Generalized Stochastic Simulation Algorithm for Artificial Chemistry

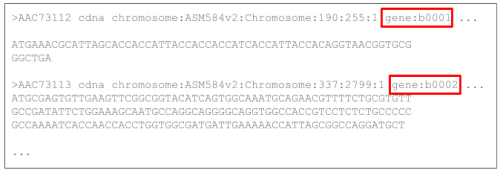
Gillespie in artificial chemistry

Hedi Soula

Sorbonne University

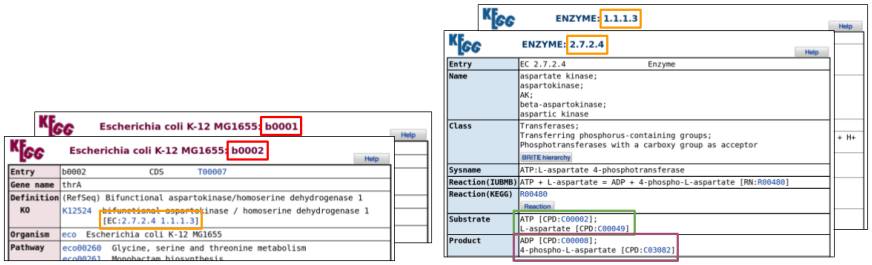
September 25 2018

Bacterial reaction graphs



cDNA fasta file

(a) Step 1: From cDNA fasta file, get gene labels.



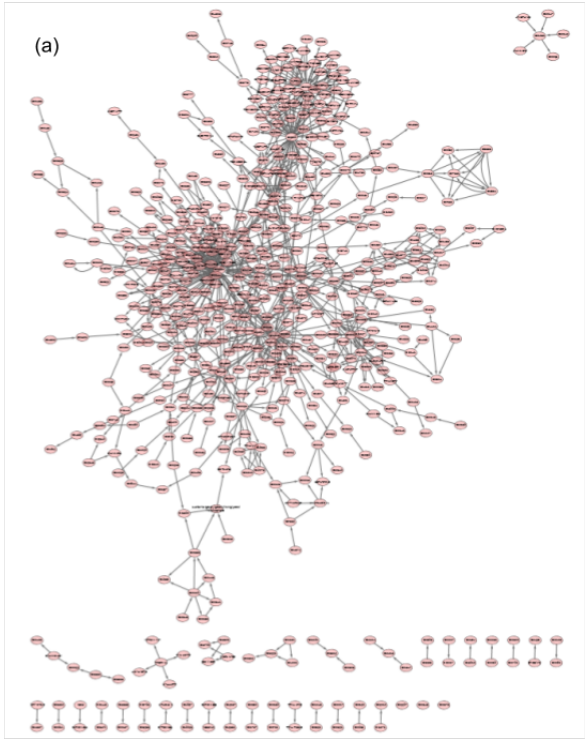
KEGG gene entries

KEGG enzyme entries

(b) Step 2: From KEGG gene entries, get EC codes.

(c) Step 3: From KEGG enzyme entries, get substrates and products.

(d) Step 4: Build graph and filter ubiquitous metabolites. Ubiquitous metabolites : H2O, ATP, ADP, NAD+, NADH, NADPH, NADP+, CO2, ammonia, sulfate, thioredoxin, phosphate, PPi, H+.



Artificial Chemistry

Motivation

- (Real) Chemistry is ... difficult
  - Maybe artificial chemistry is easier
  - can model interesting properties (for e.g. space/diffusion)
- (Real) Life is ... chemistry
  - *Artificial* Life should have ... artificial chemistry (AC)

What is done (usually)

- chemistry is prescribed : small dimension (small # of reactions)
- straightforward : chemistry graph is simple
- (somewhat) unrealistic
  - transition energies are ignored (easier)
  - mass conservation is sloppy
  - $A + B \mapsto C$ and $C \mapsto A$

AC : what is expected

1. large : should have a huge number of reactions
2. energy : all reactions are not possible
3. mass : reversibility should not be hacked
4. open : we don't know all reactions/molecules

Hutton Artificial Chemistry (1)

Hutton J.

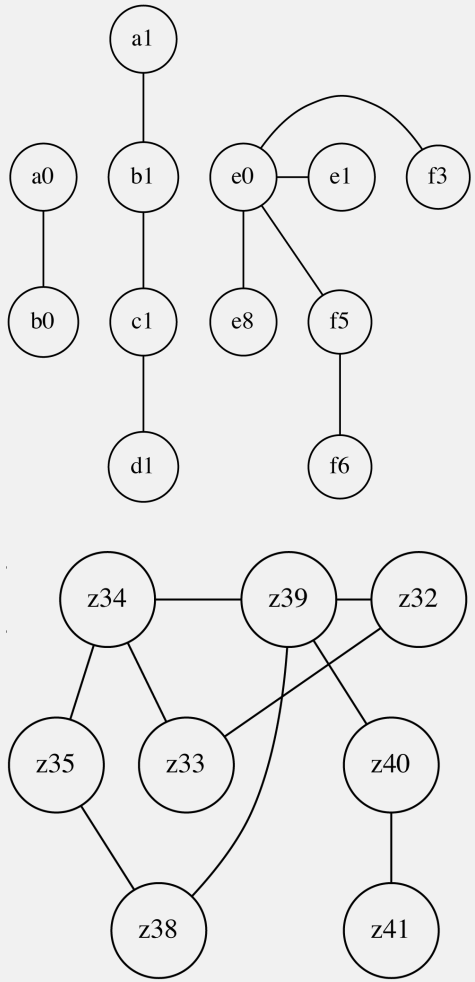
- several papers on his AC scheme
- only one follow up
- straightforward : chemistry is simple
- very complex : chemistry reactions network can be extremely complex
- however (somewhat) unrealistic

- Hutton, Tim J. Evolvable self-replicating molecules in an artificial chemistry. Artificial life 8.4 (2002) : 341-356.
- Hutton, Tim J. Evolvable self-reproducing cells in a two-dimensional artificial chemistry. Artificial life 13.1 (2007) : 11-30.
- Hutton, Tim J. A functional self-reproducing cell in a two-dimensional artificial chemistry. ALIFE9. 2004.
- Hutton, Tim J. The organic builder : A public experiment in artificial chemistries and self-replication. Artificial life 15.1 (2009) : 21-28.

Hutton Artificial Chemistry (2)

Main properties

- molecules are graphs
- nodes are domains
- edges are chemical bonding
- domains have a fixed type and a changing state
  - described by a pair $(t|s)$
  - usually a letter and a number $a0$, $b1$ ...
- any domain can have any numbers of links
- the chemistry is composed of fully connected subgraphs



Hutton Artificial Chemistry (3)

Geometry and Physics

- HAC is *2D*
- spatially resolved
- each domain has an id/position
- links are *springs* (however weird coding)
- (vaguely) Brownian in viscous medium

Hutton Artificial Chemistry (3)

Reactions

Reactions are of the form :

$$(t_1|s_1)(.|+)(t_2|s_2) \mapsto (t_1|s_3)(.|+)(t_2|s_4)$$

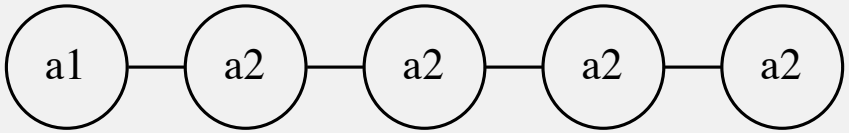
- $(t_i|s_j)$ is a domain (within a molecule)
- $(.|+)$: is either linked $.$ or no (collision) $+$
- note that type is unchanged
- modification are only local (other links unmodified)
- conformation $.$ is performed instantaneously
- reaction $+$ is performed instantaneously upon collision
- when conflict : chose at random

Hutton Artificial Chemistry (4)

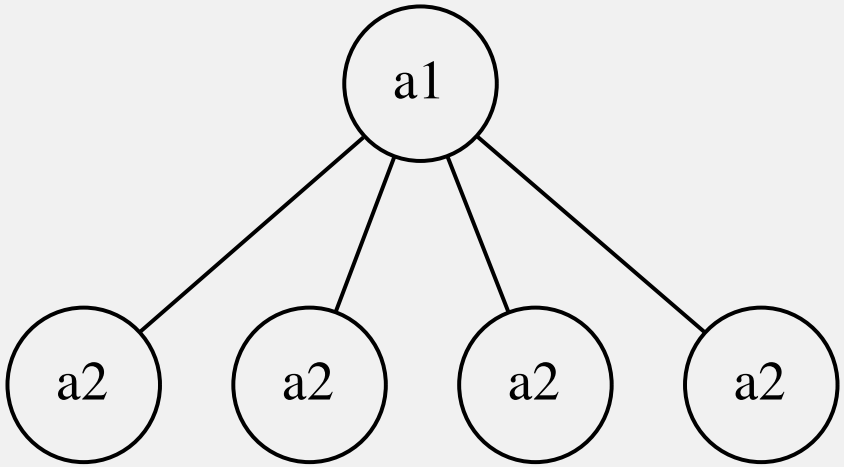
Examples

Starting a reactor with several $a0$s and 1 (one) $a1$ with

$a0 + a1 \mapsto a1.a2$ yields



whereas $a0 + a1 \mapsto a2.a1$ yields



Hutton Artificial Chemistry (5) : replication

Main properties

Using wildcards ...

R9 : $e8 + e0 \rightarrow e4e3$

R10 : $x4y1 \rightarrow x2y5$

R11 : $x5 + x0 \rightarrow x7x6$

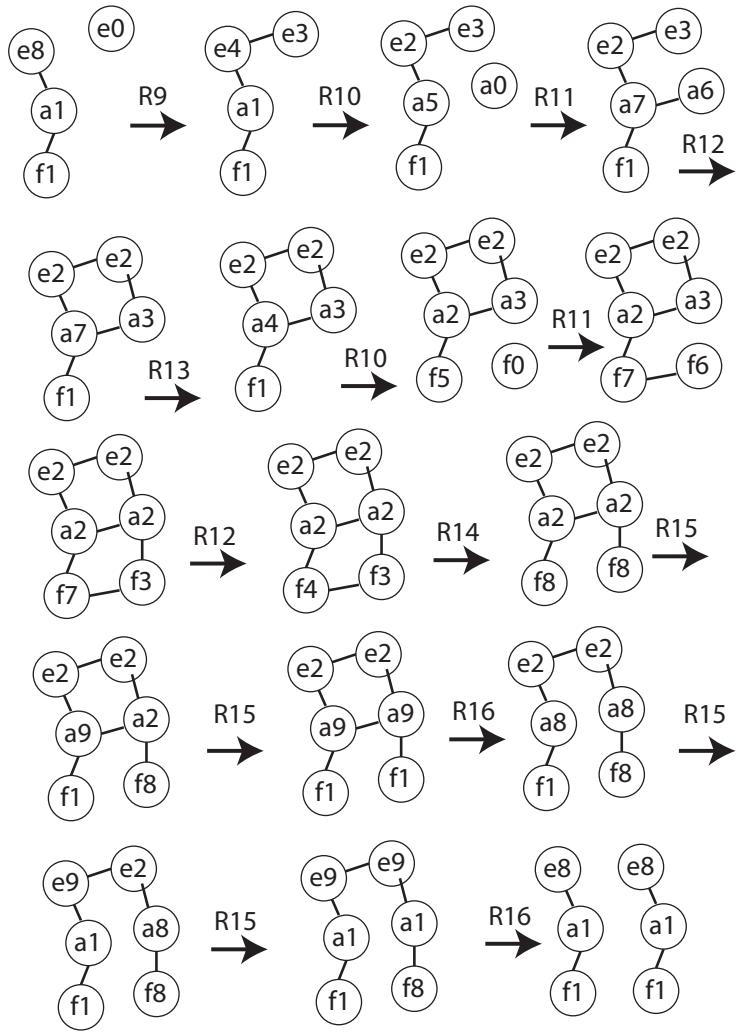
R12 : $x3 + y6 \rightarrow x2y3$

R13 : $x7y3 \rightarrow x4y3$

R14 : $f4f3 \rightarrow f8 + f8$

R15 : $x2y8 \rightarrow x9y1$

R16 : $x9y9 \rightarrow x8 + y8$



Hutton Artificial Chemistry (6)

Pros and Cons

- Pros :
  - very general
  - 2D and nice
  - mass conservation
- Cons :
  - 2D and nice
  - absurdly long
  - no reaction rates
  - *intelligently* designed

STAARC : STochastic Atom-based ARtificial Chemistry

1. Getting rid of space
2. Same data structure (without space)
3. reactions with rates
4. SSA formalism (Gillespie)
5. slightly involved (but not too much)
6. `github.com/hsoula/staarc`

Gillespie in AC

Modifications

- Reaction have a rate :
  $(t_1|s_1)(.|+)(t_2|s_2) \mapsto (t_1|s_3)(.|+)(t_2|s_4) : \lambda$
- For each $+$ reaction :
  - We count the number of $n_1 = (t_1|s_1)$ and $n_2 = (t_2|s_2)$
  - minus the number of $n_{12} = (t_1|s_1).(t_2|s_2)$ already linked
  - the propensity is $a = \lambda\,(n_1 * n_2 - n_{12})$
- For each $.$ reaction :
  - we count the number of $n_{12} = (t_1|s_1).(t_2|s_2)$ already linked
  - the propensity is $a = \lambda n_{12}$
- then classical Gillespie algorithm
- once a reaction is selected we apply the modification to a given pair (selected at random uniformly)

Gillespie

First step

compute all propensities a_i Compute the combined rates of all reactions

$$a_0 = \sum_{i=1}^{R} a_i$$

To compute the time of the next reaction, draw a random number

$$\tau = -\log(rn)/a_0$$

This answers the when

What reaction

To compute the what, we choose randomly the equation weighted by their relative weight i.e

$$\Pr(\text{next reaction is} i) = \frac{a_i}{a_0}$$

in practice, shoot a random number uniform rn in $[0, 1]$ and find r such as :

$$\sum_{i=1}^{r-1} a_i \leq a_0 rn < \sum_{i=1}^{r} a_i$$

STAARC : STochastic Atom-based ARtificial Chemistry

Properties

1. now all reactions have rates : real chemistry
2. simulated 3D well mixed medium
3. almost all HAC properties conserved
4. we can simulate diffusion (reaction rate for collision)
5. way faster computation

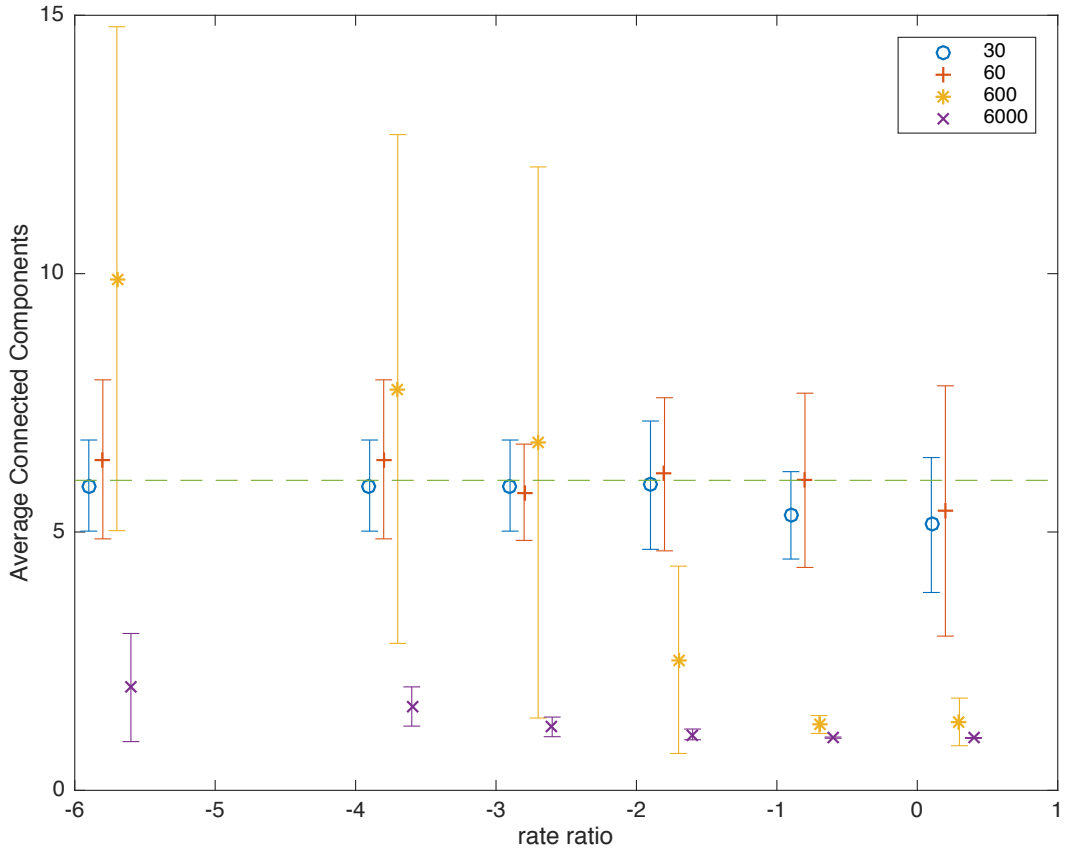
This turn a local resolution to a global one

- well mixed medium
- with infinite crowding

not intuitive ...

Replication : molecules size

Simple experiment : the replication according to the initial number of particles ($n \in \{30, 60, 600, 6000\}$) and λ ratio (between collision reaction $+$ and conformation reaction .)



STAARC : replication

Properties

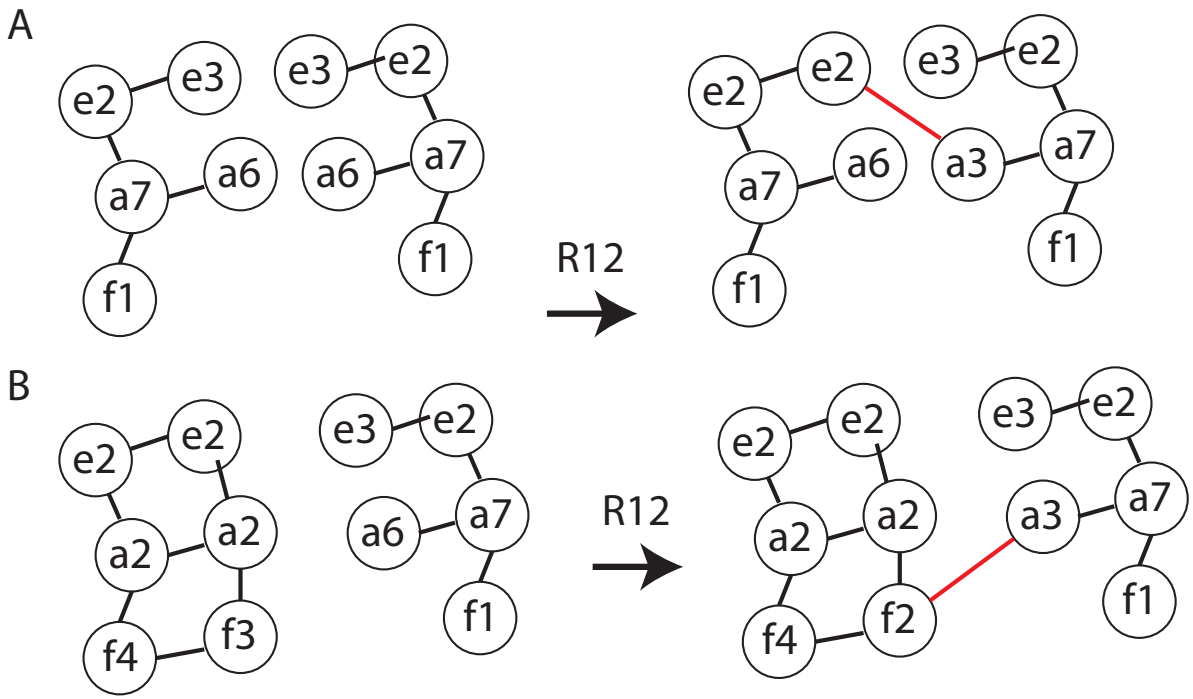
- Not very resilient - original sequence quickly lost
- Scales very badly

Not that intelligently designed :)

Race conditions

- Race condition – diffusion vs conformation
- Race condition – concurrent replications

STAARC : replication

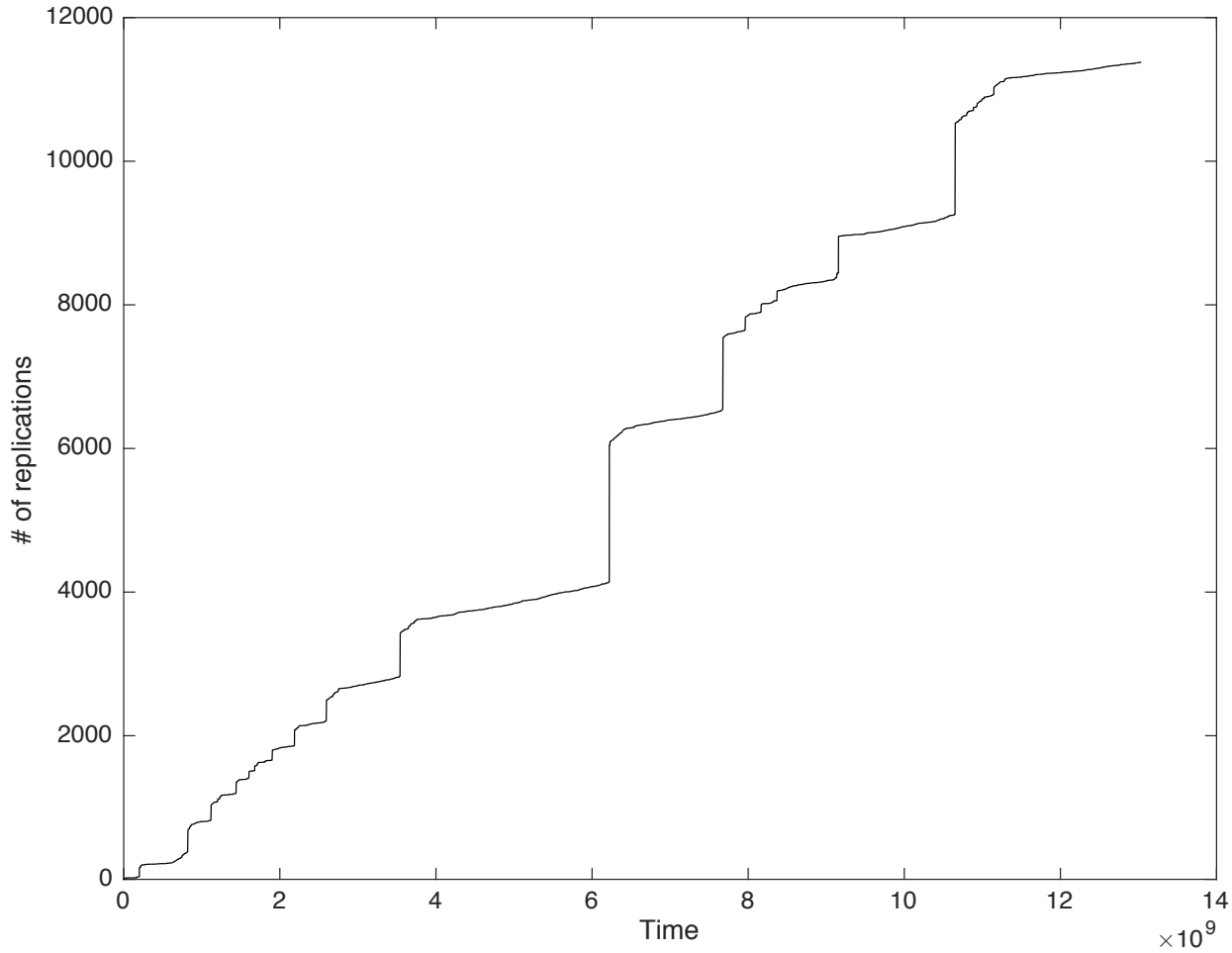


STAARC : long replication

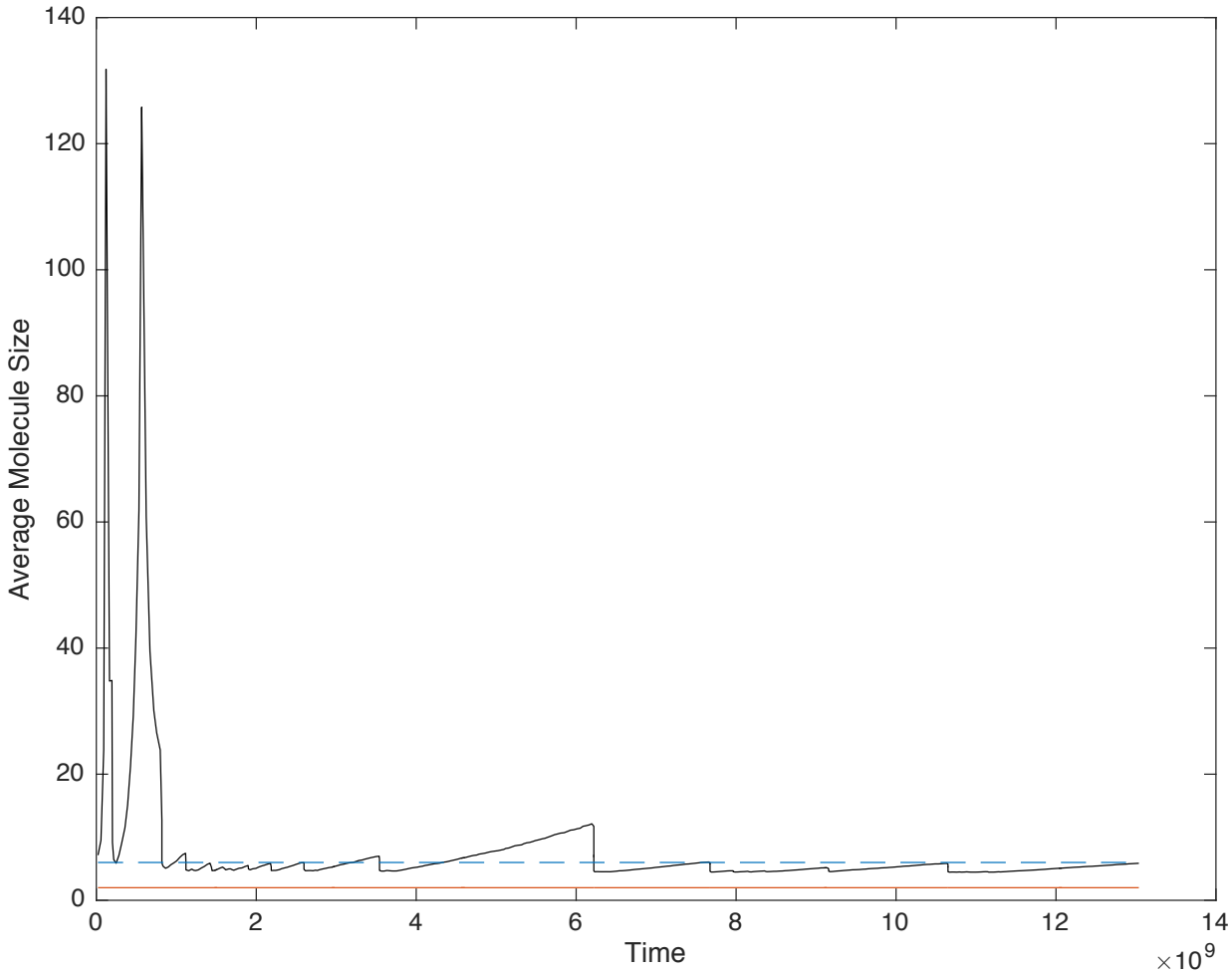
Long replication experiment

- Starting with size 6 replication seed
- Production and degradation
  - $(x|0) \rightarrow \varnothing$
  - $\varnothing \rightarrow (x|0)$
- Let the simulation for 450,000 reactions
- Compute average size and number of divisions

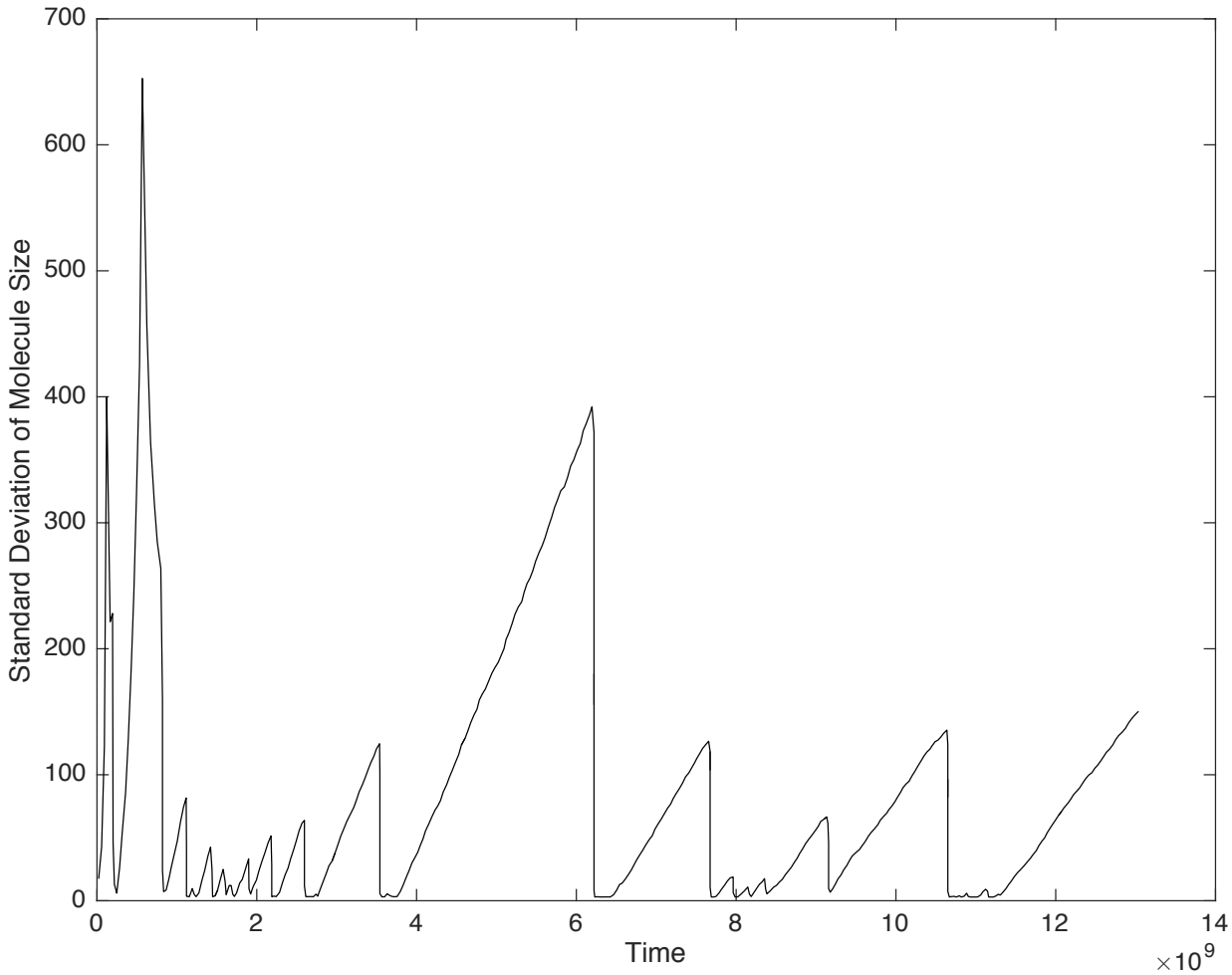
Long replication : replication event



Long replication : average size



Long replication : standard deviation



STAARC : long replication

Long replication experiment

- Replication is extremely stable
- Division occurs in bumps
- The limit size is ... 6
- Big compounds created transiently

Random Chemical World

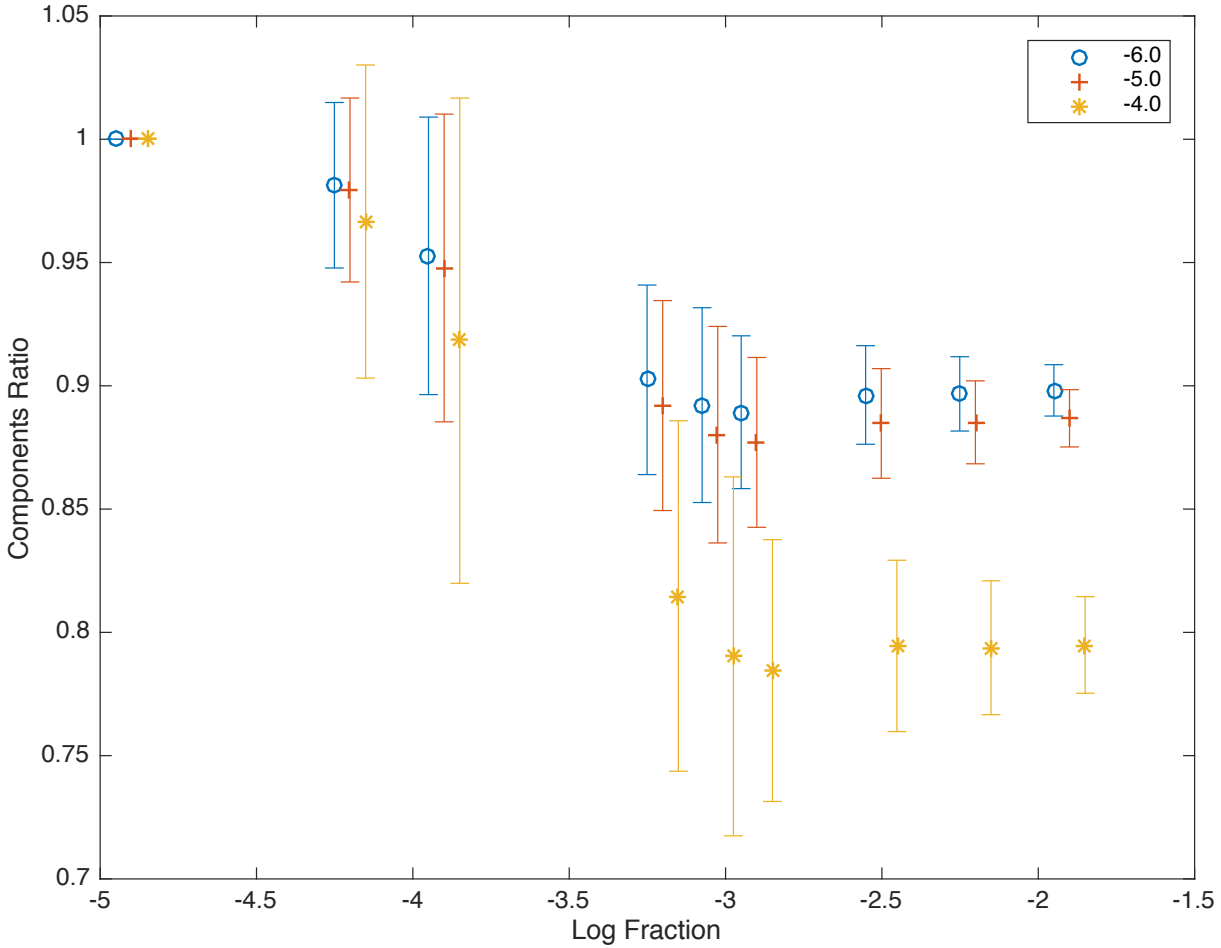
Random generation of reaction

- available types : $\{a, b, c\}$
- maximum number of state is 5.
- compute all the possible reactions
  - with no production nor degradation.
  - keeping only a fraction $p \in [0, 1]$ randomly
- starting $N = 10,000$ particles $(t|s)$ with $t \in \{a, b, c\}$ and $0 \leq s \leq 4$

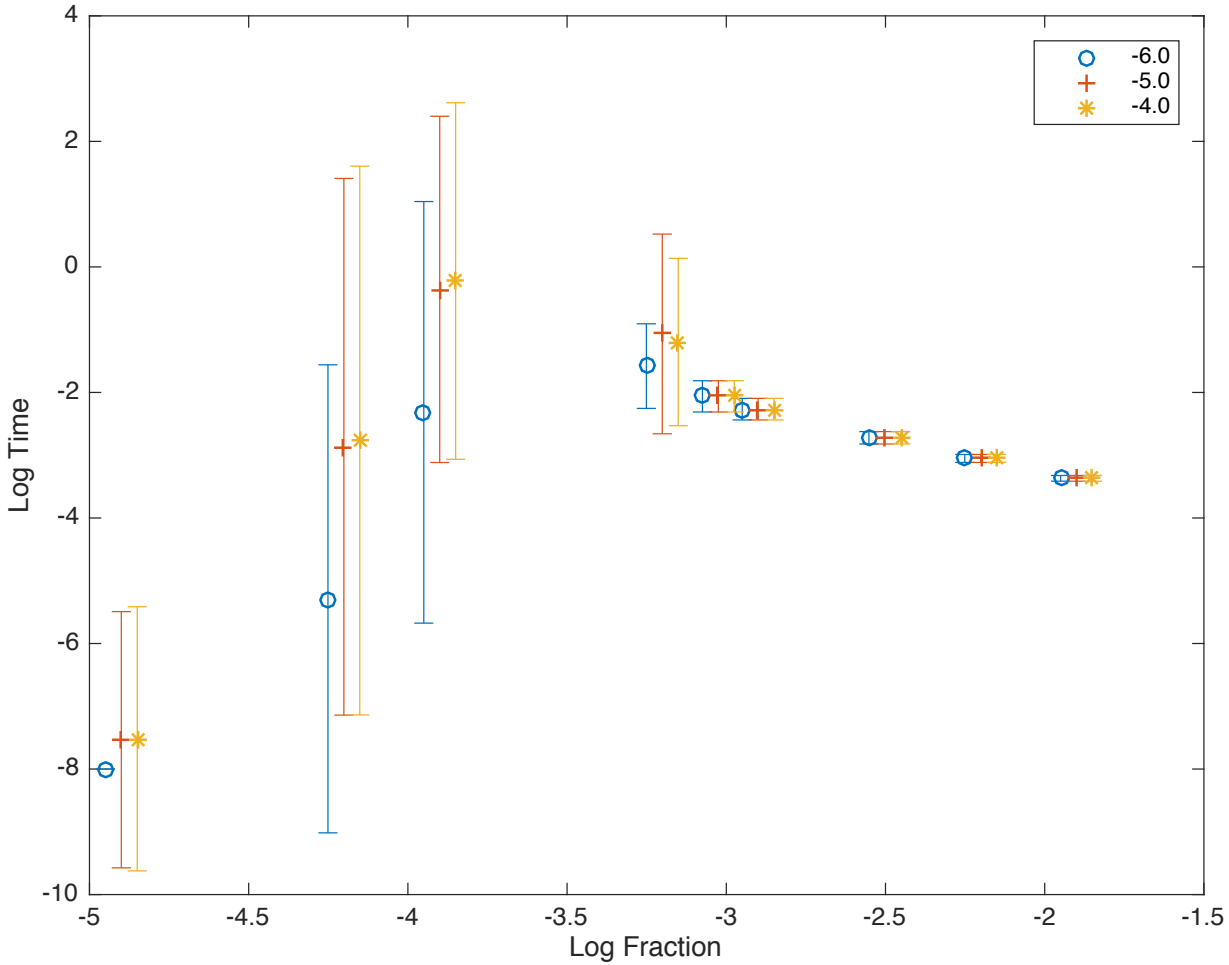
Experiment

- Maximum of 2,000 reactions.
- Compute the time needed
- The ratio of number of molecules

Random Chemical World



Random Chemical World



Random Chemical World

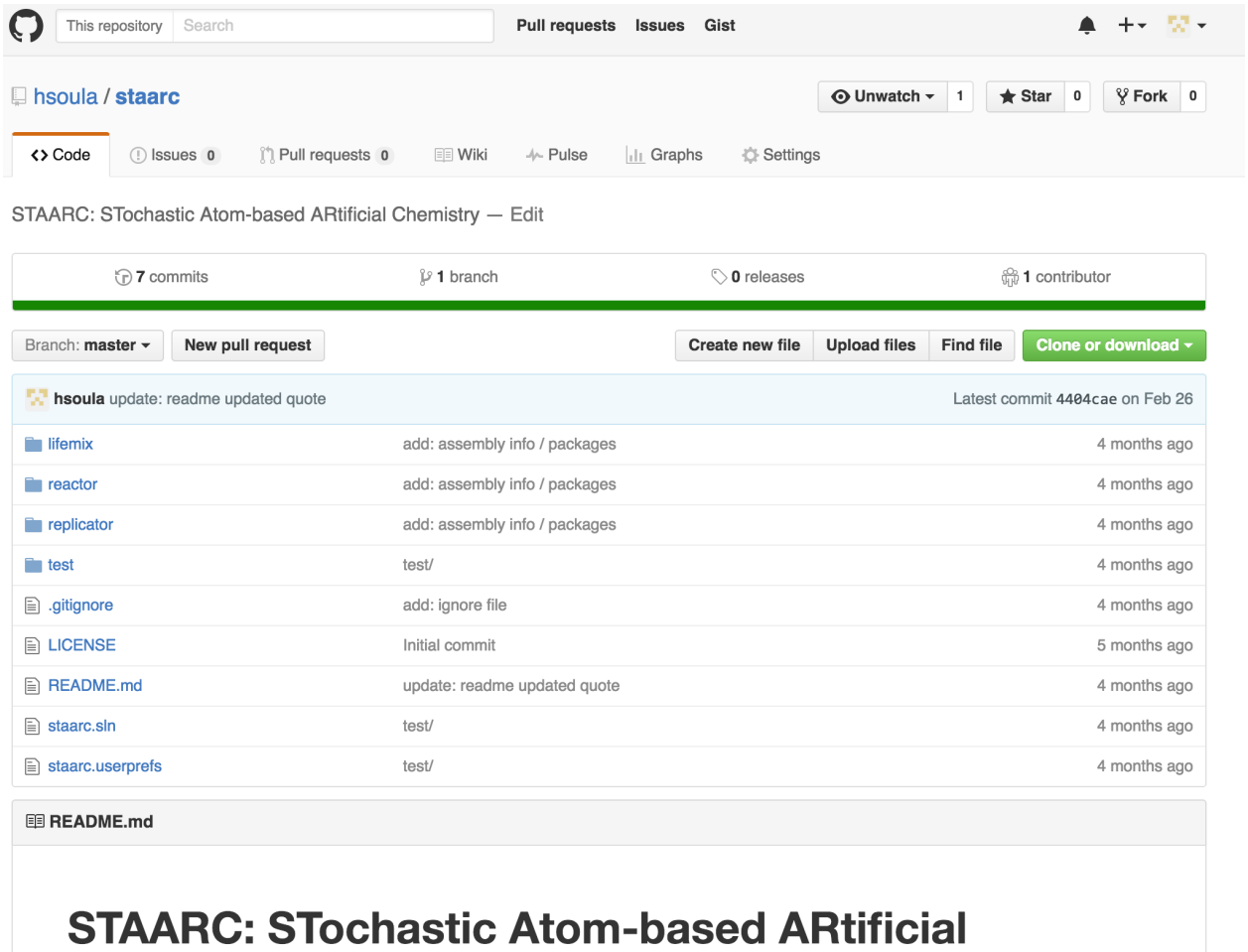
Phase transition

- Huge variability in the middle
- in molecules created
- in the time to the end of reactions
- this middle is for a low number of reactions

Drawbacks

- Results not conclusive
- Better random selection of reactions
- Common elements

STAARC



github.com/hsoula/staarc

Perspectives

Code

- Open-Endedness procedural enzyme
- Tri-molecular reactions
- Graph-based random chemistry
- 'Metabolic network'-like graph

Projects

- Explore Random Chemistry
- Use it in an evolutionnary set-up
- Artificial 'bacterial world'

thanks

hedi.soula@upmc.fr // sites.google.com/site/hsoula/

Soula HA, Generalized Stochastic simulation algorithm for Artificial Chemistry Proceedings of ALIFE XV, 2016 Cancun, Mexico