Statistical mechanical model for self replication

Jishnu- 201801128

May 14, 2022

Contents

1 Introduction

Self replication is an attribute of any system to duplicate an produce an identical copy. We study the different self replication models built on pure logic starting from Von Nuemann's automata to the recently published Micheal Brenner's colloidal replication and try to build a Isling based lattice model.

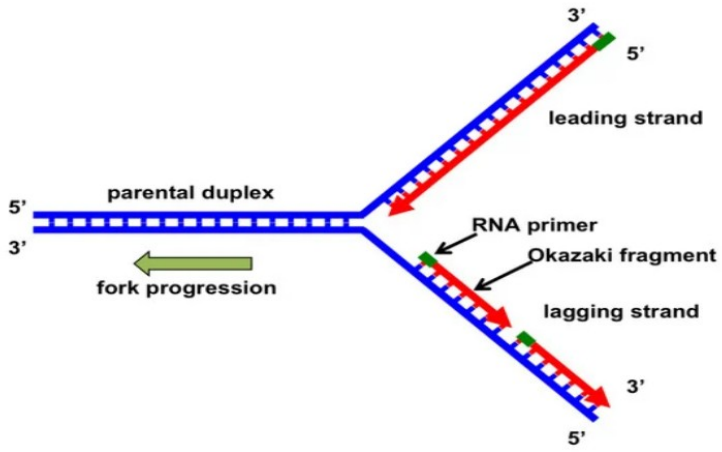
2 Replication in biological systems

2.1 DNA Replication

Dna double helix contains a sequence of nucleotides that is exactly complementary to the nucleotide sequence of it's partner strand. Each strand can serve as a template.

Fork formation- Dna Helicase enzyme unwinds the DNA and cretate a fork like structure

Primer attaching- RNA primer gets attached 3' end of the strand. Primer always binds as the strating point of replication.



Elongation-DNA polymerase III binds to the strand at the site of the primer and begins adding new base pairs complementary to the strand during replication.As replication proceeds in the 5' to 3' direction on the leading strand, the newly formed strand is continuous. But in other strand new strands are being created in the opposite direction of progression of the fork. So the newly formed strand is discontinuous and is called lagging strand

Termination Once both the continuous and discontinuous strands are formed, an enzyme called exonuclease removes all RNA primers from the original strands.The ends of the parent strands consist of repeated DNA sequences called telomeres.Telomerase catalyzes the synthesis of telomere sequences at the ends of the DNA. Once completed, the parent strand and its complementary DNA strand coils into the familiar double helix shape

3 Von neumann's models

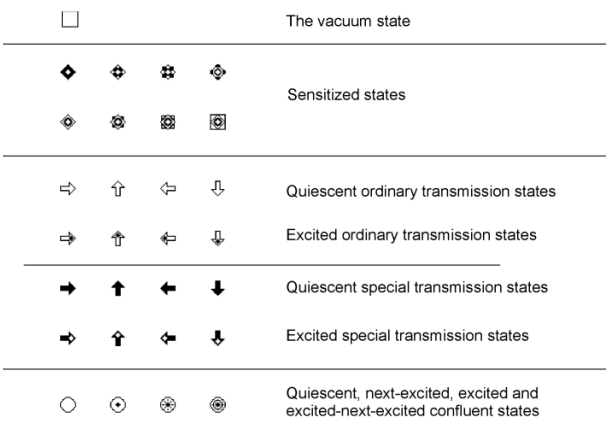
3.1 Kinetic model

This was one of the first replication models. According to this model the constructing automaton floats on a surface surrounded by unlimited supply of elemntary parts. The elemntary parts include , a stimulus organ, four logical devices, a muscle ,a fusing organ and a cutting organ.These parts move ,under the description of parent automaton, around and get assembled to form an identical copy of the automaton.In this sense motion is the basic operation underplay. Von Neumann disregards about enrgy and fuel in this model.

3.2 Cellular model/Von neumann's universal constructor

The Universal constructor of Von neumann consists of three parts-1) a description of the automation(kind of like a blue print to the machine) 2) a copy machine which copies the description and 3)a universal constructor mechanism which that can read any description and construct the machine encoded in that description To

define the constructor in detail Von neumann uses a cellular automaton. Which is basically a 2D grid of cells.Each of these cells has 29 different states. At each time the state gets changed according to deterministic transition rule,with neighbouring cell states of previous time step as parameters for this rule.



The cell state is either quiescent or excited state.The excited states are marked by a white or black dot on its symbol.The **sensitised states** are intermediate steps towards the generation of all other states from the vacuum state.(represnted by a blank square).unexcited ordinary and transmission together with confluent cells are called quiescent state cells.These are stable states in the sense that they remain unvaried unless one or more neighbouring states are in suitable excited state.**The confluent cell** is generally set at the afference of two or three ordinary transmission lines to perform logical AND , Logical OR operation.

When it comes to issues of this model, fist and foremost is the lack of crossing cells to enable intersection of two trasnmission lines without their interference. The trasnition rule also doesn't provide cell states , long lasting memory.The reason for this is that any excited cell that is not further excited by adjacent cells decay into quiescent state straightaway. Nobii and Pesavento had resolved these issues by proposing a modified transition rules, in which they brought new confluent cells which could serve as crossing elements when found at the intersecion of transmission signal trains. By the introduction of crossing elements a system simultaneously reading multiple tapes could be implemented.

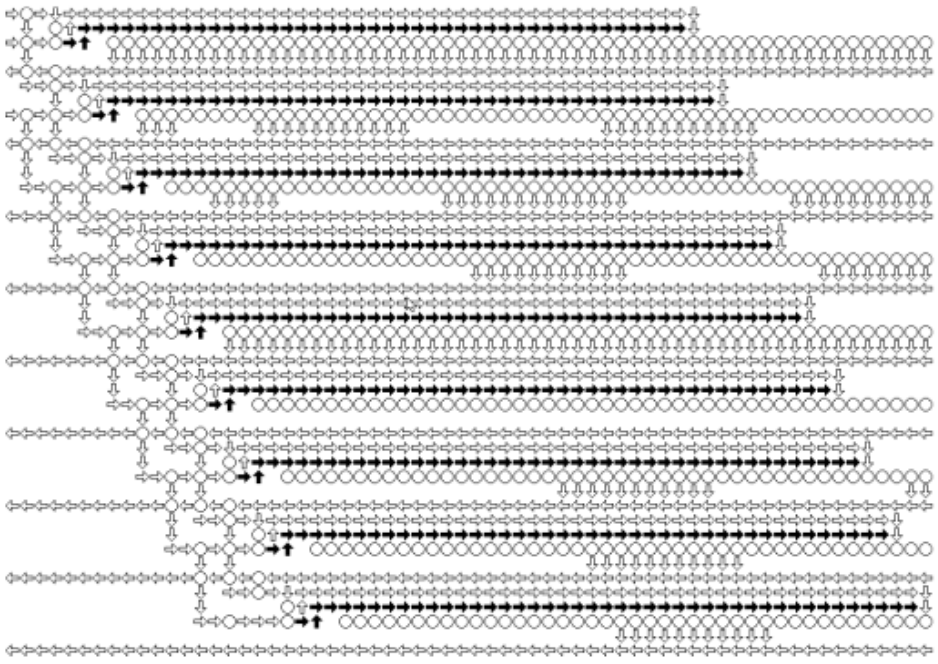


Figure 1: nine tape reading system constructed out of modified transition rules

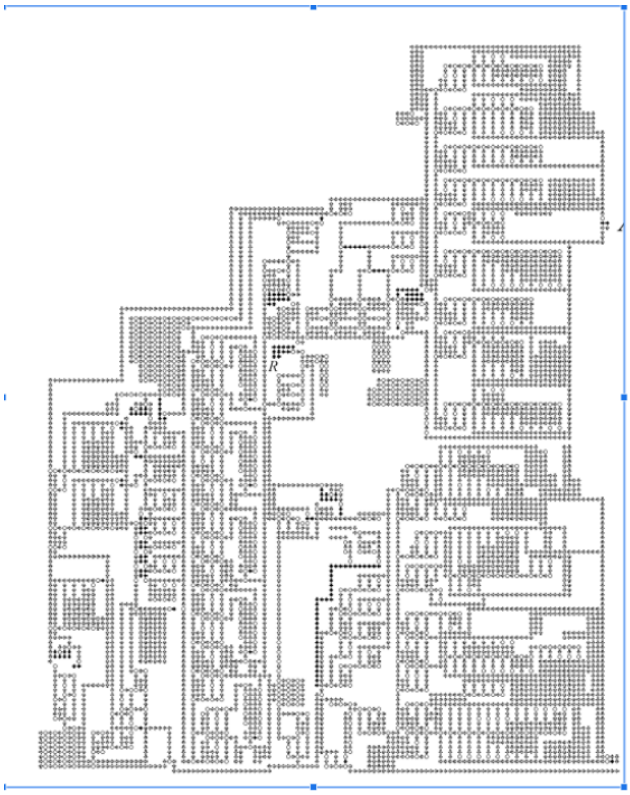


Figure 2: Von neumann's universal constructor- reading writing loop is the one which collects the information in the tape i.e the description of machine

4 Langton model

Langton had observed in 1984 that although the capacity for universal construction is a sufficient condition for self-replication, it is not a necessary one. Not only the automaton gets replicated in a universal constructor,but also the program(i.e the description tape charcterstic to machine). Thus it requires a double recursion on the program: the first to reproduce the automaton and the second to copy the description. Langton based on this observation tried to find minimally complex logical organisation that can just replicate the automaton. Langton's was successful in this attempt . His replicating structure is 'P' shaped loop with about two hundred transition rules . The signals contained in the loop cycle through it. one can use a loop of this sort to store a program dynamically rather than in static form, as on a tape. (back in the universal constructor) This is of tremendous aid in the design of a self-reproducing machine because it eliminates the complex machinery associated with moving a read head back and forth over a static tape. Furthermore, rather than having to decode the information on a tape to obtain

Figure 3: template

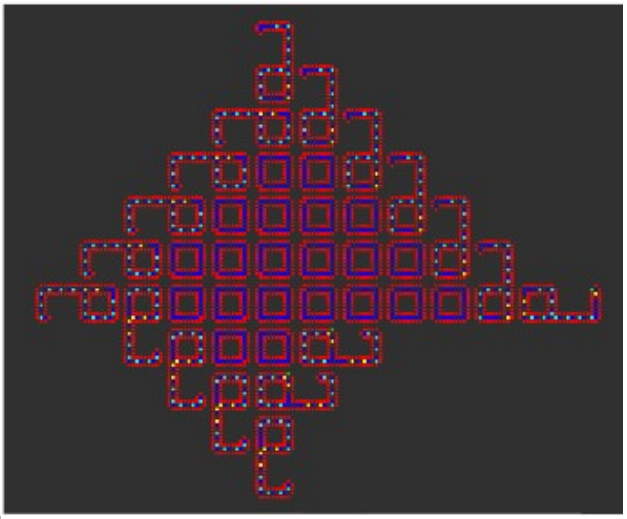


Figure 4: langton loops

the signal sequences necessary to effect the construction of a machine, the "program" which we store in the loop can simply be the proper signal sequence to effect the construction, the same sequence which would have been generated as a result of decoding asequence of 1 and 0 marks on a static tape.

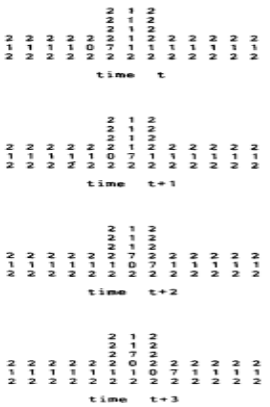


Figure 5: duplication at junction

Langton's sructure consisted of 1's which
he called data path. and 2's which acts as sheath cells. The name data path was just to implicate that these are capable trasmiting data in the form of signals. The signals contained within the loop, cycle through it indefinitely . whenever signal encounters the arm junction it is duplicated, with one copy propagating back around the loop again and the other copy propagating down the arm. using '7-0 ' signals loop tail extends by one cell. which later turns left with additional help off cell state '3'. And an identical copy is produced within 160 iterations.

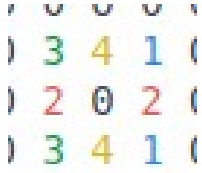


Figure 6: square template
the transition rules are[00300:5,00500:1,01010:2,00100:4,00400:3,03030:2,50301:0]

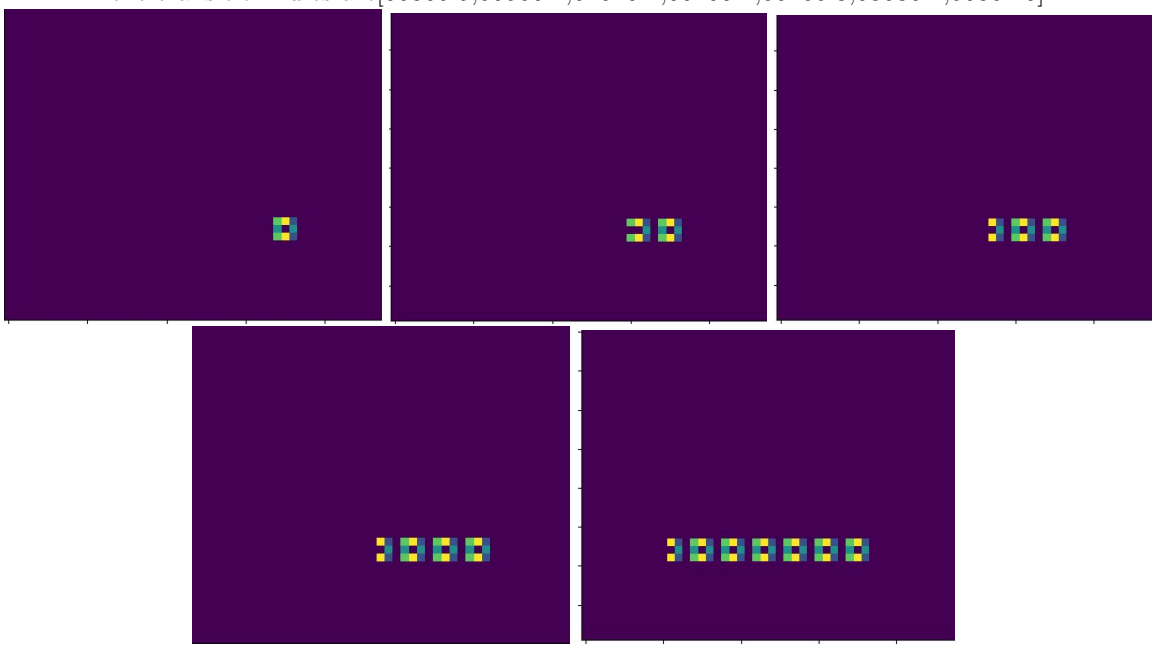


Figure 7: various stages during linear growth

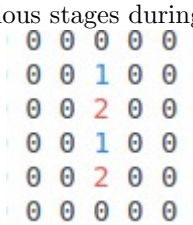


Figure 8: line template

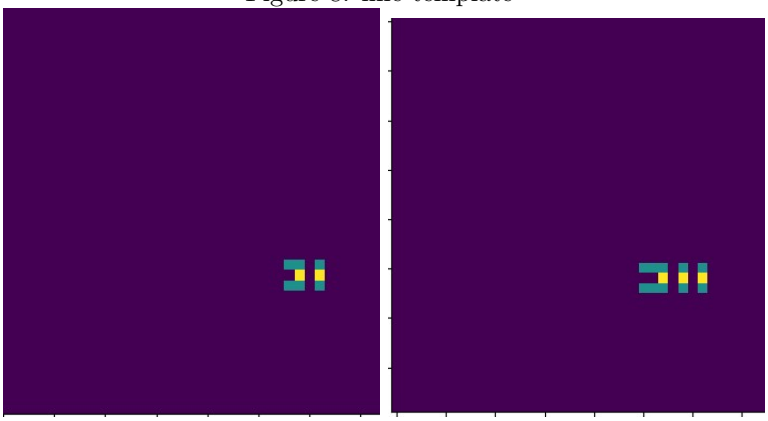


Figure 9: various stages during linear growth

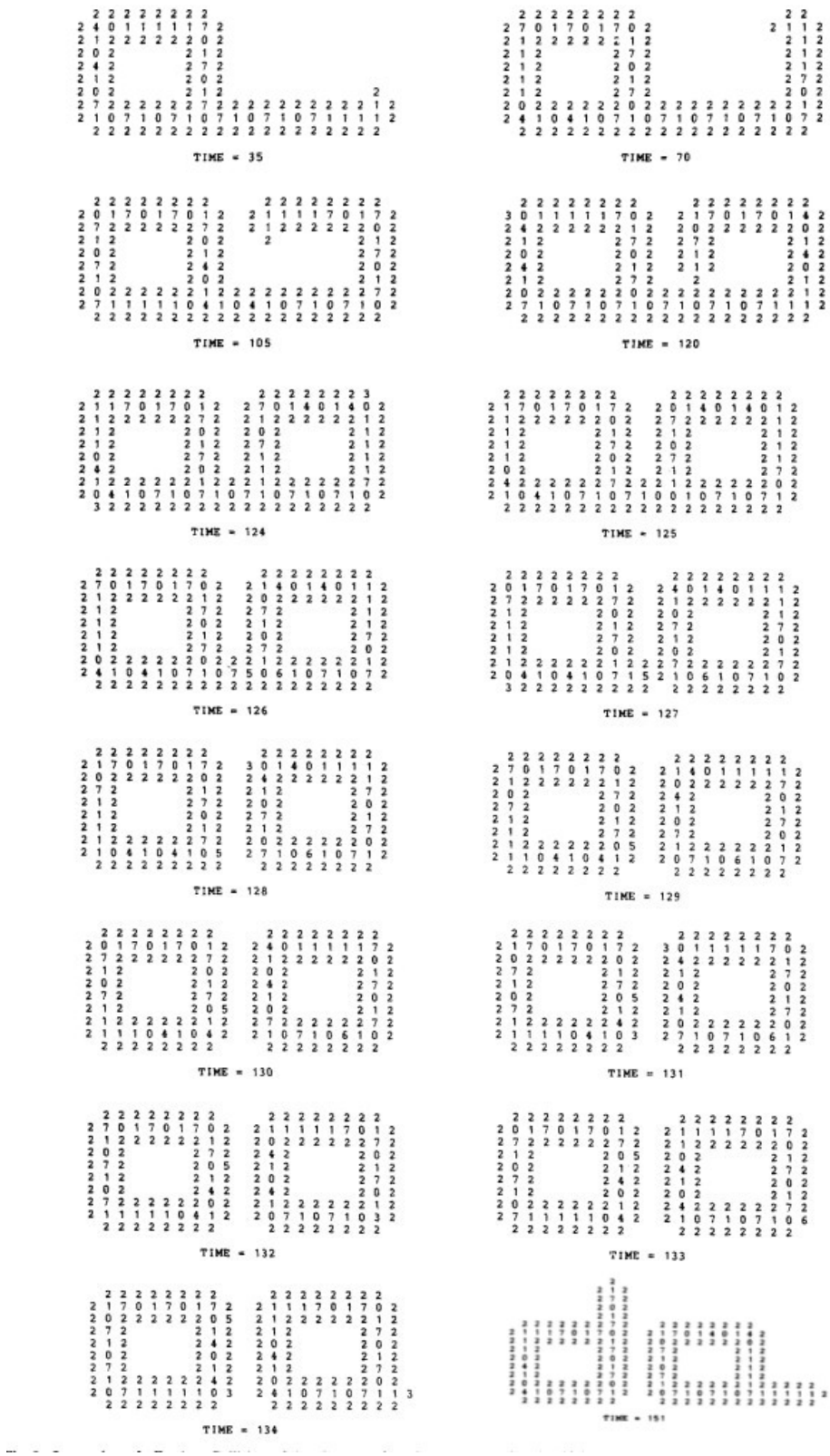


Figure 10: Langton's P shaped loop at various time steps during repliscation

5 Replication of colloidal clusters

The scheme of replication designed by Brenner and Zervacic consists of a seed octahedron which acts as a template to generate another octahedron. Each particle in the seed can interact with its complemenary particle from the monomer bath. The reaction scheme also contains a catayst with different type of particles.

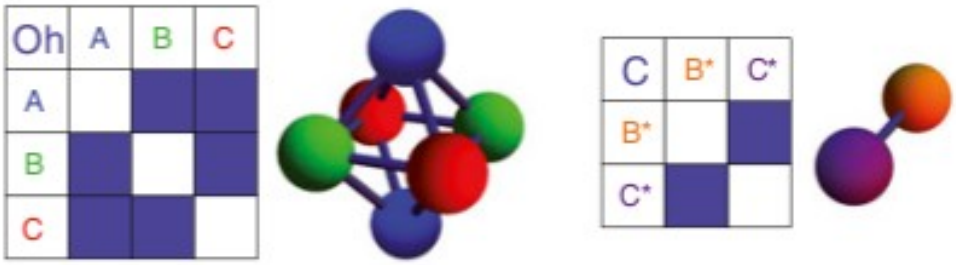


Figure 11: interaction metrices for catalyst and seed octahedron

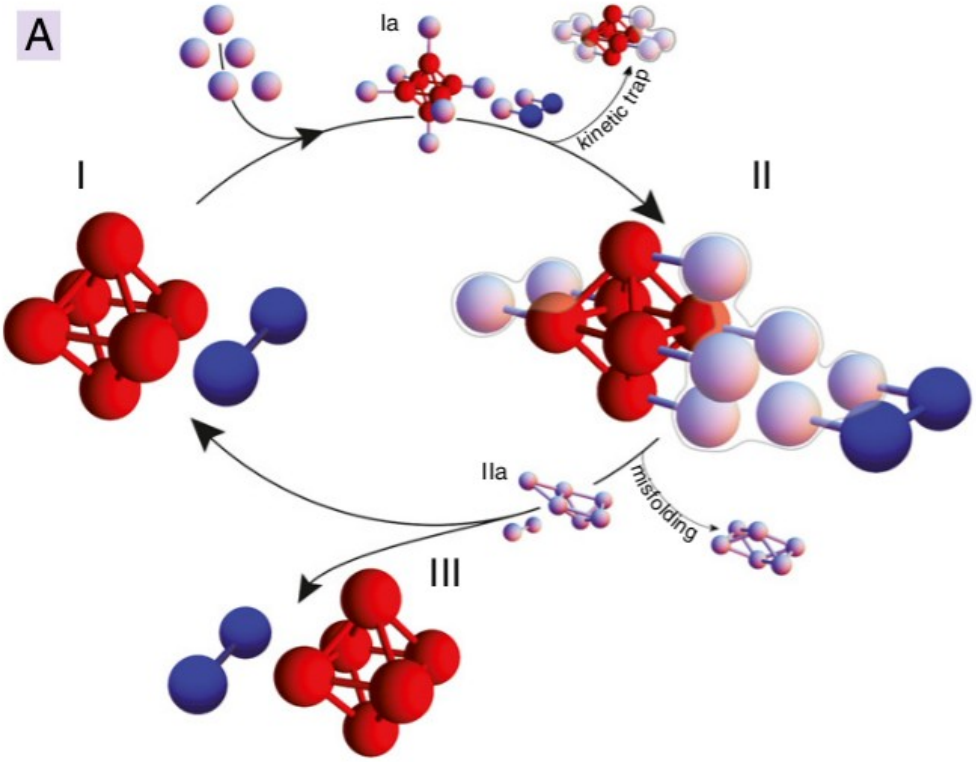


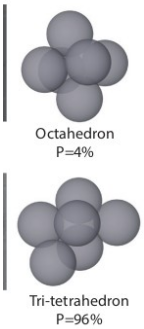
Figure 12: Only one appropriately labeled monomer (gray particles) per particle in Oh/catalyst can be attached (stage Ia). Attached monomers can now interact between themselves and form substructures templated by the Oh and the catalyst (stage II).Once a certain number of bonds between these monomers are formed, melting occurs, separating starting Oh and catalyst from attached monomers (stage IIa). New structures formed produce a new catalyst and a nonrigid cluster that folds into a new octahedron (stage III)

Along the reaction pathway there are two undesirable things (i)a kinetic trap which is formed by the monomers getting attached to form two traingles on either side. This structure is relatively frequent and stable. (ii) after melting the new structure can misfold into a local minimum. To achieve efficent replication it is enough to i) to avoid kinetic trap ii)bonds between parent and monomers should be strong but still in equilibrium iii)The newly formed octahedron bonds should be irreversible By taking into consider these they have imposed certain rules on the bond strength. With ϵas the strength between two monomers, they have put 2ϵas the bond strength between monomer and seed octahedron. To destablise the kinetic trap they choose A'-B' and B'-C' weaker compared to B'-C'(this bond strength is put as 1.5 ϵ). Further the bond strenth between particles in the seed has been put to be 5ϵ

6 Physical interpretation of partition functions of colloidal clusters

Colloidal particles are held together by short-range attractions.when the attractive interactions are weak ,oarticles can rearrange themselves different configurations.in the case of $N = 6$,the equilibrium mixture of minimum energy configurations consisted of tritetrahedron and octahedron with former occuring 24 times than the latter.

6.1 Statistical mechanical model



Probability of observing a particular structure s_k in equilibrium ensemble P_{s_k} is propotional to partition function Q_{s_k}

$$P_{s_k} = \frac{Q_{s_k}}{\sum Q_{s_l}} \tag{1}$$

6.1.1 Partition function in particle cordinate frame

In particle cordinates the position of every particle ($\vec{q_i}$) is measured from the origin of a lab frame. set of translational degree of freedom :$\mathbf{Q}$=$[q_{1_x}, q_{1_y}, ...q_{N_z}]$ .the corresponding conjuagte momenta given by the set $\mathbf{P}$. Each particle rotate through euler angles $[\phi_i, \theta_i, \psi_i]$ relative to labframe.The set of all such angles is denoted by $\mathbf{\Phi}$ and $\mathbf{L}$ denote the set of momenta conjuagte to these angles . The Hamiltonian for this system of N particles :

$$\mathbf{H} = U(\mathbf{Q}) + K(\mathbf{P}) + U(\mathbf{\Phi}) + K(\mathbf{L}) \tag{2}$$

The canonical partion function Q is :

$$\begin{aligned} Q_s &= \frac{1}{h^{6N}\sigma_s} \int_{A_s} \exp^{-\beta H}\, d\mathbf{Q} d\mathbf{P} d\phi dL \\ &= \frac{1}{h^{6N}\sigma_s} \int_{A_s} \exp^{-\beta U(\mathbf{Q})}\, d\mathbf{Q} \int \exp^{-\beta K(\mathbf{P}}\, d\mathbf{P} \int \exp^{-\beta U(\mathbf{\Phi}}\, d\phi \int \exp^{-\beta K(L)}\, dL \\ &= \frac{1}{h^{6N}\sigma_s} Q_{s,tran}(\mathbf{Q},\mathbf{P}) Q_{s,rot}(\phi, L) \end{aligned}$$

where σ_s is the symmetry factor for the structure 's'. And A_s is the adjacency matrices corresponding to structure 's' $\int_{A_s} \exp^{-\beta U(\mathbf{Q})}\, d\mathbf{Q} = V \prod_{i=2}^{i=N} V_s, i$, where the $V_{s,i}$'s are the effective volume of the i'th particle (in structure 's')can explore. Further $K(\mathbf{P}) = \sum_{i=1}^{i=N} \frac{1}{2m_i}(p_{ix}^2 + p_{iy}^2 + p_{iz}^2)$. Hence $\int \exp^{-\beta K(\mathbf{P}}\, d\mathbf{P} = \prod_{i=1}^{i=N} (\frac{2\pi m_i}{\beta})^{\frac{3}{2}}$

$$Q_{s,trans} = \prod_{i=1}^{i=N} (\frac{2\pi m_i}{\beta})^{\frac{3}{2}} Z_s \tag{3}$$

whwre $Z_s = V \prod_{i=2}^{i=N} V_s, i$ Assuiming rotational potential energy depends neither on the orientation of particles nor on their positions in the cluster it can be said that $U(\Phi) = 0$. Whereas the Rotational kinetic energy

is the sum of that of individual particles .Therefore $Q_{s,rot}$ is independant f structure 's'. i.e. $Q_{s,rot} = Q_{rot}$ So total partition function is :

$$\frac{1}{h^{6N}\sigma_s} Q_{s,tran}(\mathbf{Q},\mathbf{P}) Q_{s,rot}(\phi, L) = \frac{Q_{rot}}{h^{6N}\sigma_s} Z_s \prod_{i=1}^{i=N} \frac{2\pi m_i}{\beta}^{\frac{3}{2}}$$
$$= \frac{2\pi}{\beta}^{\frac{3N}{2}} V \prod_{i=2}^{i=N} V_s, i \prod_{i=1}^{i=N} m_i^{\frac{3}{2}}$$

6.1.2 Partition function in centre of mass frame

Rigid rotor -Harmonic potential asssumption helps in splitting the hamiltonian into terms describing the translation of centre of mss,rotations about the centre of mass and vibrations wrt to lowest energy positions.Six degrees of freedom describe the position and orientation of rotating frame relative to lab frame.The ttranslational degree of freedom $\mathbf{q'} = (\mathbf{q'_x}, \mathbf{q'_y}, \mathbf{q'_z})$ describe the position of CM wrt to labframe. Their conjuagate momenta are $\mathbf{p'} = (\mathbf{p'_x}, \mathbf{p'_y}, \mathbf{p'_z})$. The rotational degree of freedom are the standard euler angles i.e $\mathbf{\Phi'} = (\phi', \theta', \psi')$ their conjuagate angular momenta are $\mathbf{L'} = (p'_\phi, p'_\theta, p'_\psi)$. With harmonic approximation, vibrations of the cluster could be explained by the 3N-6 modes of vibrations. The set of vibrational displacements are denoted by $\epsilon' = (\epsilon'_1, \epsilon'_2, ..., \epsilon'_{3N-6})$ Under these assumptions Hamiltonian in CM frame is:

$$H'_s = U_s(\mathbf{q'}) + K_s(\mathbf{p'}) + U_s(\Phi') + K_s(\mathbf{L'}) + U_s(\epsilon') + H_{rot}(\Phi, \mathbf{L}) \tag{4}$$

The partition function Q'_s is:

$$Q'_s = \frac{1}{h^{6N}\sigma_s} \int exp^{-\beta H'} d\mathbf{q'} d\mathbf{p'} d\mathbf{\Phi'} d\mathbf{L'} d\boldsymbol{\epsilon'} d\chi'$$
$$= \frac{Q_{rot}}{h^{6N}\sigma_s} \int \exp^{-\beta H'_{trans}} d\mathbf{q'} d\mathbf{p'} \int \exp^{-\beta H'_{rot}} d\mathbf{\Phi'} d\mathbf{L'} \int \exp^{-\beta H'_{vib}} d\boldsymbol{\chi'} d\boldsymbol{\epsilon'}$$
$$= \frac{Q_{rot}}{h^{6N}\sigma_s} Q'_{s,trans}(\mathbf{q'}, \mathbf{p'}) Q'_{s,rot}(\Phi', \mathbf{L'}) Q'_{s,vib}(\boldsymbol{\epsilon'}, \boldsymbol{\chi'})$$

$Q'_{s,rot}$,$Q'_{s,vib}$ and $Q'_{s,trans}$ could be evaluated and are:

$$Q'_{s,trans} = \int \exp^{-\beta U'_{\mathbf{q'}}} d\mathbf{q'} \int \exp^{-\beta K'_{\mathbf{p'}}} d\mathbf{p'} = V(\frac{2\pi M}{\beta})^{3/2} \tag{5}$$

$$Q'_{s,rot} = 8\pi^2 (\frac{2\pi}{\beta})^{\frac{3}{2}} \sqrt{I_{s,1} I_{s,2} I_{s,3}} \tag{6}$$

$$Q'_{s,vib} = \int \exp^{-\beta U'_{\boldsymbol{\epsilon'}}} d\boldsymbol{\epsilon'} \int \exp^{-\beta K'_{\boldsymbol{\chi'}}} d\boldsymbol{\chi'} = \prod_{i=1}^{3N-6} \frac{2\pi}{\beta\omega_{s,i}} \tag{7}$$

putting these together :

$$Q'_s = \frac{Q_{rot}}{h^{6N}\sigma_s} 8\pi^2 V M (\frac{3}{2} (\frac{2\pi}{\beta})^{\frac{3}{2}}) \sqrt{I_{s,1} I_{s,2} I_{s,3}} \prod_{i=1}^{3N-6} \frac{1}{\omega_{s,i}} \tag{8}$$

Equating the partition functions obrianed from two frames :

$$\sqrt{I_{s,1} I_{s,2} I_{s,3}} \prod_{i=1}^{3N-6} \frac{1}{\omega_{s,i}} \propto \frac{\prod_{i=1}^{i=N} m_i^{\frac{3}{2}}}{M^{\frac{3}{2}}} \tag{9}$$

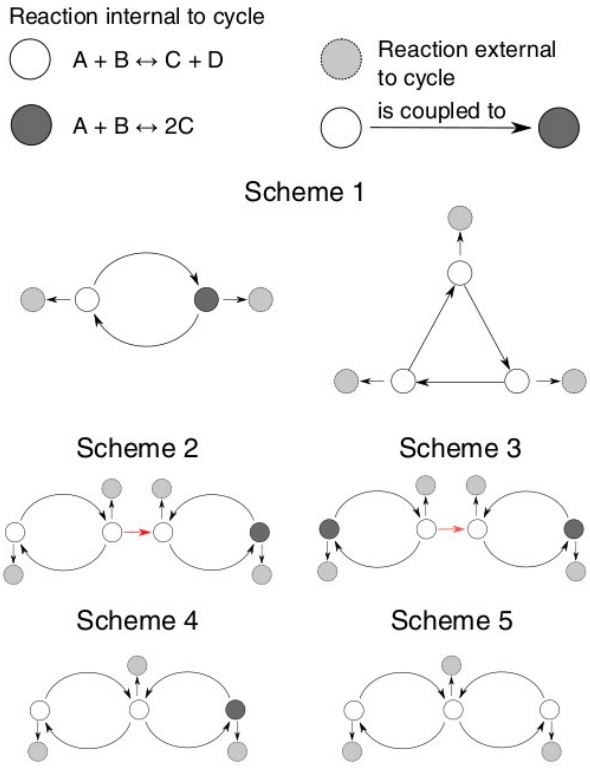
Probability of observing structure 's' ,P_s, is proportional to Q'_s. Hence Probability of observing s1 relative to s2 is proportional to $\frac{\sigma s_2}{\sigma s_1}$. As a result, the model predicts that the equilibrium probability of the octahedron relative to the tri-tetrahedron is proportional to $_{tri-tetrahedron}/_{octahedron} = 2/24 = 1/12$. Experimentally this ratio has been determined to be 1:24 . The accounting of the other factor of two comes from the product of moment of inertia and the vibrational partion function part in Q'_s.

7 Design of conditions of self replication using a toy chemical model

the model consists of two monomer types a)atom types B and G. they interact each other with energies $\epsilon_{BB}, \epsilon_{BG}, and \epsilon_{GG}$ to form molecules.The molecules take part in dissosiation or bimolecular reactions.The free energies of molecules are completely determined by ϵ parameters.

7.1 Condition for self replication

Prior works on catayltic cycles have shown that kinetic dominance of reactions can be quantified through specificity. The condition for self replication for various reaction schemes has been listed in terms of cycle specificity below. a) For isolated ACC's(auto catalytic cycles) it should be greater than .5



b)For scheme 2 ,exponential growth occurs unless specificity of ACC is greater than .5. For scheme 3 it is possible to observe exponential growth as long as one of ACC specificity greater than .5
c) For scheme 4 and scheme 5 ,it is possible to observe exponential growth when both cycles have specificity less than .5

7.2 Coarse control of exponential growth

To quantitatively establish the relationship between rate constant and transient behaviour of reaction network, it is not very practical explore the whole paramater space.Alternatively is done by using coarse features of rate constants.
a) protocol PF,the fraction of fast recations: in this protocol all the interaction energies are 0 and rate constants are chosen in such a way that a controllable fraction ,p_{fast} of reeacion may occur and the rest of them effectively forbidden.To implement this system random barriers are drawn from 0,inf corresponding to rate

constants 1 ,0.The fAst reactions with rate constant 1 are assigned a probability p_{fast}.To ensure detailed balance conditions, the barriers for the forward and the reverse reactions were set to be equal.Self-replication occurs if and only if at least one autocatalytic cycle in the reaction network has direct and exclusive access to its fuel . Hence, p_{sr} can be calculated from (a) the probability of finding at least one autocatalytic cycle with direct access to its fuel, $p_{acc}(pfast)$, and (b) the probability that all autocatalytic cycles have side reactions, $p_{loss}(pfast)$. Hence, for p_{fast} = x, $p_{sr}(x) = p_{acc}(x)(1 - p_{loss}(x))$. Due to trade off, there is optimal p_{fast} for which p_{sr} is maximised.
b)Protocol CD-The first and simplest hypothesis underlied in this protocol is that the p_{fast} can be tuned to optimality by the dispersion of the rate constants. Activation barriers were chosen from exponential distribution with varying values of dispersion c_d while keeping the interaction energy as 0.Numerical solution of equation at steady state obtained them with growth exponent,γ. And it had observed that γ= 1 the growth was exponential a nd for γ¡ 1 the growth was sub exponential. It was also observed that probability of exponential growth, p_{sr} , increased with c_d , eventually saturating at a value that is dependent on the underlying reaction network.

8 Lattice model for replication

We use a 2D grid of spins with a fixed spin structure at a specified position in the Grid.The Hamiltonian for this system is the well known 2D isling Hamiltonian, with three types of interactions being involved. Those are (i) interaction of spins part of the fixed structure ,(ii) interaction between two dynamic spins and (iii) interaction between spins part of the fixed structure and a dynamic spin.From hereforth a dynamic spin is labelled as 'D' and a fixed spin is labelled as 'F'.(Dynamic in the sense that these spin flip constntly whereas the Fixed in the sense that they do not flip at all)

the isling Hamiltonian: $H = \sum_{[i,j]} J_{ij} S_i S_j$

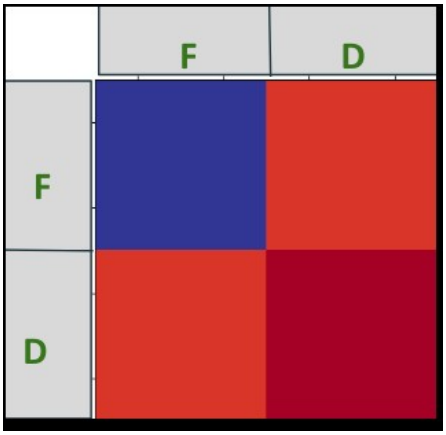


Figure 14: the interaction matrix with J_{FF}=100 ,J_{FD}=10, and J_{DD}=1

By running the monte carlo- metropolis with the Specified Hamiltonian without changing the coupling constant matrix, we could see that fluctuation of interfacial magnetisation is relatively less compared to those of spins in the bulk.

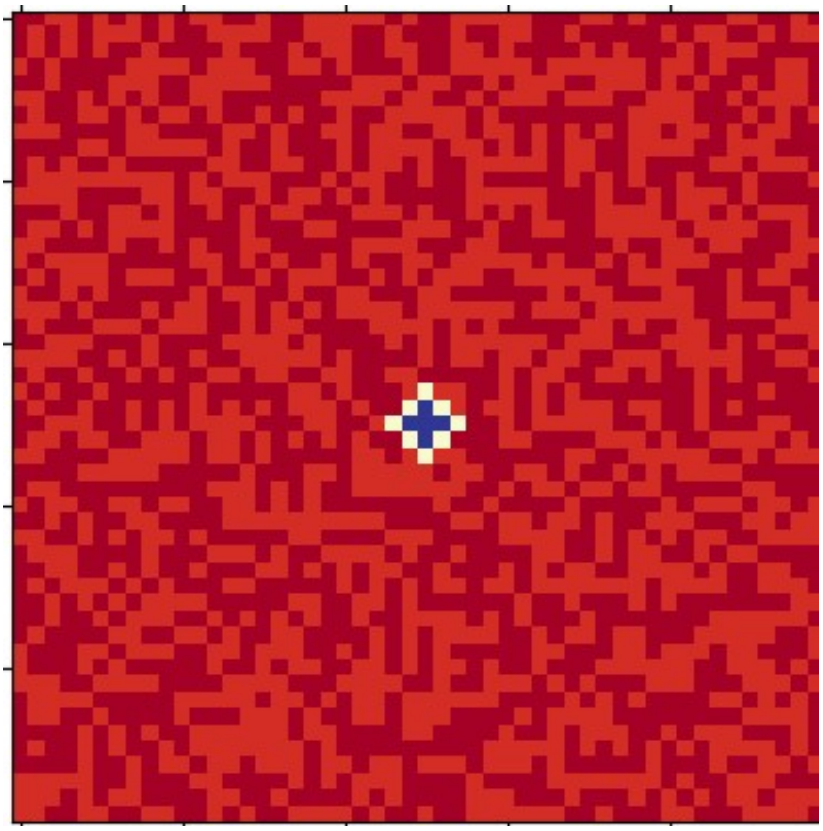


Figure 13: the blue indicates the fixed spins, the white cells indicate the interfacial spins of fixed structure and the red , orange indicate the down and up spins respectively. The latter ones are collectively called as the dynamic spin and former as fixed spins

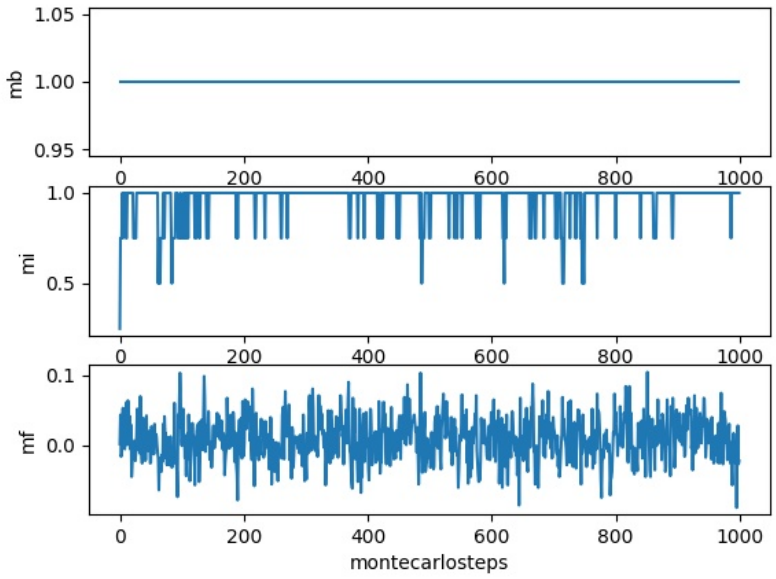


Figure 15: mb- magnetisation of bound/fixed spin//mi-magnetisation of interfacial spins,mf -magnetisation of free spins

Apart from the three interaction inserted via Hamiltonian ,we also introduced allostery so that the seed template gets replicated around its very neighbourhood. The allostery consists of fine tuning the the coupling parameter from J_{DF} to J_{FF} depending upon the the increase in the monomers(spins).(Stabilising the structure with increasing spin) This has resulted in the replication of the fixed structure and soon filling the whole lattice. With these three interaction and allostery a fixed structure in the form of '+' is replicated.

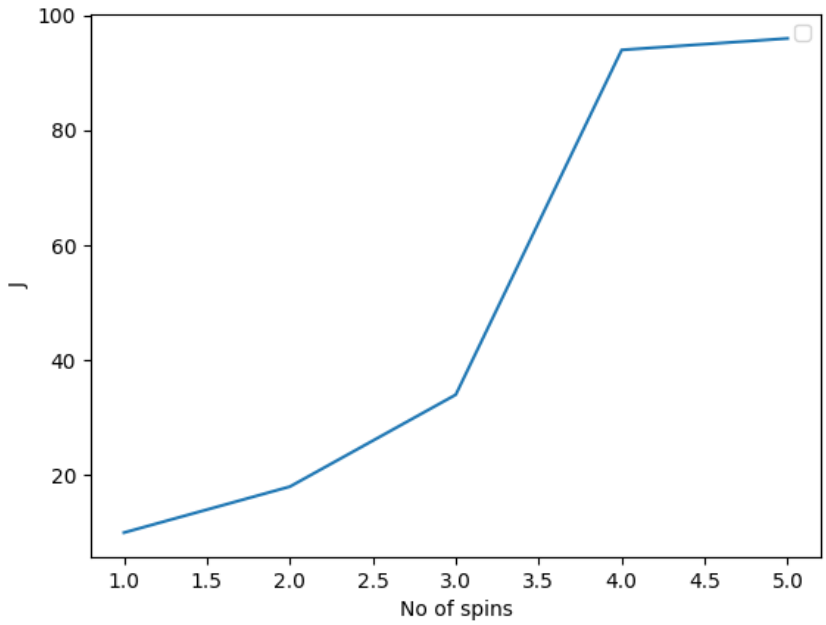


Figure 16: sigmoid allostery

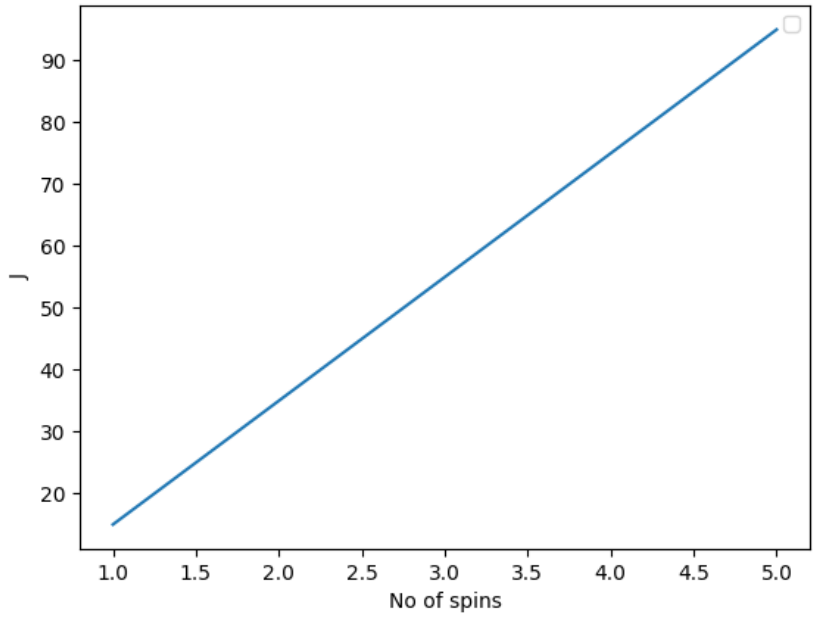
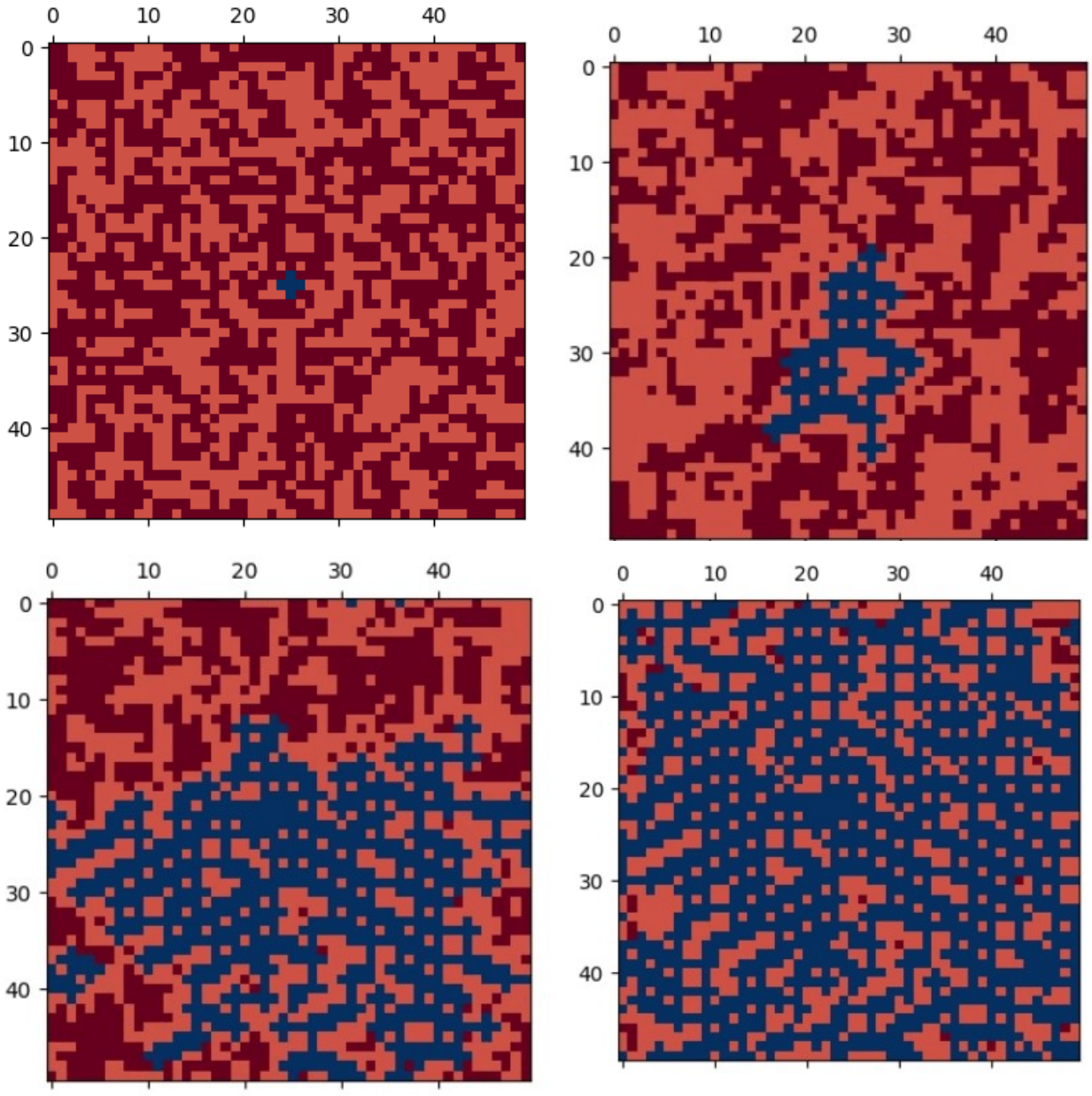


Figure 17: linear allostery



8.1 Analysis

While plotting the number of fixed structures (n_s) against the monte carlo steps(MCstep) , it could be seen that with allostery there has been an exponential growth. The growth over montecarlo steps in the case of no allostery saturates after very large number of montecarlo steps.

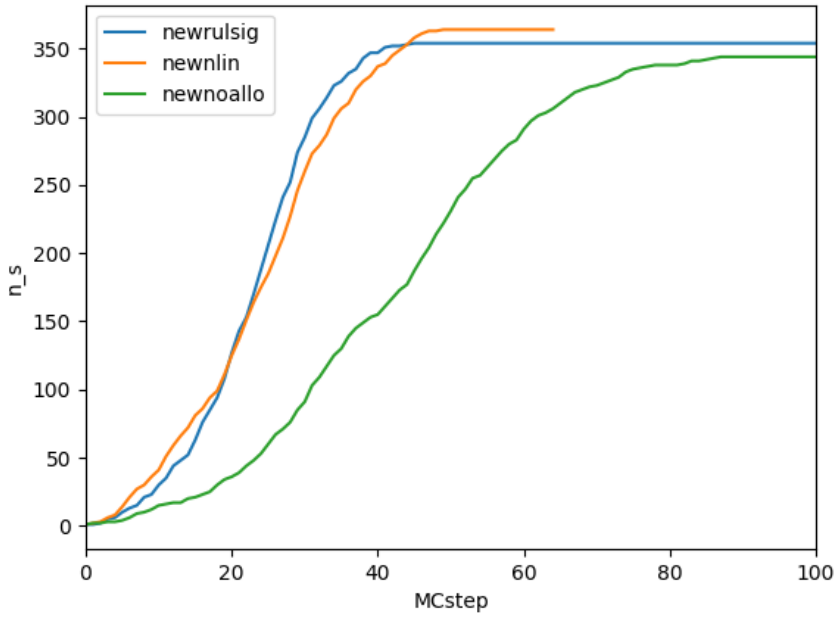


Figure 18: Growth with Linear allostery ,sigmoid allostery and No allostery For grid dimension L=50

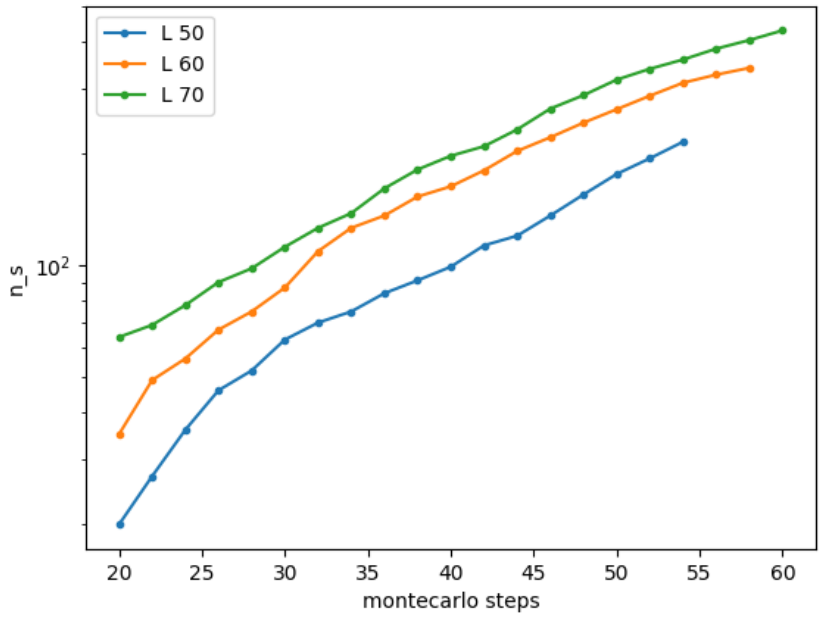


Figure 19: growth with no allostery for various grid size

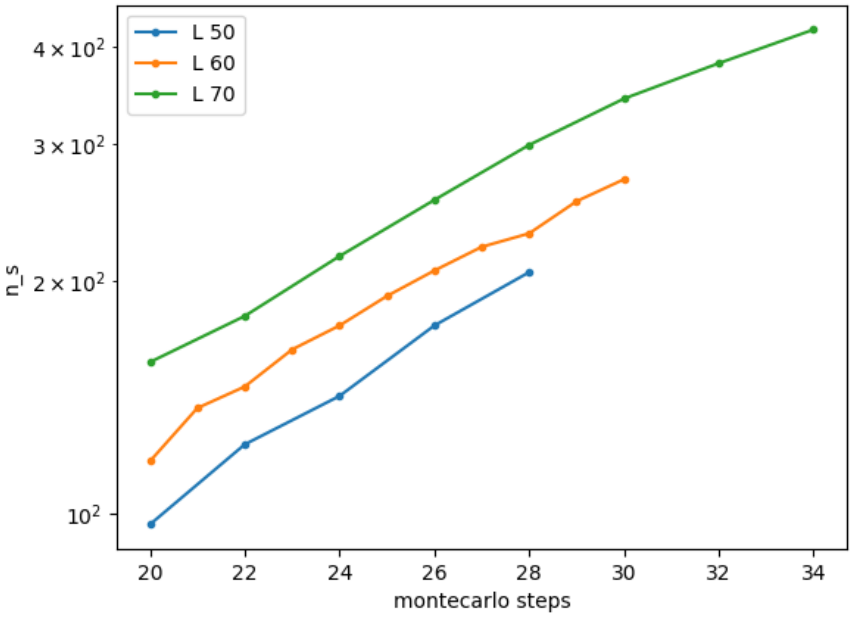


Figure 20: growth with linear allostery for various grid size

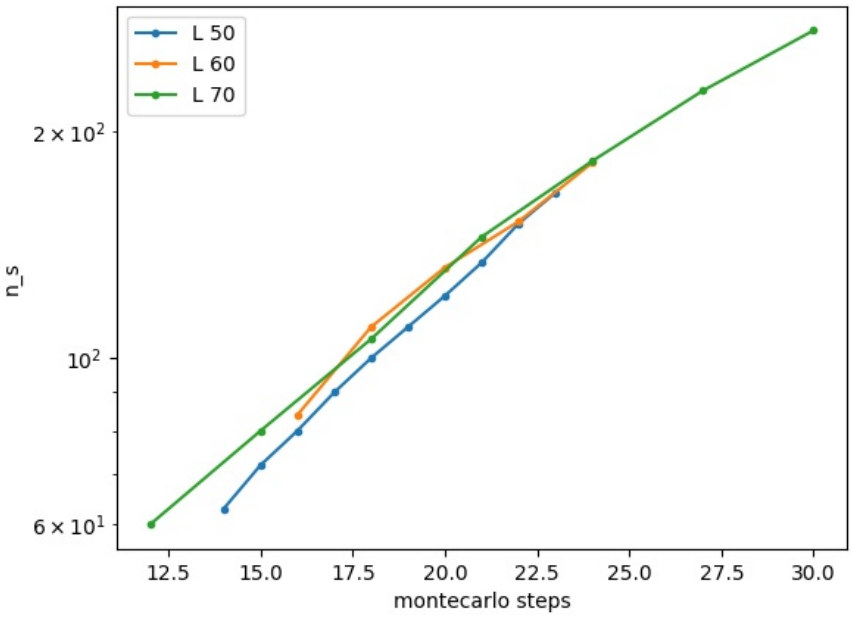


Figure 21: growth with sigmoidal allostery for various grid size

9 Appendix

Transition function table for self-reproducing loops

CTRBL->I	CTRBL->I	CTRBL->I	CTRBL->I	CTRBL->I
00000->0	02527->1	11322->1	20242->2	30102->1
00001->2	10001->1	12224->4	20245->2	30122->0
00002->0	10006->1	12227->7	20252->0	30251->1
00003->0	10007->7	12243->4	20255->2	40112->0
00005->0	10011->1	12254->7	20262->2	40122->0
00006->3	10012->1	12324->4	20272->2	40125->0
00007->1	10021->1	12327->7	20312->2	40212->0
00011->2	10024->4	12425->5	20321->6	40222->1
00012->2	10027->7	12426->7	20322->6	40232->6
00013->2	10051->1	12527->5	20342->2	40252->0
00021->2	10101->1	20001->2	20422->2	40322->1
00022->0	10111->1	20002->2	20512->2	50002->2
00023->0	10124->4	20004->2	20521->2	50021->5
00026->2	10127->7	20007->1	20522->2	50022->5
00027->2	10202->6	20012->2	20552->1	50023->2
00032->0	10212->1	20015->2	20572->5	50027->2
00052->5	10221->1	20021->2	20622->2	50052->0
00062->2	10224->4	20022->2	20672->2	50202->2
00072->2	10226->3	20023->2	20712->2	50212->2
00102->2	10227->7	20024->2	20722->2	50215->2
00112->0	10232->7	20025->0	20742->2	50222->0
00202->0	10242->4	20026->2	20772->2	50224->4
00203->0	10262->6	20027->2	21122->2	50272->2
00205->0	10264->4	20032->6	21126->1	51212->2
00212->5	10267->7	20042->3	21222->2	51222->0
00222->0	10271->0	20051->7	21224->2	51242->2
00232->2	10272->7	20052->2	21226->2	51272->2
00522->2	10542->7	20057->5	21227->2	60001->1
01232->1	11112->1	20072->2	21422->2	60002->1
01242->1	11122->1	20102->2	21522->2	60212->0
01252->5	11124->4	20112->2	21622->2	61212->5
01262->1	11125->1	20122->2	21722->2	61213->1
01272->1	11126->1	20142->2	22227->2	61222->5
01275->1	11127->7	20172->2	22244->2	70007->7
01422->1	11152->2	20202->2	22246->2	70112->0
01432->1	11212->1	20203->2	22276->2	70122->0
01442->1	11222->1	20205->2	22277->2	70125->0
01472->1	11224->4	20207->3	30001->3	70212->0
01625->1	11225->1	20212->2	30002->2	70222->1
01722->1	11227->7	20215->2	30004->1	70225->1
01725->5	11232->1	20221->2	30007->6	70232->1
01752->1	11242->4	20222->2	30012->3	70252->5
01762->1	11262->1	20227->2	30042->1	70272->0
01772->1	11272->7	20232->1	30062->2	

Neighborhoods are read as follows (rotations are not listed):

```
   T
L  C  R      ==>      I
   B
```

Figure 22: transition rules for lagton loop

References

[1] Theory of self reproducing automata -Von neumann

[2] Self replication of colloidal clusters- M Brenner, Z Zeravcic

[3] Minimal model of self-replicating nanocells: a physically embodied information-free scenario Harold Fellermann and Ricard V. Sole

[4] Physical interpretation of the partition function for colloidal clusters Ellen D. Klein, Rebecca W. Perry, and Vinothan N. Manoharan

[5] Design of conditions for self-replication Sumantra Sarkar * and Jeremy L. England