# UNIVERSITÀ DEGLI STUDI DI PADOVA 

Dipartimento di Fisica e Astronomia "Galileo Galilei"
Corso di Laurea Magistrale in Fisica

Tesi di Laurea

Emerging ecological patterns from optimal investment strategies in a randomly fluctuating environment

| Relatore | Laureando |
| :--- | :--- |
| Prof. Amos Maritan | Stefano Garlaschi |
| Correlatore |  |
| Dott. Marco Formentin |  |

# UNIVERSITÀ DEGLI STUDI DI PADOVA 

## Dipartimento di Fisica e Astronomia "Galileo Galilei"

Corso di Laurea Magistrale in Fisica

Tesi di Laurea

## Emerging ecological patterns from optimal investment

 strategies in a randomly fluctuating environmentRelatoreProf. Amos Maritan:CorrelatoreDott. Marco Formentin:
LaureandoStefano Garlaschi:
$\qquad$

## Contents

Introduction ..... 1
1 Taylor's Law in ecology ..... 5
1.1 Formulation and empirical evidences ..... 5
1.2 Open questions ..... 8
1.3 Large Deviation Theory for TL and the role of rare events ..... 8
1.3.1 The model of the paper ..... 10
1.3.2 Model results and possible explanations ..... 13
2 Introduction of an adaptive strategy in the population dy- namics ..... 17
2.1 Model dynamics ..... 18
2.1.1 Population growth ..... 19
2.2 Large Deviations Principle for the model ..... 20
2.3 Study of the generalized Taylor's Law exponent ..... 22
3 Population dynamics with two possible evolutionary strate- gies ..... 29
3.1 Model dynamics ..... 30
3.1.1 Population growth ..... 31
3.2 Study of the generalized Taylor's Law exponent ..... 32
3.3 The optimal investment strategies ..... 32
3.3.1 Maximization based on the most probable realizations of the chain ..... 33
3.3.2 Maximization considering also the contribution of rare events ..... 40
3.4 Fluctuations and extinction risk ..... 44
3.4.1 Evaluation of the fluctuations ..... 44
3.4.2 Markowitz optimization: growth and risk minimization ..... 45
4 Population dynamics with adaptive strategy and a randomly moving fraction of individuals ..... 53
4.1 Model dynamics ..... 53
4.1.1 Population growth ..... 54
4.1.2 Study of the generalized Taylor's Law exponent ..... 56
4.2 The optimal investment strategies ..... 57
4.2.1 Maximization based on the most probable realizations of the chain ..... 57
4.2.2 Maximization considering also the contribution of rare events ..... 60
4.2.3 Markowitz optimization ..... 67
Conclusions ..... 73
A Large Deviations Theory ..... 75
A. 1 Ergodic Theorem ..... 75
A. 2 Rate Function and Large Deviation Principle for the sample average ..... 76
A. 3 Gärtner-Ellis Theorem ..... 77
A. 4 Varadhan's Lemma ..... 78
A. 5 Contraction Principle ..... 79
Bibliography ..... 81

## Introduction

Systems that can be referred as complex are systems with dense and various networks of interactions among their components with the presence of collective behaviours. In this sense we can say complex systems are described by Aristotle's quote "the whole is more than the sum of its parts": the presence of non-trivial properties in these systems, such as non-linear behaviours in the response to external inputs, self organization, adaptation and feedback loops, is difficult, sometimes even impossible, to infer from the details of the constituent parts and their interactions.
In the science of complex systems the emergence of patterns showing the same statistical regularities is very interesting because it may indicate that despite of different details the key features of the dynamics are in common and are responsible for the emergence of shared recurrent motifs.
This is the case of the so called Taylor's Law (TL), also known as fluctuation scaling. It relates the variance of a certain quantity of interest $X$ to its mean via power-law, i.e.

$$
\operatorname{Var}[X]=a \mathbb{E}[X]^{b} .
$$

According to the considered system $X$ is different, in the sense that the quantities of interest change from system to system because each one describes a specific phenomenon with its proper dynamics and interaction network. Nevertheless data coming from a large class of heterogeneous systems suggest that the fluctuation scaling behaviour is quite an universal feature of the complexity.
In fact evidences of Taylor's Law validity can be found across fields and disciplines studying complex systems that seems to have nothing in common. For example this power-law relation appeared in life sciences studies. Researchers of the field computed the cell counts for several individuals of the same species and they found that the average value and the variance are related via power-law [4]. A similar work was carried out for the number of tumor cells in pulmonary metastasis in a group of mice finding the same scaling property [5].
Taylor's power-law was detected also in sequences of DNA. Data shown that the average and the variance of the numbers of Single Nucleotide Polymorphisms, which are caused by the contributions of several factors during the mechanisms of transcription and duplication, scale with the length of the DNA sequence, hence the variance scales via power-law with the mean [6]. In epidemiology TL turns out to describe quantitatively the spread of a disease: some epidemiologists analyzed the data sets of the numbers of
registered cases of measles in different communities of England and Wales from 1944 to 1997. They found a power law relationship between the average number of cases and its variance. They also divided the time window in three different eras (depending on the diffusion of vaccinations) and discovered how the exponent decreased with the diffusion of vaccinations [7]. Hence they could quantify the impact of vaccinations against the spread of the virulent agent causing this illness. Similar works were conducted for the investigations of other disease diffusions (for example HIV [8]) or for describing the spacial clustering of cancers across the population (studies regarding the distribution of children affected by leukemia in a restricted urban area [9]).
Other deeply studied complex systems are stock markets. From the large set of available data for different markets it was possible to see that the average and the variance over time of the activity of a certain stock, defined as the number of transactions performed within a certain time window, are related again by a power-law [10].
A similar behaviour has emerged also in the Information Technology: computing the average and the variance of data flow into the nodes of an Internet network once again a power-law scaling was found [11]. Moreover this relationship holds also for other similar systems where the quantity of interest is the flow of different objects between the different nodes of the network composing the system, such as web page visits, river flow, microchip logical gates, highway traffic [12] or blood flow heterogeneity in the cardio-vascular system [13].
Additionally this power-law scaling property was found in other interesting human dynamics: data sets covering a year of the emails sent by different users shown how the average and the variance of the number of emails sent in the certain time interval interval are related by a power-law [18].
Many other several examples can be provided ([14]-[17]). The point is that every system considered can be regarded as a complex system because the huge quantity of factors that have to be accounted in the description of its dynamics and the interactions among its parts. Although these examples refer to completely different phenomena, it seems to hold an universal organization of their dynamics in the same common way making Taylor's power law to emerge across all these fields.
Ecological systems are among the most investigated real world complex phenomena and they will be the topic analyzed in this thesis. Their complexity is due to the huge number of individuals interacting among themselves, within the same species or in competition with other species, and with the environment they are living in. Moreover the environmental conditions, because of external agents, may change over time affecting the growth of a population and sometime causing a deep reorganization of the ecosystem to a new set of species (i.e. some species got extinct or new species migrate into the system) and thus to a different interaction network. In spite of such a complexity the analysis of empirical data confirm that recurrent patterns emerge from the dynamics of different ecosystems.
In this scientific area Taylor's Law is a very corroborated empirical law
which applies to populations. It relates the variance of the population size to its mean via power-law, i.e.

$$
\operatorname{Var}[N]=a \mathbb{E}[N]^{b},
$$

where $N$ denotes the population of a species.
Since deterministic predictions on the evolution of the population size are not always feasible it is sometimes useful, in order to perform theoretical studies, to model the time evolution of the population as a stochastic process.
One of the reasons why, since its formulation by the biologist R.L. Taylor in 1961 [1], Taylor's Law continues to be object of great interest is the evidence of only a restricted interval of values of the exponent $b$ is assumed. In particular, for most of the data sets analyzed, the scaling exponent falls between 1 and 2 with a clustering around the value $b=2$, independently of the ecosystem.
So Taylor's Law seems to be an ecosystem-independent feature and this have led researchers to believe that an underlying common mechanism exists for its emergence. Various models has been proposed but at the moment there is no agreement on which such a mechanism could be. In particular it has been recently claimed that the restricted range of values observed for the scaling exponent may be a statistical artifact, not related to the ecosystem dynamics. To the best of our knowledge none of these models accounted for evolutionary strategies.
The aim of this thesis is to explore through mathematical modeling the role of evolutionary strategies and optimizations in shaping the Taylor's exponent for population evolving in a stochastic environment. Indeed species adapt and Taylor's Law features may be the result of optimization through evolutionary strategies.
The plan of the thesis is as follows.
In Chapter 1 we will discuss Taylor's Law and its generalized form for the moments of the random variable in ecology. We will briefly highlight their importance, the empirical evidences supporting them together with some open questions. Then we will introduce our main modeling framework for the population growth: a multiplicative model in a Markovian environment. This model has been first proposed in [22] and resumed in [41], where using a Large Deviations argument it is argued that Taylor's exponent clustering around 2 may be a statistical artifact due to under sampling. This model does not account for evolutionary strategies.
Different simple strategies will be introduced in Chapters 2, 3 and 4 . There we will consider different evolutionary lines adopted by the population and optimal investment strategies with the purpose of investigating their effects on the emergence of the Taylor's Law. When possible, we will search analytical results, otherwise numerical calculations will be implemented. Our analysis show that for each strategy analyzed and for each optimization performed there is the possibility for the exponent to assume any real value, in contrast with the empirical observation and so confirming the conclusions in [41].

## Chapter 1

## Taylor's Law in ecology

When we study ecological dynamics it is interesting to look for some statistical regularities shared among different ecosystems.
One of the most investigated patterns in the ecological dynamics is the so called Taylor's Law (TL), named after the British ecologist Lionel Roy Taylor (1924-2007), who formulated the law in 1961 [1]. TL was firstly found in ecological studies but early it started to pop up evidences of its validity across different fields and disciplines.
In this chapter we will follow closely the review of TL in [2].

### 1.1 Formulation and empirical evidences

Due to its diffusion across different disciplines, starting with an abstract point of view may be more useful. TL states that for a non-negative random variable $X$ we have

$$
\begin{equation*}
\operatorname{Var}[X]=a \mathbb{E}[X]^{b}, \tag{1.1}
\end{equation*}
$$

with $a>0$ and $b \in \mathbb{R}$, i.e. the variance scales as a power-law of the mean with exponent $b$. For this reason TL is also known as fluctuation scaling law.
TL can be also extended to the moments of the random variable and it is referred as generalized Taylor's Law. It establishes the relationship between the $k$ th moment with $j$ th one to be a power law with exponent $b_{j k}$, i.e.

$$
\begin{equation*}
\mathbb{E}\left[X^{k}\right]=a_{j k} \mathbb{E}\left[X^{j}\right]^{b_{j k}} \tag{1.2}
\end{equation*}
$$

In the following, we will look at the case $k=2$ and $j=1$, i.e the power law binding $\mathbb{E}\left[X^{2}\right]$ and $\mathbb{E}[X]$ because also in this case the exponent $b_{21}$ is a measure of the spread from the expectation value.
When Taylor formulated the law he was studying ecosystems and so the variable he was analyzing was the population size or density in a specific area. In ecology, TL states that the variance of a species population size or density (in the sense of population per unit of area) $N$ behaves like a power-law of its mean, i.e.

$$
\begin{equation*}
\operatorname{Var}[N]=a \mathbb{E}[N]^{b}, \tag{1.3}
\end{equation*}
$$

with $a>0$ and $b \in \mathbb{R}$.
Now $N$ can be regarded as a random variable since it is an evolving quantity because of the interactions between the individuals, with other species and because of the fluctuations of the environment which is not static in time or homogeneous in the space. Since the complexity of this network of interactions and the difficulties to taking into account all the features of the environment it is not possible to evaluate $N$ in a deterministic way, but we have to look at it as a stochastic object.
It is not possible to know from empirical observations the distribution of this variable, for this reason empirical studies have to introduce estimators of its expectation value and variance in a different way. So to compute them empirically we have to perform $l$ independent measurements $N_{j}$ of $N$ and in this way we can obtain

$$
\begin{gathered}
m=m_{l}=\frac{1}{l} \sum_{j=1}^{l} N_{j} \\
\sigma^{2}=\sigma_{l}^{2}=\frac{1}{l} \sum_{j=1}^{l}\left(m^{2}-N_{j}\right)^{2}
\end{gathered}
$$

Taking a large number $l$ of measurements these estimators become more and more accurate and close to the real expectation values and variance. Depending on how this statistics is computed, TL can emerge in two different ways:

- over time, in the sense that the variance and the mean are computed over the time and on the same site, so TL tells us how big the fluctuations in time are starting from the expectation value;
- over the space, i.e. the variance and the mean are computed over the space and so in this case TL quantifies the dispersion of the individuals and how much they are spread.

So with the former case we can study the evolution in time the populations and most importantly we can compute the fluctuations which are dangerous for the survival of the population since with large fluctuations extinction becomes more probable. Instead in the latter way to generate TL we quantify how the individuals are displaced in their habitat and so in this case we can have a measure of the aggregation among the individuals.
Taylor in 1961 published a paper entitled "Aggregation, variance and the mean" [1] in which he analyzed 24 data sets of population density, dated from 1936 to 1960, of different species living in heterogeneous type of environments (for example worms, birds or shellfish). He noticed that it was possible to relate the empirical variance $\sigma^{2}$ to the empirical mean $m$ of each data set using a power law in the form

$$
\sigma^{2}=a m^{b}
$$

In other words, Taylor from those data extrapolated the TL generating via space. He named originally this behaviour law of the mean, but soon it became known as Taylor's Law.
He looked at the factor $a$ as a parameter which depends on the size of the population which the data come from, thus having no physical meaning. He thought that only the exponent $b$ was worth to be studied since it can be regarded as the index of aggregation of the species due to its independence of the size of the sampling unit. In fact, thanks to the power law fort of TL, $b$ displays a scale invariance property.
Let us call $m$ and $\sigma^{2}$ the mean and variance of a certain population density $N$. Multiplying the population by a factor $c$, we have that $N$ goes in $c N$ and thanks to the linearity of the expectation value and the properties of the variance it holds for the mean and variance of scaled sampling unit:

$$
\begin{gathered}
m_{c}=\mathbb{E}[c N]=c \mathbb{E}[N]=c m \\
\sigma_{c}^{2}=\mathbb{E}\left[\left(c N-m_{c}\right)^{2}\right]=c^{2} \mathbb{E}\left[(N-m)^{2}\right]=c^{2} \sigma^{2} .
\end{gathered}
$$

From this scaling relationship we obtain

$$
\sigma_{c}^{2}=c^{2} \sigma^{2}=a c^{2} m^{b}=a_{c} m_{c}^{b}
$$

with $a_{c}=a c^{2-b}$. As we demonstrated, $b$ is invariant under scaling while $a$ change with the size of the sampling. For this reason $b$ can be regarded as index of aggregation since quantifies the spread of the individuals in the space independently of the size of the sampling unit.
After the publishing of the paper by Taylor, this law have started to be investigated more and more [19]-[24]. More data and in a more precise way have been collected and TL has always emerged becoming a very corroborated ecological pattern. For this statistical robustness TL has started to be considered like an universal regularity. In fact the data analyzed came from a large class of ecosystems with several difference and various interaction networks among the individuals (for example shellfish in the seashore [1],carabid beetles, birds and tree abundance in tropical forests or [22]).
TL is verified also to be stable for different interactions among the individuals. Researchers showed how TL exponent does not change when in bacteria culture, displaying TL when the individuals do not compete among themselves for the resources, a competition mechanisms is introduced [23].
Moreover TL with respect to time have started to be noted in some data sets collecting measurements of populations in time ([18], [28]). Evidences of TL also via time reinforced its universality and led researchers to claim the existence of an underlying common mechanisms that ruled ecological dynamics spatially and temporally.
So far we discussed TL universality but also evidences of generalized TL diffusion in a lot of different data sets can be found in nature [22].

### 1.2 Open questions

Due to its universality many theoretical efforts have started to be undertaken to explain TL. Models and methods exploiting different theoretical tools, from statistical mechanics to probability theory and statistics techniques, have started to be built and analyzed with the aim to understand the originating features of this recurrent pattern ([31]-[36]). In particular these works have tried to uncover the existence of the possible underlying common mechanisms that allows TL emerges across so many different scenarios and gives it its universal behaviour.
To answer this question researchers divided into two opposite lines of thinking:

- the first states the universality of TL to be a direct consequence of the similarity of the dynamics of the different ecosystems displaying the power law behaviour ([3], [18]);
- the second in contrast assert the power law relationship between variance and mean to be a pure statistical feature, hence without any physical meaning ([26], [27]).

Another puzzle about TL is the range of possible values for the exponent b. Since its importance (index of aggregation when TL generates via space or scale of the fluctuations with respect to the mean in time) trying to predict the possible values of the exponent of the power law is fundamental to understand the law.
From empirical studies and ecological measurements we see that the exponent falls into a restricted range of values. In particular $b$ is seen to belong to the interval $[1,2]$, clustering around $b \simeq 2$ [1], while the exponent of generalized TL is nearly the ratio of the order of the two moments, i.e. $b_{j k} \simeq k / j$. Evidences of this are shown in Figure 1.1 or Table 1.1.
In order to study theoretically the exponents population growth models and in particular multiplicative growth models, i.e. theoretical models describing population dynamics and its evolution, have been implemented ([37]-[40]). These predict that the exponents of TL and generalized TL can assume any real value.
So, to predict theoretically the range of values for the exponents, models are built up, but when their results are compared with the outcomes of the analysis of the measurements we do not see any agreement.

### 1.3 Large Deviation Theory for TL and the role of rare events

A possible explanation can be found in the work "Sample and population exponents of generalized Taylor's law" [41]. There the authors implemented a multiplicative growth model in a Markovian environment to explore the possible range of values for the generalized TL exponent $b_{12}$ relating via


Figure 1.1: Generalized TL manifestation in data referring to carabid beetles [41]. The plot is obtained using log scale in both axes and $\left\langle N^{k}\right\rangle$ is plotted as a function of $\langle N\rangle$ for different value of $k$.
The color and symbol code identifies data relative to the same year: 1961 (black open circles), 1962 (purple solid circles), 1963 (blue open squares), 1964 (green solid squares), 1965 (orange solid diamonds), and 1966 (red open diamonds).
The dashed black lines have slope $b_{1 k}=k$. This is an example of how the generalized TL exponent $b_{j k} \simeq k / j$.

| $j, k$ | $k / j$ | $b_{j k} \pm S E$ | $R^{2}$ |
| :---: | :---: | :---: | :---: |
| 1,2 | 2 | $2.14 \pm 0.12$ | 0.991 |
| 1,3 | 3 | $3.33 \pm 0.32$ | 0.973 |
| 1,4 | 4 | $4.54 \pm 0.58$ | 0.954 |
| 2,4 | 2 | $2.15 \pm 0.16$ | 0.984 |
| 2,3 | 1.5 | $1.57 \pm 0.07$ | 0.995 |
| 3,4 | 1.333 | $1.37 \pm 0.04$ | 0.997 |
| $1,1 / 2$ | 0.5 | $0.48 \pm 0.02$ | 0.997 |
| $1,1 / 4$ | 0.25 | $0.23 \pm 0.01$ | 0.993 |
| $1,2 / 3$ | 0.667 | $0.65 \pm 0.01$ | 0.999 |

Table 1.1: Generalized TL exponents coming from dataset of Black Rock Forest [22]. We can see again that $b_{j k} \simeq k / j$.
power-law the first and the second moment. Then they tried to explain why only $b_{12} \simeq 2 / 1$ is observed in nature, in contrast with the model predictions.

### 1.3.1 The model of the paper

The model aims to reproduce the dynamics of a population in a stochastic fluctuating environment. This and its time-evolution are described by a Markov chain $\left\{A_{t, t \geq 1}\right\}$ with state-space $\Gamma=\{0,1\}$ and a symmetric transition matrix

$$
P=\left(\begin{array}{cc}
1-\lambda & \lambda  \tag{1.4}\\
\lambda & 1-\lambda
\end{array}\right)
$$

with $0<\lambda<1$.
So the environment can assume two possible states and at each time it has a probability $\lambda$ of changing its state, while with probability $1-\lambda$ it remains in the same.
Now the population is modeled to be adapted only to one environmental state. Without loss of generality let us suppose that this state is the 0 state. This corresponds to a scenario in which the individuals can display only a phenotype, suited to one state but not adapt to survive in the other. So the population grows by a factor $r$ when the chain is in the state 0 , while it will be multiplied by a factor $s$ when the state 1 is assumed (in a natural situation $r>1$ since in the well-suited state the population can grow, while $0<s<1$ because in the state 1 the individuals are not able to survive and so their number has to shrink).
With this model we write the population size $N(t)$ up to time $t$ as

$$
\begin{equation*}
N(t)=N_{0} \prod_{n=1}^{t} \widetilde{A_{n}} \tag{1.5}
\end{equation*}
$$

where $N_{0}$ is the initial value and $\widetilde{A_{n}} \in\{r, s\}$ are the multiplicative factors. Thanks to the model dynamics

$$
\widetilde{A_{n}}=r \cdot \delta_{A_{n}, 0}+s \cdot\left(1-\delta_{A_{n}, 0}\right)= \begin{cases}r & \text { if } A_{n}=0  \tag{1.6}\\ s & \text { if } A_{n} \neq 0\end{cases}
$$

In this way the population is obtained as a multiplication of several factors $r$ or $s$ whose numbers depends on the specific realization of the chain. This dependence to $N(t)$ the character of a random variable. Hence we have

$$
\begin{equation*}
N(t)=N_{0} \prod_{n=1}^{t}\left[r \cdot \delta_{A_{n}, 0}+s \cdot\left(1-\delta_{A_{n}, 0}\right)\right] \tag{1.7}
\end{equation*}
$$

This equation becomes

$$
\begin{align*}
N(t) & =N_{0} \cdot r^{\sum_{n=1}^{t} \delta_{A_{n}, 0}} \cdot s^{\sum_{n=1}^{t}\left(1-\delta_{A_{n}, 0}\right)}= \\
& =N_{0} \cdot r^{t \cdot \frac{1}{t} \sum_{n=1}^{t} \delta_{A_{n}, 0}} \cdot s^{t \cdot \frac{1}{t} \sum_{n=1}^{t}\left(1-\delta_{A_{n}, 0}\right)}= \\
& =N_{0} \cdot r^{t x} \cdot s^{t(1-x)}= \\
& =N_{0} e^{t x \log r+t(1-x) \log s}=N_{0} e^{t G(x)} \tag{1.8}
\end{align*}
$$

where it is introduced

$$
\begin{equation*}
x=\frac{1}{t} \sum_{n=1}^{t} \delta_{A_{n}, 0} \tag{1.9}
\end{equation*}
$$

as the fraction of times the state 0 appears in a realization of the Markov chain up to time $t$ and

$$
\begin{equation*}
G(x)=x \log r+(1-x) \log s \tag{1.10}
\end{equation*}
$$

has the meaning of growth rate of the population.
Now let us introduce the empirical random measure

$$
\begin{aligned}
& L_{t}(\cdot): \Gamma \rightarrow[0,1] \\
& L_{t}(z)=\frac{1}{t} \sum_{n=1}^{t} \delta_{A_{n}, z} .
\end{aligned}
$$

that counts the fraction of times in a realization up to time $t$ the chain goes in the state $z \in \Gamma$. In the case of the model $z=0$ is considered. It is possible to demonstrate the family $P_{t}(x)=\mathbb{P}\left(L_{t}(0) \in[x, x+d x]\right)$ satisfies a Large Deviation Principle (LDP) on $\mathfrak{M}_{1}(\Gamma)$ with rate $t$ and rate function

$$
\begin{equation*}
I_{P}(x)=\sup _{u>0}\left[x \log \left(\frac{u_{1}}{(P u)_{1}}\right)+(1-x) \log \left(\frac{u_{2}}{(P u)_{2}}\right)\right] \tag{1.11}
\end{equation*}
$$

where $u \in \mathbb{R}^{2}$ is a positive two dimensional vector, i.e $u_{1}, u_{2}>0$.
Satisfying a LDP means that

$$
\begin{equation*}
\lim _{t \rightarrow+\infty} \frac{1}{t} \log \mathbb{P}\left(L_{t}(0) \in[x, x+d x]\right)=-I_{P}(x) \tag{1.12}
\end{equation*}
$$

for $x \in[0,1]$ and $d x$ is an infinitesimal interval.
Carrying on the calculation needed in Equation (1.11) it is possible to find
$I_{P}(x)=(x-1) \log \left\{1-\lambda\left[\frac{2(\lambda-1) x}{S_{\lambda}(x)-2 \lambda x}+1\right]\right\}-x \log \left[1-\frac{\lambda\left(S_{\lambda}(x)-2 x\right)}{2(\lambda-1) x}\right]$
with

$$
\begin{equation*}
S_{\lambda}(x)=\lambda+\sqrt{\lambda^{2}+8 \lambda(x-1) x-4(x-1) x} . \tag{1.13}
\end{equation*}
$$

In this form is more explicit the dependence of $I_{P}(x)$ on the transition matrix via the control parameter $\lambda$. We can see $I_{P}(x)$ having all the properties to be a rate function:


Figure 1.2: Plots of $I_{P}(x)$ for different values of $\lambda$ : in blue $\lambda=0.1$, in orange $\lambda=0.4$, in green $\lambda=0.6$ and in red $\lambda=0.9$.

- $I_{P}(x)$ is convex, i.e. $\frac{d^{2} I_{P}}{d x^{2}}>0$;
- $I_{P}(x)$ is non-negative, i.e $I_{P}(x) \geq 0 \forall x \in[0,1]$;
- $I_{P}(x)$ has a point of absolute minimum. This minimum is $x_{\text {min }}=\frac{1}{2}$ independently of $P$.

Some plots of $I_{P}(x)$ for different $\lambda$ are shown in Figure 1.2.
To compute the exponent $b$ the ratio between $t^{-1} \log \mathbb{E}\left[N^{2}(t)\right]$ and $t^{-1} \log \mathbb{E}[N(t)]$ is considered. The relation is looked for large $t$, i.e for $t \rightarrow+\infty$ so that the chain reaches the steady state.
With this limit, it is possible to compute $t^{-1} \log \mathbb{E}\left[N^{2}(t)\right]$ and $t^{-1} \log \mathbb{E}[N(t)]$ easily exploiting Large Deviation Theory. This in fact provides the so called Varadhan's Lemma (see Appendix A) which states

$$
\begin{equation*}
\lim _{t \rightarrow+\infty} t^{-1} \log \mathbb{E}\left[N^{k}(t)\right]=\sup _{x \in[0,1]}\left[k G(x)-I_{P}(x)\right] \tag{1.15}
\end{equation*}
$$

thanks to the relation $N^{k}(t)=e^{t k G(x)}$. In fact

$$
\begin{align*}
\lim _{t \rightarrow+\infty} \frac{1}{t} \log \mathbb{E}\left[N^{k}(t)\right] & =\lim _{t \rightarrow+\infty} \frac{1}{t} \log \mathbb{E}\left[e^{t k G(x)}\right]= \\
& =\lim _{t \rightarrow+\infty} \frac{1}{t} \log \int_{0}^{1} e^{t k G(x)} P_{t}(d x)= \\
& =\sup _{\mu \in[0,1]}[k G(x)-I(x)] . \tag{1.16}
\end{align*}
$$

So the authors obtained the exponent of the generalized Taylor's Law for the first and the second moment (we drop the subscripts when we look to $k=2$ and $j=1$ ) to be

$$
\begin{equation*}
b(\lambda)=\frac{\sup _{x \in[0,1]}\left[2 G(x)-I_{P}(x)\right]}{\sup _{x \in[0,1]}\left[G(x)-I_{P}(x)\right]} \tag{1.17}
\end{equation*}
$$



Figure 1.3: Regions of divergence for the exponent of the generalized TL. The small panels show $b(\lambda)$ for $r$ and $s$ in the different zones. Only in the highlighted gray regions ( $R_{1, \text { nostrat }}$ and $R_{2, \text { nostrat }}$ ) $\lambda_{c} \in[0,1]$ and so the exponent diverges.

More generally the exponent of the power-law relating the $k$ th to the $j$ th moment is

$$
\begin{equation*}
b_{j k}(\lambda)=\frac{\sup _{x \in[0,1]}\left[k G(x)-I_{P}(x)\right]}{\sup _{x \in[0,1]}\left[j G(x)-I_{P}(x)\right]} \tag{1.18}
\end{equation*}
$$

### 1.3.2 Model results and possible explanations

With the models the authors found an analytical expression for $b(\lambda)$ parametrized by $r$ and $s$. In particular they observed a division of the space of parameters as shown in Figure 1.3.
The division is due to the possibility of having a critical value of the probability of changing environment $\lambda_{c}$. Searching for the zero of the denominator of Equation (1.17) in fact

$$
\begin{equation*}
\log \left\{\frac{1}{2}\left[\left(1-\lambda_{c}\right)(r+s)+\sqrt{4\left(2 \lambda_{c}-1\right) r s+\left(\lambda_{c}-1\right)^{2}(r+s)^{2}}\right]\right\}=0 \tag{1.19}
\end{equation*}
$$

it holds

$$
\begin{equation*}
\lambda_{c}=\frac{1-r-s+r s}{2 r s-r-s} . \tag{1.20}
\end{equation*}
$$

The divergence occurs when $\lambda_{c} \in[0,1]$. Imposing that the regions of diver-
gence are

$$
\left\{\begin{array}{l}
R_{1, \text { nostrat }}=\left\{(r, s): r>1 \wedge 0<s<1 \wedge s<\frac{1}{r}\right\}  \tag{1.21}\\
R_{2, \text { nostrat }}=\left\{(r, s): s>1 \wedge 0<r<1 \wedge r<\frac{1}{s}\right\}
\end{array}\right.
$$

shown in Figure 1.3.
This demonstrates how the exponent can assume any value and not only bounded values near 2 as empirical evidences display.
The authors suggested that this disagreement between theoretical results and data analysis may be due to an under sampling and so $b \simeq 2$ could be only a statistical artifact, in the sense that it does not have any physical meaning. They proposed that in the data set there are not enough measurements to detect also the rare events (events with very low probability of being observed) that could have an huge impact on the right computation of the different moments and variance of $N(t)$.
In fact in Equation (1.17) $b$ is computed as the ratio of two suprema over $x \in[0,1]$. Now calling $b_{s}$ the sample exponent, i.e. the exponent coming from empirical data of the observed sampling unit, an agreement between $b$ and $b_{s}$ is possible only if the measurements contributing to $b_{s}$ included all the values in the possible range of $x$. That is Equation (1.17) gives a right estimation of $b_{s}$ only if all the values of $x \in[0,1]$ are detected by the measurements contributing to $b_{s}$ because only in this way the two suprema are correctly computed.
By definition, $x$ is the fraction of times up to the instant $t$ that the Markov chain $A_{n}$ goes in the state 0 in a realization. So the theoretical and empirical exponents can be equal only if in the sampling from which $b_{s}$ is obtained contains all the value of $x$. Thus we have to observe a number of trials of the chain large enough to observe the state 0 appears 0 times, once, twice, $\cdots, t-1$ times and $t$ times.
The problem now is that, as we said above, the probability of observing a fraction $x$ of times in a realization of $t$ step of the Markov chain satisfies a LDP with rate function (1.13) and so the probability to observe $x$ close to the extremes of the interval $[0,1]$ is exponentially low while increasing the number of steps $t$ during which we look at the chain. Thus to observe also the realizations of the chain with $x \simeq 0$ and $x \simeq 1$ we have to collect a very large number of measurements.
With this way of reasoning the authors suggested that empirically the two suprema are not computed over all the set of possible values of $x$, but on a restricted one due to the under sampling in the measurements procedure. More explicitly

$$
\begin{equation*}
b(\lambda)=\frac{\sup _{x \in[0,1]}\left[2 G(x)-I_{P}(x)\right]}{\sup _{x \in[0,1]}\left[G(x)-I_{P}(x)\right]} \neq b_{s}=\frac{\sup _{x \in\left[x_{-}, x_{+}\right]}\left[2 G(x)-I_{P}(x)\right]}{\sup _{x \in\left[x_{-}, x_{+}\right]}\left[G(x)-I_{P}(x)\right]} \tag{1.22}
\end{equation*}
$$

where $x_{-}$and $x_{+}$are the extremes of the experimentally observed interval of values for $x$ with $\left[x_{-}, x_{+}\right] \subset[0,1]$.
It the same work the authors gave an illustration on how the under sampling can be the reason of inefficiency of detecting the real value of $b$. The
sampling was reproduced by repeated direct simulations of the chain $A_{n}$. To observe the rare events in realizations of $t$ steps of the chain, the number $R$ of direct simulations of the Markovian dynamics needed as to be $R \gg e^{t}$. For this reason, due to the limited number of trials the one can perform to keep simulations feasible, for large $t$ direct simulations can not explore all the possible realizations of the chain and detect all the rare events needed to estimate precisely the generalized TL exponent $b_{s}$. The results supporting this idea are shown in Figures 1.4.


Figure 1.4: Generalized TL exponents $b_{12}$ and $b_{23}$ as a function of $\lambda$ for different values of the parameters $r$ and $s$ : in A and $\mathrm{C} r=2$ and $s=\frac{1}{4}$, hence we can see a divergence ( $r s<1$ ); instead in B and D $r=4$ and $s=\frac{1}{2}$ and the exponents don't display discontinuities ( $r s>1$ ).
The two curves have different meanings: the black continue line is the asymptotic population exponent that is the expression coming from Equation (1.17) and the dashed red line the sample exponent, i.e $k / j$. The black dots and the red squares are the sample exponents coming from the simulations in two different regimes of number of trials of the chain $R$ : for the black solid circles $1 \gg t \gg \log R\left(R=10^{6}\right.$ and $t=10$ ), i.e $R$ is large enough to detected also the low-probable realizations of the chain, while to find the red squares $t \gg \log R$ with large $t\left(R=10^{4}\right.$ and $\left.t=400\right)$, i.e in this case only a subset of possible realizations are explored and in particular the ones with highest probability.
The sample exponents are obtained via fit of $\log \mathbb{E}\left[N(t)^{k}\right]$ as a function of $\log \mathbb{E}\left[N(t)^{j}\right]$ for the last 6 (black dots) and 200 (red squares) respectively time steps. These two sets of points are obtained. Hence this panel shows the disagreement between the two regimes supporting the hypothesis of the under sampling as the source of discrepancy between $b$ and $b_{s}$.

## Chapter 2

## Introduction of an adaptive strategy in the population dynamics

We just said there is a possibility that $b_{j k} \simeq k / j$ may be not due to a some physical mechanisms but only to an accidental feature of ecological dynamics, supposing the model in 1.3.1 to be realistic.

Here the basic idea of this thesis comes: in fact in the model the population remains static and it is passively affected by the environment fluctuations, while in nature individuals perform a dynamical evolution facing the changing in their habitat and trying to be adapted as much as they can to the environment to survive and grow. So in this work we investigate the possibility of evolutionary strategies with optimal resources investments to be the mechanisms generating bounded values for TL and generalized TL introducing these ecological processes in the model dynamics.

So in this chapter we will show a new multiplicative growth model trying to reproduce the dynamics of a population living in a fluctuating environment. We call $N(t)$ the size of a population at the time $t$ and $N_{0}$ the initial value (in the following we will set $N_{0}=1$ without any lost of generality). Again we have

$$
\begin{equation*}
N(t)=\prod_{n=1}^{t} \widetilde{A}_{n} \tag{2.1}
\end{equation*}
$$

where $\widetilde{A}_{n} \in\{r, s\}$.
Instead of being passive like in the work presented in 1.3.1, the population now adopts an evolutionary strategy to face the changes in the environment states.

To accomplish the study of this model, we will exploit some techniques of the Large Deviations Theory (LDT) (see Appendix A).

### 2.1 Model dynamics

Let us consider a population of individuals living in a stochastic fluctuating environment whose evolution in time is described by a discrete-time Markov chain $\left\{A_{t, t \geq 0}\right\}$ that takes values in a two-dimensional state-space $\Gamma$. Let us call the two states 0 and 1 (so $\Gamma=\{0,1\}$ ). Let us introduce the transition matrix of this Markov chain $P$ and we assume it to be symmetric, i.e. in the form:

$$
P=\left(\begin{array}{cc}
1-\lambda & \lambda  \tag{2.2}\\
\lambda & 1-\lambda
\end{array}\right)
$$

with $0<\lambda<1$ and

$$
\begin{equation*}
p_{i j}=\mathbb{P}\left(A_{t+1}=j \mid A_{t}=i\right) \tag{2.3}
\end{equation*}
$$

We can look at parameter $\lambda \in(0,1)$ as the probability with which the environment changes state from a step to the next one.
Now the individuals can adopt two different phenotypes, labelled 0 and 1 , which are adapted respectively to the state 0 and 1 of the environment. This means that if at a certain time-step the environment assumes the state 0 the number of individuals with the phenotype 0 will grow by a factor $r$ and the number of the others with the phenotype 1 will be reduced by a factor $s$ (typically $r>1$ and $0<s<1$, but in the following we will consider the more possible generic case and so we ask only $r, s>0$ ). Again $r$ is the multiplicative factor when the phenotype matches the environment state and $s$ when this does not happen.
In this model we assume that all the individuals adopt the same phenotype at the same moment. We indicate the phenotype used at the time-step $t$ as $S_{t}$. Now we introduce a new dynamics feature with respect to the model presented above: the sequences of $S_{t}$, given a certain realization of the chain that rules the environment evolution, is fixed by the adaptive strategy. The phenotype $S_{t+1}$ that will be adopted at the step $t+1$ is the one adapted to the environment state $A_{t}$ at the step $t$,

$$
S_{t+1}=S_{t+1}\left(A_{t}\right)=A_{t} .
$$

In other words, with this strategy the population bets that the environment does not change in two consecutive steps of the chain and so the individuals will adopt for the next step the phenotype suited to the environment state they are currently living in. Here an example of what happens given a certain realization of the environment chain:

$$
\begin{array}{r}
t: 0121345 \cdots \\
A_{t}: 0011101 \cdots \\
S_{t}: x
\end{array}: 001110 \cdots \cdots .
$$

Now, the population increases, as we said above, only if the phenotype is the right one for the environment state, i.e. $A_{t+1}=S_{t+1}$, but with this strategy, since $S_{t+1}=A_{t}$ we have that the population increases by a factor
$r$ only if $A_{t+1}=A_{t}$, i.e. the environment does not change from a step to the next one. In this sense we can say that this strategy is Darwinian: the population grows only while it is adapted to its habitat, but when the environment changes the individuals are not more able to survive and so their number falls. This is intended to model the mechanism by which the natural selection operates.
So with this strategy the multiplicative factors $\widetilde{A_{n}}$ of Equation (2.1) we have

$$
\widetilde{A_{n}}=r \cdot \delta_{A_{n-1}, A_{n}}+s \cdot\left(1-\delta_{A_{n-1}, A_{n}}\right) \begin{cases}r & \text { if } A_{n-1}=A_{n}  \tag{2.4}\\ s & \text { if } A_{n-1} \neq A_{n}\end{cases}
$$

### 2.1.1 Population growth

We can now write an expression for the population according to the model dynamics. Given $N(t)$ we can find the size of the population at the time $t+1$ as

$$
\begin{equation*}
N(t+1)=N(t)\left[r \cdot \delta_{A_{t}, A_{t+1}}+s \cdot\left(1-\delta_{A_{t}, A_{t+1}}\right)\right] \tag{2.5}
\end{equation*}
$$

We call recursively Equation (2.5) and so we can write $N(t)$ as

$$
\begin{equation*}
N(t)=N_{0} \prod_{n=1}^{t} \widetilde{A_{n}}=N_{0} \prod_{n=1}^{t}\left[r \cdot \delta_{A_{n-1}, A_{n}}+s \cdot\left(1-\delta_{A_{n-1}, A_{n}}\right)\right] \tag{2.6}
\end{equation*}
$$

Now the situation simplifies a lot since the term in the square brackets of Equation (2.6) can be only equal to $r$ or $s$ thanks to the Kronecker delta as shown in Equation (2.4).
Exploiting this, we have that the size the population at the time $t$ can be expressed as $N_{0}$ multiplied by $r$ raised to the number of times that the environment did not change state in two consecutive steps of the chain $\left(\sum_{n=1}^{t} \delta_{A_{n-1} ; A_{n}}\right)$ and $s$ raised to the number of times that the environment changed state in two consecutive steps of the chain $\left(\sum_{n=1}^{t}\left(1-\delta_{A_{n-1} ; A_{n}}\right)\right)$. So we have

$$
\begin{align*}
N(t) & =N_{0} \cdot r^{\sum_{n=1}^{t} \delta_{A_{n-1}, A_{n}}} \cdot s^{\sum_{n=1}^{t}\left(1-\delta_{A_{n-1}, A_{n}}\right)}= \\
& =N_{0} \cdot r^{t \cdot \frac{1}{t} \sum_{n=1}^{t} \delta_{A_{n-1}, A_{n}}} \cdot s^{t \cdot \frac{1}{t} \sum_{n=1}^{t}\left(1-\delta_{A_{n-1}, A_{n}}\right)}= \\
& =N_{0} \cdot r^{t \mu} \cdot s^{t(1-\mu)}= \\
& =N_{0} \cdot \mathrm{e}^{t \mu \log r+t(1-\mu) \log s}= \\
& =N_{0} \cdot \mathrm{e}^{t G(\mu)}, \tag{2.7}
\end{align*}
$$

where we defined

$$
\begin{equation*}
\mu=\frac{1}{t} \sum_{n=1}^{t} \delta_{A_{n-1}, A_{n}} \tag{2.8}
\end{equation*}
$$

as the fraction of times the strategy wins the bet, i.e. the environment remains in the same state from a step to next one ( $1-\mu$ instead is the fraction of times the chain moves to a different state).

In the Equation (2.7) we introduced also the function $G(\mu)$ defined as

$$
\begin{equation*}
G(\mu)=\mu \log r+(1-\mu) \log s \tag{2.9}
\end{equation*}
$$

This function $G(\mu)$ has the meaning of the growth rate of the population and it is parametrized by $r$ and $s$.
Now from the definition in Equation (2.8) we can see that $\mu$ is not a deterministic object, but it depends on the Markov chain realization. Hence we need to study the probability of seeing each possible value of $\mu$ in order to compute the moments of the population size needed for the calculation of the generalized TL exponent.

### 2.2 Large Deviations Principle for the model

Let us introduce $\overline{L_{t}^{2}}$ as the random measure counting the fraction of times the environment chain $A_{n}$ does not change defined as

$$
\begin{equation*}
\overline{L_{t}^{2}}=\frac{1}{t} \sum_{n=1}^{t} \delta_{A_{n-1}, A_{n}} \tag{2.10}
\end{equation*}
$$

It is possible to demonstrate that the family $P_{t}(\mu)=\mathbb{P}\left(\overline{L_{t}^{2}} \in[\mu, \mu+d \mu]\right)$ satisfies a Large Deviation Principle and we will obtain its rate function. To do so, let us start from what it is known.
We introduce the empirical pair measure $L_{t}^{2}$ for Markov chain defined as

$$
\begin{equation*}
L_{t}^{2}=\frac{1}{t} \sum_{n=1}^{t} \delta_{\left(A_{n-1}, A_{n}\right)}, \tag{2.11}
\end{equation*}
$$

that counts the fraction of times each possible transition appears in a realization of the Markov chain up to time $t$. Since the chain satisfies

$$
\begin{align*}
& A_{t} \in \Gamma \subset \mathbb{N} \\
& \left\{A_{t, t \geq 0}\right\} \text { is Markov with transition matrix } P=\left(P_{i j}\right)_{i, j \in \Gamma}  \tag{2.12}\\
& P_{i j}>0, \forall i, j \in \Gamma,
\end{align*}
$$

the family $\left(P_{t}^{A}\right)$ defined as

$$
\begin{equation*}
P_{t}^{A}(\nu)=\mathbb{P}^{A}\left(L_{t}^{2} \in[\nu, \nu+d \nu]\right) \quad \text { with } \nu \in \widetilde{\mathfrak{M}}_{1}(\Gamma \times \Gamma) \tag{2.13}
\end{equation*}
$$

satisfies a LDP on $\widetilde{\mathfrak{M}}_{1}(\Gamma \times \Gamma)$ with rate $t$ and rate function

$$
\begin{equation*}
I_{P}^{2}(\nu)=\sum_{i, j \in \Gamma} \nu_{i j} \log \left(\frac{\nu_{i j}}{\bar{\nu}_{i} P_{i j}}\right) \tag{2.14}
\end{equation*}
$$

where $\bar{\nu}_{i}=\sum_{j} \nu_{i j}$ (see [42], Theorem $I V .3$ ). We introduced $\nu$ as a four components object

$$
\begin{equation*}
\nu=\left(\nu_{00}, \nu_{01}, \nu_{10}, \nu_{11}\right) \tag{2.15}
\end{equation*}
$$

where $\nu_{i j}$ counts the fraction of times the transition from $i$ to $j$ appears, i.e. $\nu_{00}, \nu_{01}, \nu_{10}$ and $\nu_{11}$ are the fraction of times in which we have a transition respectively from 0 to 0 , from 0 to 1 , from 1 to 0 and from 1 to 1 .
Explicitly Equation (2.14) becomes

$$
\begin{align*}
I_{P}^{2}(\nu) & =\nu_{00} \log \left[\frac{\nu_{00}}{\left(\nu_{00}+\nu_{01}\right) \cdot(1-\lambda)}\right]+\nu_{01} \log \left[\frac{\nu_{01}}{\left(\nu_{00}+\nu_{01}\right) \cdot \lambda}\right]+ \\
& +\nu_{10} \log \left[\frac{\nu_{10}}{\left(\nu_{10}+\nu_{11}\right) \cdot \lambda}\right]+\nu_{11} \log \left[\frac{\nu_{11}}{\left(\nu_{10}+\nu_{11}\right) \cdot(1-\lambda)}\right] \tag{2.16}
\end{align*}
$$

Now, because of the strategy used in the model, we are not interested in knowing which states the chain is moving from and to, but only if the chain is changing state or not. For this reason in our model we do not have to count separately $\nu_{00}$ from $\nu_{11}$ and $\nu_{01}$ from $\nu_{10}$ since the transitions they count provide the same contributions to the population growth. With this in mind we can notice that the quantities $\mu$ and $1-\mu$ introduced previously can be written in terms of $\nu_{i j}$ as

$$
\left\{\begin{array}{l}
\mu=\nu_{00}+\nu_{11}  \tag{2.17}\\
1-\mu=\nu_{01}+\nu_{10}
\end{array}\right.
$$

Using Equation (2.17), we can derive the rate function for $\overline{L_{t}^{2}}$ from the one in Equation (2.16) for $L_{t}^{2}$ applying the Contraction Principle (A.5. For further details see [42], Theorem III.20).
With the notation of the Theorem statement, we have

$$
\begin{gathered}
\mathcal{X}=\widetilde{\mathfrak{M}}_{1}(\Gamma \times \Gamma) \\
\mathcal{Y}=\widetilde{\mathfrak{M}}_{1}(\{(0,0) \cup(1,1),(0,1) \cup(1,0)\}) .
\end{gathered}
$$

and we apply the theorem to $P_{t}=\mathbb{P}\left(L_{t}^{2} \in[\nu, \nu+d \nu]\right)$, whose rate function $I_{P}^{2}(\nu)$ is shown in Equation (2.16) to obtain the rate function $I(\mu)$ of $Q_{t}=$ $\mathbb{P}\left(\overline{L_{t}^{2}} \in[\mu, \mu+d \mu]\right)$. The continuous map $T$ which from $\widetilde{\mathfrak{M}}_{1}(\Gamma \times \Gamma)$ goes to $\widetilde{\mathfrak{M}}_{1}(\{(0,0) \cup(1,1),(0,1) \cup(1,0)\})$ is

$$
\mu=\left(\mu_{A}, \mu_{B}\right)=T(\nu)=T\left(\nu_{00}, \nu_{01}, \nu_{10}, \nu_{11}\right)=\left\{\begin{array}{l}
\mu_{A}=\nu_{00}+\nu_{11}  \tag{2.18}\\
\mu_{B}=\nu_{01}+\nu_{10}
\end{array}\right.
$$

So it holds

$$
\begin{equation*}
I(\mu)=\inf _{\nu \in \tilde{\mathfrak{M}}_{1}(\Gamma \times \Gamma): \mu=T(\nu)} I_{P}^{2}(\nu) \tag{2.19}
\end{equation*}
$$

Computing the infimum with that constraint we obtain

$$
\begin{equation*}
I(\mu)=\mu_{A} \log \left(\frac{\mu_{A}}{1-\lambda}\right)+\mu_{B} \log \left(\frac{\mu_{B}}{\lambda}\right) \tag{2.20}
\end{equation*}
$$



Figure 2.1: Plots of $I(\mu)$ for different values of $\lambda$ : in blue $\lambda=0.1$, in orange $\lambda=0.4$, in green $\lambda=0.6$ and in red $\lambda=0.9$.
We can see that $I(\mu)$ is always non-negative, convex and it has a different absolute minimum depending on $\lambda, \mu_{\min }=1-\lambda$.

Since $1=\nu_{00}+\nu_{01}+\nu_{10}+\nu_{11}=\mu_{A}+\mu_{B}$, we have $\mu_{B}=1-\mu_{A}$. Thus we have that $I(\mu)$ depends only on one variable $\mu_{A}$ and with an abuse of notation we call this simply $\mu$. In this way Equation (2.20) becomes

$$
\begin{equation*}
I(\mu)=\mu \log \left(\frac{\mu}{1-\lambda}\right)+(1-\mu) \log \left(\frac{1-\mu}{\lambda}\right) \tag{2.21}
\end{equation*}
$$

In this way we demonstrate that $P_{t}(\mu)=\mathbb{P}\left(\overline{L_{t}^{2}} \in[\mu, \mu+d \mu]\right)$ satisfies a Large Deviation Principle on $\widetilde{\mathfrak{M}}_{1}(\{(0,0) \cup(1,1),(0,1) \cup(1,0)\})$ and we also found its rate function.
We can see from Figure 2.1 and mathematically demonstrate that this rate function has the needed good properties:

- $I(\mu)$ is convex, i.e. $\frac{d^{2} I}{d \mu^{2}}>0$;
- $I(\mu)$ is non-negative, i.e $I(\mu) \geq 0 \forall \mu \in[0,1]$;
- $I(\mu)$ has a point of absolute minimum that is $\mu_{\text {min }}=1-\lambda$.


### 2.3 Study of the generalized Taylor's Law exponent

Now we have all the tools to compute the exponent of the generalized TL

$$
\begin{equation*}
\mathbb{E}\left[N^{k}(t)\right]=a_{j k} \mathbb{E}\left[N^{j}(t)\right]^{b_{j k}}, \tag{2.22}
\end{equation*}
$$

focusing in the following only on the case $k=2$ and $j=1$ (so we will call $b=b_{12}$ ).

To calculate this exponent we follow the same path shown in 1.3.1. So we take the logarithm of both sides of Equation (2.22) and we multiply them by $t^{-1}$. Hence

$$
\begin{equation*}
b=\frac{t^{-1} \log \mathbf{E}\left[N^{2}(t)\right]}{t^{-1} \log \mathbf{E}[N(t)]} \tag{2.23}
\end{equation*}
$$

We apply Varadhan's Lemma (A.4) obtaining

$$
\begin{equation*}
b(\lambda)=\frac{\sup _{\mu \in[0,1]}[2 G(\mu)-I(\mu)]}{\sup _{\mu \in[0,1]}[G(\mu)-I(\mu)]} \tag{2.24}
\end{equation*}
$$

Carrying on the calculations to compute the two suprema in Equation (2.24), we find the explicit expression for the exponent

$$
\begin{equation*}
b(\lambda)=\frac{\log \left[(1-\lambda) r^{2}+\lambda s^{2}\right]}{\log [(1-\lambda) r+\lambda s]} \tag{2.25}
\end{equation*}
$$

In Figure 2.2 some plots of (2.25) with different $r$ and $s$ are shown. As we can see from Figures 2.2b and 2.2c, in some cases there is a divergence. This is due to the fact that there may be a critical value of $\lambda$, depending on $r$ and $s$, that we will call $\lambda_{c}$ in the following, belonging to the interval $[0,1]$ making the denominator of Equation (2.25) to go to zero.


Figure 2.2: Plots of $b(\lambda)$ for different values of $r$ and $s$ after the introduction of the adaptive strategy in the population dynamics.

To find an expression for $\lambda_{c}$ we have to solve the equation

$$
\begin{equation*}
\log [(1-\lambda) r+\lambda s]=0 \tag{2.26}
\end{equation*}
$$

finding

$$
\begin{equation*}
\lambda_{c}=\frac{r-1}{r-s} \tag{2.27}
\end{equation*}
$$



Figure 2.3: Regions of divergence for the exponent of the generalized TL after the introduction of the adaptive strategy in the population dynamics. The small panels show $b(\lambda)$ for $r$ and $s$ in the different zones. Only in the coloured regions $\lambda_{c} \in[0,1]$ making the exponent to diverge: in blue is shown $R_{1, \text { strat }}$ (the natural region of the parameters), while in orange $R_{2, \text { strat }}$ is shown.

As we said above the divergence occurs only if $\lambda_{c} \in[0,1]$ and so in the plot of $b(\lambda)$ we can see a divergence if the parameters $r$ and $s$ satisfy the inequality

$$
\begin{equation*}
0 \leq \frac{r-1}{r-s} \leq 1 \tag{2.28}
\end{equation*}
$$

In this way we find out that the regions of the parameter space with a divergence in the exponent of the generalized Taylor's Law $b_{12}$ are

$$
\left\{\begin{array}{l}
R_{1, \text { strat }}=\{(r, s): r>1 \wedge 0<s<1\}  \tag{2.29}\\
R_{2, \text { strat }}=\{(r, s): s>1 \wedge 0<r<1\}
\end{array}\right.
$$

These two regions are shown in Figure 2.3. As we can see unfortunately the first region is the one of interest, since it corresponds to the region of parameters with ecological meaning in which when the population is adapted it can grow ( $r>1$ ) and when the individuals phenotypes doesn't match the environment state their number falls $(0<s<1)$.
If we compare this model to the one presented in 1.3.1, we see that our model provides a bigger region of divergence. As matter of fact in the model in which the population does not play a strategy ([41]), the regions of the parameters $r$ and $s$ in which the exponent $b$ displays a divergence are
defined by

$$
\left\{\begin{array}{l}
R_{1, \text { nostrat }}=\left\{(r, s): r>1 \wedge 0<s<1 \wedge s<\frac{1}{r}\right\}  \tag{2.30}\\
R_{2, \text { nostrat }}=\left\{(r, s): s>1 \wedge 0<r<1 \wedge r<\frac{1}{s}\right\}
\end{array}\right.
$$

We see that the overall region of divergence in this case is smaller and in particular in the region of natural interest $(0<s<1<r)$ we can have some $r$ and $s$ that ensure the continuity in $\lambda$ of the exponent. In our model instead for all $(r, s)$ in this region $b$ displays a discontinuity.
By definition, $\lambda_{c}$ is the probability transition value for which the denominator of Equation (2.25) becomes null. So for $\lambda=\lambda_{c}$

$$
\begin{equation*}
\lim _{t \rightarrow+\infty} t^{-1} \log \mathbf{E}[N(t)]=0 \tag{2.31}
\end{equation*}
$$

This can be interpreted as an extinction of the population because it means that for great times $N(t)$ is not increased.
With the adaptive strategy the population is disadvantaged when $\lambda>\frac{1}{2}$ since in this situation is more probable that the environment changes from a step to the next one. So the individuals are more likely not to match their phenotype with the environment, hence the population can become extinct and so for there is a $\lambda_{c}$. Due to this fact in the region of divergence it is natural to expect to have a lot of points that give $\lambda_{c}>\frac{1}{2}$.
Nevertheless we have to consider that the changes in the environment in the natural scenarios are rare or at least the environment doesn't change so quickly. This situation is described by taking $\lambda<\frac{1}{2}$, i.e. it is more probable that the environment does not change than to observe a new state. In this case the adoption of an adaptive strategy can be a significant ecological move for the population because in this scenario the strategy can bring benefits for the growth. With this consideration we want to compare the two models in the natural region of the parameters.
For model without any strategy [41] the critical value of $\lambda$ is defined by

$$
\begin{equation*}
\lambda_{c, n o s t r a t}=\frac{r s-r-s+1}{2 r s-r-s}, \tag{2.32}
\end{equation*}
$$

while in the model with the strategy we have

$$
\begin{equation*}
\lambda_{c, s t r a t}=\frac{r-1}{r-s} . \tag{2.33}
\end{equation*}
$$

Searching the region $\left\{(r, s): r>1,0<s<1,0<\lambda_{c, \text { nostrat } / \text { strat }}<\frac{1}{2}\right\}$ we find

$$
\begin{equation*}
\overline{R_{1}}=\{(r, s): 1<r<2 \wedge 0<s<2-r\} \tag{2.34}
\end{equation*}
$$

in both cases. This region is shown in Figure 2.4. So considering a situation in which an adaptive strategy may be a good evolutionary move $\left(\lambda<\frac{1}{2}\right)$, the region of divergence is reduced by a lot, but this does not give any improvement with respect to the model illustrated in 1.3.1 because both models provides the same region of divergence. Hence by the introduction of a simple adaptive strategy the fundamental model predictions on the exponent


Figure 2.4: Region $\overline{R_{1}}$ of the parameter space for which $0<\lambda_{c}<\frac{1}{2}$.
range still do not match with empirical observation and actually under the natural considerations on the parameters discussed above it provides the same results.


Figure 2.5: Regions of divergence for the exponent of the generalized TL after the introduction of the adaptive strategy. The small panels show $b(\lambda)$ for $r$ and $s$ in the different zones. Only in the coloured regions $\lambda_{c} \in[0,1]$ making the exponent to diverge: in blue is shown $R_{1, \text { strat }}$ which contains the sub-region $\overline{R_{1}}$ in green, while in orange $R_{2, \text { strat }}$ is shown.

## Chapter 3

## Population dynamics with two possible evolutionary strategies

In the previous chapter we have studied a dynamics in which all the individuals of a population living in a stochastic environment play, in order to try to increase in time, the same evolutionary strategy encoding the Darwinian dynamics of the natural selection.
In nature we may not have that all the population uses the same strategy, but we can find individuals playing different evolutionary strategies in order to have the best way to grow. We can find an example of this idea in the models of financial markets and the portfolio optimization in which to have the best income in the long term we have to diversify the capital of investment in different assets. The idea now is to find a way to decide how much invest in each asset in order to find the optimal asset distribution of the capital.
To introduce this feature in ecological dynamics we implement a new multiplicative model dynamics in which the individuals can adopt two different strategies. So the population now has to choose how much individuals play the first or the second two strategy.
Biologically speaking a population has two requests: the former is to at least not to become extinct and the latter is to maximize in a certain way the growth.
The idea now is to find out how much the population is willing to invest its resources into the two strategies (i.e what are the fractions of individuals playing the first or the second strategy) looking at the possible optimization ways that try to accomplish these requests.
If we keep calling $N(t)$ the size of a population upon the time $t$ and $N_{0}$ (we can take it equal to 1 without any loss of generality) its initial value we will still have

$$
\begin{equation*}
N(t)=\prod_{n=1}^{t} \widetilde{A}_{n} \tag{3.1}
\end{equation*}
$$

but, as we will see, $\widetilde{A}_{n}$ takes values in a different set from the previous model.

### 3.1 Model dynamics

In this new dynamics we keep on looking at a population living in a stochastic environment whose evolution is described in the same way as in the previous model, i.e. the status of this environment evolves like a Markov chain $\left\{A_{t, t \geq 0}\right\}$ moving in a two dimensional state-space $\Gamma$ (we call these two states 0 and 1 again) with the same transition matrix $P(2.2)$.
Having two habitat states the individuals can adopt two different phenotypes as in the previous model, each one most adapted to one of the possible states (again we will call the two phenotypes 0 and 1 with the same notation and meaning as done above). Also in this case only the individuals with the right phenotype in use can grow (i.e. their number will be multiplied by a factor $r$, naturally $r>1$ ), while the others will be reduced by a factor $s$ (naturally $0<s<1$ ).
In this model we implement a two evolutionary strategies dynamics to determine which phenotype the different individuals will use: a fraction $p$ of the population will bet on the persistence of the environment in the same state (let us call this strategy $S_{1}$ ), i.e. these individual will bet that the environment does not change state from a time-step to the next one, while the remaining fraction $1-p$ will bet that the environment changes state from a step to the next one (let us call this strategy $S_{2}$ ) and so at the timestep $t+1$ they will assume the phenotype which was not adapted at the time-step $t$. With these two different strategies we have

$$
\begin{gathered}
S_{1, t+1}=S_{1, t+1}\left(A_{t}\right)=A_{t} \\
S_{2, t+1}=S_{2, t+1}\left(A_{t}\right)=\overline{A_{t}}=1-A_{t}
\end{gathered}
$$

where $\overline{A_{t}}=1-A_{t}$ we want to indicate the other state of $\Gamma$ different from $A_{t}$.
Here an example of what could happen with this dynamics given a certain realization of the environment chain:

$$
\begin{aligned}
& t: 012345 \cdots \\
& A_{t}: 001101 \ldots \\
& S_{1, t}: x 00110 \cdots \\
& S_{2, t}: x 11001 \cdots
\end{aligned}
$$

We can look at these two evolutionary lines as the usual Darwinian one ( $S_{1}$ ) and its exactly opposite $\left(S_{2}\right)$ : in fact the fraction playing $S_{1}$ increases the individuals by $r$ when $A_{t+1}=S_{1, t+1}$, but with this strategy $S_{1, t+1}=A_{t}$ and so we have a multiplication only if $A_{t+1}=A_{t}$, i.e the environment doesn't change and so this fraction of individuals keep on being adapted to their habitat and so able to grow; instead the other fraction using $S_{2}$ raises its users by $r$ when $A_{t+1}=S_{2, t+1}$, i.e. $A_{t+1} \neq A_{t}$. This last case means that this fraction of population bets on the changing in the environment state and so they will adopt the phenotype different from the right one at the present.

### 3.1.1 Population growth

So the population splits into

$$
\begin{equation*}
N(t)=\overbrace{p N(t)}^{S_{1}}+\overbrace{(1-p) N(t)}^{S_{2}} \tag{3.2}
\end{equation*}
$$

since now we have two sub-populations, the users of $S_{1}$ and $S_{2}$. Then the time moves from $t$ to $t+1$ and based on the strategies and on what the environment did, the numbers of individuals of the two sub-population will be multiplied by $r$ or $s$. Now the population size is $N(t+1)$ and it has to decide much it is invested on $S_{1}$ or $S_{2}$. We will keep fixed what are the fractions $p$ and $1-p$ playing the two evolutionary lines at each time-step:

$$
\begin{equation*}
N(t+1)=\overbrace{p N(t+1)}^{S_{1}}+\overbrace{(1-p) N(t+1)}^{S_{2}} \tag{3.3}
\end{equation*}
$$

We want now to write an expression for the population size at time $t$. Given $N(t)$ and remembering how the two strategies were defined we have

$$
\begin{align*}
N(t+1)= & p N(t)\left[r \cdot \delta_{A_{t}, A_{t+1}}+s \cdot\left(1-\delta_{A_{t}, A_{t+1}}\right)\right]+ \\
& +(1-p) N(t)\left[s \cdot \delta_{A_{t}, A_{t+1}}+r \cdot\left(1-\delta_{A_{t}, A_{t+1}}\right)\right] \tag{3.4}
\end{align*}
$$

In this way we find

$$
\begin{gather*}
\widetilde{A_{n}}=p\left[r \cdot \delta_{A_{n-1}, A_{n}}+s \cdot\left(1-\delta_{A_{n-1}, A_{n}}\right)\right]+(1-p)\left[s \cdot \delta_{A_{n-1}, A_{n}}+r \cdot\left(1-\delta_{A_{n-1}, A_{n}}\right)\right]= \\
= \begin{cases}\bar{r}=r p+s(1-p) & \text { if } A_{n-1}=A_{n} \\
\bar{s}=r(1-p)+s p & \text { if } A_{n-1} \neq A_{n}\end{cases} \tag{3.5}
\end{gather*}
$$

Since the fractions are kept fixed we iterate Equation (3.4) obtaining

$$
\begin{align*}
N(t)= & N_{0} \prod_{n=1}^{t}\left\{p\left[r \cdot \delta_{A_{n-1}, A_{n}}+s \cdot\left(1-\delta_{A_{n-1}, A_{n}}\right)\right]+\right. \\
& \left.+(1-p)\left[s \cdot \delta_{A_{n-1}, A_{n}}+r \cdot\left(1-\delta_{A_{n-1}, A_{n}}\right)\right]\right\}= \\
& =N_{0} \prod_{n=1}^{t}\left\{[r p+s(1-p)] \cdot \delta_{A_{n-1}, A_{n}}+[r(1-p)+s p] \cdot\left(1-\delta_{A_{n-1}, A_{n}}\right)\right\}= \\
& =N_{0} \cdot[r p+s(1-p)]^{\sum_{n=1}^{t} \delta_{A_{n-1}, A_{n}}} \cdot[r(1-p)+s p]^{\sum_{n=1}^{t} 1-\delta_{A_{n-1}, A_{n}}}= \\
& =N_{0} \cdot[r p+s(1-p)]^{t \cdot \frac{1}{t} \sum_{n=1}^{t} \delta_{A_{n-1}, A_{n}}} \cdot[r(1-p)+s p]^{t \cdot \frac{1}{t} \sum_{n=1}^{t} 1-\delta_{A_{n-1}, A_{n}}}= \\
& =N_{0} \cdot[r p+s(1-p)]^{t \mu} \cdot[r(1-p)+s p]^{t(1-\mu)}= \\
& =N_{0} \cdot \bar{r}^{t \mu} \cdot \bar{s}^{t(1-\mu)}= \\
& =N_{0} \cdot e^{t \mu \log \bar{r}+t(1-\mu) \log \bar{s}}= \\
& =N_{0} \cdot e^{t G(\mu)} \tag{3.6}
\end{align*}
$$

where we defined $\mu$ as done in Equation (2.8).
With these definitions we wrote the growth rate $G(\mu)$ as

$$
\begin{equation*}
G(\mu)=\mu \log \bar{r}+(1-\mu) \log \bar{s} \tag{3.7}
\end{equation*}
$$

which now is parametrized by $r, s$ and also $p$.

### 3.2 Study of the generalized Taylor's Law exponent

Also with the two-strategies dynamics the only thinks that matters is how many time the chain changes state. Thus the underlying random measure needed to compute the moments of the population size is the same as above and so we have to use in the following the same rate function obtained in 2.2. Looking for the exponent of the generalized Taylor's law

$$
\begin{equation*}
b=\frac{\log \mathbf{E}\left[N^{2}(t)\right]}{\log \mathbf{E}[N(t)]} \tag{3.8}
\end{equation*}
$$

we exploit Varadhan's Lemma, which gives us the exponent in the form

$$
\begin{equation*}
b(\lambda)=\frac{\sup _{\mu \in[0,1]}[2 G(\mu)-I(\mu)]}{\sup _{\mu \in[0,1]}[G(\mu)-I(\mu)]} \tag{3.9}
\end{equation*}
$$

Carrying on the calculation for (3.9) we finally find the explicit expression for the exponent

$$
\begin{equation*}
b(\lambda)=\frac{\log \left[(1-\lambda) \bar{r}^{2}+\lambda \bar{s}^{2}\right]}{\log [(1-\lambda) \bar{r}+\lambda \bar{s}]} \tag{3.10}
\end{equation*}
$$

We can notice that this is the same expression of (2.25), but with the substitution $r \longleftrightarrow \bar{r}$ and $s \longleftrightarrow \bar{s}$. This is due to the fact that in both case the population growth is ruled by the changing or the remaining of the environment in the same state, the differences between them are in the different multiplicative factors $\widetilde{A_{n}}$ of the two models.

### 3.3 The optimal investment strategies

Until now we have not fixed the value of $p$ yet. Now we want to find an expression for it and analyze the results of this choice. To do so we will look up three ways of optimization with different ecological objectives. This means that among all the possible values of $p$ the population organizes its dynamics in the best way possible choosing a certain value of $p$ (we call this the optimal $p^{*}$ ) to allow the accomplishment of a goal.
In the first place, we will maximize the growth rate on the long term considering only the most likely situation or looking at all the possible realizations (although they might have very low probability to happen). These two optimizations correspond to the application of two probability results: in the former case we appeal to the Law of Large Numbers (LLN), while in the
latter LDT and Varadhan's Lemma are used. As last optimization method we will try to minimize the entity of the fluctuations of $N(t)$, which are dangerous for the extinction of the population. We will try to do this imposing a growth, i.e. we will look for $p$ that gives the lowest risk of extinction but that at the same time can allow the population to grow in the long term.

### 3.3.1 Maximization based on the most probable realizations of the chain

We are looking for the optimal $p^{*}$ which maximizes the growth on the long term taking into account only the most probable trajectories of the environment evolution. Taking into account only the most probable realizations of the chain means to apply the LLN.
In fact we would like to maximize with respect to $p$ growth rate $G(\mu)$. Now $\mu$ is not fixed, but depend on the realization of the chain. By definition

$$
\begin{equation*}
\mu=\frac{1}{t} \sum_{n=1}^{t} \delta_{A_{n-1}, A_{n}} \tag{3.11}
\end{equation*}
$$

Now

$$
\mathbb{P}\left(A_{n-1}=A_{n}\right)=\sum_{i \in \Gamma} \mathbb{P}\left(A_{n-1}=i\right) \cdot p_{i i}=(1-\lambda) \sum_{i \in \Gamma} \mathbb{P}\left(A_{n-1}=i\right)=1-\lambda
$$

$\forall n \in \mathbb{N}$. Hence $\delta_{A_{n-1}, A_{n}}$ can be regarded as an i.i.d. random variable labeled by $n$ with expectation value equal to $1-\lambda$.
So LLN can be applied giving

$$
\begin{equation*}
\mu=1-\lambda \tag{3.12}
\end{equation*}
$$

almost surely in the limit of $t \rightarrow+\infty$ (that we consider since we want a maximization on the long term). We could also obtaining the same result noting that $\mu=1-\lambda$ is the point of absolute minimum of the rate function (2.21) for the family of probability $\mathbb{P}\left(\overline{L_{t}^{2}} \in[\mu, \mu+d \mu]\right)$, therefore the most probable value for large $t$.
Hence, Equation (3.7) becomes a function of $\lambda$ :

$$
\begin{equation*}
G(\lambda)=(1-\lambda) \log [r p+(1-p) s]+\lambda \log [r(1-p)+s p] . \tag{3.13}
\end{equation*}
$$

Looking for the maximum with respect to $p$ of Equation (3.13) we obtain the optimal investment fraction to be

$$
\begin{equation*}
p^{*}(r, s, \lambda)=\frac{r(1-\lambda)-\lambda s}{r-s} \tag{3.14}
\end{equation*}
$$

Since $0 \leq p \leq 1$, we have to set $p^{*}=1$ when Equation (3.14) is greater than 1 and $p^{*}=0$ when Equation (3.14) is smaller than 0 . To do so we have to distinguish the case $r>s$ and $r<s$.


Figure 3.1: Plot of $p^{*}(\lambda)$ optimized using LLN in the case of the two strategy dynamics with $r=4$ and $s=0.7$.

Case $r>s$
With this scenario we have $r-s>0$ and this leads us to

$$
p^{*}(r, s, \lambda)= \begin{cases}1 & \text { if } 0<\lambda \leq \frac{s}{r+s}  \tag{3.15}\\ \frac{r(1-\lambda)-\lambda s}{r-s} & \text { if } \frac{s}{r+s} \leq \lambda \leq \frac{r}{r+s} \\ 0 & \text { if } \frac{r}{r+s} \leq \lambda<1\end{cases}
$$

The plot of (3.15) with fixed $r, s$ will be as the one shown in Figure 3.1.

Case $r<s$
Now in Equation (3.14) $r-s$ changes sign and so we have

$$
p^{*}(r, s, \lambda)= \begin{cases}0 & \text { if } 0<\lambda \leq \frac{r}{r+s}  \tag{3.16}\\ \frac{r(1-\lambda)-\lambda s}{r-s} & \text { if } \frac{r}{r+s} \leq \lambda \leq \frac{s}{r+s} \\ 1 & \text { if } \frac{s}{r+s} \leq \lambda<1\end{cases}
$$

The plot of $p^{*}$ now is like the one show in Figure 3.2.
In both cases we see that there is an interval of $\lambda$ in which $0<p^{*}<1$, i.e there is a coexistence of individuals using the two different strategies, while near the extremes there is only a strategy used by the whole population. This is reasonable because near the extremes of the range for $\lambda$ the environment changes or not changes almost surely, so to maximize the growth all the individuals follow the same strategy because for these values of $\lambda$ one strategy is clearly better than the other.

## Plot of $b(\lambda)$ and divergences

Fixing $p=p^{*}$ we can insert this into the expressions of $\bar{r}$ and $\bar{s}$, which become functions of $\lambda$ parametrized by $r$ and $s$.


Figure 3.2: Plot of $p^{*}(\lambda)$ optimized using LLN in the case of the two strategy dynamics with $r=0.4$ and $s=1.6$

From the expression of the exponent shown in Equation (3.10) we can see that we may have divergences as the denominator becomes zero, i.e when

$$
\begin{equation*}
\lambda=\lambda_{c}=\frac{\bar{r}-1}{\bar{r}-\bar{s}} \tag{3.17}
\end{equation*}
$$

Since $\bar{r}$ and $\bar{s}$ are $\lambda$, we also have that $\lambda_{c}$ is a function of $\lambda$. For this reason we have to solve Equation (3.17) with the constraint $0 \leq \lambda \leq 1$ to find the regions of the space of parameters $r$ and $s$ in which the exponent $b$ of the generalized Taylor's law displays divergences.
In order to accomplish this we have to write explicitly $\lambda_{c}$ and for this reason we have to look at the case $r>s$ and $r<s$ separately since in these two different situations we have seen above that $p$ has two distinct behaviours. Let us start with the case $r>s$ and so we have

$$
\lambda_{c}=\frac{\bar{r}-1}{\bar{r}-\bar{s}}= \begin{cases}\frac{r-1}{r-s} & \text { if } 0<\lambda<\frac{s}{r+s}  \tag{3.18}\\ \frac{(r+s)(1-\lambda)-1}{(r+s)(1-2 \lambda)} & \text { if } \frac{s}{r+s}<\lambda<\frac{r}{r+s} \\ \frac{s-1}{s-r} & \text { if } \frac{r}{r+s}<\lambda<1\end{cases}
$$

Solving Equation (3.17) with $\lambda \in[0,1]$ we find two regions:

- $R_{1}=\left\{(r, s): \sqrt{-4 s^{2}+4 s+1}-2 r>-1,0<s<1, r>1\right\}$. In this region we have

$$
\begin{equation*}
\lambda_{c}=\frac{r-1}{r-s} \vee \lambda_{c}=\frac{1-s}{r-s} \tag{3.19}
\end{equation*}
$$

- $R_{2}=\left\{(r, s): \sqrt{-4 s^{2}+4 s+1}-2 r<-1, r+s<2,0<s<1, r>1\right\}$. In this region instead we have

$$
\begin{equation*}
\lambda_{c}=\frac{1}{2} \pm \frac{1}{2} \sqrt{\frac{2-r-s}{r+s}} . \tag{3.20}
\end{equation*}
$$

We can see that in the regions of divergence the exponent displays two divergences, symmetric to $\lambda=\frac{1}{2}$. These two regions are shown in Figure


Figure 3.3: Regions of the parameter space restricted to $r>1$ and $0<s<1$ in which $b$ displays divergences in the case of the two model dynamics with $p^{*}$ optimized using LLN. The blue zone represents region $R_{1}$ and in this zone the divergences are in the intervals of $\lambda$ in which $p^{*}=1$ or $p^{*}=0$, i.e. all the population plays the same strategy. The orange zone represents $R_{2}$ and in this zone the divergences appear in the interval of $\lambda$ in which there is a coexistence of both strategies, i.e. $0<p^{*}<1$ and there are individuals playing $S_{1}$ and others playing $S_{2}$.
3.3. The division of the region of divergence into two subregions is due to the possibility of having a divergence when only one strategy is used or when there is a coexistence of individuals using the two different evolutionary lines: $R_{1}$ corresponds to the former situation, while $R_{2}$ to the latter.
In Figure 3.4 are shown some plots of $b(\lambda)$ with different $r$ and $s$ with $r>s$. We can see that for $(r, s) \in R_{1}$ (see Figure 3.4b) and $(r, s) \in R 2$ (see Figure 3.4c) the exponent diverges twice as analytically found and discussed above. Now let us look at the case $r<s$. So we have

$$
\lambda_{c}=\frac{\bar{r}-1}{\bar{r}-\bar{s}}= \begin{cases}\frac{s-1}{s-r} & \text { if } 0<\lambda<\frac{r}{r+s}  \tag{3.21}\\ \frac{r+s)(1-\lambda)-1}{(r+s)(1-2 \lambda)} & \text { if } \frac{r}{r+s}<\lambda<\frac{s}{r+s} \\ \frac{r-1}{r-s} & \text { if } \frac{s}{r+s}<\lambda<1\end{cases}
$$

Looking for the solution of Equation (3.17) we find two regions:

- $R_{3}=\left\{(r, s): \sqrt{-4 r^{2}+4 r+1}-2 s>-1,0<r<1, s>1\right\}$. In this region we have

$$
\begin{equation*}
\lambda_{c}=\frac{r-1}{r-s} \vee \lambda_{c}=\frac{1-s}{r-s} . \tag{3.22}
\end{equation*}
$$



(e) $r=4$ and $s=2$.

Figure 3.4: Plots of $b(\lambda)$ in the case of the two model dynamics with $p^{*}$ optimized using LLN with different $r$ and $s(r>s)$.


Figure 3.5: Regions of the parameter space restricted to $0<r<1$ and $s>1$ in which $b$ displays divergences for the two model dynamics with $p^{*}$ optimized using LLN. The blue zone represents region $R_{3}$ and in this zone the divergences are in the intervals of $\lambda$ in which $p^{*}=1$ or $p^{*}=0$, i.e. all the population plays the same strategy. The orange zone represents $R_{4}$ and in this zone the divergences appear in the interval of $\lambda$ in which there is a coexistence of both strategies, i.e. $0<p^{*}<1$ and there are individuals playing $S_{1}$ and others playing $S_{2}$.

- $R_{4}=\left\{(r, s): \sqrt{-4 r^{2}+4 r+1}-2 s<-1, r+s<2,0<r<1, s>1\right\}$. In this region instead we have

$$
\begin{equation*}
\lambda_{c}=\frac{1}{2} \pm \frac{1}{2} \sqrt{\frac{2-r-s}{r+s}} . \tag{3.23}
\end{equation*}
$$

We can see that also in this case in the regions of divergence the exponent displays two divergences, symmetric to $\lambda=\frac{1}{2}$. These two regions are shown in Figure 3.5.
In Figure 3.6 are shown some plots of $b(\lambda)$ with different $r$ e $s$. We can see that for $(r, s) \in R 3$ (see Figure 3.6b) and for $(r, s) \in R 4$ (see Figure 3.6c) the exponent diverges.
We can now look at the parameter space globally and discuss the regions of divergence of the exponent $b$. These regions are shown in Figure 3.7. We can see that with this optimization of the growth Equation (3.17) doesn't have solution for $0<r, s<1$ or $r, s>1$ and so in this regions the exponent never diverges. In the region of natural interest (and its symmetric under exchange of $r$ and $s$ ) we still have some points causing divergence and in this case we have two values of $\lambda_{c}$ while the previous models gave only one


Figure 3.6: Plots of $b(\lambda)$ in the case of the two model dynamics with $p^{*}$ optimized using LLN with different $r$ and $s(r<s)$.


Figure 3.7: Regions of divergence for the generalized TL exponent in the case of the two strategies dynamics with $p^{*}$ optimized using LLN. The blue zone is $R 1$, the orange one is $R 2$, the green zone represents $R 3$ and the red zone is $R 4$. The small panels show $b(\lambda)$ for $r$ and $s$ in the different zones. As we can see in the coloured zones the exponent diverges.
critical value. On the other hand the situation is much improved since it is reduced a lot the region of the parameters in which the exponent displays a discontinuity.
We said that the two divergences are symmetric with respect to $\lambda=\frac{1}{2}$ and so in the region of divergence there is always $\lambda_{c}<\frac{1}{2}$. Thus, looking for the region the region $\left\{(r, s): r>1,0<s<1,0<\lambda_{c}<\frac{1}{2}\right\}$ because of its close correspondence to a real natural scenario, we automatically find $\{(r, s): r>1,0<s<1, r+s<2\}$, which is the same found in the previous model.

### 3.3.2 Maximization considering also the contribution of rare events

Above we maximized taking into account the Law of Large Numbers, i.e. we consider only the most probable trajectories of the environment states for large $t$. Now we want to consider also the rare events and to do so we have to use LDT.
For this reason we try to find the optimal investment fraction $p^{*}$ looking for the highest $\log \mathbb{E}[N(t)]$ (i.e. the highest $\mathbb{E}[N(t)]$ since the logarithm is a monotone function). In order to accomplish this we exploit Varadhan

Lemma again which gives us

$$
\begin{equation*}
\lim _{t \rightarrow+\infty} t^{-1} \log \mathbb{E}[N(t)]=\sup _{\mu \in[0,1]}[G(\mu)-I(\mu)]=\log [(1-\lambda) \bar{r}+\lambda \bar{s}] \tag{3.24}
\end{equation*}
$$

Using Equation (3.5) and writing explicitly Equation (3.24) we find

$$
\begin{equation*}
\log [(1-\lambda) \bar{r}+\lambda \bar{s}]=\log [s+\lambda \cdot(r-s)+p \cdot(r-s) \cdot(1-2 \lambda)] \tag{3.25}
\end{equation*}
$$

Now we try to maximize this quantity with respect to $p$ and to do so it is enough to find the maximum of the argument of the logarithm. The derivative is $(r-s) \cdot(1-2 \lambda)$, so the maximum is in the extremes of the interval of values allowed to $p$ (i.e. $p^{*}=0$ or $p^{*}=1$ ) depending on the sign of $r-s$ :

- if $r>s$ then

$$
p^{*}= \begin{cases}1 & \text { if } \lambda \leq \frac{1}{2}  \tag{3.26}\\ 0 & \text { if } \lambda>\frac{1}{2}\end{cases}
$$

- if $r<s$ then

$$
p^{*}= \begin{cases}0 & \text { if } \lambda \leq \frac{1}{2}  \tag{3.27}\\ 1 & \text { if } \lambda>\frac{1}{2}\end{cases}
$$

## Plot of $b(\lambda)$ and divergences

Once we fixed $p=p^{*}$ we can study $b(\lambda)$. In Figure 3.8 some plots of the exponent for different values of $r$ and $s$ are shown.
As done before we look for the potential regions of the parameters $r$ and $s$ of divergence of the exponent of generalized TL. The denominator of $b(\lambda)$ becomes zero for

$$
\lambda=\lambda_{c}=\frac{\bar{r}-1}{\bar{r}-\bar{s}}= \begin{cases}\frac{r-1}{r-s} & \text { if }\left(r-s>0 \wedge \lambda \leq \frac{1}{2}\right) \vee\left(r-s<0 \wedge \lambda>\frac{1}{2}\right)  \tag{3.28}\\ \frac{s-1}{s-r} & \text { if }\left(r-s>0 \wedge \lambda>\frac{1}{2}\right) \vee\left(r-s<0 \wedge \lambda \leq \frac{1}{2}\right)\end{cases}
$$

The regions in which $\lambda_{c} \in[0,1]$ are shown in Figure 3.9 and are defined by

$$
\left\{\begin{array}{l}
\widetilde{R_{1}}=\{(r, s): 1<r<2-s \wedge 0<s<1\}  \tag{3.29}\\
\widetilde{R_{2}}=\{(r, s): 0<r<2-s \wedge 1<s<2\}
\end{array}\right.
$$

We can see that these are the same regions obtained with the previous optimization, with the difference that there are not sub-divisions because in this case there is no more a coexistence of the two sub-populations, but only a strategy is actuated at the time by the whole population at each moment.
Due to the symmetry around $\frac{1}{2}$ of the two divergences, automatically we get the region $\left\{(r, s): r>1,0<s<1,0<\lambda_{c}<\frac{1}{2}\right\}$ is again defined by $\{(r, s): r>1,0<s<1, r+s<2\}$.


Figure 3.8: Plots of $b(\lambda)$ in the case of the two model dynamics with $p^{*}$ optimized using LDT techniques for different values of $r$ and $s$.


Figure 3.9: Regions of divergence for the generalized TL exponent in the case of the two strategies dynamics with $p^{*}$ optimized using LDT techniques to take into account even rare events. The blue zone represents $\widetilde{R_{1}}$. The orange one instead is $\widetilde{R_{2}}$. The small panels display some plots of $b(\lambda)$ for $r$ and $s$ in the different regions.

### 3.4 Fluctuations and extinction risk

Until now we asked for the fraction $p$ to maximize the growth on the long term of the population in two different ways (neglecting or considering rare events) without worrying about the possibility of extinction. But we said that a population in its dynamics would minimize the risk to become extinct. The possibility of extinction is related to the fluctuations on $\mathbb{E}[N(t)]$ so now we try to evaluate and minimize them. Then we wonder if this minimun risk condition can allow a growth of the population, i.e taking $p$ which minimize the fluctuations we wonder if it is possible to have a growth. If the answer is positive, this will be the best scenario ecologically speaking since we will have a low-risk dynamics allowing a growth. This kind of optimization in Economics and finance is called Markowitz optimization and again we exploit the similarity of our topic to the dynamics of market to study the model.

### 3.4.1 Evaluation of the fluctuations

Now we are going to evaluate the effect of the fluctuations on the expectation value $N(t)$ in two different ways and we will minimize the ratio of the them and the expectation value itself since we are interested to see if the the fluctuations are large with respect to the value of $\mathbb{E}[N(t)]$, i.e we will minimize with respect to $p$ the ratio

$$
\begin{equation*}
\frac{\sqrt{\operatorname{Var}[N(t)]}}{\mathbb{E}[N(t)]} \tag{3.30}
\end{equation*}
$$

So the first method is to compute exactly the two quantities involved. Let us consider the population in our model at the time-step $t$. In these $t$ steps the environment changed status from a step to the next one, let us say, $t-n$ times, while it didn't change the other $n$ times. In our model we have seen that when the environment doesn't change the population increases by a factor $\bar{r}=r p+(1-p) s$ and when it changes by a factor $\bar{s}=s p+(1-p) r$. So the population at the time $t$ will be:

$$
\begin{equation*}
N(t)=N_{0} \cdot \bar{r}^{n} \cdot \bar{s}^{t-n} \tag{3.31}
\end{equation*}
$$

The probability of having $n$ non-changing in the environment status in $t$ steps of the chain follows a binomial distribution with parameter $(1-\lambda)$

$$
\begin{equation*}
P(t ; n)=\frac{t!}{n!(t-n)!}(1-\lambda)^{n} \lambda^{t-n} \tag{3.32}
\end{equation*}
$$

We can easily compute the expectation value and the variance of $N(t)$ :

$$
\begin{equation*}
\mathbb{E}[N(t)]=\sum_{n=0}^{t} \frac{t!}{n!(t-n)!}(1-\lambda)^{n} \cdot \lambda^{t-n} \cdot \bar{r}^{n} \cdot \bar{s}^{t-n}=[\bar{r}(1-\lambda)+\bar{s} \lambda]^{t} \tag{3.33}
\end{equation*}
$$

$$
\begin{align*}
\mathbb{E}\left[N^{2}(t)\right]= & \sum_{n=0}^{t} \frac{t!}{n!(t-n)!}(1-\lambda)^{n} \cdot \lambda^{t-n} \cdot \bar{r}^{2 n} \cdot \bar{s}^{2(t-n)}=\left[\bar{r}^{2}(1-\lambda)+\bar{s}^{2} \lambda\right]^{t}  \tag{3.35}\\
& \operatorname{Var}[N(t)]=\left[\bar{r}^{2}(1-\lambda)+\bar{s}^{2} \lambda\right]^{t}-[\bar{r}(1-\lambda)+\bar{s} \lambda]^{2 t} \tag{3.34}
\end{align*}
$$

Inserting Equations (3.33) and (3.35) in Equation (3.30) we can look for its minimum with respect to $p$ finding $p^{*}=\frac{1}{2}$.
The same result can be achieved using Varadhan's Lemma: instead of minimizing Equation (3.30) we consider its square and then we exploit

$$
\begin{equation*}
\frac{\operatorname{Var}[N(t)]}{(\mathbb{E}[N(t)])^{2}}=\frac{\mathbb{E}\left[N^{2}(t)\right]}{(\mathbb{E}[N(t)])^{2}}-1 \tag{3.36}
\end{equation*}
$$

In this way we can look only at $\frac{\mathbb{E}\left[N^{2}(t)\right]}{\left(\mathbb{E}[N(t)]^{2}\right.}$. Taking the logarithm of this quantity, multiplying in by $\frac{1}{t}$ and considering the limit $t \rightarrow \infty$ we can apply the Lemma. So we have

$$
\begin{equation*}
\lim _{t \rightarrow \infty} \frac{1}{t} \log \left[\frac{\mathbb{E}\left[N^{2}(t)\right]}{(\mathbb{E}[N(t)])^{2}}\right]=\lim _{t \rightarrow \infty} \frac{1}{t} \log \left[\mathbb{E}\left[N^{2}(t)\right]\right]-\frac{2}{t} \log [\mathbb{E}[N(t)]]=\log \left[\frac{\bar{r}^{2}(1-\lambda)+\bar{s}^{2} \lambda}{[\bar{r}(1-\lambda)+\bar{s} \lambda]^{2}}\right] \tag{3.37}
\end{equation*}
$$

Minimizing the argument of the logarithm with respect to $p$ we find again $p^{*}=\frac{1}{2}$.
We can understand why $p^{*}=\frac{1}{2}$ is the minimum of the quantity in Equation (3.30). In fact it is the point of zero of the variance. Taking this value Equation (3.35) is zero. This can be explained looking at the definitions of $\bar{r}$ and $\bar{s}$ : with $p^{*}=\frac{1}{2}$ we have $\bar{r}=\bar{s} \forall r, s$ and so the growth of the population in time becomes deterministic. In other words with this choice of $p$ and at each step the number of individuals increases by the same factor regardless of what the environment does.
With this investment strategy we have $\bar{r}=\bar{s}$, so exploiting this fact in Equation (3.10) we find $b(\lambda)=2 \forall r, s, \lambda$ since

$$
\begin{equation*}
b(\lambda)=\frac{\log \left[(1-\lambda) \bar{r}^{2}+\lambda \bar{r}^{2}\right]}{\log [(1-\lambda) \bar{r}+\lambda \bar{r}]}=\frac{\log \left[\bar{r}^{2}\right]}{\log [\bar{r}]}=2 \tag{3.38}
\end{equation*}
$$

### 3.4.2 Markowitz optimization: growth and risk minimization

In order to minimize the fluctuations we found that the population has to invest half and half of the resources in the two strategies and we saw that this leads to a deterministic evolution. But we have not studied yet if this type of investment can allow a growth. So now we propose ourselves to find which choice of $p$ with $r, s$ and $\lambda$ fixed permits a growth in the population taking the lowest extinction-risk possible. We borrowed (and adapted to our situation) the idea of this optimization from the discipline of Finance and risk management in which is called Markowitz optimization and it applies to portfolio management.

In this field, Markowitz optimization gives the ideal portfolio diversification over a set of different assets of a market in which each asset has fixed gain or loss factors (similar to our factors $r$ and $s$ ). Then, fixing a desired income, the objective of this optimization is to provide the lowest risk of loss of capital investment that can provide this profit. Adapting this idea to an ecological framework the simplest and most realistic request is to impose a growth on the long term, without asking how big it has to be. In this way we avoid the introduction of a new parameter that has not an ecological meaning, because the individuals do not growth with purpose of reaching a certain number of individuals. Instead they can try to actually invest their living resources to make the community to survive, i.e. growing, maybe slowly, but taking the lowest risk of extinction possible.
To ask a growth means $\mathbb{E}[N(t)]>1$ or similarly $\lim _{t \rightarrow+\infty} \frac{1}{t} \log \mathbb{E}[N(t)]>0$ ( $N_{0}=1$ ). With this condition satisfied, we look for the value of $p$ that gives the smallest value of the the quantity (3.30).
Let us start looking at which values of $p$ can grant a growth on the long term. To do so we ask

$$
\begin{equation*}
\lim _{t \rightarrow+\infty} \frac{1}{t} \log \mathbb{E}[N(t)]=\log [\bar{r}(1-\lambda)+\bar{s} \lambda]>0 \tag{3.39}
\end{equation*}
$$

obtaining

$$
\begin{equation*}
\bar{r}(1-\lambda)+\bar{s} \lambda=r \lambda+s(1-\lambda)+p \cdot(r-s) \cdot(1-2 \lambda)>1 \tag{3.40}
\end{equation*}
$$

Solving this we find

- if $\lambda<\frac{1}{2}$

$$
\begin{cases}p_{\text {lim }}<p \leq 1 & \text { if } r>s  \tag{3.41}\\ 0 \leq p<p_{\text {lim }} & \text { if } s>r\end{cases}
$$

- if $\lambda>\frac{1}{2}$

$$
\begin{cases}0 \leq p<p_{\text {lim }} & \text { if } r>s  \tag{3.42}\\ p_{\text {lim }}<p \leq 1 & \text { if } s>r\end{cases}
$$

where we have defined

$$
\begin{equation*}
p_{\text {lim }}=\frac{1-r \lambda-s(1-\lambda)}{(r-s) \cdot(1-2 \lambda)} \tag{3.43}
\end{equation*}
$$

We see that this range of values is bounded by the quantity $p_{\text {lim }}$. Hence now we have to study how $p_{\text {lim }}$ changes as a function of $r, s$ and $\lambda$. We saw that $p=\frac{1}{2}$ is the absolute minimum of the fluctuations. This means that if this value is in the interval of values of $p$ allowing a growth it will be taken as optimal $p$, i.e. $p^{*}=\frac{1}{2}$. In this case the population will growth surely with a null possibility of extinction. Otherwise the population will adopt a different $p^{*}$ which can't grant a deterministic growth, but we will have a growth with non-zero fluctuations.

Case $\lambda<\frac{1}{2}$
Let us start with the case $\lambda<\frac{1}{2}$ (in the following discussion we consider $\lambda$ fixed and smaller than $\frac{1}{2}$ ). We found the interval of values for $p$ to have an increase in the population in time to be

$$
\begin{cases}p_{\text {lim }}<p \leq 1 & \text { if } r>s  \tag{3.44}\\ 0 \leq p<p_{\text {lim }} & \text { if } s>r\end{cases}
$$

So now we have to study the behaviour of $p_{\text {lim }}$ and based on that we will find the optimal value $p^{*}$. As we discussed, $p=\frac{1}{2}$ is the minimum of the fluctuations. Then we have to investigate the possibility of this value to fall into the intervals of Equation (3.44).
Let us start from $r>s$. To make $p=\frac{1}{2}$ to fall within the allowed interval of values we have to ask $p_{\text {lim }} \leq \frac{1}{2}$ that gives the condition

$$
\begin{equation*}
r+s \geq 2 \tag{3.45}
\end{equation*}
$$

Then we have to control when $p_{\text {lim }}>1$. The solution is

$$
\begin{equation*}
s<\frac{1}{\lambda}+\frac{\lambda-1}{\lambda} r \tag{3.46}
\end{equation*}
$$

When this happens there is no more an interval of $p$ allowing for a growth (on average). Hence in this case the individuals cannot growth and going to become extinct anyway. To try to live as longer as it can, the population will take $p^{*}=1$. We see that, strictly speaking, in this situation there is not an optimal value of $p$ because $\mathbb{E}[N(t)]<1$ for all $p$. The only ecological move that the individuals may carry out is trying to last for the longest time possible. So it maximizes $\log [\bar{r}(1-\lambda)+\bar{s} \lambda]$ which is the mean value of the growth rate taking $p^{*}=1$.
The last possibility to be considered is

$$
\begin{equation*}
\frac{1}{2}<p_{l i m} \leq 1 \tag{3.47}
\end{equation*}
$$

which gives the region

$$
\begin{equation*}
\frac{1}{\lambda}+\frac{\lambda-1}{\lambda} r \leq s<2-r \tag{3.48}
\end{equation*}
$$

Now there is an interval for $p$ satisfying the growth condition, but $\frac{1}{2}$ is not in there. Above we found there is only a point of minimum with respect to $p$ for the fluctuations and so minimum condition on the fluctuations cannot be imposed any more. In fact the minimum in this interval would be $p=p_{\text {lim }}$ but with this choice $\lim _{t \rightarrow+\infty} \frac{1}{t} \log \mathbb{E}[N(t)]=0$. Hence again the unique reasonable ecological move is trying to maximize the growth. This set $p^{*}=1$. The difference now is that with this $p^{*}$ a growth(on average) is possible.

In the case $r<s$ we proceed similarly. To make $\frac{1}{2}$ within the permitted interval we have to ask $p_{\text {lim }} \geq \frac{1}{2}$ finding again

$$
\begin{equation*}
r+s \geq 2 \tag{3.49}
\end{equation*}
$$

Then we check

$$
\begin{equation*}
p_{l i m}<0 \tag{3.50}
\end{equation*}
$$

finding

$$
\begin{equation*}
s<\frac{1}{1-\lambda}-\frac{1}{1-\lambda} r \tag{3.51}
\end{equation*}
$$

In this case there is no interval of values for $p$ allowing a growth. For the same way of reasoning shown above $p^{*}=0$.
Finally we look

$$
\begin{equation*}
0 \leq p_{l i m}<\frac{1}{2} \tag{3.52}
\end{equation*}
$$

providing

$$
\begin{equation*}
\frac{1}{1-\lambda}-\frac{1}{1-\lambda} r \leq s<2-r \tag{3.53}
\end{equation*}
$$

The minimum condition on the fluctuations can not be established, so to fix the fraction $p$ the population will set $p^{*}=0$ to maximize the growth and trying to live as longest as possible.
So we have that the parameter space will be divided as shown in Figure 3.10. In the blue region, defined by

$$
\begin{equation*}
\{(r, s): r+s \geq 2\} \tag{3.54}
\end{equation*}
$$

it is possible to have $p^{*}=\frac{1}{2}$ since it belongs to the allowed interval of values for $p$ which grant a growth. So with $\lambda<\frac{1}{2}$ and a couple $(r, s)$ in this area we can have a deterministic growth, i.e. surely the population will increase in time.
In the purple zone

$$
\begin{equation*}
\left\{(r, s): r<s, s<\frac{1}{1-\lambda}-\frac{1}{1-\lambda} r\right\} \tag{3.55}
\end{equation*}
$$

we have $0 \leq p<p_{\text {lim }}$ with $p_{\text {lim }}<0$. This says us that with a fixed $\lambda$ and $r, s$ in this region we can not have growth since $\bar{r}(1-\lambda)+\bar{s} \lambda<1 \forall p \in[0,1]$. Hence in this region $p^{*}=0$, i.e although taking the value for which $\mathbb{E}[N(t)]$ is maximized there will be no growth, but it can hope to survive as longest as possible.
Analogously in the green region

$$
\begin{equation*}
\left\{(r, s): r>s, s<\frac{1}{\lambda}+\frac{\lambda-1}{\lambda} r\right\} \tag{3.56}
\end{equation*}
$$

we have $p_{\text {lim }}<p \leq 1$ with $p_{\text {lim }}>1$, so we have to take $p^{*}=1$. As before this means that we can not have an increase (on average) in the number of individuals since $\bar{r}(1-\lambda)+\bar{s} \lambda<1 \forall p \in[0,1]$.

Let us now consider the orange region

$$
\begin{equation*}
\left\{(r, s): r>s, \frac{1}{\lambda}+\frac{\lambda-1}{\lambda} r<s<2-r\right\} . \tag{3.57}
\end{equation*}
$$

In this zone we have $\frac{1}{2}<p_{\text {lim }}<p \leq 1$, i.e. we can have a growth $(\bar{r}(1-$ $\lambda)+\bar{s} \lambda>1$ ) but $p=\frac{1}{2}$ would not satisfy the growth condition. So we have to choose $p^{*}$ based on another criterion (maximization of the growth).
A similar situation occurs in the red zone $(s>r)$ where we have $0 \leq p<$ $p_{\text {lim }}<\frac{1}{2}$, i.e. $\bar{r}(1-\lambda)+\bar{s} \lambda>1$ but not using $p=\frac{1}{2}$.
In these two regions the minimum of the variance can not be the absolute minimum, but it is in the extremes of the interval. Since we asked for a growth we have to take $p=1$ in the orange zone, while $p=0$ in the red one (if $p=p_{\text {lim }}$ then $\mathbb{E}[N(t)]=1=N_{0}$, i.e. there will be no increase, we invest our resources to stay steady, which is meaningless).
So with $\lambda<\frac{1}{2}$ we found

$$
p(\lambda, r, s)= \begin{cases}\frac{1}{2} & \text { if } r+s>2  \tag{3.58}\\ 1 & \text { if } r+s<2 \wedge r>s \\ 0 & \text { if } r+s<2 \wedge s>r\end{cases}
$$

Case $\lambda>\frac{1}{2}$
In the same way we can analyze the case $\lambda>\frac{1}{2}$, obtaining the division of the parameter space shown in Figure 3.11. This division is the same because, by the definition of $p_{\text {lim }}$ (which is the discriminating factor on which basically the division is based on) the exchanges

$$
\begin{gathered}
\lambda \longleftrightarrow 1-\lambda \\
r \longleftrightarrow s
\end{gathered}
$$

at the same time cannot provide any difference. Now instead the conditions (3.41) and (3.42) are the same with these exchanges. So for $\lambda>\frac{1}{2}$ we have

$$
p^{*}(\lambda, r, s)= \begin{cases}\frac{1}{2} & \text { if } r+s>2  \tag{3.59}\\ 1 & \text { if } r+s<2 \wedge r>s \\ 0 & \text { if } r+s<2 \wedge s>r\end{cases}
$$

This reasoning can be verified by direct calculation as done before.
Summarizing, with this optimization we find:

- if $\lambda<\frac{1}{2}$

$$
p^{*}(\lambda, r, s)= \begin{cases}\frac{1}{2} & \text { if } r+s>2  \tag{3.60}\\ 1 & \text { if } r+s<2 \wedge r>s \\ 0 & \text { if } r+s<2 \wedge s>r\end{cases}
$$

- if $\lambda>\frac{1}{2}$

$$
p^{*}(\lambda, r, s)= \begin{cases}\frac{1}{2} & \text { if } r+s>2  \tag{3.61}\\ 0 & \text { if } r+s<2 \wedge r>s \\ 1 & \text { if } r+s<2 \wedge s>r\end{cases}
$$



Figure 3.10: Division of the parameter space using Markowitz optimization with $\lambda=\frac{1}{4}$. In the blue region $p=\frac{1}{2}$, in the green one $p=1$, in the purple zone $p=0$. In the orange region $\frac{1}{2}<p_{\text {lim }}<p \leq 1$ and in the red one $0 \leq p<p_{\text {lim }}<\frac{1}{2}$.
The border between the blue region and the orange-red is the straight line $s=2-r$, the border between the purple zone and the green region is the straight line $r=s$, between the green and orange regions the border is defined by the straight line $s=\frac{1}{\lambda}+r \frac{\lambda-1}{\lambda}$ and between the purple and red zone is defined by the straight line $s=r \frac{\lambda}{\lambda-1}-\frac{1}{\lambda-1}$.


Figure 3.11: Division of the parameter space using Markowitz optimization with $\lambda=\frac{3}{4}$. In the blue region $p^{*}=\frac{1}{2}$, in the green one $p^{*}=0$, in the purple zone $p^{*}=1$. In the orange region $0 \leq p^{*}<p_{\text {lim }}<\frac{1}{2}$ and in the red one $\frac{1}{2}<p_{\text {lim }}<p^{*} \leq 1$.
The border between the blue region and the orange-red is the straight line $s=2-r$, the border between the purple zone and the green region is the straight line $r=s$, between the green and orange regions the border is defined by the straight line $s=\frac{1}{\lambda}+r \frac{\lambda-1}{\lambda}$ and between the purple and red zone is defined by the straight line $s=r \frac{\lambda}{\lambda-1}-\frac{1}{\lambda-1}$.


Figure 3.12: Regions of divergence for the generalized TL exponent in the case of the two strategies dynamics with $p^{*}$ optimized using Markowitz optimization. The blue zone represents $\widetilde{R_{1}}$. The orange one instead is $\widetilde{R_{2}}$. The small panels display some plots of $b(\lambda)$ for $r$ and $s$ in the different regions.

## Divergences

With the optimal $p^{*}$ found above we can look for the regions in which generalized TL exponent may diverge.
As we said, for $p^{*}=\frac{1}{2}$ there is not divergence in the exponent. Instead when $r+s<2$ we find the regions shown in Figure 3.12 and found with the previous optimization performed.
In conclusion, the divergences would disappear taking $p=\frac{1}{2}$ everywhere, but as we showed this would not allow a growth for every choice of $r$ and $s$. In these regions of the parameters the only reasonable thing to do is maximized the growth. We accomplished this maximizing $\lim _{t \rightarrow+\infty} \frac{1}{t} \log \mathbb{E}[N(t)]$ as in 3.3.2. We could use the first optimization (the one exploiting LLN), but as we verified any remarkable difference would be obtained since both of them provide the same regions of divergences.
In this way we discovered that, regardless on the optimization objective, the two strategies model predicts exactly the same regions of divergences. Moreover the region of divergence for the generalized TL, under natural hypothesis on the dynamics of the environment and the model parameters, is again $\bar{R}_{1, \text { strat }}$ as in the two previous considered models.

## Chapter 4

## Population dynamics with adaptive strategy and a randomly moving fraction of individuals

Up to now we illustrated two dynamics in which the population moves with one or two evolutionary strategies, i.e given a certain realization of the chain describing the environment the succession of the phenotypes adopted by the individuals is deterministic.
This could be seen in contrast with what is observed in nature. In fact we know from the studies of Darwin about the existence of random mutation in the genotype that could manifest as different traits in the phenotypes.
So to take into account this feature of the ecological dynamics we have to introduce in our model a fraction of population moving randomly, i.e. that takes decisions about what phenotype adopt without looking at the environment.

### 4.1 Model dynamics

Once more we consider a population living in a stochastic environment whose evolution is described by the same Markov chain $\left\{A_{t, t \geq 0}\right\}$ moving in a two dimensional state-space $\Gamma$ (we call these two states 0 and 1 again) with the same transition matrix $P(2.2)$ as in the previous models.
Also in this case the individuals can adopt two different phenotypes due to the possibility of having two different habitat states. Each phenotype is the most adapted to one of the possible environment states (again we will call the two phenotypes 0 and 1 with the same notation and meaning as above). Only the individuals with the right phenotype with respect to the environment can grow (i.e. their number will be multiplied by a factor $r$, typically $r>1$ ), while the others will be reduced by a factor $s$ (typically $0<s<1$ ).
Now we introduce a split in the population. A fraction $p$ follows the Darwinian strategy $S_{1}$ in which the individuals for the next step choose to be adapted to what they are seeing now or, in other words, they become
adapted to the environment and bet the environment won't change. Instead the other fraction $1-p$ represents that part of population which moves randomly, i.e. which adopts a phenotype without basing its decision on what the environment may do. We will call this $S_{r}$ : at each step this group of individuals choose to use the phenotype 0 or 1 with probability $\frac{1}{2}$.
An easy calculation, based on the transition matrix $P$ and the steady distribution of the Markov chain, shows how moving randomly is equal to remain always in one of the two possible phenotypes, let us say 0 without any loss of generality.
So for the fraction $p$ the phenotype adopted at the time $t+1, S_{1, t+1}$, is the one good for the environment at the step $t, A_{t}$, while for $S_{r}$ the phenotype will be always 0 , i.e. $S_{r, t}=0 \forall t \geq$ 1, i.e.

$$
\begin{gathered}
S_{1, t+1}=S_{1, t+1}\left(A_{t}\right)=A_{t} \\
S_{r, t}=0 \quad \forall t \geq 1
\end{gathered}
$$

Here an example of what could happen with this dynamics given a certain realization of the environment chain:

$$
\begin{aligned}
& t: 012345 \cdots \\
& A_{t}: 001101 \cdots \\
& S_{1, t}: x 00110 \cdots \\
& S_{r, t}: x 00000 \cdots
\end{aligned}
$$

We can look at these two evolutionary lines as the usual Darwinian one $\left(S_{1}\right)$ and the random motion naturally present $\left(S_{r}\right)$. In fact the fraction playing $S_{1}$ increases the individuals by $r$ when they become adapt to the environment until it change. When it happens they are no more able to grow since they do not find a good habitat to live in and they have to try to be adapted again adopting a new phenotype. Instead the other fraction using $S_{r}$ wins only when $A_{t+1}=0$. This is not an adaptive strategy since their user do not try to predict the evolution of the environment to remain adapted as longest as possible changing the phenotype but they remain passive.

### 4.1.1 Population growth

We can write the population size as

$$
\begin{equation*}
N(t)=\overbrace{p N(t)}^{S_{1}}+\overbrace{(1-p) N(t)}^{S_{2}} \tag{4.1}
\end{equation*}
$$

since now we have two sub-populations, the users of $S_{1}$ and $S_{r}$. When the time moves from $t$ to $t+1$, based on the strategies and on what the environment did, the numbers of individuals of the two sub-population will be multiplied by $r$ or $s$. Now the population size is $N(t+1)$ and it has to decide much it is invested in $S_{1}$ or $S_{r}$. We will keep fixed what are the fractions $p$ and $1-p$ playing the two evolutionary lines at each time-step:

$$
\begin{equation*}
N(t+1)=\overbrace{p N(t+1)}^{S_{1}}+\overbrace{(1-p) N(t+1)}^{S_{r}} \tag{4.2}
\end{equation*}
$$

We want now to write an expression for the population size as a function of the time $t$. Given $N(t)$ and remembering how the two strategies are defined we have

$$
\begin{align*}
N(t+1)= & p N(t)\left[r \cdot \delta_{A_{t}, A_{t+1}}+s \cdot\left(1-\delta_{A_{t}, A_{t+1}}\right)\right]+ \\
& +(1-p) N(t)\left[s \cdot \delta_{A_{t+1}, 0}+r \cdot\left(1-\delta_{A_{t+1}, 0}\right)\right] \tag{4.3}
\end{align*}
$$

Since the fractions are kept fixed we iterate Equation (4.3) obtaining

$$
\begin{align*}
N(t)= & N_{0} \prod_{n=1}^{t}\left\{p\left[r \cdot \delta_{A_{n-1}, A_{n}}+s \cdot\left(1-\delta_{A_{n-1}, A_{n}}\right)\right]+\right. \\
& \left.+(1-p)\left[s \cdot \delta_{A_{n}, 0}+r \cdot\left(1-\delta_{A_{n}, 0}\right)\right]\right\} \tag{4.4}
\end{align*}
$$

Now we can see concretely the difference between $S_{1}$ and $S_{r}$ : while for the first what matters is the changing or the persistence of the environment in the same state without any importance of which this is, for the last instead we have to look at the state acquires by the chain regardless of the transition just effectuated. So we have to factorize the delta in the following way to emphasize the transitions of the chain to write in more useful form Equation (4.4):

$$
\left\{\begin{array}{l}
\delta_{A_{n-1}, A_{n}}=\delta_{\left(A_{n-1}, A_{n}\right),(0,0)}+\delta_{\left(A_{n-1}, A_{n}\right),(1,1)}  \tag{4.5}\\
1-\delta_{A_{n-1}, A_{n}}=\delta_{\left(A_{n-1}, A_{n}\right),(0,1)}+\delta_{\left(A_{n-1}, A_{n}\right),(1,0)} \\
\delta_{A_{n}, 0}=\delta_{\left(A_{n-1}, A_{n}\right),(0,0)}+\delta_{\left(A_{n-1}, A_{n}\right),(1,0)} \\
1-\delta_{A_{n}, 0}=\delta_{\left(A_{n-1}, A_{n}\right),(0,0)}+\delta_{\left(A_{n-1}, A_{n}\right),(1,0)}
\end{array}\right.
$$

Equation (4.4) now becomes

$$
\begin{align*}
N(t) & =N_{0} \prod_{n=1}^{t}\left\{p \left[r \cdot\left(\delta_{\left(A_{n-1}, A_{n}\right),(0,0)}+\delta_{\left(A_{n-1}, A_{n}\right),(1,1)}\right)+s \cdot\left(\delta_{\left(A_{n-1}, A_{n}\right),(0,1)}+\right.\right.\right. \\
& \left.\left.+\delta_{\left(A_{n-1}, A_{n}\right),(1,0)}\right)\right]+(1-p)\left[s \cdot\left(\delta_{\left(A_{n-1}, A_{n}\right),(0,0)}+\delta_{\left(A_{n-1}, A_{n}\right),(1,0)}\right)+\right. \\
& \left.\left.+r \cdot\left(\delta_{\left(A_{n-1}, A_{n}\right),(0,0)}+\delta_{\left(A_{n-1}, A_{n}\right),(1,0)}\right)\right]\right\} \tag{4.6}
\end{align*}
$$

From this form we can see what are the multiplicative factors $\widetilde{A_{n}}$ and how each transition provides a different one of those

$$
\widetilde{A_{n}}= \begin{cases}r & \text { if }\left(A_{n-1}, A_{n}\right)=(0,0)  \tag{4.7}\\ s & \text { if }\left(A_{n-1}, A_{n}\right)=(0,1) \\ \bar{s} & \text { if }\left(A_{n-1}, A_{n}\right)=(1,0) \\ \bar{r} & \text { if }\left(A_{n-1}, A_{n}\right)=(1,1)\end{cases}
$$

where $\bar{r}$ and $\bar{s}$ are defined ad in the previous chapter ((3.5)). So

$$
\begin{align*}
N(t) & =N_{0} \prod_{n=1}^{t}\left\{\delta_{\left(A_{n-1}, A_{n}\right),(0,0)} \cdot[p r+(1-p) r]+\delta_{\left(A_{n-1}, A_{n}\right),(0,1)} \cdot[p s+(1-p) s]+\right. \\
& \left.+\delta_{\left(A_{n-1}, A_{n}\right),(1,0)} \cdot[p s+(1-p) r]+\delta_{\left(A_{n-1}, A_{n}\right),(1,1)} \cdot[p r+(1-p) s]\right\}= \\
& =N_{0} \prod_{n=1}^{t}\left\{\delta_{\left(A_{n-1}, A_{n}\right),(0,0)} r+\delta_{\left(A_{n-1}, A_{n}\right),(0,1)} s+\delta_{\left(A_{n-1}, A_{n}\right),(1,0)} \bar{s}+\delta_{\left(A_{n-1}, A_{n}\right),(1,1)} \bar{r}\right\}= \\
& =N_{0} \cdot r^{\sum_{n=1}^{t} \delta_{\left(A_{n-1}, A_{n}\right),(0,0)} \cdot s^{\sum_{n=1}^{t} \delta_{\left(A_{n-1}, A_{n}\right),(0,1} \cdot \bar{s}^{\sum_{n=1}^{t} \delta_{\left(A_{n-1}, A_{n}\right),(1,0)}} \cdot \bar{r}^{\sum_{n=1}^{t} \delta_{\left(A_{n-1}, A_{n}\right),(1,1}}=}} \begin{array}{l} 
\\
\end{array}=N_{0} \cdot r^{t \nu_{00}} \cdot s^{t \nu_{01}} \cdot \bar{s}^{t \nu_{10}} \cdot \bar{r}^{t \nu_{11}}= \\
& =N_{0} \cdot e^{t \nu_{00} \log r+t \nu_{01} \log s+t \nu_{10} \log \bar{s}+t \nu_{11} \log \bar{r}}= \\
& =N_{0} e^{t G_{r s}(\nu)}
\end{align*}
$$

We introduced $\nu=\left(\nu_{00}, \nu_{01}, \nu_{10}, \nu_{11}\right)$ which the four component object counting the fraction of times each transition occurs in a realization of the Markov chain up to time $t$

$$
\begin{equation*}
\nu_{i j}=\frac{1}{t} \sum_{n=1}^{t} \delta_{\left(A_{n-1}, A_{n}\right),(i, j)} \tag{4.9}
\end{equation*}
$$

with the obvious constraint $\sum_{i, j=0}^{1} \nu_{i j}=\nu_{00}+\nu_{01}+\nu_{10}+\nu_{11}=1$, i.e $\nu$ belongs to the four-dimensional simplex $\Delta^{4}$.
As we can recover from 2.2, we have already introduced the rate function for family of probabilities $\left.P_{t}(\nu)=\mathbb{P}\left(L_{t}^{2} \in \nu, \nu+d \nu\right]\right)$ for the empirical pair measure counting the same quantities

$$
\begin{equation*}
L_{t}^{2}=\frac{1}{t} \sum_{n=1}^{t} \delta_{\left(A_{n-1}, A_{n}\right)}, \tag{4.10}
\end{equation*}
$$

This rate function is

$$
\begin{align*}
I_{P}^{2}(\nu) & =\nu_{00} \log \left[\frac{\nu_{00}}{\left(\nu_{00}+\nu_{01}\right) \cdot(1-\lambda)}\right]+\nu_{01} \log \left[\frac{\nu_{01}}{\left(\nu_{00}+\nu_{01}\right) \cdot \lambda}\right]+ \\
& +\nu_{10} \log \left[\frac{\nu_{10}}{\left(\nu_{10}+\nu_{11}\right) \cdot \lambda}\right]+\nu_{11} \log \left[\frac{\nu_{11}}{\left(\nu_{10}+\nu_{11}\right) \cdot(1-\lambda)}\right] \tag{4.11}
\end{align*}
$$

The function

$$
\begin{equation*}
G(\nu)=\nu_{00} \log r+\nu_{01} \log s+\nu_{10} \log \bar{s}+\nu_{11} \log \bar{r} \tag{4.12}
\end{equation*}
$$

has the usual meaning of growth rate and it is parametrized by $r, s$ and $p$.

### 4.1.2 Study of the generalized Taylor's Law exponent

With $G(\nu)$ and $I_{P}^{2}(\nu)$ we can evaluate the exponent $b$ of the generalized Taylor's law relating the second and the first moment of the random variable
$N(t)$ exploiting again Varadhan's Lemma. In fact

$$
\begin{equation*}
b(\lambda)=\frac{\sup _{\nu \in \Delta^{4}}\left[2 G(\nu)-I_{P}^{2}(\nu)\right]}{\sup _{\nu \in \Delta^{4}}\left[G(\nu)-I_{P}^{2}(\nu)\right]} \tag{4.13}
\end{equation*}
$$

We see that we should compute two suprema over the four variables $\nu_{i j}$. Due to the constraint $\nu_{00}+\nu_{01}+\nu_{10}+\nu_{11}=1$ there are only three independent variables. We can still reduce the number of degrees of freedom noticing that, due to the dynamics of the chain, we have to count the same number of transition from 0 to 1 and from 1 to 0 and this leads us to $\nu_{01}=\nu_{10}$. Finally we have two independent variables.
Unfortunately also with these simplifications the suprema can not be obtained analytically. Although we can't find them explicitly, for sure the functions $2 G(\nu)-I(\nu)$ and $G(\nu)-I(\nu)$ have a maximum and a minimum in $\Delta^{4}$ since they are continuous functions on a compact domain. So we have to look for the two suprema via numerical calculations.

### 4.2 The optimal investment strategies

Also in this case we have to fix $p$ somehow. We will use the same optimizations used in the previous chapter, i.e. maximizing the growth considering or neglecting rare events or minimizing the risk of extinction (i.e. the fluctuations) with the constraint of a growth on the long term (not always possible as we will see). Due to the absence of an explicit form of the suprema contributing to $b$, we will not have an expression for the optimal investment fraction $p$ on the strategy $S_{1}$, with the exception of the first way of optimization since in this case we look only at $G(\nu)$.

### 4.2.1 Maximization based on the most probable realizations of the chain

Again the first optimization consists in maximizing the growth of the population on the long term taking into account only the most probable realizations of the chain describing the environment. So we will exploit the ergodic theorem.
Since we would like to look for large $t$ we will consider the limit $t \rightarrow+\infty$. In this limit the chain reaches its steady distribution which is

$$
\begin{equation*}
\pi(i)=\frac{1}{2} \quad \forall i \in \Gamma \tag{4.14}
\end{equation*}
$$

thanks to the bistochastic form of the transition matrix. So we have

$$
\begin{equation*}
\lim _{t \rightarrow+\infty} \mathbb{P}\left(A_{t}=i\right)=\pi(i)=\frac{1}{2} \forall i \in \Gamma \tag{4.15}
\end{equation*}
$$

Now $\nu_{i j}$ was defined as

$$
\nu_{i j}=\frac{1}{t} \sum_{n=1}^{t} \delta_{\left(A_{n-1}, A_{n}\right),(i, j)}
$$

The ergodic theorem (see Appendix A) states

$$
\begin{equation*}
\mathbb{P}\left(\nu_{i j} \rightarrow \pi(i) \cdot p_{i j}\right) \xrightarrow[t \rightarrow+\infty]{\longrightarrow} 1 . \tag{4.16}
\end{equation*}
$$

and so in this limit we obtain

$$
\nu=\left\{\begin{array}{l}
\nu_{00}=\nu_{11}=\frac{1-\lambda}{2}  \tag{4.17}\\
\nu_{01}=\nu_{10}=\frac{\lambda}{2}
\end{array}\right.
$$

The same result could be achieved noting that $\nu=\left(\frac{1-\lambda}{2}, \frac{\lambda}{2}, \frac{\lambda}{2}, \frac{1-\lambda}{2}\right)$ is the point of minimum of the rate function (4.11) for the family of probabilities $\mathbb{P}\left(L_{t}^{2} \in[\nu, \nu+d \nu]\right)$.
As we can see automatically we found $\nu_{01}=\nu_{10}$ and $\nu_{00}=\nu_{11}$ as we said above. In this way $G(\nu)$ becomes only a function of the control parameter $\lambda$

$$
\begin{equation*}
G(\lambda)=\frac{1-\lambda}{2} \log r+\frac{\lambda}{2} \log s+\frac{\lambda}{2} \log \bar{s}+\frac{1-\lambda}{2} \log \bar{r} \tag{4.18}
\end{equation*}
$$

Since $G$ is the growth rate, we try to maximize this quantity with respect to $p$ we obtain the optimal investment fraction to be

$$
\begin{equation*}
p^{*}(r, s, \lambda)=\frac{r(1-\lambda)-\lambda s}{r-s} \tag{4.19}
\end{equation*}
$$

which is the same found in 3.3.1. So the discussion to bound $p^{*}$ between 0 and 1 depending on the sign of $r-s$ is exactly the same. Thus we have

- if $r>s$

$$
p^{*}(r, s, \lambda)= \begin{cases}1 & \text { if } 0<\lambda \leq \frac{s}{r+s}  \tag{4.20}\\ \frac{r(1-\lambda)-\lambda s}{r-s} & \text { if } \frac{s}{r+s} \leq \lambda \leq \frac{r}{r+s} \\ 0 & \text { if } \frac{r}{r+s} \leq \lambda<1\end{cases}
$$

- if $r<s$

$$
p^{*}(r, s, \lambda)= \begin{cases}0 & \text { if } 0<\lambda \leq \frac{r}{r+s}  \tag{4.21}\\ \frac{r(1-\lambda)-\lambda s}{r-s} & \text { if } \frac{r}{r+s} \leq \lambda \leq \frac{s}{r+s} \\ 1 & \text { if } \frac{s}{r+s} \leq \lambda<1\end{cases}
$$

Depending on $r, s$ we have two different plots of $p^{*}$ shown in Figure 4.1. With the optimal investment fraction $p^{*}$ now we have to compute the two suprema of Varadhan's Lemma as functions of $\lambda$ remembering the constraint over $\nu$. To compute them we implement a numerical procedure following these steps:
(1) we fix of $r$ and $s$;
(2) we fix a value of $\lambda \in[0,1]$;
(3) with $r, s$ and $\lambda$ we compute $p^{*}$ with Equations (4.20) and (4.21);
(4) we compute numerically $\sup _{\nu \in \Delta^{4}}\left[2 G(\nu)-I_{P}^{2}(\nu)\right], \sup _{\nu \in \Delta^{4}}\left[G(\nu)-I_{P}^{2}(\nu)\right]$ and their ratio, i.e. the generalized TL exponent;


Figure 4.1: Plots of $p^{*}(\lambda)$ optimized using the ergodic theorem for the model dynamics with the adaptive strategy and the randomly moving fraction for different $r$ and $s$ with $r>s$ and $r<s$.
(5) we reiterate step 2 for other values of $\lambda$ to obtain a discrete plot for the exponent with the generated $r$ and $s$.
With this procedure we can estimate the exponent. This can be only an approximation due to the discretization of the domain of $\lambda$, but it is still useful since it allows us to have an insight on how $b$ behaves. Some plots are displayed in Figure 4.2.
As we can see from Figures 4.2c and 4.2d there are same values of the parameters $r$ and $s$ for which the exponent diverges.
Since we haven't an analytic expression for $b$, we haven't also the expression for $\lambda_{c}$, but we can add a step in the procedure to detect the points in the parameter space $(r, s)$ causing the divergence. In fact we label a couple $(r, s)$ as point of divergence if the exponent goes above a threshold (we used $8-10$ ) or becomes negative. To reconstruct the regions of divergence, shown in Figure 4.3, we sampled uniformly couples $(r, s)$ in $[0,3] \times[0,3]$ and we tested the divergence as just explained. It is possible to note that the situation is no more symmetric for the exchange of $r$ and $s$. This can be explained from the expression of $G(\nu)$ in Equation (4.12). In fact exchanging $r$ with $s$ (and consequently $p$ with $1-p$ as we can see from Equations (4.20) and (4.21)) we have that $G(\nu)$ goes into

$$
\begin{equation*}
\widetilde{G(\nu)}=\nu_{00} \log s+\nu_{01} \log r+\nu_{10} \log \bar{s}+\nu_{11} \log \bar{r} \tag{4.22}
\end{equation*}
$$

which is a different function, hence the exponent (and its region of divergence) is no more symmetric for the exchange of the parameters. This will be true also in the case of the other two optimizations
We see that this region for $r>1$ looks like the same found in the work discussed in 1.3.1, while for $r<1$ seems the one of the two strategies model.
For the usual natural hypothesis on the dynamics, we can for the region in which $\lambda_{c}<\frac{1}{2}$ finding the division shown in Figure 4.4.
Although without any expression for the critical value for $\lambda$, we can see that the red regions in Figure 4.4 look like the same regions of the two strategies model illustrated in the previous chapter and defined by the conditions

$$
\begin{equation*}
(r>1 \vee s>1) \wedge r+s<2 \tag{4.23}
\end{equation*}
$$

while the blue one look the same of the work presented at 1.3.1, i.e defined by

$$
\begin{equation*}
r>1 \wedge s<\frac{1}{r} \tag{4.24}
\end{equation*}
$$

So if we look for $0<\lambda<\frac{1}{2}$ with $0<s<1<r$ (natural interval) also in this case we find the same region of divergence predicted by all the previous analyzed dynamics.

### 4.2.2 Maximization considering also the contribution of rare events

To consider also the rare events, that may provide a great contribution to the right evaluation of mean values although their low probability of


Figure 4.2: Plots of $b(\lambda)$ for the model dynamics with the adaptive strategy and the randomly moving fraction optimized using the ergodic theorem with different values of $r$ and $s$.


Figure 4.3: Points generated by uniform sampling of the region $[0,3] \times[0,3]$ of the parameter space that make $b$ to display divergences for the model dynamics with the adaptive strategy and the randomly moving fraction optimized using the ergodic theorem.


Figure 4.4: Regions in which the generalized TL exponent diverges for the model dynamics with the adaptive strategy and the randomly moving fraction optimized using the ergodic theorem. The blue points make $b$ to diverge with $\lambda_{c}>\frac{1}{2}$. The red one instead provide $\lambda_{c}<\frac{1}{2}$. In this way we derived numerical evidences on what the regions could be. The small panels shows some plots obtained via numerical calculation of the exponent for a couple $r$ and $s$ in the different regions of the parameters obtained in such a way.
happening, we have to go beyond LLN and ergodic theorem and exploit LDT.
To maximize the growth on long term we have to look for the maximum of $\mathbb{E}[N(t)]$ with respect to $p$. Since the logarithm is monotone, in this optimization we have to maximize

$$
\begin{equation*}
\lim _{t \rightarrow+\infty} \frac{1}{t} \log \mathbb{E}[N(t)] \tag{4.25}
\end{equation*}
$$

Applying once again Varadhan's Lemma we have to look for the maximum with respect to $p$ of

$$
\begin{equation*}
\sup _{\nu \in \Delta^{4}}\left[G(\nu)-I_{P}^{2}(\nu)\right] \tag{4.26}
\end{equation*}
$$

i.e. look at the supremum as a function of $p$ and search its maximum.

Due to the absence of an analytical form, we implement a numerical procedure to compute the optimal $p^{*}$ and the exponent of generalized TL:
(1) we fix $r$ and $s$;
(2) we fix a value of $\lambda \in[0,1]$;
(3) with a fixed value of $p$ we compute $\sup _{\nu \in \Delta^{4}}\left[G(\nu)-I_{P}^{2}(\nu)\right]$;
(4) we iterate step 3 to find $p$ which maximized the supremum over $\nu$ and we use it as the optimal $p^{*}$ for the considered $r, s$ and $\lambda$;
(5) with $\lambda$ and using the optimal $p^{*}$ just found we compute also $\sup _{\nu \in \Delta^{4}}\left[2 G(\nu)-I_{P}^{2}(\nu)\right]$ and so the value of the exponent for that $\lambda$;
(6) we iterate step 2 moving $\lambda$.

With this procedure we obtain plots for $b$ shown in Figure 4.5
Instead from all the considered couples of $r$ and $s$, we have the evidence that $p *$ has the form:

- if $r>s$ then

$$
p^{*}= \begin{cases}1 & \text { if } \lambda \leq \frac{1}{2}  \tag{4.27}\\ 0 & \text { if } \lambda>\frac{1}{2}\end{cases}
$$

- if $r<s$ then

$$
p^{*}= \begin{cases}0 & \text { if } \lambda \leq \frac{1}{2}  \tag{4.28}\\ 1 & \text { if } \lambda>\frac{1}{2}\end{cases}
$$

i.e. all the population uses the strategy or moves randomly depending on the multiplicative factors and the probability of changing the environment. The two cases are shown in Figures 4.6 and 4.7.
We see from Figures 4.5c and 4.5 d that some $r$ and $s$ make $b$ to diverge. Looking for the region of divergence in the same way as done above we find the region in Figure 4.8. Dividing between $\lambda_{c}$ greater or smaller than $\frac{1}{2}$ we have Figure 4.9. As we can see we obtain the same results found above with the previous maximization.


Figure 4.5: Plots of $b(\lambda)$ for the model dynamics with the adaptive strategy and the randomly moving fraction optimized using LDT techniques with different values of $r$ and $s$.


Figure 4.6: Plot of $p^{*}(\lambda)$ for the model dynamics with the adaptive strategy and the randomly moving fraction optimized using LDT techniques with $r>s$.


Figure 4.7: Plot of $p^{*}(\lambda)$ for the model dynamics with the adaptive strategy and the randomly moving fraction optimized using LDT techniques with $r<s$.


Figure 4.8: Points generated by uniform sampling of the region $[0,3] \times[0,3]$ of the parameter space that make $b$ to display divergences for the model dynamics with the adaptive strategy and the randomly moving fraction optimized using LDT techniques.


Figure 4.9: Regions in which the generalized TL exponent diverges for the model dynamics with the adaptive strategy and the randomly moving fraction optimized using LDT techniques. The blue points make $b$ to diverge with $\lambda_{c}>\frac{1}{2}$. The red one instead provide $\lambda_{c}<\frac{1}{2}$. In this way we derived numerical evidences on what the regions could be. The small panels shows some plots obtained via numerical calculation of the exponent for a couple $r$ and $s$ in the different regions of the parameters obtained in such a way.

### 4.2.3 Markowitz optimization

The last optimization performed is Markowitz optimization. We try to minimize the ratio between the fluctuations, i.e $\sqrt{\operatorname{Var}[N(t)]}$, and the expectation value $\mathbb{E}[N(t)]$ with the constraint of having a growth or in other words $\mathbb{E}[N(t)]>1$ on the long term.
So again exploiting Varadhan's Lemma we have to minimize with respect to $p$

$$
\begin{equation*}
\lim _{t \rightarrow \infty} \frac{1}{t} \log \left[\frac{\mathbb{E}\left[N^{2}(t)\right]}{(\mathbb{E}[N(t)])^{2}}\right]=\lim _{t \rightarrow \infty} \frac{1}{t} \log \left[\mathbb{E}\left[N^{2}(t)\right]\right]-\frac{2}{t} \log [\mathbb{E}[N(t)]] \tag{4.29}
\end{equation*}
$$

asking

$$
\begin{equation*}
\lim _{t \rightarrow \infty} \frac{1}{t} \log [\mathbb{E}[N(t)]]>0 \tag{4.30}
\end{equation*}
$$

as done in 3.4.2.
Again we have to study this optimization via numerical calculation with the following procedure:
(1) we fix $r$ and $s$;
(2) we fix a value of $\lambda \in[0,1]$;
(3) with a fixed value of $p$ we compute $\sup _{\nu \in \Delta^{4}}\left[2 G(\nu)-I_{P}^{2}(\nu)\right], \sup _{\nu \in \Delta^{4}}\left[G(\nu)-I_{P}^{2}(\nu)\right]$ and $\sup _{\nu \in \Delta^{4}}\left[2 G(\nu)-I_{P}^{2}(\nu)\right]-2 \sup _{\nu \in \Delta^{4}}\left[G(\nu)-I_{P}^{2}(\nu)\right]$;
(4) we iterate step 3 to find $p^{*}$ which minimize the difference $\sup _{\nu \in \Delta^{4}}\left[2 G(\nu)-I_{P}^{2}(\nu)\right]-$ $2 \sup _{\nu \in \Delta^{4}}\left[G(\nu)-I_{P}^{2}(\nu)\right]$ with the constraint (if it possible with those $\lambda, r$ and $s) \sup _{\nu \in \Delta^{4}}\left[G(\nu)-I_{P}^{2}(\nu)\right]>0 ;$
(5) with $\lambda$ and the optimal $p^{*}$ just found we compute the value of the exponent for that $\lambda$, i.e. the ratio of the two suprema;
(6) we iterate step 2 moving $\lambda$.

Obviously, due to the inevitable discretization during the scan of values of $p$ for the search of $p^{*}$, a computational error is present. Nevertheless in this way we can estimate the optimal $p^{*}$ that minimizes the risk of extinction and that ensures a growth on the long term. Some plots are shown in Figure 4.10 .

We obtain also plots of the exponent of generalized TL displayed in Figures 4.11.

From Figures 4.11a and 4.11d we still see the presence of divergences in the exponent so we investigated the regions of the parameters as done so far finding the same results, as it is possible to see from Figures 4.12 and 4.13 Comparing the results of the three different optimization performed, we find always the same regions of divergences. Looking only at the region $0<s<1<r$ for $\lambda<\frac{1}{2}$ it is possible to see that always the same region is obtained. Moreover this region is equal to the one found in the two previous models, common also in that case among the three optimizations.


Figure 4.10: Plot of $p^{*}(\lambda)$ for the model dynamics with the adaptive strategy and the randomly moving fraction optimized using Markowitz optimization with different values of $r$ and $s$.


Figure 4.11: Plot of $b(\lambda)$ for the model dynamics with the adaptive strategy and the randomly moving fraction optimized using Markowitz optimization with different values of $r$ and $s$.


Figure 4.12: Points generated by uniform sampling of the region $[0,3] \times[0,3]$ of the parameter space that make $b$ to display divergences for the model dynamics with the adaptive strategy and the randomly moving fraction optimized using Markowitz optimization.


Figure 4.13: Regions in which the generalized TL exponent diverges for the model dynamics with the adaptive strategy and the randomly moving fraction optimized using Markowitz optimization. The blue points make $b$ to diverge with $\lambda_{c}>\frac{1}{2}$. The red one instead provide $\lambda_{c}<\frac{1}{2}$. In this way we derived numerical evidences on what the region could be. The small panels shows some plots obtained via numerical calculation of the exponent for a couple $r$ and $s$ in the different regions of the parameters obtained in such a way.

Hence, for all the evolutionary strategies we have considered in this work, the model still predict the exponent to take any real value due to the presence of regions of divergence. Moreover under some natural hypothesis on the model parameters $\left(0<s<1<r\right.$ and $\left.0<\lambda<\frac{1}{2}\right)$ the same region is found, independently of the strategies implemented or the optimization performed.

## Conclusions

The aim of this thesis was to study how simple evolutionary strategies and optimal investments of the population resources affect and maybe generate ecological patterns.
In Chapter 1 we have discussed the regularities we have been interested in, i.e. the Taylor's Law and its generalized form. The former states a powerlaw scaling property of the variance of a non-negative random variable with respect to its mean, while the latter extends the power-law relationship to all the moments of the random variable. Due to their wide diffusion across a large class of heterogeneous ecosystems (in this case the random variable is the population size), it is interesting to understand how they emerge and in particular why only a restricted range of values has been observed for the exponents of these power-laws. Indeed somehow independently of the ecosystem considered, most of the data sets analyzed in literature have provided a Taylor's Law exponent between 1 and 2 with a clustering around the value $b=2$, while the exponent of the generalized version which relates the $k$ th to the $j$ th moment acquires values near to $k / j, b_{j k} \simeq k / j$. Despite their robust corroboration through empirical data analysis, reliable explanations of how these regularities emerge with the same values for the exponents independently of the ecosystem considered have not yet been formulated. Hence the question under investigation in the study of this topic was: are these ecological regularities consequences of a some biological process or are mere statistical artifacts without any physical relevance?
In the same chapter we have illustrated a paper work in which the authors found that the exponent of the generalized TL relating the second and the first moment can acquire any real value. In fact, adopting the framework of multiplicative growth model, they showed the existence of a region of the model parameters in which the exponent displays a divergence. They also argued how only $b_{12} \simeq 2 / 1$ is observed thanks to an under sampling in the measurements procedure. That is they suggested TL and generalized TL are statistical artifacts.
In Chapters 2,3 and 4 we instead have introduced in the multiplicative growth model new dynamics features trying to accounting for the effects of ecological processes on the emergence of the studied patterns. In fact we have allowed the population to follow three different evolutionary strategies and to perform different possible optimizations of its living resources to accomplish ecological goals. With this way of reasoning we have explored the possibility of the statitical patterns to be due to some physical mechanisms.

To compute an expression of the exponent relating the second to the first moment (the same studied in the work of Chapter 1), LDT results have had to been adopted. While in Chapters 2 and 3 analytical calculations have been carried out, in Chapter 4 numerical calculations have been necessary in order to investigate the exponent. So in the first two cases we have directly analyzed the exponent handling its explicit expression, while in the last considered scenario the behaviour of the exponent has been studied via numerical procedures.
In all the dynamics analyzed and independently of the optimization objective considered, regions of the model parameter space in which the exponent diverges have been found. So also in our case, as in the model presented in Chapter 1, we predict an unbounded range of values obtaining a disagreement with the evidences coming from the data. Hence from the analysis of these model dynamics we can conclude that the simple evolutionary strategies we considered can not be the origin of the statistical regularities we have looked at and their limited possible values of the scaling exponents. Another interesting result is that, under some considerations on the model dynamics and parameters in order to reproduce in the most faithful way ecological meaningful scenarios, every considered strategy and also the model of the quoted paper have given the same (smaller) region of divergence. This may suggest us the dynamics details do not have any remarkable effect on the exponent, while Large Deviation and rare events seem to matter more making TL to appear as a meaningless statistical artifact. This assertion has to be verified and so further studied in this direction are needed. For the moment we can only exclude evolutionary strategies to be the reasons of the empirical (and still unexplained) observations.

## Appendix A

## Large Deviations Theory

In the main text we built up stochastic models based on Markov chains and we mathematically analysed exploiting Large Deviation Theory (LDT) techniques. In this appendix we briefly framed Large Deviations for Markov chains and illustrated some general results needed during the work of Thesis. For a further and detailed overview of this theory we suggest to look at the references [42], [43], [44] and [45].

## A. 1 Ergodic Theorem

The Markov chain we considered was a Markov chain with two dimensional state-space $\Gamma=\{0,1\}$ and symmetric transition matrix $P$ with $P_{i j}>0$ $\forall i, j \in \Gamma$. So we can see that $P$ is irreducible and there is only one invariant distribution $\pi$ that, due to the bistochatic form of $P$, is $\pi=\left\{\pi(i)=\frac{1}{2}\right\}$ $\forall i \in \Gamma$.
Our chain is a specific case of the family of ergodic, finite state-space Markov chains. So let be

$$
X_{i}, \quad i=1,2,3, \cdots
$$

with

$$
X_{i} \in \Gamma=\{1,2, \cdots, h\}, \quad h \in \mathbb{N}
$$

fixed and let $\pi$ be its unique invariant distribution. For such a class of chains an important theorem holds: the so called ergodic theorem.
Theorem 1 (Ergodic Theorem). Let $P$ be irreducible and $\alpha$ a generic distribution. If $\left(X_{i}\right)_{i \geq 0}$ is Markov $(\alpha, P)$ with finite state- space $\Gamma$, then for any function $f: \Gamma \rightarrow \mathbb{R}$

$$
\begin{equation*}
\mathbb{P}\left(S_{n}=\frac{1}{n} \sum_{i=1}^{n} f\left(X_{i}\right) \rightarrow \mathbb{E}_{\pi}(f)=\sum_{i=1}^{h} \pi(i) f(i)\right) \xrightarrow[n \rightarrow+\infty]{ } 1 \tag{A.1}
\end{equation*}
$$

In other words, this theorem states that for any function $f$, independently of the initial condition (i.e. the initial distribution $\alpha$ ), the empirical sample average of $f$

$$
S_{n}=\frac{1}{n} \sum_{i=1}^{n} f\left(X_{i}\right)
$$

the probabilistic average of $f$

$$
\mathbb{E}_{\pi}(f)=\sum_{i=1}^{h} \pi(i) f(i)
$$

are equal in the limit of $n \rightarrow+\infty$.
The statement of Ergodic Theorem can be formulated into the form of probability decay:

$$
\begin{equation*}
\mathbb{P}\left(\left|S_{n}-\mathbb{E}_{\pi}(f)\right| \geq x\right) \xrightarrow[n \rightarrow+\infty]{ } 0, \quad \forall x>0 \tag{A.2}
\end{equation*}
$$

In this way we see that increasing $n$ is more and more unlikely to observe $S_{n}$ taking a different value from $\mathbb{E}_{\pi}(f)$, but it is not established how fast this probability goes to 0 .
In the case in which also the low probability situations (the rare events) has to be taken into account it is important to be able to compute their probabilities and so it is mandatory to find how the decay displayed in Equation (A.2) behaves. The Large Deviations Theory provides the answer to this question.

## A. 2 Rate Function and Large Deviation Principle for the sample average

The precise statement is based on the so called rate function. To introduce the rate function a previous definition is needed.

Definition 1 (Scaled cumulant generating function). Let us consider a real random variable $A_{n}$ parametrized by the positive integer $n$. We define the scaled cumulant generating function of $A_{n}$ as

$$
\begin{equation*}
\phi(\theta)=\lim _{n \rightarrow+\infty} \frac{1}{n} \log \mathbb{E}\left[e^{n \theta A_{n}}\right] \tag{A.3}
\end{equation*}
$$

with $\theta \in \mathbb{R}$.
In the framework of Markov chain and Ergodic Theorem the random variable is the sample average $S_{n}$ and so

$$
\phi(\theta)=\lim _{n \rightarrow+\infty} \frac{1}{n} \log \mathbb{E}\left[e^{n \theta S_{n}}\right]
$$

With this function $\phi(\theta)$ it is possible to introduce the rate function.
Definition 2 (Rate Function). We define the rate function for the sample averaged of $f S_{n}$ as the Legendre-Fenchel trasformation of it scaled cumulant generating function, i.e.

$$
\begin{equation*}
I(x)=\sup _{\theta \in \mathbb{R}}[\theta x-\phi(\theta)] \tag{A.4}
\end{equation*}
$$

From this definition several properties of the rate function $I(x)$ follow:

- it is positive, $I(x) \geq 0 \forall x$;
- it is convex, $\frac{d^{2} I}{d x^{2}}>0$;
- it has an absolute point of minimum. This is $x_{\text {min }}=\mathbb{E}_{\pi}(f)$ and $I\left(x_{\text {min }}\right)=0$.

Given $I(x)$, LDT says that, given a set $B \subset \mathbb{R}$,

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} \frac{1}{n} \log \mathbb{P}\left(S_{n} \in B\right)=-\inf _{x i n B} I(x) \tag{A.5}
\end{equation*}
$$

This is called Large Deviations Principle (LDP). Starting from LDP, for $n \gg 1$ we can write

$$
\begin{equation*}
P\left(S_{n} \in B\right) \approx e^{-n \inf _{x \in B} I(x)} \tag{A.6}
\end{equation*}
$$

So finally we obtain the answer we were looking for: the decay in $n$ is exponential with rate equal to infimum of the rate function.
It is possible to remark that if $\mathbb{E}_{\pi}(f) \in B$ the Ergodic theorem is recovered and so LDT for Markov chains can be seen as an extension that result.
For the case of ergodic Markov chains with finite space state $\Gamma$ the rate function for the sample average of a real function $f$ can be computed explicitly. Let us call $P=\left(P_{i j}\right)_{i, j \in \Gamma}$ the transition matrix and $\mathcal{F}$ the set of functions from $\Gamma$ into the set of real numbers $\mathbb{R}$. We introduce the linear operator

$$
\begin{gathered}
T_{\theta}: \mathcal{F} \longrightarrow \mathcal{F} \\
g \longmapsto T_{\theta} g(i)=\sum_{j \in \Gamma} e^{\theta f(i)} g(i) M_{i j}
\end{gathered}
$$

and we look for its the largest eigenvalue

$$
\lambda(\theta)=\sup _{g:\|g\| \leq 1}\left\|T_{\theta}(g)\right\| .
$$

Then it easy to see that the rate function is

$$
I(x)=\sup _{\theta \in \mathbb{R}}[\theta x-\log \lambda(\theta)]
$$

## A. 3 Gärtner-Ellis Theorem

So far we discussed LDT for a restricted family random variables (the sample average of a real function taking value from the space state of a Markov chain). LDT can be extended to a broader class of random variables under certain assumption thanks the to Gärtner-Ellis Theorem.
Let $Y_{n}$ with $n=1,2,3, \cdots$ be a sequence of real random variables. We can define the scaled cumulant generating function as

$$
\begin{equation*}
\phi(\theta)=\lim _{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}\left[e^{\theta Y_{n}}\right] \tag{A.7}
\end{equation*}
$$

with $k \in \mathbb{R}$.

Theorem 2 (Gärtner-Ellis Theorem). Given a sequence of random variables $\left(Y_{n}\right)_{n \in \mathbb{N}}$, if $\phi(\theta)$ exists and is differentiable for all $\theta \in \mathbb{R}$, then $P_{n}=$ $\mathbb{P}\left(\frac{Y_{n}}{n} \in B\right)$ satisfies a $L D P$, i.e.

$$
\lim _{n \rightarrow \infty} \frac{1}{n} \log P_{n}=-\inf _{x \in B} I(x)
$$

with a rate function $I(x)$ given by the Legendre-Fenchel transform of $\phi(\theta)$

$$
I(x)=\sup _{\theta \in \mathbb{R}}[\theta x-\phi(\theta)] .
$$

The importance of this theorem is clear: it gives a criterion easy to evaluate to establish if a sequences of generic random variable satisfies a LDP and to calculate its rate function looking only at $\phi(\theta)$ and its features. So it is no longer necessary to compute

$$
P_{n}=\mathbb{P}\left(A_{n} \in B\right)
$$

and the limit of the logarithm of this quantity. This is a very powerful result since sometimes $P_{n}$ cannot be explicitly evaluated. With GärtnerEllis Theorem no restrictions on the sequence are made. Instead requests are made on the sequence of functions $\phi_{n}(\theta)$.

## A. 4 Varadhan's Lemma

After a general introduction to LDT, now we illustrate two results employed in the main text. The first is Varadhan's Lemma
Theorem 3 (Varadhan's Lemma). Let $P_{n}$ be a sequence of probability measures satisfying a LDP on $\Gamma$ and continuous and limited real function $F: \Gamma \rightarrow \mathbb{R}$, then

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} \frac{1}{n} \log \int_{\Gamma} e^{n F(x)} P_{n}(d x)=\sup _{x \in \Gamma}[F(x)-I(x)] \tag{A.8}
\end{equation*}
$$

Basically this Lemma is a saddle-point or Laplace's approximation exploiting LDP (i.e. $P_{n}(d x) \asymp e^{-n I(x)}$, the probability measure asymptotically in the limit $n \rightarrow+\infty$ has an exponential form) that becomes exact in the limit $n \rightarrow+\infty$.
We employed this when we needed to compute $\lim _{t \rightarrow+\infty} \frac{1}{t} \log \mathbb{E}\left[N^{k}(t)\right]$. Since we shown that $N(t)$ can be written as

$$
N(t)=e^{t G(\mu)}
$$

with $\mu$ satisfying a LDP with rate function $I(\mu)$. So we had

$$
\begin{aligned}
\lim _{t \rightarrow+\infty} \frac{1}{t} \log \mathbb{E}\left[N^{k}(t)\right] & =\lim _{t \rightarrow+\infty} \frac{1}{t} \log \mathbb{E}\left[e^{t k G(\mu)}\right]= \\
& =\lim _{t \rightarrow+\infty} \frac{1}{t} \log \int_{0}^{1} e^{t k G(\mu)} P_{t}(d \mu)= \\
& =\sup _{\mu \in[0,1]}[k G(\mu)-I(\mu)] .
\end{aligned}
$$

## A. 5 Contraction Principle

Another theorem we employed is the Contraction Principle. This theorem deals with the problem of finding the rate function of a sequence of probability measures obtained as the image through a continuous function of a sequence of another probability measures satisfying a LDP with a known rate function.

Theorem 4 (Contraction Principle). Let $\left(P_{n}\right)$ be a sequence of probability measures on a Polish space $\mathcal{X}$ that satisfies the LDP with rate $n$ and rate function I. Let

$$
\begin{cases}\mathcal{Y} & \text { be a Polish space } \\ T: \mathcal{X} \rightarrow \mathcal{Y} & \text { a continous map } \\ Q_{n}=P_{n} \circ T^{-1} & \text { an image probability measure } .\end{cases}
$$

Then $\left(Q_{n}\right)$ satisfies the LDP on $\mathcal{Y}$ with rate $n$ and rate function $J$ given by

$$
\begin{equation*}
J(y)=\inf _{x \in \mathcal{X}: T(x)=y} I(x), \tag{A.9}
\end{equation*}
$$

with the condition $\inf _{\emptyset}=\infty$.

## Bibliography

[1] L.R. Taylor; Aggregation, variance and the mean. Nature 189(4766):732-735, 1961.
[2] P. Longhin; Rare events simulation in models for ecology. Master thesis in Physics, University of Padova, a.a. 2015/2016.
[3] L.R. Taylor, R.A.J. Taylor; Aggregation, migration and population dynamics. Nature 265: 415-421, 1977.
[4] B.R.R. Azevedo, A.M. Leroi; A power law for cells. Proceedings of the National Academy of Science of the United States of America, vol. 98, number 10:5699-5704, 2001.
[5] W.S. Kendal, P. Frost; Experimental metastasis: a novel application of the variance-to-mean power function. Journal of the National Cancer Institute 79, 1113-1115, 1987.
[6] W.S. Kendal; An exponential dispersion model for the distribution of human single nucleotide polymorphisms. Molecular Biology and Evolution 20, 579-590, 2003.
[7] M.J. Keeling, B. Grenfell; Stochastic dynamics and a power law for measles variability. Philosophical Transactions of the Royal Society of London B , 769-776, 1999.
[8] R.M. Anderson, R.M.May; Epidemiological parameters of HIV transmission. Nature 333, 514-519, 1989.
[9] P.Philippe; The scale-invariant spatial clustering of leukemia in San Francisco. Journal of Theoretical Biology 199, 371-381, 1999.
[10] Z.Q. Jiang, L. Guo, W.X. Zhou; Endogenous and exogenous dynamics in the fluctuations of capital fluxes. An empirical analysis of the Chinese stock market. The European Physical Journal B 57, 347-355, 2007.
[11] J. Duch, A. Arenas; Scaling of Fluctuations in Traffic on Complex Networks. Physical Review Letters 96, 218702, 2006.
[12] M.A. de Menezes, A.L. Barabasi, Physical Review Letters 92, 28701, 2004.
[13] J.B. Bassingthwaighte; Fractal nature of regional myocardial blood flow heterogeneity. Circulation Research 65, 578-590, 1989.
[14] S. Galluccio, G. Caldarelli, M. Marsili, Y.C. Zhang; Scaling in currency exchange. Physica A 245:423-436, 1997.
[15] K. Dahlstedt, H.J. Jensen; Fluctuation spectrum and size scaling of river flow and level. Physica A 348:596-610, 2005.
[16] G. Caldarelli; Scale-Free Networks: Complex Webs in Nature and Technology. Oxford University Press, Oxford, 2007.
[17] A.M. Petersen, J. Tenenbaum, S. Havlin, H.E. Stanley; Statistical laws governing fluctuations in word use from word birth to word death. Scientific Reports 2:313, 2012.
[18] Z. Eisler, I. Bartos, J. Kertesz; Fluctuation scaling in complex systems: Taylor's law and beyond; Advances in Physics. 57.85 ,2008.
[19] M. Xu ; Taylor's power law: before and after 50 years of scientific scrutiny. ArXiv.org/abs/1505.02033v2, 2016.
[20] P.A. Marquet et al; Scaling and power-laws in ecological systems. Journal of Experimental Biology 208(Pt 9):1749-1769, 2005.
[21] A. Fronczak , P. Fronczak; Origins of Taylor's power law for fluctuation scaling in complex systems. Physical review. E, Statistical, nonlinear, and soft matter physics, 81(6 Pt 2):066112, 2010.
[22] J.E. Cohen, M. Xu, W.S.F. Schuster; Stochastic multiplicative population growth predicts and interprets Taylor's power law of fluctuation scaling. Proceedings of the Royal Society B: Biological Sciences 280(1757):20122955, 2013.
[23] J. Ramsayer, S. Fellous, J.E. Cohen J, M.E. Hochberg; Taylor's Law holds in experimental bacterial populations but competition does not influence the slope. Biology Letters, 8(2):316-319, 2012.
[24] R.M. Anderson et al; Variability in the abundance of animal and plant species. Nature 296: 245-248, 1982.
[25] A. Giometto, F. Altermatt, F. Carrara, A. Maritan, A. Rinaldo; Scaling body size fluctuations. Proceedings of the National Academy of Science of the United States of America, vol. 110, number 12:4646-4650, 2013.
[26] X. Xiao, K.J. Locey, E.P. White; A process-independent explanation for the general form of Taylor's law. The American Naturalist 186, E51-E60, 2015.
[27] W.S. Kendal, B. Jorgensen; Taylor's power law and fluctuation scaling explained by a central-limit-like convergence. Physical Review E, 83.066115, 2011.
[28] A.M. Kilpatrick, A.R. Ives; Species interactions can explain Taylor's power law for ecological time series. Nature 422, 65-68, 2003.
[29] J.R. Banavar, J. Damuth, A. Maritan, A. Rinaldo; Scaling in ecosystems and the linkage of macroecological laws. Physical Review Letters 98(6):068104, 2007.
[30] W.S. Kendal; A scale invariant clustering of genes on human chromosome 7. BMC Evolutionary Biology, 2004.
[31] B. Jørgensen; Exponential dispersion models. Journal of the Royal Statistical Society B 49(2):127-162, 1987.
[32] I. Hanski; On patterns of temporal and spatial variation in animal populations. Annales Zoologici Fennici 19(1):21-37, 1982.
[33] M.J. Keeling; Simple stochastic models and their power-law type behaviour. Theoretical Population Biology 58(1):21-31, 2000.
[34] A.M. Kilpatrick, A.R. Ives; Species interactions can explain Taylor's power law for ecological time series. Nature 422(6927):65-68, 2003.
[35] S. Engen, R. Lande, B.E. Saether; A general model for analyzing Taylor's spatial scaling laws. Ecology 89(9):2612-2622, 2008.
[36] M. Kalyuzhny et al; Temporal fluctuation scaling in populations and communities. Ecology 95(6):1701-1709, 2014.
[37] J.E. Cohen; Taylor's power law of fluctuation scaling and the growthrate theorem. Theoretical Population Biology 88:94-100, 2013
[38] J.E. Cohen; Taylor's law and abrupt biotic change in a smoothly changing environment. Theoretical Ecology 7(1):77-86, 2014.
[39] J.E. Cohen; Stochastic population dynamics in a Markovian environment implies Taylor's power law of fluctuation scaling. Theoretical Population Biology 93:30-37, 2014.
[40] J. Jiang,D.L. DeAngelis, B. Zhang, J.E. Cohen; Population age and initial density in a patchy environment affect the occurrence of abrupt transitions in a birth-and-death model of Taylor's law. Ecological modelling 289:59-65, 2014.
[41] A. Giometto, M. Formentin, A. Rinaldo, J.E. Cohen, A. Maritan; Sample and population exponents of generalized Taylor's law. Proceedings of the National Academy of Science of the United States of America, vol. 110, number 25, 2015.
[42] F. den Hollander; Large Deviations. Fields Institute Monographs, 2000.
[43] H. Touchette; The large deviation approach to statistical mechanics. Physics Reports 478(1-3):1-69, 2009.
[44] A. Dembo, O. Zeitouni; Large Deviations Techniques and Applications, Stochastic Modelling and Applied Probability. Springer, Berlin Heidelberg, 2009.
[45] M. Vidyasagar; An elementary derivation of the large deviation rate function for finite state Markov chains. 2010.

