Effects of Spatial Diffusion on a Model for Prebiotic Evolution

Ben Intoy

J. Woods Halley

Aaron Wynveen

U. of MN, Twin Cities
School of Physics and Astronomy

NASA grant:
NNX14AQ05G

A. Wynveen, I. Fedorov, and J. W. Halley
Physical Review E *89, 022725 (2014)*

B.F. Intoy, A. Wynveen, and J. W. Halley
Physical Review E *94, 042424 (2016)*

Acknowledgements:
- Computational resources of:
  - Minnesota Supercomputing Institute.
  - Open Science Grid.
  - UMN School of Physics and Astronomy Condor Cluster.
- Simon Schneider for discussions

OSG All Hands Meeting March 7, 2017

Motivations

- A protein first origin of life model might resolve Eigen's paradox (the low probability of randomly constructing a starter "naked gene").
- Assume initiating event is the formation of a network of interacting molecules assumed to be polymers (but not necessarily proteins).
- No genome, assume it comes much later.
- Unlike previous similar models, we assume here that a necessary condition for a prebiotic chemical system is that it be a stationary state out of chemical equilibrium.

Kauffman-like Binary Polymer Model

Network Formation

- Ligation and Scission: $010 + 10 \overset{11}{\rightleftharpoons} 01010$
- Given a maximum polymer length value (L_{max}) go through each possible reaction of the form: $$\mathrm{A} + \mathrm{B} \overset{\mathrm{C}}{\rightleftharpoons} \mathrm{AB}$$ and include it in the network with probability p.

Dynamics

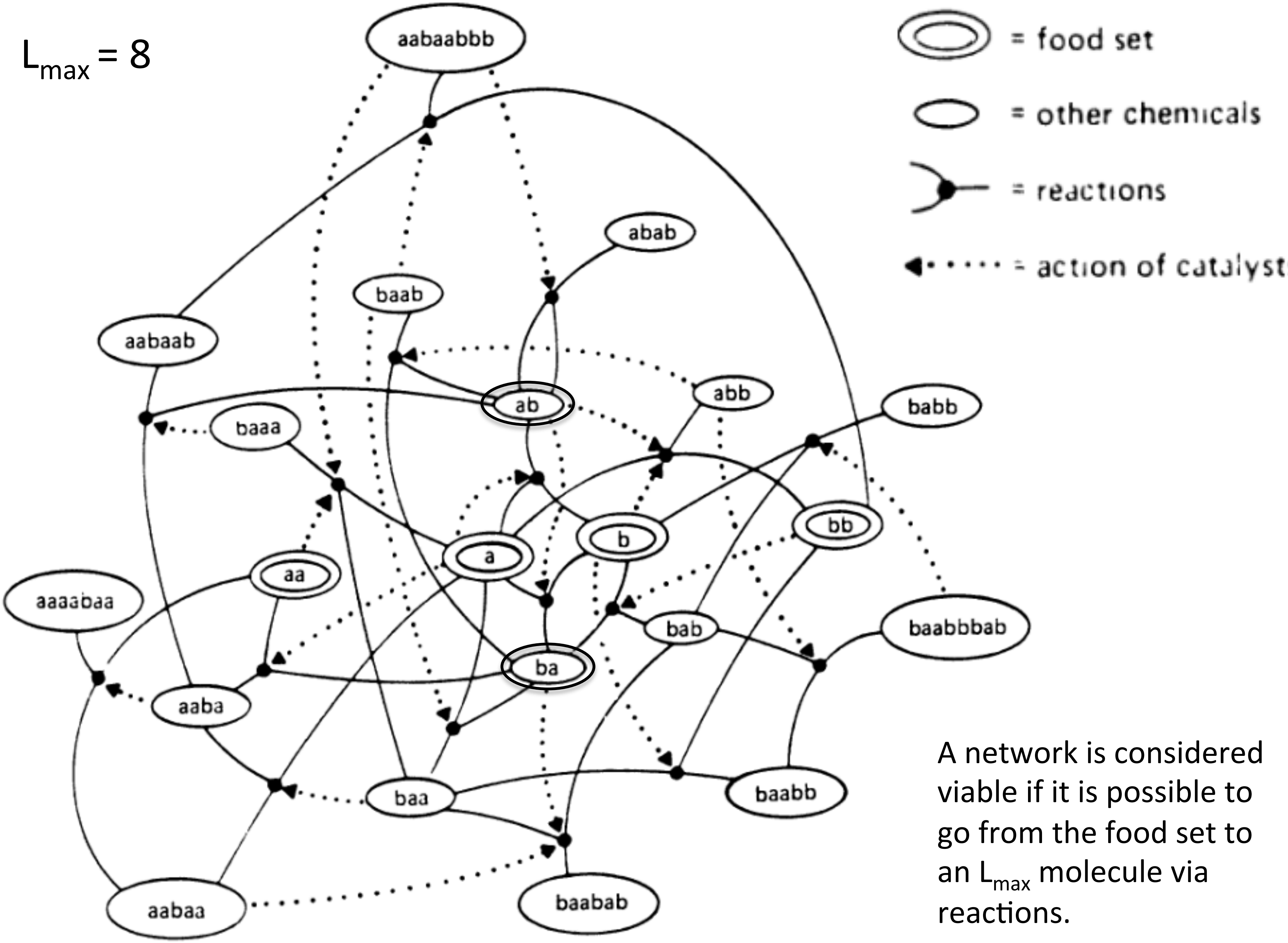
- Combine with reaction rates to generate artificial chemistry, then stochastically simulate the following master equation:

$$\frac{\mathrm{d}n_l}{\mathrm{d}t} = \sum_{l',m,e} \left[v_{l,l',m,e}(-k_d n_l n_{l'} n_e + k_d^{-1} n_m n_e) + v_{m,l',l,e}(+k_d n_m n_{l'} n_e - k_d^{-1} n_l n_e) \right]$$

- Parameters in the model: p, L_{max}, number of food particles, and maximum number of particles.

Kauffman, *The Origins of Order* (Ch. 7)

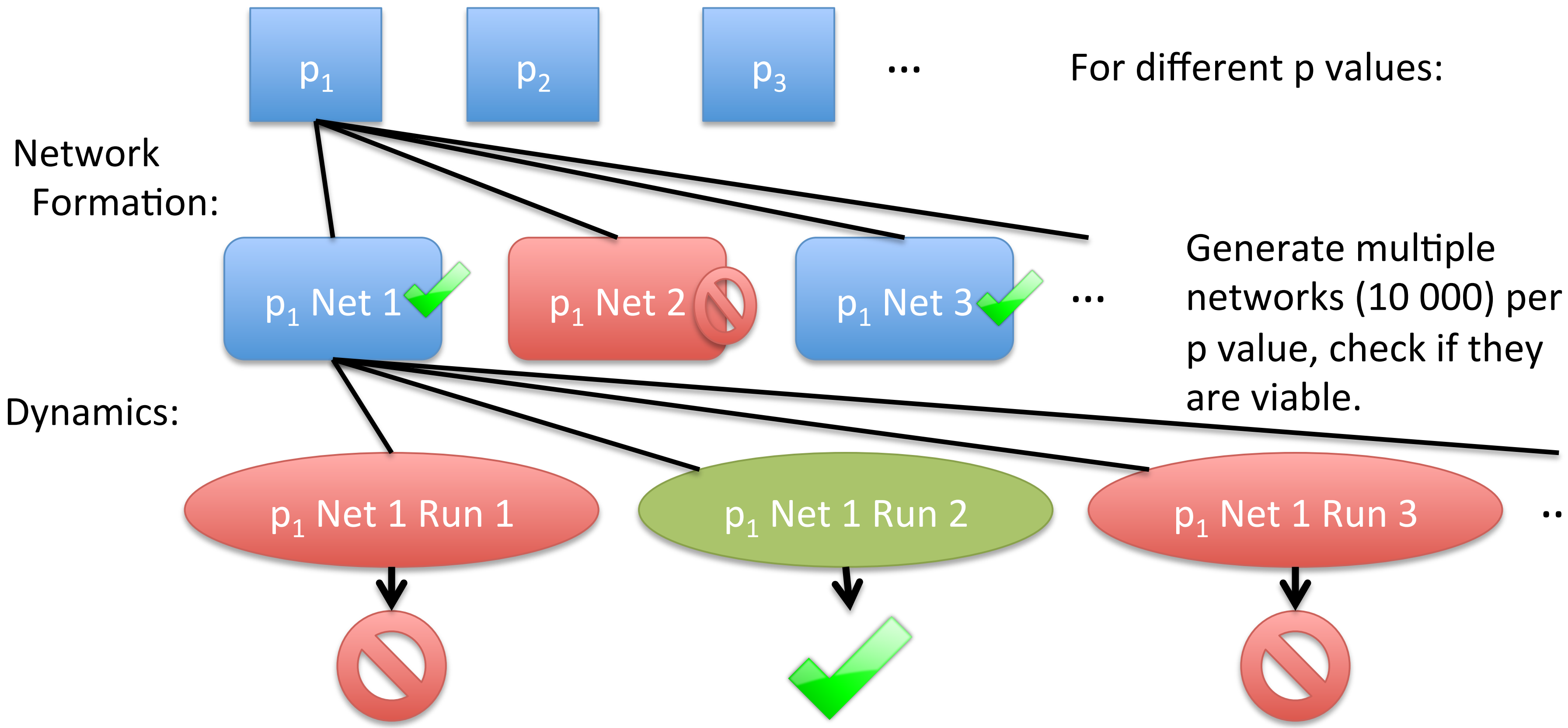
$L_{max} = 8$



A network is considered viable if it is possible to go from the food set to an L_{max} molecule via reactions.

"The Origins of Order" – Stuart A. Kauffman

General Structure



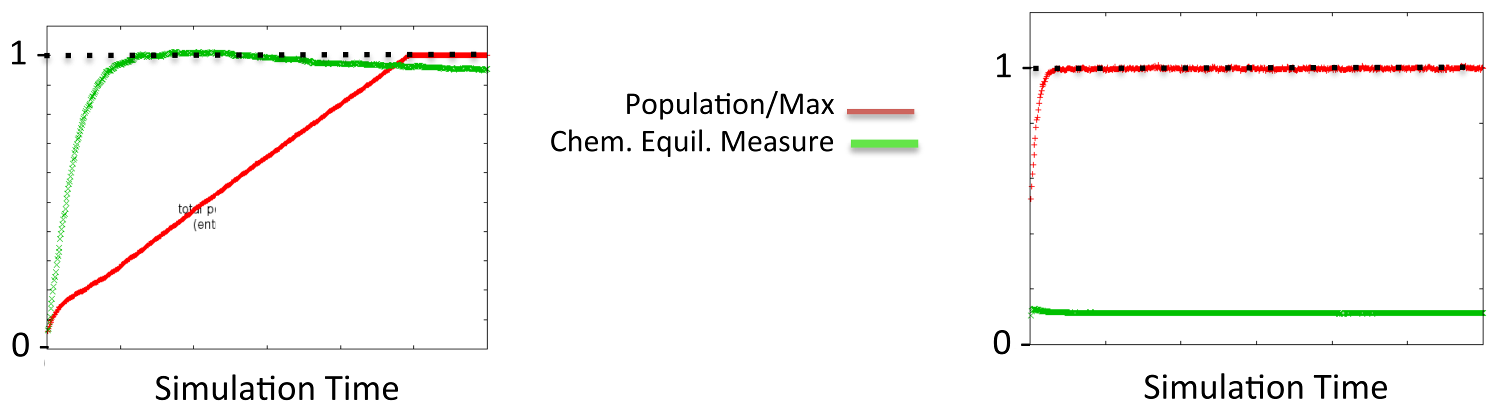
- Do multiple dynamic simulations (50) with random initial conditions using a given viable network combined with reaction rates until a steady state is reached.
- Count the number of lifelike steady states by checking if the system is out of equilibrium.
- We now have a measurement for the probability of forming a lifelike state for a value of p_i, $P_{lifelike}(p_i)$.

How Close to Chemical Equilibrium? Use Entropy

- Coarse-grain by polymer length, $\{N_L\}$.
- Given a macrostate $\{N_L\}$ the number of possible configurations is:
$$W = \prod_{L}^{L_{\max}} \frac{(N_L + 2^L - 1)!}{N_L!(2^L - 1)!}$$
- Entropy is defined as $S = k_B \operatorname{Log} W$.
- Chemical Equilibrium is reached when entropy is maximized (S_{eq}), with the constraint that there are N total molecules.
- Simulate until steady state and consider it lifelike if the entropy is less than αS_{eq}.

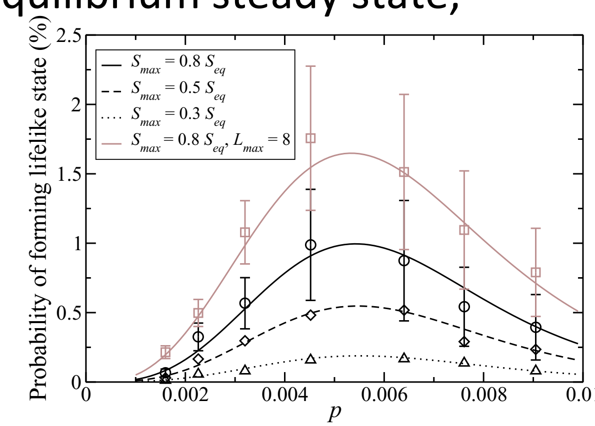
Where Kauffman and Our Group Differ

- Kauffman saw population growth with increasing p.
- System growing, but might be in chemical equilibrium.



- Same p value and artificial chemistry, two different runs. One reaches chemical equilibrium the other gets kinetically trapped in a non-equilibrium steady state, which we postulate to be a necessary condition for life.

- The non-equilibrium constraint reduces the probability of lifelike systems at large p, giving a maximum probability at a small value of p.



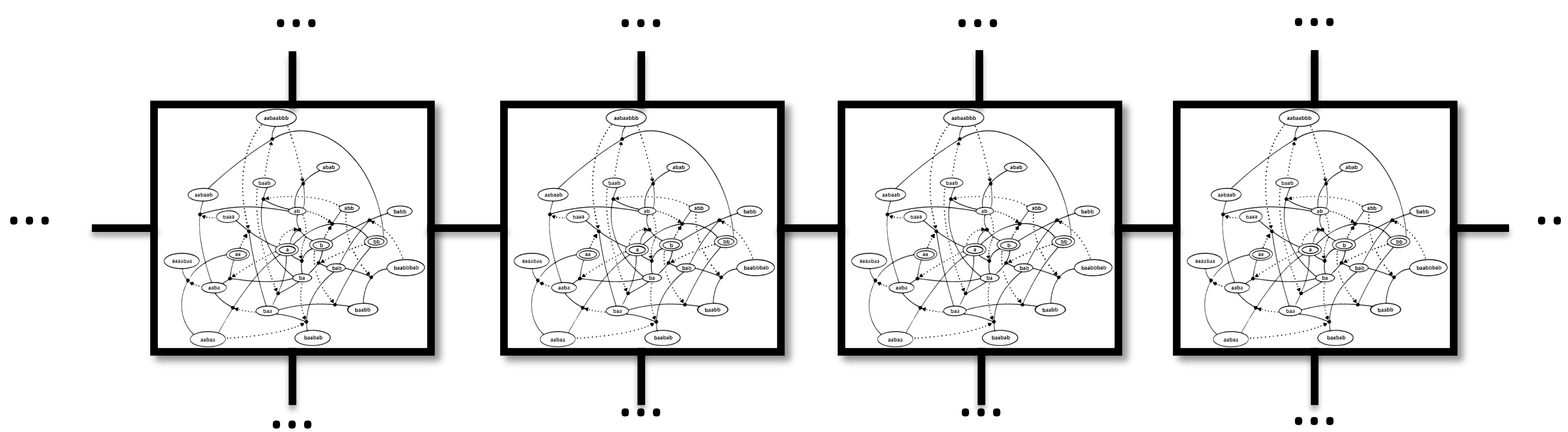
A. Wynveen, I. Fedorov, and J. W. Halley
Physical Review E *89, 022725 (2014)*

Extension to Include Diffusion Through Space

- How might spatial structure affect prebiotic evolution?
- Motivations:
  - Can the non-equilibrium states of the model without diffusion survive interaction with the environment through diffusion?
  - Are there collective effects which might suggest the beginnings of multicelluarity?
  - Space allows isolation (if at low diffusion).

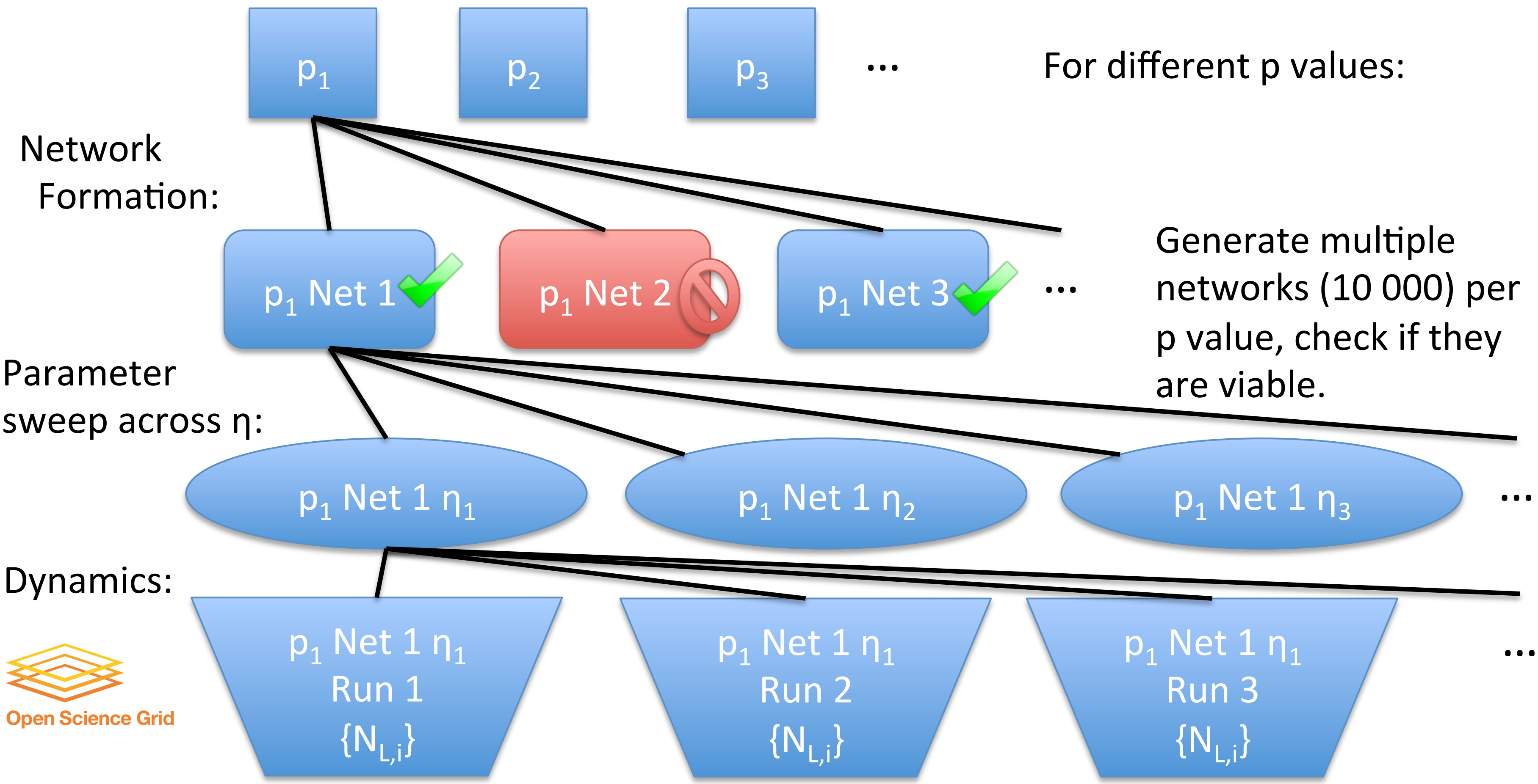
Spatial Extension

- We study M=64 sites arranged as an 8 x 8 2D periodic lattice.



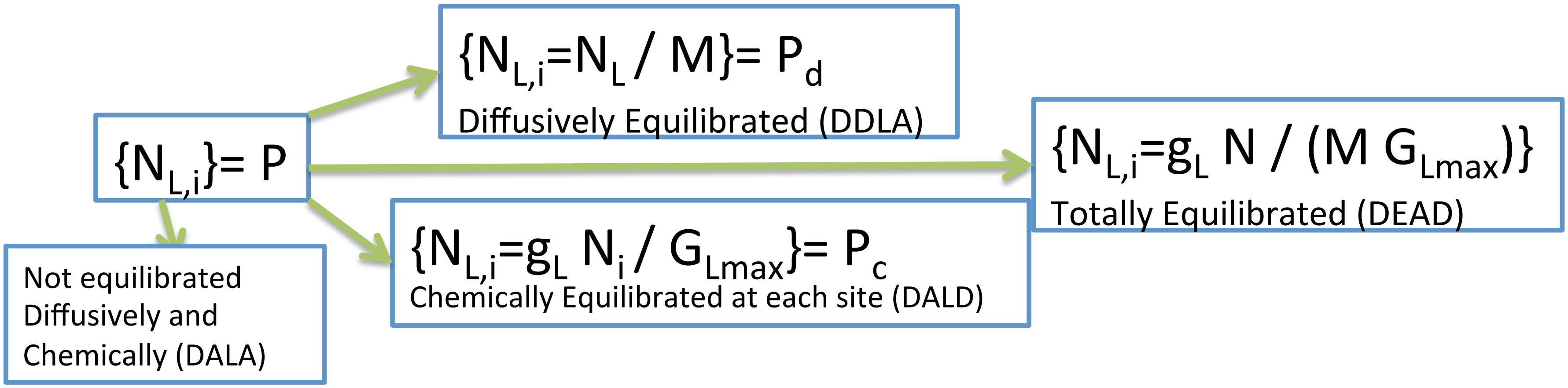
- Molecules are allowed to diffuse from site to site at a rate parameterized by η.
- Due to computational limitations we set L_{max} = 6.

Simulation General Structure

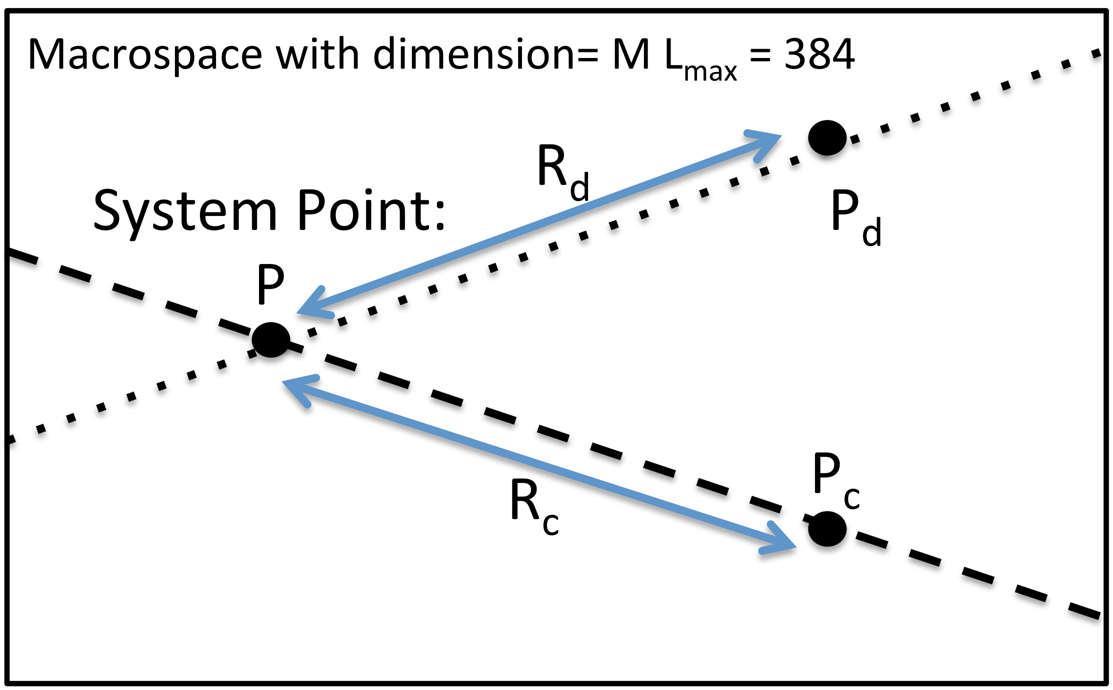


- Do multiple dynamic simulations with random initial conditions using a given viable network generated by parameter p combined with reaction rates and diffusive value η.
- A steady state is then reached with polymer length and spatial distribution $\{N_{L,i}\}$.
- Analyze the $\{N_{L,i}\}$'s to determine whether the run was lifelike or not.

Partial and Complete Equilibration



Distances From Partial Equilibria:



$$R_d = \sqrt{\sum_{L,i} (N_{L,i} - N_L/M)^2}.$$

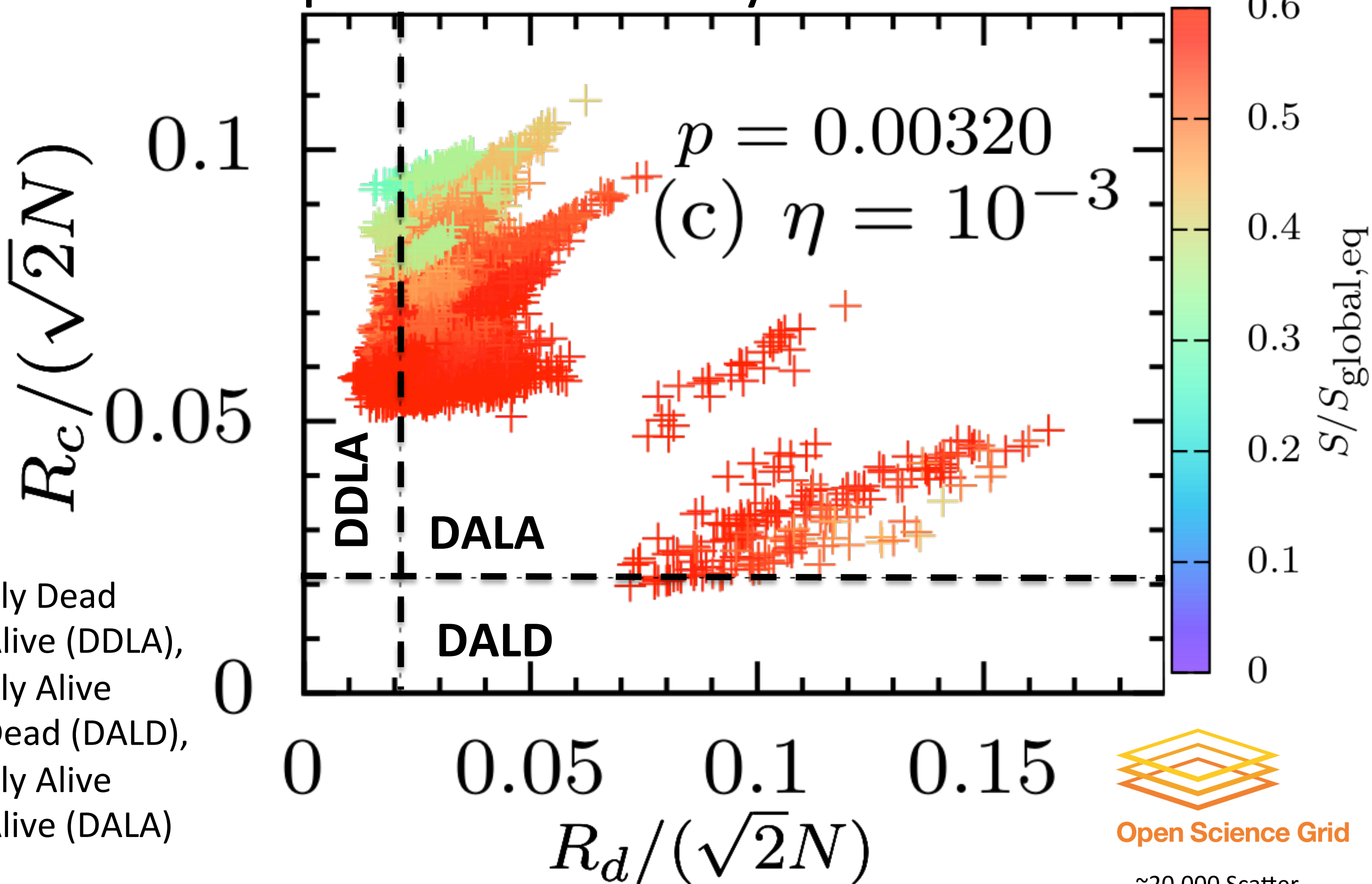
$$R_c = \sqrt{\sum_{L,i} (N_{L,i} - g_L N_i / G_{l_{\max}})^2} \ ,$$

. . . Hyperplane with fixed N_L and dimension=(M-1) L_{max}=378.

– – Hyperplane with fixed N_i and dimension=M (L_{max}-1)=320.

B.F. Intoy, A. Wynveen, and J. W. Halley
Physical Review E *94, 042424 (2016)*

Example of Results R_c and R_d in Simulated non-equilibrium Steady States



Diffusively Dead Locally Alive (DDLA), Diffusively Alive Locally Dead (DALD), Diffusively Alive Locally Alive (DALA)

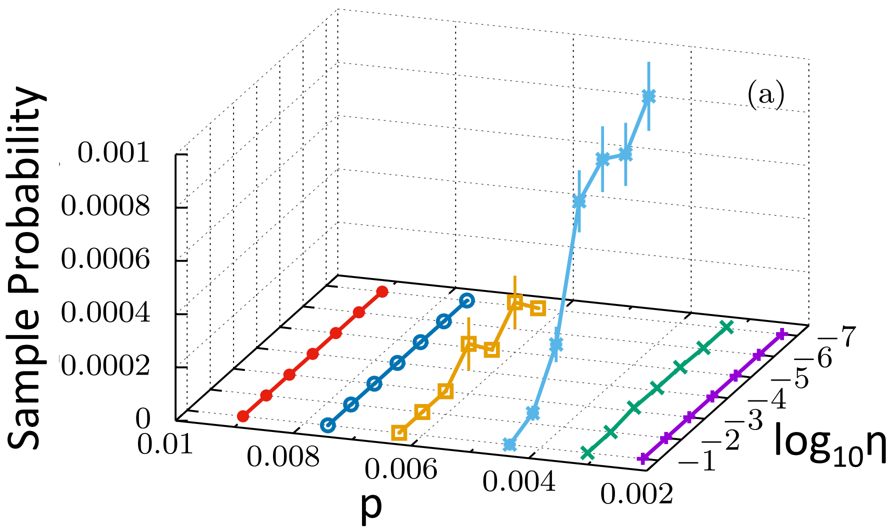
Open Science Grid

~20,000 Scatter points on this plot.

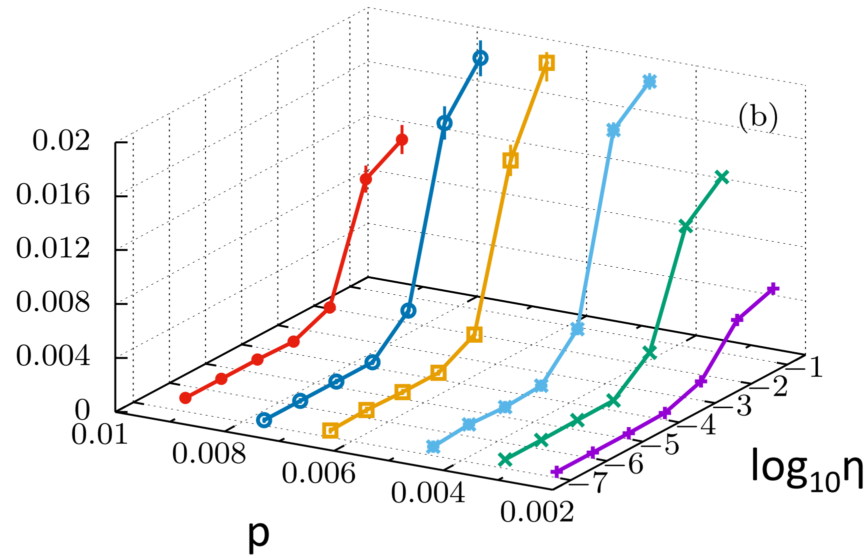
B.F. Intoy, A. Wynveen, and J. W. Halley
Physical Review E *94, 042424 (2016)*

Probabilities of DALD, DDLA, DALA states as a function of p and η

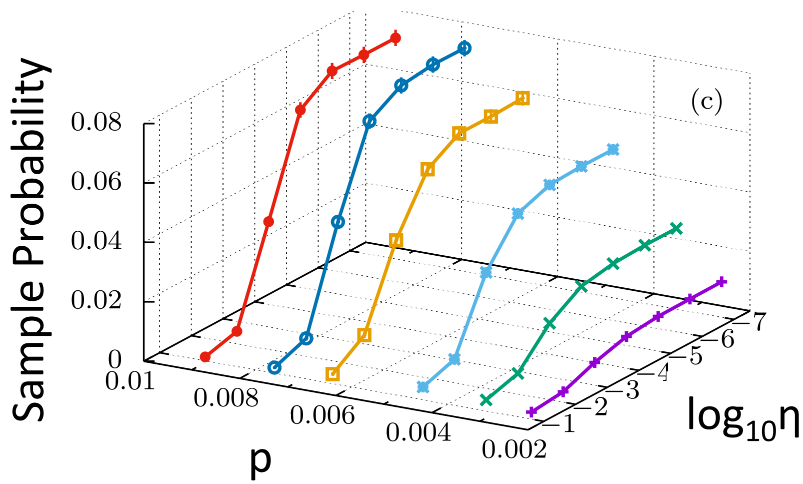
DALD



DDLA



DALA

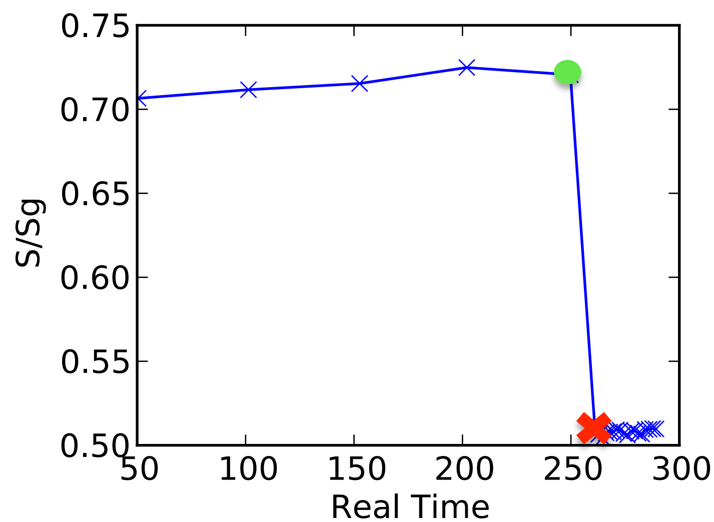


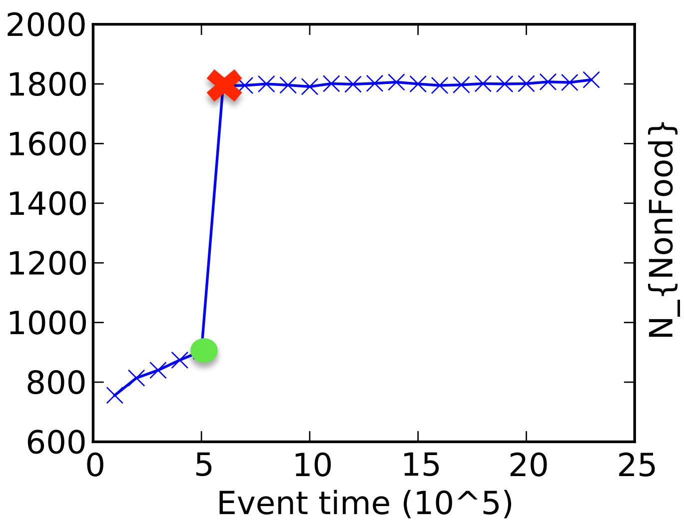
Open Science Grid

~700,000 Simulations were done to make these plots.

B.F. Intoy, A. Wynveen, and J. W. Halley
Physical Review E *94, 042424 (2016)*

DALA States Display ‘cancer-like’ Explosions

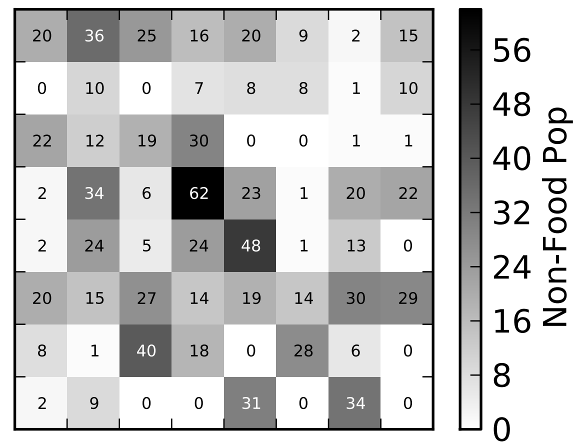




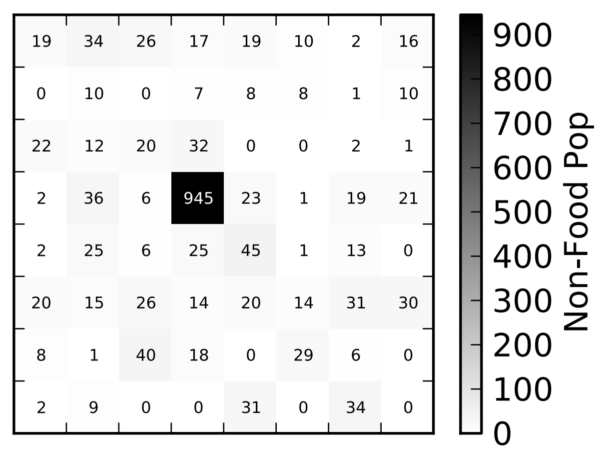
$p=0.00452$

$\eta=10^{-7}$

Before Jump (Green Dot)



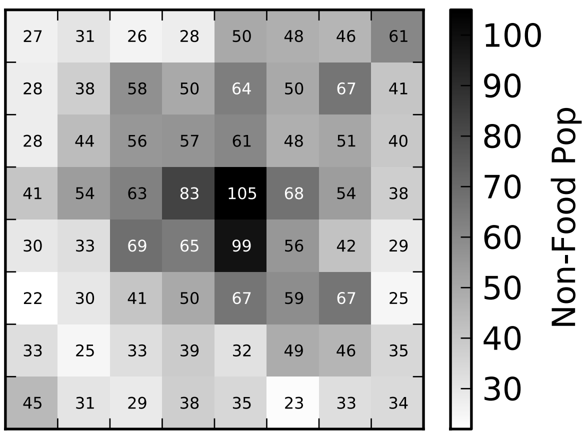
After Jump (Red X)

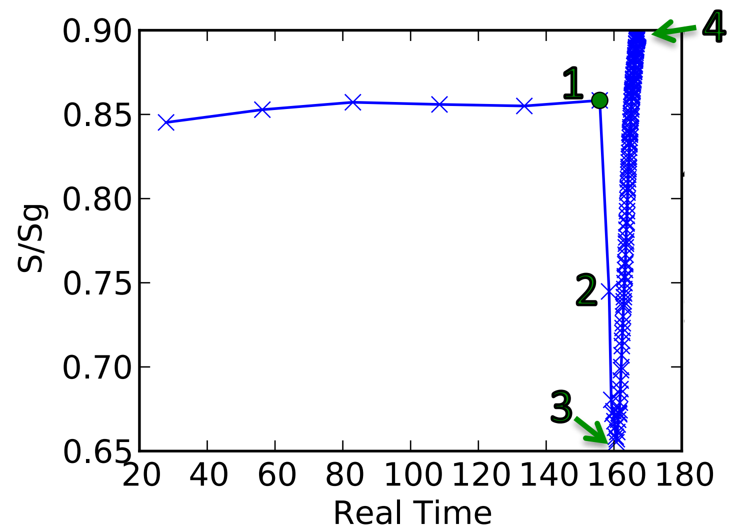


Collective Effect

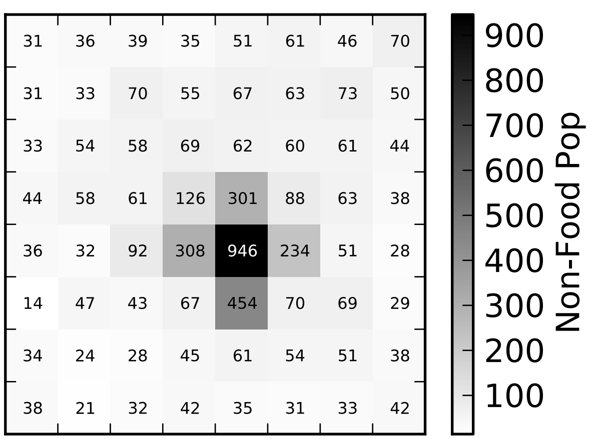
With increasing η the explosion spreads: $p=0.00452$, $\eta=10^{-1}$

1:

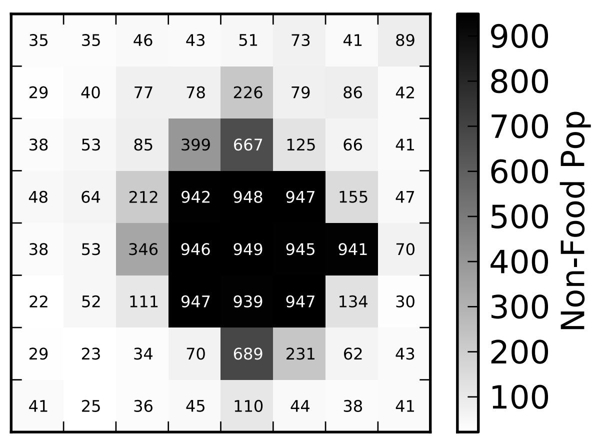




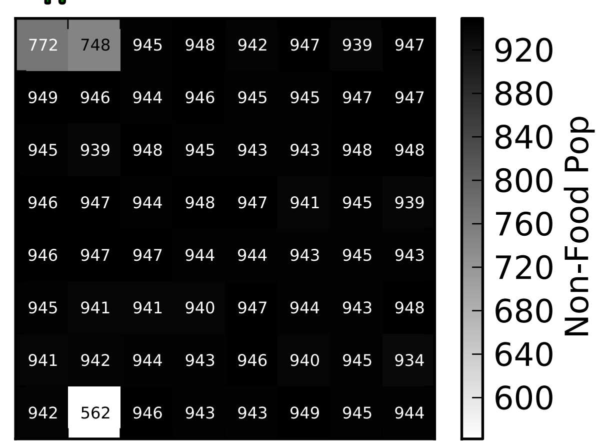
2:



3:



4:



Conclusions:

- With the inclusion of space we counted the likelihood, as functions of p and η, of lifelike states characterized unequilibrated (DALA), diffusively but not chemically unequilibrated (DALD) and chemically but not diffusively unequilibrated (DDLA).
- DDLA states closely reproduce the states in the earlier, single site, model.
- DALD are rare.
- DALA exhibit explosive growth.

OSG Computational Resources Used: Open Science Grid

- ~1.4 Million computational wall time hours was used for the resulting publication (Physical Review E *94, 042424 (2016))*.

Current Work:

- Going back to single site (well mixed) simulations:
  - Interested in the effects of bond energy and temperature on the model.
  - Exploring the sensitivity of what is in the food set (have length one and two, but could be in different proportions).
  - Interested in the effects of increasing the number of monomer types (currently only have two, biologically DNA has 4 and proteins have 20).
  - Still using OSG to perform simulations!

Thank you!

Entropy Calculations and Misc

$$S(\{N_{L,i}\}) = \sum_{i=1}^{M} S_i(\{N_{L,i}\}).$$

$$S_i(\{N_{L,i}\}) = \sum_L \ln \left[\frac{(N_{L,i} + 2^L - 1)!}{(2^L - 1)! N_{L,i}!} \right]$$

$$S_{\mathrm{global,eq}}(N) = (MG_{l_{\max}} - l_{\max}) F \left(\frac{N}{MG_{l_{\max}} - l_{\max}} \right)$$

$$N_i \ = \ \sum_L N_{L,i}$$

$$F(x) \ = \ (1 + x) \ln(1 + x) \ - \ x \ln x$$

$$N_L \ = \ \sum_i N_{L,i}$$