Universitat Politècnica de Catalunya - Barcelona Tech
Programa de Doctorat de Matemàtica Aplicada

Centre de Recerca Matemàtica

Evolutionary dynamics of populations with genotype-phenotype map

by

Esther Ibáñez Marcelo

Thesis Advised by:

Tomás Alarcón Cor

Barcelona, December 2014

Abstract

In this thesis we develop a multi-scale model of the evolutionary dynamics of a population of cells, which accounts for the mapping between genotype and phenotype as determined by a model of the gene regulatory network. We study topological properties of genotype-phenotype networks obtained from the multi-scale model. Moreover, we study the problem of evolutionary escape and survival taking into account a genotype-phenotype map.

An outstanding feature of populations with genotype-phenotype map is that selective pressures are determined by the phenotype, rather than genotypes. Our multi-scale model generates the evolution of a genotype-phenotype network represented by a pseudo-bipartite graph, that allows formulate a topological definition of the concepts of robustness and evolvability.

We further study the problem of evolutionary escape for cell populations with genotype-phenotype map, based on a multi-type branching process. We present a comparative analysis between genotype-phenotype networks obtained from the multi-scale model and networks constructed assuming that the genotype space is a regular hypercube. We compare the effects on the probability of escape and the escape rate associated to the evolutionary dynamics between both classes of graphs.

We further the study of evolutionary escape by analysing the long term survival conditioned to escape. Traditional approaches to the study of escape assume that the reproduction number of the escape genotype approaches infinity, and, therefore, survival is a surrogate of escape. Here, we analyse the process of survival upon escape by taking advantage of the fact that the natural setting of the escape problem endows the system with a separation of time scales: an initial, fast-decaying regime where escape actually occurs, is followed by a much slower dynamics within the (neutral network of) the escape phenotype. The probability of survival is analysed in terms of topological features of the neutral network of the escape phenotype.

Keywords: *genotype-phenotype map, multi-scale model, evolutionary dynamics, complex networks, escape probabilities, branching process*

Resum

En aquesta tesi es desenvolupa un model multi-escala de la dinàmica evolutiva d'una població de cèl·lules, tenint en compte la correspondència entre el genotip i el fenotip determinat per un model de la xarxa de regulació genètica. Estudiem les propietats topològiques de les xarxes genotip-fenotip obtingudes a partir del model multi-escala. D'altra banda, s'estudia el problema de la fugida evolutiva i la supervivència, tenint en compte una aplicació entre genotip i fenotip.

Una característica destacable de les poblacions amb aplicació genotip-fenotip és que les pressions selectives actuen sobre els fenotips, en lloc dels genotips. El nostre model multi-escala genera l'evolució d'una xarxa genotip-fenotip representada per un graf pseudo-bipartit, el qual permet formular una definició topològica dels conceptes de robustesa i capacitat evolutiva.

A més a més, estudiem el problema de fugida evolutiva de poblacions de cèl·lules amb una aplicació genotip-fenotip, basat en en un procés de ramificació multi-tipus. Presentem un anàlisi comparatiu entre les xarxes de genotip-fenotip obtingudes a partir del model multi-escala i les xarxes construïdes assumint un espai de genotips de tipus hipercub regular. Comparem els efectes de la probabilitat de fugida i la freqüència d'escapament associades a la dinàmica evolutiva entre ambdues classes de grafs.

Anem més enllà de l'estudi de fugida evolutiva mitjançant l'anàlisi de la supervivència a llarg plaç condicionat a fugir. Els enfocaments tradicionals per a l'estudi de la fugida o escapament suposen una taxa de reproducció en el genotip de fugida propera a infinit. Per tant, la supervivència és equivalent a la fugida. Aquí analitzem el procés de supervivència suposant fugida aprofitant el fet que l'entorn natural del problema de fugida dota al sistema amb una separació d'escales de temps: un règim inicial, de temps ràpid, on la fugida realment es produeix; seguit d'una dinàmica molt més lenta dins de la (xarxa neutra del) fenotip de fugida. La probabilitat de supervivència s'analitza en termes de les característiques topològiques de la xarxa neutra del fenotip de fugida.

Acknowledgments - Agraïments

Moltes gràcies a totes aquelles persones que han fet possible aquesta tesi, tant a aquelles que han mostrat el seu suport continu (acadèmicament, personalment o ambdues) com a aquelles que m'han ajudat puntualment, ja sigui donant idees, suport moral o polint detalls.

Cal dir que aquest projecte no hagués estat possible sense la proposta de tesi del meu director, la beca concedida pel Centre de Recerca Matemàtica i la participació en el projecte MTM2011-29342 del *Ministerio de Ciencia e Innovación.*

Thank you very much to all people which has made possible this thesis, both those who have shown their continued support (academically, personally or both) to those who helped me occasionally, whether giving ideas, moral support or polishing details.

It must be said that this project would not have been possible without the thesis design by my advisor, the grant awarded by the Centre de Recerca Matemàtica *and the participation in the project MTM2011-29342 of the Ministry of Science and Innovation.*

Me gustaría dar las gracias a mi familia la cual siempre me ha dado su apoyo en todo momento y en todas las decisiones tomadas durante estos años. Gracias por concederme esta formación a la cual mucha gente desgraciadamente no tiene acceso.

Contents

Abstract

Acknowledgments **v**

1 Introduction **1**

1.1 Background . . . 3

1.1.1 Characteristics of genotype-phenotype map . . . 3

1.1.2 Previous models and genotype-phenotype map . . . 4

1.1.3 Evolutionary escape . . . 8

1.2 Methodology . . . 10

1.2.1 Network models, properties and tools . . . 11

1.2.2 Population dynamics . . . 15

1.3 Aims . . . 19

1.4 Outline . . . 19

2 The model **21**

2.1 Model formulation . . . 21

2.1.1 Microscale: Intracellular dynamics of the GRN . . . 23

2.1.2 Mesoscale: Population dynamics . . . 24

2.1.3 Macroscale: Evolutionary dynamics of the genotype-phenotype map . . 25

2.1.4 Evolutionary dynamics of the genotype-phenotype map . . . 26

2.1.5 Remarks of genotype-phenotype network . . . 28

3 The topology of robustness and evolvability in evolutionary systems with genotype-phenotype map **31**

3.1 Network dynamics: degree distributions and parameters . . . 31

3.2 Topological characterisation of phenotypic robustness . . . 34

3.3 Topological characterisation of evolvability . . . 38

3.3.1 Phenotype network . . . 38

3.3.2 Evolution of evolvability: Emergence of giant connected component . . 38

3.3.3 The phenotype network is a small world . . . 39

3.4 Robustness of evolvability . . . 42

3.5 Conclusions . . . 45

4 Evolutionary escape on complex genotype-phenotype networks **51**

4.1 Evolutionary escape . . . 51

4.2 Mathematical model . . . 52

4.2.1 Genotype-phenotype network . . . 53

4.2.2 Population dynamics 53

4.3 Results 57

4.3.1 Connectivity structure of genotype-phenotype networks 57

4.3.2 The escape probability exhibits higher degree of heterogeneity in B graphs 59

4.3.3 Escape dynamics: asymptotic behaviour of escape time probability . . . 62

4.3.4 Escape dynamics: escape time probability conditioned to escape 65

4.4 Discussion & conclusions 65

5 Surviving evolutionary escape on complex genotype-phenotype networks 69

5.1 Surviving evolutionary escape 69

5.1.1 Population dynamics 70

5.2 Results 77

5.2.1 Fast dynamics: Escape properties 77

5.2.2 Slow dynamics: Survival probability 81

5.3 Conclusions 82

6 Summary 85

6.1 Discussion and Conclusions 85

6.2 Future work 86

7 Appendix 89

7.1 Generalization of GRN 89

7.2 Graph definitions 91

7.3 Severe selective pressure 92

7.4 Other definitions: 94

7.5 Escape probabilities in the discrete-time branching process 94

Bibliography 96

Chapter 1

Introduction

Understanding how complex global behavioural traits (phenotypes) emerge from the interactions between individual genes and their products is one of the major challenges of modern biology. Phenotypes arise from networks of interactions between genes and gene products, which ultimately regulates gene expression. These networks of gene regulation are dynamical systems whose complexity partially stems from the fact that they are non-linear, high-dimensional dynamical systems but, also and foremost, because they are shaped up by evolution by natural selection. Natural selection acts upon gene regulatory networks (GRNs) so that they evolve to exhibit properties such as robustness (i.e. resilience of the phenotype against genetic alterations [105]), canalisation (i.e. the ability for phenotypes to increase their robustness as time progresses [101, 102]) or non-uniqueness (i.e. different genotypes leading to the same phenotype [28, 27]). Moreover, they are under the influence of noise both of internal (molecular noise in the regulatory system itself) and external (unpredictable changes in the environment) characters.

The gap between genotype and phenotype poses a daunting problem in evolutionary theory. Whereas genes are the entities passed on from one generation to the next and their frequencies measured over populations (the remit of population genetics), selective pressures act at the level of phenotypes [31]. Thus, assigning fitness values to genes (mutant variants, different alleles, etc.) is not, in general, the valid approach. A more appropriate approach consists of considering models that take into account the genotype and the phenotype, or a model where selection acts at the level of phenotypes.

In order to study the properties and issues regarding the genotype-phenotype maps, several models have been studied [108]: RNA, circuits of gene regulation and metabolic networks. Concerning RNA molecules as model of the genotype-phenotype map [39], the *genotype* of each RNA molecule consists of a sequence of nucleotides. There are four such nucleotides, so for sequences of length L, the size of the genotype space is 4^L. The *phenotype* of the RNA molecule refers to the fold or three-dimensional conformation, which determines the biochemical function of the molecule. The folded structure of an RNA sequence, which is a proxy for its phenotype, but still lies far from defining its function; is determined by the sequence (genotype) in a many-to-one way, i.e. many different genotypes give rise to the same phenotype. Such non-uniqueness has led to the concept of the neutral network [67, 88]: a network whose nodes correspond to genotypes, all with the same phenotypes, with edges between those nodes which differ by only one nucleotide [105]. This system has been used extensively in the study of the genotype-phenotype map, in particular, those issues regarding its evolutionary properties, such as the role of phenotypic robustness in evolvability and adaptation [107, 108]. Recently, the topology of the RNA genotype-phenotype space, composed by an intermingled set of neutral networks, has been analysed [2].

Gene regulatory networks (GRNs) have also been used to analyse properties of the genotype-phenotype map, in particular several variants of the model introduced by Wagner to study phenotype plasticity [104]. These models are dynamical systems for the expression levels of the corresponding genes and are characterised by two elements: a matrix whose entries specify the character of the interaction between two genes (usually, activation or inhibition) and, possibly, the intensity of such interaction, and a series of rules for the time evolution of the expression levels of the genes involved. The entries of the corresponding matrix are the *genotype* of the GRN. The *phenotype* is the steady-state gene expression yielded by the dynamics. There are many such genotypes that produce the same phenotype, which allows to extend the concept of neutral network to GRN, where nodes correspond to different matrices (producing the same steady-state) and links exist between nodes if the corresponding matrices differ only in one regulatory interaction. Models of GRNs have been used to study phenotype plasticity [104], robustness and innovation in circuits of gene regulation [28, 27], and canalisation [94, 14], among other issues.

Metabolic networks are the third class of systems that have been used to assess properties regarding robustness and innovation [71, 80, 85, 81]. They are formed by thousands of enzyme-catalysed chemical reactions. These networks are responsible for supplying cells with energy (i.e. ATP) and the molecule building blocks cells need to grow. The *genotype* space for this system consists of the space of all the possible metabolic networks, whereas the *phenotype* corresponds to the secondary metabolites the metabolic network is able to synthesise, the molecules they can use as energy sources, the ability to detoxify certain waste products, etc. [108]. Innovations in these aspects not always appear as the result of gene mutations that give rise to new enzymes. They can also arise through novel combinations and utilisation of existing elements.

1.1 Background

In this Section we introduce the main characteristics of a genotype-phenotype map (Section 1.1.1). It includes properties such that, robustness, evolvability, canalisation and convergence. We then move on to describe previous models that have been used to model properties described previously (Section 1.1.2). Finally, at the end of this Section we give a brief introduction to the concept of evolutionary escape.

1.1.1 Characteristics of genotype-phenotype map

The definition of the map relating genotypes to phenotypes is one of the first problems that we have to deal with. In order to study and try to shed light on to some properties of this map, we first consider some generic properties which arise in the study of systems endowed with such structure.

1.1.1.1 Robustness and innovation (evolvability)

We can define the *robustness* of a system as the capacity of a system to retain its state in presence of perturbations (mutations, noises, environmental changes, etc).

As far as the present work is concerned, we refer to genetic robustness or mutational robustness, i.e. the ability to retain the same phenotype upon genetic mutations. Our description of robustness is based on the concept of the so-called *neutral network* [105, 106].

Another important concept is *innovation or evolvability*, which refers to the ability to adaptation, and it is defined as the capacity of a system for adaptive evolution. It is very important to produce genetic diversity but also viable individuals who will be well adapted to the system.

Evolutionary adaptation of a population to, say, a new environment requires evolutionary innovation or evolvability, namely, new, better adapted phenotypes must arise within the population. Such innovations are achieved over many generations by means of genetic mutations. Most of these mutations are known to be either detrimental or neutral [37, 86, 96]. However, rarely, one or several mutations produce new phenotypes that are better adapted to the new conditions. These rare mutations are the drivers of Darwinian evolution.

Intuitively, robustness and evolvability seem to be in conflict, since, by definition, robust phenotypes are resilient to the effects of mutations. However, it is impossible to understand robustness without innovation. There is mounting evidence which hints otherwise, i.e. phenotypic robustness facilitates evolutionary innovation [107, 108]. According to this view, the genotype space has two generic properties which allow to reconcile robustness and evolvability. The first of these properties is the existence of neutral networks, i.e. large connected sub-networks of the space of genotypes which can be navigated in small, mutation-induced steps with no change in phenotype. The second of these features regards the so-called *genotype neighbourhood*, i.e. the set of genotypes accessible from any one genotype in a prescribed number of mutations. A simple measure of how phenotypically variable is a genotype is the number of different phenotypes reached in a genotype neighbourhood, i.e. how many different phenotypes are easily accessible by mutation(s) [106]. The first of these properties allows preservation of the phenotype with changing genotypes, thus creating a substantial amount of divergence within a population, whereas the second property, neighbourhoods, allow genotypes within a neutral network to explore different phenotypes.

The argument to support that robustness favours evolvability rather than hindering it can be summarised as follows [107, 108]: Large neutral networks containing many genotypes with the same phenotype have, collectively, a much larger neighbourhood, i.e. that corresponding to the (disjoint) union of all the neighbourhoods of all the nodes within the neutral network, than non-robust genotypes, which will be typically isolated in genotype space.

These issues regarding the relation between robustness and evolvability may seem to be of a rather theoretical character. However, they are fundamental to understanding the evolutionary dynamics of systems such as cancer [58, 98]. Kitano [58] points out that cancers can be seen as robust systems with respect to its proliferate potential, which is maintained in the presence of several anti-cancer therapies and environmental factors, such as the effect of the immune system. Kitano's thesis is that, as it occurs in systems engineering, the development of robustness with respect to some properties leads to fragility with respect to others which should be exploited as therapeutic targets. Whereas this is a valid and useful point of view, it may fall short of a full picture, since it fails to take into account that also tumours highly evolve [98]. According to results recently reviewed by Tian et al. [98], the evolutionary dynamics of tumours is such that evolvability is greatly enhanced. These two results can be unified under the above paradigm that robustness favours evolvability which implies that, to achieve effective control (therapeutic) strategies, both robustness and evolvability must be tackled [98].

1.1.1.2 Canalisation and convergence

We define *canalisation* or genetic assimilation in the context of genetics as the ability of a population to produce the same phenotype regardless of variability of its environment or genotype. In other words, it means robustness, assimilation of a change that after some generations can survive without the mutation or change that we have imposed some generations before. This term, canalisation, was coined by Conrad Waddington in the 1940's, also is named by Waddington *buffering of the genotype* in [101].

In opposition to canalisation, that shows how some phenotypes are fixed in the population and become more robust as time goes by, *convergence* in evolution refers to the phenomenon whereby a number of different species of different lineages achieve the same biological traits. For example, the common ancestor of flying insects and birds does not have wings. In our case we can relate it with the fact that two or more different unrelated genotypes can have the same phenotype.

Another term introduced by Waddington is the term *epigenetics*. It can take many different meanings depending on the context. This term was first proposed by Conrad Waddington in the 1940's in [102]. He uses it to designate the study of the processes by which the genetic information of an organism, defined as genotype, interacts with the environment in order to produce its observed traits, defined as phenotype. The goal of Waddington with epigenetics was to provide insight into gene-environment interactions that influence development and embryology, but had no molecular insights to consider.

We refer to *epigenetic stability* as genetic canalisation (or assimilation). It is different than *phenotypic plasticity*, that is the ability of the genotype to change the phenotype in response to changes in the environment (environmental aspect of canalisation).

1.1.2 Previous models and genotype-phenotype map

In this Section, we state the definition of the genotype-phenotype map we use through out this thesis. Our definition is based on the idea, first proposed by Stuart Kauffman [56], according to which gene regulatory networks are dynamical systems and that phenotypes or differentiated states correspond to the stable attractors of these dynamical systems. Furthermore, previous models of robustness and canalisation, relevant to our own work are briefly described. Among these, we summarise the model of Wagner in [104], which consists of a phenotype-genotype map based on a dynamical model of the gene regulatory network. Also, we are going to comment on the work of Ciliberti et al. [27, 28] where a geometrical description of the genotype space was introduced. They studied robustness and evolvability in terms of the properties of a *metagraph* (a graph of graphs). Others authors, such as Siegal and Bergman in [94], define a more complex model introducing changes in the dynamics of the gene regulatory network. In [14], Bergman and Siegal discuss how gene silencing induces a greater range of phenotypes, that is kept so-called buffered during evolution of evolutionary capacitors, such as HSP90 [83].

1.1.2.1 The model of Wagner

Wagner [104] has formulated a model, which has been the basis of much later work on robustness and evolvability [27, 28, 14, 94], including ours. For later reference, we summarise here its main ingredients.

We consider a set of N genes, whose products mutually regulate each other's expression at the transcriptional level. These regulatory interactions are represented as a network, the so-called gene regulatory network.

A gene network is represented by a dynamical system whose state variables correspond to expression states of the genes involved in the associated regulatory system. They are denoted as,

$$\vec{g}(t) := (g_1(t), \dots, g_N(t)),$$

where $g_i(t)$ is the expression state of the ith gene at some time $t \geq 0$. It is assumed that $g_i(t)$ only can assume two values: $+1$ or -1, corresponding to the gene i is expressed or not expressed, respectively, at time t. The gene expression state $\vec{g}(0)$ at time $t = 0$ is called the initial expression state.

Starting from the initial gene-expression pattern, $\vec{g}(0)$, which can be interpreted as an inherited developmental program, cross and auto-regulatory interactions among network genes cause the expression state to change. These changes are modelled by means of the following dynamical system:

$$g_i(t+\tau) = \sigma\left[\sum_{j=1}^{N} w_{ij} g_j(t)\right] = \sigma[h_i(t)], \text{ where} \tag{1.1}$$

$$\sigma(x) = \begin{cases} -1 & \text{for } x < 0, \\ 1 & \text{for } x > 0, \\ 0 & \text{for } x = 0 \end{cases} \tag{1.2}$$

$h_i(t)$ represents the sum of the regulatory effects of all network genes on gene i and the real constants w_{ij} represents the strength of regulatory interaction of gene j with gene i. That is the degree of transcriptional activation ($w_{ij} > 0$) or repression ($w_{ij} < 0$). The set of these strengths is the connectivity matrix, w. Wagner use as a parameter in this model the fraction of entries different from zero of $w = (w_{ij})$, denoted by c, which is denoted as the *connectivity density* of the network. The dynamics of (1.1) will lead to the attainment of an equilibrium gene expression state, which may be a fixed point of (1.1) or a limit cycle. Wagner [104] only considers fixed-point equilibria, denoted by $\vec{g}(\infty)$. According to Kauffman's theory [56],

$\vec{g}(\infty)$ determines the phenotype associated to the genotype, defined by the matrix $w = (w_{ij})$. He also defines an optimal phenotype $\vec{g}^{opt}(\infty)$, so he makes the assumption that an optimal equilibrium gene-expression state exists for networks acting in a developmental process. For this reason, it is defined fitness of an individual. First, he defines a measure for the distance d between the equilibrium state $\mathbf{S}(\infty)$ attained by a network and the optimal $\vec{g}^{opt}(\infty)$ by,

$$d(\vec{g}^{opt}(\infty), \vec{g}(\infty)) := \frac{1}{2} - \frac{1}{2N}\sum_{i=1}^{N} g_i^{opt}(\infty) g_i(\infty)$$

also this measure is known as the Hamming distance. It counts the number of different states between both expressions and normalises it. Finally, based on d, the fitness of an individual is defined by a Gaussian fitness function as

$$\exp\left(-\frac{d[\vec{g}^{opt}(\infty), \vec{g}(\infty)]^2}{s}\right).$$

s represents the strength of selection, small values of s implying strong selection against deviations from the optimal state.

Wagner uses a binomial distribution, $B(N, p)$ in order to choose how many genes in the initial state are going to be expressed and connectivity matrices w are defined by a probability distribution $\rho(w)$ of regulatory interaction strengths, with a mean fraction c of connectivities different from 0.

1.1.2.2 Modelling robustness, canalisation and innovation

After the original definition of robustness and canalisation by Conrad Waddington at 1942 [101], lacking a proper mechanistic explanation, these issues lie dormant for many years. Until Rutherford and Lindquist [83], who found the first evidence for an explicit molecular mechanism that assists the process of evolutionary change in response to the environment. They found that inactivation of the molecular chaperone HSP90 induces phenotypic variation with specific variants depending on the particular genetic background. These genetic variants that prior to HSP90 inactivation remained silent, after mutation of HSP90 are able to produce phenotypic variants that after some generations, subject to evolution, rapidly became independent of the HSP90 mutation and the new phenotypes are fixed in the population even when HSP90 activity is restored.

During the development of a multicellular organism from a zygote, a large number of epigenetic interactions take place on every level of suborganism organisation. This raises the possibility that the system of epigenetic interactions may compensate or "buffer" some of the changes that occur as mutations on its lower levels, and thus stabilise the phenotype with respect to mutations. This hypothetical phenomenon will be called *epigenetic stability* [104, 103]. After long periods of stabilising selection [1], the fraction of mutations causing changes in the gene-expression patterns is substantially reduced in the model, thus leading to increased epigenetic stability. It is discussed that only *epistatic* (non-linear) gene interactions can cause such change in epigenetic stability. The relation of epigenetic stability to developmental canalisation is outlined [104, 103].

[1] *Stabilising selection*: It is fixed a final state (optimal phenotype) that a genotype have to reach, it is assigned a probability (survival probability) using the difference between the optimal phenotype and the final state of each genotype to discard or maintain this genotype alive. So, stabilising selection acts as selection of "good" genotypes in this way. It is like a kind of genotype selection.

Following the work of Rutherford and Lindquist, the interest in canalisation was rekindled which gave rise to a number of theoretical works. Among those, Wagner and co-workers proposed a theoretical model [103] to address a number of issues: How changes on the lowest, submicroscopic levels of their production system are translated onto the macroscopic level of the phenotype? How do epigenetic interactions influence the effect of mutant genes and their gene products on the phenotype? Can they absorb or buffer some such effects? If yes, can such an ability to "protect" the phenotype from mutations be subject to evolutionary change? What direction would such change take?

Wagner [103] proposed a model (see Section 1.1.2.1) of population dynamics with a phenotype-genotype map based on a dynamical model of the gene regulatory network where the phenotype is taken to be the corresponding steady state pattern of gene activation. Moreover, a fitness function was defined, such that individuals were more likely to reproduce if their phenotype is close to an (arbitrary chosen) optimal phenotype. Thus genetic variants yielding phenotypes which differ from the optimal one are very likely to go extinct [104].

The likelihood that a mutation affecting regulatory interactions changes the network's gene-expression pattern can be viewed as measure for the epigenetic stability of that segment of a developmental pathway.

Wagner [104] showed that his model exhibits the phenomenon of epigenetic stability, whereby phenotype-changing mutations become substantially rarer as time proceeds. Wagner [104] concluded that epigenetic stability is an evolvable property and that genotypes with low epigenetic stability are selected against. However, he pointed to *stabilising selection* (i.e. there exists an optimal phenotype on which fitness depend) as essential to the emergence of epigenetic stability.

He concludes mutations in networks with low stability and high fitness produce many genotypes with low fitness. However, they also produce some genotypes with high fitness and higher stability than the original network. On the other hand, networks with low fitness are eliminated and networks with high fitness and high stability accumulate because they produce fewer suboptimal variants. Thus, networks with high stability and high fitness replace networks with low stability and high fitness in a population.

Further development of the subject led to the work of Ciliberti et al. [27, 28] where a geometrical description of the genotype space was introduced. They proposed to describe the genotype space in terms of a metanetwork where nodes correspond to a particular gene regulatory network (GRNs). Edges in this metagraph link genotypes (GRNs) which are separated by one single mutation. Robustness and evolvability are studied in terms of the properties of the metagraph. The model for the genotype-phenotype map is the same as in [28].

Two types of robustness are considered (in both cases, the robust feature is the equilibrium of pattern gene expression $\vec{g}(\infty)$ in the network):

- Robustness to mutations: it corresponds to the robustness of the phenotype, defined as the steady state of the gene regulatory network ($\vec{g}(\infty)$, in the terminology of [28]) upon genetic mutations, which, following [104], correspond to changes in regulatory interactions, i.e. in the weights of the interaction matrix that defines the gene regulatory network.

- Robustness to noise (or non-genetic perturbations): it corresponds to the robustness of $\vec{g}(\infty)$ to random changes in the initial state $\vec{g}(0)$ or random changes in the trajectory from $\vec{g}(0)$ to $\vec{g}(\infty)$.

From these definitions some questions emerge:

- How does change $\vec{g}(\infty)$ from changes in $\vec{g}(0)$ (initial state of genes)?
- How many changes in $\vec{g}(0)$ we have to do in order to obtain a different $\vec{g}(\infty)$?
- How does changes in the gene expression trajectory, between $\vec{g}(0)$ to $\vec{g}(\infty)$, change $\vec{g}(\infty)$?

Recall that whereas *robustness* represents the resilience of the phenotype to gene mutations, *innovation* or *evolvability* represents the ability that a genotype can evolve to another phenotype to perturbations.

In [36] the relationship between the mutational robustness of a phenotype and the potential of a population to generate novel phenotypic variation is studied. It is found that phenotypic robustness promotes phenotypic variability in response to non-genetic perturbations, but not much in response to mutation. From the consideration of the genotype space using a variant model of Wagner's model [104] they observe how a large set of genotypes produce the same phenotype, this is called *neutral network*. The size of the neutral network is used as a proxy for robustness to mutations. In particular, the potential of different kinds of perturbations to generate phenotypic variation in populations subject to stabilising selection for many generations is analysed there [36]. Such population accumulate genetic variation that is not phenotypically visible. So, non-genetic perturbations, such as changes in the environment (e.g., changes in the initial condition, noise in the trajectory between initial state to the final state), produce new phenotypes in order to be adapted to new ambient conditions, so it is said that there is innovation. It is showed that innovation occurs more frequently in robust phenotypic populations under mutations. It is suggested that phenotypic robustness to mutations can play a positive role in phenotypic variability after non-genetic perturbations.

In Ciliberti et al. [27, 28] robustness is measured as the size of the neutral network. They have obtained two important results, namely, they observe that neutral networks evolve into more robust one (i.e. emergence of canalisation), and that long-term innovation in gene expression only emerges in the presence of the robustness induced by interconnected genotype networks.

Canalisation is also studied by Siegal and Bergman [94] using a model based on Wagner's model [104]. They show that, in the presence of an optimal phenotype, (with respect to which fitness is measured), robustness increases in time. Further to this model, Bergman and Siegal [14] have studied genetic assimilation. They observe how a genetic knock out uncovers hidden cryptic genetic variation, which becomes fixed within the population, even when the gene knock out is removed. They claim that this result implies that in this type of GRN, virtually every gene can act as an evolutionary capacitor (such as HSP90 [83]).

1.1.3 Evolutionary escape

Evolutionary escape is the process whereby a population under sudden changes in the selective pressures acting upon it try to evade extinction by evolving from previously well-adapted phenotypes to those that are favoured by the new selective pressure, the called *escape phenotypes.* This evolutionary process is driven by gene mutations. Examples of biological situations where this process is relevant include viruses evading anti-microbial therapy, emergence of drug resistance in cancer, parasites trying to infect a new host, or species attempting to invade a new ecological niche [53].

Earlier models of evolutionary escape have been formulated by Iwasa and co-workers [53, 54], their approach based on the assumption that n point mutations in some crucial parts of the genome are necessary for escape. They further assume that the genotype of the different mutants can be described by binary strings (with entries of $+1$ or -1) of length n, of which there are $2^n - 1$. It is assumed that, under the new selective pressure, most genotypes exhibit reduced proliferation ratios of sensitive mutants, $R < 1$; whereas some genotypes, the so-called *escape genotypes*, are such that $R > 1$. The corresponding evolutionary dynamics is modelled in terms of Galton-Watson multi-type branching process [57], where at each generation each individual of each type has a given (in general, mutant-dependent) probability of mutating, thus producing offspring belonging to a different type. The problem is to calculate the probability that an escape genotype is reached. The model proposed by Iwasa and co-workers has been analysed in more detail by Serra and co-workers [89, 90, 84]. These authors have thus considered the process of evolutionary escape as a random search on a genotype space modelled by a hypercube: Individuals would concentrate in a given genotype and they must reach a well-adapted genotype (the escape genotype) before the population undergoes extinction. An alternative escape mechanism have been proposed in [4] whereby escape is achieved by means of a growth-restricted (quiescent) phenotype that is insensitive to the selective pressure (e.g. a drug). This escape mechanism is relevant in cancer treatment of hypoxic tumours [3, 18, 19] and drug resistance in bacterial populations which exhibit persistence [8, 66].

1.2 Methodology

In this Section, we present a summary of tools and models used in this work, specifically, regarding complex network theory and population dynamics. We include a section on generating functions (subsection 1.2.0.1) and their properties which are used in the analysis of branching process (subsection 1.2.2.3). In Section 1.2.1 jointly with Section 7.2 in Appendix 7 we review the main properties and definitions that characterise a graph [75](degree distribution, clustering coefficient, percolation threshold, ...), usual graph models (Erdös Rényi [34],[35], Watts–Strogatz [109], Barabási–Albert [11], configuration model [16, 13]). We also review some of the literature on growing networks [74] and community detection algorithms [40], including particular properties associated to them. In Section 1.2.2, we summarise the basic properties of population genetic models and branching processes. We focus on the Wright-Fisher and Moran models and the Galton-Watson process which are extensively used in this work.

1.2.0.1 Generating functions

In mathematics, a *generating function* [110] is a formal power series in one indeterminate, whose coefficients encode information about a sequence of numbers an that is indexed by the natural numbers. Then, probability generating function, of a discrete random variable is a power series representation (the generating function) of the probability mass function of a random variable..

Suppose a discrete random variable X takes values in $0, 1, 2, ...$ and has probability function $p(x)$ $(P[X = x] = p(x))$. Then *pgf* (*probability generating function*) is

$$g_X(s) = \mathbb{E}(s^X) = \sum_{x=0}^{\infty} p(x)s^x.$$

Note: $g_X(0) = p(0)$ and $g_X(1) = 1$, pgf uniquely determines the distribution and vice-versa.

From above definition we can define the multivariate pgf. Suppose $X = (X_1, \ldots, X_n) \sim \{p_{i_1 i_2 \ldots i_n}\}_{i_1, i_2, \ldots, i_n \geq 0}$ is a finite vector of non-negative random variables, then the pgf g_X of X can be written as,

$$g_X(\vec{s}) = \mathbb{E}(s_1^{X_1} \cdot s_2^{X_2} \cdots s_n^{X_n}) = p_{i_1 i_2 \ldots i_n} s_1^{i_1} s_2^{i_2} \cdots s_n^{i_n},$$

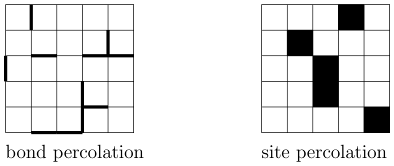
if $\vec{s} = (s_1, s_2, \ldots, s_n) \in [0, 1]^n$.

Theorem 1.2.1. *Let $Z = X_1 + X_2 + \cdots + X_n$, with X_i independent discrete random variables with pgfs $g_i(s), i = 1, \ldots, n$. Then,*

$$g_Z(s) = g_1(s)g_2(s) \cdots g_n(s).$$

In particular, if X_i are identically distributed with pgf $g(s)$, then,

$$g_z(s) = [g(s)]^n$$



bond percolation

site percolation

Figure 1.1: Different percolation models: Bond and site percolation.

Proof.

$$g_Z(s) = \mathbb{E}\left(s^Z\right) = \mathbb{E}\left(s^{X_1+X_2+\cdots+X_n}\right) = \mathbb{E}\left(s^{X_1}\cdot s^{X_2}\cdots s^{X_n}\right) \underbrace{=}_{X_i\text{independents}} \tag{1.3}$$

$$= \mathbb{E}\left(s^{X_1}\right)\cdot\mathbb{E}\left(s^{X_2}\right)\cdots\mathbb{E}\left(s^{X_n}\right) \underbrace{=}_{\text{definition}} g_1(s)\cdot g_2(s)\cdots g_n(s) \tag{1.4}$$

□

1.2.1 Network models, properties and tools

Networks or graphs are ubiquitous in many fields of active research in order to organise and represent the data as well as extracting information to represent it. In our case, we use a network representation for the genotype-phenotype space. In Appendix 7.2, we briefly define the main local and global properties of graphs defined in the field of graph theory.

1.2.1.1 Percolation

Percolation theory is a highly developed field in mathematics and statistical physics [97, 21, 17]. The best way to illustrate problems related to percolation theory applied to networks is the bond percolation model and site percolation, (see Figure 1.1). In the first model, given n isolated nodes on a regular lattice, we define the process that at each step introduces one or more edges with probability p, in this case we are interested in which is the critical value p_c to have a path between two extremes on the lattice. On the other hand, the site percolation model considers an empty lattice in which nodes are randomly filled with probability p. The percolation transition occurs when a giant connected component emerges (see Appendix Section 7.2), i.e. a connected component which contains a macroscopically large number of nodes, when the occupation probability reaches its critical value, p_c, (node percolation). In other words, it is the formation of long-range connectivity in random systems. Below the critical probability a giant connected component does not exist; while above it, there exists a giant component of the order of system size. As we can expect, the higher is p, the larger are the individual clusters. Instead of being a continuous change, we observe a phase transition.

There are some results regarding the percolation transition for different lattices and processes, where the critical parameter can be approximated accurately [24]. Another feature related to percolation is that giant connected component emerges faster in assortative networks, otherwise its emergence is delayed in dissassortative networks (see Appendix 7).

The same process could be done inversely, in this case we want to know when the system is broken or disconnected. This allows to define robustness as an inverse percolation transition. In this case, given a connected network (or a network with a giant connected component), we remove a fraction, f, of nodes (or edges) chosen at random. This parameter controls the proportion of nodes (edges) that can be removed from the network without the giant component being destroyed. In other words, how robust a network is against random failure by investigating the proportion of nodes or edges that we need to remove before long-scale connectivity is lost. As we have described before, there exists a critical threshold f_c: for $f < f_c$ we continue to have a giant component, but when $f > f_c$, the giant component breaks into disconnected components.

The critical fraction f_c of removed nodes for a network with an arbitrary degree distribution is $f_c = 1 - \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}$ (proof in Advanced Topics 8.C in [12]). It is easy to deduce $f_c = 1 - \frac{1}{\langle k \rangle}$ for a random network, with has a Poisson degree distribution. Then, as $\langle k \rangle \to 1$, it is easier to disconnect a random network (generated by the Erdös-Rényi model described later). By contrast, scale-free degree distribution networks, whose second moment $\langle k^2 \rangle$ diverges, are such that $f_c \to 1$. This result has the implication that scale-free networks are very resilient to random attacks.

A classical example of percolation processes are the epidemic processes by which diseases spread over networks of contact between population. In these processes percolation plays an important role.

1.2.1.2 Small world property

The small-world phenomenon is exhibited by those networks in which the average clustering coefficient is close to one and, simultaneously, the diameter of the network D, is small, i.e. $D \sim \log(n)$, where n is the number of nodes (see Section 7.2 in Appendix 7). In other words, the small-world effect describes those graphs whose diameter and average path length grow much more slowly than the number of nodes n, that is slower than $\log(n)$ which correspond to Erdös Rényi graphs (or random graphs).

1.2.1.3 Network models

We first summarise some of the basic models used in complex network theory.

Random networks Random graphs (Section 2.2 in [22]) are generated by fixing some of the parameters such as the number of nodes or/and edges, but the network is maximally random otherwise. The simplest model is obtained by fixing the number of nodes n and place a fixed number of edges m randomly. This model is referred to as the $G(n, m)$ model.

The best known model of random graphs is the *Erdős–Rényi model* [34, 35] denoted by $G(n, p)$, where n is the number of nodes and p, the probability of connecting each pair of nodes. This model allow a fuller characterisation than for $G(n, m)$ graphs. It is straightforward to obtain the mean number of edges, $\langle m \rangle = \binom{n}{2} p$ and the mean degree, $\langle k \rangle = (n-1)p$.

The degree distribution of graphs generated using this model follows a Poisson distribution (for n large, otherwise it follows a binomial distribution). Its clustering coefficient, c, tends to zero as n grows while mean degree stays constant ($c = \frac{\langle k \rangle}{n-1}$). This property differs from most of real-world networks which retain a high clustering coefficient with a large amount of nodes. Other departures from most real networks are that random graph models do not show correlations between degree and its second neighbourhood neither exist a community structure.

Small-world networks Another important type of network is the so-called small-world network.

Similarly to the Erdős–Rényi model, the *Watts–Strogatz model* [109] is the basic procedure to generate a small-world graph characterised by short average path lengths and high clustering. This model is specified by the number of nodes, n, the mean degree, k, and a parameter p ($0 \leq p \leq 1$), which represents the rewiring probability and controls the interpolation between the circle model and the random graph. We start from a regular ring lattice, with n nodes and each connected to k neighbours ($k/2$ on each side). The small-world graph is generated by rewiring each edge with probability p to a random chosen node. The degree distribution goes from a Dirac delta function ($p = 0$) to a Poisson distribution as $p \to 1$. In this model clustering coefficient varies in function of p by, $c = \frac{3(\langle k\rangle-2)}{4(\langle k\rangle-1)+8p\langle+4p^2k\rangle\langle k\rangle}$ [75].

Scale-free networks A usual feature of real networks is the appearance of power-law degree distributions, also called scale-free graphs. The first model which proposed a mechanism to obtain a power-law distribution is the *Barabási–Albert model* [11, 12], also known as a the Yule process, which uses the preferential attachment rule: given a number of nodes n, and an initial number of nodes n_0, the network begins with an initial connected network of n_0 nodes and with links which are chosen arbitrarily, as long as each node has at least one link. At each time step a new node is added to the network one at a time. Each new node is connected to existing nodes with a probability that is proportional to the number of links (k_i) that the existing nodes already have. Formally, the probability p_i that the new node is connected to node i is $p_i = \frac{k_i}{\sum_j k_j}$. This model produces a network with n nodes and m edges and $\langle k\rangle = 2m$. This model is limited to produce power-laws with exponent equal 3.

Average path length increases approximately logarithmically with the size of the network, $\left(D \sim \frac{\log(n)}{\log(\log(n))}\right)$ that implies a systematically shorter average path length than a random graph. Clustering coefficient of the Barabási-Albert model decays slower than expected for a random network, indicating that the obtained network is locally more clustered $\left(c \sim \frac{ln(n)^2}{n}\right)$ [59]. On the other hand, correlations between the degrees of connected nodes develop spontaneously because of the way the network is generated, but as n grows this correlation decays [12].

Configuration model A widely used algorithm to generate random graphs is the *configuration model* [72]. This model generates a maximally random graph given a degree distribution, that is: given a sequence of degrees $(k_1, k_2, \ldots, k_n)$ with length n. We generate a graph with n nodes and each node n_i has k_i stubs (half edges), by picking a pair of stubs uniformly at random and connect them.

1.2.1.4 Correlations

Most of real networks exhibit correlations in their connectivity patterns. Correlations between two nodes can be measured by the average degree (see Appendix 7) of nearest neighbours as a function of the node degree, $\bar{k}_{nn}(k) = \sum_{k'} k' P(k' \mid k)$ [22], where $P(k' \mid k)$ is the conditioned probability that a vertex of degree k is connected to a vertex of degree k'. We say that a graph is *assortative* if the correlation between the average of nearest neighbours and their degree is positive, i.e more connected nodes are preferably connected to high-degree nodes. Otherwise, if connections are preferably formed between nodes with very different degree usually are connected, implying a negative correlation between degree and number of nearest neighbours, we have a *disassortative* network. Random graph models, such as the Erdős–Rényi model and the configuration model do not exhibit degree correlations by definition. In growing networks degree correlations depend on the attachment probability A_k (sublinear, linear, superlinear). The linear case corresponds to the Barabási-Albert, which degree correlations decay as n grows. The clustering coefficient (see Appendix 7 Section 1.2.1) is another form of correlation. In this case, it is associated to correlations between three nodes [22].

Growing networks usually develop correlations between the degrees of connected nodes as the network grows and the size distributions of the *in component* and the *out component* in directed growing networks. The presence of correlations might have important consequences in dynamical processes taking place in the topology defined by the network (e.g. epidemic spreading, epidemic threshold (percolation), ...).

1.2.1.5 Growing networks

Growing networks or non-equilibrium graphs are networks that evolve, adding nodes and links as time progresses. The organisational development of growing networks has been studied in papers like [62] where an attachment rate is prescribed. Real networks do not belong to the equilibrium, i.e. static, class of networks. Their non-equilibrium character endows them with a number of interesting properties. For example, the percolation transition in growing networks is of a quite different nature in non-equilibrium networks [23] (percolation properties are well studied in [33, 73, 68, 69]). Using a simple growing model, [23] has shown that a phase transition exists at which a giant component forms whose size scales linearly with system size. In this respect, their networks resemble traditional random graphs but they differ from random graphs in many other ways. For example, the mean component size is different both quantitatively and also qualitatively, having no divergence at the phase transition. The position of the phase transition is different as well, and the transition itself appears to be infinite order rather than second order. There is, therefore, a number of features, both local and global, by which the non-equilibrium graph can be distinguished from a static one. Another usual behaviour of growing networks, specially with preferential attachment, is that in long times a single node captures a macroscopic fraction of links.

Recent approaches to the study of growing networks are based on the use of rate equations methods in order to model their evolution of growing networks, instead of probabilistic approaches or generating function techniques [60]. The rate equation approach also has the advantage that it can be adapted to other evolving graphs systems, including networks not only with addition, such as with the addition and deletion of nodes and links, and also with link rewiring [62].

For example, in [62], the following model is proposed: a growing network in which nodes are added one at time, and a link is established with a pre-existing node according to an attachment probability A_k, which depends only on the degree of the target node. This model produces a directed tree graph topology. They study the evolution of the degree distribution according different kernels of attachment: linear, sublinear and superlinear [63]. Their rate equation to model the evolution of the degree distribution is,

$$\frac{dn_k}{dt} = A^{-1}\left[A_{k-1}n_{k-1} - A_k n_k\right] + \delta_{k1} \tag{1.5}$$

where, $n_k(t)$ represents the number of nodes of degree k at time t, A_k is the probability to attach to a node of degree k and $A(t) = \sum_{j\geq 1} A_j n_j(t)$.

In growing networks the time-dependent degree distribution is determined by the rate equation approach, with different distributions arising in the growing network model depending on the asymptotic behaviour of the attachment probability as a function of node degree.

Such models also investigate the joint age-degree distribution, and it has been found that old nodes are typically more highly connected [61]. Approaches like those of [62, 32] have analysed how the degree distribution depends on the attachment rate, A_k and behaviour depends on whether A_k grows slower than linearly with k, linear or superlinear. The probability of attachment takes an important role in the generation of degree correlations, that changes the emergence of the giant component.

1.2.1.6 Communities and modularity

Communities, i.e. subsets of nodes tightly inter-connected, well above the levels of a randomly chosen set of nodes, and how to detect them is a field studied in graph theory. There are a broad range of algorithms for finding communities, whose accuracy depends on the kind of graph and the available information [40, 65]. For example, there are algorithms where *a priori* knowledge of the number of communities to detect is needed [40]. We can obtain different levels of success for all this range of algorithms applied to the same graph. A usual measure used in many community detection algorithms is *modularity*. It measures the strength of a division of a network into modules (also called groups, clusters or communities). Modularity is defined as the fraction of the edges that fall within the given groups minus the expected such fraction if edges were distributed at random: It takes positive values if the number of edges within groups exceeds the number expected on the basis of chance. Given a partition of the network into modules, modularity reflects the concentration of edges within modules compared with random distribution of links between all nodes regardless of modules. In other words, networks with high modularity have dense connections between the nodes within modules but sparse connections between nodes in different modules.

Traditional methods for community detection are: graph partitioning (Kernighan-Lin algorithm, spectral bisection method), hierarchical clustering (agglomerative algorithms and divisive algorithms; sometimes stopping conditions are given by modularity optimisation), partitional clustering (k-means clustering), spectral clustering (introduced by Donath and Hoffmann) and divisive algorithm (Girvan–Newman algorithm). An extensive review of all of these methods, as well as newer one, is given in [40].

1.2.2 Population dynamics

Modern population dynamics is the field originated by [111],[38] and [45], where short-term and long-term composition changes of populations are studied. We briefly present some models used in this thesis. Also, a very useful tool as generating functions is presented with some interesting properties and how it applies to branching processes.

1.2.2.1 Population genetic models:

Models of inheritance, mutation, and selection of genetic material in populations of individuals. Classically, these models assume a constant number of individuals related to each other through common ancestry (Wright-Fisher model, [111],[38]). Although very different from the branching processes some of these models can be approximated by branching processes (e.g., when an expanding subpopulation of mutants arises within the large population). Such a situation arises when some genetic diseases are studied.

Wright-Fisher and Moran model: *Wright-Fisher model* [111, 38], is a well-known model in population genetics where individuals within the population are picked up at random for proliferation with uniform probability. Of the rest of the individuals of the population we picked another one randomly and remove it from the population, so that the total population keeps constant.

On the other hand, in each generation of the *Moran model* [70], one individual is chosen at random to give 2 offspring and one individual is chosen to die (all other individuals survive to the next generation). In contrast to Wright-Fisher model, Moran model has overlapping generations. This model is also known as a birth-and-death model.

1.2.2.2 Multi-type Galton–Watson process:

A generalisation of the single-type *Galton–Watson process* [6, 57]. It evolves in discrete time measured by non-negative integers. Each individual belongs to one of a finite number of types. The initial population is assumed to consist of a single individual of one particular type. Processes started by individuals of different types are generally different. At each generation, the ancestors are replaced by a random number of progeny of various types. The distribution of progeny counts depends on the type of parent. Each of the first-generation progeny becomes an ancestor of an independent subprocess, distributed identically as the whole process (modulo ancestor's type). In the multi-type process, asymptotic behaviour depends on the matrix of average progeny count. Rows of this matrix correspond to the parent types and columns correspond to the progeny types. The largest positive eigenvalue of this matrix (Theorem Perron- Frobenius [78, 41]) is the Malthusian parameter of the process, provided the process is supercritical (the Perron–Frobenius eigenvalue larger than 1) and positive regular. This latter means that parent of any given type will have among its (not necessarily direct) descendants, individuals of all possible types, with non-zero probability. Galton-Watson process (or multi-type Galton-Watson) can also be generalised to a continuous branching process.

1.2.2.3 Branching processes

A *branching process* is a class of process that not need to be Markovian, but all the processes studied here are Markov processes. A *Markov process* is defined as a stochastic process with the following property:

$$P(y_n, t_n|y_1, t_1; \ldots ; y_{n-1}, t_{n-1}) = P(y_n, t_n|y_{n-1}, t_{n-1}), \forall n. \tag{1.6}$$

That is, the conditional probability density at t_n, given the value y_{n-1} at time t_{n-1}, is uniquely determined and is not affected by any knowledge of the y_s values at earlier times.

Markov process studied here models a population in which each individual in generation n produces some random number of individuals in next generation $n+1$, according, in the simplest case, to a fixed probability distribution that does not vary from individual to individual. These processes are used to model bacterial reproduction, epidemic spreading, cell mutations, etc., [57].

The oldest, simplest and best known branching process is the *Galton-Watson process* [57], [6]. It can be described as follows: A single ancestor particle lives for exactly one unit of time, and at the moment of death it produces a random number of progeny according to a given probability distribution. Each of the first-generation progeny behaves, independently of each other, as the initial particle did. It lives for a unit of time and produces a random number of progeny. Each of the second-generation progeny behaves in the identical way, and so forth.

This process can be mathematically described using a discrete-time index, identical to the number of successive generation and defining some generating functions. Generating functions are a useful tool for handling distributions of such random sums is the probability generating function (pgf) of a distribution. Methods employing pgf manipulations instead of directly dealing with random variables are called analytic. Probability generating functions are the basic analytic tool employed to deal with non-negative random variables and finite and enumerable sequences (vectors) of such variables.

The Galton-Watson process produces an equation that can be solved, providing us the number of individuals for each generation and many interesting data.

We can consider two ways to define the self-recurrence in the branching process. One is based on decomposing the process into a union of subprocesses initiated by the direct descendants of the ancestor. It can be called the “backward” approach, in an analogy to the backward Chapman–Kolmogorov equations of Markov processes. The other way is the “forward” approach. It consists of freezing the process at time t, recording the states of all individuals at that time, and predicting their future paths.

Let Z_n be a random variable used as counter of individuals in each generation (where generation 0 is composed of the single initial particle) and X be the number of offspring for an individual with $P[X = k] = p_k, k \in \mathbb{N}$, $\mathbb{E}[X] = R$ and $Var[X] = \sigma^2 < \infty$. The process can be represented as a union of the subprocesses initiated by the first-generation offspring of the ancestor particle. Let X_j be the random variable for the number of offspring for the jth individual, such that, $Z_{n+1} = \sum_{j=1}^{Z_n} X_j$. We construct the backward equation of a Galton-Watson branching process. Let the pgf of X be given by $g(s) = \sum_{k=0}^{\infty} p_k s^k$ and let $g_n(s) = \sum_{k=0}^{\infty} P[Z_n = k]s^k$ be the pgf of Z_n for $n = 1, 2, 3, \ldots$. Then, using Theorem 1.2.1 of probability generating function, follows backward equation:

$$g_{n+1}(s) = g_n(g_1(s)) \tag{1.7}$$

In order to obtain forward equation (only in Galton-Watson branching process), we only need to fix $Z_0 = 1$, it means $P[Z_0 = 1] = 1$. Then, $g_0(s) = s$, this yields the following,

$$g_n(s) = g^n(s) = g_1\{\ldots g_1(g_1(s))\ldots\}, \quad g_{n+1}(s) = g_n(g_1(s)). \tag{1.8}$$

Backward-forward equation shows that we only need to know what happens between t to $t + 1$ in order to understand all the process.

Moreover, branching processes are not restricted to one-component processes, but all properties applies for processes with r components. In the above mathematical description we only need to think s as a vector $\vec{s} = (s_1, \ldots, s_r)$, and we have a branching processes with multi-type components. One example is the multi-type Galton-Watson process.

1.2.2.4 Continuous-time branching process:

We have defined previously the discrete branching process. However it is possible to define a continuous-time (age-dependent) branching process [57] with exponential life time distributions. This process also has the Markov property and is closely related to the Galton–Watson process. Although the exponential distribution to model lifetimes of cells admitting lifetimes which are arbitrarily close to 0, whereas it is known that life cycles of organisms and cells have lower bounds of duration. The advantage of using the exponential distribution is that it leads, in many cases, to computable expressions. The latter allow us to deduce properties which then can be conjectured for more general models.

Similar to the discrete process, the age-dependent process can be described as follows. A single ancestor particle is present at $t = 0$. It lives for time τ, which is exponentially distributed with parameter λ. At the moment of death, the particle produces a random number of progeny according to a probability distribution with pgf $f(s)$. Each of the first-generation progeny behaves, independently of each other, in the same way as the initial particle. It lives for an exponentially distributed time and produces a random number of progeny. Progeny of each of the subsequent generations behave in the same way. If we denote the particle count at time t by $Z(t)$, we obtain a stochastic process $Z(t), t \geq 0$. The probability generating function $F(s,t)$ of $Z(t)$ satisfies an ordinary differential equation which is easiest to derive based on the Markov nature of the process.

Consider the process at a given time t. Any of the particles existing at this time, whatever its age is, has a remaining lifetime distributed exponentially with parameter λ. This follows from the lack of memory of the exponential distribution. Therefore, each of the particles starts, independently, a subprocess identically distributed with the entire process. Consequently, at any time $t + \Delta t$, the number of particles in the process is equal to the sum of the number of particles in all iid subprocesses, indexed by i, started by particles existing at time t, such that, $Z(t+\Delta t) = \sum_{i=0}^{Z(\Delta t)} Z^{(i)}(t)$.

So, according to the pgf theorem 1.2.1, we have the following pgf identity:

$$F(s, t+\Delta t) = F[F(s,t), \Delta t], \quad F(s,0) = s$$

.

Assuming a Δt small and then by letting $t \to 0$, leads to the following differential equation:

$$\frac{dF(s,t)}{dt} = -\lambda(F(s,t) - f(F(s,t))), \text{ with initial condition} \quad F(s,0) = s. \tag{1.9}$$

For more details see Chapter 4 of [57].

1.2.2.5 Extinction and criticality

Processes can be classified into subcritical, critical or supercritical processes in terms of their long-time asymptotic behaviour. Once process is classified we can determine its relation with an extinction process. We define $R = \mathbb{E}(X)$, the mean progeny count of a particle, Z_n random variable which counts the number of individuals in each generation and $g(s)$ is the pgf that satisfies $\mathbb{E}(X) = g'(1)$, as we have defined in 1.8, differentiating it and evaluating in $s = 1$, we obtain

$$\mathbb{E}(Z_t) = g_n'(1) = R^n.$$

Then, in the expected value sense, the process:

- grows geometrically if $R > 1$, called supercritical;
- stays constant if $R = 1$, called critical; and
- decays geometrically if $R < 1$, called subcritical.

However, this expected value sense is not in many cases a good description of the behaviour. Even in the supercritical case the probability of extinction can be very high. In terms of extinction process we have theorem 1.2.2.

Theorem 1.2.2. *The extinction probability of the Z_n process is the smallest non-negative root q of the equation,*

$$g(q) = q.$$

It is equal to 1 if $R \leq 1$, and it is less than 1 if $R > 1$. *Proof* See Section 1.5.3 in [57].

1.3 Aims

The aim of this PhD thesis is to develop a multi-scale model of biological evolution which accounts for the mapping between genotype and phenotype as determined by a model of the gene regulatory network. In order to achieve this aim, we formulate a simple model in Chapter 2 of genotype-phenotype map inspired in the model proposed in [103] and studied in [104] by Wagner.

We first proceed to generate a genotype-phenotype space (network) under such multi-scale dynamics. We then analyse several population dynamics on this genotype-phenotype graph. The objective is to study the effects of the genotype-phenotype structure on these dynamics through the complex topology of genotype-phenotype network.

More specifically in Chapter 3, we characterise the geometrical and topological properties of the genotype-phenotype space obtained from the multi-scale model which assumes a selective pressure acting at the level of phenotypes. This topological characterisation determines robustness and evolvability driven by genetic mutations, and their relation to evolutionary phenomena such as canalisation and convergence. We analyse the role of cryptic genetic variation on evolutionary processes and study how rewiring of the gene regulatory networks affect robustness and evolvability. Also, we study how robust are genotype-phenotype networks obtained from the multi-scale model against attacks.

In Chapter 4 we start our explanation of how population dynamics are affected by the complex topological features of the genotype-phenotype network. In particular, we extend the theory of evolutionary escape, a well-known evolutionary mechanism responsible for resistance, by analysing the effects on the probability of escape and the escape rate of considering that the evolutionary dynamics occurs on a genotype-phenotype network rather than on a regular hypercube. We present a comparative analysis between genotype-phenotype networks obtained from the multi-scale model and networks constructed assuming that the genotype space is a regular hypercube. We compare the effects on the probability of escape and the escape rate associated to the evolutionary dynamics between both classes of graphs.

We further our study of evolutionary escape on complex genotype-phenotype networks by introducing a continuous-time branching process in order to model evolutionary escape and survival as a process characterised by two different time-scales: A fast-decaying, initial regime in which escape actually occurs, followed by a slow, quasi-steady state regime, in which cells which have succeeded to reach the well-adapted escape phenotype, strive for survival. Our aim in Chapter 5 is to analyse the influence of topological properties associated to robustness and evolvability on the probability of escape and on the probability of survival upon escape. We analyse the role played by topological properties of escape genotype, such as, degree, clustering coefficient and a local weighted clustering coefficient, in determining escape and eventual survival.

1.4 Outline

This thesis is structured in seven Chapters. In Chapter 1 we introduce genotype-phenotype maps, giving some ideas about previous models of them and presenting the concept of evolutionary escape (Section 1.1). Then, in Section 1.2 we present a review of the main methodology used in this thesis and applied in the followings Chapters 3, 4 and 5. Chapter 2 contains a complete description of the proposed multi-scale model of biological evolution and used subsequently.

After that, the main contributions are presented in Chapters 3, 4 and 5. At the end of each of these Chapters we describe and discuss the obtained results. Finally, we summarise results and conclusions in Chapter 6 and give some ideas for future work. We provide an Appendix in Chapter 7.

Chapter 2

The model

The aim of this Chapter is to propose a multi-scale model of evolutionary dynamics inspired in the model proposed by Wagner in [103], which accounts for the mapping between genotype and phenotype as determined by a model of the gene regulatory network. The model is split in three scales: microscale, mesoscale and macroscale. Each one is associated to a particular process. The microscale models the intracellular dynamics of interaction between genes, that is the gene regulatory network (GRN), and provides a genotype-phenotype map introducing a selective pressure at the phenotype level. The mesoscale describes the dynamics of the population of cells, which we assume to be described by a multi-type Wright-Fisher model with mutation [15]. Finally, the macroscale describes the evolutionary dynamics of the genotype-phenotype space. We propose a similar, but not identical, representation of this space to the proposed by Ciliberti et al. [28, 27].

2.1 Model formulation

We consider a population of cells whose dynamics is described by a multi-scale model [3] composed of three mutually-coupled levels: the microscale, where the intracellular dynamics is modelled in terms of a gene regulatory network (GRN), the mesoscale, i.e. the population dynamics of the population where birth and death rates are assigned according to the phenotype of each cell, and the macroscale, consisting of a model of the evolutionary dynamics of the genotype-phenotype space (to be defined later).

Each of these levels have associated different characteristic time scales which become longer as we move up the hierarchy of the model: from the microscale, where the characteristic time scales correspond to the ones associated to intracellular processes (ranging from seconds to hours), to the mesoscale, characterised by the time scales of the order of the life-time of a cell (ranging from a few days to months or even years), and to the macroscale, where the time scales are those characteristic of evolutionary processes and are measured in generations [3] (see Figure 2.1).

We consider a scenario where selective pressures act on phenotypes. We therefore formulate a model with separation of time scales where we allow individuals to complete the developmental plan encoded in their genotype (to be defined in Section 2.1.1), determine the corresponding phenotype, and then decide whether it is fit to survive and thrive or not. A similar selection procedure has been used, for example, by Siegal & Bergman [94, 14]

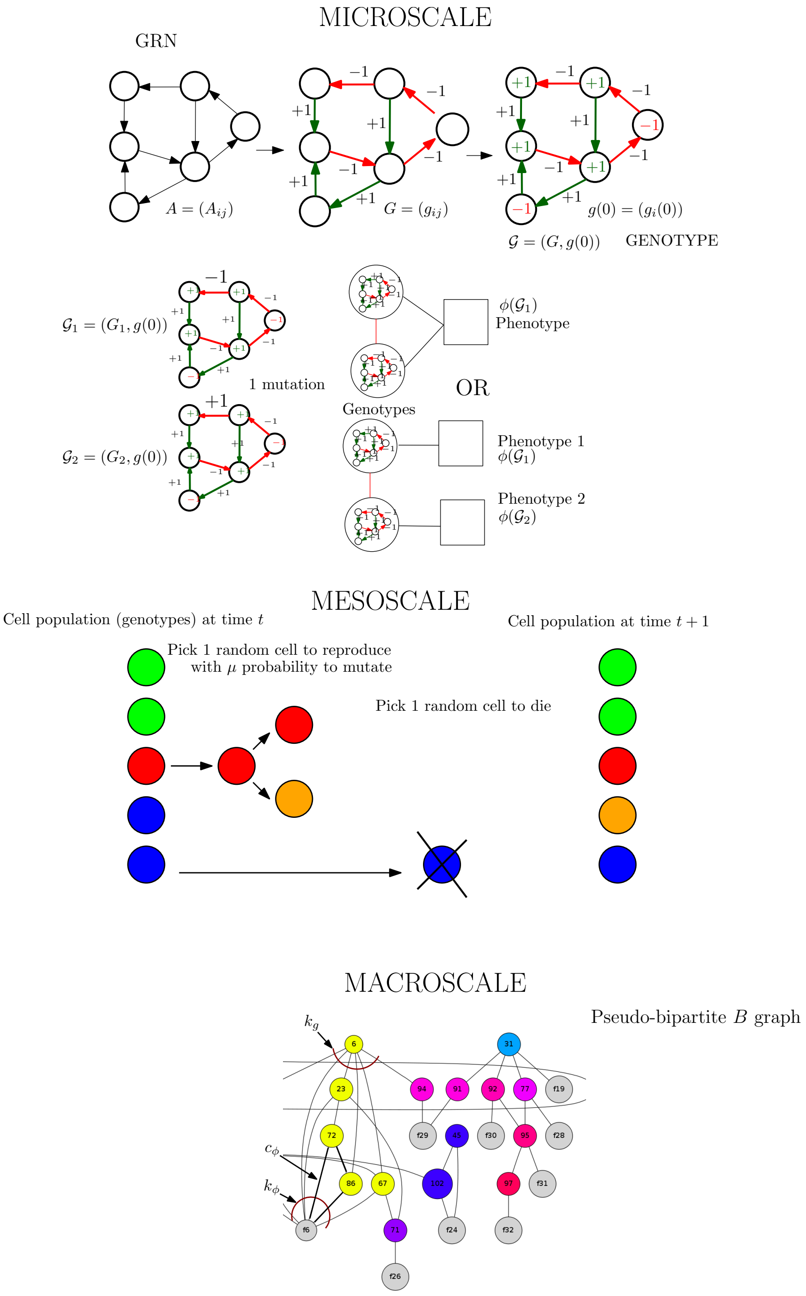


Figure 2.1: Description of multi-scale model. Time measures in microscale: intracellular process then ranging from seconds to hours, mesoscale: life-time of a cell ranging from few days, months or even years, macroscale: generations.

2.1.1 Microscale: Intracellular dynamics of the GRN

Our model considers a population of cells characterised by a pair $\mathcal{G} = (G, g(0))$. G is a matrix accounting for the interaction between genes (i.e. how a gene product affects the (in)activation of all other genes). The vector $g(0)$ corresponds to the initial pattern of gene expression, which can be interpreted as heritable genetic information: the components of $g(0) = (g_i(0))$ are the states (active or inactive) of each gene when new cells are born. This pair is referred to as the *genotype* in the remaining of this manuscript.

The aim of the microscale dynamics is to provide our model with a representation of the genotype-phenotype map, i.e., a correspondence between $(G, g(0))$ and the phenotype, ϕ. Our model is closely related to the model used by Wagner to study plasticity [104] and the model used by Bergman and Siegal to study canalisation [94, 14]. It is also based on the idea, proposed by Stuart Kauffman, that GRNs are dynamical systems and phenotypes or differentiated states correspond to the stable attractors of these dynamical systems [56]. Within this framework, we represent GRNs as graphs, where genes are the nodes and links represent a regulatory interaction between the genes corresponding to the nodes at either end of the link. This graph can be represented by a matrix A, the so-called adjacency matrix, where dim(A)=$N_G \times N_G$, with N_G equal to the number of genes. A is defined so that its entries, a_{ij}, are such that $a_{ij} = 1$ if there exists a link between nodes (genes) i and j, and $a_{ij} = 0$, otherwise.

Each gene, labelled by $i = 1, \dots, N_G$, is endowed with a state variable $g_i(t)$, which, at time t, can take two values: $g_i(t) = 1$ if the gene is being expressed and its protein product being synthesised, and $g_i(t) = -1$, otherwise. The matrix $G = (g_{ij})$, is a weighted version of A: $g_{ij} = a_{ij}$ if gene i activates gene j, and $g_{ij} = -a_{ij}$ if gene i inhibits gene j. GRNs are directed graphs and, therefore, both A and G need not be symmetric.

The phenotype is defined as the steady-state of the dynamical system defined by the following set of rules:

1. At $t = 0$, that is, at the time of birth of each cell, we fix $g(t = 0) = g_0$.

2. At each time step, and for each gene in the GRN, we determine the value of the variable $I_i(t)$, defined by:

$$I_i = \sum_{\langle j \rangle_i^{in}} g_{ji} g_j, \quad \text{where } \langle j \rangle_i^{in} \text{ is the set of } \textit{in-neighbours} \text{ of } i. \tag{2.1}$$

3. We determine the value of the state of each gene at step $t + 1$ according to:

$$\begin{cases} g_i(t+1) = 1, & \text{if } I_i(t) \geq 0 \\ g_i(t+1) = -1, & \text{if } I_i(t) < 0 \end{cases}$$

4. Steps 2 and 3 are repeated until $t = T \gg 1$. The phenotype corresponding to the genotype $\mathcal{G} = (G, g_0)$, $\phi(\mathcal{G})$, is defined by $\phi(\mathcal{G}) = g(T)$

2.1.1.1 Selective pressure: viability conditions.

We formulate conditions to discard genotypes which give rise to *unviable* phenotypes. Our viability conditions are related by those imposed by Bergman & Siegal [94, 14]. These authors consider that a gene regulatory network is in steady state if an average measure analogous to a variance is smaller than a (small) threshold value. If the GRN has not reached a steady state after a certain number of iterations, it is discarded as non-viable. The average involved in this steady state criterion is taken over a period of time of length v, which implies that solutions with periods longer than v are non-viable.. Here, we discarding oscillatory solutions if their period is longer than a threshold length, l_v. In the case of these oscillatory solutions, our definition of the phenotype is slightly different: $\phi_i(\mathcal{G}) = 0$ if g_i oscillates, and $\phi_i(\mathcal{G}) = g_i(T)$, otherwise. For example, if $\phi(\mathcal{G})$ oscillates like $(1,-1,1,1) \to (1,1,1,1) \to (1,-1,1,1)$, then $\phi(\mathcal{G})$ is defined as $(1,0,1,1)$.

From our definition of viability conditions, we can see that the shorter we choose l_v, the more GRNs are going to be disregarded as unviable. Therefore smaller l_v, the more severe the selection pressure (see Appendix 7 Fig 7.2 to see plotted genotype-phenotype networks with $l_v = 0$).

2.1.1.2 GRN topology

Finally, we need to model the topology of the GRNs. We use two different models in order to capture two topological properties. A number of protein-interaction and transcription networks have been found to exhibit the small-world phenomenon [100, 112, 5] and, in view of this, we will use the Strogatz-Watts model (Section 1.2.1) to generate the graphs underlying GRNs with the small-world property [109]. This model works as follows. Consider a regular lattice with N_G vertices and k edges per vertex such that $E = kN_G/2$. By rewiring each edge at random with probability p, Watts & Strogatz showed that, between the limits of a regular graph ($p = 0$) and a random graph ($p = 1$), there exists an intermediate regime between these two extremes where there is a transition whereby the rewired graph exhibits the small-world effect, i.e. they simultaneously show high clustering and short-path length [76].

A second topological property whose effects we are interested in exploring is the scale-free topology shown to be exhibited by GRNs [7], which we generate using the preferential attachment model [10]. Greenbury et al. [43] have shown that the scale-free topology leads to increased robustness with respect to GRN with random, Erdös-Renyi topology. Here we will examine the small-world phenomenon and the scale-free topology affect the evolvability properties of our system.

Recall that is enough to study GRN connected networks (see Appendix 7 Section 7.1).

2.1.2 Mesoscale: Population dynamics

The mesoscale describes the dynamics of the population of cells, which we assume to be described by a multi-type Wright-Fisher model with mutation [15]. In our model, cell types correspond to genotypes where n_i is the number of cells with genotype $\mathcal{G}_i$. The Wright-Fisher model is an urn model where the total number of cells $N_c = \sum_i n_i$ is kept constant. The population dynamics is as follows:

1. A cell is chosen at random for proliferation. All cells are equally likely to be chosen. Therefore, the probability of a cell with genotype $\mathcal{G}_i$ to be picked up is equal to n_i/N.

2. With probability μ, the chosen cell is subjected to mutation. In our model (see [104]), a mutation corresponds to changing the sign of one and only one randomly chosen, non-zero entry of the matrix G: i.e. $G \to G'$ where $g'_{ij} = -g_{ij}$, where the ij-edge is randomly chosen with probability $1/E$, with $E = \sum_{i,j} a_{ij}$ is the number of edges (interaction between genes) of the GRN, and $g'_{lk} = g_{lk}$ for all $i \neq l$ and $k \neq j$. We check, by running the microscale dynamics, whether $\phi(\mathcal{G}_i)$ is equal to $\phi(\mathcal{G}')$, where

$\mathcal{G}' = (G', g_0)$. We have three possibilities, namely, (i) $\phi(\mathcal{G}_i) = \phi(\mathcal{G}')$, (ii) $\phi(\mathcal{G}_i) \neq \phi(\mathcal{G}')$ and $\phi(\mathcal{G}')$ is viable, and, last, (iii) $\phi(\mathcal{G}_i) \neq \phi(\mathcal{G}')$ and $\phi(\mathcal{G}')$ is non-viable.

3. If there is no mutation (with probability $1-\mu$), $n_i \to n_i+1$. If there is a mutation which results in cases (i) or (ii), $n_i \to n_i - 1$ and $n_{i'} \to n_{i'} + 2$, where $n_{i'}$ is the population of cells with genotype $\mathcal{G}'$. In either case, another cell chosen at random is eliminated, so that the total cell population stays constant.

4. If there is a mutation which results in a non-viable genotype (case (iii)), steps 1 and 2 are repeated until a viable genotype is found

This description of the cell population dynamics shows that our viability conditions act as a selective pressure on the population: Those genotypes more prone to yield oscillations of period longer than l_v are negatively selected for.

2.1.3 Macroscale: Evolutionary dynamics of the genotype-phenotype map

The macroscale describes the evolutionary dynamics of the genotype-phenotype space. Our representation of this space is similar, but not identical, to the metagraph proposed by Ciliberti et al. [28, 27], who analysed this space using a graph which extends the concept of neutral network [105].

The model of Ciliberti et al. [28, 27] can be summarised as follows. Consider that all the cells in our population share the same inherited developmental plan, g_0. Then, genotypes $\mathcal{G} = (G, g_0)$ are characterised by the corresponding matrix G alone. In the previous section, we have defined a mutation as a sign change in one of the non-zero entries of G. To properly characterise this in mathematical terms, let us define the distance between two matrices G and G', $\text{dist}(G, G')$, as:

$$\text{dist}(G, G') = \sum_{i,j} \frac{|g_{ij} - g'_{ij}|}{2}, \tag{2.2}$$

i.e. $\text{dist}(G, G')$ is the number of entries in G and G' with differing signs.

For example, let G_1 and G_2, two genotypes,

$$G_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad G_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

then $\text{dist}(G_1, G_2) = 2$. G_1 and G_2 in the previous example differ in the sign of two entries, i.e. they are two mutations apart. According to the definition of neutral network given in Chapter 1.1 these two genotypes would not be connected.

Furthermore, consider the set of viable genotypes, $\mathcal{G} = (G, g_0)$. Based on this set, we can define a graph $\Gamma = (V_G, E_G)$, where V_G is the set of vertices and E_G is the set of edges, in the following way.

1. V_G is equal to the set of viable genotypes

2. E_G is defined as follows: for all pair $\mathcal{G}_i, \mathcal{G}_j \in V$ an edge exists linking both nodes if and only if $\text{dist}(G_i, G_j) = 1$

Ciliberti et al. [28, 27] characterise phenotypic robustness and evolvability in terms of certain topological properties of this graph, namely, its modularity structure is related to robustness: if genotypes with the same phenotype reside within modules (i.e. tightly interconnected subsets of V_G [75]), the phenotype is said to be robust as it is highly likely that gene mutations lead to genotypes producing the same phenotype.

We propose here an extension of this graph by considering a *pseudo-bipartite* graph, $B = (V_G, V_\phi, E_G, E_\phi)$, where we have two types of nodes and two types of edges:

1. V_G is equal to the set of viable genotypes
2. V_ϕ is equal to the set of viable phenotypes
3. E_G, the set of genotype-genotype links, is defined as follows: for all pair $\mathcal{G}_i, \mathcal{G}_j \in V_G$ an edge exists linking both nodes if and only if $\text{dist}(G_i, G_j) = 1$
4. E_ϕ, the set of genotype-phenotype edges, is given by the phenotype-genotype map: An edge exists between ϕ_i and $\mathcal{G}_j$ if and only if $\phi_i = \phi(\mathcal{G}_j)$

Note that this definition departs from that of a bipartite graph in that we allow edges between nodes of the same type (i.e. genotype-genotype edges).

2.1.4 Evolutionary dynamics of the genotype-phenotype map

Our aim is to analyse the evolutionary dynamics and the emergent properties of the genotype-phenotype map. We propose the following multi-scale evolutionary dynamics model.

Initial Condition At $t = 0$ we fix the initial condition for our evolutionary dynamics as follows.

1. An inheritable developmental programme, g_0, is randomly generated.
2. N_c number of GRNs are generated. This is a two step process:
   - N_c replicas of an adjacency matrix are generated at random using the Watts-Strogatz model with rewiring probability p .
   - For each of these adjacency matrices, each non-zero entry is given a positive (negative) weight with probability p_+ $(1 - p_+)$.
3. For each of the N_c genotypes $\mathcal{G}_i = (G_i, g_0)$, $i = 1, \dots, N_c$, we determine the corresponding phenotype $\phi_i = \phi(\mathcal{G}_i)$ by means of the microscale dynamics described in Section 2.1.1. If non-viable genotypes are encountered, the corresponding GRN is discarded and another one is generated. This process is repeated until we have N_c viable genotypes and produces the set of genotypes present in the population at time $t = 0$, $S_G(t = 0) = \{\mathcal{G}_i,\, i = 1, \dots, N_c\}$, and the set of phenotypes present in the population at $t = 0$, $S_\phi(t = 0) = \{\phi(\mathcal{G}_i),\, i = 1, \dots, N_c\}$. Note that repeated genotypes and phenotypes are included only once in their respective sets.
4. The initial condition for the genotype-phenotype graph, $B_0 = (V_G(t = 0), V_\phi(t = 0), E_G(t = 0), E_\phi(t = 0))$, is fixed as follows.
   - The set of genotype nodes, $V_G(t = 0) = S_G(t = 0)$

- The set of phenotype nodes, $V_\phi(t=0) = S_\phi(t=0)$. Note that $\text{card}(V_\phi) \leq \text{card}(V_G)$
- The set of genotype-genotype links, $E_G(t=0)$, is defined as follows: for all pair $\mathcal{G}_i, \mathcal{G}_j \in V_G(t=0)$ an edge exists linking both nodes if and only if $\text{dist}(G_i, G_j) = 1$
- The set of genotype-phenotype edges, $E_\phi(t=0)$, is determined by the phenotype-genotype map: An edge exists between $\phi_i \in V_\phi(t=0)$ and $\mathcal{G}_j \in V_G(t=0)$ if and only if $\phi_i = \phi(\mathcal{G}_j)$

Dynamics For $t > 0$, the system is updated according to the following dynamics.

1. An iteration of the mesoscale dynamics according to the model described in Section 2.1.2 is carried out. This iteration produces a population of cells with the corresponding set of different genotypes present in the population at iteration $t = t+1$, $S_G(t+1)$ and with the phenotype set $S_\phi(t+1)$.

2. The growth dynamics of the genotype-phenotype graph is defined as follows:

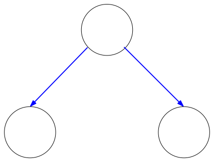
   - If $S_G(t+1) \cap V_G(t) \neq S_G(t+1)$, i.e. the population dynamics has generated new genotypes, then $V_G(t+1) = V_G(t) \cup (S_G(t+1) - S_G(t+1) \cap V_G(t))$, i.e. we update the set of genotype nodes by adding the new ones arising from the population (mesoscale) dynamics.
   - Similarly, for the set of phenotype nodes, if $S_\phi(t+1) \cap V_\phi(t) \neq S_\phi(t+1)$ then $V_\phi(t+1) = V_\phi(t) \cup (S_\phi(t+1) - S_\phi(t+1) \cap V_\phi(t))$, i.e. the set of phenotype nodes is updated by addition of the new ones arising from the population dynamics.
   - The set of genotype-genotype links, $E_G(t+1)$, is updated by adding to $E_G(t)$ the following set of edges: for all new genotype, $\mathcal{G} \in S_G(t+1) - S_G(t+1) \cap V_G(t)$, we add new links between $\mathcal{G}$ and all $\mathcal{G}_j \in V_G(t)$ such that $\text{dist}(G, G_j) = 1$, i.e. we add links from the new genotypes to all the genotypes in the graph which are one mutation apart.
   - The set of genotype-phenotype links, $E_\phi(t+1)$, is updated by applying the genotype-phenotype map (microscale dynamics) to the new genotypes, i.e. new edges are added between phenotype node ϕ_i and genotype node $\mathcal{G}_j$ if and only if $\phi_i = \phi(\mathcal{G}_j)$.

2.1.4.1 Example: Microscale to macroscale

Genes interaction:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$A = (a_{ij})$, $a_{ij} \in \{0, 1\}$, $N = 3$.



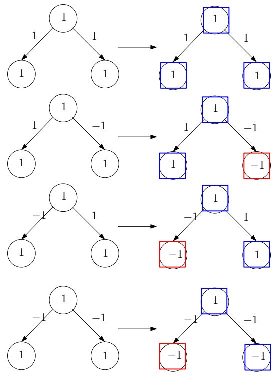
Genotype representation:

$$G = (g_{ij}), g_{ij} \in \{-1, 0, 1\},$$
$$g_{ij} \in \{-1, 1\} \iff a_{ij} = 1.$$

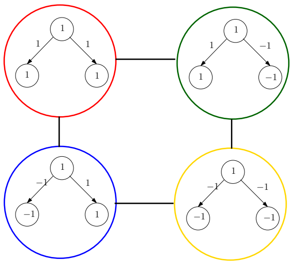
$$G = \begin{pmatrix} 0 & \pm 1 & \pm 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Initial data:

$$g(t=0) = g_0 = (1,1,1),\ g_{0_i} \in \{-1, 1\}$$
$$\Downarrow$$
Network evolves under our dynamics
$$\Downarrow$$
Final state.



Phenotype/Genotype network: $\mathcal{G} = (G, g_0)$, then the corresponding phenotype $\phi(G(g_0))$ is represented by a circle around the genotype. Each different colour represents a different phenotype.



2.1.5 Remarks of genotype-phenotype network

One of the first observations to do is that genotype-phenotype network obtained is a *pseudo-bipartite graph*. It means, we have two classes of nodes (genotypes and phenotypes) and two classes of edges (genotype-genotype and genotype-phenotype). As a difference between bipartite networks we have connections between one of the same class of nodes. This small difference give us a more advantageous way to work.

Another point is the possibility to define the clustering coefficient, that in the topological characterisation of robustness proposed by Ciliberti et al. [28, 27] in the network constructed according to the model described in Section 1.1.2.1 is not possible. Since, one can define the average clustering coefficient of a network by adding up all the nodal clustering coefficients and dividing by the number of nodes. Then, if we attempt to apply this definition to the network constructed according to the model described in Section 1.1.2.1, we immediately run into a problem: The clustering coefficient is identically zero since it is impossible to have triangles in this graph. Recall that two nodes on the genotype network are connected if only if there is only one mutation between them. If a node a is connected with node b and c, distance between a and b and a and c is 1:

$$d(a,b) = 1 \wedge d(a,c) = 1 \implies d(b,c) = 2 \tag{2.3}$$

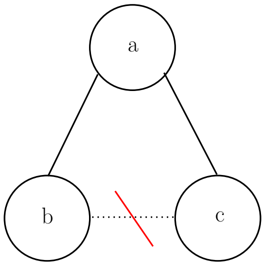


Figure 2.2: Impossible edge. Not triangles.

Eq. (2.3) implies that there cannot exist any edge between b and c, which means there are no triangles in the genotype graph. Whereas, the genotype-phenotype network obtained from the model described above, allows to apply with sense the clustering coefficient, that is going to be an useful parameter to study the structural features of genotype-phenotype networks. It can be easily verified that triangles where one of the vertex corresponds to a phenotype and the other two to genotypes exist (see Figure 2.3).

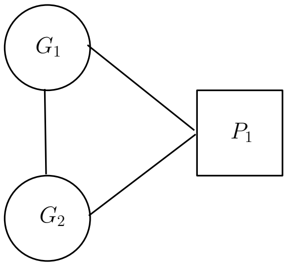


Figure 2.3: Triangle between genotypes and a phenotype.

Chapter 3

The topology of robustness and evolvability in evolutionary systems with genotype-phenotype map

In this Chapter we formulate a topological definition of the concepts of robustness and evolvability. We start our investigation by formulating a multi-scale model of the evolutionary dynamics of a population of cells. Our cells are characterised by a genotype-phenotype map: their chances of survival under selective pressure are determined by their phenotypes, whereas the latter are determined their genotypes. According to our multi-scale dynamics, the population dynamics generates the evolution of a genotype-phenotype network. Our representation of the genotype-phenotype network is similar to previously described ones, but has a novel element, namely, our network contains two types of nodes: genotype and phenotype nodes. This network representation allows us to characterise robustness and evolvability in terms of its topological properties: phenotypic robustness by means of the clustering coefficient of the phenotype nodes, and evolvability as the emergence of giant connected component which allows navigation between phenotypes. This topological definition of evolvability allows to characterise the so-called *robustness of evolvability*, which is defined in terms of the robustness against attack (i.e. edge removal) of the giant connected component. An investigation of the factors that affect the robustness of evolvability shows that phenotypic robustness and cryptic genetic variation are key to the integrity of the ability to innovate. These results fit within the framework of a number of models which point out that robustness favours rather than hindering evolvability. We further show that the corresponding phenotype network, defined as the one-component projection of the whole genotype-phenotype network, exhibits the small-world phenomenon, which implies that in this type of evolutionary system the rate of adaptability is enhanced.

3.1 Network dynamics: degree distributions and parameters

The result of the multi-scale model described in Chapter 2 is the evolutionary dynamics on the genotype-phenotype space (i.e. our pseudo-bipartite network). This space can be characterised in terms of a number of topological attributes typical from complex network theory [75]. One such metric, which conveys very useful information, is the degree distribution. Here, we consider two different degree distributions: The genotype degree distribution, $P(k_g)$, corresponding to the number of genotype neighbours of genotype nodes, and the phenotype

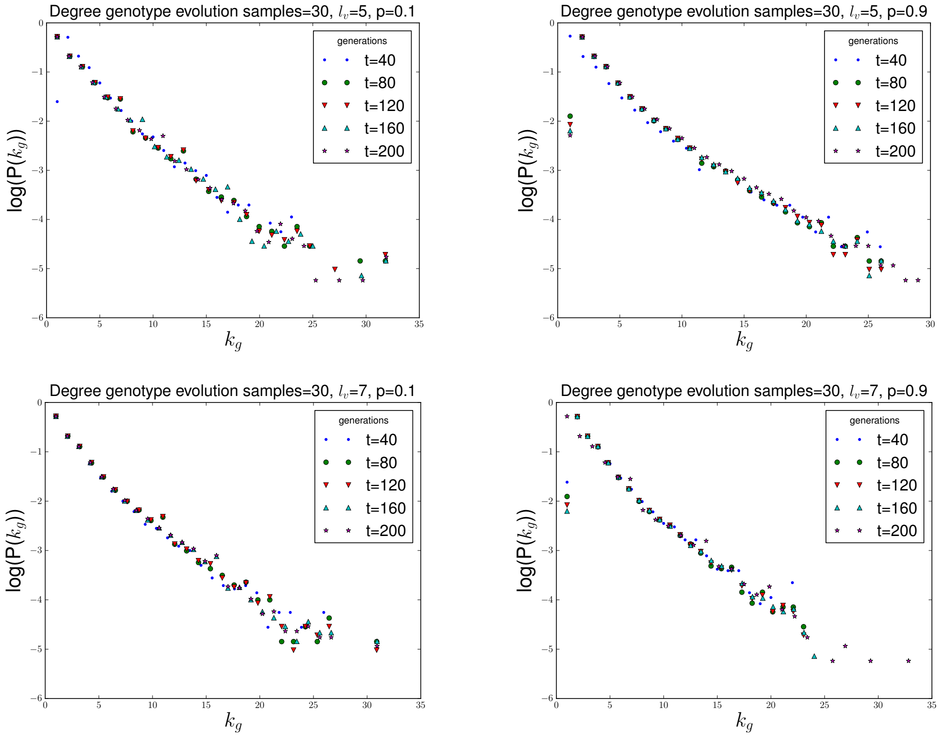


Figure 3.1: Time evolution of the degree distribution for the genotype nodes in the bipartite network. The degree of the genotype nodes is exponentially distributed. We observe that the genotype degree distribution quickly settles into its exponential distribution steady-state.

degree distribution, $P(k_\phi)$, corresponding to the number of genotype nodes connected to a phenotype node (i.e. k_ϕ is an approximate measure of the size of the basin of attraction of phenotype ϕ subjected to the selective pressure). Results for $P(k_g)$ and $P(k_\phi)$ are shown in Fig. 3.1 and Fig. 3.2, respectively.

For GRNs whose topology has been generated according to the Strogatz-Watts model (SW) (defined in Section 1.2.1), we observe that, whereas the genotype degree distribution quickly settles into an (steady-state) exponential distribution, the phenotype degree distribution evolves towards a power-law distribution (see Figs. 3.1 and 3.2). The same qualitative behaviour is observed for scale-free GRNs, although some significant differences are observed with respect to the previous case. Fig. 3.3, where we compare the genotype and phenotype degree distributions corresponding to both GRNs topologies, shows that whereas the genotype degree distribution for both scale-free (SF) generated by the Barabási-Albert model and Strogatz-Watts GRNs do not significantly differ, the phenotype degree distribution corresponding to the scale-free GRN exhibits a fatter tail than its Strogatz-Watts counterpart. The significance of this behaviour will become apparent when analysing the robustness properties of evolvability to be discussed in Sections 3.3.1 & 3.3.2.

Finally, we have considered the evolution of our multi-scale evolutionary model under two distinct regimes, namely, $\mu N > 1$ and $\mu N < 1$. These two regimes corresponding to fundamentally different dynamical regimes [105, 106]. Populations with $\mu N > 1$ are very likely to be polymorphic at any given generation, i.e. more than one phenotype coexist within the population, whereas populations with $\mu N < 1$ tend to be monomorphic for most

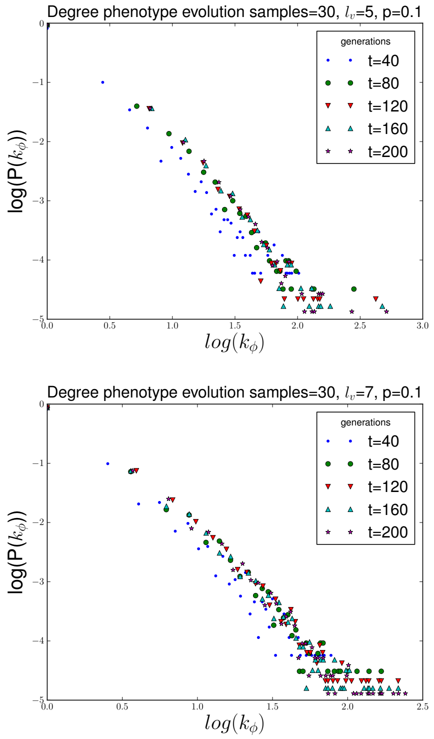


Figure 3.2: Time evolution of the phenotype degree distribution in the bipartite network. The degree of the phenotype nodes is distributed according to a power law. We observe that the phenotype degree distribution quickly settles into its power law distribution steady-state.

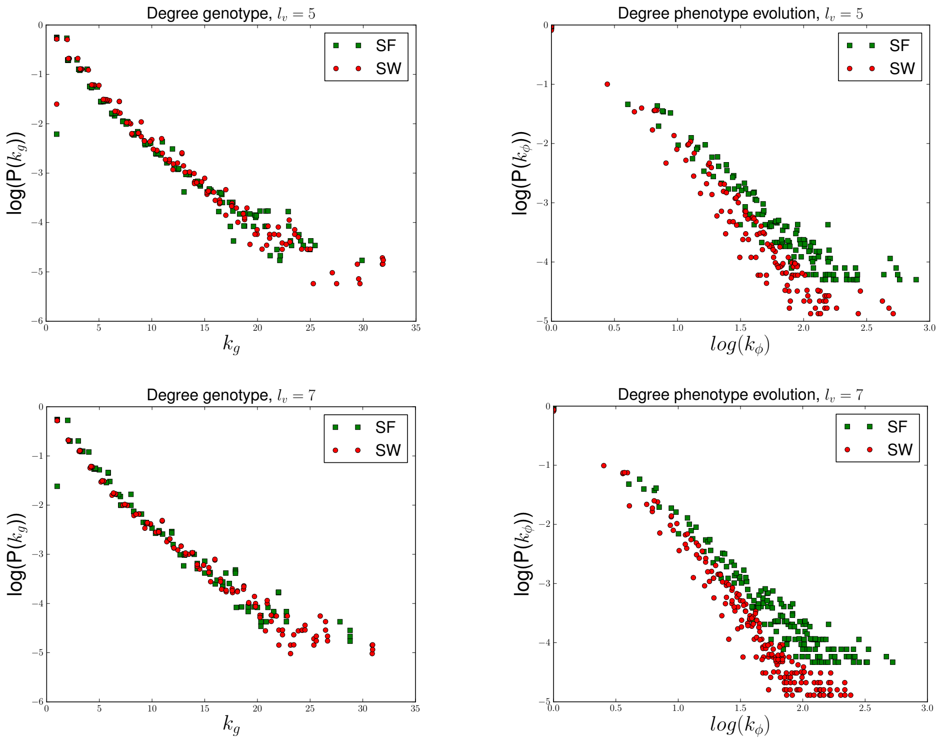


Figure 3.3: These plots show the comparison between the steady-state genotype and phenotype degree distribution for scale-free GRNs (SF) and Strogatz-Watts GRNs (SW). We observe that whereas the corresponding genotype degree distributions, $P(k_g)$, do not differ in a significant way, the phenotype degree distribution, $P(k_\phi)$, for scale-free GRNs exhibits a fatter tail than its Strogatz-Watts counterpart.

of their evolution. Consequently, whilst the latter will evolve to accumulate over time in regions of the genotype-phenotype map where strongly robust genotypes exist, the former will perform a random walk which will sample the phenotype-genotype space uniformly. Given these differences, it is important to compare the behaviour of our system in these two regimes. Regarding the genotype degree distribution, $P(k_g)$, and the phenotype degree distribution, $P(k_\phi)$, we observe that the exhibit fatter tails when $\mu N < 1$ than in the case $\mu N > 1$, as illustrated in the simulations shown in Fig. 3.4. The consequences regarding robustness and evolvability of this fact will be explored in detail in Sections 3.3.1 & 3.3.2.

3.2 Topological characterisation of phenotypic robustness

In this section, we provide a definition of a metric of phenotypic robustness based on the topological properties of the genotype-phenotype graph.

Our working definition of phenotypic robustness corresponds to the likelihood of a viable individual to retain its phenotype when gene mutations occur [105]. Our aim in this section is to propose a metric for phenotypic robustness that allows us to quantitatively analyse its evolutionary dynamics.

According to this definition, the phenotype $\phi(\mathcal{G})$ of an individual with genotype $\mathcal{G}$ is

Table 3.1: Parameter values.

Parameter	Description	Typical value
N_G	Length of the genome	20
l_v	Maximum period of viable oscillations	5 or 7
p	Rewiring probability in the Watts-Strogatz model	0.1 or 0.9
E	Number of links in the GRN	40
p_+	Probability of positive feed-back link	0.5
N_c	Number of cells in the population	50
μ	Mutation probability	0.3

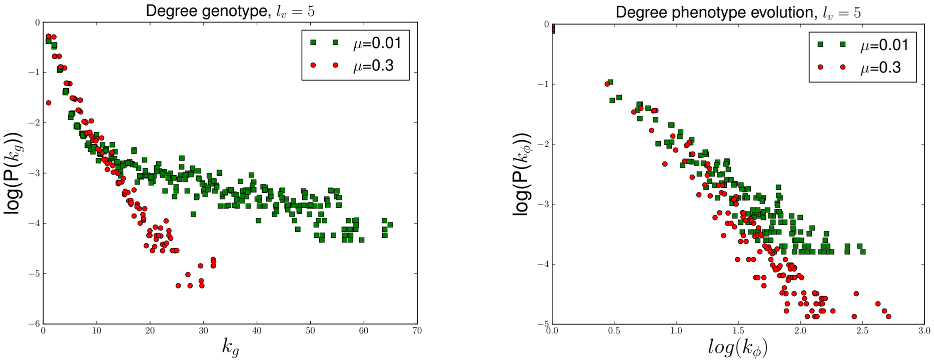


Figure 3.4: This figure shows simulation results comparing the steady-state genotype and phenotype degree distributions for different values of the mutation rate, μ. We observe that whereas the corresponding both distributions, $P(k_g)$ and $P(k_\phi)$, exhibits a fatter tail for $\mu N < 1$ than for $\mu N > 1$. We have taken $N = 50$ and $\mu = 0.3$ (red dots, $\mu N > 1$) and $\mu = 0.01$ (blue dots, $\mu N < 1$). These results correspond to genotype-phenotype maps generated using Strogatz-Watts GRNs with $p = 0.1$.

Table 3.2: Description of parameters.

Parameter	Description
t	time or iterations
$A = (a_{ij})$	Adjacency matrix of the interaction gene graph without to indicate if an interaction is positive or negative, ie, $a_{ij} \in \{0, 1\}$
$\mathcal{G} = (G, g(0))$	$G = (g_{ij})$: matrix of interaction between genes ($g_{ij} = \pm 1$), $g(0) = (g_i(0))$: vector of the initial pattern of gene expression $((g_i(0)) = \pm 1)$
ϕ	Phenotype and $\phi(\mathcal{G})$: phenotype corresponding to genotype $\mathcal{G}$
B	Pseudo-bipartite graph with: V_G: genotype nodes, V_ϕ: phenotype nodes, E_G: genotype-genotype links, E_ϕ: genotype-phenotype links. $B = (b_{ij})$ adjacency matrix of B
k_g	Genotype degree in the bipartite network
k_ϕ	Phenotype degree in the bipartite network
$P(k_g)$, $P(k_\phi)$	Genotype and phenotype degree distribution in the in the bipartite network, respectively
c_ϕ	Clustering coefficient of phenotype ϕ
s_{gcc}	Size of the giant component
$\pi_\phi(s)$	Distribution of the size of the connected components
$D_\phi(t)$	Diameter of the phenotype network at time t

deemed to be robust if a random gene mutation, transforming the genotype $\mathcal{G} \to \mathcal{G}'$ where $\text{dist}(\mathcal{G}, \mathcal{G}') = 1$, does not affect the phenotype, i.e. if $\phi(\mathcal{G}) = \phi(\mathcal{G}')$. This definition has a topological equivalence in terms the properties of the pseudo-bipartite network defined in Section 2.1.3. Since $\text{dist}(\mathcal{G}, \mathcal{G}') = 1$ these two genotypes correspond to nodes linked by a genotype-genotype edge. Moreover, if $\phi(\mathcal{G}) = \phi(\mathcal{G}')$ this means that both nodes $\mathcal{G}$ and $\mathcal{G}'$ are linked to the same phenotype node ϕ by genotype-phenotype edges. In other words, the genotype nodes $\mathcal{G}$ and $\mathcal{G}'$ and the phenotype node ϕ form a triangle within the pseudo-bipartite graph. Thus the number of such triangles within the neutral network of a given phenotype (i.e. the basin of attraction of ϕ) is a direct measure of phenotypic robustness. Phenotypic robustness can thus be quantified by means of the clustering coefficient of phenotype node ϕ, c_ϕ:

$$c_\phi = \frac{2T_\phi}{k_\phi(k_\phi - 1)} \tag{3.1}$$

where T_ϕ is the number of triangles that have ϕ as one of their vertices and k_ϕ is the degree of ϕ, i.e. the number of genotype nodes to which ϕ is connected.

Note that k_ϕ, i.e. roughly speaking, the size of the basin of attraction of ϕ, is not an appropriate measure of robustness as it does not take into account how tightly interconnected are the genotypes linked to ϕ. In this respect, our characterisation of robustness is reminiscent of the one given by Ciliberti et al. [28, 27] where robustness is characterised by the community structure within their genotype graph Γ (see Section 2.1.3). However, we found that, due to known difficulties with the algorithms currently used to determine community structure in graphs (different algorithms produce different communities, miss-assignation of nodes to clusters, etc.[40]), quantifying robustness in terms of community structure resulted problematic,

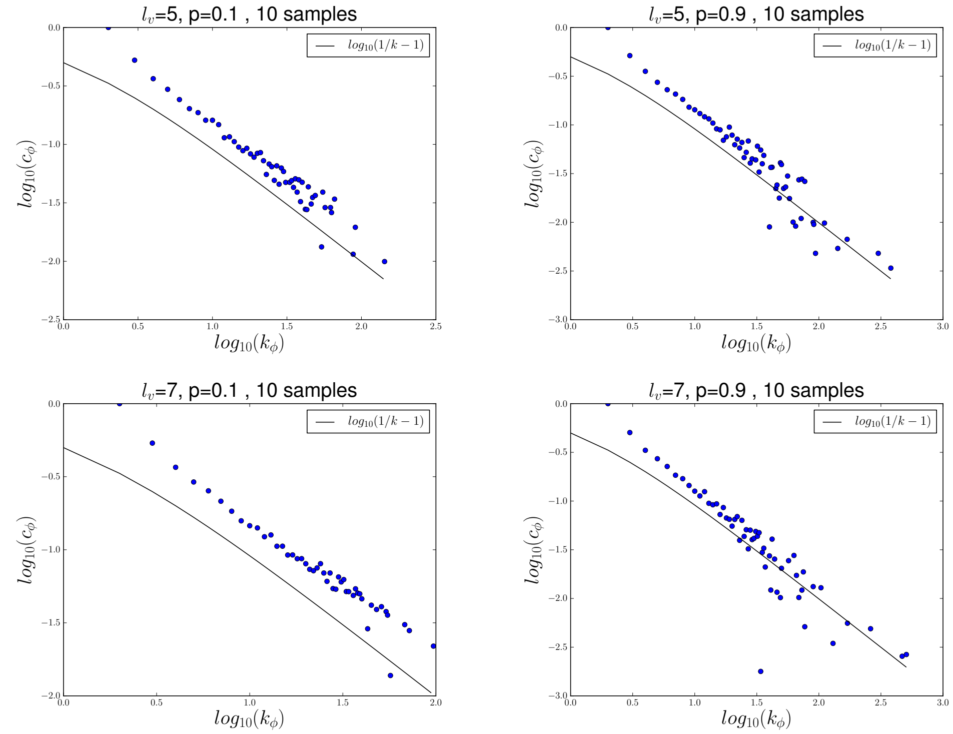


Figure 3.5: Blue dots represent simulation results for the (steady-state) phenotype clustering coefficient as a function of the degree, k_ϕ. We compare with $c(k_\phi) = (k_\phi - 1)^{-1}$ (solid black line), which according to [91, 92] establish the border between high and low clustering regimes in uncorrelated networks.

hence our introduction of this alternative description in terms of the pseudo-bipartite graph.

Fig. 3.5 shows simulation results regarding the steady-state behaviour of the relation between the clustering coefficient of phenotype nodes, c_ϕ, i.e. phenotypic robustness, as a function of their degree, k_ϕ, the size of the pool of viable genotypes with phenotype ϕ. These results show that there is an inverse correlation between c_ϕ and k_ϕ: Low-degree phenotypes exhibit much higher clustering coefficient than high-degree phenotypes.

We further observe that by tuning the rewiring probability of the GRN, p, the overall robustness can be controlled. Fig. 3.5 shows that for $p = 0.9$, regardless of the value of the of l_v, $c_\phi(k_\phi)$ is very well close to $c_\phi(k_\phi) \simeq (k_\phi - 1)^{-1}$, whereas for $p = 0.1$, the corresponding $c_\phi(k_\phi)$ curve is such that $c_\phi(k_\phi) > (k_\phi - 1)^{-1}$. This implies that the robustness of the phenotypes generated with GRNs with low values of p are intrinsically more robust. The reason for this result is that the parameter p in SW model measures how far the topology of the network is from a regular lattice and how close is to a tree. Low values of p imply that the network closely resembles a regular lattice. This means that low-p networks are, generally speaking, more heavily clustered than networks with large p values, which are more tree-like. More clustered GRNs imply that the presence of feed-back loops is more likely which, in turn, implies that oscillatory phenotypes are more common. Therefore, it is less probable to to obtain viable phenotypes for SW-GRNs with large p. However, when one is discovered, it is very likely that it is more robust.

3.3 Topological characterisation of evolvability

Evolvability is defined as the ability of an evolutionary system to innovate by generating new and better adapted behaviour [105]. Recent results have shed some light into the apparent conflict between robustness, i.e. resilience to change, and evolvability [28, 27, 106, 106, 107, 108]. These results suggest that, under very general conditions the tension is only apparent and the mechanisms that favour robustness also enforce the emergence of evolvability.

Regarding innovation, we follow the work of Ciliberti et al. [28, 27] where evolvability is related to the ability of the system to, under the effect of gene mutations, move from one neutral network (i.e. the subnetwork of genotypes with one particular phenotype) to another. This view, i.e. that evolvability is equivalent to the navigability of the network of phenotypes, has a direct mathematical translation within our model.

In this Section, we first define the phenotype network as the (weighted) one-component projection of the pseudo-bipartite genotype-phenotype graph. We then go on to characterise the emergence of a giant connected component within the phenotype network, an attribute that we use to characterise evolvability in topological terms. Finally, we conclude this section by showing that the phenotype network exhibits the small-world network property.

3.3.1 Phenotype network

In this and the following sections, we consider the properties of the phenotype network, which is defined as a one-mode projection of our phenotype-genotype network, which it is obtained in the following way. We consider that between a pair of phenotypes, ϕ_i and ϕ_j, there is an edge if there exists a path $\phi_i - \mathcal{G}_i - \mathcal{G}_j - \phi_j$ where $\phi(\mathcal{G}_i) = \phi_i$ and $\phi(\mathcal{G}_j) = \phi_j$, and $\mathcal{G}_i$ and $\mathcal{G}_j$ are connected within the pseudo-bipartite network. In other words, two phenotype nodes are connected in the one-component projection if there exists at least one length 3 path between them in pseudo-bipartite network. In our model such length 3 paths represent the minimum (two mutation) evolutionary path to reach a phenotype form another. Since there any number of different length 3 paths between any pair of phenotypes, we define the phenotype network as a weighted network where the weight of the edge between two phenotypes is the number of such paths that connect them within the full genotype-phenotype network.

Mathematically, the phenotype network is constructed as follows. If $B = (b_{ij})$ is the adjacency matrix of the pseudo-bipartite graph, then in order to calculate the number of length 3 paths between phenotypes, we use a well-known property of the adjacency matrix, namely that the entries of the matrix $B^3 = (\beta_{ij})$ are the number of length 3 paths between nodes i and j. β_{ij} is zero if no path of length 3 exists between nodes i and j. By considering only the entries corresponding to paths between phenotype nodes, we obtain the weighted adjacency matrix corresponding to our phenotype network.

3.3.2 Evolution of evolvability: Emergence of giant connected component

The onset of percolation is heralded by the formation of the so-called giant connected component. In percolation theory [24], the giant component is a connected subset of nodes that contains a macroscopic fraction of the entire set of vertices of the graph, i.e $s_{gcc} = \gamma N_v$, where s_{gcc} is the size of the giant component, N_v is the number of vertices of the graph and $\gamma \lesssim 1$.

In order to address the question of whether the evolutionary dynamics of our multi-scale system gives rise to evolvability, i.e. to a phenotype (network) space where a macroscopically large number of all viable phenotypes are mutually connected, we study the time evolution

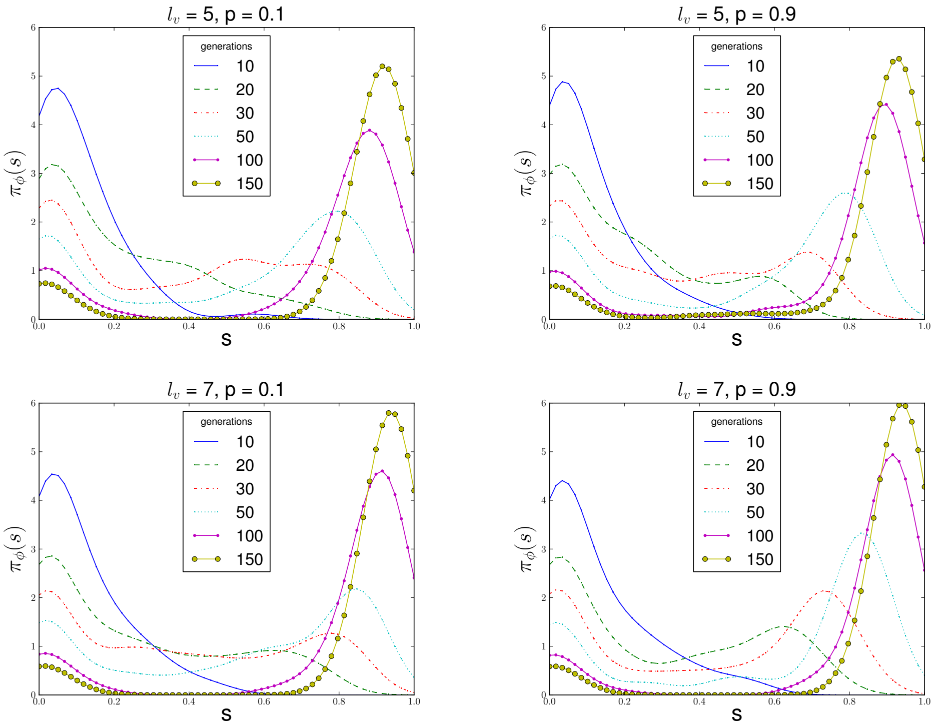


Figure 3.6: Time evolution of the distribution of cluster size in the phenotype network.

of the distribution of the size of the connected components, $\pi_\phi(s)$, in the phenotype network. The results are shown in Fig. 3.6.

We observe that, regardless the parameter values, as time progresses the formation of a giant connected component emerges. The formation of this connected component is signalled by the appearance of a peak for large size clusters in the cluster size distribution. Whereas the emergence of the giant connected component is unaffected by the GRN parameter, p, or the strength of the selective pressure, l_v, its dynamic is indeed affected by the rewiring probability: the larger the value of p, the faster the formation of the giant connected component (see Fig. 3.7).

3.3.3 The phenotype network is a small world

Complex networks often exhibit the so-called small-world phenomenon [109]. Small-world networks are such that most nodes can be reached form every other node in a small number of jumps compared to the size of the network. Mathematically, this property is defined by requiring that the diameter of the network grows proportionally to the logarithm of the number of nodes [75]. We are interested in investigating whether the evolutionary process described in Section 2.1 yields an evolving phenotype network with the small-world property, i.e. whether

$$D_\phi(t) \sim \log N_\phi(t) \tag{3.2}$$

where $D_\phi(t)$ is the diameter of the phenotype network at time t and $N_\phi(t)$ is the number of nodes at time t.

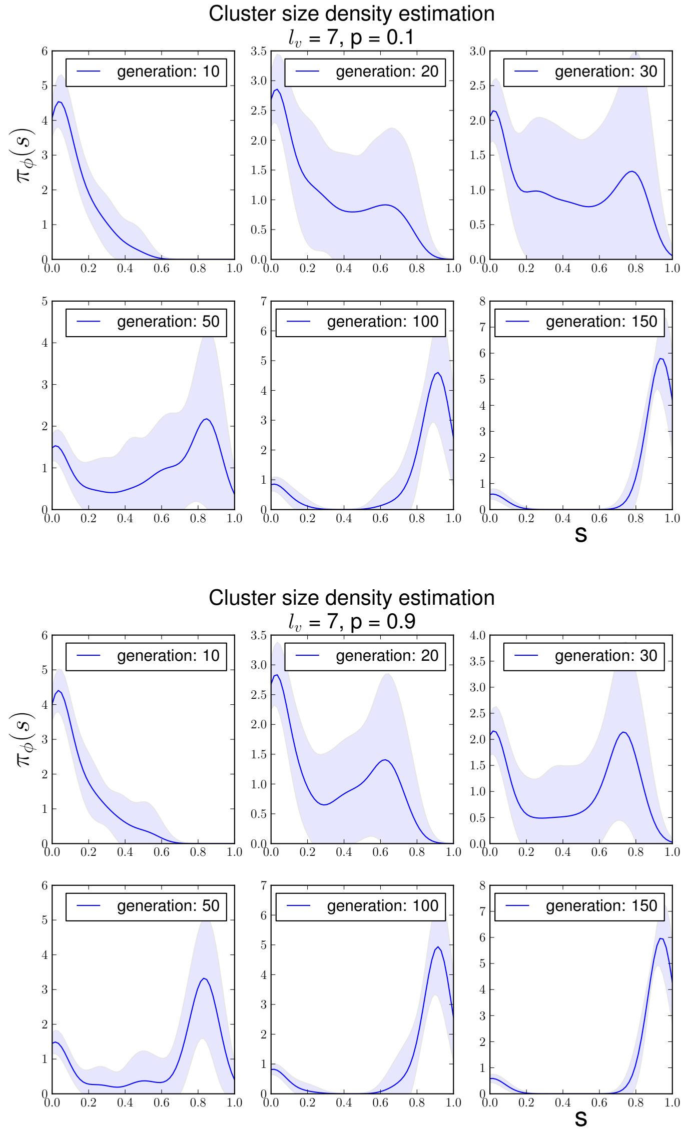


Figure 3.7: Dynamics of the formation of giant connected component as shown by the emergence of a peak in the distribution of cluster size in the phenotype network. The shadowed region is the 1-σ confidence region using a Gaussian kernel density estimation (see Appendix 7 Section 7.4).

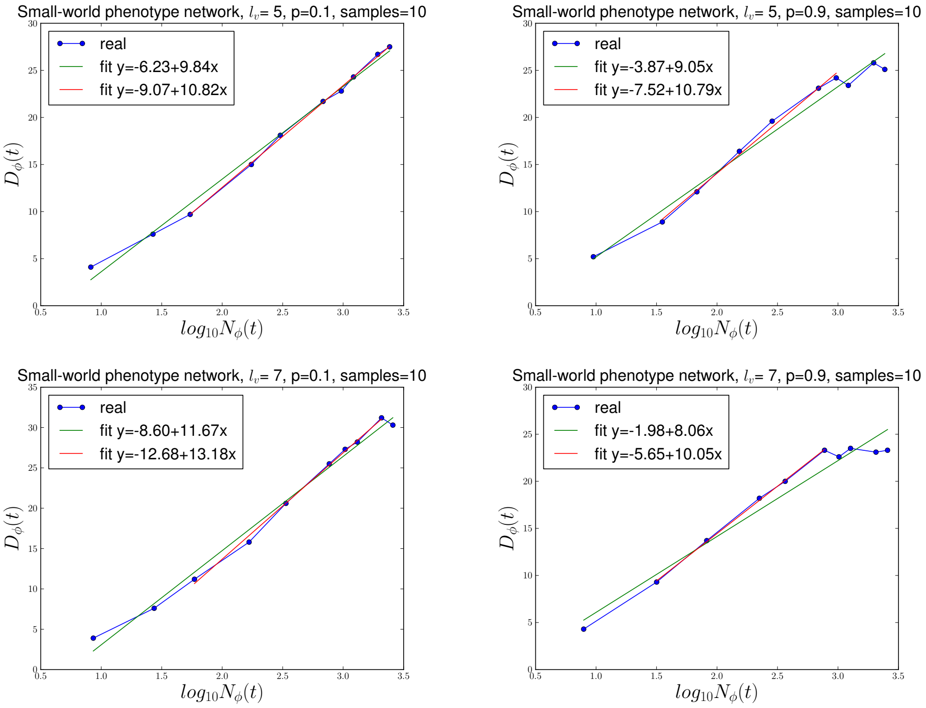


Figure 3.8: Diameter of the growing phenotype network as a function of the logarithm of the number of nodes. The linear dependence of the diameter of the phenotype network on the logarithm of the number of nodes shows that the phenotype network exhibits small-world behaviour.

Simulation results regarding the time evolution of both the diameter of the phenotype network and the number of nodes (i.e. the number of viable phenotypes that our evolutionary dynamics (see Section 2.1) allows to emerge) are shown in Fig. 3.8. We observe that the networks of phenotypes generated by our evolutionary model have the small world property as their diameter grows linearly with the logarithm of the number of phenotypes (nodes). This result is stronger than the similar one obtained recently by Aguirre et al. [2] regarding the genotype-phenotype map for RNA secondary structures. They have been demonstrated that the average path length within RNA secondary structure neutral networks scales logarithmically with the size of each independent neutral network. Here, we are able to show that our multi-scale model generates a small-world phenotype network.

From these simulation, we observe that our results regarding the small-network structure of the phenotype network is robust to changes in several parameter values. For example, according to Fig. 3.8, modifying both p, i.e. the rewiring probability of the SW model used to generate the GRN topology, and l_v, essentially, the selective pressure, leaves untouched the small-world property.

Taken together, these two properties, namely, the emergence of a giant connected component within the phenotype network and of the small-world phenomenon, have strong biological implications, in particular regarding evolvability. First, the fact that a giant connected component in phenotype space exists means that it is globally connected and, therefore, innovation by means of genetic mutations is granted regardless of the existence of robust phenotypes. This is consistent with previous results regarding innovation and robustness [105, 106, 106, 108, 28, 27]. On the other hand, the small world property of our phenotype network, implies that, in topological distance, phenotypes are very close to each other, and therefore, adaptive processes where the system must reach a privileged phenotype in order to ensure survival (as, for example, in evolutionary escape [53, 54]) could be much more efficient than previously thought.

3.4 Robustness of evolvability

We have shown in Section 3.3 that, similarly to what has been observed in other models [105, 106, 106, 108, 28, 27] evolvability is an emergent, evolved property of the dynamics. This view in which evolvability is itself an evolutionary property of the dynamics [79], together with our topological characterisation of robustness, allow us to pose the question of whether evolvability itself is a robust property.

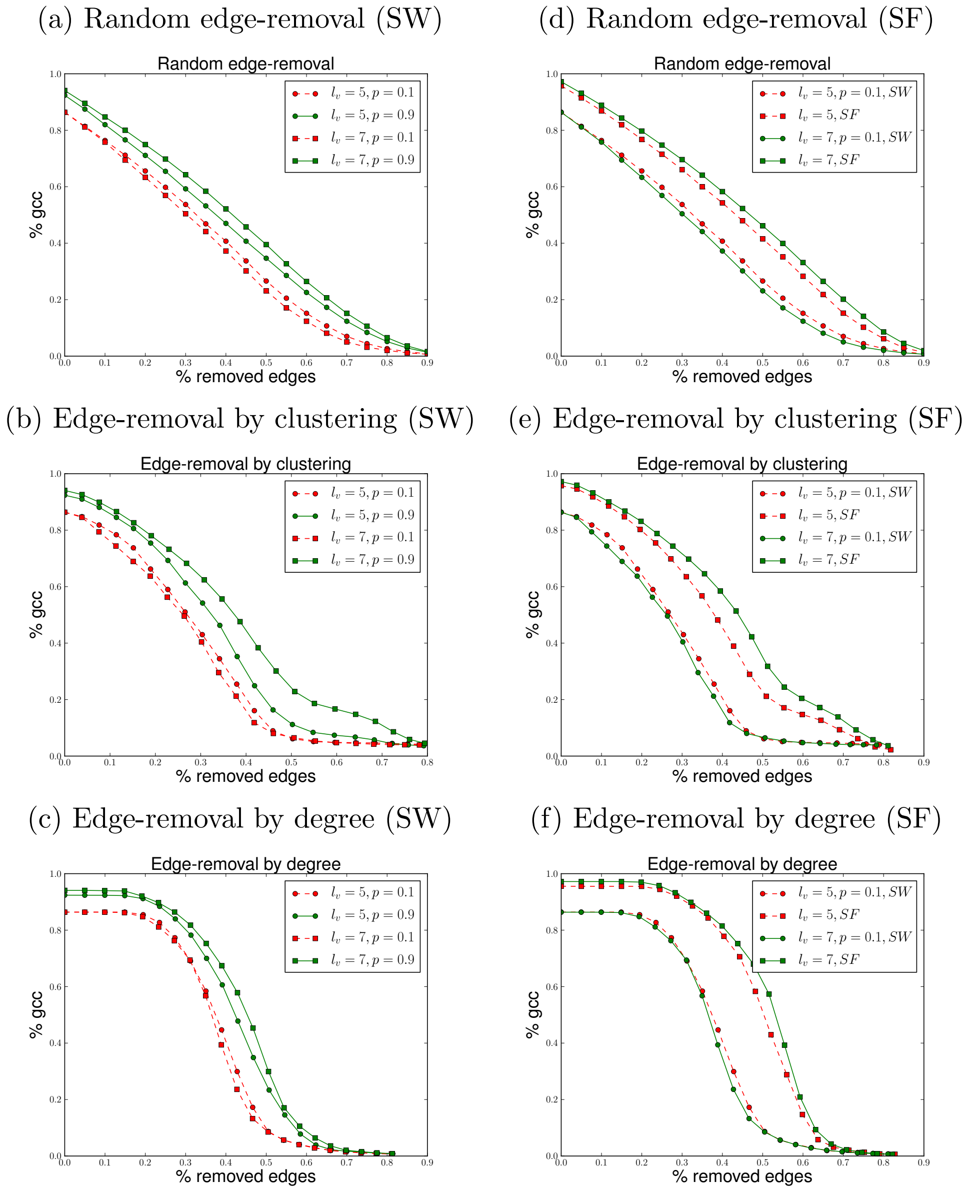
In order to address the issue of robustness of evolvability, we proceed in the following way. Ciliberti et al. [28, 27] have shown that evolvability is associated with the connectivity between neutral networks corresponding to different phenotypes. In Section 3.3.2, we have shown that this interpretation of evolvability leads to its characterisation in terms of the emergence of a giant connected component in the growing phenotype network. Within this framework, a natural definition of robustness of evolvability arises: The robustness of evolvability is associated to the resilience against edge-removal of the corresponding giant connected component.

In order to ascertain how different factors affect the robustness of evolvability, i.e. the integrity of the giant connected component, we consider three different strategies for edge-removal [23]: random removal and two targeted strategies, one in which removal of edges is done according to the clustering phenotypes to which they are connected, and the other in which removal of edges according to the degree of the phenotypes to which they are connected. Note that here clustering coefficient and degree of the phenotypes refer to the corresponding quantities in the bipartite genotype-phenotype network. Results are shown in Fig. 3.9.

Regarding the behaviour of the phenotype giant component generated by Strogatz-Watts GRNs, Fig. 3.9(a) shows that random edge removal has little effect on the integrity of the giant connected component unless a massive percentage of edges is removed. This implies that evolvability is very robust to random elimination of genotypes. By contrast, when edges are removed in targeted manner according to the degree of the phenotype, we observe that there is a sharp transition when the percentage of edges removed exceeds a critical point (see Fig. 3.9(c)). This means that evolvability is not robust to targeted attack where genes associated to high-degree phenotypes are the focus of the attack.

The behaviour described so far is very much resembling of the behaviour of other complex networks, namely, robust against, random attack but sensitive to high-degree-targeted attack [23]. We consider a third scenario in which the attack is targeted on the more robust phenotypes, i.e. those phenotypes with larger clustering coefficient, c_ϕ (Eq. (3.1)). In this case, the behaviour of the system is less clear-cut.

Fig. 3.10 shows the variance of the size of phenotype giant connected component, σ^2_{gcc}, for Strogatz-Watts GRNs. We see that, for the two targeted strategies, this quantity exhibits the sharp increase typical of a second-order phase transition. It is worth noting that the threshold/critical value for the giant connected component to disappear is not significantly different regardless of which targeting strategy is used.


(a) Random edge-removal (SW)

(d) Random edge-removal (SF)

(b) Edge-removal by clustering (SW)

(e) Edge-removal by clustering (SF)

(c) Edge-removal by degree (SW)

(f) Edge-removal by degree (SF)

Figure 3.9: Damage suffered by the phenotype giant connected component. The relative size of the giant connected, $\%gcc$, is defined as the cardinal of the set of nodes such that, for any pair of nodes, it is possible to find at least a continuous connecting path relative to the total number of nodes. Damage is measured in terms of the relative size of the remaining largest connected component after removal of a certain percentage of edges. We consider that edges are (a) removed randomly, (b) according to the clustering coefficient of the nodes (phenotypes) to which they are connected, and (c) according to the degree of the phenotype to which they are connected. Clustering coefficient and degree of the phenotypes refer to the quantities corresponding to the bipartite genotype-phenotype network. Plots (d), (e) and (f) explore the comparison between the behaviour of the phenotype giant connected component generated by a multi-scale model with Strogatz-Watts (SW) and scale-free (SF) gene regulatory networks.

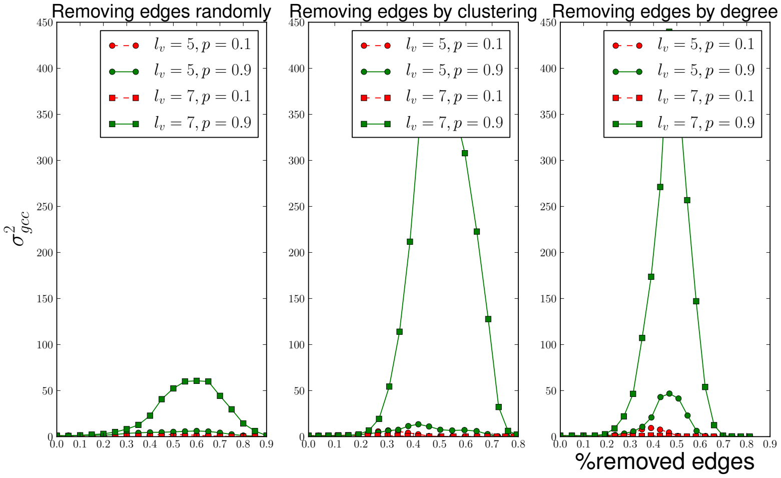


Figure 3.10: Variance of the giant component size from damage done when edges are removed randomly, according to the clustering coefficient of the nodes (phenotypes) to which they are connected, and according to the degree of the phenotype to which they are connected. Clustering coefficient and degree of the phenotypes refer to the quantities corresponding to the bipartite genotype-phenotype network.

Regarding the effects of the Strogatz-Watts GRN parameters, our simulation results show that evolvability is more robust in systems with larger values of the rewiring probability p. In this respect, it is worth noticing that, opposite to the behaviour observed in clustered uncorrelated networks, analysed in depth in [91, 92], we observe that for $p = 0.1$, for which $c_\phi(k_\phi) \leq (k_\phi - 1)^{-1}$ (as shown in Fig. 3.5), the giant connected component is less robust than for $p = 0.9$, for which $c_\phi(k_\phi) \simeq (k_\phi - 1)^{-1}$ (see in Fig. 3.5). This difference in behaviour is due to the presence of correlations.

The presence of correlations is demonstrated by analysing the average number of nearest-neighbours as function of the degree, $k_{nn}(k_\phi)$, which is shown in Fig. 3.11. These results show that $k_{nn}(k_\phi)$ is an increasing function of the phenotype degree and, therefore, the phenotype network is assortative (more connected nodes are preferably connected to each other). Note that, although this is in contrast with the common situation in biological networks lacking an underlying genotype-phenotype structure, which are commonly found to be disassortative (more connected nodes are preferably connected to low-degree nodes) [75], assortativity seems to be a systematic property of phenotype networks. A recent example is provided in [29] where it is observed that large (i.e. strongly connected) phenotypes have an inherently enhanced accessibility to new phenotypes.

It is also interesting to compare the behaviour of the robustness of the phenotype giant connected component generated by Strogatz-Watts (SW) GRNs and by scale-free (SF) GRNs. Figs. 3.9(d), (e) and (f) show that the giant connected component corresponding to SF GRNs is systematically more robust than those generated by SW GRNs.

We now address the effect of varying the mutation rate on our results regarding robustness of evolvability. We have argued in 3.1 that there are two distinct regimes, corresponding to $\mu N > 1$ and $\mu N < 1$ where the population exhibits rather different structure: whilst for $\mu N > 1$ the population is most likely to be polymorphic, for $\mu N < 1$ the population is monomorphic most of the time. We observe that, in agreement with results reported by

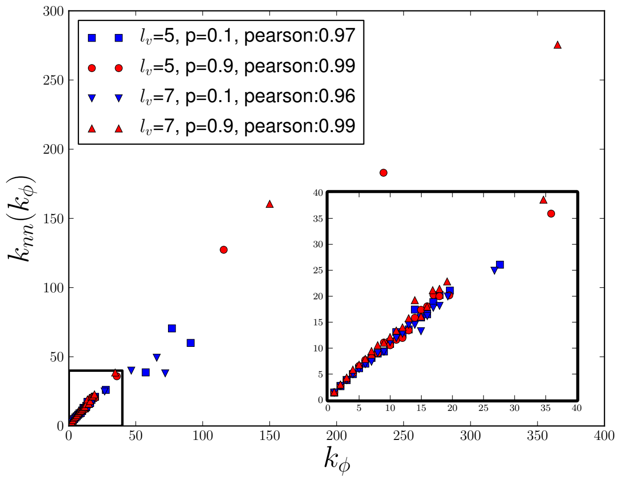


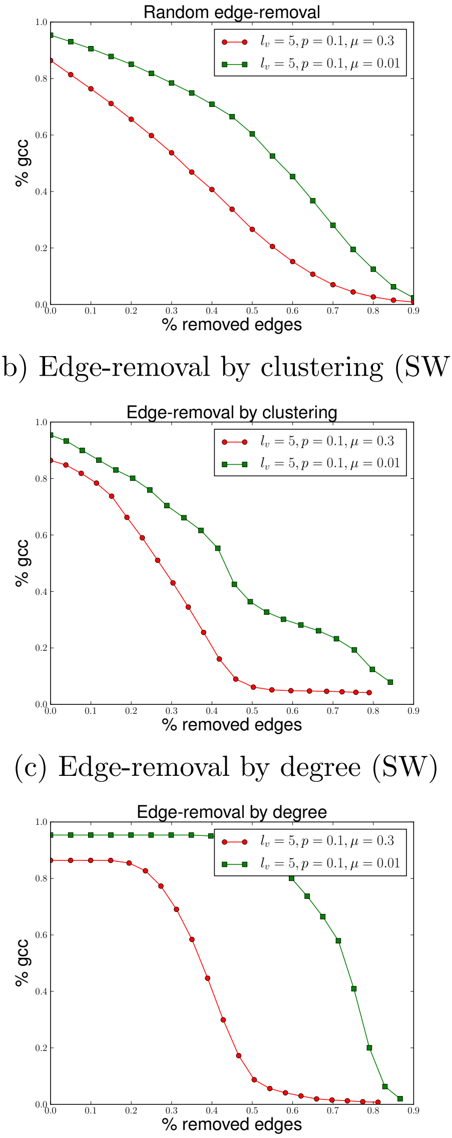
Figure 3.11: Average nearest neighbours as a function of the phenotype degree, k_ϕ, for different values of the Strogatz-Watts GRN parameters, l_v and p. Including a zoom of the figure.

Wagner [106] regarding his analysis of the RNA genotype-phenotype map, where evolvability is shown to be robust to variations in the mutation rate, our multi-scale model exhibits a similarly robustly connected phenotype network. In fact, the phenotype giant connected component, i.e. the core of mutually reachable phenotypes, is found to be more resilient to damage for small mutation rates (in the $\mu N < 1$ regime), regardless of the targeting strategy we use. Increased robustness of the phenotype giant connected component for smaller mutation rate is a consequence of the presence of fatter tails in the genotype and phenotype degree distributions as shown in Section 3.1, which implies that the giant connected component corresponding to $\mu N < 1$ is more tightly connected and, therefore, more difficult to disconnect.

3.5 Conclusions

In this chapter we have presented a novel mathematical description of the genotype-phenotype space in terms of a pseudo-bipartite graph and its associated one-component projection, the phenotype network. We have based our presentation on a model of the genotype-phenotype for circuits of gene regulation, although its extension to other genotype-phenotype models, such as RNA, should be straightforward. This new representation allows us to characterise robustness of evolvability in terms of the topological properties of these networks: phenotypic robustness is defined as the clustering coefficient of phenotype nodes in the pseudo-bipartite genotype-phenotype network, and evolvability is defined as the emergence of a giant connected component in the phenotype network, which ensures global connectedness and navigability of the space of phenotypes.

New results have been obtained regarding the small-world property of the phenotype network, as well as the characterisation of the robustness properties of evolvability. In particular, we have shown that the phenotype network exhibits, beyond global connectedness, the small-world property. We have further exploited our topological definition of evolvability, characterised in terms of the onset of percolation and the size distribution of connected components, to explore the issue of whether evolvability is a robust property, i.e. under which



(a) Random edge-removal (SW)

(b) Edge-removal by clustering (SW)

(c) Edge-removal by degree (SW)

Figure 3.12: Damage suffered by the phenotype giant connected component when edges are (a) removed randomly, (b) according to the clustering coefficient of the nodes (phenotypes) to which they are connected, and (c) according to the degree of the phenotype to which they are connected. Clustering coefficient and degree of the phenotypes refer to the quantities corresponding to the bipartite genotype-phenotype network. Green lines correspond to $\mu = 0.01$ (i.e. $\mu N = 0.5$) whereas red lines correspond to $\mu = 0.3$ (i.e. $\mu N = 1.5$), These phenotype networks are generated by a multi-scale model with Strogatz-Watts with $p = 0.1$.

conditions the giant connected component may be broken down. We observe, that under random attack, the giant connected component, and, therefore, our system's evolvability, is very resilient. However, similarly to other complex networks, when the attack is targeted rather than random, weaknesses arise and evolvability breaks down.

Our results regarding the topological characterisation of robustness and evolvability are complementary to the corresponding definitions given by Ciliberti et al. [28, 27]. By extending their description in terms of a space of genotypes to explicitly include the corresponding phenotypes, we can provide purely topological definitions in terms of easily measurable quantities for which algorithms are readily available [75], namely, the clustering coefficient of the phenotype nodes as a measure of phenotypic robustness, and the existence of a giant connected component in the phenotype network as characteristic of evolvability.

Regarding the characterisation of robustness, our results given in Section 3.2, Fig. 3.5, regarding the dependence of the phenotype clustering coefficient on the degree of the phenotype nodes, $c_\phi(k_\phi)$, show that there exists an inverse relation between clustering coefficient and degree: $c_\phi(k_\phi) \propto (k_\phi - 1)^{-\alpha_{p,l_v}}$ with $\alpha_{p,l_v} \geq 1$. This result has an important interpretation in terms of the concept of *neighbourhood* and the relation between robustness and evolvability [107, 108]. Our result implies that those phenotypes with larger cryptic variability, i.e. genotypic variability which presents no variation in phenotype and which is given by the degree of the phenotype nodes in the pseudo-bipartite genotype-phenotype network, k_ϕ, exhibit, in relative terms, higher accessibility to new phenotypes that those with smaller degree. This observation suggests that these phenotype nodes have a fundamental role to play in the ability for innovation of the system.

Our analysis of the robustness of evolvability is closely related to this issue. We have shown (see Section 3.4, Fig. 3.9) that evolvability is a robust property respect to attack (removal) of randomly selected genotypes. Recall that removal of a genotype node in the genotype-phenotype network corresponds to the removal of an edge in the phenotype network. Mathematically, this corresponds to the giant connected component in the phenotype network being robust against random removal of edges, Fig. 3.9(a). This situation is radically changed when genotype removal is targeted rather than random. In particular, we have carried out two removal strategies, namely, remove first those genotypes connected to phenotypes with bigger clustering coefficient (Fig. 3.9(b)), and remove first those genotypes connected to phenotypes with bigger degree (Fig. 3.9(c)). In other words, we target genotypes associated to more robust phenotypes or more connected phenotypes, respectively. These targeted attacks are much more efficient at breaking the giant connected component. In particular, evolvability is particularly affected by the removal of genotypes associated to phenotypes with bigger degree, k_ϕ. Removal of genotypes associated to phenotypes with bigger clustering coefficient, c_ϕ, seems to be slightly less detrimental, although its effect on evolvability appears to be much more important than that associated to random attack.

Taken together, these results support the theory that robustness and cryptic variability, rather than hindering innovation, they facilitate the ability of evolutionary systems to evolve and adapt. These results are, therefore, in agreement with recent findings by Wagner and co-workers [107, 108].

Evolvability is further quantified by the distribution of size of connected components (see Section 3.4). This measure contains a much more detailed description of evolvability: It provides the probability of a phenotype to be connected to a cluster (i.e. a set of mutually reachable phenotypes) of size s. This distribution provides a description of evolvability as it gives us information about how many new phenotypes are in reach of a given phenotype. It also provides a measure of evolvability before the giant connected component has been formed (as shown in Fig. 3.6) and, equivalently, after the damage to the phenotype network has been done. If we look at Fig. 3.6, for example, we realise that before the formation of the giant connected component some degree of evolvability is still possible. Taken together both topological measures, existence of a giant connected components and distribution of size of connected components, provide a characterisation of evolvability which allows us to explore whether evolvability is at all possible and whether a macroscopic proportion of phenotypes are mutually reachable.

A new result that emerges from our analysis is the fact that the genotype-phenotype map is such that the phenotype network exhibits the small-world phenomenon (see Section 3.3, Fig. 3.8). This result implies that the maximum average distance between two given phenotypes on the phenotype network scales as the logarithm of the number of phenotypes, which is much smaller than the distance one should expect if the phenotype-genotype space were lattice-like, as it is commonly assumed in models of evolutionary escape, for example [53, 54]. This topological result has obvious implications regarding the rate of evolutionary adaptation, which should be greatly increased by this property. This result is consistent too with recent studies suggesting that cryptic genetic variation increases the rate of evolutionary adaptation in RNA enzymes [46].

Besides the issues arising from our topological characterisation of robustness and evolvability and their biological relevance, we have also investigated their controllability, i.e. to what extent change of the parameters of the GRNs can be used to drive changes in robustness and evolvability. Although our results show that the generic properties of the genotype-phenotype space are rather insensitive to these parameters, we have found that some level of control is indeed possible. In particular, in the case of GRN topology generated using the Strogatz-Watts model, from our results shown in Fig. 3.5, we can conclude that a measure of control on the level of robustness can be achieved by tuning the characteristic parameters of the GRN. Increasing the rewiring probability p appears to lower the overall phenotypic robustness, as shown in Fig. 3.5 where we observe that the clustering coefficient as a function of phenotype degree for $p = 0.1$ is increased with respect to the corresponding values for $p = 0.9$. We have also investigated how some model details affect our results. In particular, we have explored the effect of (i) considering a scale-free GRN topology, using the Barabási-Albert model instead of the Strogatz-Watts model, and (ii) varying the value of the mutation rate, μ. We have observed that both considering scale-free GRN topology and decreasing μ render the evolvability more robust, as both these factors produce phenotype networks with giant connected components that are more resilient to damage.

A biological scenario where our topological characterisation of robustness and evolvability could be a useful tool is cancer. In a recent review, Tian et al. [98] have advocated that in order to understand the principles of tumour evolvability, which, in turn, are key to analyse issues such as resistance to therapy and cancer stem cells, a systems biology approach must be used which integrates the different factors that contribute to heritable phenotypic variability (genetic and epigenetic instability, stochastic protein dynamics, tumour microenvironment, etc.). Central to these issues is the concept of *epigenetic landscape*, which are inspired by the concept of energy landscape familiar in Physics: a space populated with diverse attractor states (i.e. minima of the landscape), corresponding to different cell states and separated by *epigenetic barriers*. In the context of our model, the epigenetic landscape is represented by the phenotype network.

In Developmental Biology, these attractors correspond to the stable-steady states of the GRN that regulate cell differentiation and represent the differentiation states of the cell [47, 49, 48, 99], and transitions between these states occur following a well-orchestrated series of events, otherwise cells are locked into the corresponding state. The epigenetic landscape is organised so that (plenipotentiary) stem cells sit atop a hierarchy of connected attractors that radiate outwards to stable states which represent distinct cell fates.

In the case of cancer cells, pathological attractors that represent different cancer states are scattered over the epigenetic landscape. In some cases, cancer cells posses a landscape with well-defined minima and correspondingly stable cancer states similar to developmental landscape. However, due to severe genetic and epigenetic disregulations occurring in aggressive forms of cancer, it appears that a de-stabilisation of the epigenetic landscape ensues where epigenetic barriers are lowered down and attractors becoming increasingly less stable, thus increasing evolvability and the ability of the tumour to evade therapy. Supporting evidence for this model (reviewed in [98]) has been recently found [82, 93] where genetic and epigenetic dysfunctions allow for the population of cancer cells to transiently visit a number of metastable states within the epigenetic landscape, thus disrupting the hierarchical organisation of normal tissues and favouring the emergence of phenomena such as transient drug resistance [93]. The application of the formalism developed to such issues would require to develop a model of the GRN governing such states as well as the corresponding epigenetic regulation [82] and then analyse using our topological description (based on percolation on and distribution of component size in the phenotype network) how disregulations at the genetic and epigenetic levels would affect evolvability, whereby targeted strategies aimed at phenotypes which cause the most disruption on evolvability and produce more static tumours could be formulated. The formulation of such models and the corroboration of whether such strategies are based on topological properties (like the ones proposed in Section 3.4) or not is beyond the scope of this work and left for future research.

Chapter 4

Evolutionary escape on complex genotype-phenotype networks

In this Chapter, we study the problem of evolutionary escape for cell populations with genotype-phenotype map. This setting involves two major variations with respect to previous models of evolutionary escape. Rather than assuming that the genotype space is a regular hypercube, we consider complex genotype-phenotype network (which we dub B-graphs), obtained from a model which assumes a selective pressure acting at the level of phenotypes, which have certain topological properties which are absent in regular hypercubes. Furthermore, the consideration of the genotype-phenotype structure allows us to associate fitness to phenotypes rather than genotypes. Our aim is to carry out a comparative analysis in which the effects that these two factors, i.e. complex genotype-phenotype topology and fitness associated to phenotype, have on evolutionary escape. Specifically, we study the effects on the probability of escape and the escape rate associated to the evolutionary dynamics occurring on a genotype-phenotype network rather than on a regular hypercube (which we denote as H-graphs). We apply a general theory, based on multi-type branching processes, in order to compute the evolutionary dynamics and probability of escape, which takes into account the structure of the genotype-phenotype space. We show that the heterogeneity observed in the distribution of distances between phenotypes in B-graphs, one of the main structural differences between both types of graphs, causes heterogeneous behaviour in all results associated to the escape problem. We further show that, due to the heterogeneity characterising escape on B-graphs, escape probability can be underestimated by assuming a regular hypercube genotype network. Similarly, it appears that the complex structure of B-graphs slows down the rate of escape.

4.1 Evolutionary escape

As we have introduced in Section 1.1.3 evolutionary escape is the process whereby a population under sudden changes in the selective pressures acting upon it try to evade extinction by evolving from previously well-adapted phenotypes to those that are favoured by the new selective pressure. In this thesis we take into account that selective pressures act on phenotypes rather than genotypes, then evolutionary escape is best described in terms of a population dynamic that accounts for the genotype-phenotype map. This modification alters the approach proposed by Iwasa and co-workers in two significant ways. First, due to evolved robustness in populations with genotype-phenotype map [104, 28, 27, 105, 106, 106, 55], not every gene

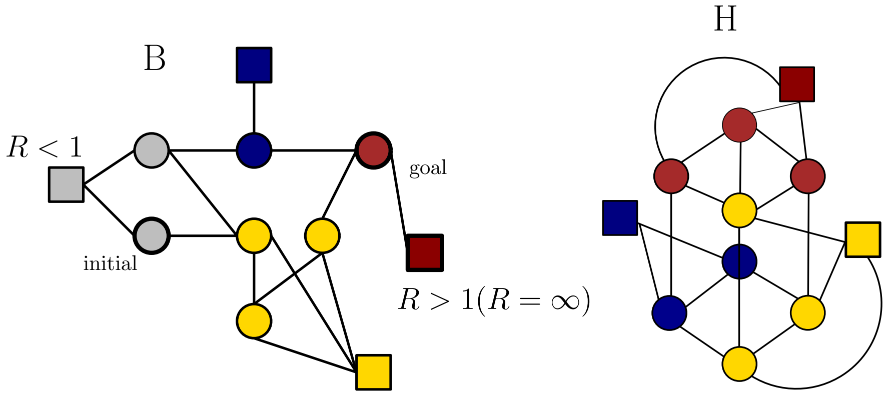


Figure 4.1: Schematic representation of the two types of networks considered in this manuscript: Plot (a) corresponds to a pseudo-bipartite genotype-phenotype graph, B, constructed according to the model [52], whereas (b) corresponds to a hypercube graph, H. In both plots, phenotype nodes are represented as squares and genotype nodes as circles. Each colour represent a different phenotype and their associated genotypes. Edges between nodes allows identify map between genotype-phenotype and feasible mutations between genotypes. Heavier strokes in some nodes indicates the initial genotype, characterised by a reproduction number such that $R < 1$, i.e. bound to long-term extinction, and the goal or escape genotype, which we consider to have a reproduction number $R = \infty$). In other words, we assume that once the system reaches a genotype corresponding to the escape phenotype, the extinction probability is vanishingly small.

mutation necessarily generates a new phenotype. As a consequence, many gene mutations are neutral as far as the evolutionary escape process is concerned. Furthermore, it has been shown that the topology of genotype-phenotype networks is far from that of the hypercube lattice assumed by Iwasa et al. [2, 52]. In fact, we have shown in Chapter 3 that the corresponding phenotype network exhibits the small-world phenomenon and that, as a consequence, accelerated evolvability (relative to that of a system with no genotype-phenotype map) may emerge. The question naturally arises as to whether these properties, i.e. phenotypic robustness and evolvability typical of genotype-phenotype networks, have an influence on the process of evolutionary escape. To address this issue, we apply a general theory, based on multi-type branching processes [57], to compute the evolutionary dynamics and probabilities of escape which takes into account the structure of the genotype-phenotype space.

4.2 Mathematical model

The classical escape model [53, 54, 89, 90, 84] can be summarised as follows. Each of the 2^n nodes of an n-dimensional hypercube is assumed to represent a genotype. Fitness values, represented here by the reproductive ratio, R, defined as the average number of offspring per individual, are assigned directly to genotypes. The population is assumed to be concentrated in one genotype which, prior to the change in selective pressure (as a consequence of e.g. the administration of a drug), was well adapted. After treatment commences, this initial genotype becomes ill-adapted to the new selective pressure, i.e. its reproductive ratio becomes $R < 1$. To avoid extinction the population needs to start a random, mutation driven search of the genotype hypercube, until it finds an *escape genotype*, i.e. a genotype such that its reproductive number satisfies $R > 1$. The reproductive number of all genotypes other than the escape genotype are such that $R < 1$. This implies that only the escape genotype has a positive probability of long-term survival [57].

The random search on the genotype hypercube is performed by an evolutionary dynamics that take the form of a multi-type Galton-Watson branching process with mutation [53, 54, 89, 90, 84]: Upon proliferation, each cell has a certain probability of mutating, the mutation rate μ, which allows the population to spread over the genotype hypercube until escape is achieved, by sequentially moving between genotypes until the escape genotype is reached, or extinction eventually occurs.

4.2.1 Genotype-phenotype network

We extend the basic framework for analysing evolutionary escape by introducing two modifications. The first one has to do with the topology of the space on which the evolutionary dynamical process occurs. We assume that space where the escape process is performing its random search is not a regular hypercube [53, 54, 89, 90, 84], but a complex genotype-phenotype network [28, 27, 2, 52].

In order to proceed further, we consider two classes of graphs (see Fig. 4.1):

- B: Genotype-phenotype networks as modelled in Chapter 2 and [52],
- H: An *artificial* genotype-phenotype graph where the genotype space is given by a hypercube. Phenotypes are assigned randomly to genotypes so that the phenotype degree distribution, i.e. the probability distribution of the number of genotypes bearing a given phenotype, is the same as that resulting from the multi-scale model in [52].

4.2.2 Population dynamics

A further modification we introduce with respect to the original model by Iwasa et al. [53] is the fact that we now associate fitness to phenotypes rather than genotypes, i.e. the value of the reproduction number depends on the phenotype: Different genotypes which exhibit the same phenotype have the same reproduction number.

In this section we summarise the methodology used to compute the escape probability, i.e. the probability of a population to reach the escape phenotype, ϕ_E, through a process of birth and mutation before extinction occurs, focusing on the effect of considering the escape process on the different genotype-phenotype networks considered in our analysis, namely, graphs of classes B or H (see Fig 4.1).

In order to model our evolutionary dynamics on B- and H-class networks, we follow the previous literature on the subject [53, 54, 89, 90, 84] and consider a Galton-Watson multi-type branching process [57]. The process takes place on the genotype network and each type corresponds to a different genotype. When we refer to the population of type i at generation t, $N_i(t)$, we mean the number of cells with genotype $\mathcal{G}_i$ at time t. Furthermore, we define

$$N_E(t) = \sum_{i \in \langle \phi_E \rangle} N_i(t)$$

as the total population of the escape phenotype. The sum in the above expression is done over all the genotypes belonging to $\langle \phi_E \rangle$, which is the set of genotype nodes whose phenotype is the escape phenotype ϕ_E, i.e. all those genotypes such that $\phi(\mathcal{G}_i) = \phi_E$.

The process is assumed to start with a clonal population, i.e. the whole initial population concentrated in one single genotype, $\mathcal{G}_0$, such that $\phi(\mathcal{G}_0) \neq \phi_E$. The evolutionary dynamics is characterised by two parameters: the birth probability and the mutation rate. Whereas the latter is assumed to have be independent on genotype/phenotype of the cell, the birth probability, λ, is assumed to be dependent on phenotype, i.e. $\lambda = \lambda(\phi_i)$. From the point of view of the evolutionary dynamics, this implies that all genotypes associated to the same phenotype have the same birth probability. The death probability, σ, is given by $\sigma = 1 - \lambda$ and it therefore depends on the phenotype. We further define the reproduction ratio [53] $R(\phi_i) = \lambda(\phi_i)/\sigma(\phi_i)$. For simplicity, we assume that $\lambda(\phi_i) = \lambda =$cnst. for all genotypes such that $\phi(\mathcal{G}_i) = \phi_i \neq \phi_E$. We further assume that $\lambda \ll \lambda_E$ where λ_E is the birth probability of those genotypes such that $\phi(\mathcal{G}_i) = \phi_E$. In fact, we consider $R(\phi_E) \to \infty$, so that once the system reaches an escape genotype, the survival probabilities $P_S(t) \to 1$.

The evolutionary dynamics is defined as follows. At each time step (generation) each individual can:

- Reproduce with no mutation with probability $\lambda(1-\mu)^2$.
- Reproduce with asymmetrical mutation, i.e. one of the descendants mutates, the other retains the genotype of its mother cell. This event occurs with probability $2\lambda\mu(1-\mu)$.
- Reproduce with symmetric mutation, i.e. both descendants mutate, which occurs with probability $\lambda\mu^2$.
- Die with probability $\sigma = 1 - \lambda$.

This dynamic is iterated until either an escape genotype is reached, upon which escape is assumed to occur with probability one, or the population undergoes extinction.

In order to proceed further, we recall the following definitions regarding multi-type Galton-Watson branching processes [57] (described in Section 1.2.2):

- We define $f_i(s_1, \ldots, s_n; t) = \mathbb{E}(s_1^{N_1(t)} \cdot s_2^{N_2(t)} \cdots s_n^{N_n(t)} \mid N_i(0) = 1, N_j(0) = 0, \quad \forall j \neq i)$ as the generating function of probability of the population to be $(N_1(t), \ldots, N_n(t))$ at time conditioned to the initial condition of the system to consist of a single individual of type i, i.e. $N_j(t=0) = \delta_{i,j}$ for $j = 1, \ldots, n$. we further define $\vec{s} = (s_1, s_2, \ldots, s_n)$. Each component, s_i, satisfies $0 \leq s_i \leq 1 \quad \forall i$.
- We consider the progeny probability generating function, $F_i(\vec{s})$, which is defined as the generating function corresponding to the probability distribution of the number of offspring of a cell with genotype $\mathcal{G}_i$

The dynamics of a multi-type Galton-Watson process is described by two equivalent functional equations, namely, the forward equation

$$\vec{f}(\vec{s}, t+1) = (\vec{F} \circ \vec{F} \circ \cdots \circ \vec{F})(\vec{s}) = \vec{F}(\vec{F}(\vec{F} \ldots F(\vec{s}))) \tag{4.1}$$

where the progeny probability generating function is composed with itself $t+1$ times, and the backward equation:

$$\vec{f}(t+1) = \vec{f}(\vec{F}(\vec{s}), t) \tag{4.2}$$

where $\vec{f} = (f_1, \ldots, f_n)$ and $\vec{F} = (F_1, \ldots, F_2)$.

Table 4.1: Description of parameters.

Parameter	Description
R	Reproduction rate
μ	Mutation rate in the branching process
ν	Mutation rate in generation of a B class graph, it is equal to parameter μ in Chapter 2
$N_i(t)$	Number of cells with genotype $\mathcal{G}_i$ at time t
$N_E(t)$	Total number of cells in the escape phenotype ϕ_E
B class graph	Pseudo-bipartite graph generated by model in Chapter 2
H class graph	Artificial genotype-phenotype graph
λ	Birth probability in genotypes $\mathcal{G}_i$ such that, $\phi(\mathcal{G}_i) \neq \phi_E$
σ	Death probability in genotypes $\mathcal{G}_i$ such that, $\phi(\mathcal{G}_i) \neq \phi_E$
λ_E	Birth probability in genotypes $\mathcal{G}_i$ such that, $\phi(\mathcal{G}_i) = \phi_E$
σ_E	Death probability in genotypes $\mathcal{G}_i$ such that, $\phi(\mathcal{G}_i) = \phi_E$
$P_S(t)$	Survival probability at time t
$(\vec{\theta}_t)_i = (f(\vec{\theta}_0, t))_i$	Probability of no individuals have reached ϕ_E, assuming an initial individual of type i
N	Initial population to generate B graphs using model in Chapter 2

According to our model evolutionary dynamics, the progeny generating function of a non-escape genotype $\mathcal{G}_i$, i.e. $\phi(\mathcal{G}_i \neq \phi_E)$, $F_i(\vec{s})$ is given by:

$$F_i(\vec{s}) = \sigma + \lambda(1-\mu)^2 s_i^2 + \sum_j 2\lambda\mu(1-\mu)\frac{a_{ij}}{d_i} s_i s_j + \sum_{j,k} \lambda\mu^2 \frac{a_{ij}a_{ik}}{d_i^2} s_j s_k \tag{4.3}$$

where $A = (a_{ij})$ is the adjacency matrix of genotype graph (defined as a sub-graph of the genotype-phenotype network) and d_i is the degree of of genotype i in the genotype network.

On the other hand, the progeny generating function of an escape genotype, i.e. $\phi(\mathcal{G}_i = \phi_E)$, is given by:

$$F_i(\vec{s}) = s_i. \tag{4.4}$$

In order to simplify notation we will define for non escape genotypes, we further define the matrix $D = (d_{ij} = d_i\delta_{i,j})$. In terms of the matrices A and D, $\vec{F}(\vec{s})$ can be re-written as:

$$\begin{aligned} \vec{F} =& \sigma\vec{1} + \lambda(1-\mu)^2\vec{s}\odot\vec{s} + 2\lambda\mu(1-\mu)(D^{-1}A\cdot\vec{s})\odot\vec{s} + \lambda\mu^2(D^{-1}A\cdot\vec{s})\odot(D^{-1}A\cdot\vec{s}) = \\ &= \sigma\vec{1} + \lambda(B\cdot\vec{s})\odot(B\cdot\vec{s}) \end{aligned} \tag{4.5}$$

where $\odot$ denotes the component-to-component product and $B = \mu D^{-1}A + (1-\mu)\mathrm{Id}$ with Id equal to the identity matrix.

Escape time probability The generating function $\vec{f}(\vec{s}, t)$ encodes all the information of the process, in particular, that pertaining to the escape probabilities. To calculate the escape probability, we need to fix the initial genotype and the escape phenotype. The escape phenotype has associated all those genotypes such that $\phi(\mathcal{G}_i) = \phi_E$. In order to proceed further, we define:

$$\vec{\theta}_0 := (1, 1, 1, \ldots, 0, 0)$$

where

$$(\vec{\theta}_0)_i = \begin{cases} 0 \text{ if } \phi(\mathcal{G}_i) = \phi_E \\ 1 \text{ otherwise} \end{cases} \tag{4.6}$$

Since $\vec{\theta}_t := f(\vec{\theta}_0, t)$, then $\vec{\theta}_t$ satisfies

$$(\vec{\theta}_t)_i = P(N_E(t) = 0 \mid N_i(0) = 1, N_j(0) = 0 \quad \forall i \neq j),$$

that is the probability of no individuals to have reached the escape phenotype, assuming that the population is initially composed by one individual of type i. Then,

$$\begin{aligned}(\vec{\theta}_t)_i = \mathbb{E}(s_1^{N_1(t)} \cdot s_2^{N_2(t)} \cdots s_n^{N_n(t)} \mid N_i(0) = 1, N_j(0) = 0 \forall j \neq i)\Big|_{\vec{s}=\vec{\theta}_0} &= \\ = P(N_{k_1}(t) = n_{k_1}, \ldots, N_{k_l}(t) = n_{k_l} \mid N_i(0) = 1, N_j(0) = 0, \forall j \neq i) & \\ = P(N_E(t) = 0 \mid N_i(0) = 1, N_j(0) = 0, \forall i \neq j) &\end{aligned} \tag{4.7}$$

where $k_i, i = 1, \ldots, l$ refer to all those genotypes such that $\phi(\mathcal{G}_{k_i}) = \phi_E$. The quantities $(\vec{\theta}_t)_i$ are obtained by iteration of Eq (4.5):

$$\overrightarrow{\theta_{n+1}} = \sigma\vec{1} + \lambda(B \cdot \vec{\theta_n}) \odot (B \cdot \vec{\theta_n}) \tag{4.8}$$

The quantities $(\vec{\theta}_t)_i$ are closely related to the escape probability at time t. Consider the probability of not reaching the escape phenotype at time $t-1$, $P(N_E(t-1) = 0)$, which can be expressed as:

$$\begin{aligned} P(N_E(t-1) = 0) = \quad & P(N_E(t-1) = 0 \mid N_E(t) > 0) P(N_E(t) > 0) \\ & + P(N_E(t-1) = 0 \mid N_E(t) = 0) P(N_E(t) = 0) \end{aligned}$$

Recalling that $P(N_E(n-1) = 0) = (\vec{\theta}_{t-1})_i$ and

$$P(N_E(n-1) = 0 \mid N_E(n) = 0) P(N_E(n) = 0) = P(N_E(n) = 0) = (\vec{\theta}_t)_i,$$

we have that:

$$(\vec{\theta}_{t-1})_i - (\vec{\theta}_t)_i = P(N_E(t-1) = 0 \mid N_E(t) = 0) P(N_E(t) > 0),$$

which, in turn, implies that the probability of reaching the escape phenotype precisely at time t, or, in other words, the escape time probability, $P(N_E(t-1) = 0 \wedge N_E(t) > 0 \mid N_i(0) = 1, N_j(0) = 0, \forall i \neq j)$ is given by:

$$P_E(t) = P(N_E(t-1) = 0 \wedge N_E(t) > 0 \mid N_i(0) = 1, N_j(0) = 0, \forall i \neq j) = (\vec{\theta}_{t-1} - \vec{\theta}_t)_i \tag{4.9}$$

Since $\mathbb{E}(N_E(t))$ is an increasing function of t, it follows that:

$$\begin{aligned} & P(N_E(t) = 0 \mid N_i(0) = 1, N_j(0) = 0, \forall i \neq j) \geq \\ & P(N_E(t+1) = 0 \mid N_i(0) = 1, N_j(0) = 0, \forall i \neq j), \end{aligned}$$

and therefore we have that $(\vec{\theta}_t)_i \geq (\vec{\theta}_{t+1})_i$. This inequality has two important consequences, namely, (i) $(\vec{\theta}_{t-1} - \vec{\theta}_t)_i \geq 0$, which guarantees that the escape probability at time t (Eq. (4.9)) is well-defined (i.e. non-negative), and (ii) $\vec{\theta}_t$ converges to $\vec{\theta}_\infty$ as $t \to \infty$.

Escape probability $P(N_E(\infty) > 0)$ From the above results, we are able to compute the asymptotic escape probability, $P(N_E(\infty) > 0)$, in terms of the escape time probability $P_E(t)$:

$$P(N_E(\infty) > 0) = 1 - (\vec{\theta}_\infty)_i = \sum_t P_E(t) \tag{4.10}$$

where $\vec{\theta}_\infty = \lim_{t\to\infty} \vec{\theta}_t$ satisfies:

$$\vec{\theta}_\infty = \sigma\vec{1} + \lambda(B \cdot \theta_\infty) \odot (B \cdot \theta_\infty) \tag{4.11}$$

Alternative recursion We consider the following recursion in order to obtain escape probability at time t. We consider the quantity $\vec{1} - \vec{\theta}_n$. By defining $\vec{\psi}_n = \vec{1} - \vec{\theta}_n$, Eq (4.8) leads to:

$$\vec{1} - \overrightarrow{\psi_{n+1}} = \sigma \cdot \vec{1} + \lambda(B \cdot\ (\vec{1} - \vec{\psi}_n)) \odot (B \cdot\ (\vec{1} - \vec{\psi}_n)) \tag{4.12}$$

After some algebra, we obtain the following recursion:

$$\vec{\psi}_{n+1} = 2\lambda B\vec{\psi}_n - \lambda(B\vec{\psi}_n \odot B\vec{\psi}_n) \tag{4.13}$$

This new recursion, Eq (4.13), is numerically better behaved than Eq (4.8). Therefore, it provides a more accurate approximation.

4.3 Results

In this section, we report the results of our comparative analysis of evolutionary escape on B and H genotype-phenotype graphs. We first focus on structural properties of both types of genotype-phenotype networks, in particular, those charactering the distance between phenotypes, which is an essential property of the networks affecting evolutionary escape. We then proceed to analyse the dynamics of escape on both types of networks, according to the model presented in Section 4.2.2. We first focus on the steady-state ($t \to \infty$) properties of the escape probability. We then move on to the study of dynamical properties of evolutionary escape as characterised by several metrics, namely, the escape time probability, the average escape time, and the escape time probability conditioned to escape.

4.3.1 Connectivity structure of genotype-phenotype networks

As a first step towards understanding the properties of the evolutionary escape phenomenon on genotype-phenotype networks, we investigate the connectivity structure of our B graphs, as this structure directly affects the properties of random walks on networks [64, 30, 87] and is therefore straightforwardly connected to the escape dynamics.

In order to proceed with our analysis, we start by investigating the distance between phenotypes on B-networks. An estimation of this quantity can be obtained via the average shortest path length between phenotypes which is computed as follows. For each genotype-phenotype graph, we consider its giant connected component (GCC). More specifically, we consider the genotype one-mode projection of the GCC. Note that, as we are not considering the phenotype nodes, the genotype one-mode projection of the GCC is not necessarily connected. Once this one-mode projection has been determined, we compute the shortest path between every pair of genotypes, which in turn allows us to compute the average distance between phenotypes by averaging over the distances between genotypes belonging to two given

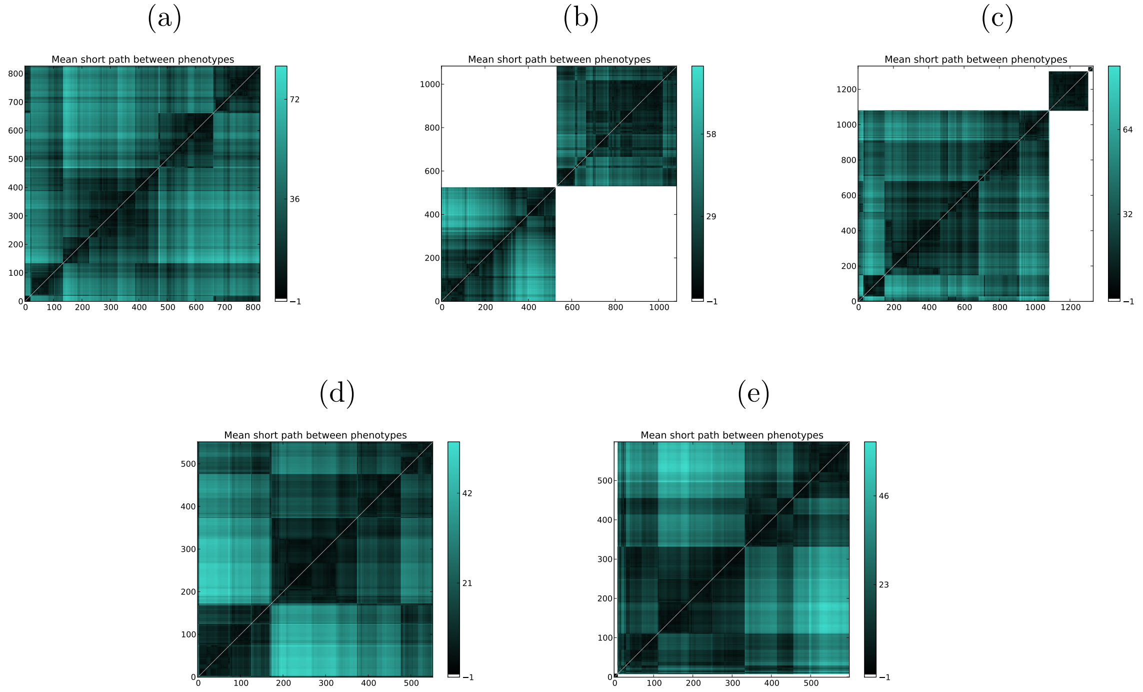


Figure 4.2: Mean short path between phenotypes ordered by clusters. Different graph structures are observed. White corresponds to no path between two phenotypes, black or really dark means distance from 1 to 10 (approx), from black to turquoise length path increases. These are graph with viability 7 and GRN modelled as Strogatz-Watts graph with parameter $p = 0.1$. Initial genotypes, $N = 50$ and for (a),(b),(c) mutation rate is $\nu = 0.3$, (d) and (e) $\nu = 0.01$.

phenotypes. In order to visualise our results of the network we make a cluster hierarchy, using a method implemented in Python: `scipy.cluster.hierarchy.linkage`. Several examples of this cluster analysis are shown in Fig. 4.2, where we have colour-coded the different phenotypes according to the distance between them: darker (lighter) colour is associated to smaller (bigger) average shortest path between corresponding phenotypes. White indicates that there is no path between phenotypes (i.e. the average shortest path is of infinite length).

Our analysis shows that genotype-phenotype networks exhibit a high degree of heterogeneity. We distinguish between two cases: genotype-phenotype networks associated to large populations (i.e. $\nu N > 1$) and genotype-phenotype networks corresponding to small populations (i.e. $\nu N < 1$). The evolutionary dynamics of populations in these two limits is fundamentally different: whereas populations with high mutation rates ($\nu N > 1$) are likely to be polymorphic at any given generation, with individuals accumulating over time in regions of the genotype space with genotypes characterised by high mutational robustness, populations with low mutation rate ($\nu N < 1$) tend to be monomorphic with individuals performing a random walk that samples the genotype network uniformly [106]. These properties are reflected in our analysis: we observe that those genotype-phenotype networks associated to $\nu N > 1$ are more likely to exhibit an structure where separated communities of phenotypes, i.e. genotype-phenotype networks where disjoint subsets phenotypes emerge (see Figs. 4.2(b) & (c)), thus enabling for polymorphic populations to appear. Such disjoint subsets of phenotypes are far less likely in genotype-phenotype maps associated to $\nu N < 1$ (see Figs. 4.2(d) & (e)). Note that, although disjoint phenotype networks are much more likely for genotype-phenotype maps with $\nu N > 1$, it is possible to observe fully connected phenotype networks associated to $\nu N > 1$. An example is shown in Fig. 4.2(a). Similarly, disjoint phenotype networks may also exist for $\nu N < 1$, although they are much less likely than the fully-connected ones. The emergence of disjoint subsets of phenotypes has an immediate consequence on evolutionary escape: since some phenotypes are unreachable, escape is not going to be possible in the initial phenotype and the escape phenotype live in different disjoint subsets.

In order to gain a more quantitative understanding connectivity between phenotypes, we have plotted the distribution of the average distance between phenotypes corresponding to the data shown in Fig. 4.2. We have further computed the same data for the associated H-type genotype-phenotype graph: according to the procedure explained in Section 4.2, for each B genotype-phenotype map, we compute a H graph by randomly linking genotype-phenotype pairs under the constraint that both B and H graphs have the same phenotype degree distribution. The results of computing the average distance between phenotypes in both types of graphs is shown in Fig. 4.3, from which a property stands out, namely, B genotype-phenotype graphs exhibit a much higher degree of heterogeneity than their H-type counterparts, as shown by the bigger width of the distributions for B-type graphs (Figs. 4.3(a), (c), and (e)) compared to that of the distributions associated to H- graphs (Figs. 4.3(b), (d), and (f)).

4.3.2 The escape probability exhibits higher degree of heterogeneity in B graphs

Using the methods of Section 4.2.2, we have proceeded to compute the long-term escape probability $P(N_E(\infty) > 0)$ for B-type genotype-phenotype graphs, i.e. those for which the nodes of the genotype network is the set of genotypes associated to viable phenotypes, and compare to escape probability associated to H-type genotype-phenotype graphs, for which the genotype network is a hypercube.

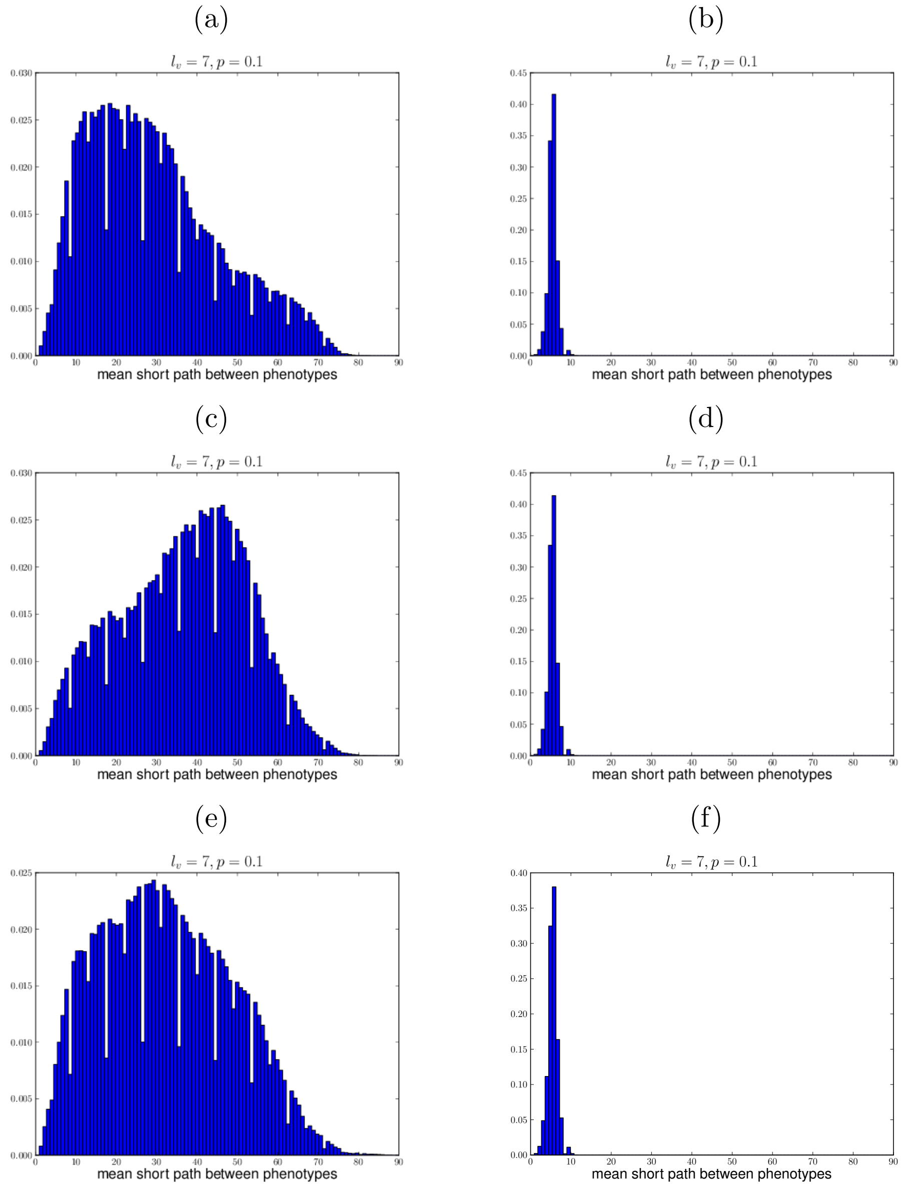


Figure 4.3: Series of plots showing the normalised histograms for the average distance between phenotypes associated to the data corresponding to the genotype-phenotype graphs shown in Fig. (4.2 (a), (b), (c), $N\nu = 15$) (plots (a), (c), and (e)), and the normalised histograms for the corresponding H-type genotype-phenotype maps (plots (b), (d), and (f)).

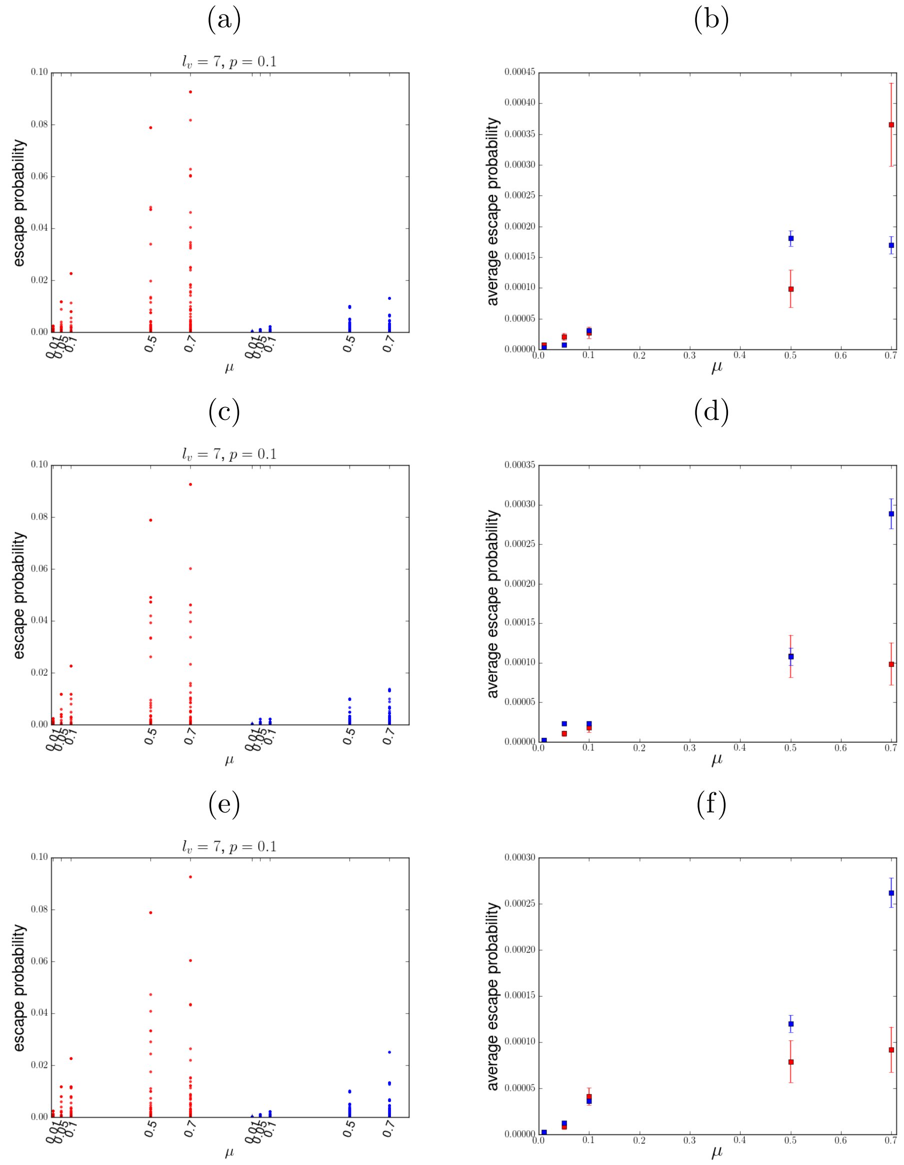


Figure 4.4: Comparison between the long-term escape probability, $P(N_E(\infty) > 0)$, for B (in red) and H (in blue) genotype-phenotype graphs as the mutation rate, μ, varies. Each row plots corresponds to a different graph from Fig 4.2 (a), (b), (c), $N\nu = 15$. For each graph, we fix an escape phenotype, and compute the escape probability for all remaining initial phenotypes. These plots show results for ten different escape phenotypes chosen at random. Plots (a), (c), and (e) show scatter plots for results for each initial condition, whilst plots (b), (d), and (e) show the escape probability averaged over initial conditions. $\lambda = 0.1$ and $\sigma = 0.9$.

Fig. 4.4 shows simulation results showing how the escape probability, $P(N_E(\infty) > 0)$, changes as the mutation rate, μ, is varied. Results are shown for both types of genotype-phenotype maps. We observe (see Figs. 4.4(a), (c), and (e)) that much higher levels of variability are obtained for B-type genotype-phenotype graphs. Such heterogeneity is directly inherited from the connectivity properties of B and H graphs (see Fig. 4.3): since the distance between phenotypes in H graphs is much more homogeneous than in B graphs so is the escape probability. Our analysis shows that the average escape probability (Figs. 4.4(b), (d), and (e)) is not informative to distinguish between escape in either type of graph. Moreover, whereas this average is representative of the behaviour of H graphs, it is likely that for B graphs the average escape probability, due to their intrinsic heterogeneity, is not characteristic of their behaviour.

Furthermore, as a consequence of the heterogeneous behaviour of B genotype-phenotype graphs, one can find cases in which the escape probability on the B graph is an order of magnitude larger than on the H graph. In these instances, the theory of escape based on regular hypercube genotype spaces seriously under-estimates the escape probability.

4.3.3 Escape dynamics: asymptotic behaviour of escape time probability

Taken a graph of set B and another of H class, we compute escape probabilities for exactly each time t and represent them in Fig 4.5. It seems clear that distribution tails, for t large, decay exponentially. Also we can note that a bigger variability of results is observed in the B genotype-phenotype graph while in the hypercube probabilities are less variable independently of the pair of genotypes chosen. This fact is given by a higher richness of structure in B graphs than in H graphs.

In order to analyse the asymptotic behaviour of $P_E(t)$, (4.8) we consider the change of variable $\vec{\theta}_t = \vec{\theta}_\infty + \vec{\epsilon}_t$ in the recursion relation Eq. (4.8):

$$\begin{aligned}\vec{\theta}_\infty + \vec{\epsilon}_{t+1} &= \sigma\vec{1} + \lambda(B\cdot(\vec{\theta}_\infty + \vec{\epsilon}_t)) \odot (B\cdot(\vec{\theta}_\infty + \vec{\epsilon}_t)) = \\ &= \sigma\vec{1} + \lambda\left(B\cdot\vec{\theta}_\infty \odot B\cdot\vec{\theta}_\infty + 2B\cdot\vec{\epsilon}_t \odot B\cdot\vec{\theta}_\infty + B\cdot\vec{\epsilon}_t \odot B\cdot\vec{\epsilon}_t\right).\end{aligned} \tag{4.14}$$

where $\vec{\theta}_\infty$ satisfies (4.11), which implies that $\vec{\epsilon}_t$ satisfies the following recursion relation:

$$\vec{\epsilon}_{t+1} = 2\lambda(B\cdot\vec{\epsilon}_t \odot B\cdot\vec{\theta}_\infty + B\cdot\vec{\epsilon}_t \odot B\cdot\vec{\epsilon}_t) \tag{4.15}$$

Moreover, asymptotically, when $t \to \infty$, $(\vec{\epsilon}_t)_i \ll 1$, so that one can linearise Eq. (4.15)

$$\vec{\epsilon}_{t+1} = 2\lambda(B\cdot\vec{\epsilon}_t \odot B\cdot\vec{\theta}_\infty). \tag{4.16}$$

Recalling that $B = \mu D^{-1}A + (1-\mu)\text{Id}$, ff $\mu \ll 1$ then $B = \text{Id}$ and, consequently, $\vec{\theta}_\infty = \vec{\theta}_0$. This implies that the asymptotic behaviour of $(\epsilon_t)_i$ is determined by the recursion relation $(\epsilon_{t+1})_i = 2\lambda(\epsilon_t)_i$, i.e.

$$(\epsilon_t)_i \approx (2\lambda)^t \tag{4.17}$$

Eq. (4.17) allows us to determine the asymptotic behaviour of $P_E(t)$ for $\mu \ll 1$. Since $P_E(t) = (\vec{\theta}_{t-1} - \vec{\theta}_t)_i = (\vec{\epsilon}_{t-1} - \vec{\epsilon}_t)_i \approx (2\lambda)^t((2\lambda)^{-1} - 1)$. This result shows that, for negligible small values of the mutation rate, the escape probability decays exponentially as $P_E(t) \approx e^{\log(2\lambda)t}$ with $2\lambda < 1$, independently of network topology.

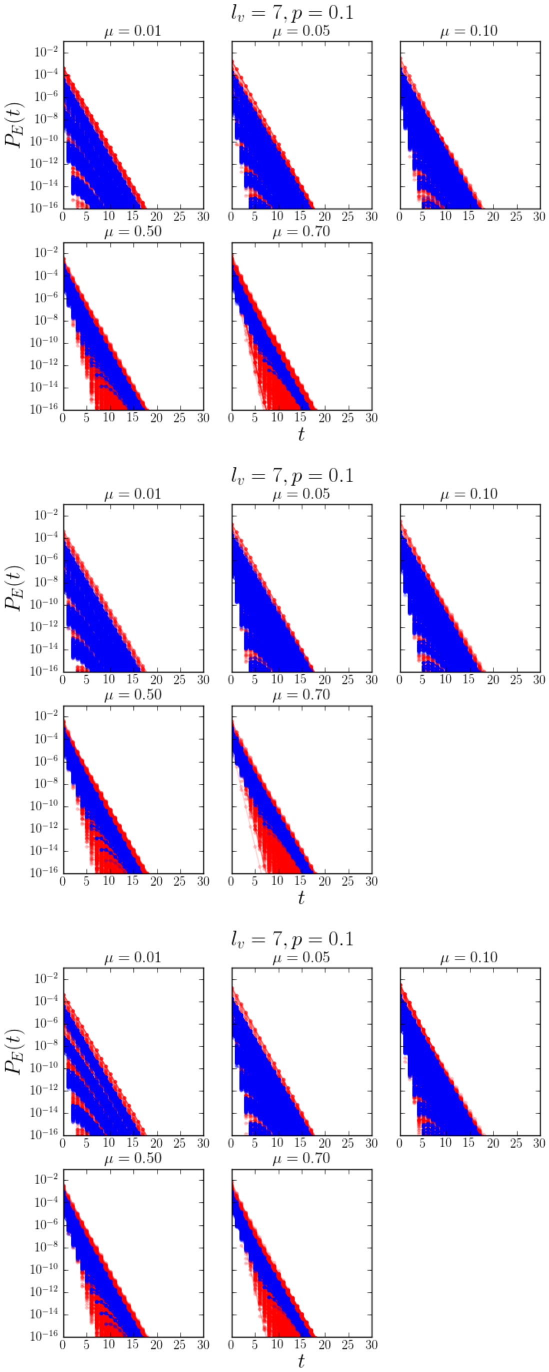


Figure 4.5: Series of plots showing the escape probability at time exactly t. $P_E(t)$ for different genotype-phenotype maps (correspond to graphs in Fig 4.2 (a), (b), (c), $N\nu = 15$) with different values of the mutation rate, μ. Red corresponds to B graphs and blue to H graphs. For each B graph, we have set an escape phenotype and calculate the probability of escape for all possible initial conditions. We have repeated this computation for ten randomly chosen escape phenotypes for each graph. Parameters: $\lambda = 0.1, \sigma = 0.9$

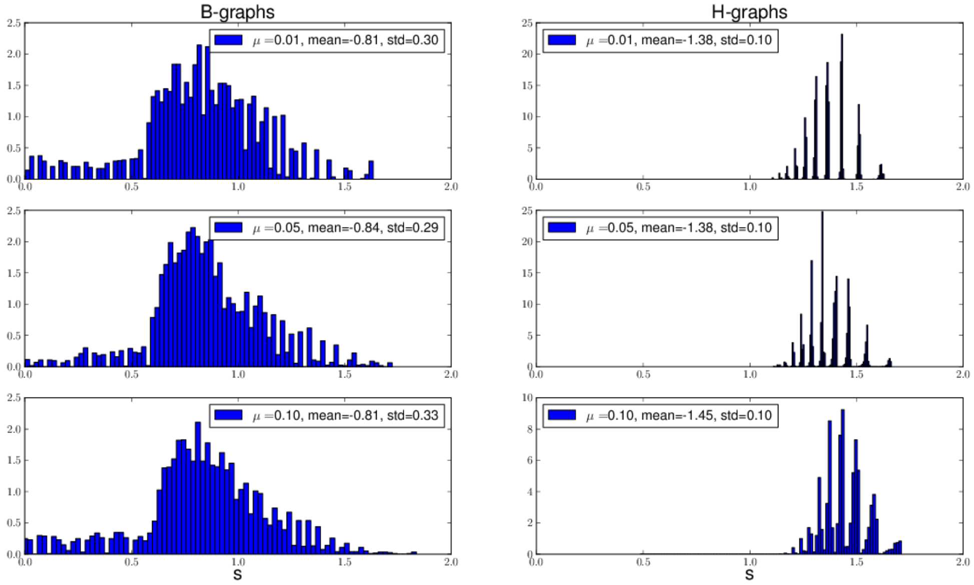


Figure 4.6: Series of plots showing the rate, s, associated to $P_E(t) \sim e^{-st}$, for different genotype-phenotype graphs of type B (left column, graphs correspond to Fig 4.2 (a), (b), (c) with $N\nu = 15$) and type H (right column). The value of s for each graph is obtained by fitting the data shown in Fig. 4.5.

Fig. 4.5 shows simulation results for different B and H genotype-phenotype networks and for different values of the mutation rate, μ. These numerical results show that, even for non-negligible values of the mutation rate μ, the asymptotic behaviour of $P_E(t)$ is exponential, i.e. $P_E(t) \approx e^{-st}$ for large t, although the rate associated to the exponential distribution is indeed dependent on the mutation rate, so that, in general, $s \neq -\log(2\lambda)$ $(2\lambda < 1)$.

We further observe a great deal of heterogeneity in the behaviour of $P_E(t)$ associated to B genotype-phenotype networks (see Figs. 4.5 and 4.6). Fig. 4.6 shows how the histograms of the distribution of rates s change as the parameter values used to generate B-type genotype-phenotype graphs are varied. These results show that, regardless of the parameter values, the dispersion observed in the rate s associated to B networks is robustly larger than for their H-counterparts.

Further information regarding the difference between evolutionary escape on B and H genotype-phenotype graphs can be obtained by recalling that, as the escape time is exponentially distributed, $P_E(t) \approx e^{-st}$, s is equal to the inverse of the average escape time, τ_E: $s = \tau_E^{-1}$. The distributions of s associated to B- and H-type graphs, Fig. 4.6, show that the values of s corresponding to escape on B genotype-phenotype graphs are shifted to the left with respect to their counterparts on H-graphs, which implies that the average escape time on B-type graphs is bigger than that on H-type graphs.

4.3.4 Escape dynamics: escape time probability conditioned to escape

We now proceed to analyse the escape time probability conditioned to eventual escape, $P_E(t|E) \equiv P(N_E(t) > 0 \wedge N_E(t-1) = 0 \mid N_E(\infty) > 0, N_i(0) = 1, N_j(0) = 0 \quad \forall j \neq i)$. The rationale for looking at the properties of this particular function follows from our results regarding the connectivity structure of B-type graphs in which we have seen that phenotypes may be disconnected, and therefore the probability of reaching certain escape phenotypes within B-type graphs may be zero. By conditioning to eventual escape, we discard this cases and focus on the escape process within connected phenotype components.

Comparative analysis between the behaviour of the conditioned escape time probability, $P_E(t|E)$, on B- and H-type graphs shows that there are striking quantitative differences between them. Whereas $P_E(t|E)$ on H-type graphs is sharply concentrated around a well-defined average escape time, $P_E(t|E)$ on B-type graphs is virtually flat, very close to a uniform distribution, where the average escape time has barely any representative.

This behaviour is a direct consequence of the distribution of average distances between phenotypes shown in Fig. 4.3 associated to both types of genotype-phenotype graph: whilst H-graphs exhibit distributions of average distances between phenotypes which are sharply concentrated around the average value, distance distributions for B-graphs are much wider with an average value that is rather un-representative of the behaviour of the ensemble of realisations of the escape process. These properties have a direct effect on the escape time subject to eventual escape.

4.4 Discussion & conclusions

In this Chapter we have studied the problem of evolutionary escape from a novel perspective, namely, since selective pressures act on phenotypes rather than genotypes and therefore fitness is determined by the former, escape problems should be analysed within the context of complex genotype-phenotype networks rather than on regular, hypercube genotype lattices where fitness is directly determined by the genotype.

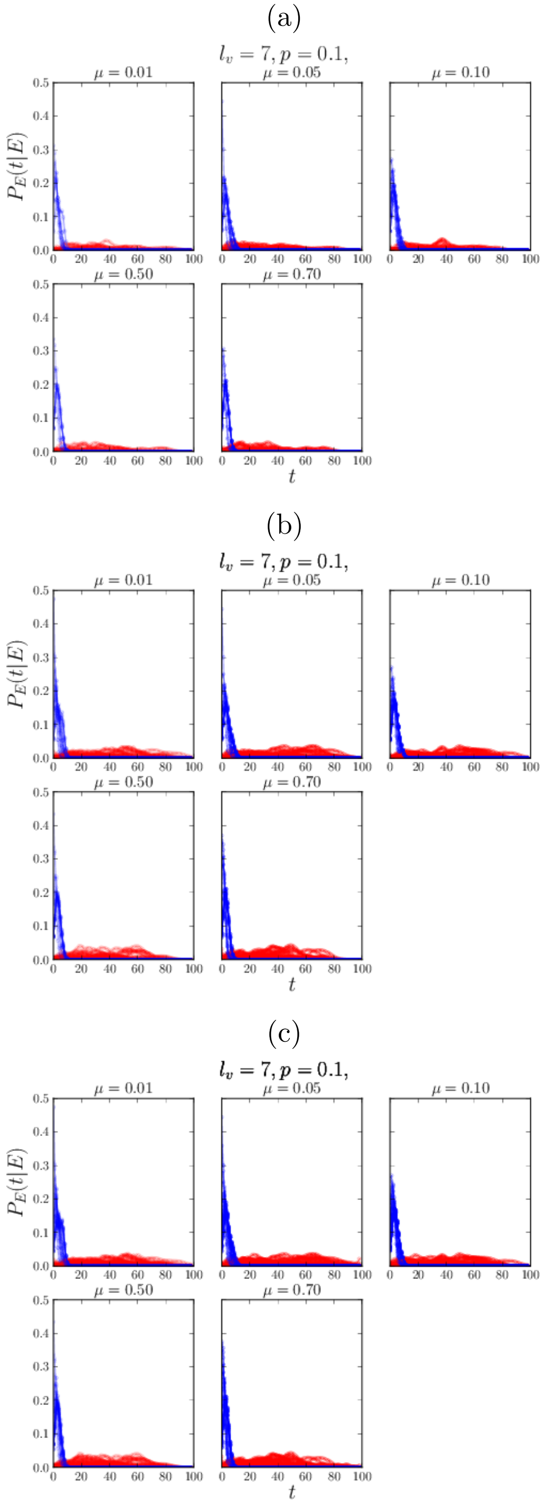


Figure 4.7: Series of plots comparing the behaviour of the conditioned escape time probability, $P_E(t|E)$, on B- (red lines) and H-type (blue lines) genotype-phenotype maps. The B-graphs considered in plots (a), (b) and (c) are the same as those whose connectivity is shown in Fig 4.2(a), (b), and (c), respectively and $N\nu = 15$. By studying how $P_E(t|E)$ changes as the mutation rate, μ, varies, we observe that $P_E(t|E)$ is rather robust to changes in these parameter.

We have carried out a comparative analysis, we have considered two types of genotype-phenotype networks. We have considered complex genotype-phenotype networks, our so-called B-type graphs, which are generated according to a multi-scale model proposed in [52] and described in Chapter 2. Associated to each B-graph, we generate a H-type genotype-phenotype graph, which, according to the procedure explained in Section 4.2, is computed by randomly linking genotype-phenotype pairs under the constraint that both B and H graphs have the same phenotype degree distribution. We have further formulated a population dynamics model, consisting of a multi-type branching process [57], where types are associated to genotypes and their proliferation probability is assigned according to the corresponding phenotype, as determined by the either B or H genotype-phenotype map.

Our comparative analysis sheds some light on the differences between the escape process on either type of genotype-phenotype network, regarding both dynamical and steady-state properties. We have started our analysis by studying the average distance between phenotypes, as this property directly affects the escape process (see Section 4.3.1). Our results (Fig. 4.3) show that, whereas the distribution of distance between phenotypes in H-type genotype-phenotype graphs is sharply peaked around its mean value, its counterpart for B-graphs exhibits a much larger degree of dispersion so that the average distance is hardly representative of individual behaviour within the statistical ensemble.

An alternative to the approach used in Section 4.3.1, which is based on computing the average shortest path between phenotypes, consists of resorting to spectral graph theory [26]. In fact, Chung [25] has proven that the diameter of a graph, i.e. the maximum of the distances among all possible pairs of vertices, is small (i.e. $O(\log n)$ where n is the number of nodes) if the modulus of the second eigenvalue of the adjacency matrix is small compared to the first eigenvalue. However, in our case, this method does not allow us to discriminate between B and H graphs, nor to identify heterogeneities among B-graphs, as per the examples shown in Fig. 4.2. This is a direct consequence of the fact that genotype subgraph for both B and H genotype-phenotype maps are bipartite graphs.

Consider an H graph whose genotype subgraph, G_H, is an n-cube, i.e. a 2^n-vertices hypercube where one can assign to each vertex a string of length n from $(-1, -1, \ldots, -1)$ to $(1, 1, \ldots, 1)$. For example, a 2-cube (or square) would be:

$$00, 01, 10, 11.$$

One can now define two disjoint subsets within the graph $G = (X, Y)$. Where X ={vertices who have an even number of 1's} and Y ={vertices who have an odd number of 1's}. Since edges connect nodes that differ by one entry in the label string, we only have edges between X and Y, thus constructing a colouring with only two colours and, therefore, the n-cube graph is bipartite. Since the genotype subgraph of a B genotype-phenotype network is a subset of G_H, then it also is a bipartite graph. Bipartite graphs happen to have rather trivial spectral properties [20]: bipartite networks have symmetric spectra and the spectral gap is equal to zero. B and H networks are not exactly bipartite, because of the presence of the phenotype nodes. However, we have checked that these nodes perturb the spectrum only very slightly. Therefore, the spectral properties associated to these two types of networks are not the right framework to analyse the differences between them, in particular, those concerning the evolutionary escape problem.

The heterogeneity observed in the distribution of distances between phenotypes in B-type graphs induces heterogeneous behaviour in all the observables associated to the escape problem that we have investigated. The escape probability in B-type graphs displays a much wider range of variability than in H-type networks, with instances in which the escape probability in the B-graph is an order of magnitude bigger than in its H-type counterpart (see Fig. 4.4). Heterogeneity is also ubiquitous when dynamical properties are examined. Fig. 4.6 shows that the average escape time is much more homogeneously distributed in H-type graphs. Furthermore, as a result of the heterogeneity in the average escape time for B-graphs, we have determined that escape is very likely to occur on considerable longer time scales on B-graphs. When studying the escape time distribution conditioned to eventual escape (see Fig. 4.7), we observe that, unlike H-graphs where escape exhibits a well-defined, characteristic time scale, conditioned escape time on B-graphs shows a much wider (almost uniform) distribution with virtually no discernible characteristic scale.

Therefore, unlike genotype-phenotype graphs where the genotype network is a regular hypercubic lattice, where averaged properties are representative of individual behaviour within the statistical ensemble, in complex genotype-phenotype networks heterogeneity is the rule. This implies that predictions regarding emergence of resistant varieties in, for example, cancer cell populations under treatment [98] based on regular hypercubic genotype spaces may be inaccurate. Rather, a more detailed study of the underlying genotype-phenotype structure is necessary to produce accurate estimates. The application of our methods to the issue of emergence of resistance in tumours under therapy would require to develop a model of the gene regulatory network governing the appearance of pathological cell states (i.e. malignant phenotype) as well as the corresponding epigenetic regulation [82], and then analyse the corresponding genotype-phenotype graph to discern how deregulations at the genetic and epigenetic levels would affect evolutionary escape, whereby targeted strategies aimed at minimising the escape probability could be formulated. The formulation of such models is beyond the scope of this work and left for future research.

Chapter 5

Surviving evolutionary escape on complex genotype-phenotype networks

In this Chapter we study the problem of evolutionary escape and survival for cell populations with genotype-phenotype map. In order to explore these issues, we formulate a population dynamics model, consisting of a multi-type time-continuous branching process, where types are associated to genotypes and their birth and death probabilities depend on the associated phenotype (non-escape or escape). We show that, within the setting associated to the escape problem, separation of time scales naturally arises and two dynamical regimes emerge: a fast-decaying regime associated to the escape process itself, and a slow regime which corresponds to the (survival) dynamics of the population once the escape phenotype has been reached (i.e. conditioned to escape). We exploit this separation of time scales to analyse the topological factors which determine escape and survival. In particular, the aim of this Chapter is to analyse the influence of topological properties associated to robustness and evolvability on the probability of escape and on the probability of survival upon escape. We show that, while the escape probability depends on size of the neutral network of the escape phenotype (i.e. its degree), the probability of survival is essentially determined by its robustness (i.e. the resilience of the escape phenotype against genetic mutations), measured in terms of a weighted clustering coefficient.

5.1 Surviving evolutionary escape

In Chapter 4 and in [50], we have considered a variation of the original framework for studying escape problems which alters the approach proposed by Iwasa and co-workers. We have shown that by incorporating the genotype-phenotype network structure into the study of evolutionary process a large degree of heterogeneity arises both in the probability of eventual escape and the dynamics of escape, which is absent in genotype-phenotype graphs associated to regular hypercube genotype networks [50]. As a consequence, we have shown that the model of escape on genotype-phenotype graphs associated to hypercube genotype networks may lead to underscoring the escape probabilities and that the topology of complex genotype-phenotype graphs slows down the rate of escape [50].

The inclusion of the complex structure of genotype-phenotype network introduces yet another issue that needs to be considered when dealing with the escape problem. Previous studies of this problem [53, 54, 89, 90, 84, 50] assume that once the escape genotype is reached, escape takes place with probability 1, i.e. the proliferation ratio of the escape genotype is $R_E = \infty$. However, one could go one step further and analyse the probability of surviving escape if $1 < R_E < \infty$. This question is particularly meaningful when genotype-phenotype structure is accounted for. On the one hand, due to the multiplicity of genotypes associated to each phenotype, there exists a non-trivial population dynamics even when the population is confined within the escape phenotype. Secondly, cell populations with genotype-phenotype map exhibit evolved properties such as robustness and evolvability which affect escape [50] and survival alike: robustness, being related to resilience of phenotypes against gene mutations, is bound to affect survival, since any mutation driving cells away from the escape phenotype can be considered deleterious. Similarly, the *degree* of the escape phenotype, k_E, which is defined as the number of viable genotypes bearing the escape phenotype [52], is a measure of accessibility to the escape phenotype (roughly speaking, the higher k_E, the more access routes for the mutation-driven random search process which drives escape to reach the target phenotype). The aim of this Chapter is to analyse the influence of this and other topological properties associated to robustness and evolvability [28, 27, 52] on the probability of survival after escape.

Remind the mathematical model in Section 4.2 and that genotype-phenotype networks in this Chapter has been modelled such as in Chapter 2. We also note, that clustering coefficient of phenotype node ϕ, c_ϕ, can be described as

$$c_\phi = \frac{2|\{e_{\mathcal{G},\mathcal{G}'} \mid \mathcal{G}, \mathcal{G}' \in \text{Neigh}_\phi\}|}{k_\phi(k_\phi - 1)} \tag{5.1}$$

where Neigh_ϕ is the neighbourhood for a node ϕ and k_ϕ is the degree of ϕ, i.e. the number of genotype nodes to which ϕ is connected. Since $\text{dist}(\mathcal{G}, \mathcal{G}') = 1$ these two genotypes correspond to nodes linked by a genotype-genotype edge, $e_{\mathcal{G},\mathcal{G}'}$.

5.1.1 Population dynamics

Evolutionary escape has been studied in the context of multi-type Galton-Watson processes [54, 90, 52]. This type of process is characterised by a lack of characteristic time scales, as its dynamic is generated by mere iteration of and individual progeny-generation process. Here, we aim to analyse the emergence of separation of time scales intrinsic to the evolutionary escape process and its associated consequences regarding the behaviour of the system. In order to account for the characteristic time scales, we resort to a description in terms of a continuous-time branching process with exponential life time distributions, which is closely related to the Galton-Watson process [57].

We first define a multi-type birth-and-death process where each type is associated to a genotype, $\mathcal{G}_i$, so that $N_i(t)$ is the number of cells with genotype $i = 1, \ldots, N_G$, where N_G is the number of viable genotypes. Our model of the genotype-phenotype map [52] assigns a phenotype to each genotype, i.e. $\phi_i = \phi(\mathcal{G}_i)$ where ϕ_i is the phenotype associated to genotype $\mathcal{G}_i$. Following to the model of evolutionary escape formulated in [50], we assume that birth and death rates of each cell type depend on the phenotype rather than on the genotype. We thus consider the birth-and-death process defined by:

$$
\begin{aligned}
&\text{Prob}(N_i(t+\Delta t) = N_i(t)+1) = \lambda(\phi_i)(1-\mu)^2\Delta t\\
&\text{Prob}(N_i(t+\Delta t) = N_i(t)-1) = \sigma(\phi_i)\Delta t\\
&\text{Prob}(N_i(t+\Delta t) = N_i(t), N_j(t+\Delta t) = N_j(t)+1) = 2\lambda(\phi_i)\mu(1-\mu)\frac{a_{ij}}{d_i}\Delta t,\\
&\text{Prob}(N_i(t+\Delta t) = N_i(t)-1, N_j(t+\Delta t) = N_j(t)+2) = \lambda(\phi_i)\mu^2\frac{a_{ij}a_{ik}}{d_i^2}\Delta t,
\end{aligned}
\tag{5.2}
$$

for all $j, k = 1, \ldots, N_G$, where a_{ij} are the entries of the adjacency matrix of the genotype network, d_i is the degree of $\mathcal{G}_i$ within the genotype network, $\lambda(\phi_i)$ and $\sigma(\phi_i)$ are the phenotype-dependent birth and death rates associated to $\mathcal{G}_i$, and μ is the mutation probability per division. For simplicity, we consider that $\lambda(\phi_i) = \lambda$ and $\sigma(\phi_i) = \sigma$ for all $\phi_i \neq \phi_E$ with $\lambda < \sigma$, and $\lambda(\phi_i) = \lambda_E$ and $\sigma(\phi_i) = \sigma_E$ for all $\phi_i = \phi_E$ with $\lambda_E > \sigma_E$. We can assume that $\lambda = \lambda_E$ and $\sigma \gg \sigma_E$.

We now define an associated embedded, continuous-time branching process [44]. This branching process is characterised by the set of type-specific generating function of the probability density of the number of progeny produced by each cell:
If $\phi_i \neq \phi_E$,

$$
F_i(\vec{s}) = \frac{\sigma}{\lambda+\sigma} + \frac{\lambda(1-\mu)^2}{\lambda+\sigma}s_i^2 + \sum_j 2\frac{\lambda\mu(1-\mu)}{\lambda+\sigma}\frac{a_{ij}}{d_i}s_is_j + \sum_{j,k}\frac{\lambda\mu^2}{\lambda+\sigma}\frac{a_{ij}a_{ik}}{d_i^2}s_js_k,
\tag{5.3}
$$

If $\phi_i = \phi_E$,

$$
\begin{aligned}
F_i(\vec{s}) =& \frac{\sigma_E}{\lambda_E+\sigma_E} + \frac{\lambda_E(1-\mu)^2}{\lambda_E+\sigma_E}s_i^2 + \sum_j 2\frac{\lambda_E\mu(1-\mu)}{\lambda_E+\sigma_E}\frac{a_{ij}}{d_i}s_is_j +\\
&+ \sum_{j,k}\frac{\lambda_E\mu^2}{\lambda_E+\sigma_E}\frac{a_{ij}a_{ik}}{d_i^2}s_js_k
\end{aligned}
\tag{5.4}
$$

where $F_i(\vec{s})$ is the probability generating function of the number of progeny generated by a cell with genotype $\mathcal{G}_i$ and $\vec{s} = (s_1, \ldots, s_G)$. Furthermore each genotype has an associated survival time which is exponentially distributed with parameter $\omega = \lambda + \sigma$ if $\phi_i \neq \phi_E$ and $\omega_E = \lambda_E + \sigma_E$ if $\phi_i = \phi_E$. Note that $\omega \gg \omega_E$.

Similarly to the Galton-Watson process, the dynamics of the continuous-time branching process is generated by iterating the progeny-generation process, which mathematically translates into the following recursive equation for the generating function, $f_i(\vec{s}, t)$ [57]:

$$
f_i(\vec{s}, t+\Delta t) = f_i\left(\vec{f}(\vec{s}, t), \Delta t\right),
\tag{5.5}
$$

where $f_i(\vec{s}, t)$ is the generating function associated to the probability distribution of the population $(N_1(t), \ldots, N_G(t))$ at time t generated from a single initial individual of (geno)type i. If Δt is small, then with probability close to one, the process consists of either the mother cell, provided that it survives, or its first-generation progeny, conditioned to the mother cell exhausts its life-span, i.e.

$$
\begin{aligned}
&f_i(\vec{s}, \Delta t) = s_ie^{-\omega\Delta t} + F_i(\vec{s})(1-e^{-\omega\Delta t}) + o(\Delta t^2) \text{ if } \phi_i \neq \phi_E,\\
&f_i(\vec{s}, \Delta t) = s_ie^{-\omega_E\Delta t} + F_i(\vec{s})(1-e^{-\omega_E\Delta t}) + o(\Delta t^2) \text{ if } \phi_i = \phi_E.
\end{aligned}
\tag{5.6}
$$

5.1.1.1 Separation of time scales: quasi-steady state regime

The setting in which evolutionary escape occurs involves separation of time scales. Within this setting, all genotypes, $\mathcal{G}_i$, associated to phenotypes $\phi_i \neq \phi_E$ are sub-critical ($\sigma > \lambda$), whereas those genotypes associated to the escape phenotype ($\phi_i = \phi_E$) are super-critical, i.e. $\sigma_E < \lambda_E$. If we can assume that, for example, $\lambda = \lambda_E$ and $\sigma \gg \sigma_E$, which implies $\omega \gg \omega_E$, separation of time scales ensues. Consider Eqs. (5.6) and the following re-scaled time variable $\tau = \omega_E t$. Carrying out this change of variables, Eqs. (5.6) read:

$$\begin{aligned} f_i(\vec{s}, \Delta\tau) &= s_i e^{-\frac{1}{\epsilon}\Delta\tau} + F_i(\vec{s})(1 - e^{-\frac{1}{\epsilon}\Delta\tau}) + o(\Delta\tau^2) \text{ if } \phi_i \neq \phi_E, \\ f_i(\vec{s}, \Delta\tau) &= s_i e^{-\Delta\tau} + F_i(\vec{s})(1 - e^{-\Delta\tau}) + o(\Delta\tau^2) \text{ if } \phi_i = \phi_E. \end{aligned} \tag{5.7}$$

where $\epsilon = \frac{\omega_E}{\omega} \ll 1$. This implies that cell types associated to non-escape genotypes evolve much faster than those associated to escape genotypes. The former will therefore settle onto their equilibrium state, i.e. extinction, whilst the latter continue to evolve at a much slower rate. This becomes clearer if we write down the system of ODEs which govern the evolution of $f_i(\vec{s}, \tau)$ (see [57] for details of their derivation):

$$\begin{aligned} \epsilon \frac{df_i(\vec{s}, \tau)}{d\tau} &= -f_i(\vec{s}, \tau) + F_i(f_1(\vec{s}, \tau), \ldots, f_G(\vec{s}, \tau)) \text{ if } \phi_i \neq \phi_E \\ \frac{df_i(\vec{s}, \tau)}{d\tau} &= -f_i(\vec{s}, \tau) + F_i(f_1(\vec{s}, \tau), \ldots, f_G(\vec{s}, \tau)) \text{ if } \phi_i = \phi_E, \end{aligned} \tag{5.8}$$

which explicitly shows that, under time re-scaling ($\tau = \omega_E t$), the population of non-escape genotypes evolves according to a fast dynamic and can be considered to be in (quasi-)steady state. The populations associated to escape genotypes follow a slow dynamic which evolves on a much slower time scale.

5.1.1.2 Separation of time scales: inner regime

Besides the long time dynamics associated to the quasi-steady state regime analysed in Section 5.1.1.1, we can study the initial, inner regime by re-scaling time: $T = \epsilon^{-1}\tau$. Under this re-scaling, Eqs. (5.6) read:

$$\begin{aligned} f_i(\vec{s}, \Delta T) &= s_i e^{-\Delta T} + F_i(\vec{s})(1 - e^{-\Delta T}) + o(\Delta T^2) \text{ if } \phi_i \neq \phi_E, \\ f_i(\vec{s}, \Delta T) &= s_i e^{-\epsilon\Delta T} + F_i(\vec{s})(1 - e^{-\epsilon\Delta T}) + o(\Delta T^2) \text{ if } \phi_i = \phi_E. \end{aligned} \tag{5.9}$$

Eqs. (5.9) imply that, since $\epsilon \ll 1$, the rate of evolution of the populations associated to escape genotypes ($\phi_i = \phi_E$) is very slow, so that the most likely event is that cells with escape genotypes do not generate progeny, i.e., during this initial regime, cells associated to escape genotypes tend to stay latent (that is, they survive without producing offspring or dying). By contrast, the populations associated to non-escape genotypes ($\phi_i \neq \phi_E$) evolve at an $O(1)$ rate. The corresponding set of ODEs for the probability generating functions is:

$$\begin{aligned} \frac{df_i(\vec{s}, T)}{dT} &= -f_i(\vec{s}, T) + F_i(f_1(\vec{s}, T), \ldots, f_G(\vec{s}, T)) \text{ if } \phi_i \neq \phi_E \\ \frac{df_i(\vec{s}, T)}{dT} &= -\epsilon\left(f_i(\vec{s}, T) - F_i(f_1(\vec{s}, T), \ldots, f_G(\vec{s}, T))\right) \text{ if } \phi_i = \phi_E, \end{aligned} \tag{5.10}$$

which further confirm that, during this initial regime, the populations associated to escape genotypes stays frozen, whereas the non-escape populations evolves at a rate $O(1)$.

These properties regarding time scale separation have the consequence that, during this initial regime, population accumulates within the escape genotypes. The source of this population is mutations occurring in cells associated with non-escape genotypes but which are first neighbours of escape genotypes. According to our analysis, these cells remain latent during the initial regime. Therefore the total number of cells accumulated within the escape phenotype during the initial regime ($T = 1$), N_0 (latter re-called $P_E(T)$), can be estimated by [42, 9]:

$$N_0 = Y\left(\frac{2\lambda}{\omega}\sum_{i\in\partial\phi_E}\int_0^1\left(\mu(1-\mu)\sum_{j\in\langle\phi_E\rangle}\frac{a_{ij}}{d_i}+\mu^2\sum_{j,k\in\langle\phi_E\rangle}\frac{a_{ij}a_{ik}}{d_i^2}\right)N_i(T)dT\right) \tag{5.11}$$

$Y(s)$ is a Poisson-distributed random number with parameter s and $\partial\phi_E$ is the sub-set of non-escape genotypes of the genotype networks which are first-neighbours with an escape genotype: $\mathcal{G}_i \in \partial\phi_E$ if $\phi_i \neq \phi_E$ and there exists at least one $\mathcal{G}_j$ such that $\phi_j = \phi_E$ and $a_{ij} = 1$. Note that the cardinal of $\partial\phi_E$ is proportional to the degree of the escape phenotype, k_E, and therefore, from Eq. (5.11), we expect N_0 be an increasing function of k_E, too: the bigger k_E, the bigger the boundary between the escape phenotype and the remaining of the network, and the more access ways for population to enter the escape phenotype.

5.1.1.3 Coarse-grained population dynamics of the escape phenotype

In view of the picture arising from the discussion of Sections 5.1.1.1 and 5.1.1.2 regarding separation of time scales, we propose a coarse-grained population dynamics of the escape phenotype, associated to the quasi-steady state regime in which the non-escape genotypes have already become extinct, and where we look at the total population of the escape phenotype, rather than looking at the populations associated to the escape genotypes. In order to formulate this coarse-grained dynamics, we invoke some of the topological properties analysed in [52].

We have shown that, within the genotype-phenotype network defined in [52], we can associate a clustering coefficient to each phenotype, where the clustering is associated to the proportion of genotypes which are first neighbours (i.e., according to the definition of the genotype network in [52], genotypes that are separated by a one-hit mutation) and which share the same phenotype. If we assume that the clustering coefficient of the escape phenotype, c_E, models the probability that a gene mutation which changes $\mathcal{G}_i$ into $\mathcal{G}_j$ without changing phenotype, i.e. $\phi_i = \phi_j = \phi_E$: the probability of a gene mutation to produce a change of phenotype is equal to $\mu(1-c_E)$. Similarly, the probability of a phenotype-preserving mutation is μc_E. Furthermore, we have shown that $c_E(k_E) \approx (k_E-1)^{-\alpha}$ with $\alpha \approx 1$ [52].

By taking into account this interpretation of the clustering coefficient, we can define the following coarse-grained continuous-time branching process:

$$\begin{array}{lll} F_E(s) = \frac{1}{\omega_E} & \big(\sigma_E + \lambda_E\mu^2(1-c_E)^2 + & \text{Death} \\ & +\lambda_E(2\mu(1-c_E)(1-\mu)+2\mu^2c(1-c))s+ & \text{Survival} \\ & +\lambda_E((1-\mu)^2+2\mu(1-\mu)c_E+\mu^2c_E^2)s^2\big) & \text{Proliferation} \end{array} \tag{5.12}$$

with a life-time which is exponentially distributed with parameter ω_E. We have further assume that cells which mutate an change phenotype die, as they fall back into a non-escape phenotype which is sub-critical.

5.1.1.4 Alternative measure of robustness

Results (see Section 5.2) show that, although c_E is a measure of robustness which subtly correlates with P_S (see Fig 5.6), a measure of robustness better suited for the coarse-grained, long-time dynamics of survival upon escape, must be defined. To do so, we proceed with a more detailed analysis of Eqs (5.8). Before proceeding we introduce the following notation: $\vec{1} = (1, \dots, 1)$, $\vec{1}_E$ is the vector with ones in the escape components and zeros otherwise, and $\vec{1}_N = \vec{1} - \vec{1}_E$. We define $\vec{g}(\tau) = (\vec{g}_N, \vec{g}_E)$ as the split vector, first components are associated to non escape genotypes, $\vec{g}_N$, and $\vec{g}_E$ corresponds to indexes associated to escape genotypes. Making some algebra we can rewrite $\vec{F}(\vec{s})$ as,

$$\vec{F}(\vec{s}) = \vec{1} + D(-\vec{1} + B \cdot \vec{s} \odot B \cdot \vec{s}) \tag{5.13}$$

where $B = \mu D^{-1} A + (1-\mu)\mathrm{Id}$, $D = \mathrm{diag}(\gamma_i)$ ($\gamma_i = \frac{\lambda}{\lambda+\sigma}$ if $\phi_i \neq \phi_E, \gamma_i = \frac{\lambda_E}{\lambda_E+\sigma_E}$ if $\phi = \phi_E$) and $\odot$ denotes the component-to-component product.

As we have shown in a previous work [50], if we define $g_i(\tau) = P(N_E(\tau) > 0 \mid \mathrm{IC}_i)$, then $\vec{f}(\vec{1}_N, \tau) = \vec{1} - \vec{g}(\tau)$ is the complementary of the escape probability. Therefore, $g_i(\tau) \ll g_j(\tau)$ if $i \notin \phi_E, j \in \phi_E$, where N_E is the number of cells in the escape phenotype. Then, Eqs (5.8) read,

$$\epsilon \frac{d\vec{g}_N(\tau)}{d\tau} = -\vec{g}_N(\tau) + \frac{\lambda}{\lambda+\sigma}(2B \cdot \vec{g}_N(\tau) - (B \cdot \vec{g}_N(\tau) \odot B \cdot \vec{g}_N(\tau))) \tag{5.14}$$

$$\frac{d\vec{g}_E(\tau)}{d\tau} = -\vec{g}_E(\tau) + \frac{\lambda_E}{\lambda_E+\sigma_E}(2B \cdot \vec{g}_E(\tau) - (B \cdot \vec{g}_E(\tau) \odot B \cdot \vec{g}_E(\tau))) \tag{5.15}$$

We split B into four sub-matrices:

$$B = \begin{pmatrix} B_{NN} & B_{NE} \\ B_{EN} & B_{EE} \end{pmatrix}$$

where B_{NN} corresponds to sub-matrix of non escape indexes, B_{EE} is the sub-matrix of escape indexes and similarly for B_{EN} and B_{NE}. In terms of these sub-matrices, Eq (5.15) can be rewritten as,

$$\begin{aligned} \frac{d\vec{g}_E}{d\tau} = -\vec{g}_E + \frac{\lambda_E}{\lambda_E+\sigma_E}[2(B_{EE} \cdot \vec{g}_E + B_{EN} \cdot \vec{g}_N) \\ - (B_{EE} \cdot \vec{g}_E + B_{EN} \cdot \vec{g}_N) \odot (B_{EE} \cdot \vec{g}_E + B_{EN} \cdot \vec{g}_N)] \end{aligned} \tag{5.16}$$

As $(\vec{g}_N)_i \ll (\vec{g}_E)_j$, we can neglect $\vec{g}_N$. Eq (5.16) then reads,

$$\frac{d\vec{g}_E}{d\tau} = -\vec{g}_E + \frac{\lambda_E}{\lambda_E+\sigma_E}[2B_{EE} \cdot \vec{g}_E - (B_{EE} \cdot \vec{g}_E \odot B_{EE} \cdot \vec{g}_E)] \tag{5.17}$$

In order to study the behaviour of $\vec{g}_E$, we first consider the simplest possible case. We assume $|\{\phi(\mathcal{G}_i) = \phi_E\}| = 1$, i.e. there is only one genotype $\mathcal{G}_i$ in ϕ_E. This allows us to write Eq (5.17) as,

$$\frac{dg_E}{d\tau} = -g_E + \frac{\lambda_E}{\lambda_E+\sigma_E}\left[2(1-\mu)g_E - (1-\mu)^2 g_E^2\right] \tag{5.18}$$

At steady state Eq (5.18) is,

$$\begin{aligned} 0 = &- g_E + \frac{\lambda_E}{\lambda_E + \sigma_E}\left[2(1-\mu)g_E - (1-\mu)^2 g_E^2\right] = \\ &= g_E\left(1 - \frac{\lambda_E}{\lambda_E + \sigma_E}(2-2\mu) + \frac{\lambda_E}{\lambda_E + \sigma_E}(1-\mu)^2 g_E\right) \end{aligned} \tag{5.19}$$

with non-trivial solution

$$g_E = \frac{2}{1-\mu} - \frac{\lambda_E + \sigma_E}{\lambda_E(1-\mu)^2}.$$

Now, consider the case of $|\{\phi(\mathcal{G}_i) = \phi_E\}| = 2$. This implies that $B = B_{EE}$,

$$B_{EE} = \begin{pmatrix} 1-\mu & a_{12}/d_1 \\ a_{21}/d_2 & 1-\mu \end{pmatrix}$$

Then,

$$\text{if } = \begin{cases} a_{12} = 0(\implies a_{21} = 0) & \implies \text{we are in case } |\{\phi(\mathcal{G}_i) = \phi_E\}| = 1, \\ & \text{as both nodes are disconnected,} \\ a_{12} = 1(\implies a_{21} = 1) & \implies \vec{g}_E \text{ will depend on } d_1 \text{ and } d_2. \end{cases}$$

where, d_i is the degree of genotype $\mathcal{G}_i$ in the genotype network. This simple case shed some light on what kind of parameters affect the value of $\vec{g}_E$. This simple case illustrates that the clustering coefficient does not characterise the evolution of $\vec{g}_E$, as we need to take the degree of the first neighbours into account.

Therefore, we define the following measure of robustness, M_E, of an escape phenotype ϕ_E:

1. Generate genotype subgraph, G_{ϕ_E}, of nodes corresponding to genotypes $\mathcal{G}_i$, such that, $\phi(\mathcal{G}_i) = \phi_E$.

2. Associate weights, w_{ij} for each (i, j) edge in G_{ϕ_E}.

3. Define $w_{ij} = \frac{1}{d_i} + \frac{1}{d_j}$.

4. Then,

$$M_E = \frac{\sum_{(ij)} w_{ij}}{|\{\phi(\mathcal{G}_i) = \phi_E\}|}, \quad M_E \in [0, 1].$$

Note that $M_E = 0$ is associated to escape phenotypes, ϕ_E, for which their escape associated genotypes are not interconnected, i.e. $M_E = 0 \implies c_E = 0$. At the other extreme, $M_E = 1$ corresponds to an isolated escape phenotype where all possible connections between escape genotypes exist. In this case, $c_E = 1$ (see Fig 5.1 and Table 5.1). In general $c_E = 1$ does not imply $M_E = 1$.

An intuitive handle on the meaning of M_E can be obtained by comparing the examples shown in Fig 5.2. Both these graphs have $c_E = 1/3$. However, M_E takes different values and, therefore, it allows to distinguish between both cases. Note that survival is less likely in graph (b) than in graph (a). From these examples, we observe that M_E provides more information than c_E because M_E accounts for inhomogeneities between nodes. In other words, M_E carries more local information regarding the genotypes than c_E.

A definition for the global weighted clustering coefficient M_E, can be thought as a local clustering coefficient for a weighted network. Given an undirected, unweighted network,

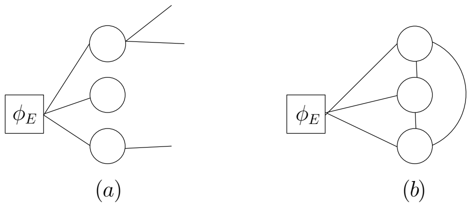


Figure 5.1: Two different graphs with extreme values of M_E. Graph (a) has $M_E = 0$, whereas graph (b) has $M_E = 1$. More details in Table 5.1.

Table 5.1: Comparison of graphs of Fig 5.1.

Parameters	(a) graph	(b) graph
k_E	3	3
c_E	0	1
G_{ϕ_E}	$\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}$	$\begin{pmatrix}0&1&1\\1&0&1\\1&1&0\end{pmatrix}$
$G_{w\phi_E}$	$\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}$	$\begin{pmatrix}0&1/2&1/2\\1/2&0&1/2\\1/2&1/2&0\end{pmatrix}$
M_E	0	1

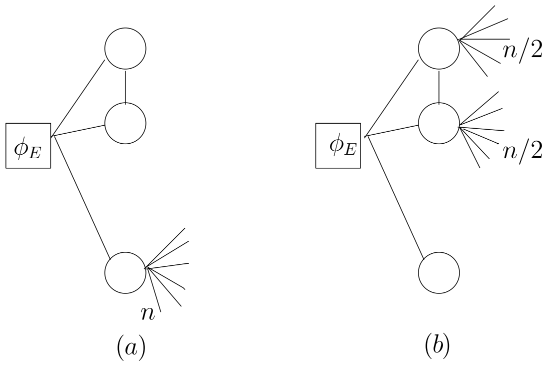


Figure 5.2: Two different graphs with very different M_E. Graph (a) has $M_E = 1/3$, whereas graph (b) has $M_E = 4/3(n + 2)$. More details in Table 5.2.

Table 5.2: Comparison parameters graphs of Fig 5.2.

Parameters	(a) graph	(b) graph
k_E	3	3
c_E	1/3	1/3
k_{nn} (second neighbours)	n	n
G_{ϕ_E}	$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$
$G_{w\phi_E}$	$\begin{pmatrix} 0 & 1/2 & 0 \\ 1/2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1/(n/2+1) & 0 \\ 1/(n/2+1) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$
M_E	$\frac{1}{3}$	$\frac{4}{3(n+2)}$

clustering coefficient of a node i, C_i is equal to the fraction of number of edges between neighbourhood of i. Then, given a undirected weighted network, we define C_i as the fraction between the sum of weights in edges between i-neighbourhood and the maximum value that can weights achieve, that is, k_i. On the other hand, these weights w_{ij} represent the sum of the probability of going from i to j and the one of going from j to i. This is equivalent to M_E which is normalised by k_E.

5.1.1.5 Redefinition of the escape phenotype coarse-grained dynamics

Using M_E as a measure of robustness, the coarse-grained dynamics of an escape phenotype is defined by the pgf:

$$\begin{array}{lll} F_E(s) = \frac{1}{\omega_E} & \big(\sigma_E + \lambda_E \mu^2 (1-M_E)^2 + & \text{Death} \\ & +\lambda_E(2\mu(1-M_E)(1-\mu) + 2\mu^2 M_E(1-M_E))s + & \text{Survival} \\ & +\lambda_E((1-\mu)^2 + 2\mu(1-\mu)c_E + \mu^2 M_E^2)s^2\big) & \text{Proliferation} \end{array} \tag{5.20}$$

with a life-time which is exponentially distributed with parameter ω_E (see Section 5.1.1.3).

5.2 Results

Having established the separation of time scales in Section 5.1.1, where an inner, faster regime associated to the escape process is followed by an outer, slower dynamics related to the within-escape phenotype dynamics, we proceed to analyse in detail the eventual survival of the population upon escape (i.e. upon the population having reached the escape phenotype). We will first study the escape process (as determined by the dynamics within the inner regime), in particular, we focus on how the escape probability and the number of cells reaching the escape phenotype depend on the degree of the escape phenotype, k_E. We then proceed to study the eventual survival conditioned to escape, which corresponds to the outer regime and it is analysed using a coarse-grained model (see Section 5.1.1.3 and 5.1.1.5).

5.2.1 Fast dynamics: Escape properties

In this Section we study escape properties during the inner regime. We show how escape probabilities can be computed and how to compute the average expected number of cells accumulated during the fast dynamics. All of these quantities are related with the degree and the clustering coefficient of the escape phenotype, we introduce in what way these correlations appears.

5.2.1.1 Escape probability

Given an escape phenotype ϕ_E we define the initial condition IC_i as $N_i(0) = 1$ and $N_j(0) = 0$ for all $j \neq i$ and we assume $\phi(\mathcal{G}_i) \neq \phi_E$. The probability of escape $P_E(T)$ is the probability to achieve the escape phenotype before extinction starting from the configuration IC_i and averaged for all possible initial conditions. This process happens in the inner regime. This probability can be computed in a similar way than it was done for the discrete model in [50] (see results in Appendix 7, Section 7.5). Briefly, it can be checked that $P_E(T)$ coincides with the evaluation of the pgf at the parameter $\vec{s} = \vec{1}_N$. As the pgf evolves according to an ODE (Eq (5.10)), we can integrate numerically it to obtain the desired values $P_E(T)$. Moreover, as we are interested in the inner regime we only compute $P_E(T)$ up to time $T = 1$.

Fig 5.3 shows results regarding how $P_E(T)$ changes as we vary the degree k_E (Fig 5.3 (a), (c)) and the clustering coefficient ϕ_E (Fig 5.3 (b)). These results agree with results in discrete branching process model described in Chapter 4 (see Appendix 7 Section 7.5). We observe that $P_E(T)$ is positively correlated with k_E, i.e. the larger the number of genotypes which belong the escape phenotype, the bigger (on average) is the probability of achieve the escape phenotype during the inner regime. Otherwise, the negative correlation between $P_E(T)$ and c_E is a direct consequence of $c_E(k_E) \approx (k_E - 1)^{-\alpha}$ [52].

5.2.1.2 Average number of cells accumulated within the escape phenotype during the initial regime

In order to calculate the average number of cells that reach the escape phenotype during the fast dynamics regime, we proceed as follows. Consider all the genotypes $\mathcal{G}_k$ such that $\phi_E = \phi(\mathcal{G}_k)$. We are interested in computing the expectation $\mathbb{E}\left(N_k(T) \mid \mathrm{IC}_i\right)$, which is given by [57]:

$$\mathbb{E}\left(N_k(T) \mid \mathrm{IC}_i\right) = \frac{\partial f_i}{\partial s_k}(\vec{1}, T), \tag{5.21}$$

therefore the ODE governing the evolution of this quantity can be obtained from Eqs. (5.10). By applying ∂_{s_k} to that system,,

$$\frac{d}{dT}\frac{\partial f_i}{\partial s_k}(\vec{s}, T) = -\frac{\partial f_i(\vec{s}, T)}{\partial s_k} + \sum_l \left.\frac{\partial F_i}{\partial s_l}\right|_{\vec{f}(\vec{s},T)} \cdot \frac{\partial f_l}{\partial s_k}(\vec{s}, T). \tag{5.22}$$

Evaluating at $\vec{s} = \vec{1}$, we obtain

$$\frac{d}{dT}\frac{\partial f_i}{\partial s_k}(\vec{1}, T) = -\frac{\partial f_i(\vec{1}, T)}{\partial s_k} + \sum_l \left.\frac{\partial F_i}{\partial s_l}\right|_{\vec{1}} \cdot \frac{\partial f_l}{\partial s_k}(\vec{1}, T) \tag{5.23}$$

In order to proceed further, we define the vector $\vec{\beta}_k(T)$ whose components are $\vec{\beta}_k(T) = (\beta_{k,i}(T), i = 1, \ldots, N_G) = \left(\frac{\partial f_i}{\partial s_k}(\vec{1}, T), i = 1, \ldots, N_G\right)$, i.e. the ith component of $\vec{\beta}_k(T)$ is the average population of cell of genotype $\mathcal{G}_k$ at time T with initial conditions given by IC_i. From Eq. (5.23) we derive a linear ODE for the time evolution of $\vec{\beta}_k(T)$:

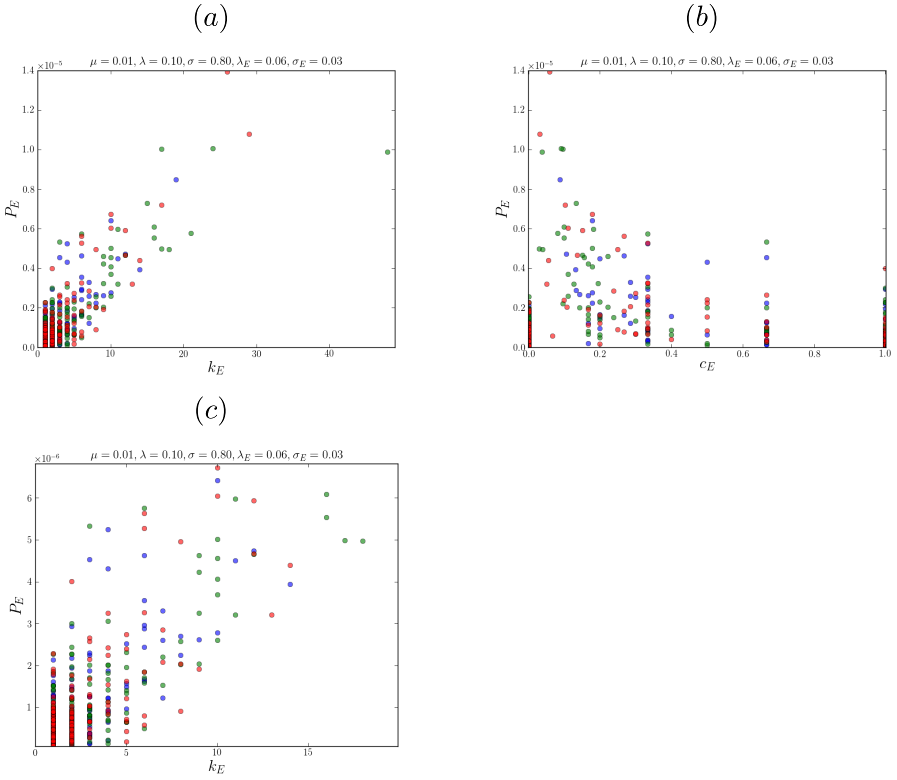


Figure 5.3: Probability of escape in the time-continuous model. Different colours represent results for different graphs. As we expected, P_E is positive correlated with k_E (plot (a)) and negatively correlated with c_E (plot (b)), in agreement with the relation that exists between clustering coefficient and degree. Plot (c) represents a zoom of (a).

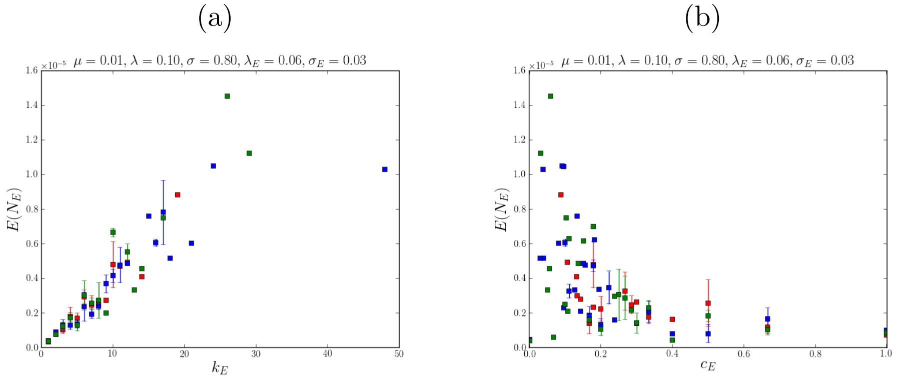


Figure 5.4: Plots showing the dependence of the unconditioned average number of cells accumulated within the escape phenotype during the initial regime, $\mathbb{E}\left(N_E\right)$. Plot (a) shows how N_E varies as the degree of the escape phenotype, k_E, changes. Plot (b) shows results regarding the behaviour of $\mathbb{E}\left(N_E\right)$ as the clustering coefficient, c_E, of the escape phenotype varies. We observe that $\mathbb{E}\left(N_E\right)$ is positively correlated with k_E and negatively correlated with c_E.

$$\frac{d}{dT}\vec{\beta}_k(T) = -\vec{\beta}_k(T) + DF(\vec{1}) \cdot \vec{\beta}_k(T) = (-\text{Id} + DF(\vec{1})) \cdot \vec{\beta}_k(T), \tag{5.24}$$

with $(DF(\vec{1}))_{ij} = \frac{\partial F_i}{\partial s_j}(\vec{1})$. The solution of Eq. (5.24) is given by:

$$\vec{\beta}_k(T) = \exp((-\text{Id} + DF(\vec{1})) \cdot T) \cdot \vec{\beta}_k(0) \tag{5.25}$$

where $\vec{\beta}_k(0) = (\delta_{1k}, \delta_{2k}, \ldots, \delta_{N_G k})$. Finally, we define $\mathbb{E}\left(N_E(T=1)\right)$ as the expected number of cells that reach the escape phenotype during the initial regime process (see Section 5.1.1.2) averaged over all possible initial conditions:

$$\mathbb{E}\left(N_E(T=1)\right) = \frac{1}{N_G - k_E} \sum_{k \in \langle \phi_E \rangle} \vec{1}_N \cdot \vec{\beta}_k(s) \tag{5.26}$$

note, $\vec{1} \cdot \vec{1}_N = N_G - k_E$.

Fig. 5.4 shows results regarding how $\mathbb{E}\left(N_E\right)$ $(N_E = N_E(T=1))$ changes as we vary the degree (Fig. 5.4(a)) and the clustering coefficient (Fig. 5.4(b)) of the escape phenotype. We observe that $\mathbb{E}\left(N_E\right)$ is positively correlated with k_E, i.e. the larger the number of genotypes which bear the escape phenotype, the bigger (on average) is the number of cells which accumulate within the escape phenotype during the initial, fast dynamics regime. This behaviour is a consequence of the fact that, in general, the larger k_E, the more likely is that cells trying to escape the ill-adapted genotypes find a route into the escape phenotype. The negative correlation between $\mathbb{E}\left(N_E\right)$ and c_E is then a direct consequence of $c_E(k_E) \approx (k_E - 1)^{-\alpha}$ [52].

In addition to $\mathbb{E}\left(N_E\right)$, since we are interested in studying survival to escape, we need to analyse the average number of cells that reach the escape conditioned to eventual escape, $\mathbb{E}\left(N_E \mid N_0 > 0\right)$. To compute this quantity, we first consider $\mathbb{E}\left(N_k(T) \mid \text{IC}_i, N_0 > 0\right)$, i.e. the expected value of the population of the genotype $\mathcal{G}_k$ such that $\phi(\mathcal{G}_k) = \phi_E$ with initial condition IC_i, which is given by:

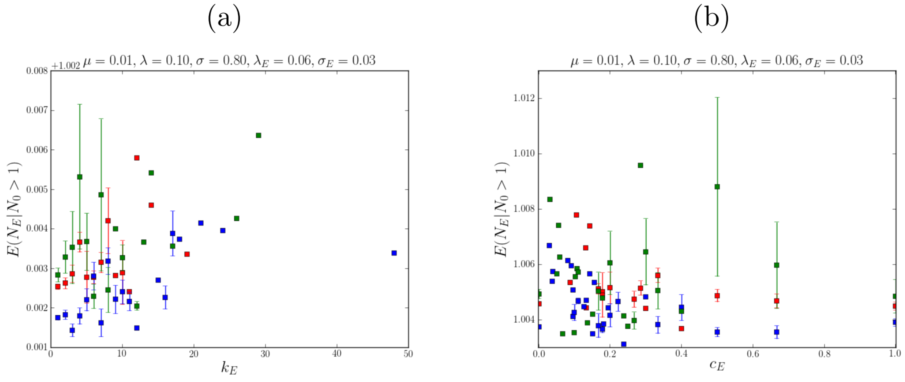


Figure 5.5: Plots showing the dependence of the average number of cells accumulated within the escape phenotype conditioned to eventual escape, $\mathbb{E}\left(N_E \mid N_0 > 0\right)$. Plot (a) shows how $\mathbb{E}\left(N_E \mid N_0 > 0\right)$ varies as the degree of the escape phenotype, k_E, changes. Plot (b) shows results regarding the behaviour of $\mathbb{E}\left(N_E \mid N_0 > 0\right)$ as the clustering coefficient, c_E, of the escape phenotype varies. We observe that $\mathbb{E}\left(N_E \mid N_0 > 0\right)$ is positively correlated with k_E and negatively correlated with c_E, although the such correlation is weaker than for the unconditioned average $\mathbb{E}\left(N_E\right)$.

$$\begin{aligned}
\mathbb{E}\left(N_k(T) \mid \mathrm{IC}_i, N_0 > 0\right) \qquad &= \sum_k n_k P(N_k(T) = n_k \mid \mathrm{IC}_i, N_0 > 0) \\
&= \frac{\sum_k n_k P(N_k(T) = n_k; N_0 > 0 \mid \mathrm{IC}_i)}{P(N_0 > 0)} \\
&= \frac{\sum_k n_k P(N_k(T) = n_k \mid \mathrm{IC}_i)}{P(N_0 > 0)} \\
&= \frac{\mathbb{E}\left(N_k(t) \mid \mathrm{IC}_i\right)}{P(N_0 > 0 \mid \mathrm{IC}_i)} \qquad (5.27)
\end{aligned}$$

Eq. (5.27) implies that to compute $\mathbb{E}\left(N_E \mid N_0 > 0\right)$ we must renormalise Eq. (5.26) by a factor which is equal to the escape probability, $P(N_0 > 0 \mid \mathrm{IC}_i)$, computed as P_E in Section 5.2.1.1, (where we only consider those initial conditions, IC_i, for which $P(N_0 > 0 \mid \mathrm{IC}_i) > 0$):

$$\mathbb{E}\left(N_E \mid N_0 > 0\right) = \frac{1}{N_G - k_E} \sum_{k \in \langle \phi_E \rangle} \sum_{i=1}^{N_G} \frac{\beta_{k,i}(s)}{P(N_0 > 0 \mid \mathrm{IC}_i)} ds \qquad (5.28)$$

The behaviour of $\mathbb{E}\left(N_E \mid N_0 > 0\right)$ as k_E and c_E vary is shown in Fig. 5.5(a) and (b), respectively. We find that $\mathbb{E}\left(N_E \mid N_0 > 0\right)$ and k_E are positively correlated (see Fig. 5.5(a)), although the observed correlation appears to be weaker for the conditioned average than for the unconditioned one (compare to Fig. 5.4(a)). Similarly, $\mathbb{E}\left(N_E \mid N_0 > 0\right)$ and c_E are negatively correlated (see Fig. 5.5(b)).

5.2.2 Slow dynamics: Survival probability

We now proceed to examine the post-escape dynamics, i.e. once the fast, transient regime in which escape occurs and described by Eqs. (5.10), a dynamical regime characterised by much longer time scales ensues in which, conditioned to the occurrence of escape, the population may still get extinct due to the within-escape phenotype population dynamics. We define $P_S(t)$ as the survival probability at time t, that is $P_S(t) = P(N_E(t) > 0 \mid \mathrm{IC}_i)$, where now IC_i corresponds to an i with belongs to the escape phenotype. This probability is computed for a large time ($t \to \infty$). We model this slow, long-time regime using two different coarse-grained models. The first model, depending on clustering coefficient, is described in Section 5.1.1.3. The second model, depending on weighted clustering, M_E, is introduced in Section 5.1.1.5. In both models, the theoretical survival probability, P_S, is the complementary of the clearance probability, P_C (see Section 5.1.1.5), $P_S = 1 - P_C$. The clearance probability is given by smaller root of (see [57]):

$$F_E(P_C) - P_C = 0, \tag{5.29}$$

This is a second order equation which can be solved exactly. Its two roots, x_1 and x_2, are given by

$$x_1 = 1, \quad x_2 = \frac{\sigma + \lambda\mu^2(1-\alpha)^2}{\lambda(1-\mu(1-\alpha))^2},$$

where $\alpha = c_E$ in the first model, and $\alpha = M_E$ in the second model. $P_C = x_2$, therefore $P_S = 1 - x_2$. Note that P_S is the survival probability conditioned to escape, as the relation $P_S = 1 - P_C$ assumes that at least one cell has reached the escape phenotype.

In the first model, the behaviour of P_S as we vary c_E and k_E using $F_E(c_E)$ is shown in Fig. 5.6(a) and (b), respectively. We find in this case that $P_S(t)$ is not well approached by theoretical P_S and its relation with c_E and k_E is not clear.

Nevertheless, in the second model, the survival probability $P_S(t)$, approached theoretically by,

$$P_S = 1 - \frac{\sigma_E + \lambda_E\mu^2(1-M_E)^2}{\lambda_E(1-\mu(1-M_E))^2} = \frac{-\sigma_e - \lambda_E + 2\lambda_E(1-\mu+\mu M_E)}{\lambda_E(1-\mu+\mu M_E)^2} \tag{5.30}$$

models perfectly $P_S(t)$ as a function of M_E (see Fig 5.7). Clearly, there are a strong positive correlation between $P_S(t)$ and M_E.

5.3 Conclusions

We have formulated a population dynamics model, consisting of a multi-type time-continuous branching process, where types are associated to genotypes and their proliferation probability is defined depending on which kind of genotype we are (non escape or escape). This model allows analyse the problem of evolutionary escape and survival for cell populations with genotype-phenotype map.

We have studied the global process of evolutionary escape and survival escape as a process with two time scales, characterised by two different regimes: an initial, fast regime, during which escape actually occurs. Then a quasi-steady regime, slow regime ensues associated to the dynamics once escape has occurred and the population has reached the escape phenotype.

We have shown that, while the escape probability depends on size of the neutral network of the escape phenotype (i.e. its degree), the probability of survival is essentially determined by its robustness (i.e. the resilience of the escape phenotype against genetic mutations). We have shown that the simple topological of measure phenotypic robustness defined in [52], i.e. the clustering coefficient, is not well-adapted to describe robustness in the context of the coarse-grained dynamics of survival. Rather, a new measure in terms of a weighted clustering coefficient is necessary to accurately account for robustness under said dynamics.

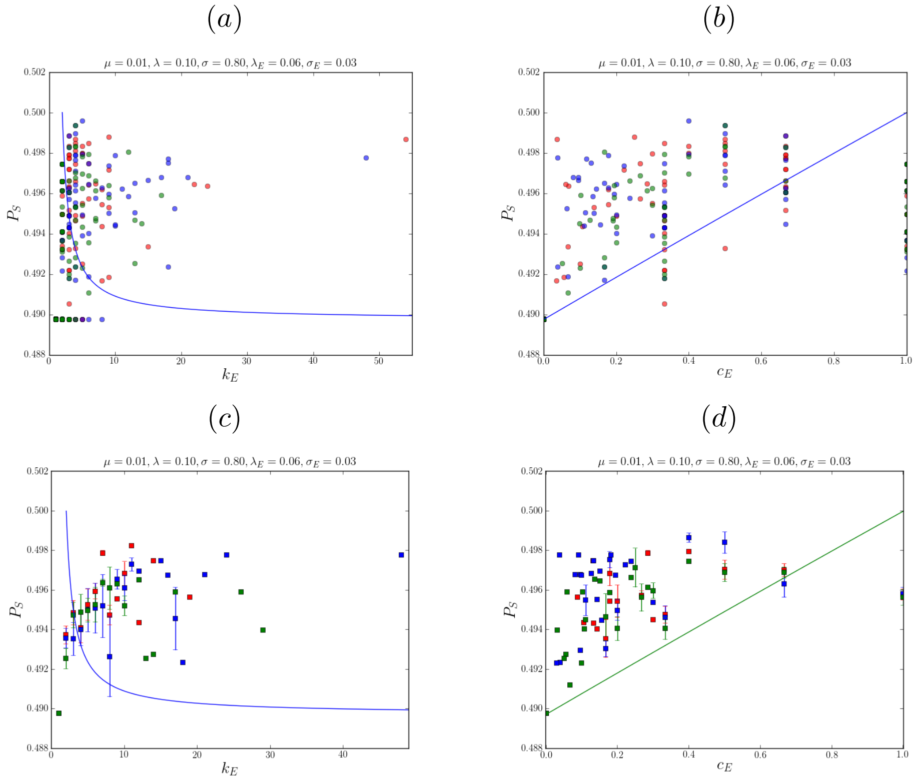


Figure 5.6: Probability of survival conditioned to escape as a function of k_E ((a),(c)) and c_E ((b),(d)). We consider an initial condition where $N_i(\tau = 0) = 1$ for a randomly chosen $\mathcal{G}_i$ so that $\phi(\mathcal{G}_i) = \phi_E$. Solid line represents the theoretical P_S from $(F_E(c_E))$, if we suppose that one individual go out from the escape phenotype, then is unable to reach again the escape. Top: Scatter plot. Bottom: Joining by degree (c), clustering (d). Different colours represent results for different graphs.

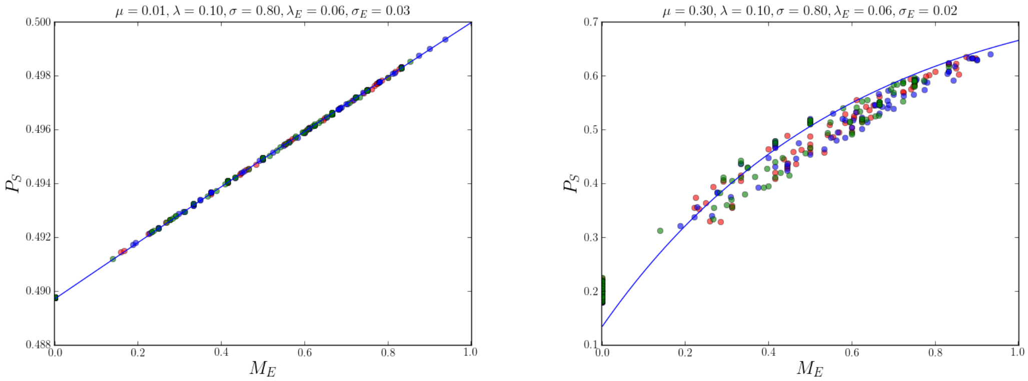


Figure 5.7: Plot showing the correlation between survival probability $P_S(t)$ and the new measure M_E for two different sets of parameters. Solid line corresponds to the theoretical P_S (approach in Eq (5.30)). Different colours of dots represent results for different graphs.

Our analysis of the fast-decaying, initial regime reveals (see Fig. 5.3) that the escape probability, P_E, is positively correlated with the size of the neutral network associated to the escape phenotype, i.e. its degree, k_E, and negatively correlated with its clustering coefficient, c_E. This is a direct consequence of the inverse relation between both quantities $c_E \approx (k_E - 1)^{-\alpha}$ with $\alpha \approx 1$ [52]. Thus, the size of borderline of escape genotypes takes an important role during the fast, initial regime, where escape actually occur. We have also calculated the average number of cells accumulated within the escape phenotype during the initial regime, both conditioned and unconditioned to eventual escape (Section 5.2.1.2, see Figs. 5.5 and 5.4, respectively). Results show that $\mathbb{E}\left(N_E\right)$ and $\mathbb{E}\left(N_E \mid N_0 > 0\right)$ are identically correlated with k_E and c_E as the escape probability.

On the other hand, our analysis of the slow dynamic regime, where escape has already occurred and we study the dynamics of the population of the escape phenotype, reveals that the probability of survival conditioned to escape is determined by the robustness of the escape phenotype. However, we have shown that the topological measure of robustness proposed in Chapter 4 and in [52], i.e. the clustering coefficient, c_E, does not accurately describe the coarse-grained, quasi-steady state dynamics of the escaped population. Instead, a new measure of robustness M_E [51], closely related to a weighted clustering coefficient [77], has been defined which is better suited to defining robustness in the context of our coarse-grained dynamics in the slow regime. Results show a strong positive correlation between the survival probability, P_S, and, M_E. Moreover, this measure provides much better accuracy for the survival probability as a function of M_E (Fig 5.7). The reason for the failure of the clustering coefficient to accurately describe robustness for the survival dynamics appears to be related to the fact that the clustering coefficient does not take into account the heterogeneity of the genotype nodes belonging to the neutral network of the escape phenotype (see Fig. 5.2). Our measure, M_E, takes into account the heterogeneity within the escape neutral network, thus providing a more detailed, better suited description of robustness.

To summarise, these results show that, while the escape probability, P_E, is essentially determined by k_E, the survival probability long-time survival probability P_S depends on the robustness of the escape phenotype, as measured by our weighted clustering coefficient M_E. This weighted clustering explains much better the population dynamics of survival within the escape neutral network than the usual local clustering coefficient c_E (Fig 5.6).

Chapter 6

Summary

6.1 Discussion and Conclusions

In this PhD thesis we have developed a multi-scale model of biological evolution which accounts for the mapping between genotype and phenotype as determined by a model of the gene regulatory network. We have formulated a simple model in Chapter 2 of genotype-phenotype map inspired in the model proposed in [103] and studied in [104] by Wagner which take into account that selective pressure acts at the level of phenotypes [52]. The theoretical basis of this genotype-phenotype map was established by Kauffman [56], where phenotypes or differentiated states are the steady states of the dynamical systems associated to the gene regulatory network.

In Chapter 3, we have characterised the geometrical and topological properties of the genotype-phenotype space obtained from the multi-scale model which assumes a selective pressure acting at the level of phenotypes. We have defined phenotypic robustness as the clustering coefficient of phenotype nodes in the pseudo-bipartite genotype-phenotype network and evolvability is defined as the emergence of a giant connected component in the phenotype network, i.e. the one-mode projection of genotype-phenotype network, which ensures global connectedness and navigability of the space of phenotypes. Further to this global definition of robustness, we have produced a local measure of robustness based on the distribution of size of connected components in the phenotype network. In particular, we have shown that, beyond global connectedness, the phenotype network exhibits the small-world property. We have further explored whether evolvability is a robust property. We have observed that under random attack, the giant connected component, and, therefore, our system's evolvability, is very resilient. However, similarly to other complex networks, when the attack is targeted rather than random, weaknesses arise and evolvability breaks down. Moreover, we show that, regarding the dependence of the phenotype clustering coefficient on the degree of the phenotype nodes, i.e. the size of the associated neutral network $c_\phi(k_\phi)$, exists an inverse relation between clustering coefficient and degree: $c_\phi(k_\phi) \propto (k_\phi - 1)^{-\alpha_{p,l_v}}$ with $\alpha_{p,l_v} \geq 1$. This result has an important interpretation in terms of the concept of *neighbourhood* and the relation between robustness and evolvability [107, 108]. It implies that those phenotypes with larger cryptic variability exhibit, in relative terms, higher accessibility to new phenotypes that those with smaller degree to the expense of robustness.

In Chapter 4 we have carried out a comparative analysis in the context of the problem of evolutionary escape. We have introduced a novel perspective using complex genotype-phenotype networks where selective pressures acts on phenotypes, obtained from the multi-scale model in Chapter 2, rather than regular, hypercube genotype lattices where fitness is directly determined by the genotype. In order to compute escape probabilities we have formulated a population dynamics model, consisting of a multi-type branching process [57], where types are associated to genotypes and their proliferation probability is assigned according to the corresponding phenotype. We have observed an heterogeneous behaviour in all the observables associated to the escape problem that we have investigated applied to complex networks in contrast with the much more uniform behaviour in hypercube genotype spaces [50]. This heterogeneity applies to the distribution of distance between phenotypes, the escape probability, the average escape time to escape and the escape time distribution conditioned to eventual escape. Moreover, this heterogeneity, in some instances, causes an order of magnitude bigger in the escape probability than in its regular hypercube associated.

Finally, in Chapter 5, the problem of evolutionary escape, but also survival probabilities have been studied for cell populations with genotype-phenotype map [51]. Prior approaches to this problem, which do not consider populations with genotype-phenotype structure and associate fitness values directly to genotypes [53, 54, 89, 90, 84], have focused on the problem of estimating the probability of reaching a well-adapted (so-called escape) genotype for an initial population entirely composed of cells with ill-adapted genotypes. This point of view implicitly assumes that, once the escape genotype has been reached, the populations survives with probability one. When the genotype-phenotype map is added to the picture, the situation becomes more complex [50]: genotype-phenotype structure endows complex structure to the escape phenotype which, in particular, provides robustness to the escape phenotype. Under genotype-phenotype structure, each phenotype has an associated neutral network which possess a rather reach topology [105, 2, 52], so that the dynamics of the system post-escape is not trivial.

In order to explore this issue, we have formulated a population dynamics model, consisting of a multi-type time-continuous branching process, where types are associated to genotypes and their birth and death probabilities depend on the associated phenotype (non-escape or escape) via the genotype-phenotype map defined in Chapter 2 and in [52]. We have shown that, within the setting associated to the escape problem, where the types associated to non-escape phenotypes are sub-critical and types associated to the escape phenotype are super-critical, separation of time scales naturally arises and two dynamical regimes emerge: a fast-decaying regime, accounting for the early stages in the evolution of the system, and associated to the escape process itself, and a slow regime which corresponds to the (survival) dynamics of the population once the escape phenotype has been reached.

We have shown that, while the escape probability depends positively on size of the neutral network of the escape phenotype (i.e. its degree), the probability of survival is essentially determined by its robustness (i.e. the resilience of the escape phenotype against genetic mutations). We have shown that clustering coefficient, is not well-adapted to describe robustness in the context of the coarse-grained dynamics of survival. In consequence, we have introduced a weighted clustering coefficient as new measure of robustness adapted to slow regime. It takes into account the heterogeneity within the escape neutral network, thus providing a more detailed, better suited description of robustness.

6.2 Future work

This thesis opens a number of interesting avenues for future work in the field of evolutionary dynamics of systems with genotype-phenotype map. In particular, we intend to extend our methodology to other, such systems, e.g. the neutral networks obtained in [1, 2] for RNA.

Regarding the particular case of RNA neutral networks, we intend to carry out a comparative analysis between the RNA genotype-phenotype networks (where selection pressure is effectively present through viability conditions on the folding structure) to regular hypercubic lattices, in order to better understand the topological properties of the former [1, 2].

Another avenue of future work is to apply our models of evolutionary escape on genotype-phenotype networks to the study of emergence of drug resistance in tumours. In these systems, a complex epigenetic landscape in which co-evolution of robustness and evolvability naturally arises [98]. Therefore, our methods for the study of escape on networks, where these properties have been shown to co-evolve, should be very relevant.

Another area in which our methods should prove useful concerns the formulation of coarse-grained population dynamics models. In Chapter 5, we have defined a weighted clustering coefficient in terms of which the population dynamics of the escape phenotype can be analysed as a single-type system (rather than as a multi-type population, where each type corresponds to a genotype). We expect this methodology be applicable to coarse-grained more general population dynamical models. This involves a more careful analysis of whether our definition of the weighted clustering coefficient is restricted to the particular situation depicted in Chapter 5, or, on the contrary has a wider range of applicability.

Another question to solve is how a change in GRN affects the full genotype-phenotype network obtained from model in 2 and to find a method to compare genotype-phenotype networks in order to decide if two networks are equivalent in any sense.

Chapter 7

Appendix

7.1 Generalization of GRN

In this Section we prove that we can assume connected GRNs. If a gene regulatory network is not connected then we can consider each connected component, study its behaviour and reconstruct the pseudo-bipartite graph from each *small* pseudo-bipartite graph. First, we show it for the simplest case: GRNs is composed by two connected components: T_1 and T_2. It can be generalised for any number of connected components. This is a consequence of the associative property of the graph product. In other words, if we have a GRN with three connected components, first study properties for two of them and after recalculate properties as you have only two connected components. Consider T_1 and T_2 networks,

- Network T_1 generates the genotype set G_1, where $|G_1| = g_1$ and $g_1 = 2^{m_1}$, $m_1 =$ #edges of T_1.
- Network T_2 generates the genotype set G_2, where $|G_2| = g_2$ and $g_2 = 2^{m_2}$, $m_2 =$ #edges of T_2.

Each component generates some phenotypes,

- Network T_1 generates the phenotype set F_1, where $|F_1| = f_1$.
- Network T_2 generates the phenotype set F_2, where $|F_2| = f_2$.

In the same way that we have constructed the pseudo-bipartite graph, we can proceed in the same way for each connected component. We define the graph product in order to obtain the complete pseudo-bipartite graph (see Figure 7.1 as an example):

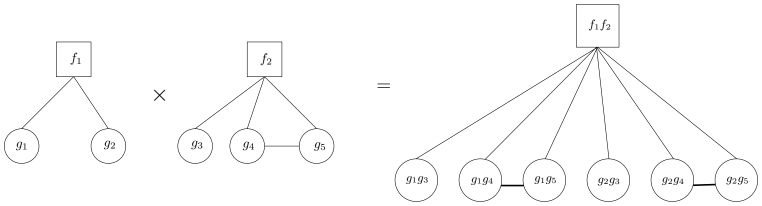


Figure 7.1: Example of graph product.

- Set of nodes: $G_1 \times G_2$ and $F_1 \times F_2$, i.e. we have $g_1 \cdot g_2$ genotypes and $f_1 \cdot f_2$ phenotypes.
- Set of edges:
  - $\exists$ edge between genotypes, $(a, b) \sim (c, d)$, $a, c \in G_1$ and $b, d \in G_2$ iff $\text{diff}(a, c) + \text{diff}(b, d) = 1$, where $\text{diff}(x, y) = 0$ iff $x = y$.
  - $\exists$ edge between genotype and phenotype, $(a, b) \in F_1 \times F_2$ (phenotype), $(c, d) \in G_1 \times G_2$ (genotype) iff $a \sim c$ and $b \sim d$.

Remark: We use $a \sim b$ if there are any edge between them in the connected component.

After introducing the definition of the product graph, we infer the properties of the product graph and of their components properties: phenotype degree, genotype degree, genotype clustering coefficient and phenotype clustering coefficient.

Phenotype degree Given $v = (v_1, v_2) \in F_1 \times F_2$. $Deg(v) = \deg(v_1) \cdot \deg(v_2)$.

Genotype degree Given $u = (u_1, u_2) \in G_1 \times G_2$.

$$\begin{aligned} Deg(u) = \sum_{v \sim u} 1 = \sum_{(v_1,v_2)\sim(u_1,u_2)} 1 = \sum_{(v_1,u_2)\sim(u_1,u_2)} 1 + \sum_{(u_1,v_2)\sim(u_1,u_2)} 1 = \\ = \deg(u_1) + \deg(u_2). \end{aligned}$$

Phenotype clustering coefficient Given $v = (a, b) \in F_1 \times F_2$. Clustering coefficient is defined as,

$$c_v = \frac{2T_v}{\deg(v)(\deg(v) - 1)}, \text{ where } T_v = \# \text{ triangles with a vertex in } v.$$

First, we must count the triangles with $v = (a, b) \in F_1 \times F_2$ as vertices, T_v.

$$T_v = T_a \deg(b) + T_b \deg(a).$$

Using $c_a = \frac{2T_a}{\deg(a)(\deg(a)-1)}$, $c_b = \frac{2T_b}{\deg(b)(\deg(b)-1)}$, $\deg(v) = \deg(a)\deg(b)$ we obtain:

$$\begin{aligned} c_v = & \frac{2T_v}{\deg(v)(\deg(v) - 1)} = \frac{2\,(T_a \deg(b) + T_b \deg(a))}{\deg(a)\deg(b)(\deg(a)\deg(b) - 1)} = \\ = & \frac{c_a \deg(a)(\deg(a) - 1)\deg(b) + c_b \deg(b)(\deg(b) - 1)\deg(a)}{\deg(a)\deg(b)(\deg(a)\deg(b) - 1)} = \\ = & \frac{c_a(\deg(a) - 1) + c_b(\deg(b) - 1)}{\deg(a)\deg(b) - 1} \end{aligned} \tag{7.1}$$

The product graph is associative, so we can generalise this procedure for any number of connected components. For example, for 3 connected components: $v = (a, b, c) \in (F_1 \times F_2 \times F_3)$, the number of triangles is:

$$T_{(a,b)} = T_a \deg(b) + T_b \deg(a) \quad \text{and } \deg((a, b)) = \deg(a)\deg(b).$$

Then,

$$T_{(a,b,c)} = (T_a \deg(b) + T_b \deg(a)) \deg(c) + T_c (\deg(a)\deg(b)).$$

Genotype clustering coefficient Given $u = (a, b) \in G_1 \times G_2$. Clustering coefficient is defined as,

$$c_u = \frac{2T_u}{\deg(u)(\deg(u) - 1)}, \text{ where } T_u = \# \text{ triangles with a vertex in } u.$$

First, we must to count how many triangles we have with $u = (a, b) \in G_1 \times G_2$ as a vertex, (T_u).

$$\begin{aligned}
T_u =& \#\{(a',b'),(a'',b'')|(a'',b'') \sim (a',b'),(a',b') \sim (a,b),(a,b) \sim (a'',b'')\} = \\
=& \#\{(a',b'),(a'',b''),(a=a')|(a'',b'') \sim (a,b'),b' \sim b,(a,b) \sim (a'',b'')\}+ \\
+& \#\{(a',b'),(a'',b''),(b=b')|(a'',b'') \sim (a',b),a' \sim a,(a,b) \sim (a'',b'')\} = \\
=& \#\{(a',b'),(a'',b''),(a=a'),(a=a'')|b'' \sim b',b' \sim b,b \sim b''\}+ \\
+& \# \underbrace{\{(a',b'),(a'',b''),(a=a'),(b=b'')|(a'',b) \sim (a,b'),b' \sim b,a, \sim a''\}}_{0} + \\
&+ \# \underbrace{\{(a',b'),(a'',b''),(b=b'),(a'=a'')|b'' \sim b,a' \sim a,(a,b) \sim (a',b'')\}}_{0} + \\
+& \#\{(a',b'),(a'',b''),(b=b'),(b=b'')|a'' \sim a,a' \sim a,a \sim a''\} = \\
=& T_b + T_a
\end{aligned} \tag{7.2}$$

Using (7.2) and $c_a = \frac{2T_a}{\deg(a)(\deg(a)-1)}, c_b = \frac{2T_b}{\deg(b)(\deg(b)-1)}$

$$\begin{aligned}
c_u =& \frac{2(T_a + T_b)}{(\deg(b) + \deg(a))(\deg(b) + \deg(a) - 1)} = \\
=& \frac{c_a\deg(a)(\deg(a) - 1) + c_b\deg(b)(\deg(b) - 1)}{(\deg(b) + \deg(a))(\deg(b) + \deg(a) - 1)}
\end{aligned} \tag{7.3}$$

Remark: We do not count edges between genotypes to phenotypes in order to count triangles.

7.2 Graph definitions

Given a undirected graph[1] G (without weights), with n nodes and m edges. It is defined [75]:

- *Degree* of a node, k: is the number of neighbours of a node. Average degree in a network $\langle k \rangle$ is the average of the degree of all nodes, and the degree distribution $P(k)$, represents the probability to pick a node with degree k. In the case of directed graphs two kind of degree are defined: *in-degree* and *out-degree*, first counts the number of incoming edges, second counts the number of outgoing edges.
- The *clustering* coefficient, c_v, of a node v is defined as the fraction between the actual number of triangles that have the node as a vertex to the corresponding maximum number of such triangles (if v has k_v neighbours, then at most $c_v = 2T_v/(k_v - 1)$, where T_v is the number of triangles that v make with its neighbours). In other words, if Neigh_v corresponds to the v neighbourhood and $e_{i,j}$ are edges between nodes in Neigh_v, then clustering coefficient is the proportion of $e_{i,j}$: $c_v = \frac{2|\{e_{i,j}|i,j \in \text{Neigh}_v\}|}{k_v(k_v-1)}$. It quantifies the

[1] *Undirected graph:* edges do not have orientation, edge (i, j) is identical to (j, i).

connectivity between neighbours of a node. The average clustering coefficient $\langle c\rangle$ is the average of the clustering coefficient for all nodes.

- *Shortest path* between two nodes, v_1 and v_2 is the shorter number of steps needed to go from v_1 to v_2 among all possible paths between both nodes. Also it is called, distance between nodes. If we have the shortest path for each pair of nodes, then its maximum is defined as the *diameter* of the graph.

- A *small-world network* is a type of graph in which most nodes are not neighbours of one another, but most nodes can be reached from every other by a small number of steps. Specifically, a small-world network is defined to be a network where the typical distance L between two randomly chosen nodes (the number of steps required) grows proportionally to the logarithm of the number of nodes n in the network: $L \propto \log n$.

- A *regular graph* is a graph where all nodes have the same degree, for example a regular lattice in d-dimensions.

- A *random graph* is a graph which probability p to have a connection between two random nodes. In this kind of graph degree distribution follows a Poisson distribution (for n large, otherwise it follows a binomial distribution).

- A *scale-free graph* is a graph whose degree distribution follows a power law ($P(k) \sim k^{\gamma}$), at least asymptotically. These graphs have hubs, few nodes with very large degree.

- *Assortativity* is showed in a graph if the correlation between the average of nearest neighbours and their degree is positive. It means, that more connected nodes are preferably connected to each other. On the other hand, if the correlation is negative we have a *disassortative* network.

- *Connected components* is the number of isolated parts that a graph is composed. A connected component is a set of nodes that for each pair of them exist a path to go from one to other. A used term is *giant component* that is a connected component that contains a significant fraction of all nodes. Other related definitions are *in component* and *out component. In component* can be defined as given a node x, the set of nodes which can reach node x via directed path of links. Likewise, *out component* is defined as, given a node x, the set of nodes which can be reached from node x via a directed path.

A *bipartite graphs* is a graph whose vertices can be divided into two disjoint sets U and V such that there are only edges between $u \in U$ and $v \in V$ nodes. It does not exist edges between nodes which belong to the same disjoint set. Equivalently, a bipartite graph is a graph that does not contain any odd-length cycles.

7.3 Severe selective pressure

As we have formulated in Chapter 2 the selective pressure (Section 2.1.1.1) we give some idea about visual genotype-phenotype graphs structure imposing that the period of time of length v, which implies that solutions with periods longer than v are non-viable is equal to 0 ($l_v = 0$). In other words, we only consider viable phenotypes if it is a fixed point. GRNs are generated using Strogatz-Watts model with $p = 0.9$ and $p = 0.1$. Figure 7.2 shows some examples.

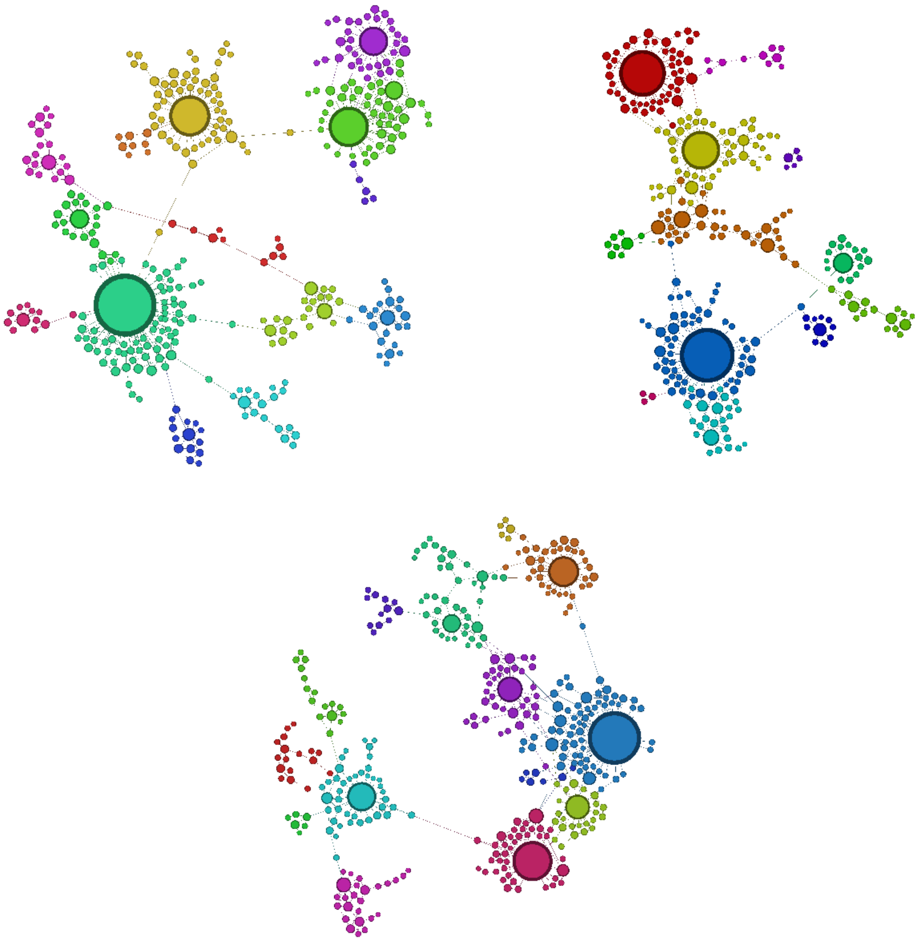


Figure 7.2: Genotype-phenotype networks. Big points represent phenotypes and each colour corresponds to all nodes with same phenotype. Top plots are generated with $p = 0.1$, bottom $p = 0.9$.

7.4 Other definitions:

Kernel density estimation: Gaussian Kernel Kernel density estimation [95] is a method to estimate the probability density function (PDF), $f(x)$, of a random variable in a non-parametric way using Eq.(7.4). Kernel density estimation is a fundamental data smoothing problem where inferences about the population are made, based on a finite data sample. Kernel density estimates are closely related to histograms, but can be endowed with properties such as smoothness or continuity by using a suitable kernel. The bandwidth of the kernel is a free parameter which exhibits a strong influence on the resulting estimate (it can produce a more picked or smoothed distribution). In other words, it is the tuning parameter which give a better estimation of the true density. As a particular case, we use a Gaussian Kernel $K(u) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}u^2)$:

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^{n} K\left(\frac{x - x(i)}{h}\right) \tag{7.4}$$

where $\int K(t)dt = 1$ to ensure that the estimates $f(x)$ integrates to 1 and where the kernel function K is usually chosen to be a smooth unimodal function with a peak at 0. Even though Gaussian kernels are the most often used, there are various choices among kernels. It is used in Figure 3.7.

7.5 Escape probabilities in the discrete-time branching process

Given an escape phenotype ϕ_E we define the initial condition IC_i as $N_i(0) = 1$ and $N_j(0) = 0$ for all $j \neq i$ and we assume $\phi(\mathcal{G}_i) \neq \phi_E$. We represent results of the probability of escape P_E, that is the probability to achieve the escape phenotype before extinction starting from the configuration IC_i and averaged for all possible initial conditions. These results correspond to the discrete model of Chapter 4 and in [50].

Fig 5.3 shows results regarding how P_E changes as we vary the degree k_E (Fig 7.3 (a)) and the clustering coefficient ϕ_E (Fig 7.3 (b)). These results agree with results in the time-continuous model model described in Chapter 5. We observe that P_E is positively correlated with k_E, i.e. the larger the number of genotypes which belong the escape phenotype, the bigger (on average) is the probability of achieve the escape phenotype. Otherwise, the negative correlation between P_E and c_E is a direct consequence of $c_E(k_E) \approx (k_E - 1)^{-\alpha}$ [52].

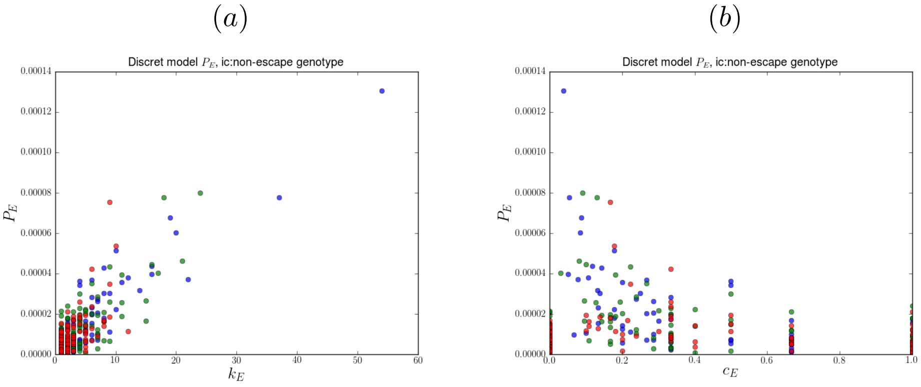


Figure 7.3: Probability of escape in the discrete model. Different colours represent results for different graphs. As we expected, P_E is positive correlated with k_E (plot (a)) and negatively correlated with c_E (plot (b)), in agreement with the relation that exists between clustering coefficient and degree.

Bibliography

[1] J. Aguirre, J. M. Buldú, and S. C. Manrubia. Evolutionary dynamics on networks of selectively neutral genotypes: Effects of topology and sequence stability. *Phys. Rev. E*, 80:066112, Dec 2009.

[2] J. Aguirre, J. M. Buldú, M. Stich, and S. C. Manrubia. Topological structure of the space of phenotypes: The case of RNA neutral networks. *PLoS One,*, 6:e26324, 2011.

[3] T. Alarcon, H. M. Byrne, and P. K. Maini. A multiple scale model of tumour growth. *Multiscale Model. Sim.*, 3:440–475, 2005.

[4] T. Alarcon and H. J. Jensen. Quiescence: a mechanism for escaping the effects of drug on cell populations. *J. R. Soc. Interface*, 8:99–106, 2010.

[5] R. Albert. Scale-free networks in cell biology. *J. Cell Sci.*, 118:4947–495, 2005.

[6] K. Athreya and P. Ney. *Branching Processes.* Dover Books on Mathematics Series. Dover Publications, 2004.

[7] M. M. Babu, N. M. Luscombe, L. Aravind, M. Gerstein, and S. A. Teichmann. Structure and evolution of transcriptional regulatory networks. *Curr. Opin. Struc. Biol.*, 14:283–291, 2004.

[8] N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler. Bacterial persistence as a phenotypic switch. *Science*, 305:1622–1625, 2004.

[9] K. Ball, T. G. Kurtz, L. Popovic, and G. Rempala. Asymptotic analysis of multi-scale approximations to reaction networks. *Ann. Appl. Prob.*, 16:1925–1961, 2006.

[10] A. L. Barabasi and R. Albert. Emergence of scaling in random netowrks. *Science*, 286:509–512, 1999.

[11] A.-L. Barabási, E. Ravasz, and T. Vicsek. Deterministic scale-free networks. *Physica A*, 299(3–4):559–564, 2001.

[12] A.-L. Barabási. *Network Science.* Barabasi Lab, 2012.

[13] E. A. Bender and E. R. Canfield. The asymptotic number of labeled graphs with given degree sequences. *J. Comb. Theory. A*, 24(3):296–307, 1978.

[14] A. Bergman and M. Siegal. Evolutionary capacitance as a general feature of complex gene networks. *Nature*, 424(6948):549–552, 2003.

[15] R. A. Blythe and A. J. McKane. Stochastic models of evolution in genetics, ecology and linguistics. *J. Stat. Mech.*, page P07018, 2007.

[16] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4):311–316, 1980.

[17] B. Bollobas and O. Riordan. *Percolation.* Cambridge University Press, 2006.

[18] F. B. Brikci, J. Clairanbault, B. Ribba, and B. Perthame. An age-and-cyclin-structured cell population model for healthy and tumoural tissues. *J. Math. Biol.*, 57:91–110, 2008.

[19] R. G. Bristow and R. P. Hill. Hypoxia, DNA repair and genetic instability. *Nature Rev. Cancer*, 8:180–192, 2008.

[20] A. R. Brower and W. H. Haemers. *Spectra of graphs.* Springer, New York, NY, USA, 2012.

[21] A. Bunde and S. Havlin. *Fractals and disordered systems.* Springer-Verlag New York, Inc., 1991.

[22] G. Caldarelli and A. Vespignani. *Large scale structure and dynamics of complex networks: from information technology to finance and natural science*, volume 2. World Scientific, 2007.

[23] D. S. Callaway, J. E. Hopcroft, J. M. Kleinberg, M. E. J. Newman, and S. H. Strogatz. Are randomly grown graphs really random? *Phys. Rev. E*, 64:041902, 2001.

[24] K. Christensen and N. R. Moloney. *Complexity and criticality*, volume 1. Imperial College Press, 2005.

[25] F. R. K. Chung. Diameter and eigenvalues. *J. Am. Math. Soc.*, 2:187–196, 1989.

[26] F. R. K. Chung. *Spectral graph theory.* American Mathematical Society, USA, 1992.

[27] S. Ciliberti, O. Martin, and A. Wagner. Innovation and robustness in complex regulatory gene networks. *Proc. Natl. Acad. Sci. USA*, 104(34):13591, 2007.

[28] S. Ciliberti, O. C. Martin, and A. Wagner. Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLoS Comp. Biol.*, 3(2):164–173, 2007.

[29] M. C. Cowperthwaite, E. P. Economo, W. R. Harcombe, E. L. Miller, and L. A. Meyers. The ascent of the abundant: how mutational networks constrain evolution. *PLoS Comp. Biol.*, 4:e1000110, 2008.

[30] J. C. Delvenne, S. N. Yaliraki, and M. Barahona. Stability of graph communities across time scales. *Proc. Natl. Acad. Sci. USA*, 107:12755–12760, 2010.

[31] D. Dennett. *Darwin's dangerous idea.* New York: Simon and Schuster, 1995.

[32] S. N. Dorogovtsev and J. F. F. Mendes. Scaling properties of scale-free evolving networks: continuous approach. *Phys. Rev. E*, 63:056125, Apr 2001.

[33] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Anomalous percolation properties of growing networks. *Phys. Rev. E*, 64:066110, Nov 2001.

[34] P. Erdős and A. Rényi. On random graphs. *Publicationes Mathematicae*, 6:290–297, 1959.

[35] P. Erdős and A. Rényi. On the evolution of random graphs. In *Publication of the Mathematical Institute of the Hungarian Academy of Sciences*, pages 17–61, 1960.

[36] C. Espinosa-Soto, O. Martin, and A. Wagner. Phenotypic robustness can increase phenotypic variability after nongenetic perturbations in gene regulatory circuits. *J. Evol. Biol.*, 24(6):1284–1297, 2011.

[37] A. Eyre-Walker and P. D. Keightley. The distribution of fitness effects of new mutations. *Nat. Rev. Genet.*, 61:610–618, 1990.

[38] R. Fisher. *The Genetical Theory of Natural Selection.* Clarendon Press, Oxford, 1930.

[39] W. Fontana and P. Schuster. Shaping space: The possible and the attainable in RNA genotype-phenotype mapping. *J. Theor. Biol.*, 194:491–515, 1998.

[40] S. Fortunato. Community detection in graphs. *Phys. Rep.*, 486:75–174, 2010.

[41] G. F. Frobenius. *Über Matrizen aus nicht negativen Elementen.* Königliche Akademie der Wissenschaften, 1912.

[42] D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.*, 155:1716–1733, 2001.

[43] S. F. Greenbury, I. G. Johnston, M. A. Smith, J. P. K. Doye, and A. A. Louis. The effect of scale-free topology on the robustness and evolvability of genetic regulatory networks. *J. Theor. Biol.*, 267:48–61, 2010.

[44] G. R. Grimmett and D. R. Stirzaker. *Probability and random processes.* Oxford University Press, 2nd edition, 1992.

[45] J. B. S. Haldane. *The causes of evolution.* 1932.

[46] E. J. Hayden, E. Ferrara, and A. Wagner. Cryptic genetic variation promotes rapid evolutionary adptation in an RNA enzyme. *Nature*, 474:92–95, 2011.

[47] S. Huang. Reprogramming cell fates: reconciling rarity and robustness. *BioEssays*, 31:546–560, 2009.

[48] S. Huang. The molecular and mathematical basis of Waddington's epigenetic landscape: A framework for post-Darwinian biology. *BioEssays*, 34:149–157, 2011.

[49] S. Huang, I. Ernberg, and S. Kauffman. Cancer attractors: A systems view of tumors from a gene network dynamics and developmental perspective. *Sem. Cell Dev. Biol.*, 20:869–876, 2009.

[50] E. Ibáñez-Marcelo and T. Alarcón. Evolutionary escape on complex genotype-phenotype networks. *Submitted*, 2014.

[51] E. Ibáñez-Marcelo and T. Alarcón. Surviving evolutionary escape on complex genotype-phenotype networks. *Submitted*, 2014.

[52] E. Ibáñez-Marcelo and T. Alarcón. The topology of robustness and evolvability in evolutionary systems with genotype-phenotype map. *J. Theor. Biol.*, 356:144–162, 2014.

[53] Y. Iwasa, F. Michor, and M. A. Nowak. Evolutionary dynamics of escape from biomedical intervention. *Proc. Roy. Soc. B*, 270:2573–2578, 2003.

[54] Y. Iwasa, F. Michor, and M. A. Nowak. Evolutionary dynamics of invasion and escape. *J. Theor. Biol.*, 226:205–214, 2004.

[55] J. Jaeger and N. Monk. Bioattractors: dynamical systems theory and the evolution of regulatory processes. *J. Physiol.*, 592:2267–2281, 2014.

[56] S. A. Kauffman. *The origins of order*. Oxford University Press, New York, U.S.A., 1993.

[57] M. Kimmel and D. E. Axelrod. *Branching processes in Biology*, volume 19 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, NY, USA, 2002.

[58] H. Kitano. Cancer as a robust system: Implications for cancer therapy. *Nature Rev. Cancer*, 4:227–235, 2004.

[59] K. Klemm and V. M. Eguiluz. Growing scale-free networks with small-world behavior. *Physical Review E*, 65(5):057102, 2002.

[60] P. Krapivsky and S. Redner. Rate equation approach for growing networks. In R. Pastor-Satorras, M. Rubi, and A. Diaz-Guilera, editors, *Statistical Mechanics of Complex Networks*, volume 625 of *Lecture Notes in Physics*, pages 3–22. Springer Berlin Heidelberg, 2003.

[61] P. Krapivsky, S. Redner, and F. Leyvraz. Connectivity of growing random networks. *Phys. Rev. Lett.*, 85(21):4629–4632, 2000.

[62] P. L. Krapivsky and S. Redner. Organization of growing random networks. *Phys. Rev. E*, 63:066123, May 2001.

[63] P. L. Krapivsky, S. Redner, and E. Ben-Naim. *A kinetic view of statistical physics*. Cambridge University Press, 2010.

[64] R. Lambiotte, J. C. Delvenne, and M. Barahona. Laplacian Dynamics and Multiscale Modular Structure in Networks. *arxiv:0812.1770*, 2009.

[65] A. Lancichinetti and S. Fortunato. Community detection algorithms: a comparative analysis. *Phys. Rev. E*, 80(5):056117, 2009.

[66] K. Lewis. Persister cells, dormancy and infectious disease. *Nat. Rev. Microbiol.*, 5:48–56, 2007.

[67] D. Lipman and W. Wilbur. Modelling neutral and selective evolution of protein folding. *Proc. Roy. Soc. B*, 245:7–1, 1991.

[68] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Struct. Alg.*, 6:161–180, 1995.

[69] M. Molloy and B. Reed. The size of the giant component of a random graph with a given degree sequence. *Combin. Probab. Comput.*, 7:295–305, 1998.

[70] P. A. P. Moran. The statistical processes of evolutionary theory. *The statistical processes of evolutionary theory.*, 1962.

[71] W. Ndifon, J. B. Plotkin, and J. Dushoff. On the accessibility of adaptive phenotypes of a bacterial metabolic network. *PLoS Comp. Biol.*, 5:e1000472, 2009.

[72] M. Newman, S. Strogatz, and D. Watts. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, 64(2):026118, 2001.

[73] M. Newman and D. Watts. Scaling and percolation in the small-world network model. *Phys. Rev. E*, 60(6):7332–7342, Dec. 1999.

[74] M. E. J. Newman. Clustering and preferential attachment in growing networks. *Phys. Rev. E*, 64:025102, Jul 2001.

[75] M. E. J. Newman. *Networks: An introduction.* Oxford University Press, Oxford, UK, 2010.

[76] M. E. J. Newman and D. J. Watts. Renormalisation group analysis of the small-world network model. *Phys. Rev. Lett.*, 260:341–346, 1999.

[77] T. Opsahl and P. Panzarasa. Clustering in weighted networks. *Social networks*, 31(2):155–163, 2009.

[78] O. Perron. Zur theorie der matrices. *Mathematische Annalen*, 64(2):248–263, 1907.

[79] M. Piggliucci. Is evolvability evolvable? *Nat. Rev. Genet.*, 9:75–82, 2008.

[80] J. F. Rodrigues and A. Wagner. Evolutionary plasticity and innovations in complex metabolic reaction networks. *PLoS Comp. Biol.*, 5:e1000613, 2009.

[81] J. F. Rodrigues and A. Wagner. Genotype networks in sulfur metabolism. BMC Systems Biology. *BMC Syst. Biol.*, 5:39, 2011.

[82] A. Roesch, M. Fukunaga-Kalabis, E. C. Schmidt, S. E. Zabierowski, P. A. Brafford, A. Vultur, D. Basu, P. Gimotty, T. Vogt, and M. Herlyn1. A Temporarily Distinct Subpopulation of Slow-Cycling Melanoma Cells Is Required for Continuous Tumor Growth. *Cell*, 141:583–594, 2010.

[83] S. Rutherford, S. Lindquist, et al. Hsp90 as a capacitor for morphological evolution. *Nature*, 396(6709):336–342, 1998.

[84] S. Sagitov and M. C. Serra. Multitype Bienayme-Galton-Watson processes escaping extinction. *Adv. Appl. Prob.*, 41:225–246, 2009.

[85] A. Samal, J. F. Rodrigues, J. Jost, O. C. Martin, and A. Wagner. Genotype networks in metabolic reaction spaces. *BMC Syst. Biol.*, 4:30, 2010.

[86] S. A. Sawyer, J. Parsch, Z. Zhang, and D. L. Hartl. Prevalence of positive selectio among nearly neutral amino acid replacement in *Drosophila*. *Proc. Natl. Acad. Sci. USA*, 104:6504–6510, 2007.

[87] M. T. Schaub, J. C. Delvenne, S. N. Yaliraki, and M. Barahona. Markov Dynamics as a Zooming Lens for Multiscale Community Detection: Non Clique-Like Communities and the Field-of-View Limit. *PLoS One*, 7:e32210, 2010.

[88] P. Schuster, W. Fontana, P. Stadler, and I. Hofacker. From sequences to shapes and back: a case study in RNA secondary structures. *Proc. Roy. Soc. B*, 255:279–284, 1994.

[89] M. C. Serra. On the wainting time to escape. *J. Appl. Prob.*, 43:296–302, 2006.

[90] M. C. Serra and P. Haccou. Dynamics of escape mutants. *Theor. Popul. Biol.*, 72:167–178, 2007.

[91] M. Á. Serrano and M. Boguñá. Clustering in complex networks I: General formalism. *Phys. Rev. E*, 74:056114, 2006.

[92] M. Á. Serrano and M. Boguñá. Clustering in complex networks II: Percolation properties. *Phys. Rev. E*, 74:056115, 2006.

[93] S. V. Sharma, D. Y. Lee, B. Li, M. P. Quinlan, F. Takahashi, S. Maheswaran, U. McDermott, N. Azizian, L. Zou, M. A. Fischbach, K.-K. Wong, K. Brandstetter, B. Wittner, S. Ramaswamy, M. Classon, and J. Settleman. A Chromatin-Mediated Reversible Drug-Tolerant State in Cancer Cell Subpopulations. *Cell*, 141:69–80, 2010.

[94] M. Siegal and A. Bergman. Waddington's canalization revisited: developmental stability and evolution. *Proc. Natl. Acad. Sci. USA*, 99(16):10528, 2002.

[95] B. W. Silverman. *Density estimation for statistics and data analysis*, volume 26. CRC press, 1986.

[96] M. Soskine and D. S. Tawfik. Mutational effects and the evolution of new protein functions. *Nat. Rev. Genet.*, 11:572–582, 2010.

[97] D. Stauffer and A. Aharony. *Introduction to percolation theory.* Taylor and Francis, 1991.

[98] T. Tian, S. Olson, J. M. Whitacre, and A. Harding. The origins of cancer robustness and evolvability. *Integrative Biology*, 3:17–30, 2011.

[99] C. Villarreal, P. Padilla-Longoria, and E. Alvarez-Buylla. General Theory of Genotype to Phenotype Mapping: Derivation of Epigenetic Landscapes from N-Node Complex Gene Regulatory Networks. *Phys. Rev. Lett.*, 109:118102, 2012.

[100] V. von Noort, B. Snel, and M. A. Huynen. The yeast coexpression network has a small-world, scale-free architecture and can be explained by a simple model. *EMBO Reports*, 5:280–284, 2004.

[101] C. H. Waddington. Canalization of development and the inheritance of acquired characters. *Nature*, Vol. 150, No. 3811:563–565, 1942.

[102] C. H. Waddington. The epigenotype. *Endeavour*, 1:18–20, 1942.

[103] A. Wagner. Evolution of gene networks by gene duplications: a mathematical model and its implications on genome organization. *Proc. Natl. Acad. Sci. USA*, vol. 91 no. 10:4387–4391, 1994.

[104] A. Wagner. Does evolutionary plasticity evolve? *Evolution*, 50(3):1008–1023, 1996.

[105] A. Wagner. *Robustness and evolvability in living systems.* Princeton University Press Princeton, NJ:, 2005.

[106] A. Wagner. Neutralism and selectionism: a network-based reconciliation. *Nat. Rev. Genet.*, 9:965–974, 2008.

[107] A. Wagner. Genotype networks shed light on evolutionary constraints. *Trends Ecol. Evol.*, 26:577–583, 2011.

[108] A. Wagner. The role of robustness in phenotypic adaptation and innovation. *Proc. Roy. Soc. B*, 279:1249–1258, 2012.

[109] D. J. Watts and S. H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393(6684):440–442, June 1998.

[110] H. S. Wilf. *generatingfunctionology.* Academic Press, Inc., 2 edition, 1994.

[111] S. Wright. Evolution in mendelian populations. *Genetics*, 16:97–159, 1931.

[112] S.-H. Yook, Z. N. Oltvai, and A. L. Barabasi. Functional and topological characterization of protein interaction networks. *Proteomics*, 4:928–942, 2004.