



Evolution of Metabolic Networks

Eleni Giannoulatou¹ and Jotun Hein²
 Life Sciences Interface Doctoral Training Centre¹,
 Department of Statistics²,
 University of Oxford
 {giannoul,hein}@stats.ox.ac.uk



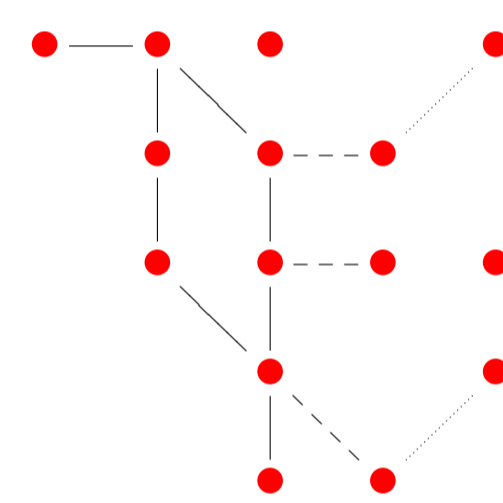
1. Introduction

METABOLISM is one of the most complex cellular processes and a basal system for maintaining life of all organisms. Most of the components of a metabolic network can be identified and analysed due to large-scale sequencing projects and genome annotation efforts. However, understanding the organisational principles of the metabolism of living organisms is a major challenge in network biology [1]. The process of evolution of these networks can be used in order to justify their principal design. Although models of network growth have been introduced, they do not apply directly to metabolic networks where the dynamics are determined by the gain or loss of enzymatic reactions.

2. A stochastic model of evolution

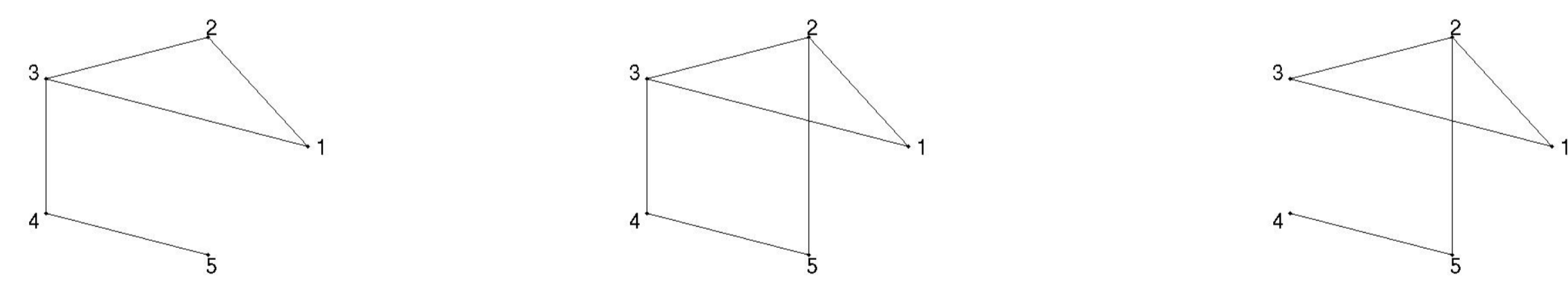
A metabolic network is assumed to evolve down the branches of a phylogenetic tree according to a continuous time Markov process with finite states.

- Given a set of fixed metabolites, and a set of present reactions between them, M (solid and dashed edges) one of the edges of the graph can be either created with rate μ if it was absent or deleted with rate λ .
- The metabolic network contains a core of reactions (solid edges), which cannot be changed under evolution.
- An edge cannot be deleted if it is a bridge in the graph; its removal will create two disconnected subgraphs.
- The set of removable edges, $D(M)$, is defined by the dashed edges.
- The set of addable edges, $A(M)$, is defined by the dotted edges.



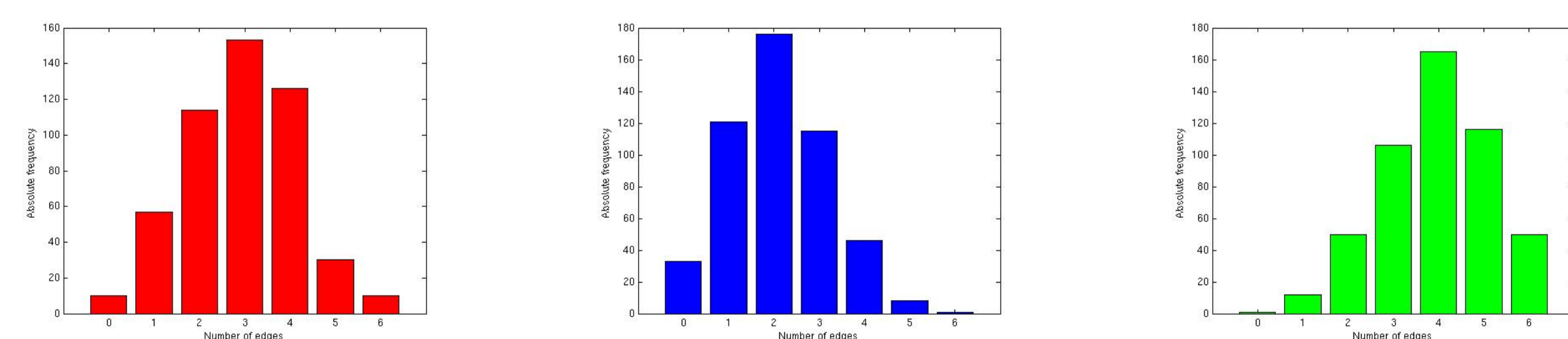
$$\frac{dP(M)}{dt} = \lambda \sum_{M' \in A(M)} P(M') + \mu \sum_{M'' \in D(M)} P(M'') - P(M)(\lambda A(M) + \mu D(M))$$

3. Simulation results: Evolution of connectivity



(a) Initial network at t_m (b) Intermediate state (c) Evolved network at t_{m+1}

Figure 1: $\lambda = 3 * 10^{-3}$, $\mu = 5 * 10^{-3}$, $t_m = 0$ and $t_{m+1} = 10$.



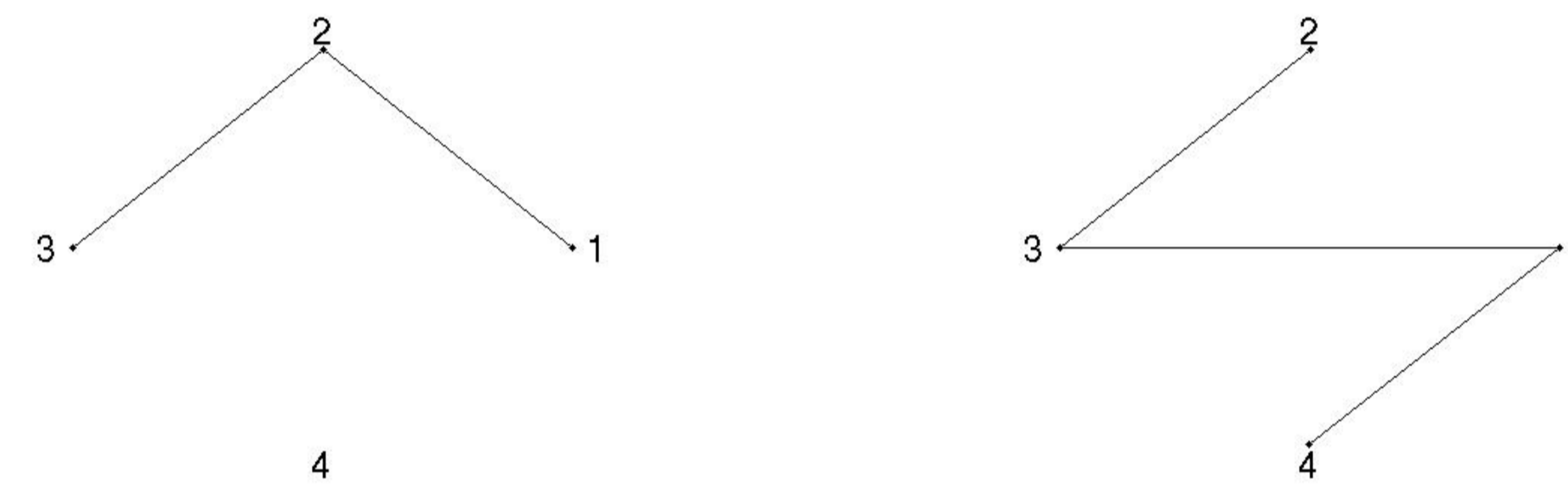
(a) $\lambda = \mu = 1$ (b) $\lambda = 1, \mu = 2$ (c) $\lambda = 2, \mu = 1$

Figure 2: Frequency distribution for network of size 4. $t_m = 0$ and $t_{m+1} = 100$.

4. Maximum Likelihood Estimates of phylogeny of networks

The probability of obtaining a given set of networks at the tips of a given tree can be computed using the above time-reversible model that specifies the probability that network N_1 changes to network N_2 during evolution along a segment of the tree of length (in time) t . In the case of metabolic networks:

- The phylogenetic tree between organisms can be given by the sequence analysis, therefore it can be assumed.
- The likelihood approach is used to estimate parameters concerning the gain and loss of enzymatic interactions [2].



(a) N_1 (b) N_2

Figure 3: Two simple networks used for maximum likelihood estimation of parameters.

The likelihood of a pair of networks, N_1 and N_2 , evolved from a common ancestor N by divergence time t is equivalent to the probability of one network evolving to the other, in twice the time that separates the ancestor from the two descendants [3]:

$$P_t(N_1, N_2) = P_\infty(N_1)P_{2t}(N_2|N_1)$$

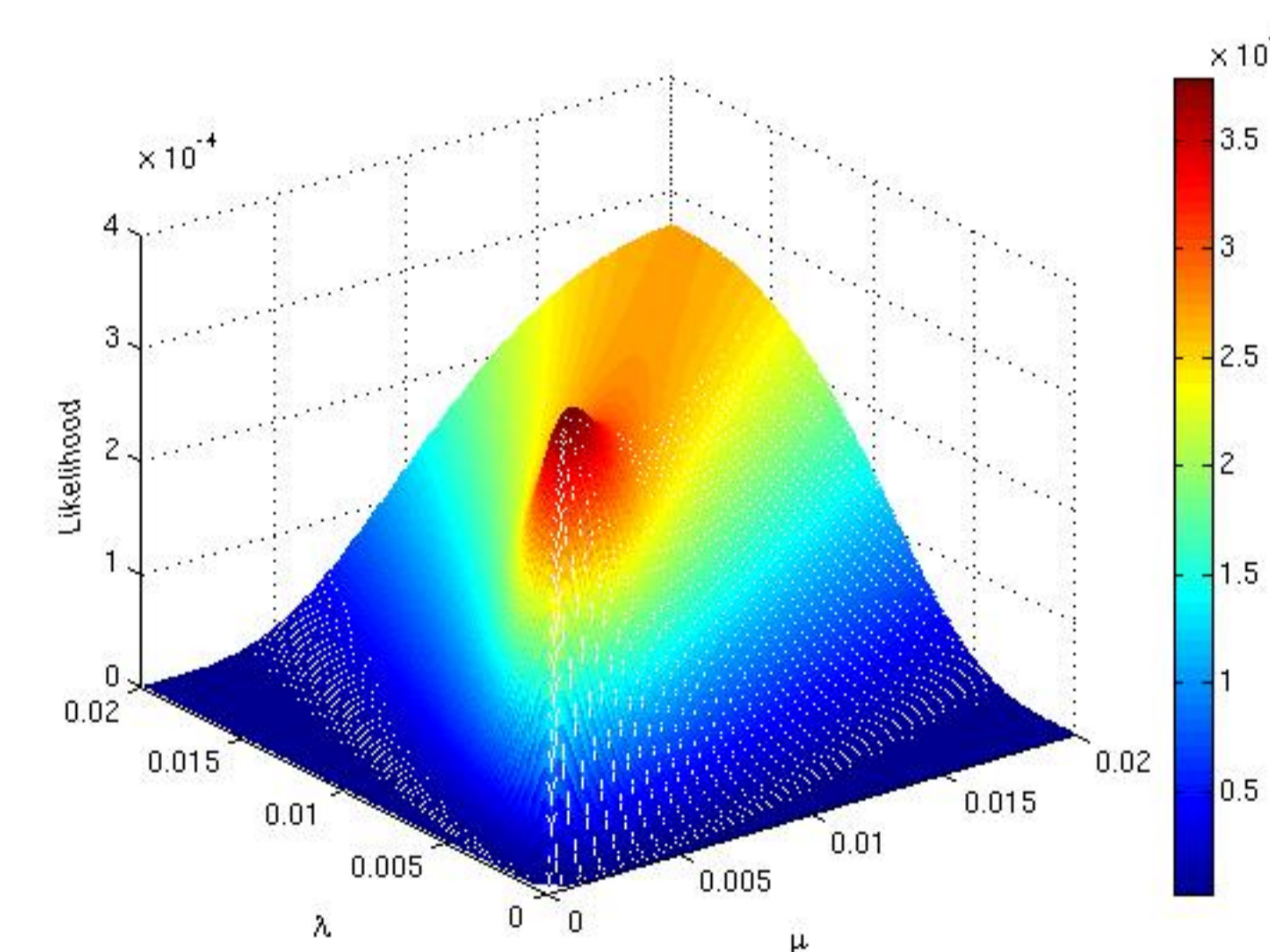


Figure 4: Likelihood curve for estimating insertion and deletion rates for the evolution between networks N_1 and N_2 .

References

[1] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai and A. Barabasi, The large-scale organization of metabolic networks. *Nature* 407(6804):651, 2000.
 [2] C. Wiuf, M. Brameier, O. Hagberg and M. Stumpf. A likelihood approach to analysis of network data. *Proc Natl Acad Sci USA* 103(20):7566, 2006.
 [3] J. Felsenstein, Evolutionary trees from DNA sequences: A maximum likelihood approach. *J Mol Evol* 17(6):368, 1981.

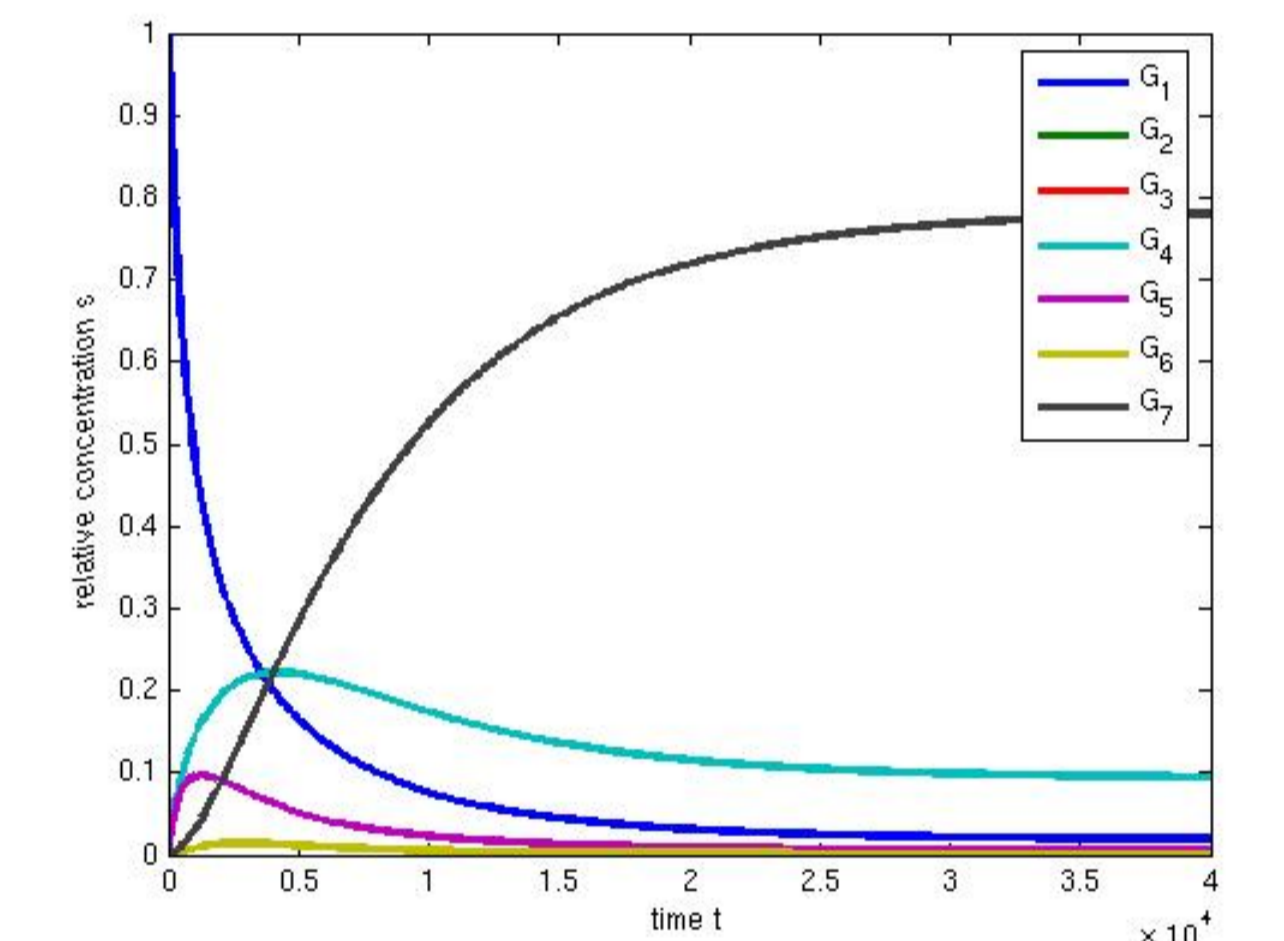
5. Model with Fitness-dependent link dynamics

By quantifying the concentration of species that have a specific structure of metabolism, we are interested in its rate of change. A species is assumed to mutate to other species by insertions or deletions of edges in their metabolism, causing the concentration of the rest of the species to change. We define:

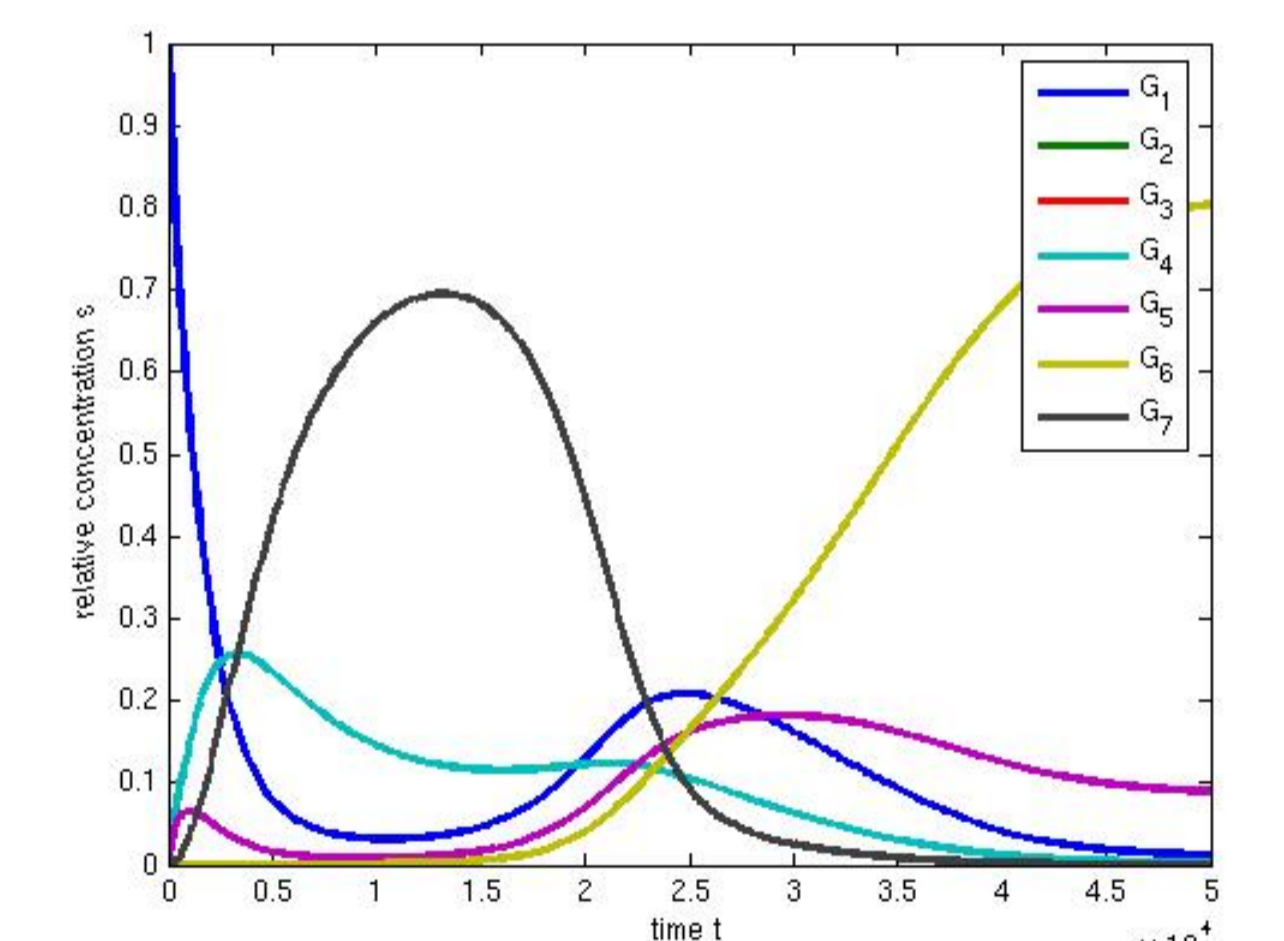
- $m(s, s')$: the instantaneous mutation rate from species with network s to species with network s' that depends on the structure of s and s' and the rates of insertion and deletion of edges.
- $f(s)$: the fitness of species with network s that depends on the structure of s and the environment.

$$\frac{ds}{dt} = f(s)s(1 - \sum_{s'} m(s, s')) + \sum_{s'} m(s', s)f(s')s'$$

Network	Graph Structure
G_1	
G_2	
G_3	
G_4	
G_5	
G_6	
G_7	



(a) $A = 0.001, B = 0.002, \lambda = 0.01, \mu = 0.1$



(b) $A = 10^{-7}t, B = 0.002, \lambda = 0.01, \mu = 0.1$

Figure 5: Relative concentration of species with different metabolisms over time.

6. Conclusion

By modelling the evolutionary process of the metabolic networks, important parameters can be inferred, such as the rate of gain or loss of enzymatic reactions. A simple stochastic model is presented in order to make estimates for the phylogeny of simple metabolisms based on likelihood approaches. As environment plays a key role for adaptation and evolution, we also presented a model that accounts for the fitness of the organism changing its metabolism. Using this model we can see how the environment can affect the dynamics by changing the concentration of key metabolites.