

Competition and space structure in an efficient population genetics model of carcinogenesis

Tommaso Tarchi

September 23, 2023

University of Trieste

Introduction

Title: Efficient simulation under a population genetics model of carcinogenesis

Authors: Zhu Tianqi, Hu Yucheng, Ma Zhi-Ming, Zhang De-Xing, Li Tiejun, Yang Ziheng

Year: 2011

Journal: Bioinformatics

Volume: 27

Number: 6

Pages: 837-843

Genotypic space

From the paper: "It is now generally accepted that the rate-limiting step in the process of carcinogenesis is the accumulation of new mutations. [...] Here we develop a hybrid algorithm that can efficiently simulate the evolutionary processes of large populations."

Genotypic space

From the paper: "It is now generally accepted that the rate-limiting step in the process of carcinogenesis is the accumulation of new mutations. [...] Here we develop a hybrid algorithm that can efficiently simulate the evolutionary processes of large populations."

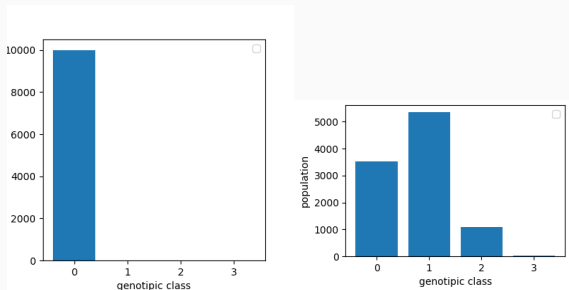


Figure 1: example of population distribution in the genotypic space at $t = 0$ (left) and $t = \text{waiting time}$ (right).

Fixed-population assumptions

- Fixed population.

Fixed-population assumptions

- Fixed population.
- Fitness landscape: $f_i = 1.01^i$.

Fixed-population assumptions

- Fixed population.
- Fitness landscape: $f_i = 1.01^i$.
- Constant and uniform mutation rate (μ).

Fixed-population assumptions

- Fixed population.
- Fitness landscape: $f_i = 1.01^i$.
- Constant and uniform mutation rate (μ).
- Possible events:
 1. a type j cell is replaced by a type j' one, with rate:

$$a_{jj'} = \frac{x_j f_{j'} x_{j'}}{\sum_{l=0}^{m-1} f_l x_l};$$

2. a type j cell mutates into a type $j + 1$ one, with rate:

$$a_j = \mu x_j.$$

Dynamic-population assumptions

- Variable population.

Dynamic-population assumptions

- Variable population.
- Fitness landscape: $f_i = 1.01^i$.

Dynamic-population assumptions

- Variable population.
- Fitness landscape: $f_i = 1.01^i$.
- Constant and uniform mutation rate (μ).

Dynamic-population assumptions

- Variable population.
- Fitness landscape: $f_i = 1.01^i$.
- Constant and uniform mutation rate (μ).
- Possible events:
 1. a type j cell is born, with rate:

$$a_j^b = \frac{\tilde{N}(t) f_j x_j}{\sum_{l=0}^{m-1} f_l x_l},$$

where

$$\tilde{N}(t)$$

is a parameter capable of controlling the population size;

2. a type j cell dies with rate:

$$a_j^d = x_j;$$

3. a type j cell mutates into a type $j + 1$ one, with rate:

$$a_j^m = \mu x_j.$$

Pseudocode:

1. Compute events rates.
2. Apply Gillespie algorithm to sample time to next event e and next event.
3. Update time to $t' = t + e$ and system's state.
4. If population of the m -th genotypic class is different from zero, set waiting time to t' and stop, otherwise go back to step 1.

Hybrid algorithm

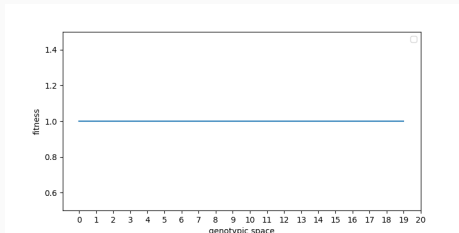
Pseudocode:

1. Compute events rates.
2. Partition events in critical and non-critical ones.
3. Determine tau-leaping step length τ .
4. Determine (by Gillespie's method) time to next critical event e .
5. If $e < \tau$, update system's state according to Gillespie algorithm, otherwise go directly to next step.
6. Let $h = \min(e, \tau)$; update system's state applying tau-leaping algorithm over time h .
7. Update time to $t' = t + h$.
8. If population of the m -th genotypic class is different from zero, set waiting time to t' and stop, otherwise go back to step 1.

My humble contribute

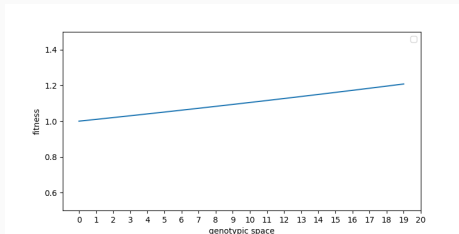
Fitness landscapes

1. Flat: $f_i = 1$



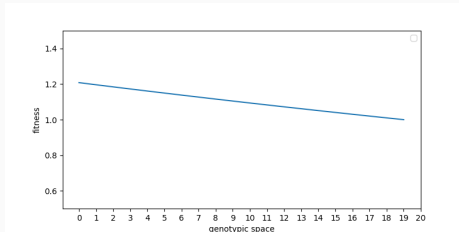
Fitness landscapes

1. Flat: $f_i = 1$
2. Static increasing: $f_i = 1.01^i$



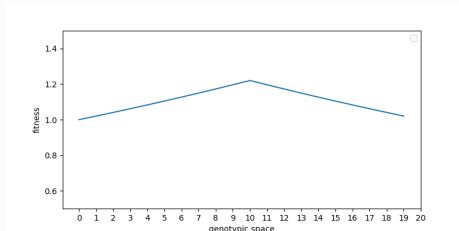
Fitness landscapes

1. Flat: $f_i = 1$
2. Static increasing: $f_i = 1.01^i$
3. Static decreasing: $f_i = 1.01^{m-i-1}$



Fitness landscapes

1. Flat: $f_i = 1$
2. Static increasing: $f_i = 1.01^i$
3. Static decreasing: $f_i = 1.01^{m-i-1}$
4. Static 'mountain': $f_i = 1.01^{m-|m-2i|}$



Modelling competition

Two **assumptions**:

- the closer two organisms in the phenotypic space, the more the competition for resources
- the phenotype is strongly correlated to the genotype

Modelling competition

Two assumptions:

- the closer two organisms in the phenotypic space, the more the competition for resources
- the phenotype is strongly correlated to the genotype

Result:

The more 'crowded' a portion of the genotypic space, the lower the fitness of the organisms in that region.

1. Flat: $f_i = 1$
2. Static increasing: $f_i = 1.01^i$
3. Static decreasing: $f_i = 1.01^{m-i-1}$
4. Static 'mountain': $f_i = 1.01^{m-|m-2i|}$
5. Static 'mountain' + dynamic: $f_i = 1.01^{m-|m-2i|} - \sum_{j=0}^{m-1} \frac{x_j}{N} \frac{|i-j|}{m}$
6. static decreasing + dynamic: $f_i = 1.01^{m-i-1} - \sum_{j=0}^{m-1} \frac{x_j}{N} \frac{|i-j|}{m}$

Spatial structure (naive)

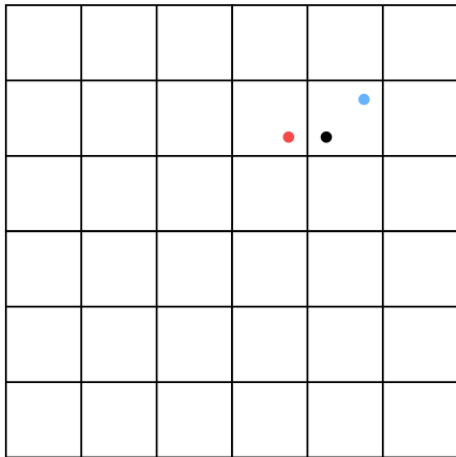


Figure 2: 'world' of the spatial model with resolution=16.

Spatial structure (less naive)

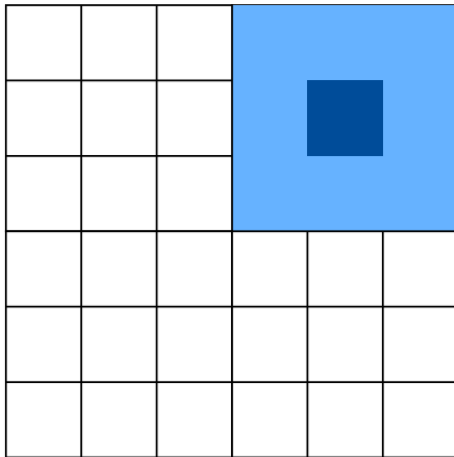


Figure 3: 'world' of the spatial model with resolution=16.

Spatial structure (less naive)

Simulation results

Reproducing results of the original paper

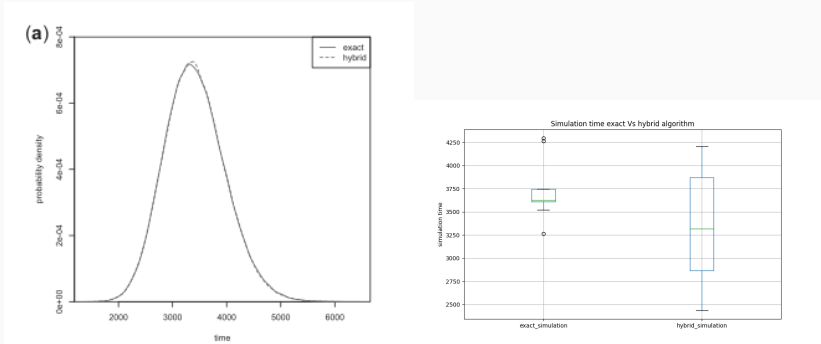


Figure 4: comparison of waiting time obtained with fixed-population exact and hybrid algorithms in original paper (**left**) and our work (**right**); parameters: $N = 10^3$, $m = 20$, $\mu = 10^{-3}$, fitness: static increasing.

Reproducing results of the original paper

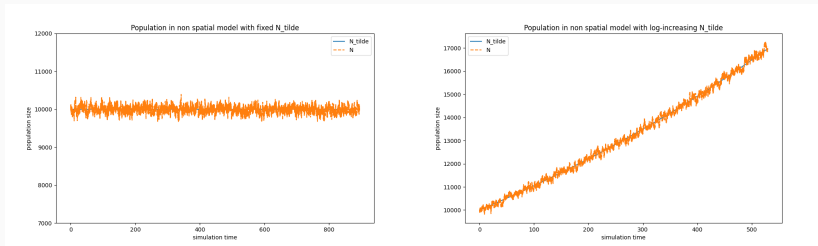


Figure 5: behaviour of N and N_{tilde} in time in our dynamic-population model with constant (**left**) and increasing (**right**) N_{tilde} ; parameters: $N = 10^4$, $m = 4$, $\mu = 10^{-4}$, fitness: static increasing.

Changing the fitness landscape

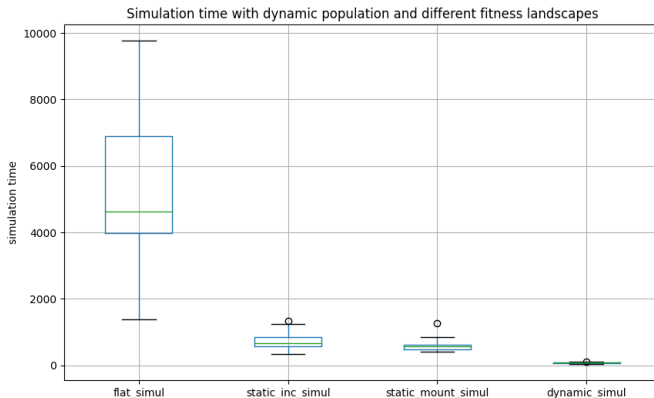


Figure 6: waiting time obtained with fixed-population hybrid model for different fitness landscapes; parameters: $N = 10^4$, $m = 4$, $\mu = 10^{-4}$; **note:** static *increasing* fitness did not converge within an acceptable time.

Changing the fitness landscape

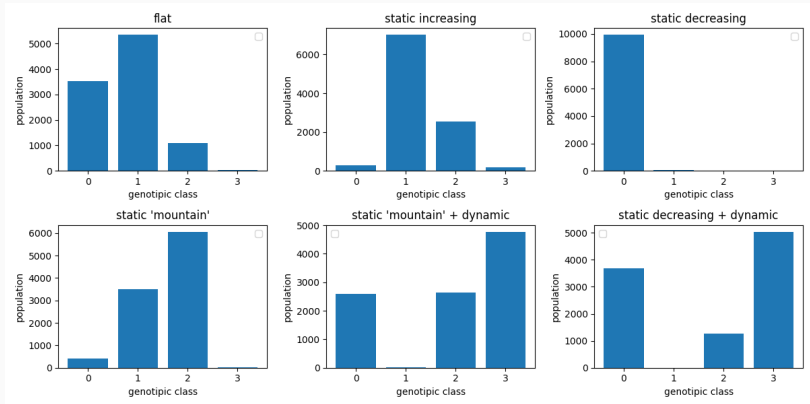


Figure 7: final state obtained with fixed-population hybrid model for different fitness landscapes; parameters: $N = 10^4$, $m = 4$, $\mu = 10^{-4}$.

Waiting time in spatial model

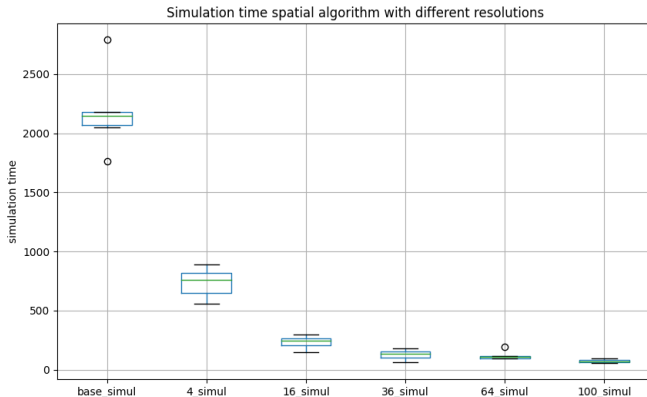


Figure 8: waiting time obtained with spatial model for different resolutions (from 1 to 100); parameters: $N_{\text{tilde}} = 10^6$ (constant), $m = 4$, $\mu = 10^{-6}$, fitness: static increasing.

Population in spatial model

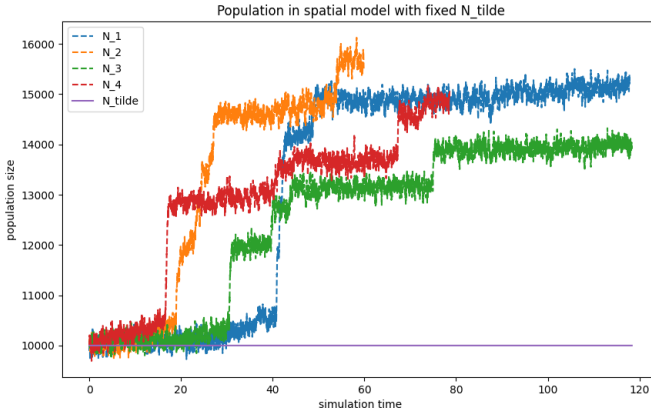


Figure 9: N 's behaviour in time for four simulations of the same spatial model; parameters: $N_{\text{tilde}} = 10^4$, $m = 4$, $\mu = 10^{-4}$, resolution = 16, fitness: static increasing.

An efficiency study

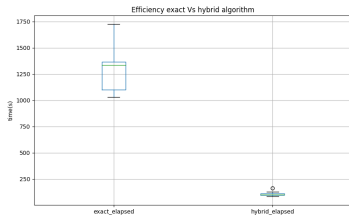


Figure 10: efficiency (in seconds) for fixed-population exact and hybrid models; parameters: $N = 10^3$, $m = 20$, $\mu = 10^{-3}$, fitness: static increasing.

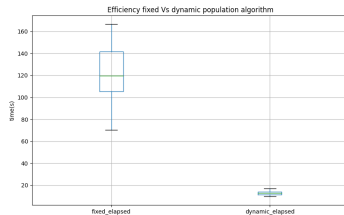


Figure 11: efficiency (in seconds) for fixed-population and dynamic-population hybrid models; parameters: $N(N_tilde) = 10^3$, $m = 20$, $\mu = 10^{-3}$, fitness: static increasing.

An efficiency study

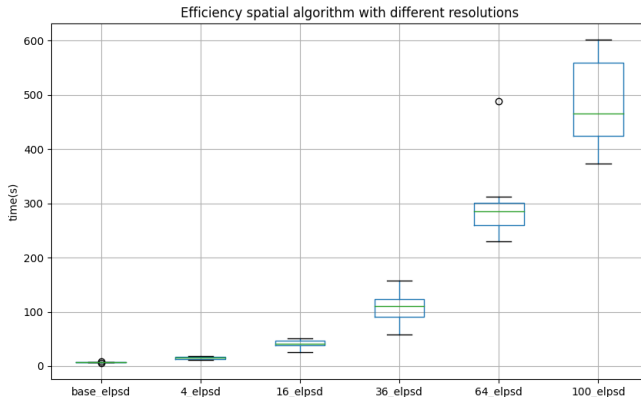


Figure 12: efficiency (in seconds) of spatial model for different resolutions (from 1 to 100); parameters: $N_{\text{tilde}} = 10^6$ (constant), $m = 4$, $\mu = 10^{-6}$, fitness: static increasing.

