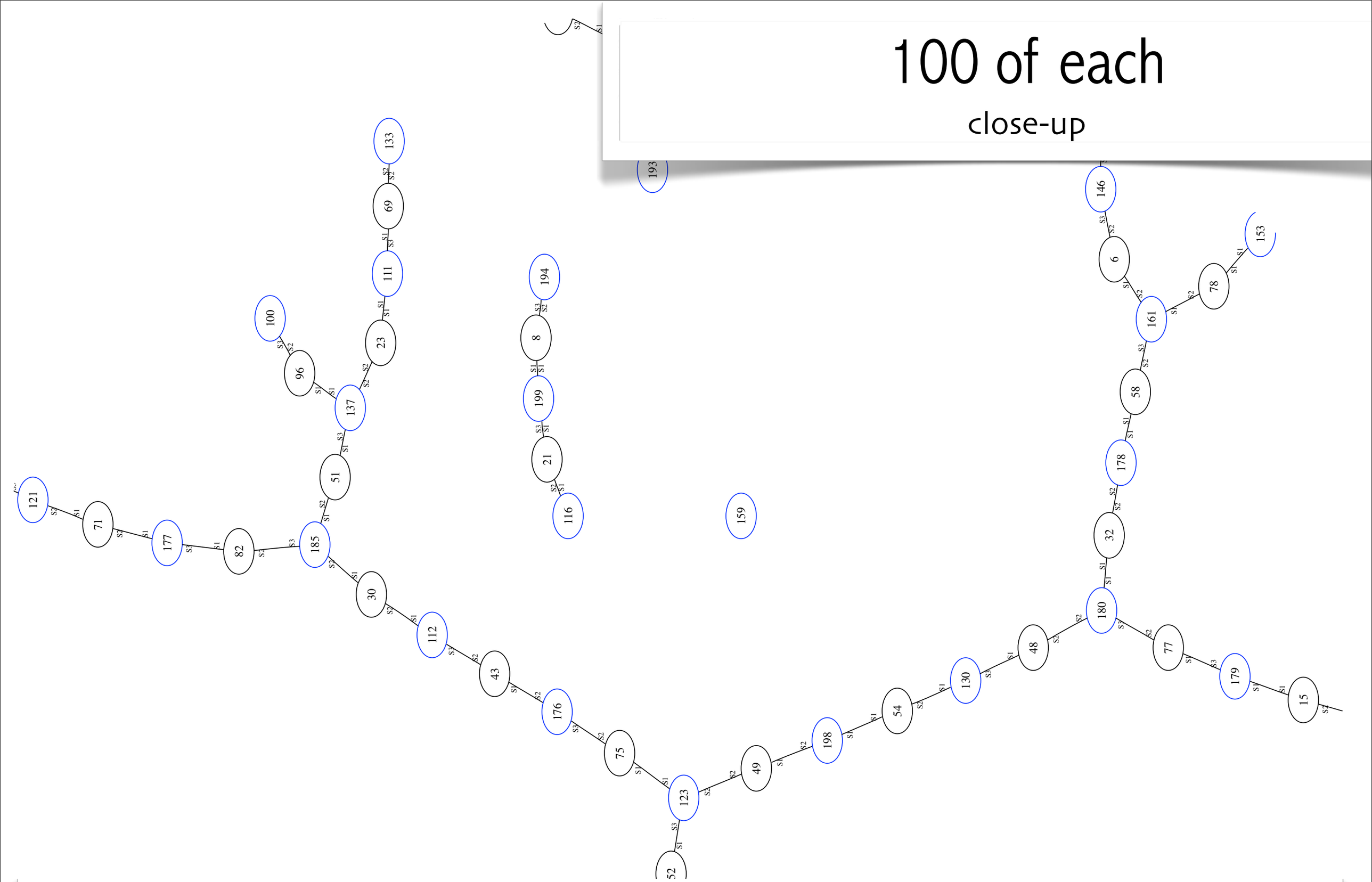
100 of each

(2*s_val2<3*s_val3)



100 of each

close-up



giant cluster! phase transition

150 val2 vs 100 val3

2*s_val2=3*s_val3

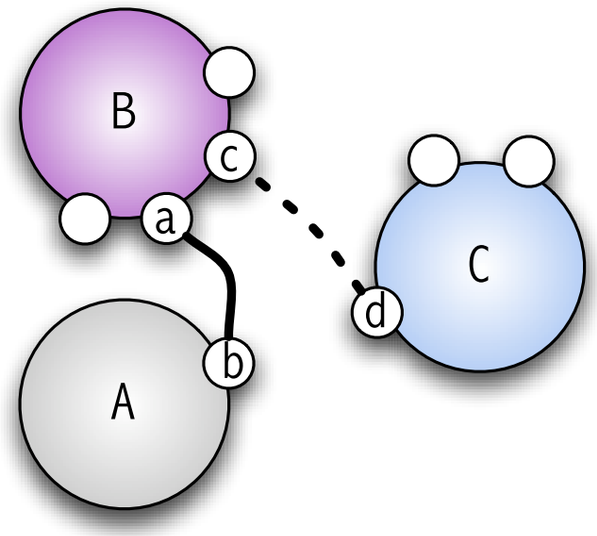
criticality happens easily and early (not just at ∞!)

sensitive to T structure = dissociation rate

sensitive to E structure := local stub correlations

liquidity index

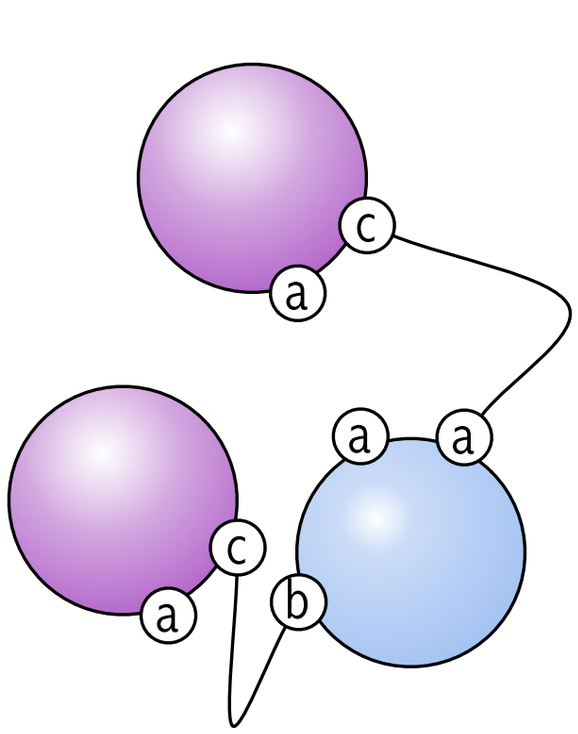
the dRG model

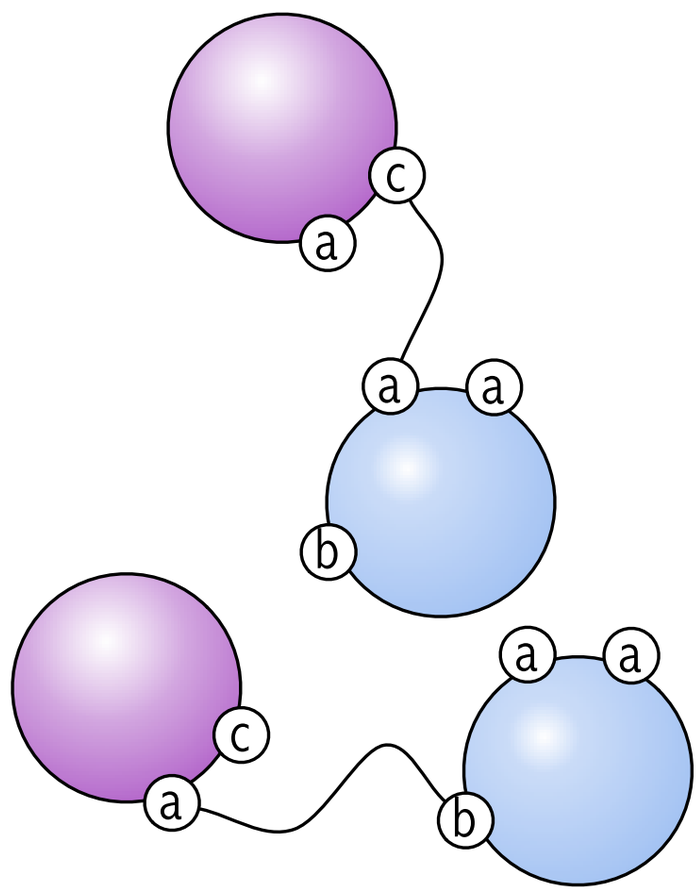


A *random graph with sites* consists of the following data:
- n the set of nodes
- K the (finite) set of *colours*
- Z the *node* random variable with values in $\mathbb{N}^K$
- for each a, $b \in K$ a *dissociation* constant $\Gamma_{ab} \in [0, \infty]$

the dRG model II

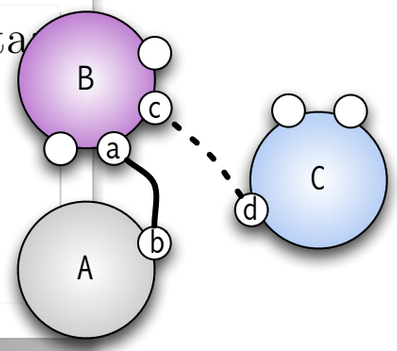
- *[binding]* two free sites x, y of respective colours a, b bind each other with a probability proportional to γ^+_{ab};
- *[unbinding]* two sites x, y of respective colours a, b, and already bound together, unbind with a probability proportional to γ^-_{ab}.





the dRG model steady state

A *random graph with sites* consists of the following data
- n the set of nodes
- K the (finite) set of *colours*
- Z the *node* random variable with values in $\mathbb{N}^K$
- for each a, $b \in K$ a *dissociation* constant $\Gamma_{ab} \in [0, \infty]$



$$n_a = n_a^f + \sum_b e_{ab}$$

$$\Gamma_{ab} e_{ab} = n_a^f n_b^f$$

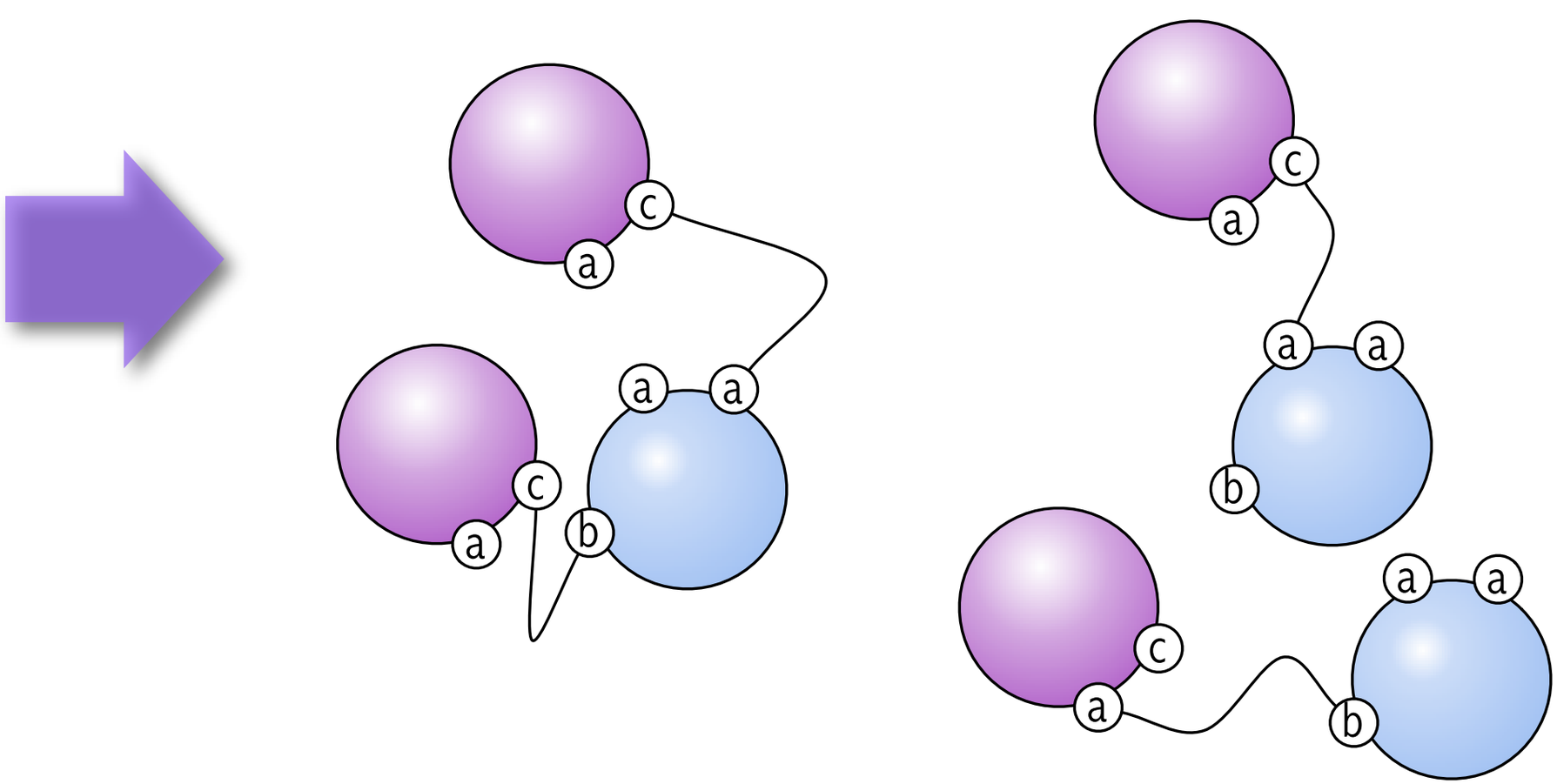
steady state equation is node independent (E structure), scale-less too!

$$\Gamma_{ab} \cdot e_{ab} = (n_a - \sum_c e_{ac})(n_b - \sum_d e_{db})$$

$$K_{ab} \epsilon_{ab} = (\langle m_a \rangle - \sum_c \epsilon_{ac})(\langle m_b \rangle - \sum_d \epsilon_{bd})$$

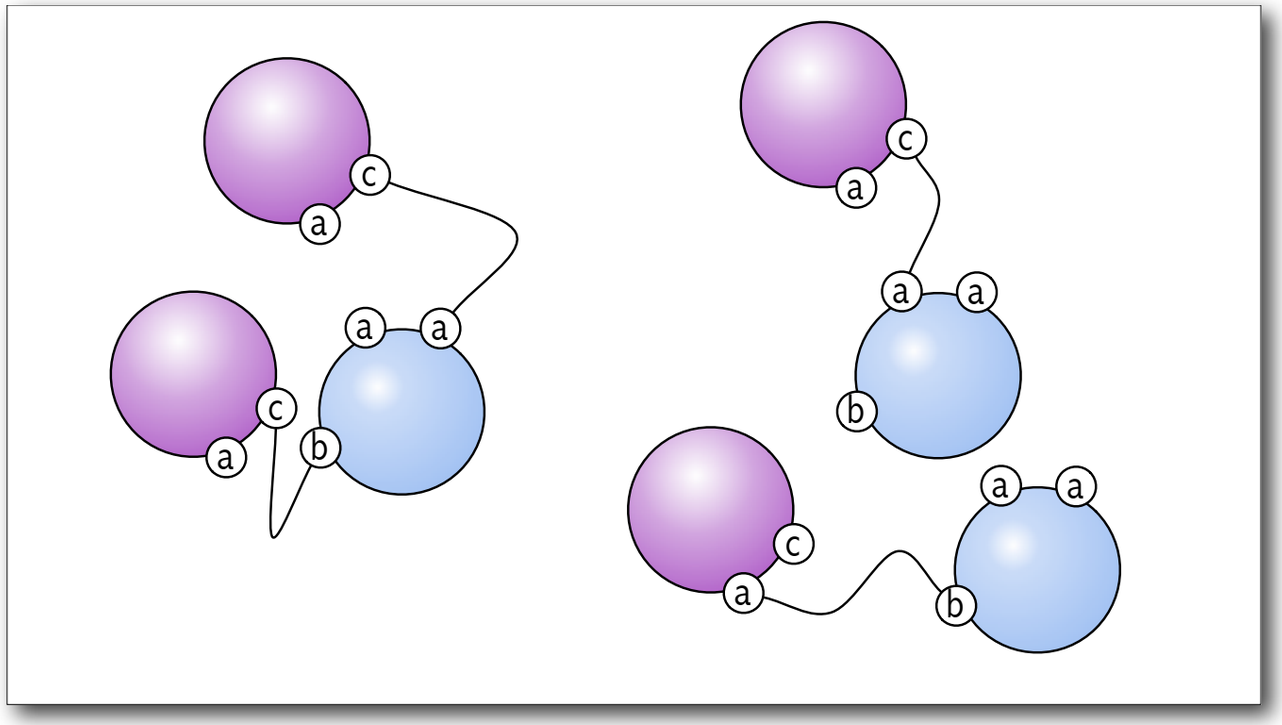
an sRG model -after Soderberg Act. Phys. Polonica (2003)

- n the set of nodes
- K the set of colours together with $*$ a special value not in K
- Z the *node* random variable with values in $\mathbb{N}^K$
- for each $a \in K$, Y_a the *edge* random variable with values in $K + \{*\}$



from dRG to sRG

limit distribution of dRG gives Y_a/T_{ab} as a function of e_{ab}



$$p(Y_a = b) := \epsilon_{ab}/\langle m_a \rangle$$

$$T_{ab} := \epsilon_{ab}/\langle m_a \rangle \langle m_b \rangle$$

size and gf-ology

$$\epsilon_{ab}/\langle m_a \rangle \cdot m_b p_m / \langle m_b \rangle$$

$$T_{ab} m_b p_m$$

$$S_p^a(z) := \sum_n p(S_p^a = n) z^n$$

$$Z(x_c; c \in K) := \sum_{m \in \mathbb{N}^K} p(Z = m) \prod_{c \in K} x_c^{m_c}$$

inductive size gf

$$S_p^a(z) - p(Y^a = *)$$

$$= z \sum_{n>0} p(S_p^a = n) z^{n-1}$$

$$= z \sum_{n>0,m} \sum_{b \in K} T_{ab} m_b p_m \, p(\sum_{c \in m-b} S_{p-1}^c = n-1) z^{n-1}$$

$$= z \sum_{b \in K} T_{ab} \sum_m m_b p_m \, (\sum_{n>0} p(\sum_{c \in m-b} S_{p-1}^c = n-1) z^{n-1})$$

$$= z \sum_{b \in K} T_{ab} \sum_m m_b p_m \prod_{c \in m-b} S_{p-1}^c(z)$$

$$= z \sum_{b \in K} T_{ab} \partial_b Z(S_{p-1}^c(z); c \in K)$$

liquidity index

$$\psi_a(x_b; b \in K) := p(Y^a = *) + \sum_{b \in K} T_{ab} \partial_b Z(x_b; b \in K)$$

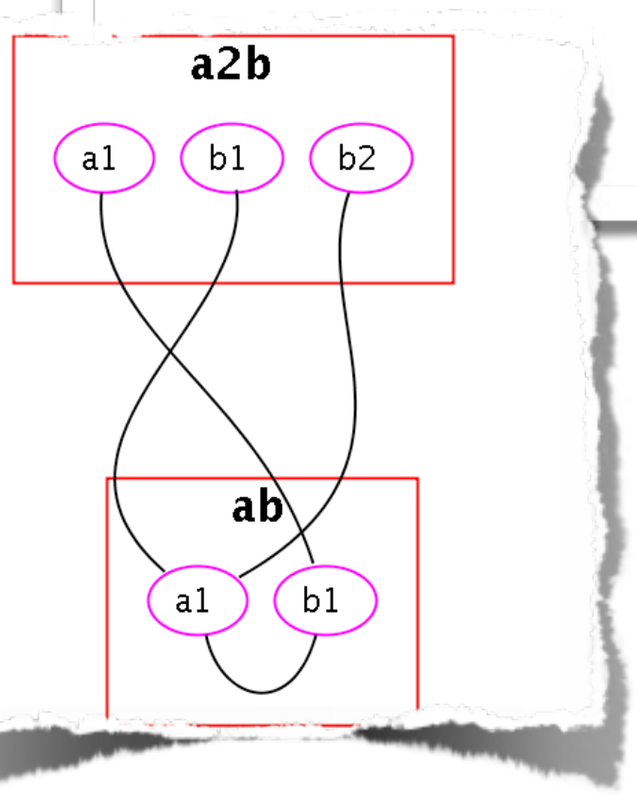
$$\partial_c \psi_a(x_b; b \in K) = \sum_{b \in K} T_{ab} \partial_c \partial_b Z(x_b; b \in K)$$

$$\partial_c \psi_a(\mathbf{1}) = \sum_{b \in K} T_{ab} E_{bc} = (TE)_{ac}$$

liquidity index=highest modulus eigenvalue of TE

$$\lambda_1(TE)$$

bicolor case



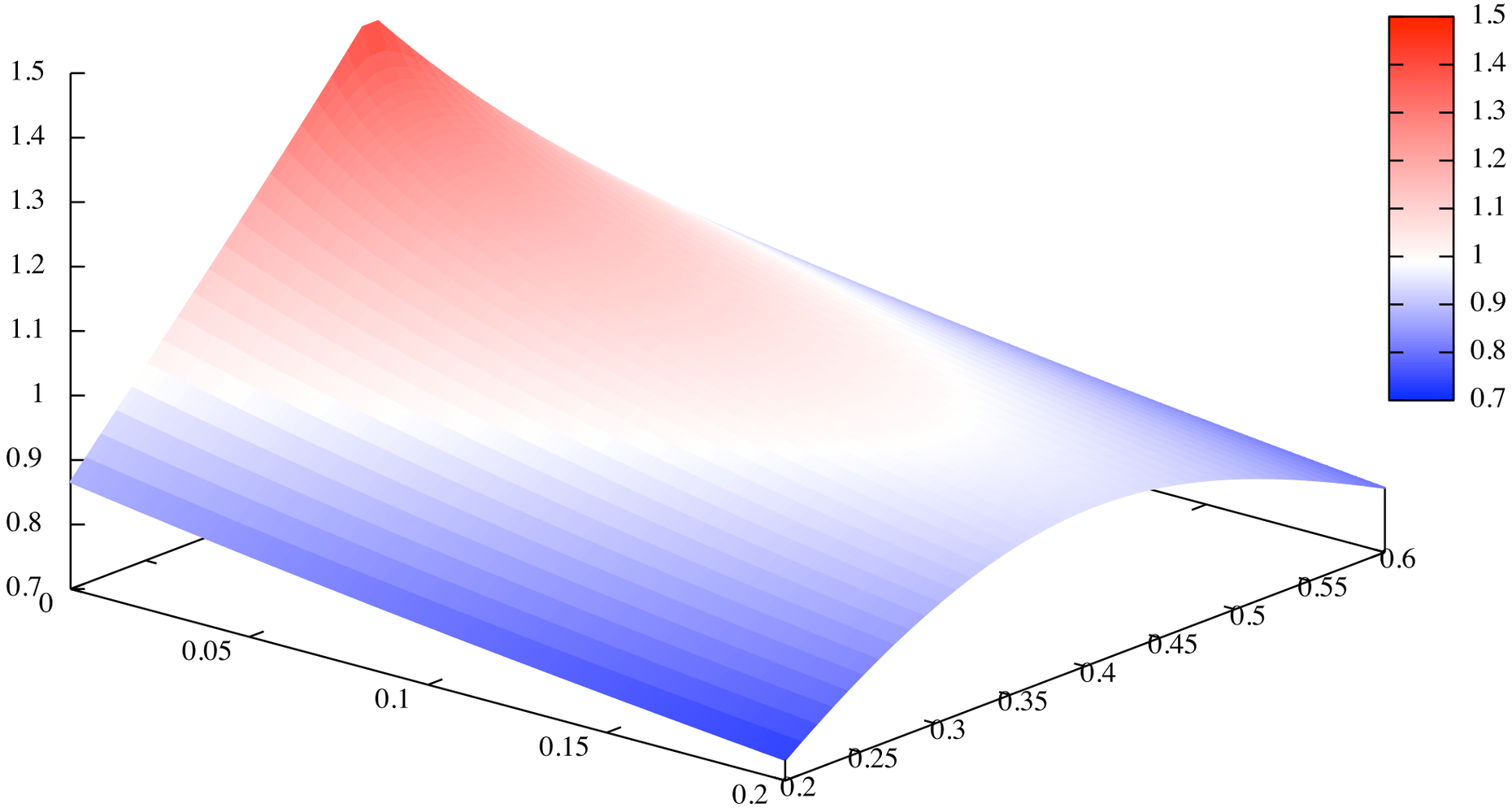
$$N := \frac{\langle m_a m_b \rangle + \sqrt{\langle m_a(m_a-1)\rangle \langle m_b(m_b-1)\rangle}}{\langle m_a \rangle \langle m_b \rangle}$$

$$\epsilon := \frac{\langle m_a \rangle + \langle m_b \rangle + K - \sqrt{(\langle m_a \rangle + \langle m_b \rangle + K)^2 - 4\langle m_a \rangle \langle m_b \rangle}}{2}$$

2a vs. 3b

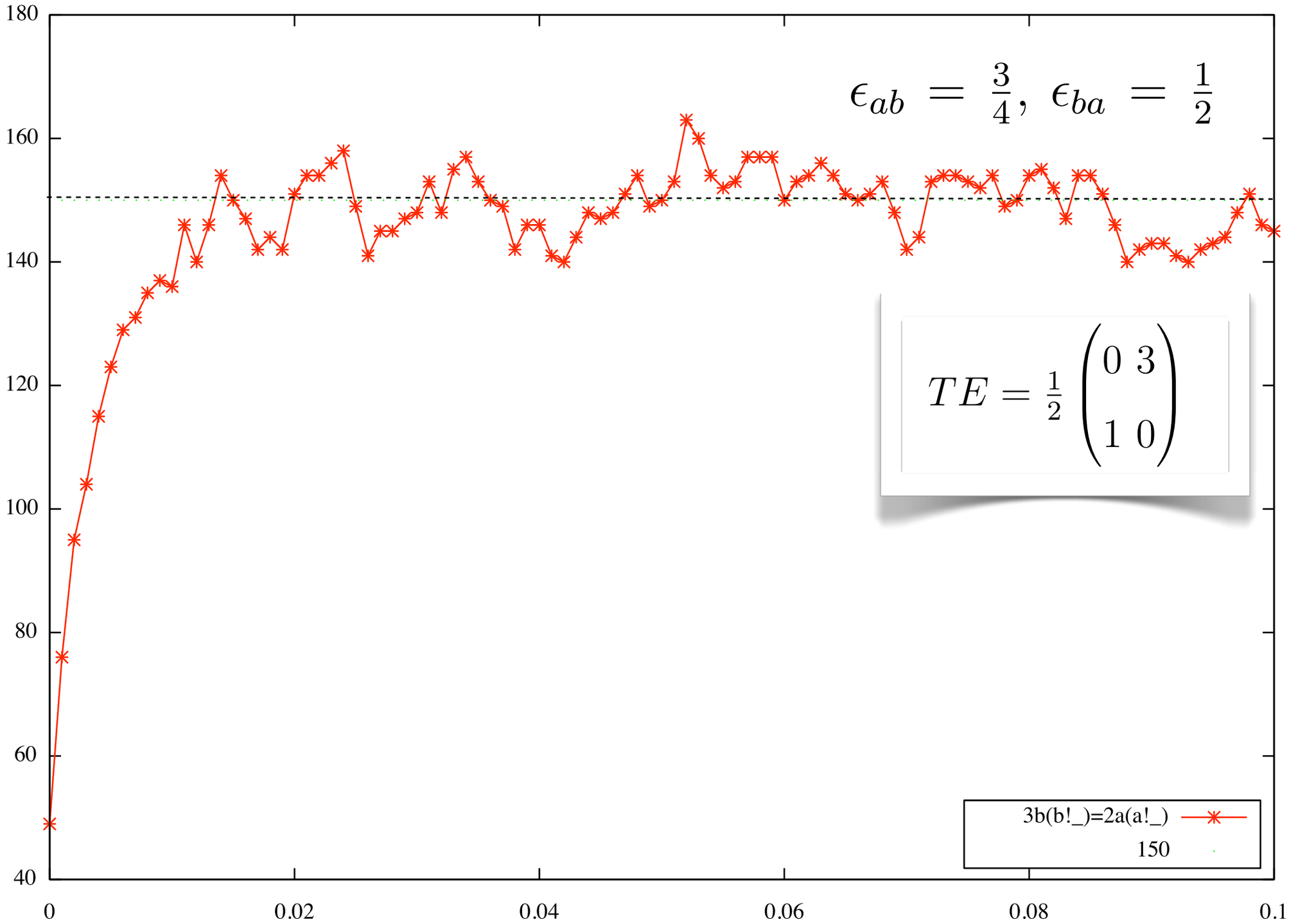
p = proportion of 3b

$$\lambda(p, K) := \frac{2 + p + K - \sqrt{(2 + p + K)^2 - 24p(1 - p)}}{2\sqrt{3p(1 - p)}}$$

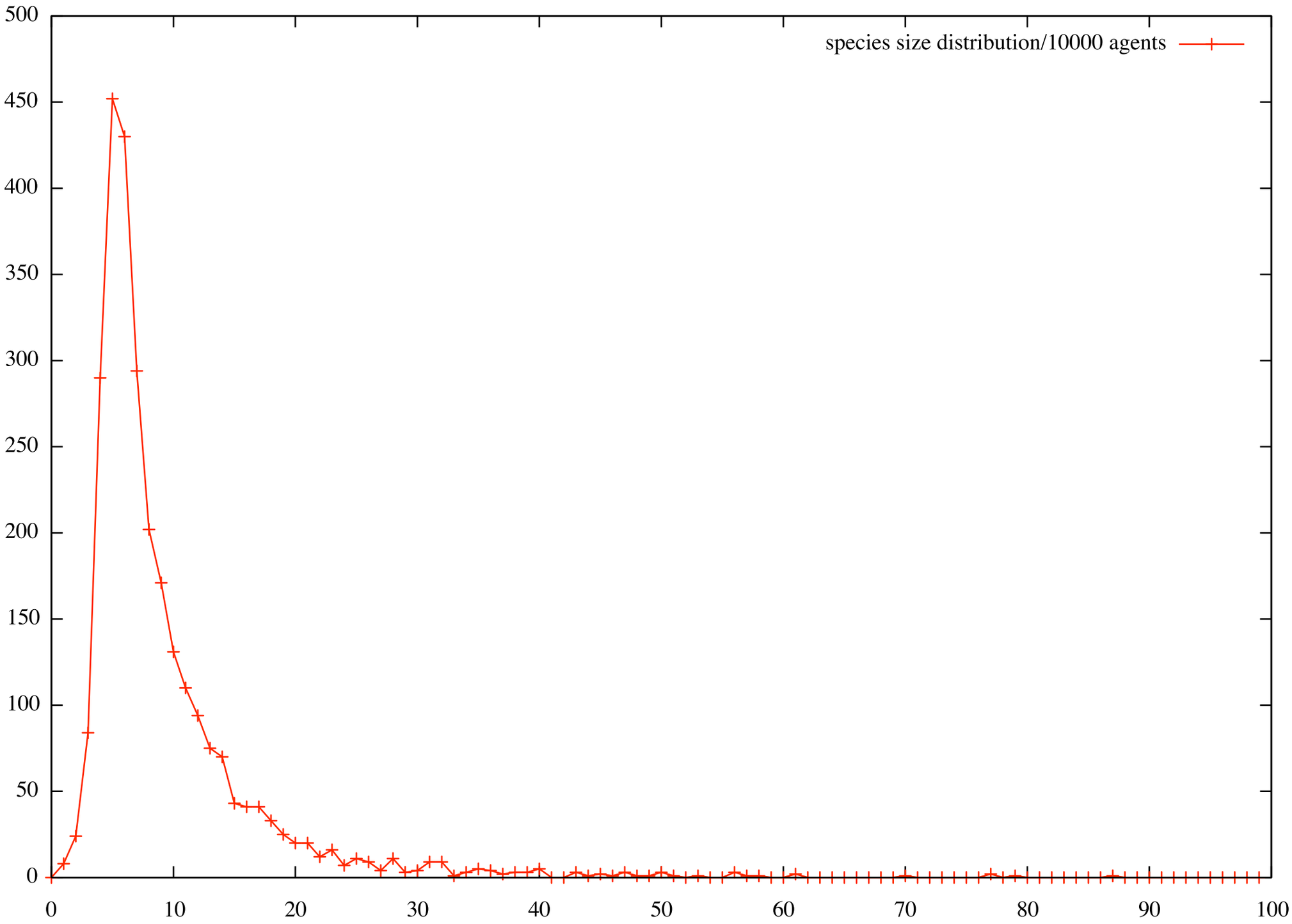


simulations

n=200, $K_{ab=1/4,}$


$\epsilon_{ab} = \frac{3}{4}, \epsilon_{ba} = \frac{1}{2}$
$TE = \frac{1}{2}\begin{pmatrix} 0 & 3 \\ 1 & 0 \end{pmatrix}$
180
160
140
120
100
80
60
40
0
0.02
0.04
0.06
0.08
0.1
3b(b!_)=2a(a!_)
150

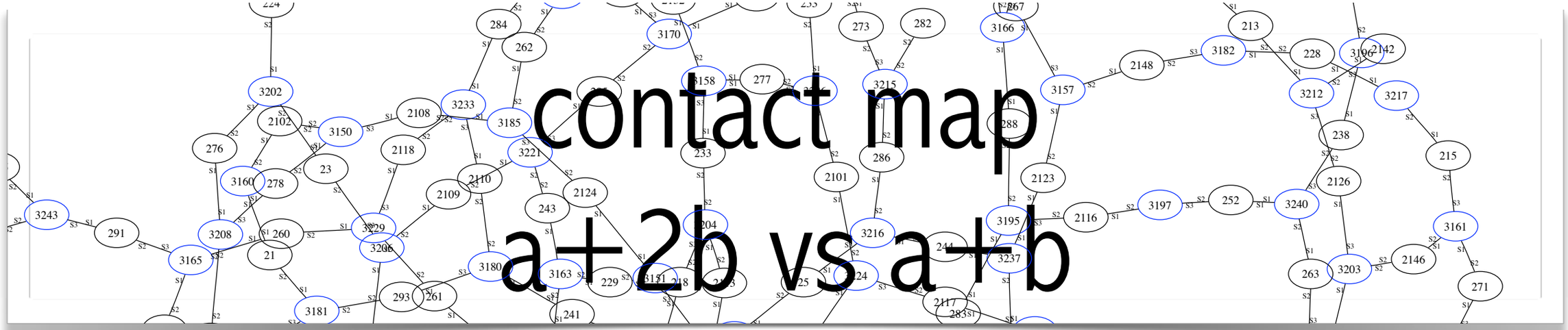
n=1000 size distribution



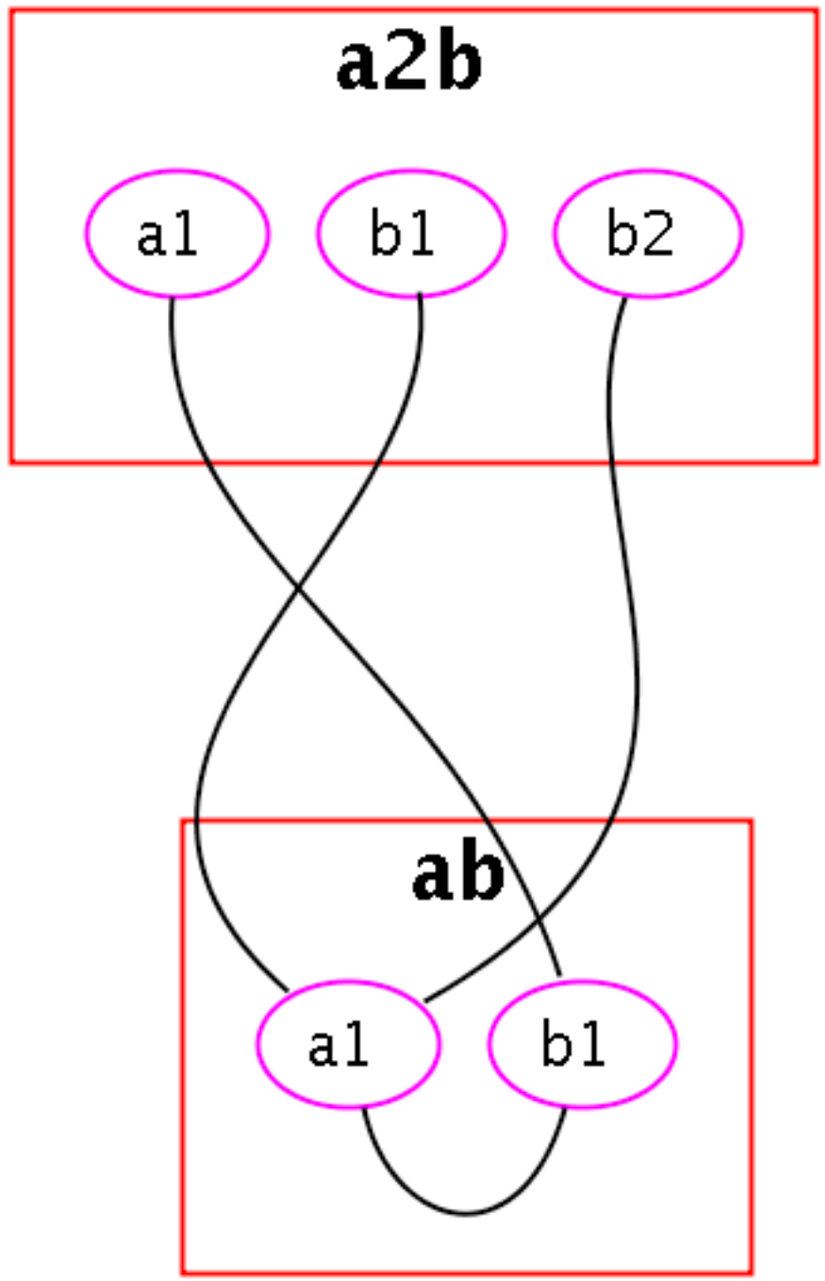
n=0.2 10^6 -max size

```
simplx --sim 2a_3b.ka --time 0.03 --no-measure --no-maps --output-final-state --rescale 1000

'a1-b1' 3b(b1),2a(a1) <-> 3b(b1!1),2a(a1!1) @ 1.0,50
'a1-b2' 3b(b2),2a(a1) <-> 3b(b2!1),2a(a1!1) @ 1.0,50
'a1-b3' 3b(b3),2a(a1) <-> 3b(b3!1),2a(a1!1) @ 1.0,50
'a2-b1' 3b(b1),2a(a2) <-> 3b(b1!1),2a(a2!1) @ 1.0,50
'a2-b2' 3b(b2),2a(a2) <-> 3b(b2!1),2a(a2!1) @ 1.0,50
'a2-b3' 3b(b3),2a(a2) <-> 3b(b3!1),2a(a2!1) @ 1.0,50

%init: 100 * (2a(a1,a2))
%init: 100 * (3b(b3,b1,b2))

%obs: 2a(a1!_)
%obs: 3b(b1!_)
%obs: 3b(b2!_)
%obs: 3b(b3!_)
%obs: 2a(a2!_)
```

the above takes ($n = 0.2\ 10^6$):
- Initialization: 27.5 sec. CPU
- Simulation: 213.5 sec. CPU the final state is 2MB -takes ages to write in a file!- max size is 168, of relative size $< 0.1\%$.

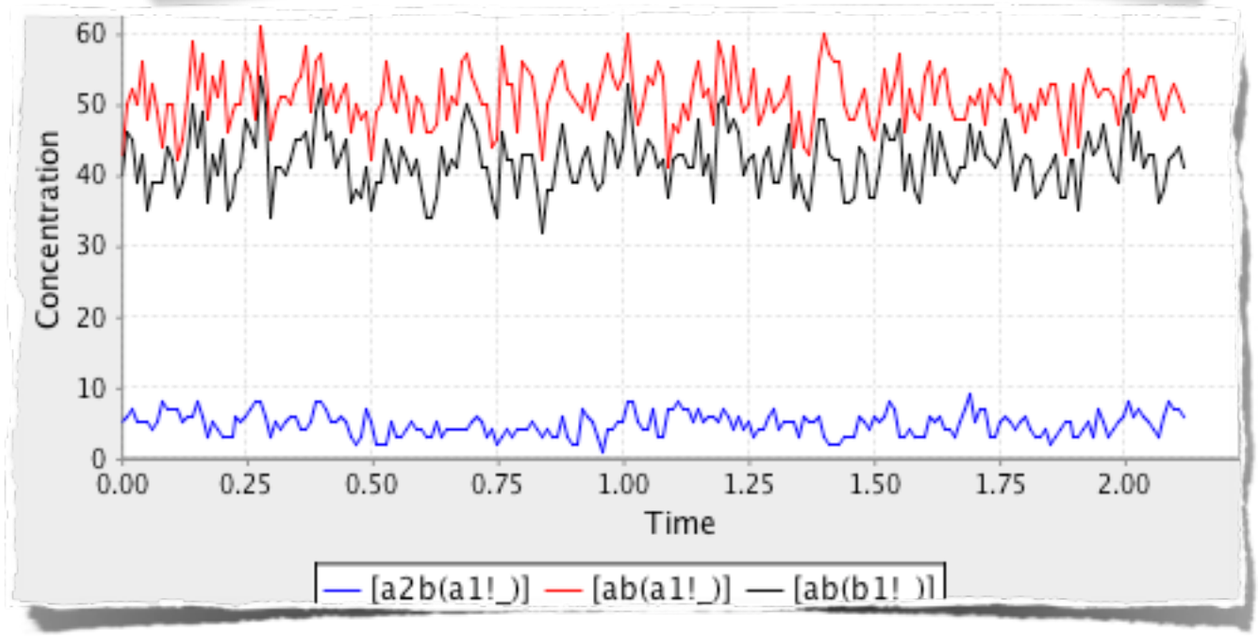


contact map
a+2b vs a+b



multisets are enumerated

a+2b allows branching




```
Gamma=50, Z=90*ab+10*a2b
```

conclusions

- TE have non monotonic effects in the case of conflicting contact maps
- how good is liquidity a proxy for the size distribution
- what is the influence of other forces (extend to view-local systems)
- compute liquidity of yeast!
- where is information (more later)