

A survey of the modified Moran process and evolutionary graph theory

Josep Díaz* Dieter Mitsche†

Abstract

Historically, the field of evolutionary dynamics has been a relevant tool in modeling the evolution of processes in areas as biology, sociology and others. In this work we first present a survey of the field from the XVIII century to mid XX century, until the genetic models of Wright-Fisher and Moran. In the second part we survey recent mathematical and algorithmic results on the adaptation of the Moran process due to Lieberman, Hauert and Nowak [41]. We survey the results related to the absorption probability and the absorption time.

1 Introduction

Evolutionary dynamics is a wide field dealing with the mathematical foundations that model the evolution of social, demographic and biological systems, see for example [58]. In the present work, we survey a particular model, the modified Moran model, or as we denote it from now on, the LHN-model. The main reason to study that model is due to the fact that María José Serna has contributed to that specific area (see [14, 15, 16]). To frame historically the model, we begin by presenting a brief sketch of the historical development of population dynamic models. Next, we survey the basic definitions and main results for the LHN model. For completeness, in Section 2, we give a minimalist bibliographical footnote of mentioned researchers. For the reader interested in learning more about the historical development of population dynamics, there are several nice books, see for example [3]. Regarding the technical survey on the LHN model, there is a previous nice survey on evolutionary graph theory [62], which makes emphasis on the applications to

*Departament of Computer Science, Universitat Politècnica de Catalunya. Email: diaz@cs.upc.edu Supported by TIN2017-86727-C2-1-R.

†Institute Camille Jordan (UMR 5208), Université Jean Monnet, Email: dmitsche@univ-st-etienne.fr Supported by IDEXLYON of Université de Lyon Programme Investissements d’Avenir (ANR16-IDEX-0005).

game theory. We tried to minimize the intersection between both surveys, although it is impossible not to have an intersection. Also, in the section of related work in some of the recent papers on the topic, there are nice surveys of more recent results. In particular, Giakkoupis [24] and Goldberg et al [25] give updated information on the recent progress on absorption of the Lieberman et al. model.

2 A historical perspective of the mathematical models in evolutionary biology

During the XVIII century there were two important contributions to the study of population dynamics. The first contribution is the book by T.R. Malthus¹ [44], where he modeled the social effects of the population growth, by showing the inexorability of poverty, as populations increase faster than food. The second contribution is a deterministic mathematical model for the growth of a population, proposed by Euler² [19]. The model is a deterministic one that uses ordinary differential equations. As we will see, both works would have a non-negligible influence in the posterior development of the field of population dynamics.

The XIX century was a period of important advances in population dynamics and in particular genetics. Lamarck³ in his book *Philosophie Zoologique* [40], developed a particular theory of inheritance of evolutionary trends for humans and animals. He put forward his belief that the "environment" could produce evolutionary changes in genetic inheritance. Later Verhulst⁴ introduced his *logistic equation* to represent the evolution in the number of individuals in a population: Initially the growth of individuals follows approximately a geometric progression, then as the population starts to be saturated, the growth follows an arithmetic progression, until the growth stops [65]. The deterministic mathematical models for the evolution of populations produced good approximations for large populations, but for the dynamics of small population models it turned out to be necessary to resort to use discrete and random models.

Bienayme in 1845 [5], and in 1987 Galton and Watson [23, 22]⁵ developed a

¹Thomas Robert Malthus (1766-1834), English scholar

²Leonhard Euler (1707-1783), Swiss mathematician

³Jean-Baptiste Lamarck (1744-1829), French naturalist

⁴Pierre-François Verhulst (1804-1849), Belgian mathematician

⁵Irene-Jules Bienayme (1796-1878), French statistician; Francis Galton (1822-1911), English polymath and statistician; Henry W. Watson (1927-1903), English mathematician

useful probabilistic technique, that became a lasting powerful technique in probability theory, the *branching process* (a.k.a. BGW-branching process). The aim was to take into consideration the fact that individuals produce a random number of offsprings, and that a population can become extinct. For a more detailed account about the genesis of the BGW-process, see Chapter 9 in [3].

One of the best known figures in the development of evolutionary dynamics is Darwin⁶. His main contribution is the proposal that all species of life have descended over time from common ancestors [13]. Darwin's theory of evolution has become one of the important cornerstones of modern science. Darwin arrived to his theory of evolution in the 1830's after reading Malthus' book, but he was hesitant to publish it because of the possible crash with the religious authorities and public in general in the Victorian England. In the late 1840's, another English naturalist A. Russell⁷ [60], independently came up with the same idea of evolution by natural selection, communicated his writings to Darwin, and apparently convinced him to publish the manuscript *On the Origin of Species*, which appeared in 1859 [59].

At approximately the same time, in the present Czech town of Brno, Mendel⁸ published the conclusions of his genetic experiments with peas, establishing one of the fundamental laws in genetics: Genes come in pairs and are inherited as distinct units, one from each parent [48] (see [56] for a survey of the mathematical contents of Mendel's work). Due to several different causes, mainly that at the time, communications with Brno were rather slow, and Mendel work did not get the deserved recognition, see for example Ch.8 in [3]. Today, his work is recognized as one of the important contributions to the development of genetics.

From the mid XIX century, there was an exponential growth in the research effort on evolutionary dynamics, at the beginning mostly in population biology, but in the beginning of the XX century the effort extended to other fields dealing with evolutionary processes. We only mention the most relevant work to the Moran process.

A result that had important repercussions for future mathematical models of population genetics, was the *Hardy-Weinberg principle*⁹, which states that *in the absence of certain external disturbing factors, the genetic variation in a population will remain constant from one generation to the next* [30, 67].

⁶Charles Robert Darwin (1809-1882), English naturalist and biologist

⁷Alfred Russel Wallace (1823-1913) English naturalist

⁸Gregor Johann Mendel (1822-1884), Czech abbot and scientist

⁹Godfrey Harold Hardy (1877-1947), English mathematician; Wilhelm Weinberg (1862-1937), German obstetrician

The Hardy-Weinberg principle describes how starting from an initial equilibrium state, genetic variations can be measured as changes from that state. However, as in reality, the disruptive factors in the model do not occur in nature, the Hardy-Weinberg equilibrium rarely applies in reality. But the principle was a seed for more realistic models. Another important result, inspired by the work of Verhulst, were the Lotka-Volterra¹⁰ partial differential equations for studying the survival in a framework of competing species [43, 66]. Their system of equations analyzes the parameters of a predator-prey situation for getting a coexistence equilibrium, where the prey species survive. The PDE system models the principle *that organisms that survive and prosper are those that have a greater capacity to capture and use energy than their competitors*. Today, the Lotka-Volterra equations are one of the important deterministic math tools used in many different areas of evolutionary dynamics, for example ecology or oncology (see for ex Ch. 4 in [58]).

In 1927, Kermack and McKendrick¹¹ produced a notable series of papers on a general theory for transmission of infectious diseases, using deterministic ordinary differential equations [37, 38, 39]. They introduced the SIR epidemiology model; Susceptible-Infected-Retired, where retired means that the healed agent does not participate more¹² in the computation of the evolution of people infected with a contagious illness in a closed population. Today there are many epidemiology models, which divide the infections on classes according to the characteristics of the infection. The most famous of the classes is the SIS model; Susceptible-Infected-Susceptible (to be re-infected). Most of those models use differential equations as the modeling tool, see for example [33].

The work of Kermack and McKendrick also gave rise to a fruitful area of modeling the epidemiological models on a graph, where vertices represent agents and edges represent a proximity relationship. Those models are known by the generic name of the *contact process*, see [18] or [57]. The contact process is a Markovian process¹³ used to model the spread of infection on a usually infinite graph $G = (V, E)$, where each $v \in V$ can be in

¹⁰Alfred J. Lotka (1880-1949), US mathematician; Vito Volterra (1860-1940), Italian mathematician and physicist

¹¹Anderson G. McKendrick (1876-1943), Scottish physician; William Ogilvy Kermak (1898-1970), Scottish biochemist

¹²because it becomes immune or because it dies

¹³Recall that an stochastic process $\{X(t)\}_t$ is said to be *Markovian* if the process does not remember the history of past events, i.e. the value of at step $X(t+1)$ only depends on $X(t)$.

two states, 0 (infected) or 1 (healthy). An infected $v \in V$ heals (changes from 0 to 1) with a rate of 1, independently of their neighbors. Site v gets infected (changes from 1 to 0) at a rate proportional to the number of infected neighbors, where the constant of proportionality is λ . Depending on the value of λ the system could behave in two different ways; the infection could survive forever, or the infection could die out. For instance if $\lambda = 0$, the infection will die rather quickly, even when initially starting from a large but finite number of infected vertices. One of the main questions in the contact process is to find the critical constant λ_c , which separates the two regimes. Notice that the fact that an agent can self-heal, makes the contact process a SIS model. The seminal work on continuous contact process was due to Harris¹⁴ [32], where he studied the contact process on infinite k -dimensional grids. Since then, there have been hundreds of research papers on the contact process and its different variants, with most of the work done in the mathematical, physics and computer science communities (see for example [42, 18, 27]). Notice that there is a similarity between the contact process and percolation¹⁵.

A significant event in early XX century was the probabilistic model for *genetic drift* of finite population due to Wright and Fisher¹⁶ [68, 20]. It describes the sampling of alleles in a population with no selection, no mutation, no migration, non-overlapping generation times and random mating. In spite of its unrealistic assumptions, the Wright-Fisher model has been used as a tool for modeling the effect of more complex evolutionary forces. In the same way that most geneticists recognize Mendel as the father of modern genetics, they also would agree that the Wright-Fisher model can be considered the main significant pioneer work of the field of population genetics.

Genetic drift is a change that occurs in alleles¹⁷ from generation to generation in a population due to random sampling of organisms. For instance, if there is a population of 5,000 people, then are 10,000 copies of each gene. Assume a gene has 3,000 copies of one particular allele (type), in the next generation, there will not necessarily be exactly 3,000 copies again. Instead there may be 3,050 or 2,960 copies instead [46]. Mathematically, the Wright-Fisher model is a discrete-time Markov chain that describes the evolution

¹⁴Theodore Edward Harris (1919 - 2005), US mathematician

¹⁵For a gentle introduction to percolation see [36], for a more technical treatment see [28]

¹⁶Sewall Green Wright (1889-1988), US geneticist, Ronald A. Fisher (1890 - 1962), English statistician and geneticist

¹⁷The human specie only have two types of alleles, each human inherits one allele from the father and one from the mother

of the count of one of these alleles over time, where at each generation the population does not change, i.e. say, it is always $2N$ ¹⁸. Let A and B be two alleles appearing in genes in a given population with N individuals at generation t . Let X_t be the count of the A allele in the population at generation t . The state space of the Markov chain is the set of possible counts of A , i.e. $X_t \in \{0, \dots, 2N\}$. To go from population t to $t + 1$, we sample with replacement a collection of alleles from the population at t to form a new population at generation $t + 1$. This process is a binomial sampling of alleles for each generation, where the probability of going from $X_t = i$ to $X_{t+1} = j$ can be computed from binomial probability mass function with size $2N$, and with probability $p = \frac{i}{2N}$ of success in choosing A , i.e. $\Pr[X_{t+1} = j | X_t = i] \sim \text{Bin}(2N, \frac{i}{2N})$. Therefore, the transition probability $P_{i,j}$ in the Markov chain is given by:

$$P_{i,j} = \binom{2N}{j} \left(\frac{i}{2N}\right)^j \left(1 - \frac{i}{2N}\right)^{2N-j}.$$

It is easy to prove that for any $t > 0$, $\mathbf{E}(X_t) = X_0$, so the expected value of the Wright-Fisher process at any time t is X_0 , i.e., the number of A alleles at the starting generation.

From the biology point of view, another important feature of the Wright-Fisher Markov chain is that the two absorbing states are the cases $X_t = 0$ (all the alleles are B in generation t) or $X_t = 2N$ (all the alleles are A in generation t). Notice the main characteristics of the Wright-Fisher model are:

- The model evolves in discrete generations (the generations do not overlap).
- Some genes in the generation have no offsprings, others could have up to N offsprings.
- The model can be generalized to $K > 2$ alleles.

Taking inspiration in Wright [69] to do further analysis extension from the Wright-Fisher model, Patrick Moran¹⁹ defined a new model for genetic drift, which today is known as the *Moran Process* [54, 55]. At the beginning of his article, Moran explain his motivation for creating a "different" model than Wright: "*In the present paper we obtain a more manageable theory* (than the

¹⁸For the definitions of Markov Chain, see for example chapters 7 and 11 in [51]

¹⁹Patrick Alfred Pierce Moran (1917-1988), Australian statistician

one in [69]) *by making a slight modification of the model. Instead of assuming that all the individuals die at the same time we assume that at each instant at which the state of the model may change, one of the individuals, chosen at random, dies and is replaced by a new individual . . .*

Moran's model considers a population of size N , which is *well-mixed* (no structure between the population), which evolves in the following way: A pair of genes are sampled with replacement from the population, then with an exponential rate of $\binom{N}{2}$, one dies and the other splits into two new ones. In particular, the above reproductive system means that in the labeled population, each pair of labels has an alarm clock which goes off at intervals that are independently exponentially distributed with parameter 1, and when a clock goes off one gene dies and the other reproduces (with equal probabilities). Notice that there are three main differences of the Moran model with respect to the Wright-Fisher model:

- In the Moran model, generations overlap.
- The time between off-springs generation in the Moran model is continuous.
- In the Moran model an individual always has zero or two offspring, which again is a difference with the Wright-Fisher model.

Notice the Markov chain for the Moran process is a continuous Markov chain. Moreover, recall that in continuous Markov chains, we could use the *embedded Markov chain*, which is a discrete stochastic process obtained from a Markov chain when sampled at instants where something happens, see for example Section 8.5 in [51].

For the Moran process, the embedded Markov chain has state space $\{0, \dots, N\}$, and transition probabilities for going from $X_t = i$ to $X_{t+1} = j$ are

$$p_{i,j} = \begin{cases} \frac{i(N-i)}{N^2} & \text{if } j = i \pm 1, \\ \frac{i^2}{N^2} + \frac{(N-i)^2}{N^2} & \text{if } j = 1, \\ 0 & \text{otherwise.} \end{cases}$$

As in the case of the Wright-Fisher model, 0 and N are the absorbing states in the Moran embedded Markov chain.

Finally, let us remark that the exact calculations for the Wright-Fisher model are intractable, even in the case of infinite population limits, while for the Moran model, it is more tractable to get exact results (it does not mean it is easy).

3 The LHN model: Definitions and intuitions

In this section, we present an adaptation of the Moran process by Lieberman, Hauert, and Nowak [41], see also Chapter 6 and 8 in [58]. As Lieberman et al. write in the introduction of their paper: *We want to provide a general account of how population structure affects evolutionary dynamics. We study the simplest possible question: what is the probability that a newly introduced mutant generates a lineage that takes over the whole population?*

From now on, we denote the modified model as the LHN model. In the LHN model, individuals are represented as the vertices of a finite graph and, at each step of the discrete-time process, an individual is selected at random to reproduce. Then, the chosen vertex selects uniformly at random one of its neighbors and replaces that neighbor with a copy of itself (its offspring). The probability that any given individual is chosen to reproduce is proportional to its *fitness*. In biology, *fitness* denotes the reproductive success of an "individual". There are two types of vertices, "mutants" and "non-mutants"²⁰. Mutants have fitness $r > 0$ and non-mutants have fitness 1. The process iterates by choosing randomly a vertex. If it is a mutant with a certain probability he will replace a non-mutant neighbor by a new mutant, and if the vertex is a non-mutant, with a certain probability it can replace a mutant neighbor by a non-mutant. Note that the neighborhood requirement makes the population structured. See Fig. 1 for the details of the LHN algorithm. The input graph to the LHN process can be a digraph or it can be an undirected graph, depending on the constraints of the modeling.

Previously, there had been other graph dynamics models. For instance, Bramson and Griffeath had a process to model the spread of cancerous cells placed on the nodes of a bidimensional finite grid $\mathbb{Z}^2$ [7].

The main component of the LHN model with respect to the original Moran model is the reproductive idea, but even this idea is different between both models. In the Moran model we sample any two nodes according to an exponential distribution, while in the LHN first we sample a vertex v according to its "fitness" and afterwards we sample a second from the v neighbors. In Section 4 it will be shown that a discrete version of the Moran's drift process can be modeled as a particular case of the LHN process on a complete graph with size N , with fitness $r = 1$ and probability $1/(N - 1)$. When $r = 1$, the process is said to be *neutral*.

Explaining more in detail the LHN model, given as input a graph $G = (V, E)$, with $|V| = N$, at step 0 all vertices are non-mutant (i.e. with fitness 1).

²⁰It could have been black and white, or healthy and sick, depending on the application

Then u.a.r.²¹ one vertex is selected and changed to a mutant, with new fitness $r > 0$. The evolution of the LHN process defines a simple random walk on a discrete, transient Markov chain, on states $\{0, 1, \dots, N-1, N\}$, where each state indicates the number of mutants in the graph. The Markov chain has two absorbing states; the *fixation* state N , when all the nodes are mutants, and the *extinction* state 0 , where all the mutants disappear. Fig. 2 shows the Markov chain for a process on a graph of size $N = 4$. Notice that state 1 corresponds to having only one mutant, and therefore the random walk starts in that state. Also, in Fig. 2 we introduce the notations for the transition probabilities; if we are in state i , p_i denotes the probability of going "forward" to state $i+1$, s_i is the probability of remaining in state i , and q_i is the probability of going "backwards" to state $i-1$.

LHN $G(V, E), r > 0$

At $t = 0$ sample uniformly at random a $v \in V$ which becomes a mutant.

for all $t \geq 1$ **do**

Assume at step t we have k mutants and $N - k$ non-mutants

Define the total fitness at time t by $W_t = kr + (n - k)$

Choose $v \in V$ with probability $\frac{r}{W_t}$ if u is mutant,

and $\frac{1}{W_t}$ if v is non-mutant.

Let $\mathcal{N}(v)$ be the set of neighbors of v .

Choose uar a $u \in \mathcal{N}(v)$, and replace u with a copy of v .

end for

Figure 1: Basic randomized algorithm for the LHN model.

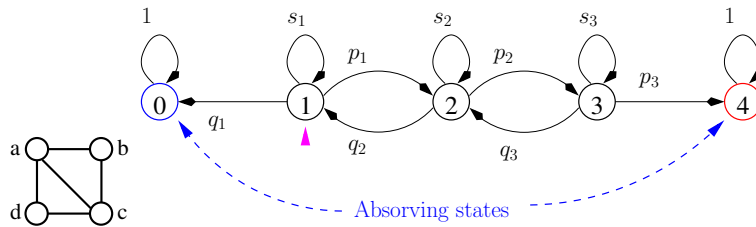


Figure 2: To the left, the graph G , with vertices $\{a, b, c, d\}$, to the right the Markov chain of the LHN process on G for any fitness $r > 0$.

Given a graph G with a mutant fitness of $r > 0$, define the *fixation probability*

²¹uniformly at random, taken from a uniform distribution

$f_{G,r}$ as the probability that a single mutant will take over the whole graph G . In a similar way define the *extinction probability* $g_{G,r} = 1 - f_{G,r}$, as the probability that the process will stop in the extinction absorbing state. From the computational point of view, there are several interesting problems associated with the fixation probability: Given a graph with a particular topology and a specific value for r , compute the value of $f_{G,r}$. How long does it take to arrive to an absorbing state? If computing $f_{G,r}$ is difficult, could we give a good approximation to it?

Another stochastic process that has some resemblance with the LHN is the SIS model for the contact process: Each vertex in the graph represents an individual, one gets randomly infected and the infection transmits through the directed or undirected edges, as in the LHN model. However, there is a substantial difference between the LHN process and the contact process, in particular between the LHN and the *voter system*, which is a finite and discrete version of the contact process (see for example Part II in [42], or Sect. 6.9 in [17] for more information on the voter system model). The main difference is that in the contact process, there is no fitness, i.e. fitness is 1, and therefore the probabilities of transmitting and healing from infection are totally different in the contact process and in the LHN process, which makes that the techniques needed to deal with both models are quite different. For a more detailed explanation of the specific differences between voter systems and LHN, see Sect. 1.4 in [21].

4 Computing the fixation probability

Given G and r , a necessary condition to have non-trivial values of $f_{G,r}$ is that if G is undirected then G must be connected, and if $\vec{G}$ is a digraph it should be strongly connected. For instance, consider a directed digraph with at least one vertex with in-degree 0, for example the directed line or the directed star, as in Figure 4. The only way that those graphs can arrive to fixation is that the first mutant is the vertex with in-degree 0, so $f_{G,r} = 1/N$, and the absorption probability is $g_{G,r} = 1/N$. If we have more than two vertices with in-degree 0, $f_{G,r} = 0$.

For general graphs and digraphs, working with the chain where the states are the total number of mutant vertices could be difficult, as given a particular topology, the transition probabilities p_i, s_i and q_i depend on the specific positions of the mutant vertices. This is easy to see: Consider the N star graph, i.e. a tree on N vertices, where one vertex has degree $N - 1$ and the remaining vertices have degree 1. To compute p_1 , i.e. the probability of

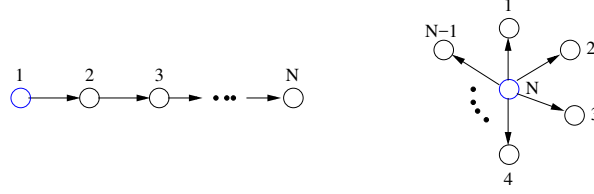


Figure 3: The directed line and the centre-out directed star

going from one vertex mutant to two mutants is very different depending if the mutant is the vertex with degree N or if the mutant is one of the vertices with degree 1.

To simplify the computation of the fixation probability we use the Markov chain of configurations, which is the same chain as before, but considering as states all the possible configurations throughout the process. A configuration C is a specific $C \subseteq V(G)$, where all vertices in C are the mutants, at that step in process. The number of configurations of size i is $\binom{N}{i}$, and the total number of states in the Markov chain of configurations is 2^N . Given a configuration C of size i , let $p_C(i)$ denote the probability of going to a configuration C' of size $i + 1$, let $q_C(i)$ the probability of going to a configuration C'' of size $i - 1$, and let s_C the probability of remaining in the same configuration. In Figure 4 we present the configuration chain to the example in Figure 3. In the Markov chain of configurations, the fixation probability can be determined by standard techniques, but we have to solve a set of 2^N linear equations, which even for not too large N is not computationally feasible.

To survey known results for the computation of $f_{G,r}$ let us first state some useful result about the monotonicity of the fixation probability function. For any configuration C of mutant vertices, let $f_{G,r}(C)$ denote the fixation probability from a configuration C of mutants, i.e. $f_{G,r}(C) = \sum_{v \in C} f_{G,r}(\{v\})$. Shakarian et al. [63] proved that for a given graph, and any configuration C , the fixation probability under neutral drift is a lower bound for the fixation probability when $r > 1$, i.e. $f_{G,1}(C) \leq f_{G,r}(C)$, for $r > 1$. Serna et al. [15] answered a conjecture stated in [62], showing that *for any digraph $\vec{G}$, if $0 < r \leq r'$ and for configurations $C \subseteq C' \subseteq V$, we have that $f_{G,r}(C) \leq f_{G,r'}(C')$* . In the same paper, a stronger condition on the monotonicity of the fixation probability for digraphs is proved as well, namely that *if $0 < r \leq r'$, then for any $\vec{G}$ we have that $f_{G,r} \leq f_{G,r'}$* .

Regarding the computation of $f_{G,r}$, an intuitive idea is that if $r > 1$ then

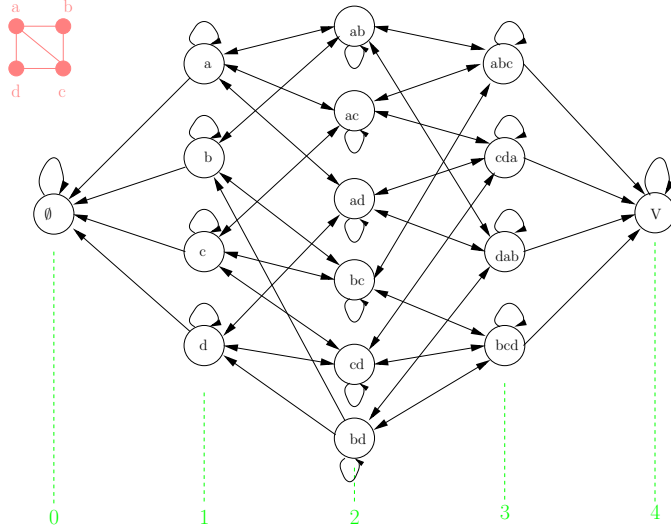


Figure 4: The configuration Markov chain.

the process is going to favor absorption, while if $0 < r < 1$, the process is going to favor extinction.

Let us consider an easy example to compute $f_{G,r}$. We compute the fixation probability for the LHN algorithm, which models the original Moran process [54]. In this example, the input graph is the complete graph on N vertices $G = K_N$, with neutral drift ($r = 1$). Due to the highly symmetrical topology of the graph, we will argue on the Markov chain with $N + 1$ states, each one counting the number of mutants. We start in state 1: to go to state 0 we have to choose a non-mutant with probability $\frac{N-1}{N}$ and then choose the infected neighbor with probability $\frac{1}{N-1}$, therefore $q_1 = \frac{1}{N}$. To go from state 1 to state 2, we have to choose the mutant with probability $\frac{1}{N}$ and then the probability of choosing a neighbor which is non mutant is $\frac{N-1}{N}$, therefore $p_1 = \frac{1}{N} = q_1$. Notice the probability s_1 of having self-loops is $1 - 2p_1$. The same argument shows that if we are in state i (i mutants), we can choose a non-mutant with probability $\frac{N-i}{N}$ and choose a mutant neighbor with probability $\frac{i}{N-1}$, so $q_i = \frac{(N-i)i}{N(N-1)} = p_i$. Moreover, the probability s_i of staying at the same state is $1 - 2p_i$.

To compute the fixation probability $f_{K_N,1}(i)$ from state i , we can normalize by making $s_i = 0$ for all self-loops except the ones for N and for 0. As

$s_i + 2p_i = 1$, then we have $p_i = 1/2 = q_i$, and we get

$$f_{K_N,1}(i) = \begin{cases} \frac{1}{2}f_{K_N,1}(i-1) + \frac{1}{2}f_{K_N,1}(i+1) & \text{if } 0 < i < N, \\ 0 & \text{if } i = 0, \\ 1 & \text{if } i = N. \end{cases}$$

An easy induction yields that $f_{K_N,1}(i) = \frac{i}{N}$ and $f_{K_N,1}(1) = \frac{1}{N} = f_{K_N,1}$.

As said, for general graphs (digraphs) with any value of $r > 0$, if the graph is connected (strongly connected) depending on the value of r , there could be a large number of iterations until having convergence towards one of the two absorbing states. From a given configuration C with $|C| = i$ we could go to any of the $\binom{N}{i-1}$ configurations with $i-1$ mutants, or to any one of the $\binom{N}{i+1}$ configurations with $i+1$ mutants.

Let $\vec{G} = (V, \vec{E})$ be an input digraph for the LHN process, and let $\forall v \in V$, $\delta^+(v)$ be its outgoing degree, i.e. the number of edges going out of vertex v . Define the stochastic matrix $W = [w_{vu}]$, associated to $\vec{G}$, where

$$w_{vu} = \begin{cases} 1/\delta^+(v) & \text{if } \overrightarrow{(v,u)} \in \vec{E}, \\ 0 & \text{if } \overrightarrow{(v,u)} \notin \vec{E}. \end{cases}$$

Therefore, regarding a specific mutant $v \in C$, we can go to a new configuration with $i+1$ mutants by choosing a non-mutant vertex $u \in \mathcal{N}(v)$ with probability $r \cdot w_{vu}$, and we can go to a configuration with $i-1$ mutants by choosing $u \notin C$ such that $v \in \mathcal{N}(u)$ with probability w_{uv} . The equation of the transitions' configuration for a given C is given by

$$f_{G,r}(C) = \frac{\sum_{v \in C} \sum_{u \notin C} (r \cdot w_{vu}) \cdot f_{G,r}(C \cup \{u\})}{\sum_{v \in C} \sum_{u \notin C} (r \cdot w_{vu}) + \sum_{u \notin C} \sum_{v \in C} w_{uv}} + \frac{\sum_{u \notin C} \sum_{v \in C} w_{uv} f_{G,r}(C \setminus \{v\})}{\sum_{v \in C} \sum_{u \notin C} (r \cdot w_{vu}) + \sum_{u \notin C} \sum_{v \in C} w_{uv}},$$

with $f_{G,r}(0) = 0$, and $f_{G,r}(N) = 1$, see for example [61]. Therefore, in general to compute $f_{\vec{G},r}$ we may have to solve 2^N transition equations.

Lieberman et al. [41], computed the fixation probability for a class of graphs with a specific property, which basically is the class of regular digraphs. The authors deal with undirected graphs by considering each edge as doubly directed.

To see their specific definition, define the *temperature* of a vertex $i \in V$ as $T_v = \sum_{u \in V} w_{uv}$. We say that $\vec{G}$ is *isothermal* if $\forall u, v \in V$, $T_u = T_v$. Notice

that a graph is isothermal if its matrix W is bi-stochastic, all rows and columns add to 1. The intuitive idea behind the temperature is that "hot" vertices, with high temperature T , tend to be replaced often, while "cold" vertices are seldomly replaced, see Fig. 5. Notice that the same definition of W applies to undirected G , with $w_{vu} = 1/d(v)$, where $d(v)$ is the degree of v .

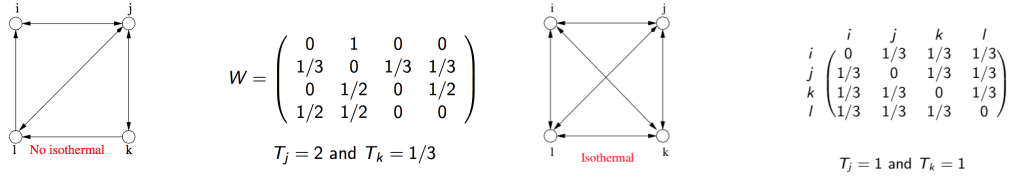


Figure 5: No-isothermal and isothermal graphs with their W matrices.

The main theorem in [41] is the following

Isothermal Theorem: [LHN]

1. For a strongly connected digraph $\vec{G}(V, \vec{E})$, with $|V| = N$, and such that $\forall i, j \in V$ we have $T_v = T_u$ (i.e. W is bi-stochastic), then for $r \neq 1$

$$f_{\vec{G},r} = \frac{1 - 1/r}{1 - 1/r^N}.$$

2. If G is undirected, connected and d -regular then for $r \neq 1$

$$f_{G,r} = \frac{1 - 1/r}{1 - 1/r^N}.$$

The authors denoted this specific fixation probability by $\rho = \frac{1-1/r}{1-1/r^N}$, and call it *the Moran fixation probability*. Since then, it has been used as a benchmark for other fixation probabilities to compare with.

A particular case of the LHN-Theorem was earlier obtained by Maruyama [45], who gave a proof that the fixation probability of a mutant is independent of the geographical structure of a population, in a model that was a geographically structured version of Moran's overlapping generation model.

The proof of the LHN-theorem that is given in [41] is a bit complicated. A more intuitive one is given in Section 6.3 of [58]. For the specific case

where the graph is undirected, Broom and al. [10], gave an alternative statement and proof of the LHN-theorem: *A simple connected undirected graph is isothermal if and only if it is regular.*

After the publication of [41] there were several papers doing experimental and analytical work for the fixation probability: For instance, Masuda and Ohtsuki [47] investigated the evolutionary dynamics in directed and/or weighted networks; Broom et al. [8, 10] studied the LHN process for the case of small graphs and graph classes where a high degree of symmetry reduces the size of the set of equations, as paths, cycles, stars and complete graphs; Rychtář et al. [61] produced experimental results on fixation probabilities for grid-like random graphs.

There also has been research in different aspects of the LHN process. For instance, Adlema et al. [1] studied the amplification with respect to an adversarial placement of the initial mutant. Mertzijs et al. [50] also consider the extent to which the number of favorable initial mutants for fixation can be bounded.

A direct consequence of the LHN-theorem is the following statement for different values of $r > 0$:

Given a connected N -vertex regular digraph $\vec{G}_N$ with fitness $r > 0$,

$$f_{\vec{G}_N, r} = \begin{cases} f_{\vec{G}_N, r} = \frac{1-(1/r)}{1-(1/r^N)} > 1 - \frac{1}{r} = \Theta(1) & \text{if } r > 1, \\ f_{\vec{G}_N, 1} = \frac{1}{N} \rightarrow 0 \text{ as } N \rightarrow \infty & \text{if } r = 1, \\ f_{\vec{G}_N, r} = \frac{1-(1/r)}{1-(1/r^N)} \rightarrow 0 \text{ as } N \rightarrow \infty & \text{if } r < 1. \end{cases}$$

Goldberg et al. [25] established a *phase transition* on $f_{G, r}$, depending on the fitness parameter r :

Theorem *For all sufficiently large connected undirected graphs G with N vertices and for all $r > 0$ and real constant $c > 0$, the fixation probability has a phase transition behavior:*

$$f_{G, r} \begin{cases} < 1/N^c & \text{if } r < 1, \\ = 1/N & \text{if } r = 1, \\ > \frac{(\log N)^c}{N} & \text{if } r > 1. \end{cases}$$

In the same work the authors prove that for directed graphs there is no such phase transition.

Theorem *There exist a large family $\mathcal{G}$ of strongly connected graphs, such that, if $\vec{G} \in \mathcal{G}$ is a graph in the family, with sufficiently large number of*

vertices N , for every $r > 0$ there exists a function $\psi(r) : \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}$ such that

$$\frac{1}{\psi(r)N} \leq f_{\vec{G},r} \leq \frac{\psi(r)}{N}. \quad (1)$$

To define $\mathcal{G}$, in the previous theorem, let $k, a \geq 1$ be integers, let $w_1, \dots, w_{ka}, v_1, \dots, v_k$ be distinct vertices, for $0 \leq i \leq k$, let $I_i = \{w_{(i-1)a+1}, \dots, w_{ia}\}$. Form $G_{k,a}$ by taking the directed cycle $w_1, \dots, w_{ka}, v_k, \dots, v_1, w_1$ and adding the edge set $\cup_{0 < i \leq k} (I_i \times \{v\})$, see Fig. 4. Finally, for all $0 < i \leq k$, set $W_i = \cup_{l=j}^k I_l$, $V_j = v_j, \dots, v_k$ and $X_j = W_j \cup V_j$, with $W_{k+1} = V_{k+1} = X_{k+1} = \emptyset$. In order to obtain the theorem, for every $r > 1$ choose $k = \lceil 7 + \log(5r^2) \rceil$ and $a = \lceil 4r \rceil$, and then set $\mathcal{G} = G_{k,a}$.

Using the result that $f_{\vec{G},r} \geq 1/N$ for all N -vertex digraphs [15] together with the bounds given in (1), it is direct to get that for any digraph $\vec{G}$, we must have that $f_{\vec{G},r} = \Theta(1/N)$.

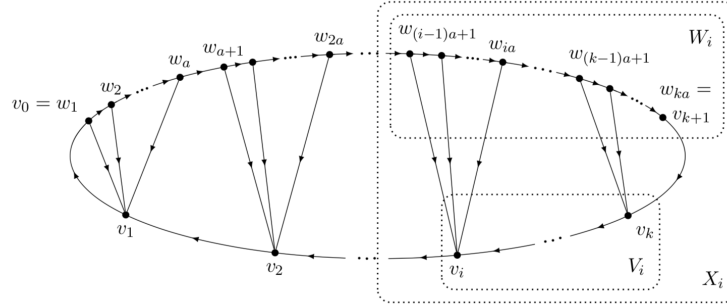


Figure 6: A graph in $\mathcal{G}$ as constructed in [25]

Using ρ as a mark of fixation probability, Lieberman et al. [41] classified the graphs according to their fixation probability, into three major subclasses: Those with $f_{G,r} = \rho$, those with $f_{G,r} > \rho$, which they call *amplifiers*, and those with $f_{G,r} < \rho$, which they call *suppressors*.

4.1 Amplifiers

Given a digraph $\vec{G}$ (or undirected graph) and $r > 1$, G is said to be an *amplifier* if $f_{G,r} > \rho$.

Notice that because of monotonicity, we know that for $0 < r < 1$, we have $f_{G,r} \leq f_{G,1}$ [14]. For this reason, in the definition of amplifiers, it suffices to have the case $r > 1$.

Lieberman et al. [41] proved that the undirected star has fixation probability $f_{G,r} = \frac{1 - \frac{1}{r^2}}{1 - \frac{1}{r^{2N}}}$, which as N grows tends to $(1 - \frac{1}{r^2})$, so it is an amplifier. In the same paper, the authors introduced the directed *superstar* (see Fig. 7), and empirically estimated the fixation probability as $f_{\vec{G},r} = 1 - \frac{1}{r^{-(k+2)}} > \rho$, where k is a parameter of the input graph. Serna et al. [16] showed that the claim was not correct for $k = 3$, getting that for any $r \geq 1.42$, $f_{\vec{G},r} \leq 1 - \frac{r+1}{2r^5+r+1}$, which is less than the claimed $1 - r^{-5}$ (for $k = 3$). Jamieson-Lane et al. [35] gave a useful heuristic argument together with a revised bound. Section 9 in [26] explains why it is difficult to convert that heuristic argument into a rigorous proof.

A family of graphs $\mathcal{G}$ is said to be *strong amplifying* if for all $G \in \mathcal{G}$, with $|V(G)| = N$ the fixation probability $f_{G,r} \rightarrow 1$ as $N \rightarrow \infty$ when the LHN-process is run on G . That is equivalent to saying the extinction probability tends to 0, as N grows large. Notice that strong amplifiers guarantee that a mutant will almost surely reach fixation, even if its fitness is small.



Figure 7: Amplifiers: The undirected star and the directed superstar.

Galanis et al. [21] gave a rigorous analytical answer to an open problem posed in [41], whether there exists an infinite family of digraphs such that the LHN-algorithm has a fixation probability of $1 - o(1)$, as a function of input size N . In [21] Galanis et al. defined a new family of digraphs called *megastars*, see Fig. 8. Megastars are a generalization of the superstars introduced in [41]. Their main result of the paper states that:

Theorem *There exists an infinite family of megastars $\mathcal{M}_N$, such that any graph $G \in \mathcal{M}_N$ has fixation probability $1 - O(N^{-1/2} \log^{23} N)$, for constant*

$r > 1$. Therefore the graphs in the family are strong amplifiers. The paper also provides formal proofs and counter-examples to some other heuristic assertions in [41]. Technically the paper is difficult to grasp, but the techniques introduced for the analysis of the fixation probability should be mandatory for people wishing to work in the field.

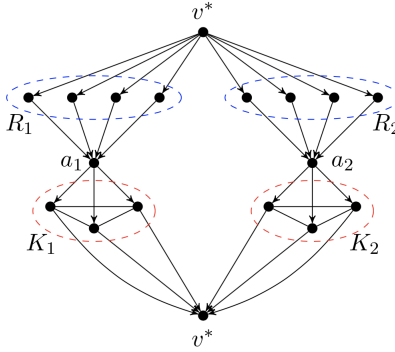


Figure 8: The megastar graphs introduced in [21]

Goldberg et al. [26] showed that the family of megastars $\mathcal{M}_N$ defined in [21] are optimal, in the sense that the extinction limit grows as slowly as possible up to logarithmic factors, since every strongly connected N -vertex digraph $\vec{G}$ has extinction probability $g_{\vec{G},r} = \Omega(1/\sqrt{N})$.

For a long time, it was an open problem to find **undirected graphs** that were strong amplifiers. Giakkoupis [24] gave a simple family of undirected graphs that are strong amplifiers, but they have worse fixation probability than the directed megastars in [21]. Goldberg et al. [26] showed the existence of a strong amplifying family of undirected graphs, with better fixation probability:

Theorem *There is an infinite family of undirected graphs, which they call dense incubators, whose extinction probability is bounded by $O(N^{1/3})$ (see Fig 9).*

The authors also show that up to constant factors, depending on N , the dense incubator family is optimal, in the sense that the extinction limit grows as slowly as possible.

Basically this completes quite well the picture of amplifiers and strong am-

plifiers.

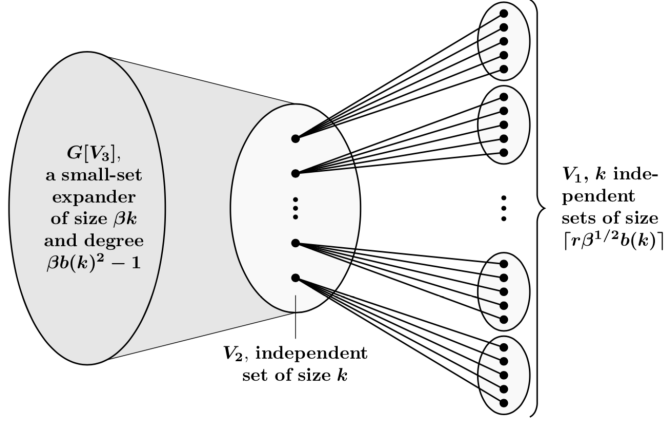


Figure 9: The family of dense incubators. Let β be a constant depending on r , let b be a function such that $b(k) \leq \sqrt{k}$. As $G[V_2, V_3]$ is a bi-regular graph with $\Theta(kb(k)^2)$ edges, each $v \in V_2$ sends $\Theta(b(K)^2)$ edges to each vertex in V_3 and each vertex in V_3 sends $\Theta(b(K)^2)$ edges to V_2 (from [26]).

4.2 Suppressors

Given a digraph $\vec{G}$ (or undirected undirected) and $r > 0$, G is said to be a *suppressor* if $f_{G,r} < \rho$.

As we saw, the directed star and the directed line in Fig 3 are clear examples of suppressors, but they are not strongly connected.

Mertzios et al. [49] introduced a family of undirected graphs $\{G_N\}$ the *axinos* (a.k.a. urchin), which consists of a perfect matching between an undirected cycle C_N and the undirected clique K_N , see Fig. 10. The authors proved that for $1 < r < 4/3$, $f_{G,r} \leq \frac{1}{2}(1 - \frac{1}{r}) < \rho$, as N grows large.

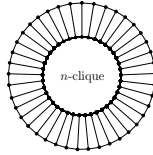


Figure 10: A suppressor: The "axinos" graph (from [49]).

As with the amplification, we say that a family of graphs is *strong suppressor* if for $N \rightarrow \infty$, the extinction probability tends to 1, when the LHN-process is run on any graph in this family. The property of strong suppressors guarantees that a mutant will almost surely reach fixation, even if its fitness is not small.

For a long time it was an open problem to find a family of undirected graphs that were strong suppressors. As we mentioned before, using the fact that for any strongly connected $\vec{G}$, $f_{\vec{G},r} \geq 1/n$ (see [15]) together with the previous equation (1), even the stronger directed suppressors have fixation probability $\Theta(1/N)$ (see [25]).

Regarding undirected graphs, Giakkoupis [24] constructed a family of simple undirected graphs $\mathcal{B}$ parametrized by N , see Fig. 11:

Theorem For $G \in \mathcal{B}$ with N vertices, $f_{G,r} = O(\frac{r^2 \log N}{N^{1/4}})$.

When $r = \Theta(1)$, then $f_{G,r} = O(\frac{\log N}{\sqrt{N}})$ and if $r = o(N^{1/8} \log^{-1/2} N)$ then $f_{G,r} = o(1)$, i.e. it goes to 0 as N grows, so it is a strong suppressor.

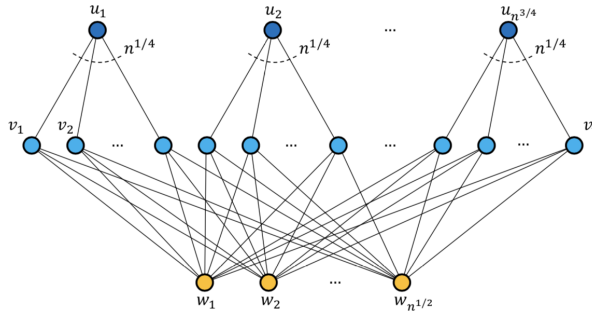


Figure 11: A strong suppressor graph [24]

The proof introduced an interesting general lower bound on the fixation probability of a given graph. Recall that given an undirected graph, if C is a set of mutants, Lieberman et al. [41] introduced the concept of temperature $w(C)$ (harmonic volume, in the nomenclature of Giakkoupis). Explicitly, Lemma 6 in [24] states: *Let G be a connected undirected graph with N vertices and minimum degree δ , and let C be a configuration of mutants. If the initial set of mutants on G is C , then*

$$f_{G,r}(C) \geq \frac{1 - r^{\delta w(C)}}{1 - r^{\delta N}}. \quad (2)$$

This result implies a convenient stopping criterion that can be used to speed up numerical approximations of the fixation probability via simulations. Similar *early termination* experimental heuristic was used in [4] for empirically bounding the fixation probability in digraphs.

Goldberg et al. [25] proved a stronger result:

Theorem *There is a family $\mathcal{H}$ of undirected graphs with N vertices, such that for all $r > 1$, and $G \in \mathcal{H}$, as N grows large, $f_{G,r} \leq \frac{10r^2}{N^{1/2}}$.*

Note that this bound is closer to the known fixation probability of the phase transition for undirected graphs than the bound in the previous theorem by Giakkoupis. On the other hand, the family $\mathcal{H}$ is quite more *complex* than the family $\mathcal{B}$.

To define the family $\mathcal{H}$ (see Fig. 12) let $\mathcal{H} = \{H_{a,k}\}$ for values $r > 1$, $a, k \geq 1$ integers, let $V_0, \dots, V_3$ disjoint vertex sets with accompanying weights $\sigma_0, \dots, \sigma_3$, where the weights σ_i are as given in Fig. 11. Let $H_{a,k}$ be the graph with vertex set $V_0 \cup \dots \cup V_3$ which is the union of a complete bipartite graph between V_0 and V_1 ; a^2k vertex-disjoint k -leaf stars between V_1 and V_2 ; and a perfect matching between V_2 and V_3 . In order to obtain the previous theorem, one may take $\mathcal{H} = \{H_{a,k}\}$, with $a = \lceil 3.5r^2 \rceil$ and $k = \lceil 36r \rceil$.

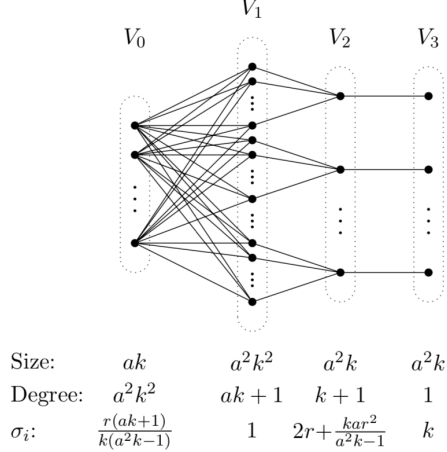


Figure 12: A strong suppressor graph $G = H_{k,a}$ [25]

Therefore, basically most of the interesting open problems for suppressors have been solved.

5 Expected absorption time of the LHN process

Another quantity of interest in the analysis of the LHN process is the random variable τ counting the number of steps until the process arrives to absorption time, that is, the time until one of two states of absorption, fixation or extinction, is reached.

Given an undirected connected G , with $|V(G)| = N$, the LHN process on G is a random walk $\{C_t\}_{t \geq 0}$, where $\{C_t\}$ is the configuration with the vertices which are mutants at time step t , so that $f_{G,r} = \sum_{v \in V} f_{G,r}(\{v\})$.

Define the *absorption time* as the value $\tau = \min\{t \mid C_t = \emptyset \vee C_t = V\}$.

Taylor et al. [64] studied absorption times for variants of the LHN process but their results only apply to the process on regular graphs, where it is equivalent to a biased random walk on the line Markov chain with absorbing barriers. Broom et al. [9] also do an absorption time analysis, but restricted to cliques, cycles and stars. Serna et al. [14] considered the problem of the expected absorption time, for all undirected graphs. Their main result is that there is a polynomial bound for the expected absorption time:

Theorem *Given an undirected and connected G , together with a fitness $r > 0$ if we run the LHN process $\{C_t\}_{t=0}$, starting with $|S_1| = 1$, then:*

1. *If $r < 1$, then $\mathbf{E}(\tau) \leq \frac{1}{1-r}n^3$,*
2. *if $r > 1$, then $\frac{r}{1-r}n^3 \leq \mathbf{E}(\tau) \leq \frac{r}{r-1}n^4$,*
3. *if $r = 1$, then $\mathbf{E}(\tau) \leq n^6$.*

The proof technique used to prove that theorem uses a *potential function* argument. It was the first time that this kind of argument was used to prove a result in this particular field.

The sketch of the proof is as follows: Given an undirected and connected G with N vertices let $\{C_t\}_{t \geq 0}$, be the states chosen in the random walk on the Markov chain corresponding to the LHN process, where $C_t \subset V$ are the configurations of mutant vertices at a given step t . Define the potential function $\phi(C_t) = \sum_{v \in C_t} \frac{1}{d(v)}$. Then the proof proceeds by showing that the potential strictly increases in expectation for $r > 1$, and strictly decreases for $r < 1$, i.e. given any set of vertices $S \subseteq V$, we have that:

- $\mathbf{E}(\phi(C_{t+1}) - \phi(C_t) \mid C_t = S) < \frac{r-1}{N^3} < 0$, if $r < 1$
- $\mathbf{E}(\phi(C_{t+1}) - \phi(C_t) \mid C_t = S) \geq (1 - \frac{1}{r})\frac{1}{N^3} > 0$, if $r > 1$, and
- $\mathbf{E}(\phi(C_{t+1}) - \phi(C_t) \mid C_t = S) = 0$, if $r = 1$.

In order to bound the absorption time, the paper uses well known martingale techniques with a potential function that decreases in expectation until absorption, see [29]. Breaking the proof into different values of the fitness, the statement of the theorem is obtained.

As Serna et al. mention in [14], *the method of bounding is crude. For example consider G_N to be the double star formed from two undirected $N/2$ -stars with an extra edge between the two centers. Let $S = N/2$ be the set of all vertices in one of the stars, then it is easy to show that $\mathbf{E}(\phi(C_{t+1}) - \phi(C_t) | C_t = S) = \Theta(N^{-3})$, which is far from the bound for the expected change of ϕ .*

As a side curiosity, Goldberg et al. [25] gave the best absorption time for the *double star* on N vertices, and $r > 1$, their expected absorption time is $\mathbf{E}(\tau) \leq \frac{(r-1)^2 N^3}{2^5 r^4}$.

Goldberg et al. [25] improved the expected absorption time for undirected connected graphs:

Theorem *Given an undirected connected graph G with N vertices, assuming that $r \neq 1$, then for all $\epsilon > 0$, $\mathbf{E}(\tau) = o(N^{3+\epsilon})$. Moreover, there is a family of undirected graphs $\mathcal{G}$ such that for any $G \in \mathcal{G}$ whenever $\mathbf{E}(\tau) = \Omega(N^3)$.*

The case of bounding $\mathbf{E}(\tau)$ for digraphs is very different. It is proved in [15] that *there is an infinite family of strongly connected digraphs with N vertices such that their expected absorption time is exponential, i.e. $\mathbf{E}(\tau) = 2^{\Omega(N)}$ (see for example Fig. 13).*

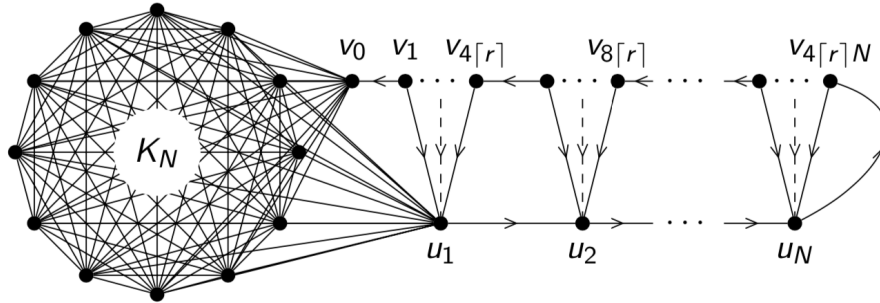


Figure 13: A digraph with exponential absorption time [15]

Besides the important fact that for digraphs the expected absorption time not necessarily is polynomial in N , it is also an indication that the techniques

used to approximate fixation in undirected graphs, do not work for directed graphs.

Let us consider a particular case of digraphs, the d -regular digraphs. Recall that in a strongly connected d -regular digraph every vertex has in-degree d and out-degree d . Regularity simplifies some calculations because the detailed topology of the graph is irrelevant.

Recall that in [41] it is proved that for any regular graph on N vertices and fitness $r > 0$, a single randomly placed mutant reaches fixation with probability $\rho = \frac{1-(1/r)}{1-(1/r^n)}$.

The main theorem in [15] is the following:

Theorem *Let G be a strongly connected d -regular N -vertex digraph. Then the expected absorption time satisfies*

$$\left(\frac{r-1}{r^2}\right)NH_{N-1} \leq \mathbf{E}(\tau) \leq N^2d, \quad (3)$$

where H_N is the N -th harmonic number²². Moreover, the upper bound is tight up to a constant, which depends on d and r , i.e. there are d -regular digraphs with $\mathbf{E}(\tau) = \Theta(N^2d)$.

Notice the first immediate consequence of this theorem is that in the particular case of regular digraphs absorption is reached in expected polynomial time, exactly like in undirected graphs.

We can improve the upper bound on the absorption time for certain families of undirected graphs, using the notion of *isoperimetric number*.

Given an undirected graph $G = (V, E)$, the *isoperimetric number* is defined as $i(G) = \min_S \left\{ \frac{|\delta S|}{|S|} \mid S \subset V, 0 < |S| \leq |V|/2 \right\}$, where δS is the set of edges in the cut between S and $V \setminus S$. Isoperimetric theory was developed in the 1960's by different researchers, see for example [31, 52].

Let us start by stating an alternative version of the previous theorem for undirected d -regular graphs:

Theorem [15] *Let G be a d -regular undirected graph with N vertices. The expected absorption time value for the LHN process on G satisfies*

$$\mathbf{E}(\tau) \leq \frac{2dNH_N}{i(G)}. \quad (4)$$

A direct application of (4) yields several interesting implications to improve the expected absorption time for some undirected graph families:

²² $\ln N + \gamma + o(\frac{1}{N})$

1. Since K_N has $i(G) = \Theta(1/\sqrt{N})$ [52], we have $\mathbf{E}(\tau) = \Theta(N \log N)$, which improves the previous bound $\mathbf{E}(\tau) = O(N^3)$.
2. Since the $\sqrt{N} \times \sqrt{N}$ -grid has $i(G) = \Theta(1/\sqrt{N})$ (see [2]), it follows that $\mathbf{E}(\tau) = O(N^{3/2} \log N)$, which improves the previously obtained bound of $\mathbf{E}(\tau) = O(N^2)$.
3. Since the hypercube on N vertices has $i(G) = \Theta(1/\sqrt{N})$ (see [52]), it follows that $\mathbf{E}(\tau) = O(N^{3/2} \log^2 N)$, which improves the obtained bound $\mathbf{E}(\tau) = O(N^2 \log N)$ from (3).
4. Since for $d \geq 3$ there exists a $0 < \nu < 1$ s.t. as $N \rightarrow \infty$, for almost all undirected d -regular graph G , $i(G) = \nu d/2$ [6], it follows that *for almost all undirected d -regular graph G* , $\mathbf{E}(\tau) = O(N \log N)$.

5.1 Lurking through the proof of equation (3)

A useful proof technique to analyze the behavior of a Markov chains is to show that the process is stochastically dominated by that of a related process that is easier to analyze. Therefore, it would be useful to *dominate* the behavior of the LHN process on a digraph $\vec{G}$, with that of another similar process on the same digraph, which is easier to analyze. For instance, in the second process we allow transitions that create new mutants but we forbid some of the transitions removing mutants. Recall that the LHN process is a Markov chain $(C_t)_{t \geq 0}$ where C_t is the set of mutants at step t . To explain the coupling process, we will use instead of C_t the notation $\{X_t\}_{t \geq 0}$ for the LHN process and $\{Y_t\}_{t \geq 0}$ for the dominating process. To establish a coupling between LHN process $(X_t)_{t \geq 0}$ and process $(Y_t)_{t \geq 0}$, we must have that if $X_1 \subseteq Y_1$ then for all $t \geq 1$, $X_t \subseteq Y_t$.

Unfortunately for the discrete LHN process, the coupling technique does not is always work. In Fig 14 we show a simple example: let G be the undirected path with $V(G) = \{1, 2, 3\}$ and $E(G) = \{(1, 2), (2, 1), (2, 3), (3, 2)\}$ and let $(X_t)_{t \geq 1}$ and $(Y_t)_{t \geq 1}$ be LHN processes on G with $X_1 = \{2\}$ and $Y_1 = \{2, 3\}$. With probability $\frac{r}{2(r+2)}$, we have $X_2 = \{1, 2\}$. Therefore, the only possible value for Y_2 to contain Y_2 is $\{1, 2, 3\}$, which occurs with probability $\frac{r}{2(2r+1)}$. Therefore, any coupling between the two processes fails because

$$\Pr[Y_2 \not\subseteq Y'_2] \geq \frac{r(r-1)}{2(r+2)(2r+1)} > 0 \text{ for any } r > 1.$$

An essential and common trick in the area of stochastic processes is moving to continuous time. That is what is done in [15], given a set of mutants at time t , each vertex waits an amount of time before reproducing. For each vertex, the period of time is chosen according to an exponential distribution

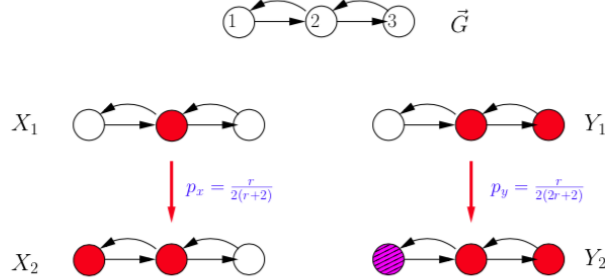


Figure 14: Counterexample for coupling for LHN processes.

with parameter equal to the vertex's fitness, independently of the other vertices. If the first vertex to reproduce is v then, as in the discrete-time version of the process, one of its out-neighbours w is chosen uniformly at random, then w is replaced by a copy of v , and the time at which w will next reproduce is exponentially distributed with parameter given by its new fitness r' . The discrete-time process is recovered by taking the sequence of configurations each time a vertex reproduces. Thus, the fixation probability of the discrete-time process is exactly the same as the fixation probability of the continuous-time process. Therefore, moving to the continuous-time model causes no harm. The analysis is easier in the continuous-time model because certain natural stochastic domination techniques apply in the continuous-time setting but not in the discrete-time setting.

The authors of [15] use the following result:

Coupling Lemma: *Given $\vec{G}$, let $X \subseteq Y \subset V(G)$ and $1 \leq r \leq r'$. For $t \geq 0$, let $\tilde{X}[t]$ and $\tilde{Y}[t]$ be the continuous-time process on $\vec{G}$ with mutant fitness r and r' , and with $\tilde{X}[0] = X$ and $\tilde{Y}[0] = Y$. Then, there exists a coupling between the two processes such that $\tilde{X}[t] \subseteq \tilde{Y}[t]$, $\forall t \geq 0$.*

The following three results follow from the coupling lemma:

- For any $\vec{G}$, if $0 < r \leq r'$ and $X \subseteq Y$, we have $f_{\vec{G},r}(X) \leq f_{\vec{G},r'}(Y)$.
- (Fixation probability is monotone with respect to mutant fitness) For any $\vec{G}$ and $0 < r \leq r'$, $f_{\vec{G},r} \leq f_{\vec{G},r'}$.
- (Subset domination) For any $\vec{G}$ and $0 < r \leq r'$, if $X \subseteq Y$ then $f_{\vec{G},r}(X) \leq f_{\vec{G},r'}(Y)$.

Then, using basic techniques from Markov chains, equation (3) is obtained.

6 Approximating the fixation probability

A Monte-Carlo algorithm for the LHN process basically simulates the process, and from the obtained data one gets a ϵ approximation to the fixation probability. This $\epsilon > 0$ is denoted as the tolerance (or error). In this section we consider the computational problem of using Monte Carlo algorithms to approximate the fixation probability.

Let us review some basic concepts on approximation algorithms (for further information see for example Chapters 9 and 13 in [53]). For all tolerance $\epsilon > 0$, and any input $x \in \{0, 1\}^*$, let $A(x)$ be a function. We say that $A(x)$ has *randomized approximation scheme* (RAS) if there is a polynomial-time algorithm (polynomial in $|x|$) that with probability at least $2/3$ returns a function $f(x)$ such that $(1 - \epsilon)f(x) \leq A(x) \leq (1 + \epsilon)f(x)$. A finer approximation scheme is the *fully polynomial randomized approximation scheme* (FPRAS), where as before we ask if there is a polynomial-time algorithm (polynomial in $|x|$, but this time we also require that it is also polynomial in $1/\epsilon$), that with probability at least $2/3$ returns a function $f(x)$ such that $(1 - \epsilon)f(x) \leq A(x) \leq (1 + \epsilon)f(x)$.

Since the expected absorption time of a LHN process on an N -vertex undirected graph is polynomial in N , and $f_{G,r} \geq 1/N$ for all undirected and connected graphs, we know the existence of a FPRAS for computing the value of $f_{G,r}$, using a Monte-Carlo method [14].

Given an undirected graph G on N vertices with tolerance $\epsilon \in (0, 1)$, Serna et al. [14] used a simple Monte Carlo method: for a given $\epsilon > 0$, they simulated $O(N^8 \epsilon^{-4})$ iterations of the LHN process until absorption. Therefore, they present a FPRAS for fixation probability in the LHN process.

Nowak et al. [12] used the same basic approach, but with two important differences. In [14] the simulation uses all steps on the Markov chain, including the ones that the random walk stays in the same state. In [12] the authors discard "ineffective steps", where the walk does not move from the state. Afterwards, they perform a sampling of each of those "effective steps", in time $O(\Delta)$. Their sampling algorithm is more involved and efficient than the sampling performed in [14]. The main result of Nowak et al. is that for undirected connected G with maximum degree Δ , their algorithm samples $O(n^2 \Delta \epsilon^{-2} \log(\Delta \epsilon^{-1}))$ effective steps, so the total running time is $T = O(n^2 \Delta \epsilon^{-2} \log(\Delta \epsilon^{-1}))$, and therefore they get a faster FPRAS and a more precise algorithm than the one in [14].

Goldberg et al. [25], using the stopping criterion given by Giakkoupis (see (2) and also [24]), gave a further improvement to approximating the fixation probability:

Theorem *Given an undirected graph G , with maximum degree Δ , average degree d , and tolerance $\epsilon \in (0, 1)$, for any $r > 1$, there exists a RAS for $f_{G,r}$ that simulates $T = O(\Delta d \epsilon^{-2} \log(d \epsilon^{-1}))$ effective steps of the LHN process.*

The authors also present an improved sampling algorithm that yields the following corollary to the previous theorem: *Let $r > 1$. Then there is an FPRAS for $f_{G,r}$ whose running time is $O(\Delta^2 d \epsilon^{-2} \log(d \epsilon^{-1}))$.*

Their new statement means that a single advantageous mutant can have a very high probability of reaching fixation, despite being heavily outnumbered in the initial configuration.

7 Conclusion and open problems

As stated in the abstract, we focus on the issues dealing with absorption for the LHN model on graphs. We did not cover the evolutionary games variant, which from the research point of view also presents nice open problems.

For the basic LHN model, in the last years quite a bit of work has been done. The basic important questions have been answered. There are some open questions as the effect of different types of mutants with different fitnesses. Also there is the issue of mobility: one may ask what happens if the topology of the network evolves in time, for instance edges or vertices appear or disappear.

Martin Nowak proposed different variations of his original model for different epidemiology situations. In particular his I&R graph evolutionary model for ecological problems is a nice generalization of the basic model, see for example [34]. There are quite a few theoretical possibilities for further research that emerge from those models.

References

- [1] Ben Adlam, Krishnendu Chatterjee, Martin A. Nowak. Amplifiers of selection. *Proceedings of the Royal Society*, A471: 20150114. 2015.
- [2] Cemil Azizoğlu, M and Egecioğlu, Ömer. The bisection width and the isoperimetric number of arrays. *Discrete Applied Mathematics*, 138 (2), 3–12, 2004.

- [3] Nicolas Bacaër. *A short history of mathematical population dynamics*. Springer, 2011.
- [4] Vladimir C. Barbosa, Raul Donangelo and Sergio R. Souza. Early appraisal of the fixation probability in directed networks. *Physical Review E*, **82**, 4, 046114, 2010.
- [5] Irénée-Jules Bienaymé. De la loi de multiplication et de la durée des familles. *Soc. Philomath. Paris Extraits*, 5, 37–39, 1845.
- [6] Béla Bollobás. The isoperimetric number of random regular graphs. *European Journal of Combinatorics*, 9(3):241–244, 1988.
- [7] Maury Bramson, David Griffeath. On the Williams–Bjerknes tumour growth model. *Annals of Probability*, 9, 173–185, 1981.
- [8] Mark Broom, Jan Rychtář and Brian Stadler. Evolutionary dynamics on small order graphs. *Journal Interdisciplinary Mathematics*, 12, 129–140, 2009.
- [9] Mark Broom, Cristoforos Hadjichrysanthou, and Jan Rychtář. Evolutionary games on graphs and the speed of the evolutionary process. *Proceedings of the Royal Society A*, 466(2117):1327–1346, 2010.
- [10] Mark Broom and Jan Rychtář. An analysis of the fixation probability of a mutant on special classes of non-directed graphs. *Proceedings of the Royal Society A*, 464(2098):2609–2627, 2008.
- [11] William E. Castle. The laws of Galton and Mendel and some laws governing race improvement by selection. *Proceedings of the American Academy of Arts and Sciences*, 35: 233–242, 1903.
- [12] Krishnendu Chatterjee, Rasmus Ibsen-Jensen, and Martin A. Nowak. Faster Monte-Carlo algorithms for fixation probability of the Moran process on undirected graphs. In: *Proc. 42nd International Symposium on Mathematical Foundations of Computer Science (MFCS 2017)*, vol. 83 of LIPIcs, 61:1–61:13. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2017.
- [13] Charles Darwin. *On the Origin of Species by Means of Natural Selection*. John Murray, 1850.

- [14] Josep Díaz, Leslie Ann Goldberg, George B. Mertzios, David Richerby, Maria J. Serna, and Paul G. Spirakis. Approximating fixation probabilities in the generalized Moran process. *Algorithmica*, 69, 78–91, 2014.
- [15] Josep Díaz, Leslie Ann Goldberg, David Richerby, Maria J. Serna. Absorption Time of the Moran Process. *Random Structures and Algorithms*, 59:1, 137–159, 2016.
- [16] Josep Díaz, Leslie Ann Goldberg, George B. Mertzios, David Richerby, Maria J. Serna, and Paul G. Spirakis. On the fixation probability of superstars. *Proceedings of the Royal Society A*, 469(2156):20130193, 2013.
- [17] Richard Durrett. *Random Graph Dynamics*. Cambridge University Press, 2007.
- [18] Richard Durrett. Some features of the spread of epidemics and information on a random graph. *Proceedings of the National Academy of Sciences*, 107(10):4491–4498, 2010.
- [19] Leonhard Euler. Recherches generales sur la mortalité et la multiplication : la multiplication du genre humain. *Memoires de l’Academie des Sciences de Berlin*, 16, 144–164, 1767.
- [20] Ronald A. Fisher. On the dominance ratio. *Proc. Royal Society of Edinburgh*, 42: 321–341, 1922.
- [21] Andreas Galanis, Andreas Göbel, Leslie Ann Goldberg, John Lapinskas and David Richerby. Amplifiers for the Moran Process. *Journal of the ACM*, 64 (1): 5:1–5:90, 2017.
- [22] Francis Galton. *Natural Inheritance*. McMillan, 1889.
- [23] Francis Galton and Henry W. Watson. On the probability of the extinction of families. *Journal of the Royal Anthropological Institute*, 4: 138–144, 1875.
- [24] George Giakkoupis. Amplifiers and suppressors of selection for the Moran process on undirected graphs. CoRR abs/1611.01585, 2016.
- [25] Leslie Ann Goldberg, John Lapinskas and David Richerby. Phase Transitions of the Moran Process and Algorithmic Consequences. *Random Structures and Algorithms*, 56, 3, 597–647, 2020.

- [26] Leslie Ann Goldberg, John Lapinskas, Johannes Lengler, Florian Meier, Konstantinos Panagiotou and Pascal Pfister. Asymptotically optimal amplifiers for the Moran process. *Theoretical Computer Science*, 758, 73–93, 2019.
- [27] Peter Grassberger and Alberto de la Torre. Reggeon field Theory (Schlogl’s first model) on a lattice: Monte Carlo calculations of critical behaviour. *Annals of Physics*, 122, 373–396, 1979.
- [28] Geoffrey Grimmett. *Percolation*. (2nd. edition), Springer, 1999.
- [29] Bruce Hajek. Hitting-time and occupation-time bounds implied by drift analysis with applications. *Advances in Applied Probability*, 14(3): 502—525, 1982.
- [30] Goffrey H. Hardy. Mendelian Proportions in a Mixed Population. *Science*, 28 (706): 49—50, 1908.
- [31] Larry H. Harper. Optimal numberings and isoperimetric problems on graphs. *Journal of Combinatorial Theory*, 1(3): 385–393, 1966.
- [32] Theodore E. Harris. Contact interactions on a lattice. *The Annals of Probability*, 2:969–988, 1974.
- [33] Herbert W. Hethcote. The Mathematics of Infectious Diseases. *SIAM Review*, 42 (4): 599–653, 2000.
- [34] Rasmus Ibsen-Jensen, Krishnendu Chatterjee, and Martin A. Nowak. Computational complexity of ecological and evolutionary spatial dynamics. *Proc. National Academy of Science*, 112 (51), 15636–15641, 2015. Proofs available at <http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1511366112/-/DCSupplemental>.
- [35] Alastair Jemieson-Lane and Christoph Hauert. Fixation probabilities on superstars, revised and revised. *Journal of Theoretical Biology*, 382:44–56, 2015.
- [36] Harry Kesten. What is percolation? *Notices of the AMS* , 53(5): 572–573, 2006.
- [37] William O. Kermack and Anderson G. McKendrick. A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society of London*, Series A, Vol. 115, 700–721, 1927.

- [38] William O. Kermack and Anderson G. McKendrick. A Contribution to the Mathematical Theory of Epidemics II. The Problem of Endemicity. *Proceedings of the Royal Society of London*, Series A, Vol. 138, 55–83, 1932.
- [39] William O. Kermack and Anderson G. McKendrick. A Contribution to the Mathematical Theory of Epidemics III. Further Studies of the Problem of Endemicity. *Proceedings of the Royal Society of London*, Series A, Vol. 141, 94–122, 1933.
- [40] Jean-Baptiste Lamarck. *Philosophie zoologique*. Musée d’Histoire Naturelle, 1809.
- [41] Erez Lieberman, Christoph Hauert, and Martin A. Nowak. Evolutionary dynamics on graphs. *Nature*, 433(7023):312–316, 2005. Proofs available at <http://www.nature.com/nature/journal/v433/n7023/full/nature03204.html>.
- [42] Thomas M. Liggett. *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Springer, 1999.
- [43] Alfred J. Lotka. *Elements of Physical Biology*. Williams and Wilkins Company, 1925.
- [44] Thomas R. Malthus. *An Essay on the Principle of Population*. London, 1798.
- [45] Takeo Maruyama. A simple proof that certain quantities are independent of the geographical structure of population. *Theoretical Population Biology*, 5(2): 148–154, 1974.
- [46] Joanna Masel. Genetic drift. *Current Biology*, 21(20): R837–R838, 2011.
- [47] Naoki Masuda and Hisashi Ohtsuki. Evolutionary dynamics and fixation probabilities in directed networks. *New Journal of Physics*, 11(3): 33012, 2009.
- [48] George Mendel. Versuche über Pflanzen-Hybriden. Verhandlungen des naturforschenden Vereines in Brünn, IV: 3-47, 1865. (English translation in: <http://www.mendelweb.org/Mendel.html>)
- [49] George B. Mertzios, Sotiris E. Nikolettseas, Christoforos Raptopoulos, and Paul G. Spirakis. Natural models for evolution on networks. *Theoretical Computer Science*, 477:76–95, 2013.

- [50] George B. Mertzios and Paul G. Spirakis. Strong bounds for evolution in networks. *Journal of Computer and System Science*, 97: 60–82, 2018.
- [51] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.
- [52] Bojan Mohar. Isoperimetric numbers of graphs. *Journal of Combinatorial Theory, Series B*, 47(3), 274–291, 1989.
- [53] Cristopher Moore and Stephan Mertens. *The Nature of Computation*. Oxford University Press, 2011.
- [54] Patrick P. Moran. Random processes in genetics. *Proceedings of the Cambridge Philosophical Society*, 54(1):60–71, 1958.
- [55] Patrick P. Moran. *The statistical process of evolutionary theory*. Clarendon Press, Oxford, 1962.
- [56] Jaroslav Nešetřil and Helena Nešetřilová. Remarks on Mendel’s mathematics. *Folia Mendeliana*, 53(1-2): 5–22, 1917.
- [57] Mark E. J. Newman. Spread of epidemic disease on networks. *Physical Review E*, 66 (1): 16128–16139, 2002.
- [58] Martin A. Nowak. *Evolutionary Dynamics: Exploring the Equations of Life*. Harvard U.P., 2006.
- [59] John Parrington. Charles Darwin: Revolution of evolution. *Socialist Review*, 333, Feb. 2009.
- [60] Alfred Russell. On the Tendency of Species to form Varieties. *Journal of the Proceedings of the Linnean Society: Zoology*, Vol. 3: 53–62, 1858.
- [61] Jan Rychtář and Brian Stadler. Evolutionary dynamics on small-world networks. *International Journal of Computer and Mathematical Sciences*, 2(1), 1–4, 2008.
- [62] Paulo Shakarian, Patrick Roos, and Anthony Johnson. A review of evolutionary graph theory with applications to game theory. *Biosystems*, 107:66–80, 2012.
- [63] Paulo Shakarian and Patrick Roos. Fast and deterministic computation of fixation probability in evolutionary graphs. In: *Proceedings of the*

6th IASTED International Conference on Computational Intelligence and Bioinformatics, 97–104, 2011.

- [64] Christine Taylor, Yoh Iwassa and Martin Nowak. A symmetry of fixation times in evolutionary dynamics. *Journal Theoretical Biology*, 243(2):245–251, 2006.
- [65] Pierre F. Verhulst. Notice sur la loi que la population suit dans son accroissement. *Corr. Math.et Phys.*, X, 268–281, 1838.
- [66] Vito Volterra. Variazioni e fluttuazioni del numero d’individui in specie animali conviventi. *Mem. R. Accad. Naz. dei Lincei*, 2: 31–113, 1926.
- [67] Wilhelm Weinberg. Über den Nachweis der Vererbung beim Menschen. *Jahreshefte des Vereins für vaterlandische Naturkunde in Württemberg*, 64: 368–382, 1908.
- [68] Sewall G. Wright. Evolution in Mendelian populations. *Genetics*, 16: 97–159, 1931.
- [69] Sewall G. Wright. Statistical genetics and evolution. *Bulletin American Mathematical Society*, 48 (4): 223–246, 1942.