

# The role of fluctuations in ecological patterns and processes

THÈSE N° 6541 (2015)

PRÉSENTÉE LE 24 FÉVRIER 2015

À LA FACULTÉ DE L'ENVIRONNEMENT NATUREL, ARCHITECTURAL ET CONSTRUIT  
LABORATOIRE D'ÉCOHYDROLOGIE  
PROGRAMME DOCTORAL EN GÉNIE CIVIL ET ENVIRONNEMENT

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Andrea GIOMETTO

acceptée sur proposition du jury:

Prof. R. Bernier-Latmani, présidente du jury  
Prof. A. Rinaldo, Prof. F. Altermatt, directeurs de thèse  
Prof. O. Cordero, rapporteur  
Prof. P. De Los Rios, rapporteur  
Prof. R. Solé, rapporteur



ÉCOLE POLYTECHNIQUE  
FÉDÉRALE DE LAUSANNE

Suisse  
2015



*Io stimo più il trovar un vero,  
benché di cosa leggiera,  
che 'l disputar lungamente  
delle massime questioni  
senza conseguir verità nissuna*  
Galileo Galilei, in *Le Opere*, IV, 1894, p.738

To my family and mentors



# Acknowledgements

First and foremost I wish to thank Andrea Rinaldo and Florian Altermatt for their enthusiastic support and continuous encouragement during my PhD. Since my first visit to EPFL and Eawag in May 2011 it was clear that I had a great opportunity in front of me, and it proved to be so. The last three years contributed immensely to my personal and professional development and I am particularly grateful to Andrea and Florian for having taught me that the key to good research is finding the right questions and being eager to answer them. I am indebted to them for their unconditional support and contribution in pursuing the curiosity-driven research presented here and I am glad that I could interact so closely with them. I gratefully acknowledge the funding by Eawag and the ECHO lab that supported my PhD.

I wish to thank the Galilean School of Higher Education. The School supported my bachelor and master studies, set the foundations for my choice of pursuing an interdisciplinary PhD and allowed me to get into contact with Andrea Rinaldo, Amos Maritan, and many other inspiring professors working at the interface between the physical and the life sciences. The five years in Padova were instrumental for the development of my *forma mentis* and the School played a big part in it.

I am grateful to my friends and colleagues Enrico Bertuzzo, Francesco Carrara, Marco Formentin, Lorenzo Mari and Amos Maritan for their constructive suggestions and precious help during my research. A special thank you goes to my friends and colleagues in Lausanne and at Eawag, who made the everyday life both interesting and fun. Thank you Roman Alther, Emeline Archambault, Luca Carraro, Serena Ceola, Stephanie Cohen, Kristy Deiner, Gözde Ergin, Flavio Finger, Emanuel Fronhofer, Pravin Ganesanandamoorthy, Melany Gilis, Lorenzo Gorla, Michèle Heeb, Catherine Hoffman, Regula Illi, Allyn Knox, Elvira Mächler, Rob Modini, Iñaki Navarro, Holly Oldroyd, Javier Perez-Saez, Pierre Queloz, Lorenzo Righetto, Daniele Rinaldo, Anna Rothenbühler, Giulia Ruggeri, Bettina Schaefli, Mat Seymour, Bernard Sperandio and Samir Suweis.

Finally, a big thank you goes to my family, who always supported and encouraged me, to my friends in Italy who are always next to me and to Francesca for her love and presence.

*Lausanne, 18 December 2014*

A. G.



# Preface

The present Thesis is submitted in fulfillment of the requirements of the degree of Docteur ès Sciences at the École Doctorale in Civil and Environmental Engineering (EDCE) of the École Polytechnique Fédérale de Lausanne (EPFL). It contains the result of outstanding scientific work carried out by the candidate Andrea Giometto. The candidate conducted the work both at the Laboratory of Ecohydrology (ECHO) within the School of Architecture, Civil and Environmental Engineering (ENAC), EPF Lausanne under the supervision of Prof. Andrea Rinaldo and at the Department of Aquatic Ecology at Eawag under the co-supervision of Prof. Florian Altermatt. The Thesis work spans a period of slightly more than three years (October 2011 to January 2015). The Thesis is a combination of both extensive theoretical and experimental results, and this combination results in a broad range of topics addressed, and leads to important advancements in science. The theoretical work was supervised by Prof. Andrea Rinaldo and the experimental work was supervised by Prof. Florian Altermatt. The main studies, on which all results of the thesis herein are based, have been carried out either at EPFL in Lausanne or at Eawag in Dübendorf under the joint supervision of the thesis director and co-director.

The Thesis is organized in five independent Chapters, preceded by a general introduction. The general introduction outlines the conceptual framework that embeds the various issues studied. The five chapters will each stand as independent peer-reviewed publication. Results of Chapter 1 and 5 are both already published in two different issues of the *Proceedings of the National Academy of Sciences of the United States of America*. The Chapters are ordered in a sequential form, looking at different aspects of invasion dynamics in homogeneous (Chapter 1) to heterogeneous landscapes (Chapters 2 and 3), addressing population fluctuations induced by temporal environmental stochasticity in the framework of Taylor's law of fluctuation scaling (Chapter 4), and investigating the source of heterogeneity (here: body size) in the search for general properties across different levels of biological organization, commonly described by scaling laws (Chapter 5). The Thesis ends with general conclusions and an outlook to future related research.

The Thesis uses tools from statistical mechanics in combination with state-of-the-art protist microcosm experiments. As such, the Thesis assesses various longstanding

## Preface

---

issues in ecology, getting original and relevant results that have already attracted the interest of the field. The main topic of the Thesis is the characterization and impact of the main sources of fluctuations on relevant ecological patterns and processes. Specifically, the Thesis gives the experimental test of predictions on the spatial spread of organisms in the so-called Fisher-Kolmogorov framework and studies the significance of demographic and environmental stochasticity on a very timely topic in ecology, i.e. biological invasions, possibly in heterogeneous environments, resourcewise or as far as their support matrix is concerned. The combination of extensive theoretical models and four independent microcosm experiments result in an exceptionally comprehensive and broad Thesis work. The Thesis reflects the candidate's skills to conduct experiments and link them to complex theoretical models and simulations across a broad range of topics in ecology in the wider sense. The candidate also exhibits highest talents in writing and formulating complex circumstances and data in a very scholarly and educative manner. Each Chapter contains independent sets of conclusions, putting forth perspectives and further possible developments. The original references for the independent Chapters 1 to 5, unambiguously attributable to a leading role of the candidate, are:

- **Chapter 1:** Giometto, A., Rinaldo, A., Carrara, E., and Altermatt, F. (2014). Emerging predictable features of replicated biological invasion fronts. *Proceedings of the National Academy of Sciences of the United States of America*, 111:297–301.
- **Chapter 2:** Giometto, A., Altermatt, F., Maritan, A., Stocker, R., and Rinaldo, A. A generalized receptor law governs phototaxis in phytoplankton. Manuscript submitted.
- **Chapter 3:** Giometto, A., Altermatt, F., and Rinaldo, R. Biological invasions in autocorrelated heterogeneous landscapes. Manuscript in preparation.
- **Chapter 4:** Giometto, A., Formentin, M., Rinaldo, A., Cohen, J. E., and Maritan, A. Sample and population exponents of generalized Taylor's law. Manuscript submitted.
- **Chapter 5:** Giometto, A., Altermatt, F., Carrara, E., Maritan, A., and Rinaldo, A. (2013). Scaling body size fluctuations. *Proceedings of the National Academy of Sciences of the United States of America*, 100:4646–4650.

The candidate, with a background in statistical mechanics, expressed his talents in bridging disciplines and integrating questions as well as tools from both physics and biology. Such an interdisciplinary approach is rare in general, and exceptional for a single Thesis work. The candidate has an extensive overview of the relevant ecological literature. He managed to blend theoretical and conceptual findings with empirical



data in a noteworthy manner, and to present the results, both in the figures as well as in the writing, in a scholarly manner. With the submission of this Thesis, the candidate and excellent scientist Andrea Giometto will receive the worthy academic honors.

### **Originality**

The present Thesis fulfills the requirements regarding originality as well as relevance required by the academic standards of EPFL and of the Doctoral School EDCE. The wide range of questions addressed under one common framework (characterization and impact of main sources of fluctuations on ecological patterns and processes), the diversity and technical demands of the methods employed, and integration with the relevant current literature result in an exemplary Thesis work that we recommend for acceptance without any reservations and with the highest expectations for the ensuing academic career<sup>1</sup>.

Lausanne and Dübendorf, 10 December 2014

Prof. Dr. Andrea RINALDO  
Thesis Director

and

Prof. Dr. Florian ALTERMATT  
Thesis Co-Director

---

<sup>1</sup>The research and the doctoral position of the candidate at EPFL have been funded by the Eawag discretionary funds (October 2011 to September 2014) and by ECHO funding (October 2014 to February 2015).



# Abstract

Fluctuations are ubiquitous in nature and are relevant for nearly every ecological process. The main sources of fluctuations in population abundances are demographic and environmental stochasticity, whose effect on local population dynamics, metapopulations and metacommunities have attracted much interest in the ecological literature. A third source of stochasticity is demographic heterogeneity, which is the variability of demographic traits within a population. Despite the large body of literature dedicated to fluctuations in ecology, their role in some relevant ecological patterns and processes is still rather unexplored. For example, the effect of demographic and environmental stochasticity on species spread is poorly understood, mostly due to a scarcity of experimentation linking theoretical models with replicated experiments. Additionally, environmental stochasticity can induce population fluctuations and has been shown theoretically to determine the exponent of one of the most widespread scaling laws in nature, Taylor's law of fluctuation scaling. However, empirical observations point towards the existence of a single universal Taylor's law exponent, in contrast with such model predictions. Here, experiments with protist microcosms and methods from statistical physics are used to investigate the role of fluctuations and heterogeneity on relevant ecological patterns and processes. The effect of demographic and environmental stochasticity on the propagation of biological invasions is studied in microcosm experiments with *Tetrahymena* sp. and *Euglena gracilis* and with stochastic generalizations of the Fisher-Kolmogorov equation. Demographic stochasticity is shown to induce fluctuations in the position of the propagating front and the statistical structure of the environmental heterogeneity is shown to cause a slowing-down of the invasion front at large autocorrelation lengths. The investigation of biological invasions in environments with heterogeneous distribution of resources is performed experimentally by manipulating light, the energy resource for photosynthetic organisms. Such experimental setup is further used to study phototaxis, the directed motion of phytoplankton towards or against light sources, a process that is important for relevant ecological phenomena such as diel vertical migration. A model for phototaxis is derived from the experiments in the generalized Keller-Segel framework. Large deviations theory is used to derive a generalized Taylor's law and to elucidate the origin of a universal scaling exponent

## Abstract

---

as due to sampling rather than to the population growth process. The framework of finite-size scaling is used to characterize the demographic heterogeneity in a relevant ecological trait, the body size of individuals. Intra-specific body size distributions measured experimentally are shown to be described by a universal scaling distribution across different taxa and over four orders of magnitude in body size. Mathematical models of cell growth and division are shown to be compatible with the observed universal body size distribution.

**Key words:** Fluctuations, Heterogeneity, Environmental Stochasticity, Demographic Stochasticity, Biological Dispersal, Biological Invasions, Biological Fronts, Fisher-Kolmogorov Equation, Phototaxis, Taylor's law, Scaling, Body Size, Body Mass, Protist, Microcosm

## Sommario

Le fluttuazioni sono onnipresenti in natura e sono rilevanti per quasi ogni processo ecologico. Le principali sorgenti di fluttuazioni nell'abbondanza di popolazioni ecologiche sono la stocasticità demografica e la stocasticità ambientale, i cui effetti sulla dinamica di popolazione locale, sulle meta-popolazioni e sulle meta-comunità hanno attratto molto interesse nella letteratura ecologica. Un'ulteriore sorgente di fluttuazioni è data dalla eterogeneità demografica, che consiste nella variabilità dei tratti demografici all'interno di una popolazione. Nonostante le fluttuazioni siano state oggetto di un gran numero di studi nella letteratura ecologica, il loro effetto in alcuni processi e pattern ecologici di grande rilevanza è pressoché inesplorato. Per esempio, l'effetto della stocasticità demografica ed ambientale sulla dinamica delle invasioni biologiche è ancora poco compreso, perlopiù a causa di una scarsità di studi sperimentali che mettano in relazione i modelli matematici di dispersione biologica con i risultati sperimentali. Inoltre, la stocasticità ambientale può indurre fluttuazioni nell'abbondanza di popolazione e modelli teorici prevedono che tale stocasticità possa determinare l'esponente di una delle leggi di scala più note in natura, la legge di Taylor, in disaccordo con risultati empirici che suggeriscono l'esistenza di un singolo esponente universale. In questa tesi si fa utilizzo di esperimenti in microcosmi con protisti e di metodi della fisica statistica per studiare il ruolo delle fluttuazioni e dell'eterogeneità in processi e pattern ecologici di grande rilevanza. L'effetto della stocasticità demografica e della stocasticità ambientale sulla propagazione di invasioni biologiche è studiata tramite esperimenti con *Tetrahymena* sp. ed *Euglena gracilis* e tramite generalizzazioni stocastiche della nota equazione di Fisher-Kolmogorov. Si dimostra che la stocasticità demografica induce fluttuazioni nella posizione del fronte e che la struttura statistica dell'eterogeneità ambientale causa un rallentamento del fronte di propagazione all'aumentare della lunghezza di correlazione. Sperimentalmente, lo studio delle invasioni biologiche in ambienti con distribuzione eterogenea delle risorse si effettua tramite la manipolazione della luce, che è utilizzata come risorsa di energia per la specie autotrofa *E. gracilis*. Il setup sperimentale adoperato è utilizzato per lo studio della fototassi, il movimento di organismi in direzione della luce o in direzione contraria, un processo importante per fenomeni ecologici rilevanti quali la migrazione verticale giornaliera. I risultati sperimentali sono utilizzati

## Sommario

---

per derivare un modello matematico per la fototassi di tipo Keller-Segel. La teoria delle grandi deviazioni è utilizzata per derivare una legge generalizzata di Taylor e per dimostrare che l'origine di un singolo esponente di scala universale osservata nei dati empirici può essere dovuta ad un problema di campionamento e non essere riconducibile alle proprietà del processo di crescita. Concetti di finite-size scaling (il comportamento di leggi di scala a taglia finita) vengono utilizzati per caratterizzare l'eterogeneità demografica di una delle variabili ecologiche più importanti, la taglia degli individui. Si dimostra che le distribuzioni intraspecifiche misurate sperimentalmente in taxa differenti e in oltre quattro ordini di grandezza sono caratterizzate da una distribuzione universale. Si dimostra che modelli di crescita e divisione cellulare sono compatibili con la distribuzione universale osservata.

**Parole chiave:** Fluttuazioni, Eterogeneità, Stocasticità ambientale, Stocasticità demografica, Dispersione biologica, Invasioni biologiche, Fronti biologici, Equazione di Fisher-Kolmogorov, Fototassi, Legge di Taylor, Leggi di scala, Taglia corporea, Massa corporea, Protista, Microcosmo

# Contents

<b>Acknowledgements</b>	<b>i</b>
<b>Preface</b>	<b>iii</b>
<b>Abstract (English/Italian)</b>	<b>vii</b>
<b>List of figures</b>	<b>xiii</b>
<b>List of tables</b>	<b>xv</b>
<b>Introduction</b>	<b>1</b>
<b>1 Invasion fronts: intrinsic fluctuations</b>	<b>9</b>
1.1 Introduction . . . . .	10
1.2 Theoretical background . . . . .	13
1.2.1 Reaction-diffusion models . . . . .	13
1.2.2 The chemical Langevin equation . . . . .	17
1.3 Results . . . . .	19
1.4 Methods . . . . .	26
1.4.1 Experiments . . . . .	26
1.4.2 Stochastic model and analysis . . . . .	29
1.4.3 The split-step integration scheme . . . . .	30
1.5 Discussion . . . . .	32
<b>2 Phototaxis in phytoplankton</b>	<b>35</b>
2.1 Introduction . . . . .	36
2.2 Theoretical background . . . . .	38
2.3 Results . . . . .	39
2.4 Methods . . . . .	51
2.4.1 Algal culture . . . . .	51
2.4.2 Density estimates and video recording . . . . .	51
2.4.3 Linear landscapes . . . . .	51
2.4.4 Light sources and light intensity profile . . . . .	52
	xi

## Contents

---

2.4.5	Numerical integration . . . . .	52
2.4.6	The expansion of the Fokker-Planck equation . . . . .	52
2.5	Discussion . . . . .	53
<b>3</b>	<b>Invasion fronts: environmental heterogeneity</b>	<b>59</b>
3.1	Introduction . . . . .	60
3.2	Experimental setup . . . . .	61
3.3	Results . . . . .	63
3.3.1	Total number of cells in the experiment . . . . .	68
3.4	Methods . . . . .	68
3.4.1	Model . . . . .	68
3.4.2	Experiment . . . . .	71
3.5	Discussion . . . . .	72
3.6	Appendix . . . . .	75
<b>4</b>	<b>Taylor's law of fluctuation scaling</b>	<b>79</b>
4.1	Introduction . . . . .	80
4.2	Results . . . . .	81
4.3	Discussion . . . . .	97
4.4	Appendix . . . . .	99
4.4.1	Analysis of the discontinuity in $b$ as a function of $r$ and $s$ . . . . .	99
4.4.2	Compatibility of Eq. 4.7 here and Eq. 8 in Cohen (2014a) . . . . .	99
4.4.3	Comparison with other demographic models . . . . .	100
<b>5</b>	<b>Scaling body size fluctuations</b>	<b>105</b>
5.1	Introduction . . . . .	106
5.2	Results . . . . .	107
5.3	Methods . . . . .	114
5.3.1	Protist cultures . . . . .	114
5.3.2	Size distributions . . . . .	115
5.3.3	Finite-size scaling distributions . . . . .	116
5.3.4	Data collapses . . . . .	120
5.3.5	Community size spectra . . . . .	121
5.4	Discussion . . . . .	122
	<b>Conclusions and perspectives</b>	<b>125</b>
	<b>Bibliography</b>	<b>146</b>
	<b>Curriculum Vitae</b>	<b>147</b>



# List of Figures

1.1	Biological invasions in uniform landscapes: schematic representation of the experiment . . . . .	20
1.2	Density profiles in the experiment and in the stochastic model . . . . .	21
1.3	Front propagation in the experiment and in the stochastic model . . . . .	23
1.4	Front position in the stochastic model . . . . .	26
1.5	Mean square displacement $\langle x^2 \rangle$ of individual trajectories . . . . .	29
2.1	Phototactic accumulation patterns of <i>E. gracilis</i> . . . . .	40
2.2	Temporal dynamics of accumulation and relaxation . . . . .	42
2.3	Computation of the phototactic potential $\phi(I)$ . . . . .	44
2.4	<i>E. gracilis</i> trajectories in uniform light and fit to the OU process . . . . .	48
2.5	<i>E. gracilis</i> trajectories in a light gradient . . . . .	50
3.1	<i>E. gracilis</i> growth curves in the presence and absence of light . . . . .	62
3.2	Front propagation in the model Eq. 3.1 . . . . .	64
3.3	LED sequences and light intensity profiles in the experiment . . . . .	66
3.4	Front propagation in the experiment . . . . .	67
3.5	Total number of cells during the experimental invasions . . . . .	69
3.6	Mean front speed in the model Eq. 3.1 with different parameters . . . . .	75
3.7	Experimental front propagation for different density thresholds, A . . . . .	77
3.8	Experimental front propagation for different density thresholds, B . . . . .	78
4.1	TL exponent $b_R = b_{12}$ and generalized exponent $b_{23}$ for different values of the transition probability $\lambda$ . . . . .	85
4.2	Plot of $I_{\Pi}(x)$ , $G(x)$ and $2G(x)$ . . . . .	86
4.3	Time evolution of the sample TL exponent $b_R$ . . . . .	88
4.4	Generalized TL: Intra-specific patterns of carabid beetles abundance . . . . .	93
4.5	Frequency histogram for the exponent $b_{jk}$ in the intra-specific generalized TL in the carabid beetles dataset . . . . .	94
4.6	Generalized TL: Inter-specific patterns of abundance of carabid beetles . . . . .	95
4.7	The critical transition probability $\lambda_c$ as a function of $r$ (with $s$ fixed) . . . . .	100
4.8	Existence of a critical transition probability $\lambda_c$ . . . . .	101

## List of Figures

---

5.1	Linear size distributions of 13 protist species . . . . .	109
5.2	Evidence for a universal single-species size distribution . . . . .	111
5.3	Robustness of the universal form of the single-species size distribution to biotic and abiotic forcings . . . . .	112
5.4	Collapse of rescaled size distributions with superimposed the size distri- bution of the multicellular organism <i>Cephalodella</i> sp. . . . .	123

# List of Tables

1.1	Best fit estimates of the wavefront speed in the experiment . . . . .	25
1.2	Experimentally measured species' traits . . . . .	25
2.1	AIC values for the best fit of all model combinations to the phototactic potential . . . . .	45
2.2	Parameters describing the phototactic response and movement dynamics of <i>E. gracilis</i> . . . . .	49
3.1	Mixed-effect test statistics . . . . .	68
3.2	Bartlett test statistics . . . . .	70
3.3	Mixed-effect test statistics for all density thresholds . . . . .	76
4.1	Generalized TL: Sample exponents in the Black Rock Forest dataset . .	91
4.2	Inter-specific generalized TL: Sample exponents for carabid beetles in woodland sites . . . . .	92
4.3	Inter-specific generalized TL: Sample exponents for carabid beetles in heath sites . . . . .	94
4.4	Statistics of estimated sample exponents in the inter-specific generalized TL on carabid beetles abundances . . . . .	96



# Introduction

Population fluctuations are amongst the most fascinating phenomena in ecology. Despite decades of investigation, the causes and consequences of such fluctuations constitute a large portion of the 100 fundamental ecological questions recently identified on the occasion of the 100th anniversary of the British Ecological Society (Sutherland et al., 2013), and thus are still at the core of ecological research. Along with promoting the study of population fluctuations, the list in Sutherland et al. (2013) underlines the need for concepts and methods to link local ecological processes to global patterns and dynamics, an issue that received increasing attention since the publication of the influential McArthur Award lecture Levin (1992). The implications of fluctuations on ecological patterns and processes and the problem of pattern and scale in ecology (Levin, 1992) are the main focus of this thesis.

Two main sources of stochasticity have been traditionally advocated to cause fluctuations in population abundances. The first is demographic stochasticity, which is due to the random processes of birth and death of individuals. The importance of demographic stochasticity in ecological theory is easily understood as it is at the core of the neutral theory of ecology (Hubbell, 2001) and is contemplated among the main drivers of species diversity and composition (Vellend, 2010). Demographic stochasticity is most relevant when the population size is small and can have deleterious effects such as extinction or genetic drift. The dynamics of small population is relevant for several ecological processes, such as dispersal and persistence. For example, population abundance at the edge of a species range is low and the local extinctions caused by demographic stochasticity contribute to the definition of the edge and may affect the dynamics of biological invasions (Giometto et al., 2014, Hallatschek and Korolev, 2009, Melbourne and Hastings, 2009). The study of the effect of demographic stochasticity on ecological processes is also important in view of human-induced fragmentation, which contributes to the creation of small and isolated populations that are more prone to extinction (Hanski, 1999, Hanski and Ovaskainen, 2000, Holyoak et al., 2005).

The second main source of stochasticity is the environment, which can affect the vital

## Introduction

---

rates of species and thus induce fluctuations in their population abundances. For example, spatially correlated density-independent fluctuations induced by the climate can cause the synchronization of two spatially separated populations, a phenomenon known as the Moran effect (Moran, 1953). Natural heterogeneity may affect species vital rates, for example via resource fluctuations in space and time, and ecological processes may interact with the heterogeneous geometrical substrate in which they take place (Bertuzzo et al., 2011). Unlike demographic stochasticity, environmental stochasticity is relevant for all populations sizes and not only for small ones. Both theoretical (Roy et al., 2005, Vasseur, 2007) and experimental (Fontaine and Gonzalez, 2005, García-Carreras and Reuman, 2011) studies have highlighted the relevance of the autocorrelation structure of environmental fluctuations for ecological dynamics. For example, Gonzalez and Holt (2002) showed that red (i.e., positively autocorrelated) temporal temperature fluctuations allowed the persistence of sink populations in an experimental protist microcosm, while white fluctuations did not. The study of ecological processes in the presence of environmental stochasticity with different levels of autocorrelation is of interest not only because environmental fluctuations are typically positively correlated (Benincà et al., 2011), but also in view of the global shift towards 'bluer' climate variables (i.e., more fluctuating) that took place in most continents over the last century (García-Carreras and Reuman, 2011).

One source of stochasticity that received less attention in the literature is demographic heterogeneity, that is, the variation of traits affecting fecundity and survival across individuals (Kendall and Fox, 2003). The difference between demographic heterogeneity and demographic stochasticity is subtle, and is best understood observing that for the former birth and death rates vary according to heterogeneous individual traits, while for the latter births and deaths are random events originating from a stochastic process with fixed demographic rates (Melbourne and Hastings, 2008). Analogous to demographic stochasticity, demographic heterogeneity is most important for small populations, where unbalances between births and deaths can have profound and possibly detrimental consequences.

This thesis investigates the implications of demographic stochasticity (Chapters 1 and 3), environmental stochasticity (Chapters 3 and 4) and demographic heterogeneity (Chapter 5) on important ecological patterns and process, mostly through a combination of laboratory experiments with protist microcosms (Holyoak and Lawler, 2005) and theoretical models inspired by statistical physics (Solé and Bascompte, 2006) and by the theory of stochastic processes (Bailey et al., 2000, Gardiner, 2006). Protist microcosms have been instrumental in the development and testing of ecological theory. In Gause (1934), a pioneering set of microcosm experiments aimed at the understanding of competition and predation in the natural environment were pre-

sented. Microcosms allow establishing a close connection between experiments and simple mathematical models and thus allow testing specific predictions of ecological models in simplified ecosystems containing solely the processes of interest. Such features were already present in Gause (1934), where the experimental results were related to the Lotka-Volterra equation for predation and competition. Because microcosms are devoid of the inevitable complexity of the natural environment, they allow establishing causal relationships in idealized ecological settings that may be impossible to investigate in the field. For example, microcosm experiments with protists allowed establishing a cause-effect relationship between landscape connectivity and biodiversity patterns (Carrara et al., 2012) by suitably manipulating the landscape connectivity. To what extent the results of microcosm experiments can be extrapolated to the natural environment is a debated open question (Gardner et al., 2001, Holyoak and Lawler, 2005). Suffice here to note that protist microcosms have already proved instrumental for testing predictions of metacommunity theory, investigating food-web and source-sink dynamics and driving the development of theoretical models (Holyoak and Lawler, 2005). Notably, experiments with protist microcosms have long been at the core of the development of population dynamics theory (Benton et al., 2007, Cadotte et al., 2005, Jessup et al., 2004, Srivastava et al., 2004).

Concepts and methods from statistical physics have found widespread application in ecology. Among the most successful examples are the concepts of scaling and universality, which had a profound impact on the understanding of macroecological patterns. Scaling laws have been observed in a wide variety of ecosystems across several levels of organization and have found applied relevance, for example in the estimation of extinction risk (Marquet et al., 2005). The observation of scaling laws in many descriptors of ecosystem dynamics has led many authors to suggest that biological systems might operate close to criticality (Hidalgo et al., 2014, Mora and Bialek, 2011). As a consequence, such systems display scale invariance, that is, they cannot be described at a single scale (Levin, 1992), and despite their apparent unrelatedness they might share similar quantitative features that would allow to catalogue them within a limited set of universality classes, akin to several systems in physics and other disciplines (Stanley et al., 2000). Theoretical frameworks borrowed from statistical physics, such as the finite-size scaling, allow to link seemingly unrelated ecological laws (Banavar et al., 2007, Southwood et al., 2006), deepening the understanding of the joint occurrence of ecological patterns. The study of ecological phenomena has contributed to the development of stochastic processes (Bailey, 1964), which now find widespread application in ecology. Most importantly, stochastic models allow reducing the dimensionality of the problem by accounting for processes operating at finer scales via suitable stochastic terms (Rosindell et al., 2012). Furthermore, the

## Introduction

---

transition from deterministic to stochastic models allows accounting for the various sources of stochasticity described above and investigating fluctuations in the process of interest. The work presented in this thesis combines the capabilities of protist microcosm experiments with tools from statistical physics that are now widely applied in the science of complex systems. The combination of these two approaches is deemed as a fruitful avenue for future research, as it allows investigating ecological processes at the mesoscopic scale by deriving theoretical predictions based on the smallest set of viable assumptions and validating them in microcosm experiment tailored to the mathematical models.

The thesis is organized as follows. The first three chapters investigate the source of variance in a fundamental ecological process: biological dispersal. The search for the processes that affect biological dispersal and the sources of variability observed in ecological range expansions is fundamental to the study of invasive species dynamics (Hastings et al., 2005), shifts in species ranges due to climate or environmental change (Parmesan et al., 1999) and, in general, the spatial distribution of species (Skellam, 1951). Dispersal is the key agent that brings favorable genotypes or highly competitive species into new ranges much faster than any other ecological or evolutionary process (Fisher, 1937, Hallatschek et al., 2007). Understanding the potential and realized dispersal is thus key to ecology in general (Sutherland et al., 2013). When organisms' spread occurs on the timescale of multiple generations, it is the byproduct of processes that take place at finer spatial and temporal scales, that are the local movement and reproduction of individuals (Andow et al., 1990, Hastings et al., 2005). The main difficulty in causally understanding dispersal is thus to upscale processes that happen at the short-term individual level to long-term and broad-scale population patterns (Andow et al., 1990, Bascompte and Solé, 1995, Levin, 1992, Sutherland et al., 2013). Furthermore, the large fluctuations observed in range expansions have been claimed to reflect an intrinsic lack of predictability of the phenomenon (Melbourne and Hastings, 2009). Whether the variability observed in nature or in experimental ensembles might be accounted for by demographic or environmental stochasticity affecting basic vital rates of the organisms involved is an open research question (Hastings et al., 2005, Melbourne and Hastings, 2009, Sutherland et al., 2013).

**Chapter 1** investigates the role of demographic stochasticity in the propagation of invasion fronts in uniform landscapes. Replicated invasions are performed in experiments with the ciliate *Tetrahymena* sp. in uniform linear landscapes. Density profiles are measured during the invasion and the mean dynamics and fluctuations of the propagating front are compared to the prediction of a generalized Fisher-Kolmogorov equation accounting for demographic stochasticity (Dornic et al., 2005, Hallatschek and Korolev, 2009), with the species' traits estimated locally in independent exper-



iments. The original contributions of this chapter are the setup and realization of the experiments and data analysis, the experimental corroboration of the Fisher-Kolmogorov prediction (Fisher, 1937, Kolmogorov et al., 1937) for the speed of the invasion front and the identification of demographic stochasticity as the main driver of fluctuations in the front position, with quantitative agreement between the experiments and the generalized Fisher-Kolmogorov equation (Dornic et al., 2005, Halatschek and Korolev, 2009).

**Chapter 2** is dedicated to the measurement and modeling of phototaxis (Jékely et al., 2008), the directed movement of algae towards or away from light, in an effort to develop experimental and mathematical techniques to subsequently investigate biological invasions in heterogeneous environments. Beyond the methodological interest of exploiting phototaxis to generate heterogeneous microcosms, the directed movement of phytoplankton towards light is of relevance for important ecological processes such as diel vertical migration and the vertical distribution of phytoplankton (Häder and Griebenow, 1988, Jékely et al., 2008). The accumulation patterns and dynamics of *E. gracilis* in the presence of external light gradients are measured in a broad range of light intensities in replicated experiments. An advection-diffusion equation describing the measured phototactic dynamics is derived in the Keller-Segel framework (Keller and Segel, 1970a, Tindall et al., 2008) and the phototactic response function is shown to obey to a generalized receptor law (Lapidus and Schiller, 1976, Tindall et al., 2008). The original contributions of this chapter are the setup and realization of the experiments and data analysis, the identification of the phototactic response function and the corresponding Keller-Segel model.

**Chapter 3** capitalizes on the results of Chapter 2 by exploiting the process of phototaxis to investigate the propagation of invasion fronts in environments with heterogeneous distribution of resources, obtained via the manipulation of the external light. The Keller-Segel framework derived in Chapter 2 is coupled to a demographic term accounting for demographic stochasticity as in Chapter 1 and the resulting stochastic reaction-diffusion equation is used to study front propagation dynamics in such heterogeneous environments. A replicated experiment with the alga *E. gracilis* is performed to compare front propagations in landscapes with different resource autocorrelation lengths. The original contributions of this chapter are the setup and realization of the experiments and data analysis and the investigation of the interplay between demographic stochasticity and environmental heterogeneity in the generalized Fisher-Kolmogorov framework.

In **Chapter 4**, population fluctuations induced by temporal environmental stochasticity are studied in the context of Taylor's law (Taylor, 1961). Taylor's law (TL), also known

## Introduction

---

as fluctuation scaling, states that the variance  $V$  of a non-negative random variable scales with the mean  $M$  via a power-law,  $V = aM^b$ . In the last 50 years, more than a thousand publications have verified TL very widely in ecology (where it describes the abundance or density of populations), physics, other natural sciences, information technology, and finance (Eisler, 2008). Empirical estimates of the exponent  $b$  often cluster around the value  $b = 2$ , which is equivalent to a constant coefficient of variation or to a constant signal-to-noise ratio. The surprisingly broad range of natural and artificial systems described by TL suggests that a context-independent mechanism may be at work. However, recent theoretical studies of population dynamics using multiplicative models predict that environmental stochasticity can affect the scaling exponent of Taylor's law and possibly cause abrupt transitions between positive and negative exponents, following smooth changes of the environmental autocorrelation (Cohen, 2014b). In Chapter 4, a broad class of population growth models are studied in combination with a stochastic environment with different degrees of temporal autocorrelation. The seemingly universal emergence of the scaling law pattern in population size fluctuations with exponent  $b = 2$ , irrespectively of the details of the population dynamics and of the forcings that drive the system, is rationalized in the context of the theory of large deviations (den Hollander, 2008). The original contributions of this chapter are the analytical derivation of a generalized TL for both sample and population exponents and the analytical computation of the sample exponents and their dependence on the total number of replicates in the sample, which highlights the universal character of the TL exponent  $b = 2$  as due to sampling rather than to the population dynamics. The data analyses on the generalized TL are also original.

**Chapter 5** studies the fluctuations of one of the most relevant ecological traits, the body size of individuals (Marquet et al., 2005, Peters, 1983). The importance of body size in shaping ecological processes can be easily understood given the plethora of scaling relationships where body size features as the independent variable (Brown and West, 2000, Marquet et al., 2005). In fact, despite the incredible complexity of living organisms (Brown and West, 2000, Labra et al., 2007) and ecosystems (Levin, 1992, Marquet et al., 2005), vital rates and ecosystem properties are often found to be simple power functions of body size, allowing the delineation of a quantitative framework bridging ecology, biogeography and physiology (Brown and West, 2000, Marquet et al., 2005). To cite a few examples, metabolic rates have been found to scale with body size with exponent  $3/4$  (Kleiber (1947), although the value of the exponent is still debated), the maximum number of individuals in a given area has been shown to scale with body size with exponent  $-3/4$  (Damuth, 1981) and the total number of species in a given area has been argued to scale with body size with exponent  $-3/4$  (Marquet et al., 2005). Chapter 5 is dedicated to the characterization of the demographic heterogeneity in body size, moving from the observation that the scaling body

size spectra typically encountered in aquatic ecosystems (Sheldon et al., 1972) require regularities of the component parts, namely the intra-specific body size distributions. Such intra-specific distributions are measured in laboratory experiments with protist microcosms, and their universal properties across different taxa and over more than four orders of magnitude in body size are discussed and interpreted in the framework of finite-size scaling (Banavar et al., 2007, Fisher and Barber, 1972). In the investigation, a new facet of Taylor's law is uncovered through the observation that the variance of intra-specific body size distributions scales quadratically with their mean body mass. The original contributions of this chapter are the setup and realization of the experiments and data analysis, the observation of the lognormal scaling of body size distributions and the observation that a simple model of cell growth and division (Diekmann et al., 1983) is compatible with the observed scaling form for the body size distributions.



# 1 Invasion fronts: intrinsic fluctuations

## Abstract

Biological dispersal shapes species' distribution and affects their coexistence. The spread of organisms governs the dynamics of invasive species, the spread of pathogens and the shifts in species' ranges due to climate or environmental change. Despite its relevance for fundamental ecological processes, however, replicated experimentation on biological dispersal is lacking and current assessments point at inherent limitations to predictability, even in the simplest ecological settings. Here, replicated experimentation on the spread of the ciliate *Tetrahymena* sp. in linear landscapes shows that information on local unconstrained movement and reproduction allows to predict reliably the existence and speed of traveling waves of invasion at the macroscopic scale. Furthermore, a theoretical approach introducing demographic stochasticity in the Fisher-Kolmogorov framework of reaction-diffusion processes captures the observed fluctuations in range expansions. Therefore, predictability of the key features of biological dispersal overcomes the inherent biological stochasticity. Such results establish a causal link from the short-term individual level to the long-term, broad-scale population patterns and may be generalized, possibly providing a general predictive framework for biological invasions in natural environments.

### 1.1 Introduction

Modeling of biological dispersal established the theoretical framework of reaction-transport processes (Méndez et al., 2010, Murray, 2004), which now finds common application in dispersal ecology (Andow et al., 1990, Holmes, 1993, Lubina and Levin, 1988, Méndez et al., 2011) and in several other fields (Campos et al., 2006, Fort, 2012, Fort and Solé, 2013, Méndez et al., 2010, Murray, 2004). Reaction-transport models are typically expressed in the form of partial differential equations that embed the two key components of the phenomena described: the movement of particles and their production. In ecology, the movement or transport of organisms has been most often described via the diffusion equation (Okubo and Levin, 2002), sometimes in the presence of advective terms (Bertuzzo et al., 2007, Lubina and Levin, 1988). The diffusion assumption defines the set of reaction-diffusion models and, in the case of logistic growth, reduces to the Fisher-Kolmogorov equation (Fisher, 1937, Kolmogorov et al., 1937). The adoption of the diffusion equation to describe animal movement embodies a number of approximations that may or may not be realistic depending on the study problem (Méndez et al., 2010). In most cases, however, even when organisms can make informed movement decisions at the local scale, the statistical patterns at the macroscopic scale are consistent with the diffusive approximation (Andow et al., 1990, Murray, 2004) and thus support the adoption of reaction-diffusion models as a broadly applicable framework. One of the drawbacks of the diffusive approximation that can influence the model predictions, especially for fast-growing and slow-moving species, is that it assumes an infinite velocity of individuals (Holmes, 1993, Méndez et al., 2010), causing local perturbations to spread instantly through the landscape. Correcting for such unrealistic behavior turns the parabolic reaction diffusion equations into hyperbolic equations, which involve a second order time-derivative that is absent in the former set (Méndez et al., 2010). However, because the consequences of the infinite-velocity assumptions are often negligible (Holmes, 1993, Méndez et al., 2010), reaction-diffusion equations are most commonly employed. The classical prediction of reaction-diffusion models (Méndez et al., 2010, Volpert and Petrovskii, 2009) is the propagation of an invading wavefront traveling undeformed at a constant speed, which can be expressed in terms of the parameters characterizing the local movement and growth. For example, in the Fisher-Kolmogorov equation (Fisher, 1937, Kolmogorov et al., 1937) the rate  $r$  of exponential increase of a population and its diffusion coefficient  $D$  are constant and the front propagation occurs at the asymptotic speed  $v = 2\sqrt{rD}$ . Generalizations of the Fisher-Kolmogorov equation introduce corrections to such speed that can depend on additional species traits (Hallatschek and Korolev, 2009, Holmes, 1993, Méndez et al., 2010) or on features of the environmental matrix (Bertuzzo et al., 2007, Campos et al., 2006, Campos and

Méndez, 2005).

Within the domain of ecology, reaction-diffusion models have been adopted to describe the spread of organisms in a variety of comparative studies. Relevant early studies of ecological data include Skellam (1951), who showed that the spread of muskrats in central Europe was consistent with the theoretical prediction that the area covered by the population increases quadratically with time, and Lubina and Levin (1988), which allowed relating the Fisher-Kolmogorov model to an essentially one-dimensional spread, that of sea otters along the coast of California. These works were instrumental in linking the model predictions concerning the spread dynamics and species distribution patterns at the large scale to real-case scenarios of range expansions. In such examples, the parameters that enter the reaction-diffusion model could not be estimated independently and the authors had to rely on estimates drawn from the onset of the spread. In few field studies (Andow et al., 1990), such parameters could be measured independently and were used to test the model predictions. The extensive use of these models (Elton, 1958, Grosholz, 1996, Hastings et al., 2005) and the good fit to observational data favored their common endorsement as a paradigm for biological dispersal. Despite the large amount of comparative studies that made use of such models, however, the study of invasion fronts received little replicated and controlled experimentation (Melbourne and Hastings, 2009, Volpert and Petrovskii, 2009), in particular concerning the link between the different scales involved (from local to global, from individuals to populations, Andow et al. (1990), Levin (1992)) and the study of variability across replicated invasions (Hastings et al., 2005, Melbourne and Hastings, 2009). Furthermore, reaction-diffusion models have been traditionally formulated in a deterministic framework that cannot provide information on the fluctuations that are embedded in the propagation process and that originate from the movement and reproduction of individuals. The need for stochastic generalizations of the classical Fisher-Kolmogorov framework is accentuated by recent experimental assessments (Melbourne and Hastings, 2009) that pointed at inherent limitations to the predictability of the phenomenon, due to its intrinsic stochasticity. Because single realizations of a dispersal event (as those addressed in comparative studies) might deviate significantly from the mean of the process, replicated experimentation is required to characterize such intrinsic fluctuations and to relate them to model predictions.

This chapter is dedicated to the study of biological invasion fronts propagating in uniform linear landscapes. The microcosm experiments performed are aimed at establishing the simplest ecological settings, thus avoiding the contribution of external sources of variability to the invasion process. In real case scenarios, such superimposed external variability might be due to environmental stochasticity (Méndez et al.,

## Chapter 1. Invasion fronts: intrinsic fluctuations

---

2011) or to the specificities of the landscape, for example the existence of physical barriers or the landscape connectivity structure (Bertuzzo et al., 2007). The idealized settings employed in experimental microcosms allow to establish a close link between the experiments and the mathematical description in terms of deterministic (Méndez et al., 2010, Murray, 2004, Volpert and Petrovskii, 2009) and stochastic (Bonachela et al., 2012, Dornic et al., 2005, Méndez et al., 2011) reaction-diffusion equations. Herein, the classical Fisher-Kolmogorov equation is shown to provide a satisfactory quantitative link between the processes at the local and short-time scale and the speed of propagating invasion fronts at the global scale. Furthermore, the fluctuations that are intrinsic to the propagation of invasion fronts and that emerge from the two processes underlying the spread, movement and demography, are characterized and modeled. A generalized Fisher-Kolmogorov equation accounting for the variability induced by demographic stochasticity gives a satisfactory quantitative prediction for the fluctuations of the front position measured in the experiments.

This chapter is organized as follows. First, a theoretical background on the mathematical modeling of biological invasions with reaction-diffusion processes is provided. Second, the classical Fisher-Kolmogorov prediction (Fisher, 1937, Kolmogorov et al., 1937) on the existence and the mean speed of traveling wavefronts is substantiated experimentally by measuring the individual components of the process at the local and short-time scale in microcosm experiments with the ciliate *Tetrahymena* sp. Third, the inclusion of demographic stochasticity in the model is shown to reproduce the observed variability in range expansions.



## 1.2 Theoretical background

### 1.2.1 Reaction-diffusion models

This Box contains known results on the mathematical modeling of biological invasions and is included here to aid the comprehension of the chapter. Please note that the results presented within this Box are not original.

Reaction-diffusion models have been shown to accurately describe the spread of organisms in many comparative studies (Andow et al., 1990, Grosholz, 1996, Lubina and Levin, 1988) and here experimentally confirmed. Models other than reaction-diffusion equations, such as integro-difference equations involving dispersal kernels, are best suited to describe dispersal of organisms that exhibit distinct reproductive and dispersive phases (Hastings et al., 2005, Melbourne and Hastings, 2009). However, for many organisms, especially those with continuous, non-overlapping generations and temporally unstructured dispersal-reproductive dynamics, the reaction-diffusion approach is highly appropriate (Grosholz, 1996, Lubina and Levin, 1988, Murray, 2004). Refinements of reaction-diffusion models have also been put forward in the literature. For instance, reaction-telegraph models were introduced to account for the finite movement speed of individuals, resulting in a correction to the wavefront speed (Holmes, 1993), which is however negligible for *Tetrahymena* sp., as it is shown in the following sections. Here, the theoretical framework of reaction-diffusion processes is briefly reviewed and the relationship to reaction-telegraph processes is discussed, with reference to the experiments described above. The interested reader can refer to specialized texts for further investigation and generalizations (Berg, 1993, Gardiner, 2006, Méndez et al., 2010).

#### The Fisher-Kolmogorov equation

The diffusion equation:

$$\frac{\partial \rho}{\partial t}(x, t) = D \frac{\partial^2 \rho}{\partial x^2}(x, t) \quad (1.1)$$

describes the evolution of the density of an ensemble of independent random walkers (Gardiner, 2006). The diffusion coefficient  $D$  can be measured as the proportionality constant that links the mean square displacement to time as

(Gardiner, 2006):

$$\langle x_t^2 \rangle = 2Dt, \quad (1.2)$$

with  $D$  as in Eq. 1.1. Macroscopically, or phenomenologically, the continuity equation in the presence of a reaction term reads:

$$\frac{\partial \rho}{\partial t} = -\frac{\partial J}{\partial x} + F(\rho). \quad (1.3)$$

Assuming proportionality between the flux  $J$  and the density gradient  $\partial \rho / \partial x$  via the diffusion coefficient one finds the so-called *reaction-diffusion* equation (Fisher, 1937, Kolmogorov et al., 1937, Méndez et al., 2010, Murray, 2004):

$$\frac{\partial \rho}{\partial t} = D \frac{\partial^2 \rho}{\partial x^2} + F(\rho). \quad (1.4)$$

If the reaction term  $F(\rho)$  is logistic one finds the Fisher-Kolmogorov equation:

$$\frac{\partial \rho}{\partial t} = D \frac{\partial^2 \rho}{\partial x^2} + r\rho \left[ 1 - \frac{\rho}{K} \right], \quad (1.5)$$

where  $\rho$  is the density of organisms,  $D$  is the diffusion coefficient of the species,  $r$  is its growth rate and  $K$  its carrying capacity. The Fisher-Kolmogorov Eq. 1.5 is probably the best known example of equation that accepts traveling wave solutions. A traveling wave is a wave that travels without change of shape, that is, the density profile along a line moves rigidly in time without deformation (Fig. 1E). Mathematically, this means that if  $u(x, t)$  is a traveling wave solution of a reaction-diffusion equation, then  $u(x, t)$  is a function of  $x - vt$ , where  $v$  is the speed of the wave, that is  $u(x, t) = u(x - vt)$ .

Dimensional analysis of Eq. 1.5 shows that the speed is  $v \propto \sqrt{rD}$ . Fisher (1937) proved that traveling wave solutions can only exist with speed  $v \geq 2\sqrt{rD}$  and Kolmogorov et al. (1937) demonstrated that, with suitable and realistic initial conditions, the verified speed of the wavefront is equal to the lower bound, that is:

$$v = v_{FK} = 2\sqrt{rD}. \quad (1.6)$$

For any concave  $F(\rho)$  in Eq. 1.4, that is,  $F(\rho) \leq \rho F'(0)$ , the front velocity has been shown to be equal to  $v_{RD} = 2\sqrt{DF'(0)}$  (Méndez et al., 2010).

### The reaction-telegraph equation

The diffusion equation has been widely employed to describe the movement of organisms (Murray, 2004, Okubo and Levin, 2002). It is clear, however, that individuals do not perform exact random walks at the microscopic scale (where ‘microscopic’ here is used to refer to the typical length scale of an organism). What is implied when adopting diffusion equations to describe movement behavior is that there exists an appropriate mesoscopic scale in which the collective behavior of organisms is indistinguishable from that of an ensemble of random walkers (Andow et al., 1990). In this section the adoption of a reaction-diffusion equation is justified and follows from the fact that the correlation time in the trajectories performed by individuals of the species *Tetrahymena* sp. is much smaller than the typical timescale of the dispersal process or, more precisely, the growth rate of the species (see Eq. 1.14).

One can describe the movement of an individual (particle) as a sequence of jumps of length  $\Delta x$  and duration  $\Delta t$ . A model for a correlated random walk was introduced in Fürth (1920) and assumes that particles move along an infinite line at a constant speed  $\gamma$ , with a probability  $\mu$  per unit time to reverse their direction of motion. Precisely, the probability for the particle to continue in the direction of motion is given by  $1 - \mu\Delta t$  and the probability to reverse its direction is  $\mu\Delta t$ , in such a way that the speed  $\lim_{\Delta x, \Delta t \rightarrow 0} \Delta x / \Delta t = \gamma$  is constant. With these assumptions (Méndez et al., 2010) one obtains the *telegraph equation* for the density of particles:

$$\frac{1}{2\mu} \frac{\partial^2 \rho}{\partial t^2} + \frac{\partial \rho}{\partial t} = \frac{\gamma^2}{2\mu} \frac{\partial^2 \rho}{\partial x^2}, \quad (1.7)$$

which is rewritten as:

$$\tau \frac{\partial^2 \rho}{\partial t^2} + \frac{\partial \rho}{\partial t} = D \frac{\partial^2 \rho}{\partial x^2}, \quad (1.8)$$

where  $\tau^{-1} = 2\mu$  is the correlation time of the turning process and  $D = \gamma^2 / (2\mu)$ . Note that Eq. 1.8 differs from the diffusion equation for the additional term  $\tau \frac{\partial^2 \rho}{\partial t^2}$ . Eq. 1.8 is an hyperbolic equation and therefore information cannot travel faster than the speed of particles  $\gamma$ . Therefore, the telegraph equation is physically more appropriate than the diffusion equation, as for the diffusion equation the probability density of finding a particle in an infinitesimal interval around  $(x, t)$  is larger than zero for all  $x$  and  $t > 0$ , that is, signals can travel at infinite speed.

For *Tetrahymena* sp., the correlation time  $\tau$  is very small, so it is argued that the term  $\tau \frac{\partial^2 \rho}{\partial t^2}$  is negligible compared to the other terms in the equation and thus the system is well described by the diffusion equation. To estimate the parameters  $\tau$  and  $D$  from experimental data one can compute the value for the mean square displacement along the line, that is:

$$\langle x^2 \rangle = \int_{-\infty}^{+\infty} dx x^2 \rho(x, t). \quad (1.9)$$

Multiplying Eq. 1.8 by  $x^2$  and integrating one has:

$$\tau \int_{-\infty}^{+\infty} dx x^2 \frac{\partial^2 \rho}{\partial t^2}(x, t) + \int_{-\infty}^{+\infty} dx x^2 \frac{\partial \rho}{\partial t}(x, t) = D \int_{-\infty}^{+\infty} dx x^2 \frac{\partial^2 \rho}{\partial x^2}(x, t), \quad (1.10)$$

which reads:

$$\tau \frac{d^2 \langle x^2 \rangle}{dt^2} + \frac{d \langle x^2 \rangle}{dt} = 2D, \quad (1.11)$$

assuming that  $\rho(x, t)$ ,  $\frac{\partial \rho}{\partial t}(x, t)$  and  $\frac{\partial^2 \rho}{\partial t^2}(x, t)$  go to zero sufficiently fast for  $x \rightarrow \pm\infty$ . Assuming further that  $\rho(x, 0) = \delta(x)$  (where  $\delta$  is the delta function) and  $\partial \rho / \partial t(x, t)|_{t=0} = 0$ , one has  $\langle x^2 \rangle|_{t=0} = d \langle x^2 \rangle / dt|_{t=0} = 0$  and:

$$\langle x^2(t) \rangle = 2Dt - 2D\tau [1 - e^{-t/\tau}]. \quad (1.12)$$

One can amend the instantaneous adjustment of the flux to the density gradient implied by Fick's first law with the introduction of a relaxation time  $\tau$ , which leads to the *reaction-telegraph equation* (Holmes, 1993, Méndez et al., 2010):

$$\tau \frac{\partial^2 \rho}{\partial t^2} + [1 - \tau F'(\rho)] \frac{\partial \rho}{\partial t} = D \frac{\partial^2 \rho}{\partial x^2} + F(\rho). \quad (1.13)$$

Eq. 1.13 can be obtained combining the telegraph Eq. 1.8 with kinetics (Méndez et al., 2010). Eq. 1.13 differs from the reaction-diffusion Eq. 1.4 for the additional term  $\tau \frac{\partial^2 \rho}{\partial t^2} - \tau F'(\rho) \frac{\partial \rho}{\partial t}$ . Solutions of Eq. 1.13 converge to solutions of the reaction-diffusion equation as  $\tau \rightarrow 0$  (Méndez et al., 2010, Zauderer, 1989). In the experiments performed in this chapter the correlation time  $\tau$  is sufficiently small to consider the process as well described by the reaction-diffusion equation. Quantitative support to this statement is given below. One can also show that the introduction of reactions in the persistent random walk equation leads to the reaction-telegraph Eq. 1.13 (Méndez et al., 2010) with  $\tau^{-1} = 2\mu$  and  $D = \gamma^2 / (2\mu)$  as

in Eq. 1.8.

Holmes (1993) studied the propagation of traveling wavefronts in the reaction-telegraph equation with logistic reaction  $F$ , estimating model parameters for several case studies. The reaction-telegraph Eq. 1.13 with logistic growth was shown (Méndez and Camacho, 1997) to accept traveling wave solutions with speed:

$$v_{RT} = \frac{2\sqrt{rD}}{1 + \tau r} = \frac{v_{FK}}{1 + \tau r} \quad (1.14)$$

if  $r\tau < 1$ , otherwise  $v_{RT} = (D/\tau)^{\frac{1}{2}}$ . In this perspective it is claimed that, in this specific case, the ratio between the speed of a reaction-telegraph and that of a reaction-diffusion equation is practically one. In fact, one finds that for *Tetrahymena* sp.  $r\tau = (2.2 \pm 0.3) \cdot 10^{-4}$  so that  $v_{RT}/v_{RD} > 0.999$ , which makes the two processes experimentally undistinguishable. Therefore, the experiments performed are compared to a reaction-diffusion equation in the rest of the chapter.

## 1.2.2 The chemical Langevin equation

This section is introduced here to justify the form of the multiplicative noise term that is used in this chapter to model demographic stochasticity at the front of the invasion. The stochastic dynamics of chemical reactions are described by the so-called chemical master equation (Gardiner, 2006). Under specific conditions that are investigated in Gillespie (2000), such chemical master equation leads to a Langevin equation in the continuum limit. Such result is used in this chapter to model demographic stochasticity in the continuum Fisher-Kolmogorov equation, by interpreting the duplication of an individual cell  $A$  as the reaction  $A \rightarrow 2A$  and its death as the reaction  $A \rightarrow \emptyset$ . To aid the comprehension of this chapter, we report here the main result of Gillespie (2000), formulated in the ecological context of species interactions. Following Gillespie (2000), consider a well-mixed ensemble of  $N$  species  $S_i$ , with  $i = 1, \dots, N$ . Such species interact through the interactions  $R_j$ , with  $j = 1 \dots, M$ . Let  $\bar{X}(t) = (X_1(t), \dots, X_N(t))$  where  $X_i$  indicates the number of individuals of species  $S_i$ . Define  $a_j(\bar{x})dt$  as the probability that one  $R_j$  interaction occurs in the time interval  $[t, t + dt]$  given  $\bar{X}(t) = \bar{x}$  and define  $v_{ji}$  as the change in the number of  $S_i$  individuals due to the interaction  $R_j$ . The

Langevin equation associated to such process is given by:

$$\frac{dX_i}{dt}(t) = \sum_{j=1}^M v_{ji} a_j(\bar{X}(t)) + \sum_{j=1}^M v_{ji} \sqrt{a_j(\bar{X}(t))} \Gamma_j(t), \quad (1.15)$$

for  $i = 1, \dots, N$ , with  $\Gamma_j(t)$   $\delta$ -correlated statistically independent Gaussian white noises. For the case of interest here (i.e. at the low densities characteristic of the propagating front),  $N = 1$  (one species) and  $M = 2$ , the interactions are the birth of individuals (here, cell duplication)  $R_1: A \rightarrow 2A$  and the death of individuals  $R_2: A \rightarrow \emptyset$ . One has  $v_{11} = +1$  and  $v_{21} = -1$ . Therefore, the corresponding chemical Langevin equation assuming duplication (birth) and death rates  $a_1(x) = bx$  and  $a_2(x) = dx$  reads (let  $\bar{X} = X_i = x$  denote the abundance of cells):

$$\begin{aligned} \frac{dx}{dt}(t) &= a_1(x(t)) - a_2(x(t)) + \sqrt{a_1(x)} \Gamma_1(t) - \sqrt{a_2(x)} \Gamma_2(t) \\ &= rx(t) + \sigma \sqrt{x} \Gamma(t), \end{aligned} \quad (1.16)$$

with  $r = b - d$ ,  $\sigma = \sqrt{b + d}$  and  $\Gamma(t)$  a  $\delta$ -correlated Gaussian white noise.

## 1.3 Results

The Fisher-Kolmogorov equation for the density of organisms  $\rho(x, t)$  (Fisher, 1937, Kolmogorov et al., 1937, Méndez et al., 2010, Volpert and Petrovskii, 2009) reads:

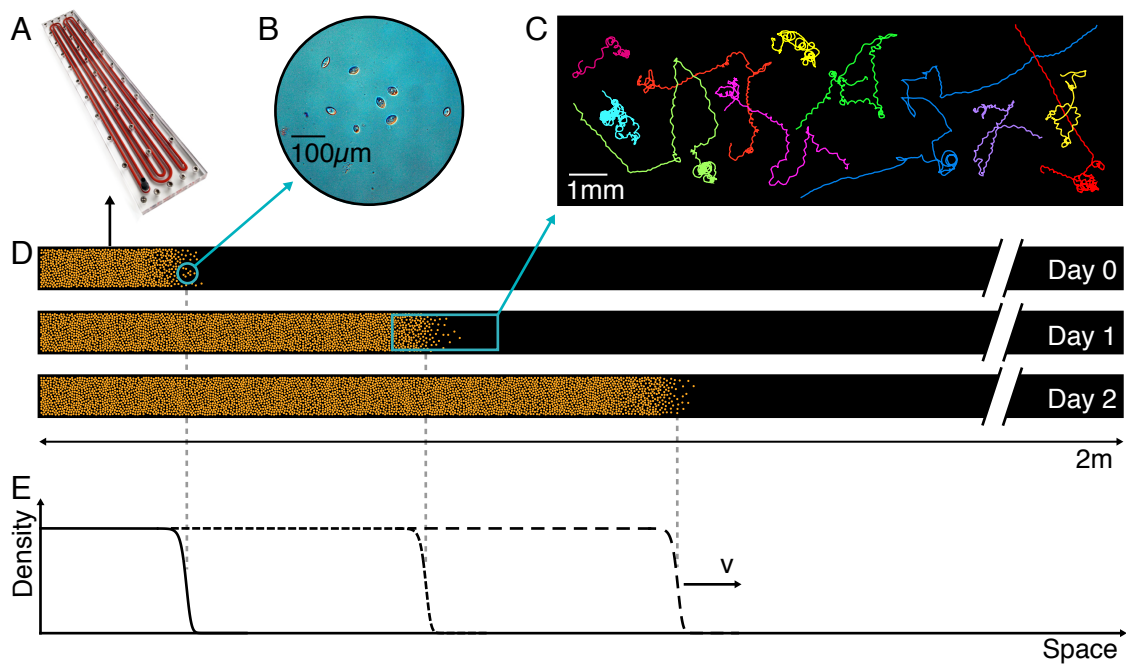
$$\frac{\partial \rho}{\partial t} = D \frac{\partial^2 \rho}{\partial x^2} + r \rho \left[ 1 - \frac{\rho}{K} \right]. \quad (1.17)$$

The equation couples a logistic term describing the reproduction of individuals with growth rate  $r$  (dimensions  $[T^{-1}]$ ) and carrying capacity  $K$   $[L^{-1}]$  and a diffusion term accounting for local movement, epitomized by the diffusion coefficient  $D$   $[L^2 T^{-1}]$ . These species' traits define the characteristic scales of the dispersal process. In this framework, a population initially located at one end of a linear landscape is predicted to form a wavefront of colonization invading empty space at a constant speed  $v = 2\sqrt{rD}$  (Fisher, 1937, Kolmogorov et al., 1937, Méndez et al., 2010, Volpert and Petrovskii, 2009), which was measured in the front propagation experiment (Fig. 1.1D and Methods).

Experiments were performed with the freshwater ciliate *Tetrahymena* sp. (Methods), which was chosen because of its short generation time (Carrara et al., 2012) and its history as a model system in ecology (Altermatt et al., 2014). The experimental setup consisted of linear landscapes (Fig. 1.1), filled with a nutrient medium, kept in constant environmental conditions and of suitable size to meet the assumptions about the relevant dispersal timescales (Methods). Replicated dispersal events were conducted by introducing an ensemble of individuals at one end of the landscape and measuring density profiles throughout the system at different times, through image analysis (Methods, Altermatt et al. (2014)). Density profiles in the six replicated dispersal events, at successive times, are shown in Fig. 1.2 (panels A-F). Collected data were binned in 5 cm bins, which corresponds to the typical length scale of the dispersal process ( $\sqrt{D/r} \approx 5$  cm). Individuals of *Tetrahymena* sp., initially localized at one end of the landscape, colonized the whole system in 4 d.

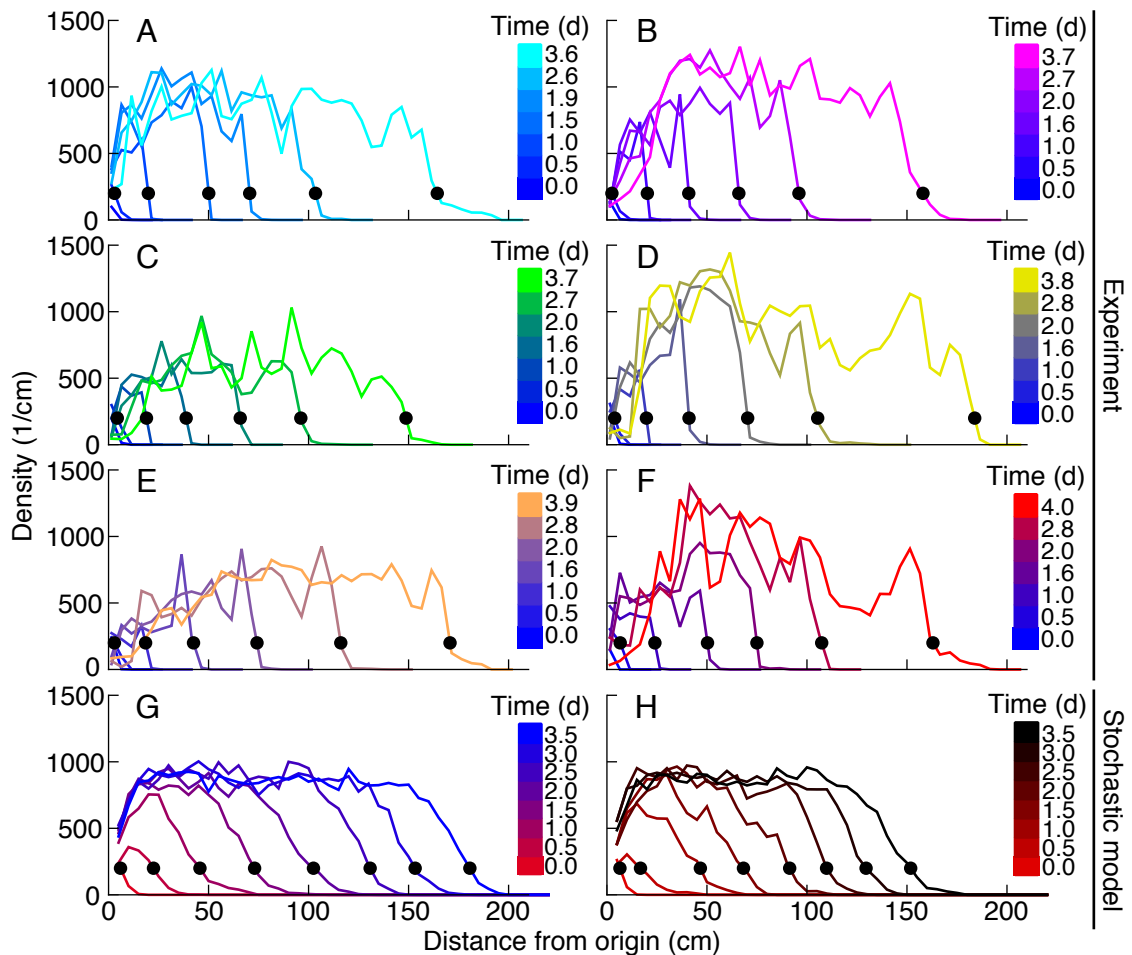
The front position at each time was calculated as the first occurrence, starting from the end of the landscape, of a fixed value of the density  $\rho^*$  (Fig. 1.3;  $\rho^* = 200 \text{ cm}^{-1}$ ). A linear model was fitted (least-squares fit) to each replicate in the linear region (day 1 to 4) to compute the front propagation speed. Table 1.1 reports the observed velocities in the six replicates (see Fig. 1.3). As for traveling waves predicted by the Fisher-Kolmogorov equation, the mean front speed in the experiment is notably constant for different choices of the reference density value (Fig. 1.3C).

The species' traits  $r$ ,  $K$  and  $D$  that enter the Fisher-Kolmogorov equation were mea-



**Figure 1.1:** Schematic representation of the experiment. (A) Linear landscape. (B) Individuals of the ciliate *Tetrahymena* sp. move and reproduce within the landscape. (C) Examples of reconstructed trajectories of individuals. (D) Individuals are introduced at one end of a linear landscape and are observed to reproduce and disperse within the landscape (not to scale). (E) Illustrative representation of density profiles along the landscape at subsequent times. A wavefront is argued to propagate undeformed at a constant speed  $v$  according to the Fisher-Kolmogorov equation.

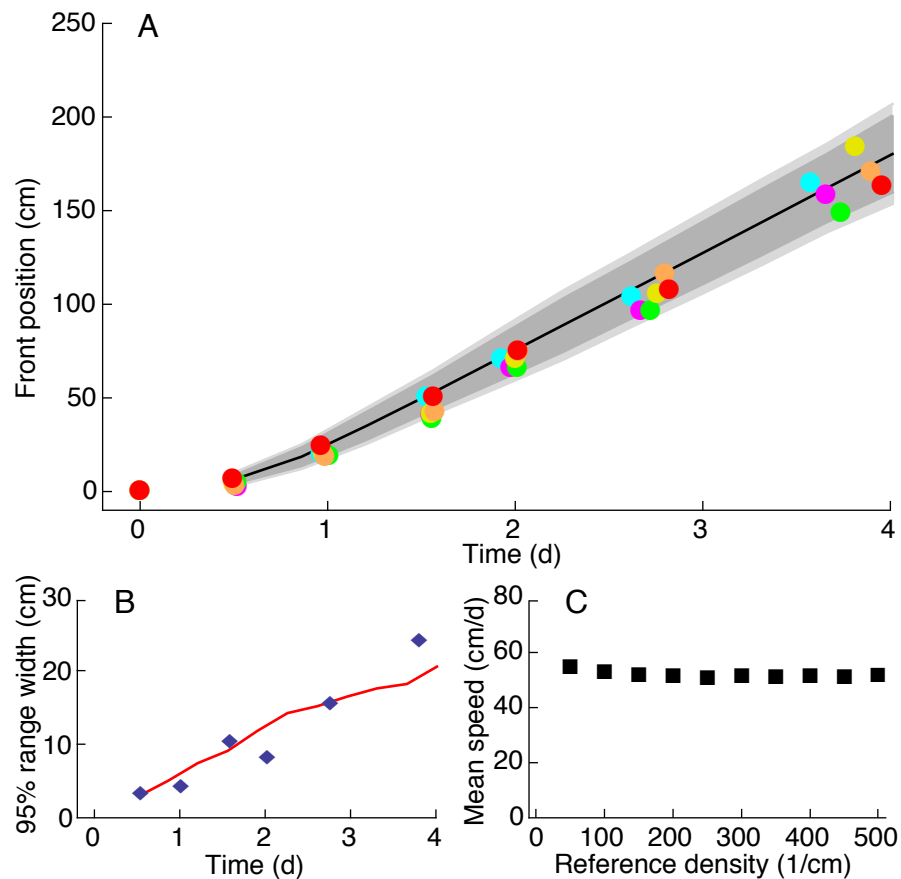




**Figure 1.2:** Density profiles in the dispersal experiment and in the stochastic model. (A-F) Density profiles of six replicated experimentally measured dispersal events, at different times. Legends link each color to the corresponding measuring time. Black dots are the estimates of the front position at each time point. Organisms were introduced at the origin and subsequently colonized the whole landscape in 4 d ( $\sim 20$  generations). (G-H) Two dispersal events simulated according to the generalized model equation, with initial conditions as at the second experimental time point. Data are binned in 5 cm intervals, typical length scale of the process.

sured in independent experiments (Table 1.2). In the local growth experiment, a low-density population of *Tetrahymena* sp. was introduced evenly across the landscape and its density was measured locally at different times. Recorded density measurements were fitted to the (deterministic) logistic growth model, which gave the estimates for  $r$  and  $K$  (Table 1.2). In the local unimpeded movement experiment, the mean square displacement (Methods) of individual trajectories (Berg, 1993) was computed to estimate the diffusion coefficient  $D$  in density-independent conditions (Table 1.2 and Methods). The diffusion coefficient was estimated by looking at individuals at the front of the traveling wave, as these are the individuals responsible for the colonization of the empty landscape. During the dispersal experiment the diffusion coefficient of *Tetrahymena* sp. was also measured in the bulk of the wave, that is, where the population was at high density. Trajectories differed qualitatively between the bulk and the front of the wave and this reflects in a much smaller diffusion coefficient estimate where the population is at carrying capacity. In fact, in the bulk of the wave, the mean diffusion coefficient was  $D_{\text{bulk}} = 0.003 \pm 0.001 \text{ mm}^2/\text{s}$ , much smaller than at the wavefront. Such density-dependent effects, however, are not assumed to be operating at the low densities that determine the speed of the front. The growth and movement measurements were performed in the same linear landscape settings as in the dispersal experiment and therefore are assumed to accurately describe the dynamics at the front of the traveling wave in the dispersal events. Whereas the microscopic movement underlying the Fisher-Kolmogorov Eq. 1.17 is brownian motion (Gardiner, 2006, Méndez et al., 2010), investigation of the movement behavior of *Tetrahymena* sp. shows that cells' trajectories are consistent with a persistent random walk with an autocorrelation time  $\tau = 3.9 \pm 0.4 \text{ s}$ . The macroscopic equation corresponding to the persistent random walk is the reaction-telegraph equation (Méndez et al., 2010) in place of the Fisher-Kolmogorov Eq. 1.17. However, as the autocorrelation time for *Tetrahymena* sp. is much smaller than the growth rate  $r$  ( $\tau r \sim 10^{-4}$ ), Eq. 1.17 provides an excellent approximation to the reaction-telegraph equation. Following the principle of parsimony, the simpler Eq. 1.17 is adopted here.

The comparison of the predicted front speed  $v = 2\sqrt{rD}$  to the wavefront speed measured in the dispersal experiment,  $v_o$ , yields a compelling agreement. The observed speed in the dispersal experiment was  $v_o = 52.0 \pm 1.8 \text{ cm/day}$  (mean $\pm$ SE), which was compared to the predicted one  $v = 51.9 \pm 1.1 \text{ cm/day}$  (mean $\pm$ SE). The two velocities are compatible within one standard error. A  $t$ -test between the replicated observed speeds and bootstrap estimates of  $v = 2\sqrt{rD}$  (the quantity  $2\sqrt{rD}$  was computed for all possible combinations of the  $r$  and  $D$  values measured in the growth and movement experiments) gives a  $p$ -value of  $p = 0.96$  ( $t = 0.05$ ,  $\text{df} = 9$ ). Thus, the null hypothesis that the mean difference is 0 is not rejected at the 5 percent level and there is no indication that the two means are different. As the measurements of  $r$  and  $D$  were per-



**Figure 1.3:** Front propagation in the dispersal experiment and in the stochastic model. (A), Front position of the expanding population in six replicated dispersal events, colors identify replicates as in Fig. 1.2. The dark and light grey shadings are respectively the 95% and 99% confidence intervals computed by numerically integrating the generalized model equation, with initial conditions as at the second experimental time point, in 1020 iterations. The black curve is the mean front position in the stochastic integrations. (B), The increase in range variability between replicates in the dispersal experiment (blue diamonds) is well described by the stochastic model (red line). (C), Mean front speed for different choices of the reference density value at which the front position was estimated, error bars are smaller than symbols.

formed in independent experiments, at scales that were orders of magnitude smaller than in the dispersal events, the agreement between the two estimates of the front velocity is deemed remarkable.

Although the Fisher-Kolmogorov equation correctly predicts the mean speed of the experimentally observed invading wavefront, its deterministic formulation prevents it to reproduce the variability that is inherent to biological dispersal (Melbourne and Hastings, 2009). In particular, it cannot reproduce the fluctuations in range expansion between different replicates of the dispersal experiment (Fig. 1.3A). A generalization of the Fisher-Kolmogorov equation (Bonachela et al. (2012), Dornic et al. (2005), Hallatschek and Korolev (2009), Methods) accounting for demographic stochasticity is able to capture the observed variability of invasion fronts:

$$\frac{\partial \rho}{\partial t} = D \frac{\partial^2 \rho}{\partial x^2} + r \rho \left[ 1 - \frac{\rho}{K} \right] + \sigma \sqrt{\rho} \eta, \quad (1.18)$$

where  $\eta = \eta(x, t)$  is a gaussian, zero-mean white noise (i.e., the noise has correlations  $\langle \eta(x, t) \eta(x', t') \rangle = \delta(x - x') \delta(t - t')$ , where  $\delta$  is the Dirac's delta distribution) and  $\sigma > 0$  is constant. The strength of demographic stochasticity is embedded in an additional species' trait  $\sigma [T^{-1/2}]$ . The Itô's stochastic calculus (Gardiner, 2006) is appropriate in this case. Note, in fact, that the choice of the Stratonovich framework (Gardiner, 2006) would make no sense here, as the noise term would have a constant non-zero mean that would allow an extinct population to possibly escape the zero-density absorbing state. Such non-zero mean can be calculated via Novikov's theorem for spatially extended systems (Méndez et al., 2011, Novikov, 1965) and would be equal to  $\langle \sigma \sqrt{\rho} \eta \rangle = \sigma/2$  (in the Stratonovich framework, not adopted here). The square-root multiplicative noise term in Eq. 1.18 originates from the chemical Langevin equation (Gillespie, 2000) associated with the two microscopic reactions  $A \rightarrow 2A$  and  $A \rightarrow \emptyset$  (here,  $A$  represents an individual) and is thus interpreted as describing demographic stochasticity in the population (Bonachela et al., 2012). Eq. 1.18 needs extra-care in simulations (Dornic et al., 2005, Moro, 2004). In particular, standard stochastic integration schemes fail to preserve the positivity of  $\rho$ . The split-step method developed in Dornic et al. (2005) was adopted to numerically integrate Eq. 1.18 (Methods). This method allows to perform the integration with relatively large spatial and temporal steps maintaining numerical accuracy.

In this stochastic framework, the demographic parameters  $r$ ,  $K$  and  $\sigma$  were estimated from the local growth experiment with a maximum likelihood approach (Table 1.2 and Methods), while the diffusion coefficient  $D$  was left unchanged. These local independent estimates were then used to numerically integrate Eq. 1.18 with initial conditions as in the dispersal experiment. The front positions measured in the experiment are

**Table 1.1:** Best fit estimates ( $\pm$ SE) of the wavefront speed in six replicated dispersal events.

Replica	Speed (cm/day)
1	$54.6 \pm 1.9$
2	$51.7 \pm 2.8$
3	$48.0 \pm 1.5$
4	$58.0 \pm 4.0$
5	$53.4 \pm 1.8$
6	$46.3 \pm 1.0$

**Table 1.2:** Experimentally measured species' traits (mean $\pm$ SE). Demographic traits were estimated both in the framework of the deterministic logistic equation and in the framework of the stochastic logistic Eq. 1.23. Demographic stochasticity strongly affects the dynamics at low densities, thus a different value for the growth rate  $r$  is obtained in the stochastic model, compared to the deterministic one.

**Movement traits**

$$D = 0.17 \pm 0.01 \text{ mm}^2 \text{ s}^{-1}$$

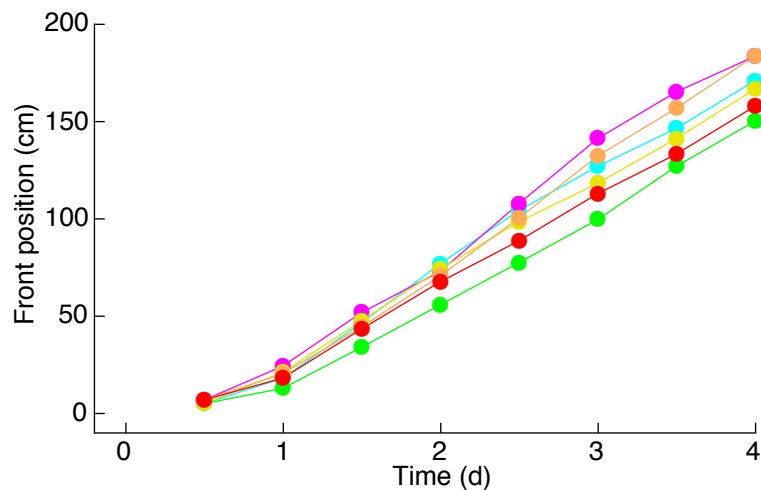
$$\tau = 3.9 \pm 0.4 \text{ s}$$

**Demographic traits**

Deterministic model	Stochastic model
$r = 4.9 \pm 0.5 \text{ day}^{-1}$	$r = 6.1 \pm 0.8 \text{ day}^{-1}$
$K = 901 \pm 130 \text{ cm}^{-1}$	$K = 903 \pm 135 \text{ cm}^{-1}$
	$\sigma = 25 \pm 5 \text{ day}^{-\frac{1}{2}}$

in accordance with simulations (Fig. 1.3A). In particular, most experimental data are within the 95% confidence interval for the simulated front position and the observed range variability is well captured by the stochastic model (Fig. 1.3B). Accordingly, the estimate for the front speed and its variability in the experiment are in good agreement with simulations.

The experimentally measured position of the front and the speed of front propagation were compared with the prediction of the stochastic model. The black curve in Fig. 1.3A is the mean position of the front over 1020 integrations of Eq. 1.18, with 170 iterations starting from each experimental density profile at the second measurement time point (Fig. 1.3A and Methods). The dark and light grey shadings in Fig. 1.3A represent respectively the 95% and 99% intervals for the front position. The increase in width for the front position is captured by the red curve in Fig. 1.3B that represents the 95% interval width for the front position at each time step. Simulations are in quantitative agreement with data (Fig. 1.3B). Examples of the front position in different simulations of the stochastic equation are shown in Fig. 1.4. The speed of



**Figure 1.4:** Front position in six integrations of the stochastic model Eq. 1.18 (compare with experimental positions in Fig. 1.3A). Each simulation was initialized from each of the measured density profiles at the second experimental time point.

the front in the stochastic simulations was estimated by fitting the front position at 8 equally-spaced time points in the time interval [1 – 4] day, over 1020 integrations of Eq. 1.18. The resulting mean speed of the front was 52.1 cm/day, the standard deviation was 4.2 cm/day. The mean speed in the dispersal experiment was 52.0 cm/day and the measured standard deviation was 4.3 cm/day.

## 1.4 Methods

### 1.4.1 Experiments

The species used in this study is *Tetrahymena* sp. (Fig. 1.1B), a freshwater ciliate, purchased at Carolina Biological Supply (Burlington, NC, USA). Individuals of *Tetrahymena* sp. have typical linear size (equivalent diameter) of 14  $\mu\text{m}$  (Giometto et al., 2013). Freshwater bacteria of the species *Serratia fonticola*, *Breviacillus brevis* and *Bacillus subtilis* were used as a food resource for ciliates, which were kept in a medium made of sterilized spring water and protozoan pellets (Carolina Biological Supply) at a density of 0.45  $\text{g L}^{-1}$ . The experimental units were kept under constant fluorescent light for the whole duration of the study, at a constant temperature of 22  $^{\circ}\text{C}$ . Experimental protocols are well established (Altermatt et al., 2014, Holyoak and Lawler, 2005) and the contribution of laboratory experiments on protists to the understanding of population and metapopulation dynamics proved noteworthy (Holyoak and Lawler, 2005).

Experiments were performed in linear landscapes (Fig. 1.1A) filled with a nutrient medium and bacteria of the three species above mentioned. The linear landscapes were 2 m long, 5 mm wide and 3 mm deep, respectively  $10^5$ , 350 and 200 times the size of *Tetrahymena* sp. (Giometto et al., 2013). Landscapes consisted of channels drilled on a plexiglass sheet, a second sheet was used as lid and a gasket was introduced to avoid water spillage (Fig. 1.1A). At one end of the landscapes, an opening was placed for the introduction of ciliates. The plexiglass sheets were sterilized with a 70% alcohol solution and gaskets were autoclaved at 120 °C before filling the landscape with medium. As plexiglass is transparent, the experimental units could be placed under the objective of a stereomicroscope, to record pictures (for counting of individuals) or videos (to track ciliates). Individuals were observed to distribute mainly at the bottom of the landscape, whose length was three orders of magnitude larger than its width ( $w$ ) and depth ( $d$ ) and two orders of magnitude larger than the typical length scale of the process ( $\sqrt{D/r} \simeq 5$  cm).

Three independent and complementing experiments were performed, specifically:

- i) A dispersal experiment was carried out to study the existence and the propagation of traveling invasion wavefronts in replicated dispersal events;
- ii) A growth experiment was run to obtain estimates of the demographic species' traits, that are  $r$  and  $K$  in the deterministic framework of Eq. 1.17 and  $r$ ,  $K$  and  $\sigma$  in the stochastic framework of Eq. 1.18;
- iii) A local movement experiment was performed to study the local unimpeded movement of *Tetrahymena* sp. over a short timescale (in a time window  $t \ll r^{-1}$ ), in order to estimate the diffusion coefficient  $D$  for the study species, independently from the dispersal and growth experiments.

### **Dispersal experiment**

Six replicated dispersal events were performed in the linear landscapes. After filling the landscapes with medium and bacteria, a small ensemble of *Tetrahymena* sp. was introduced at the origin. Subsequently, the density of *Tetrahymena* sp. was measured at 1 cm intervals, five times in the first 48 h and twice in the last 48 h. The whole experiment lasted for about 20 generations of the study species.

### **Local growth experiment**

Five replicated growth measurements were performed in the linear landscapes, in order to measure the demographic species' traits, in the same environmental conditions as in the dispersal experiment, but independently from it. A low density culture

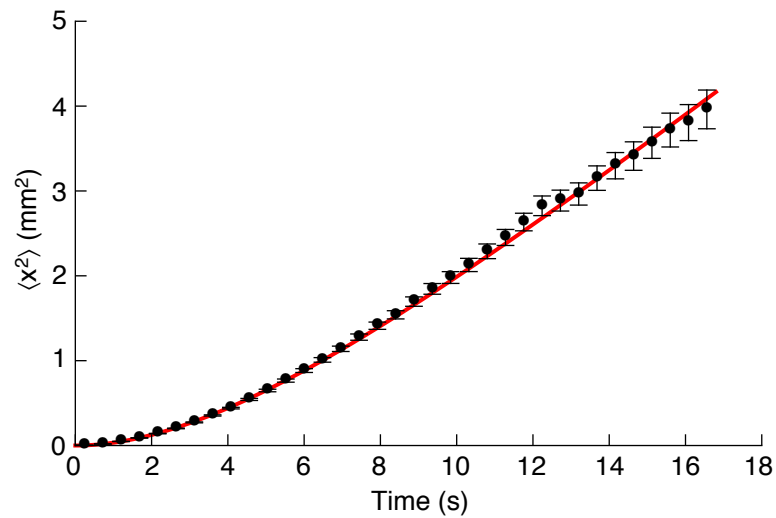
of *Tetrahymena* sp. was introduced in the whole landscape and its density was measured by taking pictures and counting individuals, covering a region of 7 cm along the landscape. Density measurements were performed at several time points for each of the five replicates, in a time window of 3 d.

### Local movement experiment

Four additional, replicated dispersal events were performed, initialized in the same way as in the dispersal experiment, in order to measure the diffusion coefficient of *Tetrahymena* sp. The diffusion coefficient  $D$  is the proportionality constant that links the mean square displacement of organisms' trajectories to time (Berg, 1993, Méndez et al., 2010). Macroscopically, it relates the local flux to the density of individuals. To estimate the diffusion coefficient several videos of individuals moving at the front of the traveling wave (at low density) were recorded, the corresponding trajectories (Berg, 1993, Sbalzarini and Koumoutsakos, 2005) were reconstructed and their mean square displacement  $\langle x^2(t) \rangle = \langle [x(t) - x(0)]^2 \rangle$  was computed. Videos of *Tetrahymena* sp. at the front of the traveling wave were recorded in four replicated dispersal events, at various times over 4 d. The area covered in each video was of 24 mm in the direction of the landscape and 5 mm orthogonal to it. Each video lasted for 12 min. For each recorded video, individuals' spatial coordinates in each frame were extracted and the *MOSAIC* plugin for the software *ImageJ* was employed to reconstruct trajectories (Sbalzarini and Koumoutsakos, 2005). The goodness of the tracking was checked on several trajectories by direct comparison with the videos. Examples of reconstructed trajectories can be seen in Fig. 1.1C.

For each video, the square displacement of each trajectory in the direction parallel to the landscape was computed at all time points and then averaged across trajectories. Precisely, for each trajectory  $i$  the quantity  $x_i^2(t) = [X_i(t) - X_i(0)]^2$  was computed, where  $X_i(t)$  is the 1-dimensional coordinate of organism  $i$  at time  $t$  in the direction parallel to the landscape and  $X_i(0)$  is its initial position. The mean square displacement in a video was then computed as the mean of  $x_i^2(t)$  across all trajectories, that is,  $\langle x^2(t) \rangle = \frac{1}{N} \sum_i x_i^2(t)$  (where  $N$  is the total number of trajectories). A typical measurement of  $\langle x^2(t) \rangle$  is shown in Fig. 1.5. As shown in the figure, there exists an initial auto-correlated phase. To estimate the diffusion coefficient from the mean square displacement, the measured  $\langle x^2(t) \rangle$  were fitted to the function  $\langle x^2(t) \rangle = 2Dt - 2D\tau [1 - e^{-t/\tau}]$  with the two parameters  $D$  (diffusion coefficient) and  $\tau$  (correlation time). The total number of recorded videos was 28, that is, 7 for each replicate.





**Figure 1.5:** Mean square displacement  $\langle x^2 \rangle$  of individual trajectories versus time, for a representative video. The red curve is the best fit of the data to the equation  $\langle x^2(t) \rangle = 2Dt - 2D\tau [1 - e^{-t/\tau}]$ . After an initial auto-correlated phase, the mean square displacement increases linearly with time. Error bars are  $\pm \text{SE}$ .

## 1.4.2 Stochastic model and analysis

### Spatial discretization

Eq. 1.18 is interpreted as the continuum limit of a set of coupled Itô equations resulting from a discretization of space (Dornic et al., 2005). Let  $\Delta x$  be the step of spatial discretization on a 1D lattice. The discretization reads:

$$\begin{aligned} \frac{d\rho_i}{dt}(t) = & \frac{D}{(\Delta x)^2} [\rho_{i+1}(t) + \rho_{i-1}(t) - 2\rho_i(t)] + \\ & + r\rho_i(t) \left(1 - \frac{\rho_i(t)}{K}\right) + \frac{\sigma}{\sqrt{\Delta x}} \sqrt{\rho_i(t)} \eta(t) \end{aligned} \quad (1.19)$$

where  $i$  identifies the lattice site and the term  $\sqrt{\Delta x}$  ensures proper normalization in the continuum limit (Doering et al., 2005).

### 1.4.3 The split-step integration scheme

This Box contains known results on the numerical integration of Langevin equations with multiplicative noise and is included here to aid the comprehension of the chapter. Please note that the results presented within this Box are not original and can be found in Dornic et al. (2005).

The integration of the stochastic Eq. 1.18 was performed with the split-step integration scheme developed in Dornic et al. (2005) via the spatial discretization Eq. 1.19. The split-step method consists in separating the terms in Eq. 1.19 and integrating the stochastic part by sampling the solution of the associated Fokker-Plank equation instead of generating Gaussian random numbers, which may not preserve the positivity of the solution. The method exploits the fact that the conditional transition probability density function for the equation  $d\rho/dt = \alpha + \beta\rho + \gamma\sqrt{\rho}\eta$  in the time interval  $[t, t + dt]$  is known (Dornic et al., 2005):

$$P(\rho, t + \Delta t | \rho_t, t) = \lambda e^{-\lambda(\rho_t e^{\beta\Delta t} + \rho)} \left[ \frac{\rho}{\rho_t e^{\beta\Delta t}} \right]^{\mu/2} I_\mu \left( 2\lambda \sqrt{\rho_t \rho e^{\beta\Delta t}} \right), \quad (1.20)$$

where  $I_\mu$  is the Bessel function of order  $\mu$ ,  $\lambda = 2\beta / [\gamma^2 (e^{\beta t} - 1)]$  and  $\mu = -1 + 2\alpha/\gamma^2$ . The integration is performed as follows. For each lattice site  $i$ , let  $\alpha = D/(\Delta x)^2 [\rho_{i+1}(t) + \rho_{i-1}(t)]$ ,  $\beta = r - 2D/(\Delta x)^2$  and  $\gamma = \sigma/\sqrt{\Delta x}$ . The stochastic value  $\rho^*$  is obtained by sampling the conditional transition probability density function Eq. 1.20 and is used as the initial condition for the deterministic remaining part of Eq. 1.19  $d\rho_i/dt(t) = -r/K\rho_i^2(t)$ , which is integrated as  $\rho_i(t + \Delta t) = \rho^*/(1 + \rho^*r/K\Delta t)$ . Following Dornic et al. (2005), the sampling of Eq. 1.20 is performed by Taylor expanding  $I_\mu$  in Eq. 1.20 and rewriting it as:

$$P(\rho, t + \Delta t | \rho_t, t) = \sum_{n=0}^{\infty} \frac{(\lambda \rho_t e^{\beta\Delta t})^n e^{-\lambda \rho_t e^{\beta\Delta t}}}{n!} \frac{\lambda e^{-\lambda \rho} (\lambda \rho)^{n+\mu}}{\Gamma(n + \mu + 1)} \quad (1.21)$$

and thus sampling  $\rho^*$  from the mixture:

$$\rho^* = \text{Gamma} \left[ 1 + \mu + \text{Poisson} \left( \lambda \rho_t e^{\beta\Delta t} \right) \right] / \lambda, \quad (1.22)$$

as proposed in Dornic et al. (2005), where further details are provided. Except for the stability requirement due to the discretized Laplacian, such split-step method was claimed to display no limitation on  $\Delta t$  for the numerical stability of the integration scheme.

The spatial discretization Eq. 1.19 allows to compare the noise term in Eq. 1.18 to the local noise acting at a lattice site of size  $\Delta x$ . In particular, it allows to estimate the noise strength  $\sigma$  by parameter identification on the growth experiment data, where the cell density was measured in a single site  $i$  of size  $l = 7$  cm (a length comparable to the step size adopted in the numerical integration, that is  $\Delta x = 5$  cm). The equation governing the density of the single site  $i$  in the growth experiment is thus:

$$\frac{d\rho}{dt}(t) = r\rho(t) \left(1 - \frac{\rho(t)}{K}\right) + \frac{\sigma}{\sqrt{l}} \sqrt{\rho(t)} \eta(t), \quad (1.23)$$

where the diffusion term is neglected because of the well-mixed settings and the  $i$  subscript is dropped as there is only one site. The maximum likelihood approach described in the following section allows to estimate  $r$ ,  $K$  and  $\sigma$  from the experimental growth data.

### Parametric inference

One key issue in this chapter is to estimate the stochastic model parameters from the independent experiments. To obtain estimates for the demographic parameters  $r$ ,  $K$  and  $\sigma$ , data from the growth experiment were fitted to Eq. 1.23. The likelihood function for Eq. 1.23 can be written as:

$$L(\theta) = \prod_{j=2}^n P[\rho(t_j), t_j | \rho(t_{j-1}), t_{j-1}; \theta], \quad (1.24)$$

where  $n$  is the total number of observation in the growth time series,  $\theta = (r, K, \sigma)$  is the vector of demographic parameters and  $P(\rho, t | \rho_0, t_0; \theta)$  is the transitional probability density of having a density of individuals  $\rho$  at time  $t$ , given that the density at time  $t_0$  was  $\rho_0$  (for a given  $\theta$ ). The transitional probability density  $P(\rho, t | \rho_0, t_0; \theta)$  satisfies the Fokker-Planck equation associated to Eq. 1.23, that is:

$$\begin{aligned} \frac{\partial}{\partial t} P(\rho, t | \rho_0, t_0; \theta) = & - \frac{\partial}{\partial \rho} \left[ r\rho \left(1 - \frac{\rho}{K}\right) P(\rho, t | \rho_0, t_0; \theta) \right] + \\ & + \frac{\sigma^2}{2l} \frac{\partial^2}{\partial \rho^2} [\rho P(\rho, t | \rho_0, t_0; \theta)]. \end{aligned} \quad (1.25)$$

Maximization of the likelihood is equivalent to the minimization of the negative log-likelihood  $-\log L(\theta)$ , which is computationally less expensive. To compute the likelihood for a fixed set of parameters  $\theta$  one has to solve numerically the Fokker-Planck Eq. 1.25 for all observed transitions, with the  $[t_j, \rho(t_j)]$  as measured in the experiment. It is computationally more accurate to solve Eq. 1.25 in terms of the cumulative distribution function (CDF), as its initial condition in the transition  $[t_{j-1}, \rho(t_{j-1})] \rightarrow [t_j, \rho(t_j)]$  can be expressed as a step function instead of a delta function, the first one being more

accurate in the numerical approximation (Hurn et al., 2007). The transitional probability densities (solutions of Eq. 1.25) can then be recovered by numerical differentiation. The numerical integration was performed adopting the implicit Crank-Nicolson scheme (Hurn et al., 2007), the minimization was performed with the software *Mathworks Matlab*, adopting the active-set algorithm in a large domain. Different initial conditions for the parameters led to the same estimate for the minimum, which is thus interpreted as the global minimum of the negative log-likelihood function, that is, the global maximum for the likelihood function. The set of demographic parameters that maximized the likelihood function is reported in Table 1.2.

### 1.5 Discussion

The results of this chapter suggest that measuring and suitably interpreting local processes allows to accurately predict the main features of biological invasions. The deterministic Fisher-Kolmogorov equation is shown to correctly predict the mean speed of invasion, but cannot capture the observed variability. Instead, characterizing the inherent stochasticity of the biological processes involved allows to predict both the mean and the intrinsic variability of range expansions, which is of interest for practical purposes, such as the delineation of worst-case scenarios for the spread of invasive species and the projection of confidence intervals for the position of the front.

The use of the generalized Fisher-Kolmogorov equation proposed here has been insofar limited to theoretical investigations ranging from non-equilibrium phase transitions in statistical physics (Dornic et al., 2005) to the propagation of invasion fronts (Hallatschek and Korolev, 2009) and patchy distribution patterns (Bonachela et al., 2012) in theoretical ecological studies. Here, it was shown that such equation can be applied to experimental investigations of biological dispersal, yielding quantitative predictions for the internal fluctuations of the process via a proper parametric inference of the demographic traits.

Our phenomenological approach allows to make predictions on the spread of organisms without the need to introduce all details on the movement behavior, biology or any other information. Such details are synthesized in three parameters describing the density-independent yet stochastic behavior of individuals riding the invasion wave. The parsimony of the model allows generalization to organisms with different biology (e.g., growth rates and diffusion coefficients are available for several species in the literature, Grosholz (1996)) and supports the view that the protocol adopted here may provide a general predictive framework for biological invasions in natural

environments.

It is likely that the internal fluctuations investigated here are only one of many components affecting invasion processes in natural environments. However, demographic fluctuations have the unique feature of being proportional to the square root of the population density, which makes them dominate over environmental fluctuations (typically proportional to the population density, Méndez et al. (2011)) at low densities and thus at the front of the invasion. Therefore, demographic stochasticity is argued to be among the main drivers of fluctuations in the position of propagating fronts. The interplay between demographic and environmental fluctuations and their combined effect on the propagation of biological invasion fronts is an exciting avenue for research, which is explored both experimentally and theoretically in Chapter 3, where the demographic noise term introduced in the Fisher-Kolmogorov equation will be shown to lead to a slowing-down of the front propagation when environmental stochasticity is superimposed on the system.

In conclusion, at least in simple ecological settings, predictability remains notwithstanding biological fluctuations. This investigation supports the use of deterministic models to describe the mean dynamics of invasion fronts and provides an experimental link between the processes at the local and short-time scale (movement and growth) and the invasion process at the large scale, in a topic where experimentation is the exception rather than the rule. Information on the stochasticity acting at the mesoscopic scale allows to estimate fluctuations at the macroscopic scale via the stochastic treatment devised. The results of this chapter might have implications for the dynamics of phenomena other than species' invasions, such as morphogenesis (Murray, 2004, Turing, 1952), tumor growth (Fort and Solé, 2013, Méndez et al., 2010, Murray, 2004) and the spreading of epidemics (Bertuzzo et al., 2010, Murray, 2004), which have been traditionally modeled with reaction-diffusion equations and where demographic stochasticity is expected to occur.



## 2 Phototaxis in phytoplankton

### Abstract

Phototaxis, the process through which motile organisms direct their swimming towards or away from light, is implicated in key ecological phenomena (including algal blooms and diel vertical migration) that shape the distribution, diversity and productivity of phytoplankton and thus energy transfer to higher trophic levels in aquatic ecosystems. Phototaxis also finds important applications in biofuel reactors and micro-bio-propellers and is argued to serve as a benchmark for the study of biological invasions in heterogeneous environments owing to the ease of generating stochastic light fields. Despite its ecological and technological relevance, an experimentally tested, general theoretical model of phototaxis seems unavailable to date. Here, accurate measurements of the behavior of the alga *Euglena gracilis* exposed to controlled laboratory light fields are performed. Analysis of *E. gracilis*' phototactic accumulation dynamics over a broad range of light intensities proves that the classic Keller-Segel mathematical framework for taxis provides an accurate description of both positive and negative phototaxis when phototactic sensitivity is modeled by a generalized 'receptor law', a specific nonlinear response function to light intensity that drives algae towards beneficial light conditions and away from harmful ones. The proposed phototactic model captures the temporal dynamics of both cells' accumulation towards light sources and their dispersion upon light cessation. The model could thus be of use in integrating models of vertical phytoplankton migrations in marine and freshwater ecosystems, and in designing bioreactors and other technological applications.

### 2.1 Introduction

Microorganisms possess a variety of sensory systems to acquire information about their environment (Hazelbauer et al., 1993), including the availability of resources, the presence of predators, and the local light conditions (Stocker, 2012). For any sensory system, the system's response function determines the organism's capability to process the available information and turn it into a behavioral response. Such response function is shaped by the natural environment and its fluctuations (Celani and Vergassola, 2010, Kussell and Leibler, 2005, Laughlin, 1981) and affects the search strategy (be it mate search, food search, etc., Mesibov et al. (1973), Shoval et al. (2010)) and the swimming behavior of microorganisms (Lazova et al., 2011). Gradient-sensing is particularly important in marine and freshwater ecosystems, where the distribution of resources is highly heterogeneous (Azam, 1998, Levin, 1994) and the ability to move towards resource hotspots can provide a strong selective advantage to motile organisms over non-motile ones (Celani and Vergassola, 2010, Stocker, 2012). Spatio-temporal patterns of light underwater contribute considerably to the heterogeneity of the aquatic environment. Because light is a major carrier of energy and information in the water column (Ragni and Ribera D'Alcalà, 2004), phototaxis is a widespread example of directed gradient-driven locomotion (Bhaya, 2004, Jékely et al., 2008), found in many species of phytoplankton and zooplankton. Phototaxis strongly affects the ecology of aquatic ecosystems, contributing to diel vertical migration of phytoplankton, one of the most dramatic migratory phenomena on Earth and the largest in terms of biomass (Hays, 2003). Diel vertical migration is crucial for the survival and proliferation of plankton (Jékely et al., 2008, Kingston, 1999, Ringelberg and Flik, 1994), allows them to escape from predation by filter-feeding organisms and may affect the structuring of algal blooms (Smayda, 1997). Because phytoplankton are responsible for half of the global photosynthetic activity (Behrenfeld et al., 2006, Field et al., 1998) and are the basis of marine and freshwater food webs (Chassot et al., 2010), their behavior and productivity have strong implications for ocean biogeochemistry, carbon cycling, and trophic dynamics (Boyce et al., 2010, Falkowski, 1998).

The quantitative understanding and the associated development of mathematical models for the directed movement of microorganisms has been largely limited to chemotaxis, while other forms of taxis have received considerably less attention despite their ecological importance. For chemotaxis, quantitative experiments have led to a comprehensive characterization of the motile response of bacteria to chemical gradients (Adler et al., 1973, Barbara and Mitchell, 2003) and this knowledge has been distilled into detailed mathematical models (Tindall et al., 2008). Continuum approaches such as the Keller-Segel model (Keller and Segel, 1970a, 1971), and its generalizations (Tindall et al., 2008), have been used extensively to describe the behavior



of chemotactic bacterial populations in laboratory experiments. However, while a limited number of models for phototaxis exist (Burkart and Häder, 1980, Torney and Neufeld, 2008, Vincent and Hill, 1996, Williams and Bees, 2011a), an assessment of the phototactic response function is lacking. Existing models rely on untested working hypotheses concerning the cell response to light originating from the scarcity of experimental work linking controlled light conditions to measured organism responses.

Here, quantitative experimental observations of the phototactic response of the flagellate alga *Euglena gracilis* to controlled light gradients are presented. *E. gracilis* is a common freshwater phytoplankton species that swims via an anterior flagellum and uses a paraflagellar body and red stigma (a red eyespot, Jékely (2009)) to respond to light gradients. *E. gracilis* has been used extensively as a model organism in both the ecological (Carrara et al., 2012, Giometto et al., 2013) and the ecophysiological literature (Vallee and Falchuk, 1993, Wolken, 1961) and has been used as a candidate species for technological applications such as photo-bioreactors (Ooka et al., 2014) and micro-propellers (Itoh, 2004, Itoh and Tamura, 2008). The experimental results are used to identify a general mathematical model for phototaxis. It is found that a Keller-Segel type model (Keller and Segel, 1970a, 1971) accurately describes cell accumulation patterns at all light intensities tested and that the light sensitivity of *E. gracilis* is described by a generalized receptor law (Lapidus and Schiller, 1976, Tindall et al., 2008), a non-linear function of light intensity that displays a maximum at the light intensity at which most cells accumulate and is negative at higher intensities, where negative phototaxis occurs.

This chapter is organized as follows. First, a theoretical background on the Keller-Segel model for chemotaxis is provided. Second, a set of experiments aimed at measuring the stationary distribution patterns in the presence of light gradients and the temporal dynamics of accumulation and relaxation of density peaks is described. Third, a mathematical model for phototaxis is proposed and is shown to describe the experimental results, allowing the inference of the functional form of the phototactic response function from the data. Fourth, experiments at the individual cell level are described and a mathematical model for phototaxis at the individual level is proposed. Finally, the implications for ecological processes and technological applications are discussed.

## 2.2 Theoretical background

This Box contains known results on the mathematical modeling of chemotaxis in the Keller-Segel framework (Keller and Segel, 1970a, 1971, Tindall et al., 2008). Please note that the results presented within this Box are not original. The model equations presented here are used in the chapter, adapted to the case of phototaxis.

The Keller-Segel continuum model (Keller and Segel, 1970a,b, 1971) and its generalizations (Tindall et al., 2008) have been used extensively to model the behavior of chemotactic bacterial populations. The simplest form of the Keller-Segel model (Tindall et al., 2008) considers a population of organisms of density  $\rho(x, t)$  in the presence of an attractant that influences cell motion through chemical interaction with the organism's receptor (Keller and Segel, 1971). Let  $\chi$  be the chemotactic coefficient that characterizes the response of the population to different concentrations of the attractant. To start with,  $\chi$  is assumed to be independent of  $c$  (Rosen, 1976). In one dimension, the flux of  $\rho(x, t)$  reads:

$$J(x, t) = -D \frac{\partial \rho}{\partial x}(x, t) + \chi \frac{\partial c}{\partial x}(x, t) \rho(x, t), \quad (2.1)$$

where  $D$  is the diffusion coefficient typical of the study species. The equation governing the dynamics of the density of organisms  $\rho(x, t)$ , neglecting the reproduction of organisms, reads:

$$\frac{\partial \rho}{\partial t}(x, t) = -\frac{\partial J}{\partial x}(x, t) = \frac{\partial}{\partial x} \left[ D \frac{\partial \rho}{\partial x}(x, t) - \chi \frac{\partial c}{\partial x}(x, t) \rho(x, t) \right]. \quad (2.2)$$

The quantity  $\chi \partial c / \partial x$  is known in the literature as the 'chemotactic drift velocity'.

If the attractant diffuses in the medium with diffusion coefficient  $D_c$  and is consumed by the organisms with rate  $g(\rho, c)$ , the equation for  $c(x, t)$  reads:

$$\frac{\partial c}{\partial t}(x, t) = D_c \frac{\partial^2 c}{\partial x^2}(x, t) - g[\rho(x, t), c(x, t)]. \quad (2.3)$$

The coupled Eqs. 2.2 and 2.3 constitute the canonical Keller-Segel model (Tindall et al., 2008). Such model can be readily generalized to account for the reproduction of organisms, for example by adding a logistic growth term  $r\rho(1 - \rho/K)$  on the right hand side of Eq. 2.2. One important generalization of Eq. 2.2 is the introduction of a non-constant chemotactic coefficient  $\chi(c)$  dependent on the attractant's

concentration. In such generalization, the flux  $J$  is equal to  $J = -D \frac{\partial \rho}{\partial x} + \chi(c) \frac{\partial c}{\partial x} \rho$ , and thus Eq. 2.2 is modified in:

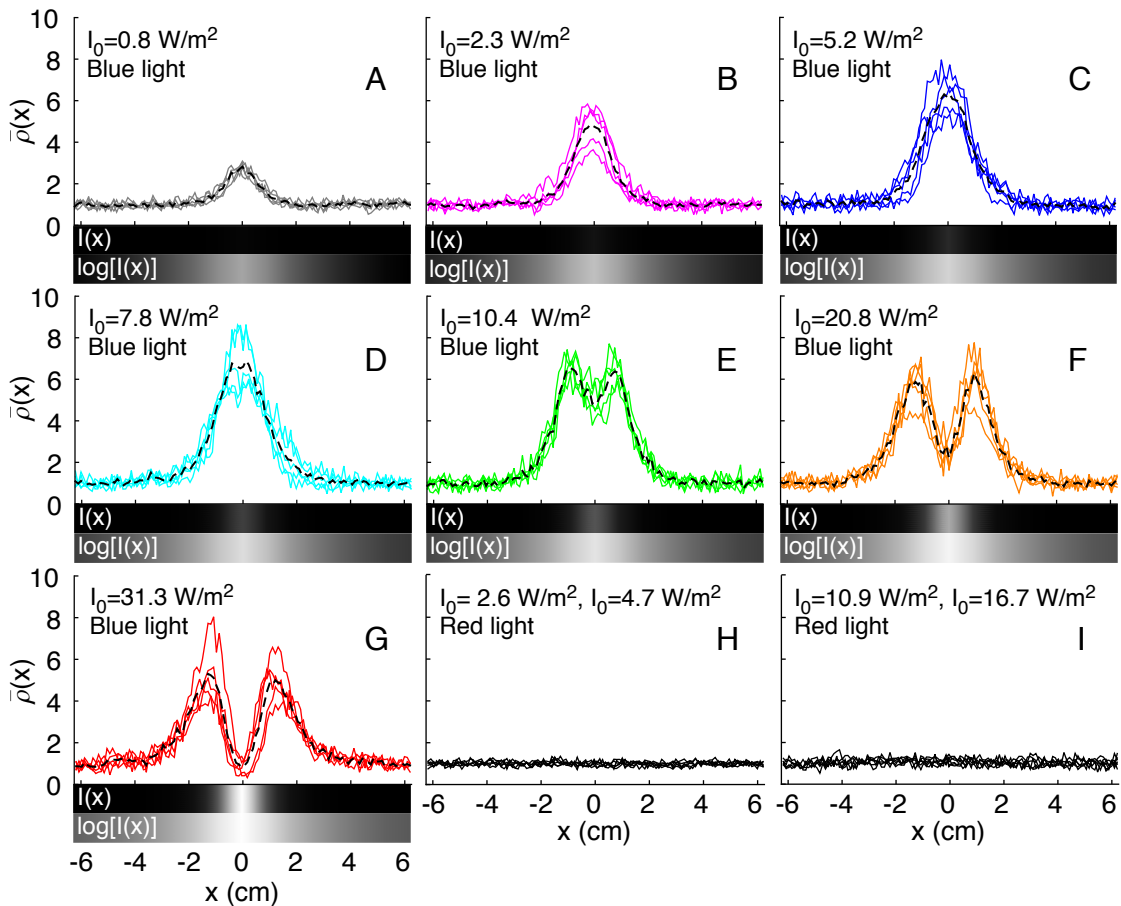
$$\frac{\partial \rho}{\partial t}(x, t) = -\frac{\partial J}{\partial x}(x, t) = \frac{\partial}{\partial x} \left[ D \frac{\partial \rho}{\partial x}(x, t) - \chi(c) \frac{\partial c}{\partial x}(x, t) \rho(x, t) \right]. \quad (2.4)$$

The generalization to a non-constant chemotactic coefficient  $\chi(c)$  allows to account for possible saturating effect at high concentrations of the attractant or, in general, to provide better fits to experimental data. Several forms for the chemotactic coefficient  $\chi$  have been proposed in the literature and are reviewed in Tindall et al. (2008) (Appendix A). Typical functional forms for  $\chi(c)$  are  $\chi(c) = \chi$  constant (Rosen, 1976),  $\chi(c) = \chi/c$  (Keller and Segel, 1970b),  $\chi(c) = \chi/(K + c)^2$  (Lapidus and Schiller, 1976).

## 2.3 Results

Two sets of experiments were performed, one at the population scale and one at the individual cell level, to track the response of *E. gracilis* to imposed light conditions. Experiments were conducted in linear channels (5 mm wide  $\times$  3 mm high  $\times$  2 m long) filled with cells ( $2100 \pm 200$  cells  $\text{mL}^{-1}$ ) suspended in nutrient medium (Methods). Light conditions were controlled by Light-Emitting Diodes (LED) illuminating the channels from below and operated via Arduino Uno boards (Methods). In the population-scale experiments, cell distributions were measured in response to localized light sources of different intensity and wavelength  $\lambda$  in the blue ( $\lambda = 469$  nm) and red ( $\lambda = 627$  nm) regions of the visible spectrum. The light intensity profile  $I(x) = I_0 i(x)$  ( $i(0) = 1$ , units are retained in  $I_0$ ) in the linear channels (Methods) was measured and LEDs were programmed to produce the following peak intensities within the channel, at  $x = 0$  cm (above the LED):  $I_0 = 0.8, 2.3, 5.2, 7.8, 10.4, 20.8, 31.3$   $\text{W m}^{-2}$  for  $\lambda = 469$  nm and  $I_0 = 2.6, 4.7, 10.9, 16.7$   $\text{W m}^{-2}$  for  $\lambda = 627$  nm. The light profile  $i(x)$  was determined by the experimental setting geometry and was invariant for all values of  $I_0$ .

Stationary *E. gracilis* accumulation patterns in blue light are shown in Fig. 2.1A-G. Fig. 2.1 shows that by increasing the peak light intensity  $I_0$  from  $I_0 = 0.8$   $\text{W m}^{-2}$  to  $I_0 = 5.2$   $\text{W m}^{-2}$ , cell density peaks increase in magnitude (shown are the density profiles normalized by the value at the boundary) and are aligned with the position of the light source ( $x = 0$  cm). Then, for larger values of  $I_0$ , cell density peaks are approximately constant in magnitude, but shift to the left and right of the source.



**Figure 2.1:** Accumulation patterns of *E. gracilis* in blue and red light of different intensities. Shown are normalized stationary cell density profiles  $\bar{\rho}(x)$  around a light source located at  $x = 0$  cm for various peak intensities  $I_0$  in the blue (A-G;  $\lambda = 469$  nm) and red (H-I;  $\lambda = 627$  nm) regions of the visible spectrum. The colored curves in panels A-G are the experimental cell density distributions (five replicates for each value of  $I_0$ ) and the dashed black lines denote the mean. The grayscale plots below the panels A-G show the imposed blue light intensity profiles, where the gray level scales linearly (upper panels) or logarithmically (lower panels) with the intensity  $I$ ; white corresponds to  $I = 31 \text{ W m}^{-2}$  and black to  $I = 0.001 \text{ W m}^{-2}$ . Positive phototaxis towards blue light is observed up to  $I \approx I_m = 5.5 \text{ W m}^{-2}$ . Negative phototaxis (the directed movement towards regions with lower values of  $I$ ) is observed for blue light intensity higher than  $I_m$ . No phototactic behavior is discernible with red light (H-I) (three replicates for each value of  $I_0$ ).

The largest value of blue light intensity for which positive phototaxis (the directed movement towards regions with larger values of  $I$ ) was observed was  $I \simeq I_m = 5.5 \text{ W m}^{-2}$  (Fig. 2.1A-G;  $\lambda = 496 \text{ nm}$ ). In contrast, light intensities higher than  $I_m$  elicited consistent negative phototaxis, indicating a biphasic response to light (Fig. 2.1G). Such biphasic responses are common in phototaxis, because they allow cells to increase their photosynthetic activity by migrating towards light while preventing damage to the photosynthetic apparatus and cell pigments at excessive light intensities (Häder and Lebert, 1998, Lebert et al., 1999). The experiments showed clearly no response to red light (Fig. 2.1H-I), in line with the reported weak absorption of *E. gracilis*' eyespot at these wavelengths (Strother and Wolken, 1960). Red light experiments thus serve as a control that allows to exclude that the observed cell accumulations towards blue light were due to factors other than phototaxis.

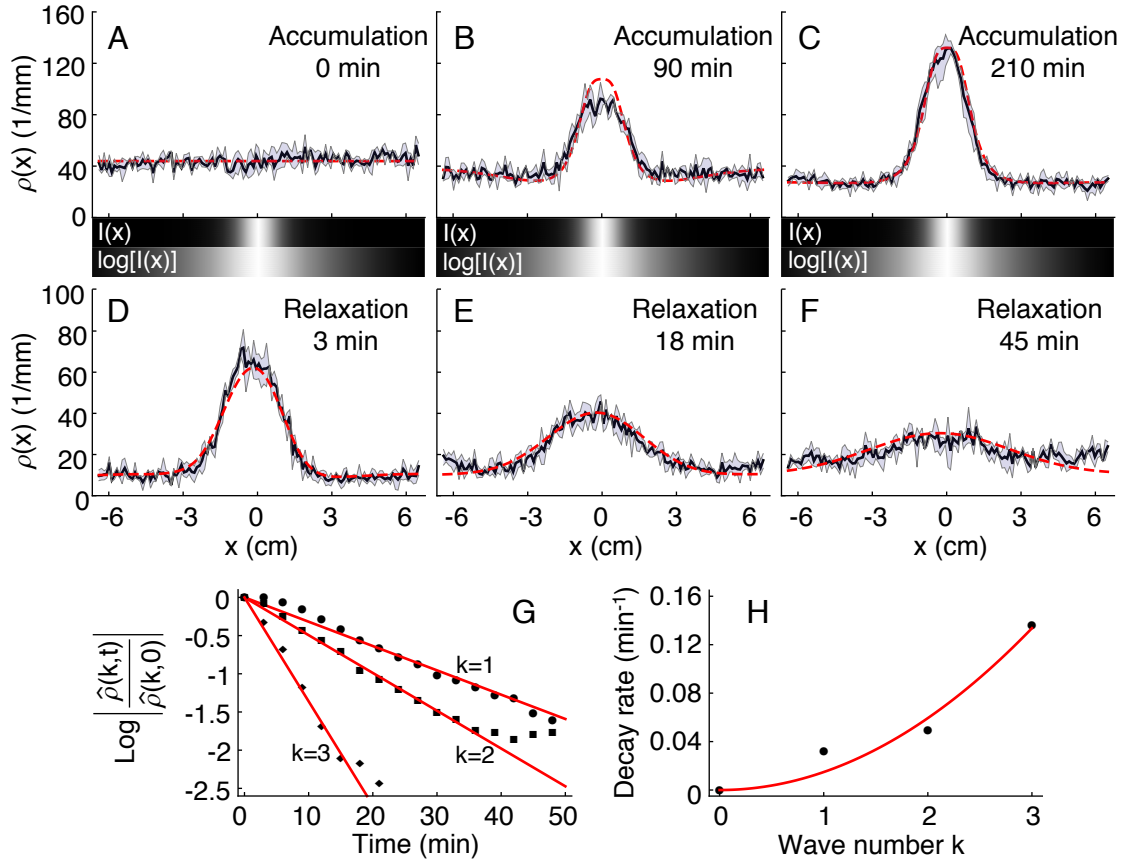
The formation of cell density peaks was measured in time (Fig. 2.2A-C), starting from a homogeneous suspension of cells (Fig. 2.2A), in the presence of a light source of peak intensity  $I_0 = 5.2 \text{ W m}^{-2}$  at  $x = 0 \text{ cm}$ . Then, the relaxation of the stationary density peaks after the removal of light was measured (Fig. 2.2D-F). This allowed to measure robustly the cell diffusion coefficient,  $D$ , by fitting the decay rate of the spectral log-amplitudes  $\log|\hat{\rho}(k, t)|$  to the square of the wave number, resulting in the estimate  $D = 0.13 \pm 0.04 \text{ mm}^2 \text{ s}^{-1}$  (the standard error represents the variability across the first three Discrete Fourier Transform modes).

The experimental results allowed to derive a model of phototaxis in *E. gracilis*. A Keller-Segel framework was adopted and consists of an advection-diffusion equation for the cell density  $\rho(x, t)$  (Tindall et al., 2008) (neglecting cell division owing to the relatively short duration of the experiments):

$$\frac{\partial \rho}{\partial t}(x, t) = \frac{\partial}{\partial x} \left[ D \frac{\partial \rho}{\partial x}(x, t) - \frac{d\phi}{dx}[I(x)]\rho(x, t) \right], \quad (2.5)$$

where  $D$  is the diffusion coefficient of the cells due to the random component of their motility and  $v_p = d\phi/dx$  is the drift velocity or 'phototactic velocity' of the population in the direction of the light gradient. The phototactic velocity was written as the derivative of a phototactic potential,  $\phi$ , which is solely a function of the light intensity  $I(x)$ . Such reformulation of the Keller-Segel model allows to express the stationary density distribution as a function of  $I(x)$ . In fact, the steady state accumulation of cells that satisfies Eq. 2.5, computed over the spatial extent of the imaging windows ( $-L \leq x \leq L$ ;  $L = 6.25 \text{ cm}$ ), is:

$$\bar{\rho}(x) = \frac{\rho(x)}{\rho(-L)} = \exp \left[ \frac{\phi[I(x)]}{D} \right], \quad (2.6)$$



**Figure 2.2:** Temporal dynamics of accumulation around a light source at  $x = 0$  cm (A-C) and relaxation of cell density peaks upon removal of light (D-H). (A-F) Experimental cell density profiles at different times. The shaded gray area is delimited by the maximum and minimum cell densities of three replicate experiments and the black line denotes the mean. The red dashed line shows the theoretical prediction from the phototaxis model, Eq. 2.5, using the experimentally determined  $\phi(I)$  and  $I(x)$  (Fig. 2.3A-B) and  $D$  (Table 2.2) determined experimentally from the relaxation of density peaks (D-H). Density profiles are renormalized to display the same mean abundance. The grayscale plots below panels A-C show the light intensity profile imposed during the accumulation; the gray level scales linearly (upper panels) or logarithmically (lower panels) with the intensity  $I$ , with white corresponding to  $I = 5.2 \text{ W m}^{-2}$  and black to  $I = 0.001 \text{ W m}^{-2}$ . The temporal decay of Fourier modes (G) during the relaxation of density peaks (D-F) is exponential ( $\log |\hat{\rho}(k, t) / \hat{\rho}(k, 0)| = -Dk^2 t$ , data in black and linear fit in red) and the decay rate is a quadratic function of the wave number  $k$  (H, data in black and parabolic fit in red), proving the diffusive behavior in the absence of light gradients.

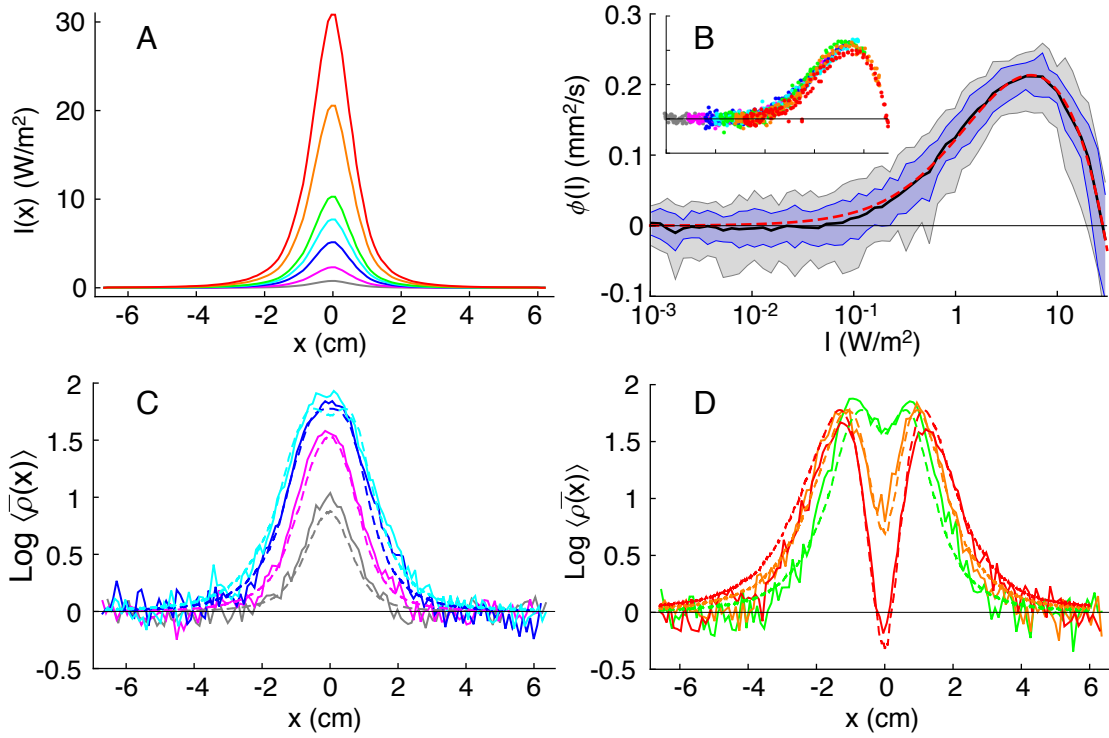
where  $\bar{\rho}(x)$  is a normalized cell density appropriate for comparison with experimental observations. Note that, in general, the exponent should be  $\phi[I(x)] - \phi[I(-L)]$ , but because  $\phi$  is defined only up to an additive constant  $\phi[I(-L)]$  was set to 0. Thus,  $\phi$  is set to zero for  $I = 0$ . Experimental cell density profiles,  $\bar{\rho}(x) = \rho(x)/\rho(-L)$  in Fig. 2.1, are normalized by their value at the edge of the imaging window,  $\rho(-L)$ , to account for small variations ( $< 15\%$ ) in the total number of cells across replicates.

The stationary cell density distributions under blue light (Fig. 2.1A-G) together with the measured light intensity profiles (Fig. 2.3A) were used to derive the phototactic potential  $\phi(I)$  from the data. First, the ability of the Keller-Segel model (Eq. 2.5) to capture the observed phototactic responses in different light regimes was tested. Fig. 2.3B (inset) shows that the mean cell density profiles  $\bar{\rho}(x)$  collapse on the same curve when plotted together as a function of the light intensity (via Eq. 2.6), thus supporting the applicability of Eq. 2.5 and the computation of  $\phi$  via Eq. 2.6, i.e.,  $\phi(I) = D \log \bar{\rho}[x(I)]$ . Second, the functional form of the phototactic potential  $\phi(I)$  was determined. To investigate the functional form of the phototactic potential, a set of models that have been used to describe sensing in chemotaxis (Tindall et al., 2008) was combined with a set of monotonically decreasing functions aimed at reproducing the photophobic behavior at high light intensity. The resulting functional forms were compared via the Akaike Information Criterion (AIC) to compare their performance in predicting the data. The first set of models, which consists of monotonically increasing functions of light intensity, is:

- $\phi_1(I) = aI$
- $\phi_2(I) = a \frac{I}{1+bI}$
- $\phi_3(I) = a \log(1 + bI)$

These models have been used extensively to describe chemotactic responses (for comparison with the theoretical background provided in this chapter and with Tindall et al. (2008), such models should be compared with the quantity  $\int \chi(c) \partial c / \partial x dx$ ). The second set of models consists of monotonically decreasing functions of  $I$ :

- $\phi_A(I) = -\log(1 + cI)$
- $\phi_B(I) = -c\sqrt{I}$
- $\phi_C(I) = -cI$



**Figure 2.3:** Computation of the phototactic potential  $\phi(I)$ . (A) Light intensity profiles, for different peak intensities,  $I_0$ . (B) Phototactic potential  $\phi(I)$  computed from Eq. 2.6 via inversion of the light intensity profile  $I(x)$  (panel A). The solid black line is the mean value of  $\phi(I)$  over the stationary density profiles for the various  $I_0$ , while the blue and gray regions represent the 68% and 95% confidence intervals, respectively. The dashed red line is the best fit of the phototactic potential predicted by the modified receptor law, Eq. 2.7. Inset: the phototactic potential calculated from each of the stationary density profiles (color-coded by light intensity regime, see panel A and Fig. 2.1A-G) at different  $I_0$  collapse on the same curve (displayed on the  $y$ -axis is the quantity  $\phi(I) = D \log \langle \bar{\rho}(I) \rangle$ , where the mean is over the five replicates with same  $I_0$ ), proving the applicability of the Eq. 2.5. Axes labels and ticks are as in the enclosing figure. (C-D) Mean cell density profiles measured at steady state (solid lines) and predicted from Eq. 2.6 (dashed lines), color-coded according to the light intensity regime (see panel A and Fig. 2.1A-G).



The functional forms in the second set were chosen to allow  $\lim_{I \rightarrow \infty} \phi = -\infty$  (some of the combinations do not satisfy this limiting behavior, resulting in poor fits). In fact, experimental observations show that  $\bar{\rho}(x) = 0$  if the light intensity in  $x$  is too high. In such situation,  $\phi(x) = D \log \bar{\rho}(x) = -\infty$ . Models from the first set were combined with models from the second set both in a multiplicative (e.g.,  $\phi_{1A} = \phi_1 \cdot (1 + \phi_A) = aI[1 - \log(1 + cI)]$ ) and additive (e.g.,  $\phi_{1A} = \phi_1 + \phi_A = aI - \log(1 + cI)$ ) fashion. All models were fitted to the data and the corresponding AIC values were computed; they are reported in Table 2.1. The best model according to the AIC is  $\phi_{2C} = aI(1 - cI)/(1 + bI)$ ; all other models have a  $\Delta$ AIC value (compared to the best model) larger than 7 and are thus unlikely (Burnham and Anderson, 2002). The AIC is unable to distinguish between the additive and multiplicative form of the model combination  $\phi_2$  and  $\phi_C$ , because the  $\Delta$ AIC difference between the additive combination  $\phi_{2C} = aI/(1 + bI) - cI$  and the multiplicative one  $\phi_{2C} = aI(1 - cI)/(1 + bI)$  is only  $\Delta$ AIC = -0.0005. The combination yielding the smallest AIC value, the multiplicative one, was thus assumed as the best model.

By using the AIC to formally quantify the relative performance of models in reproducing the experimental patterns discounting the number of parameters, the proposed generalization of the receptor law modified to account for the photophobic behavior shown at high light intensities reads:

$$\phi(I) = aI \frac{I_c - I}{I_r + I}, \quad (2.7)$$

where  $a = (1.4 \pm 0.04) \cdot 10^{-8} \text{ m}^4 \text{ W}^{-1} \text{ s}^{-1}$ ,  $I_r = 1.7 \pm 0.1 \text{ W m}^{-2}$  and  $I_c = 28.0 \pm 0.3 \text{ W m}^{-2}$  (standard errors are calculated via nonlinear least-squares fitting). The phototactic potential displays a maximum ( $\phi = 1.8 \text{ mm}^2/\text{s}$ ) at  $I_m = 5.5 \text{ W m}^{-2}$  (the light intensity value that separates the positive and negative phototaxis regimes) (Fig. 2.3B) and is equal to zero at  $I_c = 28.0 \text{ W m}^{-2}$ . Eq. 2.7 yields the best model for phototaxis in *E. gracilis* in reproducing the measured stationary cell density profiles (Fig. 2.3C-D).

**Table 2.1:** AIC values for the best fit of all model combinations, both multiplicative, e.g,  $\phi_{1A} = \phi_1(1 + \phi_A)$ , and additive, e.g.,  $\phi_{1A} = \phi_1 + \phi_A$ .

	Multiplicative			Additive		
	$\phi_A$	$\phi_B$	$\phi_C$	$\phi_A$	$\phi_B$	$\phi_C$
$\phi_1$	58	53	76	137	-34	137
$\phi_2$	-172	-174	-181	-162	-167	-181
$\phi_3$	-163	-164	-123	44	-166	-141

## Chapter 2. Phototaxis in phytoplankton

---

The proposed phototaxis model, though derived from stationary distributions, is shown to correctly capture also the temporal dynamics of phototaxis (Fig. 2.2). The formation of density peaks in the presence of light and their subsequent dissipation following light removal are reproduced well by the model (red dashed lines in Fig. 2.2A-F; note that Eq. 2.5 reduces to the diffusion equation in the absence of light stimuli). Small deviations from the model prediction during cell accumulation (Fig. 2.2A-C) are observed. They are possibly due to the repeated transfers of the channel from the illumination setup to the stereomicroscope for algal density measurements.

To characterize the swimming behavior at the single-cell level, trajectories of individual *E. gracilis* cells (Fig. 2.4) were recorded, both in a uniform light field and within a light gradient. 330 trajectories of cells were recorded through dark field microscopy, placing the recording window at the center of the microscope stage to minimize light gradients and thus bias in the direction of motion. The recorded trajectories were analyzed by computing the mean square displacement, mean square velocity, velocity autocorrelation and velocity distribution along the direction of the channel (Fig. 2.4). The  $x$ -coordinates of the recorded trajectories (i.e., the coordinate of each individual in the direction of the linear landscape, which was also the direction of the light gradient under the non-uniform light conditions) were analyzed. The statistics (Fig. 2.4) of cells' motion in uniform light are in good agreement with the Ornstein-Uhlenbeck (OU) process (Gardiner, 2006, Méndez et al., 2014):

$$\begin{cases} \dot{x} = v \\ \dot{v} = -\gamma v + \sigma \eta(t) \end{cases} \quad (2.8)$$

where  $x$  is the (one-dimensional) position of the cell,  $v$  is its instantaneous velocity and  $\eta(t)$  is a Gaussian white noise. Specifically, the mean square displacement  $\langle \Delta x^2(t) \rangle$ , mean square velocity  $\langle v^2(t) \rangle$  and velocity autocorrelation  $\langle v(t)v(0) \rangle$  were fitted to their analytical expressions (Méndez et al., 2014):

$$\langle \Delta x^2(t) \rangle_{OU} = \frac{\sigma^2}{\gamma^2} \left[ t - \frac{2}{\gamma} (1 - e^{-\gamma t}) + \frac{1}{2\gamma} (1 - e^{-2\gamma t}) \right] + \frac{v_0^2}{\gamma^2} (1 - e^{-\gamma t})^2, \quad (2.9)$$

$$\langle v^2(t) \rangle_{OU} = v_0^2 e^{-2\gamma t} + \frac{\sigma^2}{2\gamma} [1 - e^{-2\gamma t}], \quad (2.10)$$

$$\langle v(t)v(0) \rangle_{OU} = v_0^2 e^{-\gamma t}. \quad (2.11)$$

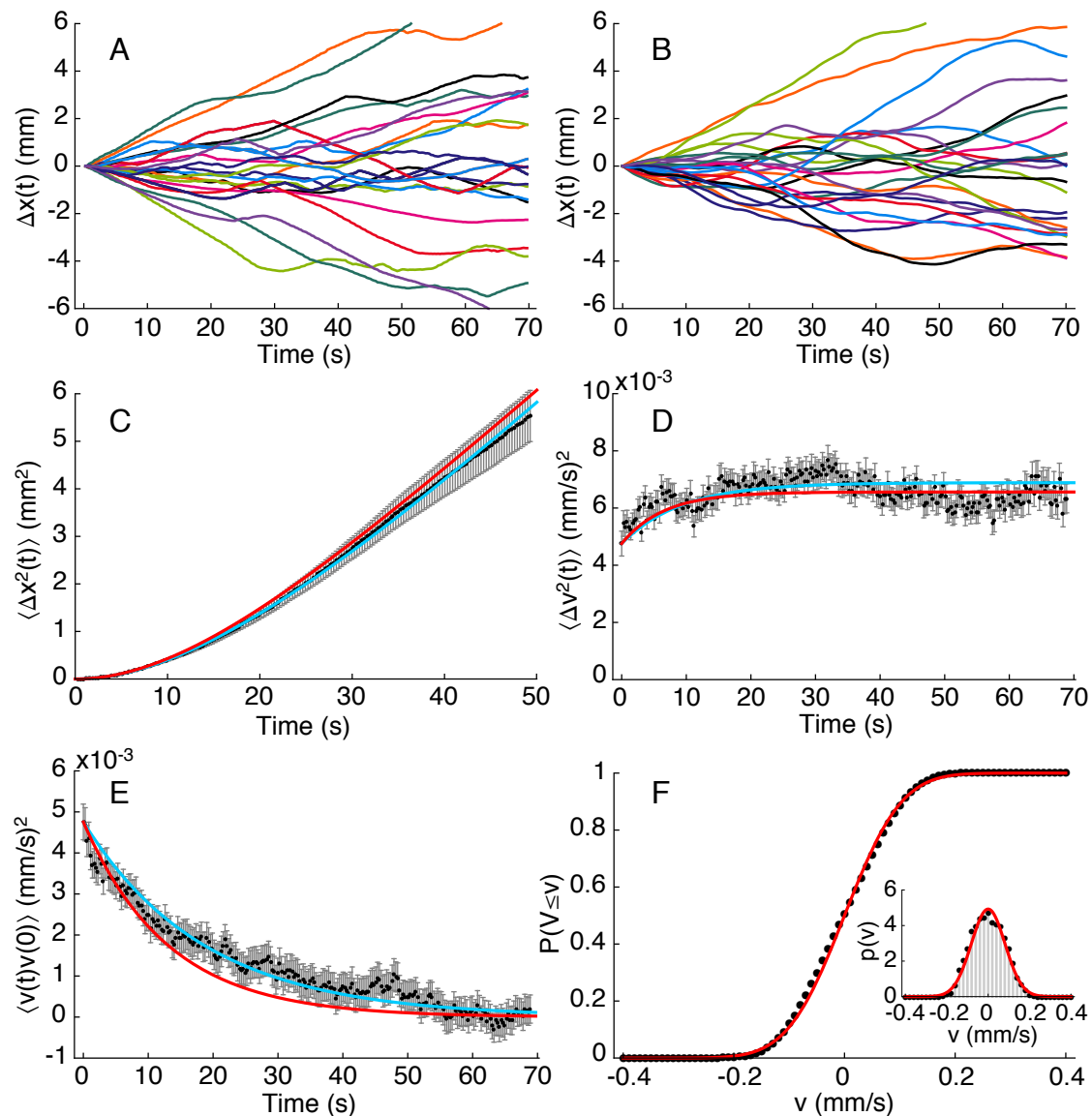
The quantity  $\gamma^{-1}$  measures the typical timescale of the velocity autocorrelation, while  $\sigma$  describes the degree of stochasticity of the motion. The fit was performed simultaneously for the three curves (red lines in Fig. 2.4), that is, the best fit parameters for  $\gamma$  and  $\sigma$  were those that minimized the chi-square  $\chi^2 = 1/2[\sum_{t=1}^T (\langle \Delta x^2(t) \rangle_{\text{data}} - \langle \Delta x^2(t) \rangle_{OU})^2 / \sigma_{\langle \Delta x^2(t) \rangle}^2 + \sum_{t=1}^T (\langle v^2(t) \rangle_{\text{data}} - \langle v^2(t) \rangle_{OU})^2 / \sigma_{\langle v^2(t) \rangle}^2 + \sum_{t=1}^T (\langle v(t)v(0) \rangle_{\text{data}} - \langle v(t)v(0) \rangle_{OU})^2 / \sigma_{\langle v(t)v(0) \rangle}^2]$ , where  $\sigma^2$  indicates the standard error (SE) of the mean in the data. The parameters' errors are given by the square root of the diagonal elements of the Hessian matrix, which is evaluated at the minimum. This fitting procedure provided the estimates  $\gamma = 0.077 \pm 0.014 \text{ s}^{-1}$  and  $\sigma = 0.032 \pm 0.004 \text{ mm/s}^{3/2}$  (mean $\pm$ SE). The cyan lines in Fig. 2.4 were obtained by fitting the velocity autocorrelation to its analytical expression Eq. 2.11 and subsequently fitting  $\sigma$  separately to the mean square displacement via Eq. 2.9 and to the mean square velocity via Eq. 2.10. This fitting procedure provided the estimates  $\gamma = 0.054 \pm 0.014 \text{ s}^{-1}$  and  $\sigma = 0.025 \pm 0.003 \text{ mm/s}^{3/2}$  (mean $\pm$ SE). The diffusive behavior observed at the population level finds additional confirmation at the individual level (at times  $t > \gamma^{-1} = \tau$ , Fig. 2.4C), with quantitative agreement between the diffusion coefficients at the two scales (for the trajectories data,  $D = \sigma^2 / (2\gamma^2) = 0.09 \pm 0.04 \text{ mm}^2 \text{ s}^{-1}$ , were the mean values of  $\gamma$  and  $\sigma$  obtained with the two fitting procedures were used). The mean (instantaneous) swimming speed of *E. gracilis* cells was  $\bar{v} = 0.10 \pm 0.05 \text{ mm s}^{-1}$ , mean $\pm$ SE.

130 trajectories of individual organisms were recorded in the presence of an imposed light gradient (Fig. 2.5), obtained by placing a LED with peak light intensity  $I_0 = 5.2 \text{ W m}^{-2}$  at the right border of the imaging window. No net displacement towards the light source (Fig. 2.5B) was found. The mean (instantaneous) swimming speed of *E. gracilis* cells was the same in the presence ( $\bar{v} = 0.10 \pm 0.04 \text{ mm s}^{-1}$ , mean $\pm$ SE) and in the absence of light gradients. The mean phototactic velocity  $\bar{v}_P = \langle d\phi/dx \rangle$  in the imaging window (the mean is computed over space) (Eq. 2) is  $\bar{v}_P = 0.007 \text{ mm s}^{-1}$ ; therefore, the directionality of swimming  $\bar{v}_P / \bar{v} = 0.07$  is very small.

Single-cell statistics allowed to derive a microscopic model of phototaxis, applicable to the motility behavior of individual *E. gracilis* cells. Despite the difficulty of discerning phototaxis at the single-cell level, the good agreement of trajectory statistics with the Ornstein-Uhlenbeck model in the absence of light gradients and the observation of accumulation dynamics around light sources at the population level suggest the following Langevin model for the phototaxis of individual cells,

$$\begin{cases} \dot{x} = v \\ \dot{v} = -\gamma v + \sigma \eta(t) + \gamma \frac{d\phi[I(x)]}{dx} \end{cases} \quad (2.12)$$

where  $x$  is the (one-dimensional) position of the cell,  $v$  is its instantaneous velocity,



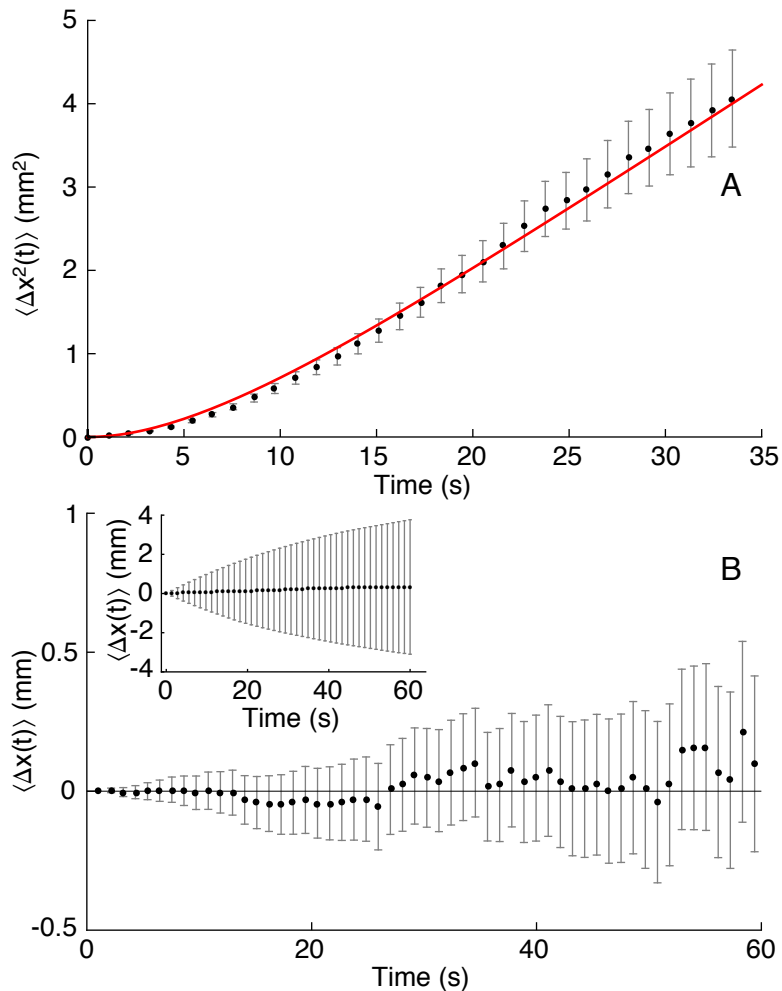
**Figure 2.4:** Single *E. gracilis* swimming trajectories in uniform light and fit to the OU process. The recorded swimming trajectories of algae (A) along the direction of the linear landscape resemble the OU process (B) both qualitatively (A,B) and quantitatively (C-F). (C-F) Statistics of the measured trajectories (black dots, mean $\pm$ SE) and fit to the OU process (red and cyan lines). The simultaneous fit (red lines) of the mean square displacement (C), mean square velocity (D) and velocity autocorrelation (E) shows that the OU process represents a good description of the movement behavior of individual algae. The cyan lines were obtained by fitting the parameter  $\gamma$  in the velocity autocorrelation data (E) and subsequently fitting  $\sigma$  in the mean square displacement (C) and mean squared velocity (D) data separately. (F) The simultaneous fit also provides a very good prediction for the stationary velocity distribution.

$\eta(t)$  is a Gaussian white noise, and  $\phi[I(x)]$  is the phototactic potential. In the presence of a light gradient, the term  $\gamma d\phi[I(x)]/dx$  in Eq. 2.12 drives the accumulation of individuals around the light source in the long term. This single-cell model is consistent with the Keller-Segel model at the population scale (Eq. 2.5) and reduces to the OU model in the absence of external gradients.

To interpret the failure to detect a bias towards the light source, 1000 integrations of Eq. 2.12 were performed, with initial positions drawn uniformly at random in the range  $[-10.5 \text{ mm}, -1 \text{ mm}]$  (i.e., the region measured in the experiment) and with initial velocities drawn according to the stationary velocity distribution of the OU process. Fig. 2.5B (inset) shows a plot of the computed mean displacement  $\langle \Delta x(t) \rangle$  and standard deviation in the simulations. Fig. 2.5B elucidates why no discernible net displacement towards the light source is appreciable in the data, that is, the random motion of *E. gracilis* dominates over the drift towards the source at these spatial and temporal scales. Accordingly, phototactic accumulation of density peaks takes place in a timeframe much larger than the typical persistence time  $\tau = 1/\gamma = 15 \text{ s}$ . Therefore, the model Eq. 2.12 provides interpretation for the impossibility to observe a net bias towards the source in the experimental trajectories.

**Table 2.2:** Parameters describing the phototactic response and movement dynamics of *E. gracilis* (mean $\pm$ SE). The parameters  $a$ ,  $I_r$  and  $I_c$  define the phototactic potential  $\phi$ . The diffusion coefficient  $D$  was estimated via the relaxation of density peaks (Fig. 2.2D-H).  $\tau$  is the typical autocorrelation time of swimming trajectories and was measured by fitting an exponential decay to the velocity autocorrelation of the experimental trajectories measured in uniform light. The mean instantaneous cell swimming velocity  $\bar{v}$  was measured in the recorded trajectories and was the same in uniform and non-uniform light.

<b>Phototactic response</b>	
$a$	$(1.4 \pm 0.04) \cdot 10^{-8} \text{ m}^4 \text{ W}^{-1} \text{ s}^{-1}$
$I_r$	$1.7 \pm 0.1 \text{ W m}^{-2}$
$I_c$	$28.0 \pm 0.3 \text{ W m}^{-2}$
<b>Movement dynamics</b>	
$D$	$0.13 \pm 0.04 \text{ mm}^2 \text{ s}^{-1}$
$\tau$	$15 \pm 3 \text{ s}$
$\bar{v}$	$0.10 \pm 0.05 \text{ mm s}^{-1}$



**Figure 2.5:** *E. gracilis* trajectories in a light gradient. (A) The statistics of the trajectories are consistent with a persistent random walk. (B) Mean displacement ( $\pm$ SE): at these spatial and temporal scales there is no discernible net displacement towards the light source ( $\Delta x > 0$ ). The inset in panel (B) shows the mean displacement and standard deviation of 1000 trajectories, which were simulated according to Eq. 2.12. The random component of the motion is much stronger than the force term in Eq. 2.12 and thus hides the mean net displacement of the individuals towards the light.

## 2.4 Methods

### 2.4.1 Algal culture

The species used in the experiments, *E. gracilis*, was purchased from Carolina Biological Supply (Burlington, North Carolina, USA) and maintained in a nutrient medium (Altermatt et al., 2014) composed of sterilized spring water and Protozoan Pellets (Carolina Biological Supply) at a density of  $0.45 \text{ g L}^{-1}$ , filtered through a  $2 \mu\text{m}$  filter. Algal cultures were initialized 2 weeks before the start of the experiment and kept at a constant temperature of  $22 \text{ }^\circ\text{C}$  under constant LED light at  $\lambda = 469 \text{ nm}$ . *E. gracilis* individuals have a typical linear size of  $14 \mu\text{m}$  (Giometto et al., 2013) and the duplication time is approximately 20 h (Carrara et al., 2012); thus, reproduction can be neglected in the experiments.

### 2.4.2 Density estimates and video recording

Density estimates were obtained by placing the linear landscape under the objective of a stereomicroscope, taking pictures and counting individuals through image analysis (Altermatt et al., 2014). Density profiles were measured at the centre of the linear landscape across one entire period of the light intensity profile. Stationary density profiles were measured after 210 min from the introduction of cells in the landscape. Border effects were neglected here because the measurements were performed at the centre of the landscape, which had a total length (2 m) that was much larger than the length of one period (12.5 cm). To reconstruct the trajectories, videos were recorded with a stereomicroscope and particle tracking was performed automatically with the MOSAIC (Sbalzarini and Koumoutsakos, 2005) plugin for ImageJ under homogeneous light conditions (Altermatt et al., 2014) and manually with the MTrackJ plugin for ImageJ in the presence of a light gradient (automatic tracking was not possible in the light gradient setup because of the low quality of pictures, due to the use of only one LED light for the microscopy).

### 2.4.3 Linear landscapes

The linear landscapes used in the experiments were channels drilled on a plexiglass sheet and were 2 m long, 5 mm wide and 3 mm deep (Giometto et al., 2014). A second plexiglass sheet was used as a cover, and a gasket prevented water spillage. Before the introduction of the algal culture in the linear landscapes, the plexiglass sheets were sterilized with a 70% (vol/vol) ethanol solution, and the gaskets were autoclaved.

### 2.4.4 Light sources and light intensity profile

A linear array of LEDs was developed to control the light intensity profile along the linear landscapes. RGB (Red Green Blue) LED strips were controlled via Arduino Uno boards. The LED strips consisted of individually addressable LEDs separated by a distance of 3.12 cm. The light intensity for the B (Blue) and R (Red) color channels (wavelength of 463 – 475 nm and 619 – 635 nm respectively) could be controlled. The total radiant flux emitted by LEDs at the different intensities and wavelengths employed was measured with a calibrated photodiode. The relative light intensity profiles, with the LEDs set at the different intensities employed, was measured by placing a white paper sheet in the linear channels and measuring the irradiance on the sheet with a digital camera operated in grayscale at fixed aperture, exposure and distance from the LED. This relative measure of light intensity was converted to absolute values via the total radiant flux measured. In the experiments, periodic light intensity profiles were established with one LED switched on every 12.5 cm.

### 2.4.5 Numerical integration

To compute the time evolution of algal accumulation according to the Keller-Segel model (Fig. 2.2) Eq. 2.5 was integrated numerically with the method of lines (Schiesser, 1991) in the domain  $x \in [-6.25, 6.25]$  cm, whose total length of 12.5 cm corresponds to one period of the periodic light intensity profile established in the experiments. The initial condition was uniform and equal to the mean cell density. Reflecting boundary conditions were set at the border of the domain. Eq. 2.12 was integrated numerically with the Euler-Maruyama method of order 1/2 to compute the model predicted mean displacement and standard deviation for the experimental settings (inset in Fig. 2.5B). The light intensity profile used in the numerical integrations of Eqs. 1, 2.12 was the best fit of the Eq.  $I(x) = c_0/(x^2 + c_1^2)^2$ , which approximates very well the measured profile.

### 2.4.6 The expansion of the Fokker-Planck equation

The expansion of the Fokker-Planck equation for the Langevin Eq. 2.12 in  $\gamma^{-1}$  is acceptable because the typical persistence time  $\tau = \gamma^{-1} = 15$  s of the trajectories is much smaller than the typical timescale for the macroscopic dynamics. An intuitive derivation of the expansion can be obtained by neglecting the inertial term  $\dot{v}$  in Eq. 2.12 (a technique known as adiabatic elimination of fast variables, Gardiner (2006)), which results in the Langevin Eq.  $\dot{x} = \frac{\sigma}{\gamma}\eta(t) + \frac{d\phi}{dx}[I(x)]$ . The corresponding Fokker-Planck equation describing the time evolution of the probability density function



$\rho(x, t)$  is then  $\frac{\partial \rho}{\partial t}(x, t) = \frac{\sigma^2}{2\gamma^2} \frac{\partial^2 \rho}{\partial x^2}(x, t) - \frac{\partial}{\partial x} \left[ \frac{d\phi}{dx}[I(x)]\rho(x, t) \right]$ , which is equivalent to Eq. 2.5 for  $D = \sigma^2/(2\gamma^2)$ .

## 2.5 Discussion

To compare the experimental setup employed with the natural environment, the ASTM G-173 reference terrestrial solar spectral irradiance (ASTM International, 2008) was integrated in a wavelength window of 10 nm centered at  $\lambda = 469$  nm (10 nm is the typical width of emission for the LEDs, Methods) and was found to give a typical irradiance of approximately  $13 \text{ W m}^{-2}$  at sea level. Wavelengths in the blue region of the visible spectrum are among the most transmitted in natural aquatic habitats (Jeffrey, 1984, Ragni and Ribera D'Alcalà, 2004) and penetrate the farthest in the water column, whereas red light is the most attenuated. Thus, the experimental values of light intensities and the wavelengths adopted here are typical of natural conditions and suggests that the experimental and theoretical results might have implications for the behavior of phytoplankton in natural environments.

The response of cells to light of different intensities, here expressed in terms of the phototactic potential  $\phi(I)$ , was inferred from observed stationary cell density profiles. However, the model was shown to capture also the temporal dynamics of cell accumulation around a light source and the diffusive relaxation following light removal. This is suggested to be an important result in view of studying the effect of stochastic light fields in the laboratory on the migration of cells (Chapter 3). In particular, the experimental approach to phototaxis adopted here provides a template for the study of ecological processes in shifting and fluctuating resource availability. The convenient use of programmable LEDs allows one to create microbial microcosms in which light conditions can be accurately controlled to generate a boundless variety of spatiotemporal patterns of environmental stochasticity, affecting both the growth and the movement behavior of cells. Hence, the study system developed here is suggested to be a promising candidate for quantitative microcosm experiments on biological invasions along ecological corridors, range expansions and source-sink dynamics under environmental noise (Bertuzzo et al., 2007, Gonzalez and Holt, 2002, Méndez et al., 2011, Rodriguez-Iturbe et al., 2009).

Two main phenomenological approaches have been adopted in the literature to model phototaxis of motile algae (Vincent and Hill, 1996, Williams and Bees, 2011a,b), both acknowledging the lack of experimental verification. The first is a photokinetic approach (Burkart and Häder, 1980, Vincent and Hill, 1996, Williams and Bees, 2011a), where the average swimming velocity of cells is assumed to be a function of the light

## Chapter 2. Phototaxis in phytoplankton

---

intensity  $I$ . Vincent and Hill (1996) assumed that the average cell velocity in a vertical light gradient is given by  $v_p(I) = vT(I)$ , where  $v$  is the average cell swimming speed and  $T(I)$  is a suitable taxis response function (see Fig. 2.1 in Vincent and Hill (1996)). Analogously, Williams and Bees (2011a) proposed a photokinetic model where the average cell velocity is a linear function of  $I$ , i.e.,  $v_p(I) = v/I_c \cdot (I - I_c)$ , equal to  $v$  in the dark and equal to zero at  $I_c$ . Such model assumptions were taken as working hypotheses due to the lack of direct or indirect empirical validation (Vincent and Hill, 1996, Williams and Bees, 2011a). The need for experimental characterization of the phototactic response to move beyond the linear assumption was clearly stated therein (Williams and Bees, 2011a). Such models described phototaxis in relatively shallow settings (a few centimeters) where light came uniformly from above and heterogeneity in the light availability would be solely induced by cell shading. In Vincent and Hill (1996), Williams and Bees (2011a), light was assumed to control the magnitude and the sign of the cell swimming velocity, while the directional bias in the vertical direction is provided by gyrotaxis and gravitaxis. Thus, cells were assumed to exhibit no net average displacement in the absence of gravity. Burkart and Häder (1980) performed a light-trap experiment with the alga *Phormidium uncinatum* and used experimental observations to derive an advection model with average cell velocity  $v_p[I(x)] = \alpha \log[\beta I(x)]$  (for  $I \geq \beta^{-1}$  and with  $\alpha > 0$ ), neglecting cells diffusion. Such model cannot reproduce the negative phototaxis behavior observed in *E. gracilis*, because  $v_p$  is assumed to be a monotonically increasing function of  $I$ . Thus, cells would accumulate at the highest available light intensity value, contrary to the experimental results (Fig. 2.3D).

The second modeling approach consists of assuming an advection flow proportional to the light intensity gradient,  $v_p[dI/dx] = \chi dI(x)/dx$  ( $\chi > 0$ ), such as in Torney and Neufeld (2008), where cell aggregation in turbulent flows was investigated theoretically. The proportionality of the advection flow to  $dI/dx$  ensures the existence of a net flow towards regions of higher light intensity. However, because  $v_p$  is not a function also of the light intensity  $I$ , possible saturating effects or photophobic behavior at high light intensities are neglected. The response function chosen in Torney and Neufeld (2008) corresponds to the choice  $\phi(I) = \phi_1(I) = aI$ , which gives a bad fit for the data shown in Fig. 2.3B, even if only the data points up to  $I_m = 5.5 \text{ W/m}^2$  (maximum of the experimental  $\phi(I)$ ) are used. Specifically, if the data for  $\phi$  (Fig. 2.3B) are truncated at  $I_m = 5.5 \text{ W/m}^2$  the functional form  $\phi_2$  performs best compared to  $\phi_1$  and  $\phi_3$  ( $\Delta\text{AIC} = 145$  and  $26$ , respectively). Other mechanistic approaches have been explored to model phototaxis in combination with gyrotaxis. For example, Williams and Bees (2011a) considered other two models: one where cells exhibited a centre-of-mass offset dependent on light intensity; and another in which a reactive phototactic torque was introduced. However, because several additional processes are

included, comparison of such models with data on photo-bioconvection patterns are difficult and mostly qualitative. Williams and Bees (2011b) provided the first replicated experimental investigation of photo-bioconvection patterns. They found qualitative agreement between the experimental data and the model predictions (Williams and Bees, 2011a) for the dominant initial wavelength of bioconvective patterns obtained via linear stability analysis (Williams and Bees, 2011b). Limitations to the quantitative comparison of the models and the experiments are discussed therein (Williams and Bees, 2011b).

In the experiments, light came from below and cells were observed to distribute mostly at the top of the channels. Therefore, shading was neglected (unlike in Vincent and Hill (1996), Williams and Bees (2011a)) and light gradients are present by design owing to the experimental setup. The assumption that the phototactic velocity  $v_P$  is a function of the light intensity  $I$  (and not of  $I$  and  $dI/dx$ ) such as in Burkart and Häder (1980), Vincent and Hill (1996), Williams and Bees (2011a) is unfeasible here, because it would induce a net phototactic movement in uniform light settings, without the existence of a preferential direction in the horizontal plane. The fact that the advection velocity is a function of both  $I$  and  $dI/dx$  is a common feature of Keller-Segel models, compare e.g. Eq. 1 in Tindall et al. (2008) and Eq. 2.2 here, and ensures that no net movement is induced in homogeneous distributions of the stimulus (be it a chemical for chemotaxis or light for phototaxis). The framework presented here differs from previous attempts to model phototaxis for the fact that the phototactic velocity here is a function of both the light intensity  $I$  and its spatial gradient  $dI/dx$ , i.e.,  $v_P = d\phi/dx = d\phi/dI \cdot dI/dx$  ( $\phi$  is a function of  $I$ ). A model capable of describing both the positive and negative phototaxis regimes (a feature present in Vincent and Hill (1996), Williams and Bees (2011a) without direct experimental validation, and absent in Burkart and Häder (1980), Torney and Neufeld (2008)) is deemed desirable because negative phototaxis has important ecological consequences as it allows cells to avoid harmful radiation. Moreover, negative phototaxis contributes to determine the vertical positioning of sensible organisms in the water column (Häder and Griebenow, 1988), is a widespread behavior in phytoplankton (Jékely, 2009), and its exploitation was suggested as an efficient technique to enhance harvesting in photobioreactors (<http://www.google.com/patents/US20100237009>) and to control microorganisms for the transport of colloidal cargo (Itoh, 2004).

Compared to previous research efforts, the present investigation allowed a quantitative experimental determination of the phototactic response function (embedded in the potential  $\phi$  and, by derivation, in the advection velocity  $v_P$ ) directly from cell density patterns in a broad range of light intensities, thus allowing the characterization of both negative and positive phototaxis regimes within a unified mathematical framework.

## Chapter 2. Phototaxis in phytoplankton

---

The simple experimental settings devised allowed a direct quantitative comparison of the model with the experiments and invites further experimental investigation in more complex scenarios.

All things considered, it is suggested that the literature lacked an experimentally tested mathematical framework comprising a measure of the phototactic response function of individual organisms and their associated population behavior. This work is thus suggested to provide the blueprint for the characterization of the collective response of phytoplankton to light availability and its migration strategies in aquatic ecosystems. The identification of the light intensity window where positive and negative phototaxis occurs pinpoints the regions of the water column where phototactic effects affect the vertical distribution of phytoplankton. Currently, models of phytoplankton growth in contrasting gradients of light and nutrients aimed at reproducing the vertical distribution of phytoplankton either ignore phototaxis (Greenwood and Craig, 2014) or rely on untested assumptions for the phototactic advection velocity  $v_p = d\phi[I(x)]/dx$  (Klausmeier and Litchman, 2001, Mellard et al., 2011). The identification of the functional form for  $\phi(I)$  provided here can be used directly to integrate realistic predictions for the phytoplankton vertical distribution, which is known to be relevant for global biogeochemical cycles, ecosystem functioning and the diversity and coexistence of plankton species (Lampert et al., 2003, Steele and Yentsch, 1960). In fact, the inter-specific variability of the optimal light intensity (defined by  $d\phi(I)/dI = 0$ ) and nutrient requirements have been argued (Yoshiyama et al., 2009) to translate into a sectoring of the water column into separate niches, allowing the coexistence of competitive species.

Other direct applications of the mathematical framework outlined above can be envisioned. For example, phototaxis of swimming algae, sometimes in combination with other directional behaviors such as gravitaxis (the directed swimming in response to gravity) and gyrotaxis (gravitaxis in the presence of ambient velocity gradients), is speculated to have noteworthy implications for the design of algal photobioreactors. The phototaxis model proposed here (Eq. 2.5) may be used directly to refine existing models for photo-gyrotactic (Williams and Bees, 2011a) and photo-gyro-gravitactic (Williams and Bees, 2011b) bioconvection, which currently rely on educated guesses for the phototactic advection term, and may be applied to identify optimal designs for cell accumulation far from the reactor surface to avoid biofouling and to achieve enhanced harvesting, a strategy that has been investigated empirically (Ooka et al., 2014). The fact that the phototactic potential  $\phi$  is much steeper for light intensities above  $I_m = 5.5 \text{ W/m}^2$  than below such value, and hence the phototactic velocity is larger for  $I > I_m$ , suggests that the exploitation of negative phototaxis might be a more effective strategy than the use of positive phototaxis to achieve optimal harvesting.

Algae are also increasingly employed in micro-bio-machine research, for example as micro-propellers for the transport of colloidal cargo (Itoh, 2004, Itoh and Tamura, 2008), where light is often used as the external driver of the motion. Although this research is yet to translate into practice, it represents an exciting avenue to harness microbial motility for controlled microscale applications, and phototaxis represents one of the most controllable processes because of the ease of accurately imposing and rapidly modulating external light gradients. The algorithms that are currently employed to control such micro-bio-machines are mostly empirical, and the model Eq. 2.5 may indeed serve to render machine control more robust and accurate. While much attention is currently dedicated to understanding the swimming behavior in these artificial environments (Garcia et al., 2013), the characterization of collective phototactic dynamics provided here might be exploited to optimize existing technological applications or design new ones.



### 3 Invasion fronts: environmental heterogeneity

#### Abstract

Biological invasions typically spread in heterogeneous and fluctuating environments. However, most of the current knowledge on biological invasions is based on theoretical models and a handful of experiments that assumed a homogeneous environment and thus neglect the heterogeneity that is ubiquitous in nature. Recently, theoretical models have shown that spatially and temporally uncorrelated environmental fluctuations can cause a slowing-down in the propagation of invasion fronts, but the role of autocorrelation in the landscape spatial heterogeneity has not been investigated. Here, a generalization of the Fisher-Kolmogorov model including demographic stochasticity and biased movement of populations towards favorable regions is proposed and the front propagation dynamics is studied in landscapes with different autocorrelation lengths of the resource distribution. The speed of the invasion spread is shown to be a monotonically decreasing function of the resource autocorrelation length and demographic stochasticity is found to play a key role in such slowing-down effect. The prediction that biological invasions spread slower in landscapes with large autocorrelation length is verified and confirmed in a microcosm experiment with the alga *Euglena gracilis*. Such results may have implications for the control of invasive species in the natural environment.

### 3.1 Introduction

Environmental fluctuations and heterogeneity are ubiquitous in nature and are thought to affect nearly all aspects of ecology, ranging from species coexistence to population synchrony, driving range shifts and potentially causing abrupt biotic change. Local population dynamics in fluctuating and heterogeneous environments have been studied extensively in recent years (Chesson and Huntly, 1997, Duncan et al., 2013, Gonzalez and Holt, 2002), mainly with respect to population synchrony (Benton et al., 2001, Fox et al., 2011, Post and Forchhammer, 2002, Vasseur and Fox, 2009). However, the implications on spatial dynamics (Duncan et al., 2013, Gonzalez and Holt, 2002) and especially on front propagation (Méndez et al., 2011) received little experimental attention. Accordingly, most experimental investigations were dedicated to the study of temporal fluctuations, whereas spatial heterogeneity received surprisingly little attention (Melbourne et al., 2007). Both theoretical (Roy et al., 2005, Vasseur, 2007) and experimental (Fontaine and Gonzalez, 2005, García-Carreras and Reuman, 2011, Gonzalez and Holt, 2002) studies have highlighted the relevance of the autocorrelation structure of environmental fluctuations for ecological dynamics. For example, Fontaine and Gonzalez (2005) showed that positively autocorrelated (i.e. red) temporal prey fluctuations induced synchrony in the predator population dynamics in a microcosm composed of the alga *Chlorella vulgaris* and the rotifer *Brachionus calyciflorus*, whereas uncorrelated fluctuations did not.

Despite the fact that most natural environments are inevitably heterogeneous, much of the current understanding of species spread is based on theoretical models (Hastings et al., 2005, Melbourne et al., 2007) and only on a handful of experiments (Giometto et al., 2014, Melbourne and Hastings, 2009) that investigated spread in uniform environments. Melbourne et al. (2007) stated the need for experiments and theory addressing spread and invasions in more complex, heterogeneous environments. From the theoretical viewpoint, integrating environmental heterogeneity in models of spread is a challenging task and a modeling framework that allows drawing general conclusions is lacking to date (Hastings et al., 2005). In the search for such a framework, the study of biological invasions in heterogeneous and fluctuating environments has been addressed in the context of the Fisher-Kolmogorov equation (Kinezaki et al., 2003, Méndez et al., 2003, 2011, Nelson, 1999, Nelson and Schnerb, 1998), by embedding various sources of environmental stochasticity in the original deterministic equation (Fisher, 1937, Kolmogorov et al., 1937). However, most of these studies modeled spatially heterogeneity and fluctuations through noise terms that were uncorrelated in space (Nelson, 1999, 2012, Nelson and Schnerb, 1998), periodic in space (Kinezaki et al., 2003, Shigesada et al., 1986) or characterized by a gaussian spatial correlation function with a fixed correlation length (Méndez et al.,



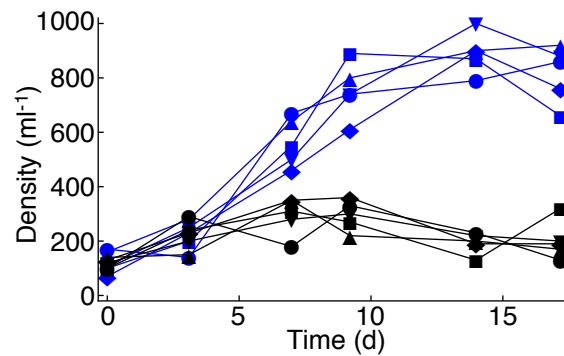
2011). Therefore, whereas the importance of the autocorrelation structure of temporal environmental fluctuations for local ecological processes is now widely recognized (Fontaine and Gonzalez, 2005, García-Carreras and Reuman, 2011, Gonzalez and Holt, 2002, Vasseur, 2007), the effect of the autocorrelation structure of environmental fluctuations on biological spread rates is rather unexplored. A limited number of empirical works (Bailey et al., 2000, Bergelson et al., 1994, Hastings et al., 2005, Williamson and Harrison, 2002) measured spread rates in heterogeneous and diverse habitats and compared the realized spread distances in patchily distributed sites (Bailey et al., 2000, Bergelson et al., 1994). However, microcosm experiments testing specific predictions of Fisher-Kolmogorov models with embedded environmental stochasticity are lacking to date (Méndez et al., 2011).

Motivated by the above considerations, this chapter studies biological invasions in the presence of spatial heterogeneous resource distributions. This, for example, could reflect the spatial composition and quality of soil or topographically determined habitat elements such as exposure or elevation. Here, focus is placed on the effect of the spatial autocorrelation structure of the resource distribution on the propagation of biological invasion fronts. Giometto et al. (2014) (Chapter 1) showed that, even in uniform landscapes, demographic stochasticity introduces noise terms in the reaction-diffusion equation describing the propagation of invasion fronts, leading to a quantifiable variability of the process across replicated invasions (Fig. 1.3). Therefore, both environmental and demographic stochasticity are thought to affect biological invasions and the interplay between these two sources of stochasticity will be investigated here.

This chapter is organized as follows. First, the experimental setup is discussed. Second, a mathematical model that embeds the processes of positive phototaxis and demographic stochasticity is introduced and its predictions are used to formulate the experimental hypothesis. Third, the experimental hypothesis is tested in a microcosm experiment with the species *Euglena gracilis*.

## 3.2 Experimental setup

The setup developed in Chapters 1 and 2 allows to study experimentally biological invasions in continuous one-dimensional landscapes. The use of the photosynthetic organism *E. gracilis* allows to impose environmental heterogeneity along the landscapes by controlling the light intensity profile via a linear array of Light Emitting Diodes (LEDs). Such light intensity was shown in Chapter 2 to affect the movement behavior of *E. gracilis*, introducing a bias on motility towards light in the positive pho-



**Figure 3.1:** Replicated growth curves of *E. gracilis* in the presence of blue light (blue curves) and in the dark (black curves). Replicated populations starting at identical initial conditions are shown to grow in the presence of light, but not in the dark.

totaxis regime and causing the accumulation of algae in the presence of light. Growth curves of *E. gracilis* were measured in the presence of light by placing a cell culture flask with 10 mL of *E. gracilis* on top of two LEDs operated at a total flux of 1 mW each. Growth curves in the absence of light were measured in the same conditions, but covering the LEDs with black tape. Fig. 3.1 shows that *E. gracilis* grows in the presence of light, but not in the dark. Therefore, light in our experimental setup affected both the growth and the movement behavior of *E. gracilis*.

The experimental setup allowed to impose the desired distribution of resources (i.e., light) along the linear landscapes via the control of linear arrays of LEDs. Thus, resources in the experimental landscapes were centered at regularly spaced lattice sites (the fixed distance between two successive LEDs was  $\Delta L = 3.12$  cm) and different resource distributions could be cast in place by switching on or off each LED in the linear array. For a fixed sequence of switched-on LEDs, the light intensity profile was given by the superposition of the spatial light intensity profile of each switched-on LED in the sequence. Light intensity profiles with the desired autocorrelation length were designed by imposing the probability  $\lambda$  of LED  $i + 1$  in the LED array to be switched-on if LED  $i$  was switched-off, i.e.,  $\mathbb{P}[\text{LED}(i + 1) = \text{ON} \mid \text{LED}(i) = \text{OFF}] = \lambda$ . Such Markov Chain was imposed to be symmetric such that  $\mathbb{P}[\text{LED}(i + 1) = \text{OFF} \mid \text{LED}(i) = \text{ON}] = \lambda$ . Small and large values of  $\lambda$  generate resource distributions with long and small autocorrelation lengths (approximately equal to  $\Delta L/\lambda$ ), respectively. Because landscapes were of finite total length in both the experiment and in the simulations, the above procedure can generate by chance resource profiles with autocorrelation different from the desired one and with a mean frequency of switched-on LEDs different from 1/2. Therefore, in both the experiment and the simulations, the set of resource profiles obtained with the Markov Chain procedure was restricted to those with a mean frequency of switched-on LEDs equal to 1/2 and in a narrow window

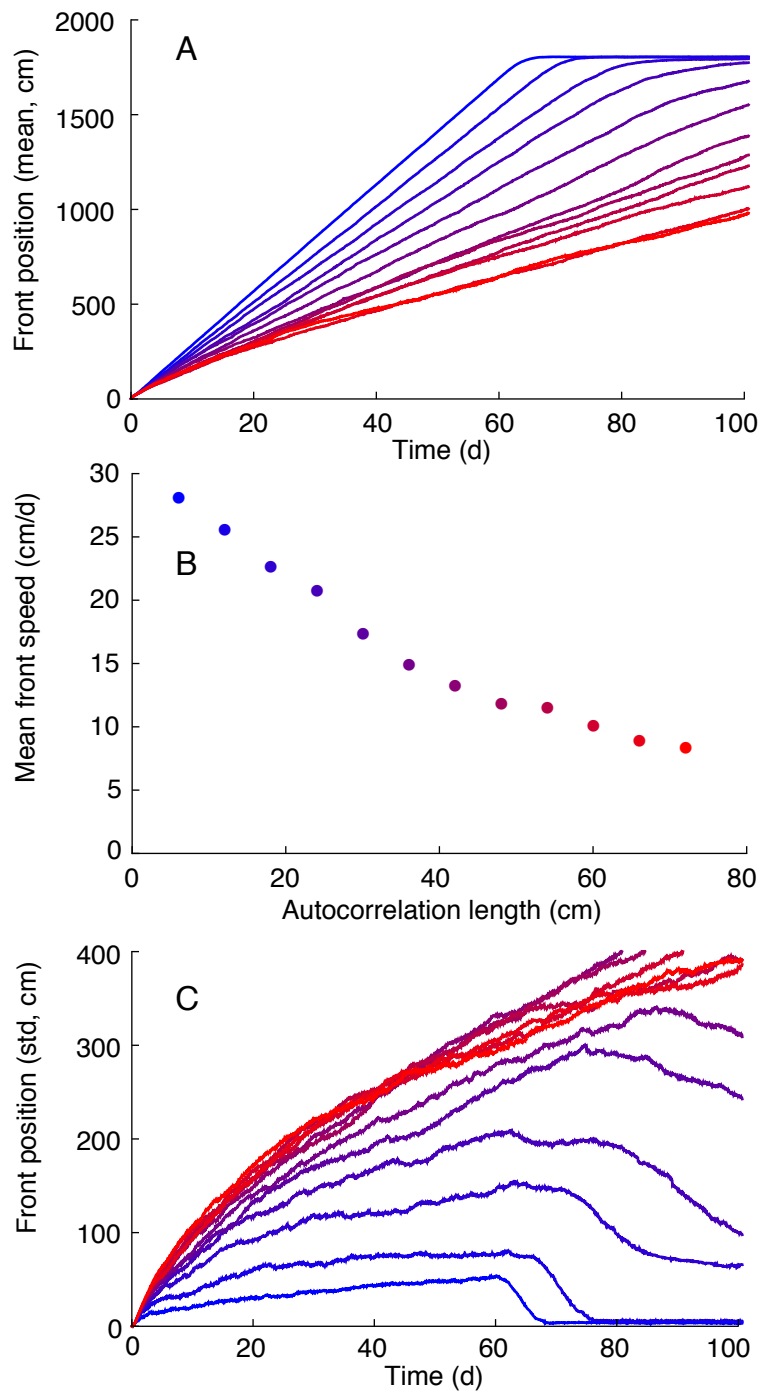
of autocorrelation length around the desired one. The LED sequences employed in the experiment are shown in Fig. 3.3. The choice of the large autocorrelation length in the experiment was limited by the total finite length of the experimental setup and was chosen to be less than 1/10 of the total setup length.

### 3.3 Results

A mathematical model was developed to assess the contribution of demographic stochasticity and of the directed movement of organisms to the front propagation in heterogeneous environments. The model was then used to identify the experimental hypothesis. The proposed mathematical model is a combination of Eqs. 1.18 and 2.5, that is, it is a generalization of the Fisher-Kolmogorov model that accounts for the directed movement towards the resource (Eq. 2.5) and for demographic stochasticity (Eq. 1.18). The model reads:

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x} \left[ D \frac{\partial \rho}{\partial x} - \frac{d\phi}{dx}(I)\rho \right] + r(I)\rho \left[ 1 - \frac{\rho}{K} \right] + \sigma \sqrt{\rho} \eta, \quad (3.1)$$

where  $\rho$  is the density of organisms,  $D$  is the diffusion coefficient of the cells due to their active movement ( $D$  for *E. gracilis* was measured in Chapter 2),  $\phi$  is the phototactic potential as in Eq. 2.7,  $r$  is the species growth rate,  $K$  is the carrying capacity,  $\eta = \eta(x, t)$  is a gaussian, zero-mean white noise (i.e., the noise has correlations  $\langle \eta(x, t)\eta(x', t') \rangle = \delta(x - x')\delta(t - t')$ , where  $\delta$  is the Dirac's delta distribution) and  $\sigma > 0$  is constant. The landscape heterogeneity is embedded in the resource profile  $I(x)$ . The Itô's stochastic calculus (Gardiner, 2006) was adopted, as appropriate for the demographic noise term (see Chapter 1). The growth rate  $r$  is assumed to be a function of the light intensity, as suggested by the experimental growth curves (Fig. 3.1). The assumption made here is that  $r$  follows Monod kinetics (the assumption is customary for phytoplankton, Diehl (2002)), i.e.,  $r(I) = r_{\max}I/(I + K_I)$ , where  $r_{\max}$  is the maximum growth rate and  $K_I$  is the half-saturation constant. Eq. 3.1 was used to simulate biological invasions in linear landscapes with resource distributions  $I(x)$  exhibiting various autocorrelation lengths. The profiles  $I(x)$  employed in the simulations were chosen to reproduce closely the experimental light intensity profiles, although landscapes were much longer in the simulations than in the experiments; details are provided in the Methods. 100 landscapes for each resource autocorrelation length were generated and front propagations according to Eq. 3.1 were simulated on such landscapes, with parameters suitable to describe *E. gracilis* (Methods). The position of the front was measured in the simulations by fixing a threshold value of the density  $\bar{\rho}$  and recording the furthest point from the origin where the cell density was higher than such value. The mean propagation speed for each value of the resource

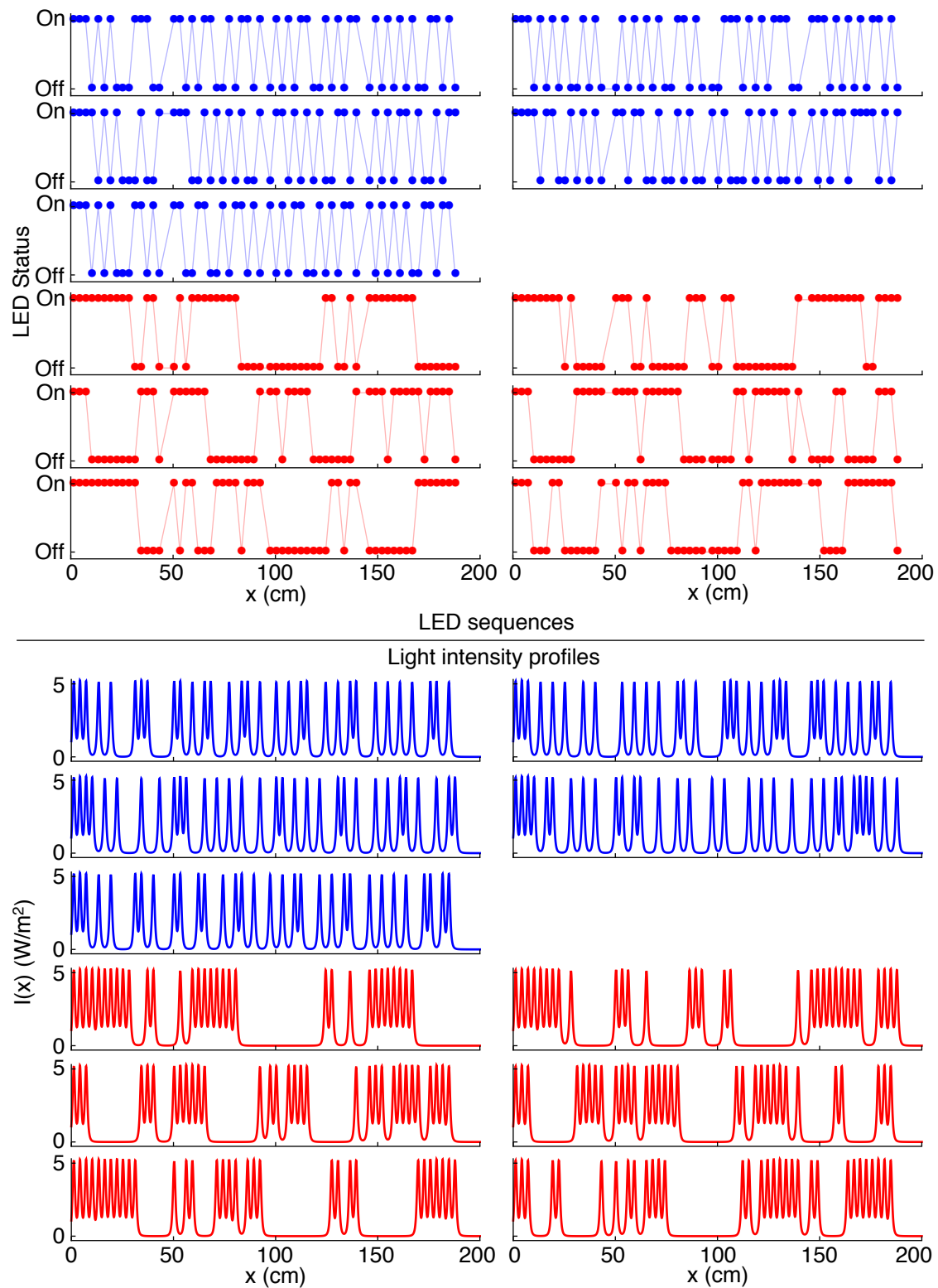


**Figure 3.2:** Mean front position (A), mean propagation speed (B) and standard deviation (C, std) of the front position in landscapes with different autocorrelation length, according to Eq. 3.1. Points and curves are color coded from blue (weak autocorrelation) to red (strong autocorrelation). The mean front positions and standard deviations were calculated across 150 integrations of Eq. 3.1 for each value of resource autocorrelation length, with initial condition localized at the origin. The front propagation speed was computed via least-squares linear fits of the mean front position versus time. The standard deviation drops to zero when the landscape is fully colonized.

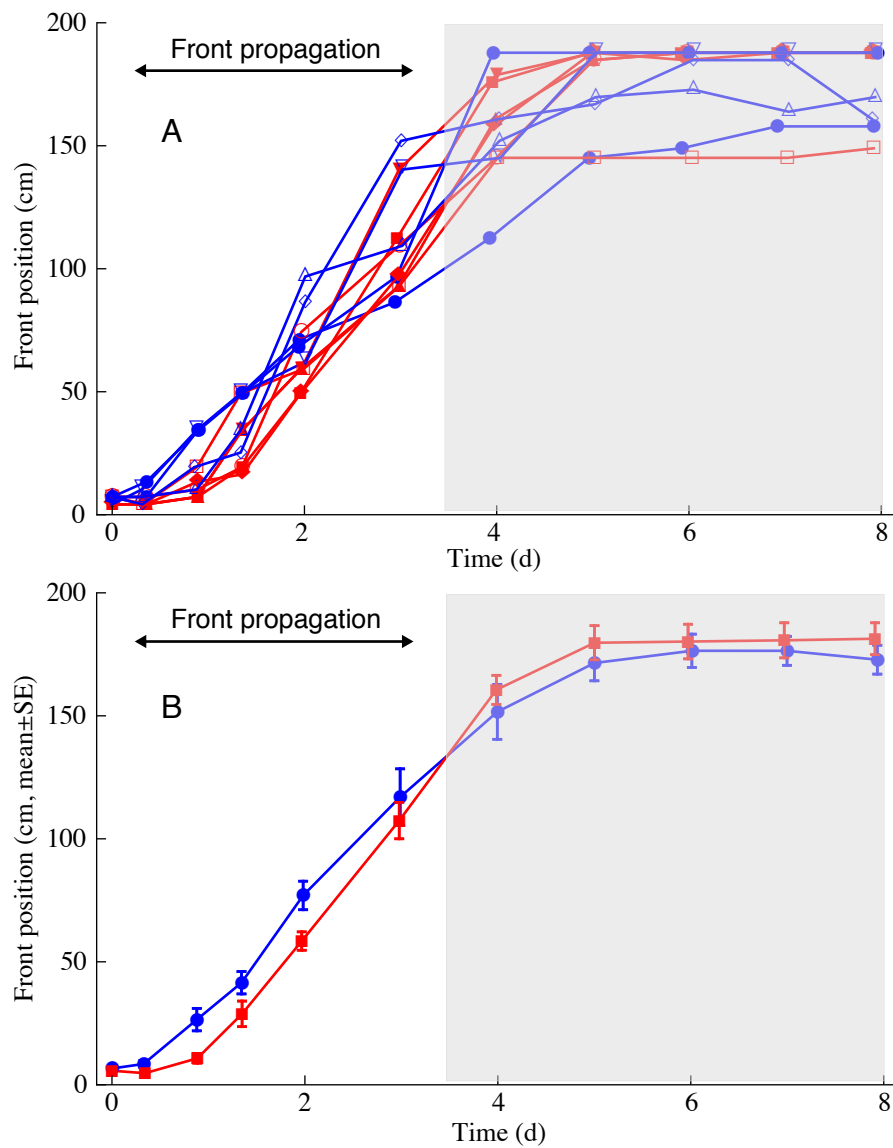
autocorrelation length was computed by fitting a straight line (least-squares fit) in the asymptotic propagation regime. The mean front position and the mean propagation speed, which is found to be a decreasing function of the resource autocorrelation length, are shown in Fig. 3.2A-B. Therefore, large resource autocorrelation lengths cause a slowing-down of the propagating front with respect to small autocorrelation lengths. This model prediction was tested in the microcosm experiment with *E. gracilis* discussed in the following section. Increasing the autocorrelation length, two landscapes with identical autocorrelation appear increasingly dissimilar (see, e.g., Fig. 3.3). Such increasing dissimilarity translates into an enhanced variability of the front position for larger values of the autocorrelation length. Specifically, plotting the standard deviation of the front position in 150 integrations of Eq. 3.1 shows that such standard deviation increases with the resource autocorrelation length (Fig. 3.2C).

In the experiment, two treatments were performed to study the effect of autocorrelation of the resource distribution on the propagation of biological invasion fronts. In the first treatment, LED sequences (Fig. 3.3, in blue) were set to reproduce resource distributions with small autocorrelation length ( $\approx 4$  cm, Methods). In the second treatment, LED sequences (Fig. 3.3, in red) were set to reproduce resource distributions with long autocorrelation length ( $\approx 12$  cm, Methods). The total number of switched-on LEDs was the same in both treatments. Within each treatment, different sequences of switched-on LEDs were chosen for each experimental replicate. Five replicates with small autocorrelation length and six replicates with large autocorrelation length were run. The LED sequences and the corresponding light intensity profiles along the linear landscapes are shown in Fig. 3.3. In all the figures, the color blue is used for environments with small autocorrelation length and red is used for those with large autocorrelation length. Linear landscapes identical to those employed in Chapter 1 and 2 (i.e., channels drilled in plexiglass and filled with a nutrient medium) were posed on top of the LED arrays. At the start of the experiment, an ensemble of *E. gracilis* was introduced at one end of the landscape and cell density profiles were then measured at various times by recording images with a stereomicroscope and counting cells via image analysis (Altermatt et al., 2014). The position of the propagating front at each measurement time was estimated by fixing a threshold value of cell density  $\bar{\rho}$  and recording the furthest point from the origin where the cell density was higher than such threshold value. The front position in each replicate at each measurement time, for a fixed value of the threshold, is shown in Fig. 3.4. Visual inspection of Fig. 3.4 suggests that strongly autocorrelated resource distributions (i.e., those with large autocorrelation length) cause a slowing-down of the propagating front compared to weakly autocorrelated ones. A mixed-effect model (Crawley, 2007) supports this statement for almost any choice of the threshold value used to trigger the front position in the experiment (Figs. 3.7, 3.8). The test statistics are reported in Table 3.1 for the

### Chapter 3. Invasion fronts: environmental heterogeneity



**Figure 3.3:** LED sequences (upper panels) and corresponding light intensity profiles (lower panels) used in the experiment. The blue color indicates environments with small autocorrelation length and the red color indicates environments with large autocorrelation length.



**Figure 3.4:** Front propagation in the experiment. Fronts propagating in environments with large resource autocorrelation length are colored in red, those propagating in environments with small resource correlation length are colored in blue. (A) Front position in all experimental replicates at different times. (B) Mean ( $\pm$ SE) position of the front across replicates with identical autocorrelation. In some replicates, the front reached the end of the landscape at day 4. Thus, the front propagation analysis was performed only with the data up to day 3 to avoid spurious effects due to the finite size of the system. The density threshold value used to trigger the front position was  $\bar{\rho} = 60 \text{ cm}^{-1}$ .

density threshold value  $\bar{\rho} = 60 \text{ cm}^{-1}$  and in Table 3.3 for all values of  $\bar{\rho}$  considered. The autocorrelation treatment was the single fixed effect and time/replicate was the random effect. The treatment with small autocorrelation length had 5 replicates, the treatment with large autocorrelation length had 6 replicates.

**Table 3.1:** Mixed-effect test statistics, with the autocorrelation treatment as single fixed effect and time/replicate as random effect. The treatment with small autocorrelation length had 5 replicates, the treatment with large autocorrelation length had 6 replicates. The front position was measured at the density threshold value  $\bar{\rho} = 60 \text{ cm}^{-1}$ .

	Value	Std. Error	df	<i>t</i> -value	<i>p</i> -value
Intercept	45.98	3.27	44	14.04	$p < 10^{-4}$
Treatment	-11.61	4.43	9	-2.62	0.0279

### 3.3.1 Total number of cells in the experiment

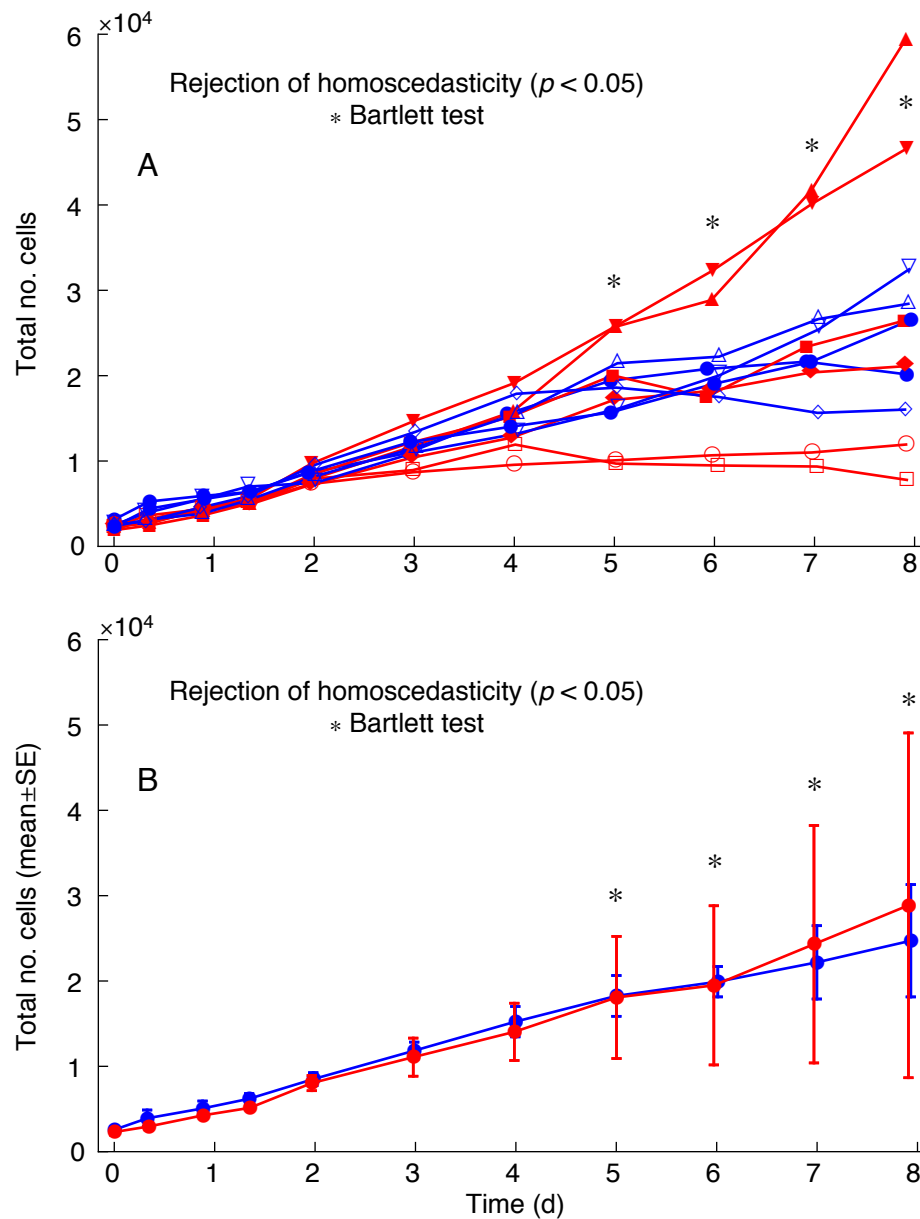
The experimental data suggest that the resource autocorrelation structure also affects the total number of individuals throughout the invasion. Although the mean number of individuals is the same in landscapes with large and small resource autocorrelation length (the mean is performed across all landscapes with identical autocorrelation), the variance in the total number of individuals across replicates with identical autocorrelation is much larger in autocorrelated landscapes, asymptotically in time (Fig. 3.5). In other words, replicated landscapes with identical resource autocorrelation length are much more alike with respect to the total number of individuals in the case of small rather than large resource autocorrelation lengths. The Bartlett test performed by grouping the total number of cells in two sets corresponding to the two treatments, separately for each measurement time, shows that the variances of the two sets were different from day 5 onwards (Fig. 3.5; the test statistics are reported in Table 3.2), that is, after the landscapes were fully colonized.

## 3.4 Methods

### 3.4.1 Model

The model Eq. 3.1 was integrated with parameters  $D = 0.08 \text{ cm}^2 \text{ min}^{-1}$ ,  $r_{\max} = 0.005 \text{ min}^{-1}$ ,  $K_I = 1 \text{ W m}^{-2}$ ,  $K = 300 \text{ cm}^{-1}$  and  $\sigma = 0.4 \text{ min}^{-1/2}$ . The light intensity profiles in the simulations were generated with the desired autocorrelation length as described in the text. The linear landscapes used in the simulations were 18 m long to avoid border





**Figure 3.5:** Total number of cells during the experimental invasions. Replicates with large resource autocorrelation length are colored in red, those with small resource autocorrelation length are colored in blue. The upper panel shows the total number of cells in each replicate, the lower panel shows the mean ( $\pm$ SD) total number of cells across replicates belonging to the same treatment. Visual inspection suggests that the variance in the total number of replicates is asymptotically (in time) larger in environments with large autocorrelation length than in those with small autocorrelation length. Accordingly, the Bartlett test rejects the equality of variances in the two treatments for the last four time points.

### Chapter 3. Invasion fronts: environmental heterogeneity

---

**Table 3.2:** Bartlett test statistics. The hypothesis of equal variances in the total number of cells between the two treatments (small and large autocorrelation lengths) is rejected from day 5 onwards.

Time (d)	$\chi^2(\text{df}=1)$	$p$ -value
0	0.026	0.872
0.3	3.044	0.081
0.9	3.578	0.059
1.3	3.447	0.063
2.0	0.106	0.745
3.0	2.279	0.131
4.0	1.499	0.221
5.0	3.950	0.047
6.0	7.639	0.006
7.0	4.438	0.035
8.0	4.104	0.043

effects, to allow a clear identification of the propagating invasion front and to allow the investigation of propagating fronts in environments with very large autocorrelation length, which could not be investigated experimentally because of the finite size of the experimental setup. The light intensity profile generated by a single LED (centered in  $x = 0$ ) was assumed equal to the best fit of Eq.  $I(x) = c_0/(c_1^2 + x^2)^2$  to the measured light intensity profile (Chapter 2).

Integration of Eq. 3.1 requires a fine spatial discretization to resolve the light intensity gradients. The spatial step was chosen equal to  $\Delta x = 0.6$  cm, while the temporal step was chosen equal to  $\Delta t = 0.5 \text{ min}^{-1}$ . The spatial discretization of Eq. 3.1 reads (Dornic et al., 2005):

$$\begin{aligned} \frac{d\rho_i}{dt}(t) = & \frac{D}{(\Delta x)^2} [\rho_{i+1}(t) + \rho_{i-1}(t) - 2\rho_i(t)] - \frac{1}{2\Delta x} [g_{i+1}\rho_{i+1}(t) - g_{i-1}\rho_{i-1}(t)] \\ & + r_i\rho_i(t) \left(1 - \frac{\rho_i(t)}{K}\right) + \frac{\sigma}{\sqrt{\Delta x}} \sqrt{\rho_i(t)}\eta(t), \end{aligned} \quad (3.2)$$

where  $i$  identifies the lattice site,  $g = \frac{d\phi}{dx}[I(x)]$  and the term  $\sqrt{\Delta x}$  ensures proper normalization in the continuum limit (Doering et al., 2005). For the numerical integration, Eq. 3.2 was divided by  $K$  and the equation for  $\rho/K$  was solved. The split-step method proposed in Dornic et al. (2005) (see also Chapter 1) was modified to solve Eq. 3.2. Eq. 3.2 contains an advection term that might cause an artificial loss of mass if the step sizes are too coarse. Such issue does not occur with the step sizes  $\Delta x$  and  $\Delta t$  chosen

here. Various alternatives to the spatial discretization Eq. 3.2 were investigated, including higher-order discretizations of the spatial derivatives and upwind/downwind schemes for the advection term. Their performances were compared by imposing  $r = 0$  and small  $\sigma$  and inspecting the conservation of mass in a spatial window of length  $L = 12$  cm with a light intensity profile corresponding to that generated by one LED located at the center of the spatial window. The discretization in Eq. 3.2 proved to be the most effective in conserving the total mass ( $\sum_i \rho_i$ ) in the system and is therefore employed here. The CFL condition for the diffusion equation  $2D\Delta t/\Delta x^2 < 1$  is satisfied and  $\Delta t/\Delta x < 1$ . Note, however, that except for the stability requirement due to the discretized Laplacian, the split-step method described in Dornic et al. (2005) was claimed to display no limitation on  $\Delta t$ .

#### 3.4.2 Experiment

The experiment was performed with the alga *E. gracilis*, purchased from Carolina Biological Supply (Burlington, North Carolina, USA). A culture of *E. gracilis* was initialized in a 500 mL Schott flask in a nutrient medium (Altermatt et al., 2014) two weeks prior to the start of the experiment and was kept at a temperature of 22 °C under constant blue light provided by the LEDs. The experiment was performed in linear landscapes identical to those used in Giometto et al. (2014) and were 1.9 m long. Linear arrays of LEDs identical to those used in Chapter 2 were used to generate the light intensity profiles along the landscapes. Switched-on LEDs were set to produce a peak light intensity of  $5.2 \text{ W m}^{-2}$  within the channel above the LEDs, so that only positive phototaxis occurred (see Chapter 2). The experimental replicates were kept in a climatized room at 22 °C for the whole duration of the experiment.

Five replicated landscapes with identical small autocorrelation length ( $\sim 4$  cm) but different switched-on LED sequences were generated as described in the Results section via the Markov Chain procedure with  $\lambda = 0.75$ . Analogously, six replicated landscapes with identical large autocorrelation length ( $\sim 12$  cm) but different switched-on LED sequences were generated via the Markov Chain procedure with  $\lambda = 0.25$ . All 11 landscapes had the same total number of switched-on LEDs and are shown in Fig. 3.3. The stated values of autocorrelation length are based on the first-order autocorrelation of the Markov Chain that generated the landscape (see the Results section). The first three LEDs in every landscape were switched-on to allow the local establishment of the inoculated *E. gracilis* population. Thus, the landscapes generated via the Markov Chain procedure described in the text started at the fourth LED. For the landscapes with large autocorrelation length, three landscapes were chosen so that the fourth LED was switched on and the other four were chosen so that the fourth LED was

switched off. In other words, the realized Markov Chain started from its stationary distribution.

In some replicates the propagating front reached the end of the landscape at day 4. Thus, the front propagation analysis was performed only with the data up to day 3 to avoid spurious effects due to the finite size of the system. The first time point was also excluded from the analysis because it was measured immediately after the inoculation of *E. gracilis* in the landscape and thus is identical for all replicates.

### 3.5 Discussion

This chapter studied the propagation of biological invasion fronts in landscapes with different heterogeneous distributions of resources. Focus was placed on the effect of the resource autocorrelation structure on the mean speed of the invasion. Both the model Eq. 3.1 and the experiment showed that the speed of biological invasions is a decreasing function of the resource autocorrelation length. Additionally, the experiments showed that the variance in the total number of cells when the landscape was fully colonized was much larger in the treatment with long resource autocorrelation length. Further theoretical investigation is required to understand the cause of such difference between the two treatments concerning fluctuations in the total number of individuals. Relating the low and high total number of individuals observed in some autocorrelated landscapes to specific features of the environment might translate into recommendations to minimize the impact of threatening invasive species in the field.

Numerical investigation of Eq. 3.1 allows to discern the effect of each process included in the equation on the slowing-down of the propagating front. In particular, simulating Eq. 3.1 with  $\phi = 0$  allows to isolate the net contribution of the biased movement towards favorable patches to the slowing down of the front. Fig. 3.6 shows that such biased movement contributes to the slowing down of the front, that is, the front propagates faster without such directed movement. This effect can be understood by considering an autocorrelated environment where a long stretch of favorable patches (light) is followed by a long stretch of unfavorable ones (dark). In such a situation, taxis acts as a spring at the border between the favorable and unfavorable regions, keeping the population in the former one. Different choices of the functional form of  $r(I)$  (for example  $r$  linear in  $I$ ) were investigated and do not change the picture significantly. Simulating Eq. 3.1 with very small  $\sigma$  allows to isolate the net contribution of demographic stochasticity. Integrating Eq. 3.1 with  $\sigma = 0.001 \text{ min}^{-1}$  shows that almost no slowing-down of the front (in strongly autocorrelated environments

compared to weakly autocorrelated ones) is observed for small values of demographic stochasticity (Fig. 3.6). No slowing-down would be observed in the complete absence of demographic stochasticity (i.e., with  $\sigma = 0$ )<sup>1</sup>. Therefore, the local extinctions caused by demographic stochasticity in unfavorable regions of the landscape are crucial to produce the slowing-down of the front that was observed in the experiment. Conversely, larger values of  $\sigma$  cause a more pronounced slowing-down of the invasion front in strongly autocorrelated environments compared to weakly autocorrelated ones (Fig. 3.6). Integration of Eq. 3.1 showed that the variability of the front position is larger for larger values of the resource autocorrelation length (Fig. 3.2). Differently from the intrinsic variability in the front position studied in Chapter 1, however, the variability observed here is due to the fact that two landscapes with identical resource autocorrelation length appear increasingly dissimilar for increasing values of the autocorrelation length. Therefore, such variability does not concern replicated invasions in a single heterogeneous landscape, but rather invasions in landscapes with different distribution of resources but identical autocorrelation length.

Among the few existing empirical studies of biological spread in heterogeneous environments (Hastings et al., 2005), two are worth mentioning for being conceptually related to our investigation, although not directly linked to the Fisher-Kolmogorov model or to its generalizations. The finding that larger autocorrelation lengths reduce the spread rate of invading species is compatible with the results of Bergelson et al. (1994). Bergelson et al. (1994) performed a field experiment with the invading weed *Senecio vulgaris* and found that the average spatial distance between two generations along linear transects increased when favorable patches were uniformly distributed in space (in the language of this chapter, the transect had small autocorrelation length), compared to transects with clumped patches (i.e., with large autocorrelation length). Bailey et al. (2000) performed spread experiment with the fungal plant pathogen *Rhizoctonia solani*. Bailey et al. (2000) provides a complementing view to the present investigation, by investigating the effect of the inter-distance between favorable patches on the spread, identifying the existence of a percolation threshold at a critical level of inter-patches distance. In the framework addressed here, the analogue of such percolation threshold corresponds to a typical autocorrelation length much larger than the typical distance traveled by the front in one generation. As opposed to Bailey et al. (2000), this work provides a general theoretical framework to interpret the dynamical processes acting behind the realized invasions and to link the species' traits to the front propagation (Giometto et al., 2014). The theoretical investigation of Eq. 3.1 allowed isolating the net effect of each process embedded therein and

---

<sup>1</sup>Here, it was chosen to show the mean propagating speed for  $\sigma = 0.001 \text{ min}^{-1}$  (as opposed to  $\sigma = 0 \text{ min}^{-1}$ ) in order to compare speeds calculated with the same integration scheme, which cannot be adopted to integrate the deterministic equation with  $\sigma = 0 \text{ min}^{-1}$ .

### Chapter 3. Invasion fronts: environmental heterogeneity

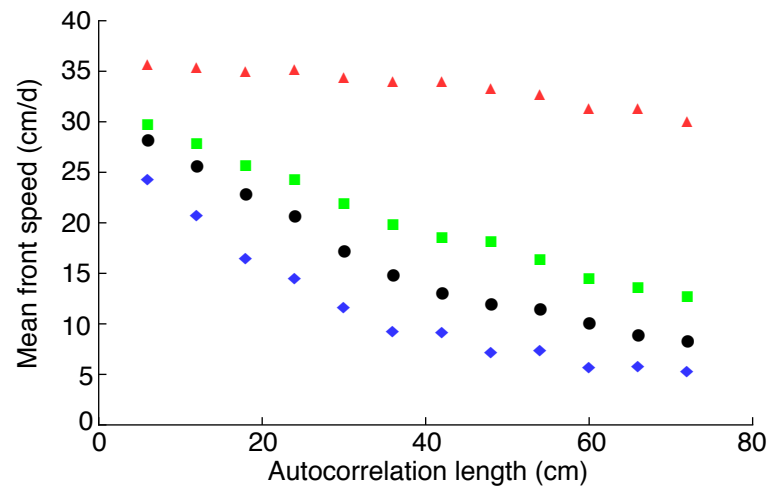
---

notably the key contribution of demographic stochasticity to the slowing-down of the propagating front. There are considerable differences in the experimental setup and the study system between this investigation and Bailey et al. (2000), Bergelson et al. (1994). Most importantly, biased active movement towards favorable patches was present in the experiment performed here and is embedded in Eq. 3.1, while passive dispersal occurred in Bergelson et al. (1994). Both Bergelson et al. (1994) and Bailey et al. (2000) differ from this study for the fact that the landscape and the distribution of resources here are continuous, while Bergelson et al. (1994) and Bailey et al. (2000) only considered discrete spatial distributions of favorable patches. Although a discrete distribution of favorable patches might be a good approximation in certain environments, a continuous distribution is deemed more likely to occur in nature. Because of the widespread application of the Fisher-Kolmogorov framework in the ecological literature (Elton, 1958, Grosholz, 1996, Lubina and Levin, 1988, Skellam, 1951), the generalization proposed and experimentally tested here is regarded as a significant step towards a more comprehensive understanding of species spread in heterogeneous environments.

The results derived in this chapter might have implications for the natural environment. The typical autocorrelation length of the resource distribution should be relatively easily inferred from environmental data (Turner, 2005) and thus appears as a concise indicator for the propagation success of the study species. Human impact on ecosystems might be seen, in the framework developed here, as affecting the typical autocorrelation length of favorable patches. Habitat fragmentation causes a reduction in the total habitat area, but is also argued to decrease the autocorrelation length of the landscape through the introduction of qualitatively different patches in the natural environment (Hanski, 1999, Holyoak et al., 2005). Such reduced autocorrelation length might translate into a landscape more prone to biological invasions.

### 3.6 Appendix

Eq. 3.1 was integrated imposing  $\phi = 0$  to investigate the contribution of phototaxis to the slowing-down of the front. Fig. 3.6 shows that the front propagation was larger in the absence of phototaxis (green squares), indicating that the directed movement of organisms towards favorable regions of the landscape contributes to the slowing-down of the front. Additionally, the mean speed of front propagation is shown to be a decreasing function of  $\sigma$  (Fig. 3.6).



**Figure 3.6:** Mean front propagation speed in landscapes with different resource autocorrelation length. The mean front speeds are calculated integrating Eq. 3.1 with different parameters for different colors and symbols. Each data point reports the mean front speed computed across 125 replicates with identical autocorrelation. Black dots report the mean front speed computed with parameters as in Fig. 3.2 ( $\sigma = 0.4 \text{ min}^{-1}$ ). Green squares report the mean front speed computed without phototaxis (i.e., imposing  $\phi = 0$ ). Blue diamonds report the mean front speed computed with a large value of  $\sigma$  (i.e.,  $\sigma = 0.7 \text{ min}^{-1}$ ). Red triangles report the mean front speed computed with a very small value of  $\sigma$  (i.e.,  $\sigma = 0.001 \text{ min}^{-1}$ ). The front propagation speed was computed via least-squares linear fits of the mean front position versus time. Fronts propagate faster in the absence than in the presence of phototaxis, that is, the directed movement of organisms towards favorable regions of the landscape contributes to the slowing-down of the front. Large values of  $\sigma$  cause increased slowing-down of the front. Conversely, small values of  $\sigma$  cause a reduced slowing-down of the front.

Figs. 3.7 and 3.8 show the front position in the experiment, computed with different density thresholds used to identify the front position. The  $p$ -value displayed refers to the hypothesis tested with the mixed-effect model (Results) that the autocorrelation treatment has an effect on the front propagation. The only threshold density value for which the null hypothesis is not rejected is the lowest value  $30 \text{ cells cm}^{-1}$ . Such threshold value, however, is most likely too low and introduces much noise in the estimation of the front position. As can be seen in the plot of the front position in

### Chapter 3. Invasion fronts: environmental heterogeneity

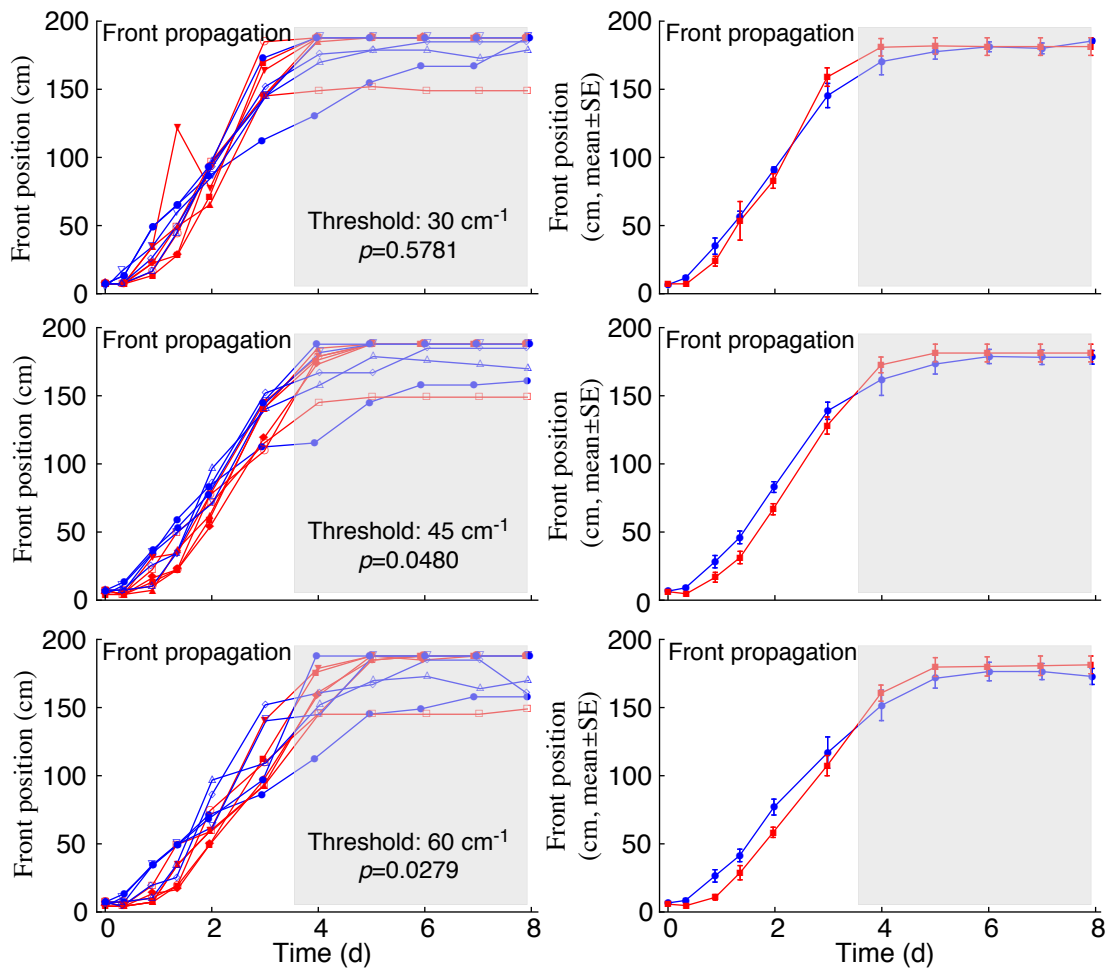
---

individual replicates, at such threshold value one replicate is observed to advance 90 cm from day 1 to day 1.5 and subsequently pull back 50 cm from day 1.5 to day 2. For all the other density threshold values, the effect of the autocorrelation treatment on the front propagation is significant.

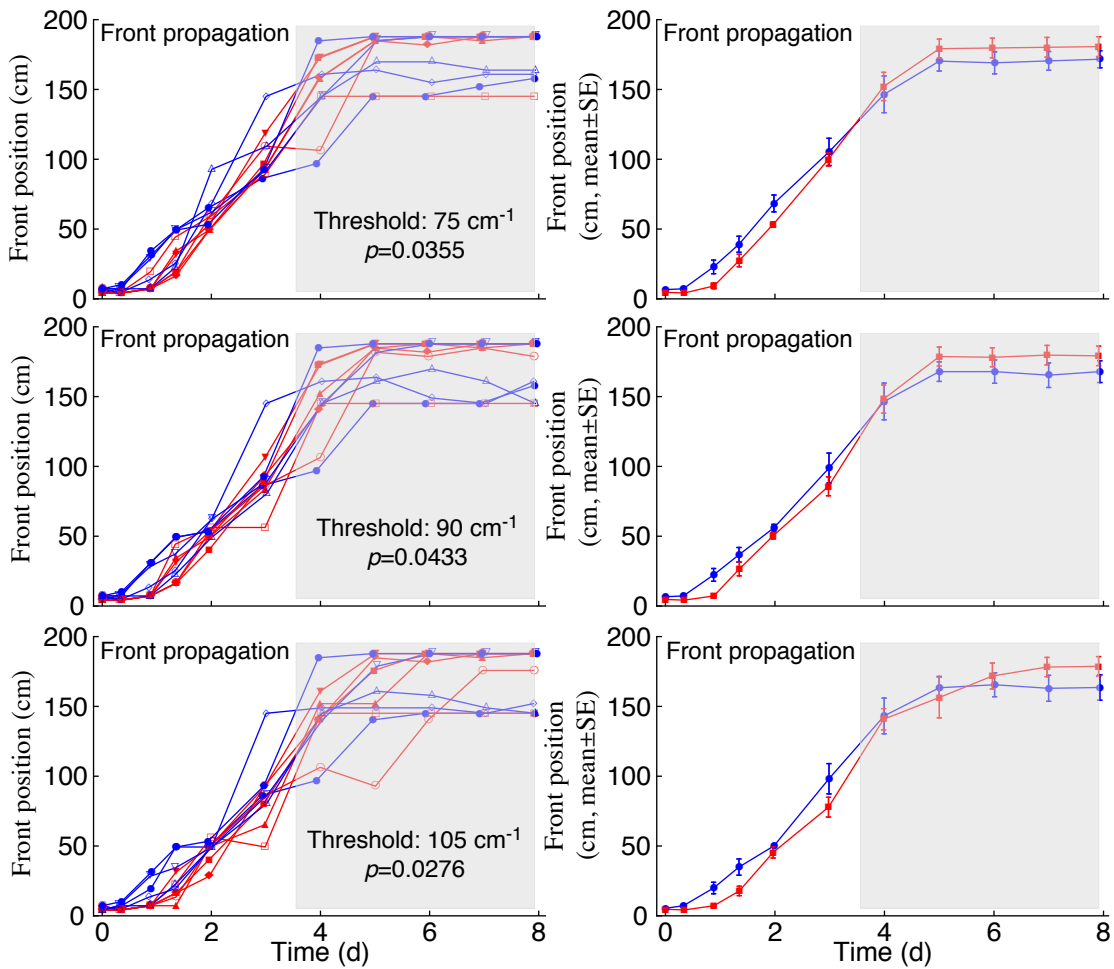
**Table 3.3:** Mixed-effect test statistics for all density threshold used to trigger the front position, with the autocorrelation treatment as single fixed effect and time/replicate as random effect. The treatment with small autocorrelation length had 5 replicates, the treatment with large autocorrelation length had 6 replicates.

Threshold $\bar{\rho}$		Value	Std. Error	df	<i>t</i> -value	<i>p</i> -value
30 cm <sup>-1</sup>	Intercept	61.77	4.33	44	14.27	$p < 10^{-4}$
	Treatment	-3.38	5.86	9	-0.58	$p = 0.5781$
45 cm <sup>-1</sup>	Intercept	57.15	3.65	44	15.65	$p < 10^{-4}$
	Treatment	-11.31	4.94	9	-2.29	$p = 0.0480$
60 cm <sup>-1</sup>	Intercept	45.98	3.27	44	14.04	$p < 10^{-4}$
	Treatment	-11.61	4.43	9	-2.62	$p = 0.0279$
75 cm <sup>-1</sup>	Intercept	45.27	2.88	44	15.70	$p < 10^{-4}$
	Treatment	-9.65	3.90	9	-2.47	$p = 0.0355$
90 cm <sup>-1</sup>	Intercept	36.65	2.84	44	12.91	$p < 10^{-4}$
	Treatment	-9.04	3.85	9	-2.35	$p = 0.0433$
105 cm <sup>-1</sup>	Intercept	35.91	3.04	44	11.83	$p < 10^{-4}$
	Treatment	-10.79	4.11	9	-2.62	$p = 0.0276$





**Figure 3.7:** Front propagation in the experiment. Different rows refer to different values of the density threshold used to trigger the front position. Fronts propagating in environments with large resource autocorrelation length are colored in red, those propagating in environments with small resource correlation length are colored in blue. The left panels show the position of the front in all replicates, while the right panels show the mean ( $\pm$ SE) position of the front across replicates with identical autocorrelation. The  $p$ -value refers to the hypothesis that the autocorrelation treatment has a significant effect on the front propagation.



**Figure 3.8:** Front propagation in the experiment. Different rows refer to different values of the density threshold used to trigger the front position. Fronts propagating in environments with large resource autocorrelation length are colored in red, those propagating in environments with small resource correlation length are colored in blue. The left panels show the position of the front in all replicates, while the right panels show the mean ( $\pm$ SE) position of the front across replicates with identical autocorrelation. The  $p$ -value refers to the hypothesis that the autocorrelation treatment has a significant effect on the front propagation.

## 4 Taylor's law of fluctuation scaling

### Abstract

Taylor's law (TL) states that the variance  $V$  of a non-negative random variable is a power function of its mean  $M$ , i.e.  $V = aM^b$ . The ubiquitous empirical verification of TL in ecology, physics and other natural sciences, typically displaying exponents  $b \simeq 2$ , suggests a context-independent mechanism. However, theoretical studies of population dynamics predict a broad range of values of  $b$ . Here, this apparent contradiction is rationalized by using large deviations theory to derive a generalized TL in terms of a population exponent  $b_{jk}$  for the scaling of the  $k$ -th vs the  $j$ -th cumulant (conventional TL is recovered for  $b = b_{12}$ ) in a broad class of population growth models. The sample exponents  $b_{jk}$  are found to depend predictably on the number of observed samples. For finite numbers of observations one observes sample exponents  $b_{jk} \simeq k/j$  (thus  $b \simeq 2$ ) asymptotically in time and independently of population exponents. Empirical analyses on two datasets support the theoretical results. In the broad settings investigated here, the sample TL exponent appears to be governed by sampling limitations rather than being determined by the underlying population growth process.

### 4.1 Introduction

Taylor's law (TL) (Taylor, 1961), also known as fluctuation scaling in physics, is one of the most verified patterns in both the biological and physical sciences (Eisler, 2008). TL states that the variance of a non-negative random variable  $V = \text{Var}[X]$ , is approximately related to its mean  $M = \mathbb{E}[X]$  by a power law, that is,  $\text{Var}[X] = a\mathbb{E}[X]^b$ , with  $a > 0$  and  $b \in \mathbb{R}$ . In biological context, TL was confirmed in bacterial populations (Ramsayer et al., 2012), forests (Cohen et al., 2013), gene distribution in human chromosomes (Fronczak and Fronczak, 2010) and other examples. In Chapter 5, Taylor's law will be shown to arise in intraspecific cell size distributions (Giometto et al., 2013). In the physics literature, this scaling pattern was found in cosmic radiation fluxes (Uttley and McHardy, 2001), heavy ion collisions (Botet et al., 2001), river flow (Dahlstedt and Jensen, 2005), traffic of Internet routers (Eisler and Kertész, 2005), traded value of stocks (Eisler and Kertész, 2006) and several other scenarios (Eisler, 2008). In ecology, the random variable of interest is generally the size or density  $N$  of a censused population and TL can arise in time (i.e., the statistics of  $N$  are computed over time) or in space (i.e., the statistics are computed over space).

The widespread verification of TL has led many authors to suggest the existence of a universal mechanism for its emergence, although there is currently no consensus on what such mechanism would be. Various approaches have been employed in the attempt of identifying such mechanism, ranging from the study of probability distributions compatible with the law (Jørgensen, 1987, Kendal and Jørgensen, 2011, Tweedie, 1946) to phenomenological and mechanistic models (Hanski, 1982, Kalyuzhny et al., 2014, Keeling, 2000, Kilpatrick and Ives, 2003). Although most empirical studies on spatial TL report an observed sample exponent  $b$  in the range 1–2 (Anderson et al., 1982, Taylor, 1961), mostly around  $b \simeq 2$  (Anderson et al., 1982) (see also Fig. 4.10(g) in Taylor and Woiod (1982)), population growth models (Cohen, 2013, 2014a,b, Cohen et al., 2013, Jiang et al., 2014) can generate TL with any real value of the exponent. Moreover, theoretical investigations of multiplicative growth models in correlated Markovian environments (Cohen, 2014a,b) have shown that the exponent  $b$  can undergo abrupt transitions following smooth changes in the environmental autocorrelation.

Here, a distinction is made between values of  $b$  derived from empirical fitting (sample exponents) and values obtained via theoretical models that pertain to the probability distribution of the random variable  $N$  (population exponents). It is shown that in a broad class of multiplicative growth models, the sample and population exponents coincide only if the number of observed samples or replicates is greater than an exponential function of the duration of observation. Among the relevant consequences, the sample TL exponent is demonstrated to settle on  $b \simeq 2$  for almost any Markovian

environment observed for a duration that is larger than a logarithmic function of the number of replicates. Accordingly, abrupt transitions in the sample TL exponent can only be observed within relatively short time windows when the number of observations is limited.

This chapter is organized as follows. First, the multiplicative growth model is described in its simplest formulation and the result on the scaling of the sample and population exponents is derived. Second, the generalized Taylor's law for the scaling of any pair of moments is derived. Third, generalizations of the growth model are discussed and the same results are proved in such settings. Fourth, the prediction of the generalized Taylor's law for the sample exponents is tested on two datasets. The Appendix provides additional details on the multiplicative growth model and the derivation of the generalized Taylor's law in other growth models.

## 4.2 Results

This investigation starts from the multiplicative growth model in a Markovian environment introduced in Cohen (2014a,b), which includes as a special case (in the absence of autocorrelation) the Lewontin-Cohen model (Cohen et al., 2013, Lewontin and Cohen, 1969). Let  $N(t)$  be the density of a population at time  $t$  and assume that the initial density is  $N_0 > 0$ .  $N(t)$  is assumed to undergo a multiplicative growth process such that:

$$N(t) = N_0 \prod_{n=1}^t A_n. \quad (4.1)$$

The values of the multiplicative growth factors  $A_i$  are determined via a finite-state homogeneous Markov chain with state space  $\chi = \{r, s\}$  (the state space is labeled  $\chi = \{1 \leftrightarrow r, 2 \leftrightarrow s\}$  and, without loss of generality, assume  $r > s$  and  $N_0 = 1$ ) and transition matrix  $\Pi$ , with  $\Pi(i, j) > 0$  for all  $i, j \in \chi$ . In this notation,  $\Pi(i, j)$  is the one-step probability to go from state  $i$  to state  $j$ , i.e.,  $\Pi(i, j) = \text{Prob}(A_{n+1} = j | A_n = i)$ . For the sake of clarity this initial investigation is restricted to symmetric transition matrices, with  $\Pi(i, j) = \lambda$  for  $i \neq j$ , but all results hold with minor changes also for non-symmetric matrices and in the case of a more general state space, as discussed in the subsequent paragraphs. The stationary distribution  $\pi$  of the chain is unique and in the symmetric case satisfies  $\pi(i) = 1/2$ ,  $i \in \chi$ , for all  $\lambda \in (0, 1)$ . It is assumed that the chain starts at equilibrium.

Under the broad assumptions just stated, for any choice of  $\Pi$  and  $\chi$ , the sample mean and variance in a finite set of  $R$  independent realizations of the process (in an ecologi-

## Chapter 4. Taylor's law of fluctuation scaling

---

cal example, e.g.,  $R$  sufficiently separated regions in space) obey TL asymptotically with exponent  $b \simeq 2$ , even when the population moments of the distribution of  $N(t)$  satisfy TL with exponent  $b \neq 2$ . To distinguish between the two exponents, the exponent of TL calculated with the sample mean and variance will be referred to as the sample exponent and the exponent calculated via the population mean and variance will be referred to as the population exponent. Correspondingly, a distinction is placed between the sample and the population TL.

The empirical mean  $L_t(z) : \chi \rightarrow [0, 1]$  is defined as:

$$L_t(z) = \frac{1}{t} \sum_{n=1}^t \delta_{A_n, z}, \quad (4.2)$$

where  $\delta$  is the Kronecker's delta. The random measure  $L_t(r)$  gives the fraction of times that  $r$  appears in a realization of the Markov chain up to time  $t$ .  $L_t$  satisfies a Large Deviation Principle (LDP) (den Hollander, 2008) with rate function:

$$I_{\Pi}(x) = \sup_{u > 0} \left[ x \log \left( \frac{u_1}{(\Pi u)_1} \right) + (1-x) \log \left( \frac{u_2}{(\Pi u)_2} \right) \right] \quad (4.3)$$

where  $x$  ( $x \in [0, 1]$ ) is the proportion of  $r$  in a realization of the Markov chain up to time  $t$  (correspondingly, the proportion of  $s$  is  $1-x$ ) and  $u$  is a strictly positive vector in  $\mathbb{R}^2$  (i.e.,  $u_1, u_2 > 0$ ). Stating that  $L_t$  satisfies a LDP means that  $\lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{P}(L_t(r) \in [x, x+dx]) = -I_{\Pi}(x)$ . The rate function  $I_{\Pi}(x)$  is convex ( $d^2 I_{\Pi}/dx^2 > 0$ ), attains its minimum at  $x_{\min} = 1/2$  with  $I_{\Pi}(x_{\min}) = 0$  and is symmetric around  $x_{\min}$  (Lemma IV.10 of den Hollander (2008), Theorems 3.1.2, 3.1.6 of Dembo and Zeitouni (2009), Section 4.3 of Touchette (2009)). The subscript  $\Pi$  is used to indicate that the rate function depends on the transition matrix. Additionally, Eq. 4.3 depends on  $u_1$  and  $u_2$  only through  $u \equiv u_2/u_1$ ; thus, by standard one-variable calculus, a long but explicit form of  $I_{\Pi}(x)$  can be computed:

$$I_{\Pi}(x) = (x-1) \log \left[ 1 - \lambda \left( \frac{2(\lambda-1)x}{\lambda + \sqrt{\lambda^2 + 8\lambda(x-1)x - 4(x-1)x - 2\lambda x}} + 1 \right) \right] - \\ - x \log \left[ 1 - \frac{\lambda \left( \lambda + \sqrt{\lambda^2 + 8\lambda(x-1)x - 4(x-1)x - 2\lambda x} \right)}{2(\lambda-1)x} \right]. \quad (4.4)$$

The rate function does not depend on the values of the multiplicative factors  $r$  and  $s$ . As in Cohen (2014a), the ratio between  $t^{-1} \log \text{Var}[N(t)]$  and  $t^{-1} \log \mathbb{E}[N(t)]$  is considered, but here the LDP, adopting Varadhan's lemma (Theorem III.13, den Hollander (2008)), is exploited to perform such computation. First, since  $\Pi$  is positive and  $r \neq s$ ,

it holds true that:

$$\lim_{t \rightarrow \infty} t^{-1} \log \text{Var}[N(t)] = \lim_{t \rightarrow \infty} t^{-1} \log \mathbb{E}[N(t)^2], \quad (4.5)$$

see the Appendix in Cohen (2014a) for a proof. Then, for the population moments of the population density  $N(t)$ , applying Varadhan's lemma:

$$\lim_{t \rightarrow \infty} t^{-1} \log \mathbb{E}[N(t)^k] = \sup_{x \in [0,1]} [kG(x) - I_{\Pi}(x)], \quad (4.6)$$

where  $G(x) = x \log r + (1-x) \log s$ . The population TL exponent  $b$  (which depends on  $\lambda$ ) can thus be computed as:

$$b(\lambda) = \frac{\sup_{x \in [0,1]} [2G(x) - I_{\Pi}(x)]}{\sup_{x \in [0,1]} [G(x) - I_{\Pi}(x)]}. \quad (4.7)$$

For certain values of  $r$  and  $s$ ,  $b(\lambda)$  can show a discontinuity at a critical value of the transition probability  $\lambda$  (black line in Fig. 4.1A, see also Fig. 4.8). The existence of such discontinuity was discovered and discussed in Cohen (2014b). An analysis of the critical transition probability is also available in the Appendix. A generalized TL can be derived by adapting Eq. 4.6 to compute the scaling exponent for any pair of population moments as:

$$b_{jk}(\lambda) = \frac{\lim_{t \rightarrow \infty} t^{-1} \log \mathbb{E}[N(t)^k]}{\lim_{t \rightarrow \infty} t^{-1} \log \mathbb{E}[N(t)^j]} = \frac{\sup_{x \in [0,1]} [kG(x) - I_{\Pi}(x)]}{\sup_{x \in [0,1]} [jG(x) - I_{\Pi}(x)]}. \quad (4.8)$$

Discontinuities can also arise for these population exponents (Fig. 4.1). In the following,  $b$  refers to the conventional TL population exponent ( $b_R$  for the conventional TL sample exponent), while the generalized TL exponents are indicated with  $b_{jk}$  (the distinction between sample and population exponents will be clear from the context).

Eqs. 4.7, 4.8 hold true when one considers an infinite number of realizations of the multiplicative process, which ensures visiting the whole region  $x \in [0, 1]$ . Here, the sample exponent  $b_R$  that is based on the sample mean and variance calculated over a finite set of  $R$  realizations of the multiplicative process is estimated. First, a heuristic derivation of the sample exponent is presented. A more rigorous calculation of  $b_R$  is given in the subsequent paragraph. Let  $x_+$  be the value in  $[0, 1]$  such that the probability of a larger frequency  $x$  of  $r$  in  $R$  runs of the Markov chain up to time  $t$  is  $1/R$  (Redner, 1990):

$$\mathbb{P}[L_t(r) \in (x_+, 1]] = \frac{1}{R}. \quad (4.9)$$

## Chapter 4. Taylor's law of fluctuation scaling

---

With this definition,  $x_+$  can be interpreted (Redner, 1990) as the typical maximum frequency of  $r$  in  $R$  realizations of the chain. Analogously, let  $x_-$  be the value such that smaller values of the frequency of  $r$  are observed with probability  $1/R$ , namely  $\mathbb{P}[L_t(r) \in [0, x_-]] = 1/R$ . For large  $t$ , one can adopt Varadhan's lemma (or Laplace's method of integration) to obtain, as a function of  $t$ , the approximate number of replicas  $R$  needed to explore rare events (i.e., to compute  $\mathbb{P}[L_t(r) \in (x_+, 1)] = R^{-1}$ ). Approximately:

$$R \simeq \exp[t I_{\Pi}(x_{\pm})]. \quad (4.10)$$

Inversion of this formula (by taking the logarithm on both sides and expanding  $I_{\Pi}$  in Taylor series around  $x = x_{\min}$ ) gives  $x_{\pm} \simeq \frac{1}{2} \pm \sqrt{\frac{1-\lambda}{2\lambda} \frac{\log R}{t}}$ . Consequently, the sample TL exponent in an ensemble of  $R$  realizations of the process can be approximated as:

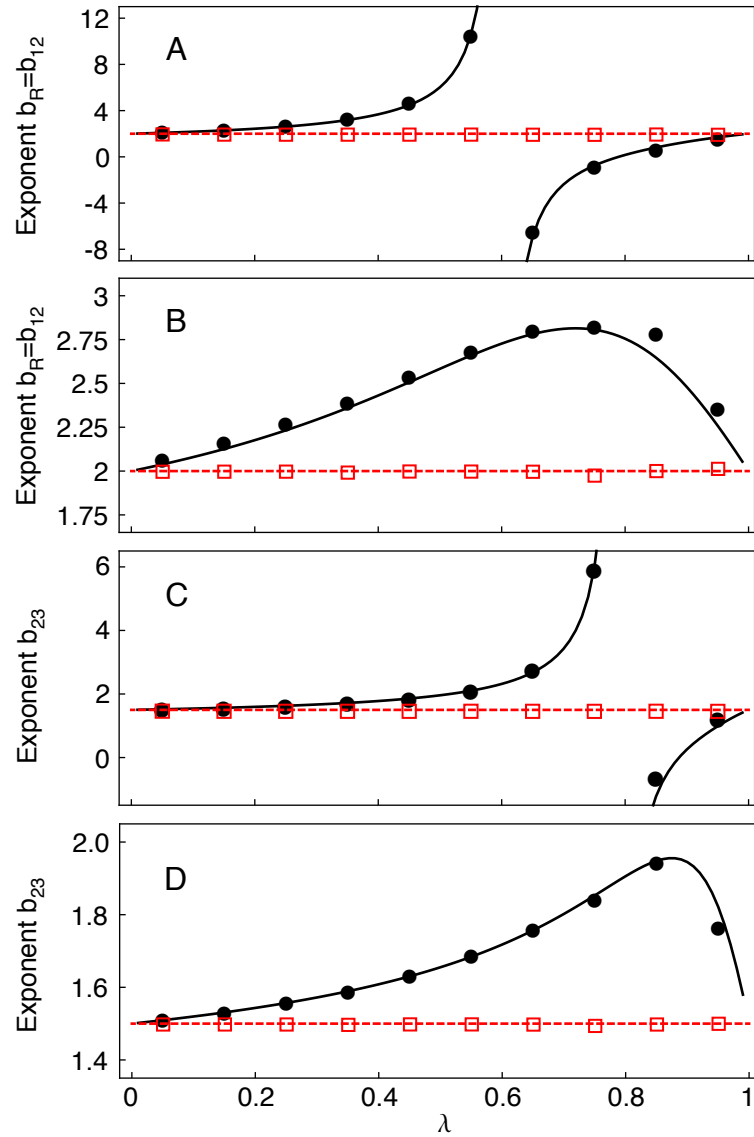
$$b_R(\lambda, t) \simeq \frac{\sup_{x \in [x_-, x_+]} [2G(x) - I_{\Pi}(x)]}{\sup_{x \in [x_-, x_+]} [G(x) - I_{\Pi}(x)]}, \quad (4.11)$$

where the dependence on  $t$  is through  $x_+$  and  $x_-$ . The zero of the rate function,  $x_{\min} = 1/2$ , corresponds to the most probable value of the product in Eq. 4.1. Because  $x_{\pm} \simeq \frac{1}{2} \pm \sqrt{\frac{1-\lambda}{2\lambda} \frac{\log R}{t}}$ , for fixed  $R$  the suprema in Eq. 4.11 are computed over an increasingly narrower set around  $x_{\min}$  (with  $I_{\Pi}(x_{\min}) = 0$ ) as  $t$  increases. Fig. 4.2 exemplifies such computation. Thus, for any finite number of realizations  $R$ , the sample exponent will approximate  $\lim_{t \rightarrow \infty} b_R(\lambda, t) \simeq 2$  after a time  $t^*$  that increases only logarithmically with  $R$  (Eq. 4.10 and Fig. 4.3), for any choice of  $\lambda$ ,  $r$  and  $s$ . For example, with  $\lambda = 0.5$ , when  $t = 100$ , in order to access to the extreme event  $x_+ = 0.9$  (and  $x_- = 0.1$ ) one needs about  $R \simeq 10^{16}$  replicates of the process. Analogous considerations hold for the asymptotic sample exponent describing the scaling of the sample moments  $\mathbb{E}[N(t)^k]$  with  $\mathbb{E}[N(t)^j]$ , which can be approximated as:

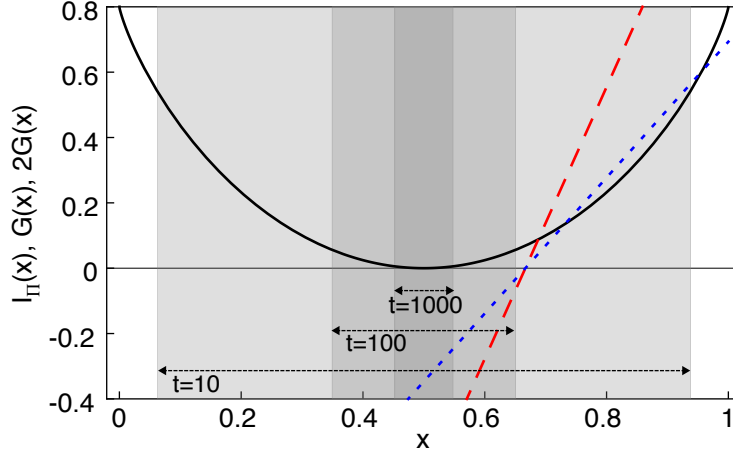
$$b_{jk}(\lambda, t) \simeq \frac{\sup_{x \in [x_-, x_+]} [kG(x) - I_{\Pi}(x)]}{\sup_{x \in [x_-, x_+]} [jG(x) - I_{\Pi}(x)]}, \quad (4.12)$$

which is the analogue of Eq. 4.11 for any pair of sample moments. Fig. 4.1 illustrates typical behaviors of sample and population exponents as a function of the transition probability  $\lambda$  for the 2-state multiplicative model with symmetric transition matrix. The black and red lines portray respectively the predicted asymptotic population and sample exponent (Eqs. 4.7 and 4.11), computed for different values of  $\chi = \{r, s\}$  in the two panels. Dots and squares illustrate the sample exponents  $b_R$  and  $b_{23}$  calculated via simulations in the regimes  $t \ll \log R$  and  $t \gg \log R$ , respectively. Simulations results in the two regimes (dots and squares) and theoretical predictions (solid and dashed lines) show excellent agreement.





**Figure 4.1:** TL exponent  $b_R = b_{12}$  and generalized exponent  $b_{23}$  for different values of the transition probability  $\lambda$ . The sample exponents (dots and squares) computed in simulations of a 2-state multiplicative process with symmetric transition matrix in the two regimes  $1 \ll t \ll \log R$  (black filled dots,  $R = 10^6$  up to time  $t = 10$ ) and  $t \gg \log R$  (red open squares,  $R = 10^4$  up to time  $t = 400$ ) are in good agreement with predictions for the asymptotic population (black solid line, Eq. 4.7) and sample (red dashed line,  $b_R = b_{12} = 2$  and  $b_{23} = 3/2$ ) exponents. In the simulations, the sample exponent  $b_R = b_{12}$  was computed by least-squares fitting of  $\log \text{Var}[N(t)]$  as a function of  $\log \mathbb{E}[N(t)]$  for the last 6 (black dots) and 200 (red squares) time steps. The sample exponent  $b_{23}$  was computed by least-squares fitting of  $\log \mathbb{E}[N(t)^3]$  as a function of  $\log \mathbb{E}[N(t)^2]$  in the same fashion. In panel (A), which reproduces Cohen (2014b), and (C)  $\chi = \{r, s\} = \{2, 1/4\}$  (the population exponents  $b = b_{12}$  and  $b_{23}$  display discontinuities); in panel (B) and (D)  $\chi = \{r, s\} = \{4, 1/2\}$  (in such a case, the population exponents  $b_{12}$  and  $b_{23}$  display no discontinuities).



**Figure 4.2:** Plot of  $I_{\Pi}(x)$  (black curve),  $G(x)$  (dotted blue line) and  $2G(x)$  (dashed red line). Marked in gray are the regions  $[x_-, x_+]$  at times  $t = 10, 100$  and  $1000$  (from light to dark gray) for fixed  $R = 100$ . These gray regions are the intervals over which the supremum in Eq. 4.11 is computed. In this example,  $r = 2, s = 1/4, \lambda = 0.55$ . The quantities  $x_+$  and  $x_-$  are computed by solving numerically Eq. 4.10.

The above calculations identify the logarithmic dependence of  $x_+$  on the number of realizations  $R$ , but rely on a number of approximations: the definition of  $x_+$  (which, in a given realization, is a random variable), the computation of Laplace integrals (Eq. 4.9) and the expansion of the rate function around  $x_{\min}$  (Eq. 4.10). Such calculations can be made more rigorous if one considers the independent identically distributed random variables  $X^i(t) = L_t^i(r)$ , that is,  $X^i(t)$  is the frequency of occurrence of the first state up to time  $t$  in the  $i$ -th realization of the Markov chain ( $i = 1, \dots, R$ ). Let  $x^+ = \max\{X^1(t), \dots, X^R(t)\}$ ; it holds true that:

$$\frac{1}{t} \log \mathbb{P}(X^1(t) > x) \leq \frac{1}{t} \log \mathbb{P}(x^+ > x) \leq \frac{1}{t} \log(R) + \frac{1}{t} \log \mathbb{P}(X^1(t) > x). \quad (4.13)$$

For fixed  $R$  (or, more generally,  $\log R = o(t)$ ) and  $x > 1/2$ , taking the limit ( $\lim_{t \rightarrow \infty}$ ) in Eq. 4.13 and knowing that  $L^t(r)$  satisfies a LDP, one has:

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{P}(x^+ > x) = \sup_{y \in (x, 1]} -I_{\Pi}(y) = -I_{\Pi}(x). \quad (4.14)$$

Because  $0 < I_{\Pi}(x) \leq \infty$ , Eq. 4.14 implies that  $\lim_{t \rightarrow \infty} \mathbb{P}(x^+ > x) = 0$  for any  $x > 1/2$ . An analogous calculation for  $x^- = \min\{X^1(t), \dots, X^R(t)\}$  shows that  $\lim_{t \rightarrow \infty} \mathbb{P}(x^- < x) = 0$  for any  $x < 1/2$ . In this context, one can approximate the sample exponent at time  $t$

with an analogue of Eq. 4.11:

$$b_R(\lambda, t) \simeq \frac{\sup_{x \in [x^-, x^+]} [2G(x) - I_\Pi(x)]}{\sup_{x \in [x^-, x^+]} [G(x) - I_\Pi(x)]}. \quad (4.15)$$

In the narrow interval  $[x^-, x^+]$  centered around  $x_{\min}$ ,  $I_\Pi(x) \simeq 0$  and as a consequence  $b_R(\lambda, t) \simeq 2$  (Fig. 4.2). More precisely,  $|b_R(\lambda, t) - 2|$  goes to 0 in probability as  $t$  tends to infinity. In fact, for every  $\epsilon > 0$ :

$$\mathbb{P}(|b_R(\lambda, t) - 2| > \epsilon) \leq \mathbb{P}\left(x^+ > \frac{1}{2} + \eta(\epsilon)\right) + \mathbb{P}\left(x^- < \frac{1}{2} - \eta(\epsilon)\right), \quad (4.16)$$

where  $\eta(\epsilon)$  is a function that goes to zero for  $\epsilon \rightarrow 0$ . Because of Eqs. 4.13 and 4.14, it follows that:

$$\lim_{t \rightarrow \infty} \mathbb{P}(|b_R(t) - 2| > \epsilon) = 0. \quad (4.17)$$

Analogous considerations hold for the generalized TL describing the scaling of any pair of moments.

Eqs. 4.11 and 4.12 give the estimated sample exponent of TL asymptotically, ignoring the constant term in the scaling of the variance  $V$  versus the mean  $M$  as  $\log V = b \log M + \log a$ . For small  $t$ ,  $\log a$  can be of the same order of magnitude of  $\log V$ . Fig. 4.3 shows the crossover of the sample exponent (for fixed  $R$ ,  $\lambda$ ,  $r$  and  $s$ ) from the population exponent  $b = b(\lambda)$  as in Eq. 4.7 (observed when  $t \ll \log R$ ) to  $b \simeq 2$  (when  $t \gg \log R$ ), where the sample exponent is calculated as the slope of the curve  $\log \mathbb{E}[N(t)^2]$  versus  $\log \mathbb{E}[N(t)]$  at time  $t$  (thus not neglecting the constant term  $\log a$ ). The sample moments are computed as:

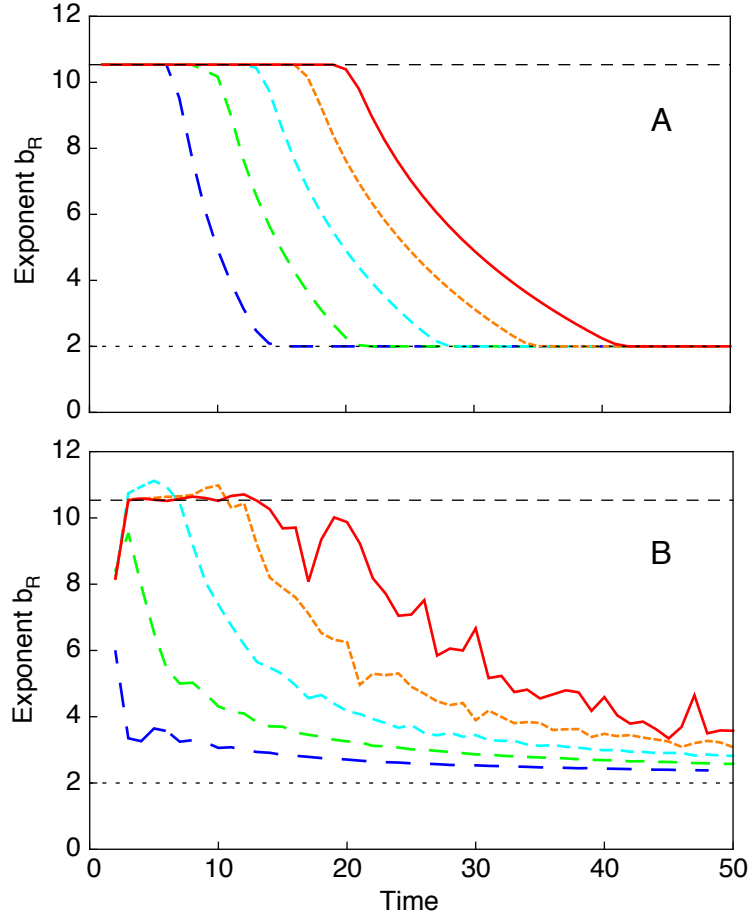
$$t^{-1} \log \mathbb{E}[N(t)^k] \simeq \sup_{x \in [x^-, x^+]} [kG(x) - I_\Pi(x)] \quad (4.18)$$

(cf. Eq. 4.6) in panel A and as the sample moments in simulations in panel B.

Some generalizations of the stochastic multiplicative process considered above are now investigated. The sample exponent in a finite set of  $R$  independent realizations of the process is  $b \simeq 2$  also for non-symmetric transition matrices  $\Pi$ . In the asymmetric case, the transition matrix is:

$$\Pi = \begin{pmatrix} 1 - \lambda & \lambda \\ \mu & 1 - \mu \end{pmatrix}, \quad (4.19)$$

with  $0 < \lambda, \mu < 1$ . The rate function  $I_\Pi(x)$  is convex, attains its minimum at  $x_{\min} =$



**Figure 4.3:** Time evolution of the sample TL exponent  $b_R$ . The sample exponent (computed as the slope of the curve  $\log \mathbb{E}[N(t)^2]$  versus  $\log \mathbb{E}[N(t)]$ ) crosses over from the approximate population exponent (Eq. 4.7, dashed upper horizontal line) at small times to  $b \simeq 2$  (dotted lower horizontal line) at larger times. The number of replicates  $R = 10^n$  increases exponentially from  $10^2$  (blue dashed lines) to  $10^6$  (red solid lines), while the crossover time increases approximately linearly. Here,  $\chi = \{r, s\} = \{2, 1/4\}$  and the transition probability in the symmetric  $\Pi$  is  $\lambda = 0.55$ . Panel (A) shows the theoretical prediction via Eq. 4.18. Panel (B) shows simulation results and the curves are averaged over  $10^8/R$  simulations (apart for the blue curve, which was averaged over  $10^5$  simulations). Mismatches between panel (A) and (B) are due to the necessity to have  $t$  and  $R$  not too large to keep simulations feasible, while Eqs. 4.6, 4.7 and 4.11 hold true asymptotically in  $t$ .

$\pi(1) = \mu/(\lambda + \mu)$ , where  $\pi = (\pi(1), \pi(2) = \lambda/(\lambda + \mu))$  is the invariant measure for  $\Pi$  and  $I_\Pi(x_{\min}) = 0$ . Only the value of the rate function at  $x_{\min}$  and not the value of  $x_{\min}$  is relevant for the argument made here. Due to asymmetries of  $I_\Pi$ , ‘left’ (i.e.,  $x < x_-$ ) rare events could be easier to see than ‘right’ (i.e.,  $x > x_+$ ) rare events or vice versa. In all cases, an exponentially large in  $t$  number of replicates is needed to sample the tails with the correct weights. In this context, Eqs. 4.7, 4.8 and Eqs. 4.11, 4.12, 4.23 are still valid and give, respectively, the asymptotic population and sample exponents.

The previous considerations can also be extended to multiplicative processes  $N(t)$  in more general Markovian environments with  $w$  states and state space  $\chi = \{r_1, \dots, r_w\}$ , where all  $r_i$  are strictly positive and at least two  $r_i$  are different. The state space is labeled  $\chi = \{1 \leftrightarrow r_1, \dots, w \leftrightarrow r_w\}$ . Let the transition matrix  $\Pi$  be two-fold irreducible (i.e.,  $\Pi$  irreducible and  $\Pi \Pi^\top$  irreducible, where  $\Pi^\top$  is the transpose of  $\Pi$ ). The rate function in Eq. 4.3 reads (Theorem IV.7 and Section IV.3 of den Hollander (2008), or Theorem 3.1.6 of Dembo and Zeitouni (2009)):

$$I_\Pi(\mu) = \sup_{u>0} \left[ \sum_{v=1}^w \mu_v \log \frac{u_v}{(\Pi u)_v} \right], \quad (4.20)$$

where  $u$  is a strictly positive vector in  $\mathbb{R}^w$ . Here,  $\sum_{v=1}^w \mu_v = 1$ , and  $\mu_v$  represents the proportion of  $v$  after  $t$  steps (for large  $t$ ). The rate function is convex and  $I_\Pi(\mu_{\min}) = 0$ , with  $\mu_{\min}$  the most probable state for large  $t$  (Theorems 3.1.2, 3.1.6 of Dembo and Zeitouni (2009), Section 4.3 of Touchette (2009)). Eq. 4.7, with  $x$  in the standard  $w - 1$  simplex in  $\mathbb{R}^w$  and  $G(x) = \sum_{i=1}^w x_i \log r_i$ , gives the population scaling exponent of  $\mathbb{E}[N(t)^2]$  with  $\mathbb{E}[N(t)]$ . The two-fold irreducibility of  $\Pi$  plus the condition that  $r_i \neq r_j$  for some  $i \neq j$  is the sharpest sufficient assumption that is presently known (Cohen, 2014a) to guarantee that the limiting growth rate of the second moment equals the limiting growth rate of the variance; thus, Eq. 4.7, with  $x$  in the standard  $w - 1$  simplex in  $\mathbb{R}^w$  and  $G(x) = \sum_{i=1}^w x_i \log r_i$ , gives the population scaling exponent of  $\text{Var}[N(t)]$  with  $\mathbb{E}[N(t)]$ . Analogously, Eq. 4.8, with  $x$  in the standard  $w - 1$  simplex in  $\mathbb{R}^w$  and  $G(x) = \sum_{i=1}^w x_i \log r_i$ , gives the population scaling exponent of  $\mathbb{E}[N(t)^k]$  with  $\mathbb{E}[N(t)^j]$ . As far as the scaling of moments is of interest, the ergodicity (i.e., irreducibility and aperiodicity) of  $\Pi$  (as opposed to the two-fold irreducibility) and  $G(x)$  not identically equal to zero (which happens only if  $r_i = 1 \forall i$ ) are sufficient to compute the scaling exponents via Eqs. 4.7, 4.8, modified as stated above. This is true because the ergodicity of  $\Pi$  ensures that the empirical measure  $L_t$  satisfies a LDP (Theorems 3.1.2 and 3.1.6 of Dembo and Zeitouni (2009)). Therefore, one can apply Varadhan’s lemma (Theorem III.13, den Hollander (2008)) to compute the limiting growth rate of the moments of  $N(t)$  via Eq. 4.6, with  $x$  in the standard  $w - 1$  simplex in  $\mathbb{R}^w$  and  $G(x) = \sum_{i=1}^w x_i \log r_i$ . The computation of the sample exponents  $b_R$  and  $b_{jk}$  is similar

## Chapter 4. Taylor's law of fluctuation scaling

---

to that in the 2-state case and the sample exponents approximate  $b_R = 2$  and  $b_{jk} = k/j$  asymptotically in time; the proof is as follows. Consider the independent identically distributed random variables  $Y^i(t) = |L_t^i - \mu_{\min}|$ , where  $L_t^i = (L_t^i(r_1), \dots, L_t^i(r_w))$  and the superscript  $i$  indicates the  $i$ -th independent realization of the chain ( $i = 1, \dots, R$ ). Let  $y^+ = \max\{Y^1(t), \dots, Y^R(t)\}$ ; for every  $\epsilon > 0$ :

$$\mathbb{P}(y^+ > \epsilon) \leq R \mathbb{P}(Y^1(t) > \epsilon). \quad (4.21)$$

For fixed  $R$  and  $\epsilon$ , taking the limit ( $\lim_{t \rightarrow \infty}$ ) in Eq. 4.21 and knowing that  $L_t^1$  satisfies a LDP (in particular,  $\lim_{t \rightarrow \infty} \mathbb{P}(Y^1(t) > \epsilon) = 0$ ), one has:

$$\lim_{t \rightarrow \infty} \mathbb{P}(y^+ > \epsilon) = 0. \quad (4.22)$$

In this context, one can approximate the sample exponent with:

$$b_R(\lambda, t) \simeq \frac{\sup_{|\mu - \mu_{\min}| < y^+} [2G(\mu) - I_{\Pi}(\mu)]}{\sup_{|\mu - \mu_{\min}| < y^+} [G(\mu) - I_{\Pi}(\mu)]}. \quad (4.23)$$

In the narrow region  $|\mu - \mu_{\min}| < y^+$  centered around  $\mu_{\min}$ ,  $I_{\Pi}(\mu) \simeq 0$  and as a consequence  $b_R(\lambda, t) \simeq 2$ . More precisely,  $|b_R(\lambda, t) - 2|$  goes to 0 in probability as  $t$  tends to infinity. In fact, for every  $\delta > 0$ :

$$\mathbb{P}(|b_R(\lambda, t) - 2| > \delta) \leq \mathbb{P}(y^+ > \eta(\delta)), \quad (4.24)$$

where  $\eta(\delta)$  is a function that goes to zero for  $\delta \rightarrow 0$ . Because of Eq. 4.22, it follows that:

$$\lim_{t \rightarrow \infty} \mathbb{P}(|b_R(t) - 2| > \delta) = 0. \quad (4.25)$$

Analogous considerations hold for the generalized TL describing the scaling of any pair of moments. A standard saddle-point calculation suggests that the limiting growth rate of the variance is equal to the limiting growth rate of the second moment also for ergodic transition matrices, apart from peculiar cases (see Cohen (2014a) for a discussion of a counterexample). The same argument suggests that the limiting growth rate of the  $k$ -th cumulant equals that of the  $k$ -th moment ( $t^{-1} \log \mathbb{E}[N(t)^k]$ ) for large  $t$ . The suggested equivalence between the scaling exponents of cumulants and moments for ergodic  $\Pi$  would allow extending the result on the sample TL ( $b_R = 2$ ) and generalized TL ( $b_{jk} = k/j$ ) to the scaling of cumulants in  $m$ -step Markov chains, whose transition matrix is ergodic but not two-fold irreducible. However, pathological counterexamples may exist.

In ecological contexts, the number of realizations  $R$  that determine the possible convergence of sample and population TL exponents could refer, for instance, to independent patches experiencing different realizations of the same climate (Cohen, 2014b). In an established ecosystem, species have been present for several generations, and one might assume that the system is in the asymptotic regime  $t \gg \log R$ . Within this perspective, the prediction that for large  $t$  sample exponents satisfy the relation  $b_{jk} = k/j$  (including the conventional TL) was tested on two datasets.

A first example is drawn from a long-term census of six plots within the Black Rock Forest (BRF) (Cohen et al., 2013). It was shown that the Lewontin-Cohen (LC) model (a particular case of the multiplicative model studied here) describes the population dynamics of trees in the BRF (Cohen et al., 2013). The interpretation of the six plots as distinct and independent replicates of the LC model is supported by statistical analysis (Cohen et al., 2013) and allowed relating the model predictions to the spatial TL. Here, the same dataset was used to show that the generalized TL holds with sample exponent  $b_{jk} = k/j$ . The moment ratios  $\langle [N(t)/N_0]^k \rangle$  were computed, where the symbol  $\langle \cdot \rangle$  identifies the sample mean across the six plots of BRF and  $N_0$  is the number of trees at the start of the census in 1931. Following Cohen et al. (2013), it was tested whether the moments of the spatial density ratio  $N(t)/N_0$  in the five most recent censuses satisfied TL and the generalized TL with  $b_{jk} = k/j$ . Table 4.1 reports the slopes of the least-squares linear regressions of  $\langle [N(t)/N_0]^k \rangle$  versus  $\langle [N(t)/N_0]^j \rangle$ , which are all compatible with the model prediction  $b_{jk} = k/j$ . The BRF dataset thus provides an empirical example where the multiplicative model satisfactorily describes the underlying dynamics and the generalized TL holds asymptotically as the model predicts.

**Table 4.1:** Sample exponents for the generalized TL in the Black Rock Forest dataset, data from Cohen et al. (2013).

(j,k)	$k/j$	$b_{jk} \pm \text{SE}$	$R^2$
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

## Chapter 4. Taylor's law of fluctuation scaling

A second example uses the data in Den Boer (1977), where abundances of carabid beetles were measured in various sites across The Netherlands within a 200-km<sup>2</sup> area for 8 consecutive years. The dataset was shown to support the conventional spatial TL (Hanski, 1982). The sample moments of carabid beetles abundance,  $\langle N^k(t) \rangle$ , were computed across similar sites (either woodland or heath), for each species separately and year  $t$ . In the intra-specific analysis (Fig. 4.4), linear regressions of  $\log\langle N^k(t) \rangle$  vs  $\log\langle N^j(t) \rangle$  for  $t = 1, \dots, Y$  ( $Y$  is the total number of years) gave the estimate of the sample exponent  $b_{jk}$  for each species. Frequency histograms of empirical exponents  $b_{jk}$  are shown in Fig. 4.5. A one-sample  $t$ -test does not reject the null hypothesis that the sample mean of  $b_{jk}$  does not differ significantly from the theoretically predicted mean  $k/j$  (see Fig. 4.5). In the inter-specific analysis (Fig. 4.6), the least-squares slope  $b_{jk}$  (for  $j = 1$ ) of  $\log\langle N^k \rangle$  versus  $\log\langle N \rangle$  across all species at a given year and site type was computed (Tables 4.2, 4.3). The empirical exponents  $b_{jk}$  for all years are compatible with the asymptotic model prediction  $b_{jk} = k/j$ , as are the mean (across years and site type) exponents  $b_{jk}$  (Table 4.4).

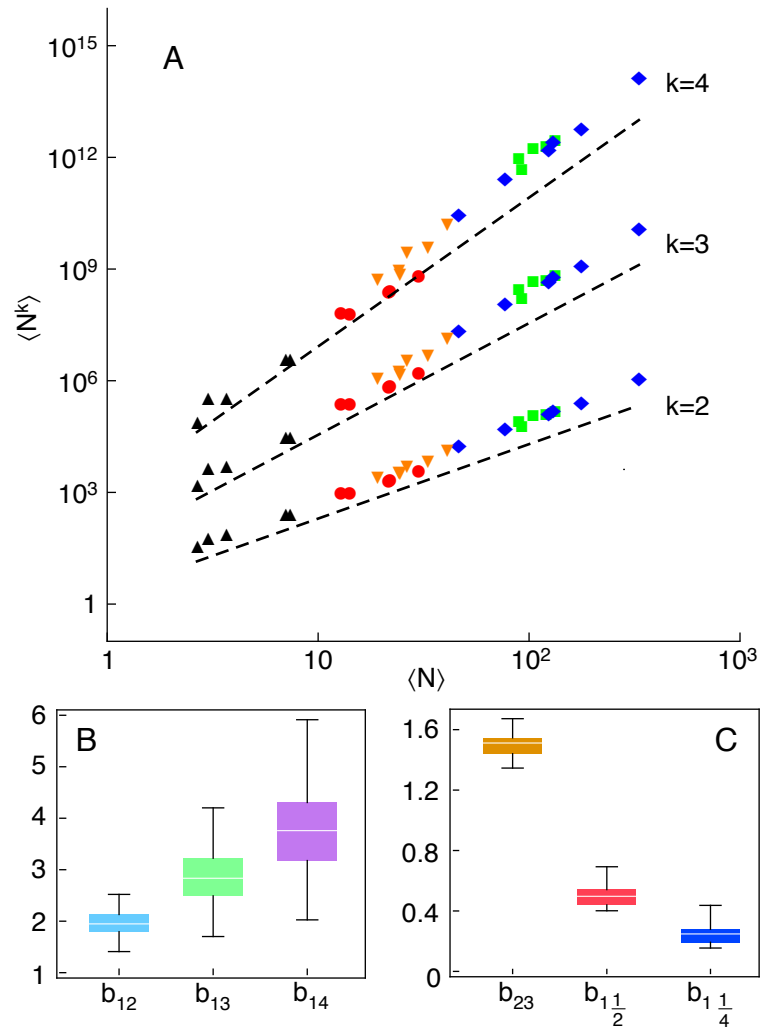
**Table 4.2:** Sample exponents for the inter-specific generalized TL on carabid beetles abundances in woodland sites, data from Den Boer (1977). The column  $k/j$  gives the asymptotic model prediction for the exponent  $b_{jk}$ . The estimates  $b_{jk}$  (mean $\pm$ SE) are the least-squares slopes of  $\log\langle N^k \rangle$  vs  $\log\langle N \rangle$ .  $R^2$  is the squared correlation coefficient. Nonlinearity was checked with least-squares quadratic regression on log-log coordinates. The coefficient of the second power term did not differ significantly from 0 in any of the regressions; hence, the null hypothesis of linearity was not rejected.

		1961		1962		1963	
$j, k$	$k/j$	$b_{jk} \pm \text{SE}$	$R^2$	$b_{jk} \pm \text{SE}$	$R^2$	$b_{jk} \pm \text{SE}$	$R^2$
1,2	2	2.03 $\pm$ 0.09	0.988	2.07 $\pm$ 0.04	0.995	2.00 $\pm$ 0.07	0.988
1,3	3	3.04 $\pm$ 0.18	0.976	3.13 $\pm$ 0.09	0.991	3.00 $\pm$ 0.15	0.977
1,4	4	4.03 $\pm$ 0.28	0.968	4.20 $\pm$ 0.14	0.988	4.01 $\pm$ 0.23	0.971
No. points		9		13		11	

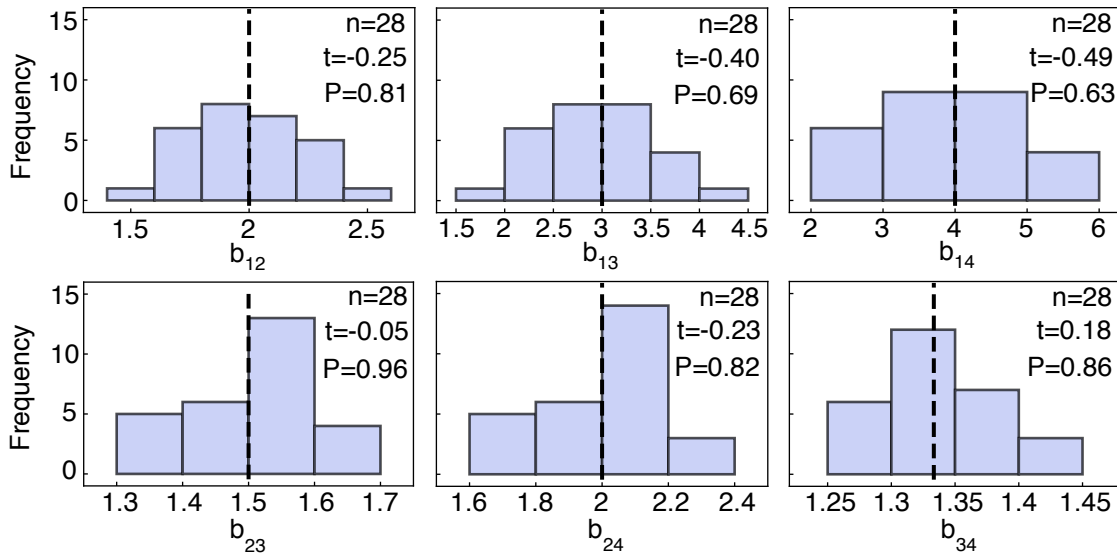
		1964		1965		1966	
$j, k$	$k/j$	$b_{jk} \pm \text{SE}$	$R^2$	$b_{jk} \pm \text{SE}$	$R^2$	$b_{jk} \pm \text{SE}$	$R^2$
1,2	2	1.96 $\pm$ 0.09	0.977	2.01 $\pm$ 0.07	0.989	1.97 $\pm$ 0.06	0.995
1,3	3	2.94 $\pm$ 0.20	0.957	3.00 $\pm$ 0.16	0.976	2.90 $\pm$ 0.12	0.989
1,4	4	3.92 $\pm$ 0.29	0.947	4.00 $\pm$ 0.24	0.967	3.83 $\pm$ 0.18	0.985
No. points		12		11		9	





**Figure 4.4:** Generalized TL for intra-specific patterns of carabid beetles abundance. (A) Double logarithmic plot of  $\langle N^k \rangle$  vs  $\langle N \rangle$  for different species (identified by different colors and symbols), for consecutive years (each symbol refers to a single year  $t$ ). For visual clarity, only 5 species are shown. Dashed black lines of slopes  $b_{1k} = k$  (asymptotic model prediction) are shown. Vertical offsets are introduced to aid comparison of slopes. (B-C) box and whisker plots for the empirical distribution of intra-specific generalized TL exponents  $b_{1k}$ , showing the median (white horizontal line) and the 25% and 75% quantiles.

## Chapter 4. Taylor's law of fluctuation scaling



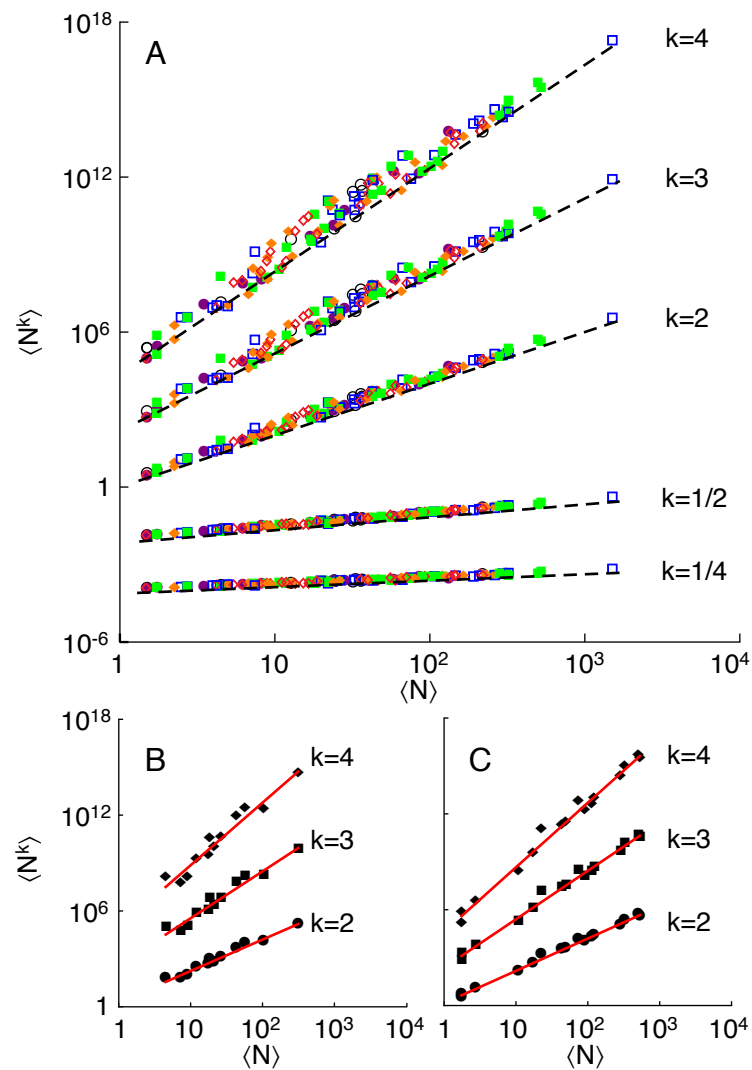
**Figure 4.5:** Frequency histogram for the exponent  $b_{jk}$  in the intra-specific generalized TL  $\langle N^k \rangle = a \langle N^j \rangle^{b_{jk}}$ , computed for each species (carabid beetles, Den Boer (1977)) across similar sites (woodland or heath). The dashed black line shows the value of the exponent  $b_{jk} = k/j$  as the asymptotic model predicted. The binning of data points is determined by using Scott's rule (Scott, 1979). Shown in each panel are the number of observations  $n$  of  $b_{jk}$ , the test statistic for the  $t$ -test of the null hypothesis that the sample mean of the values of  $b_{jk}$  did not differ significantly from the theoretically predicted mean  $k/j$  and the corresponding  $p$ -value.

**Table 4.3:** Sample exponents for the inter-specific generalized TL on carabid beetles abundances in heath sites, data from Den Boer (1977). The Table is organized as Table 4.2.

		1963		1964	
$j, k$	$k/j$	$b_{jk} \pm \text{SE}$	$R^2$	$b_{jk} \pm \text{SE}$	$R^2$
1,2	2	$1.99 \pm 0.05$	0.993	$2.02 \pm 0.04$	0.995
1,3	3	$2.98 \pm 0.09$	0.987	$3.03 \pm 0.08$	0.990
1,4	4	$3.83 \pm 0.18$	0.985	$3.96 \pm 0.14$	0.983
No. points		16		16	

		1965		1966	
$j, k$	$k/j$	$b_{jk} \pm \text{SE}$	$R^2$	$b_{jk} \pm \text{SE}$	$R^2$
1,2	2	$1.98 \pm 0.08$	0.982	$2.02 \pm 0.06$	0.986
1,3	3	$2.97 \pm 0.17$	0.965	$3.05 \pm 0.13$	0.974
1,4	4	$4.04 \pm 0.12$	0.987	$3.98 \pm 0.26$	0.956
No. points		13		17	



**Figure 4.6:** Generalized TL for inter-specific patterns of abundance of carabid beetles. (A) Double logarithmic plot of  $\langle N^k \rangle$  vs  $\langle N \rangle$  for all species, years and site type. Each data point refers to a single species in one year and site type. The color and symbol code identifies data relative to the same year: 1961 (black open circles), 1962 (purple filled circles), 1963 (blue open squares), 1964 (green filled squares), 1965 (orange filled diamonds), 1966 (red open diamonds). Dashed black lines of slope  $b_{1k} = k$  (asymptotic model prediction) are plotted next to the corresponding data series. Vertical offsets are introduced to aid comparison of slopes. (B-C) Examples of inter-specific moments scaling (each data point refers to a single species) for a single year and site type (B, woodland 1964 - C, heath 1964) used for the statistical analysis (Tables 4.2, 4.3, 4.4). The red lines are the least-squares regressions of  $\log \langle N^k \rangle$  vs  $\log \langle N \rangle$  across species.

**Table 4.4:** Statistics of estimated sample exponents in the inter-specific generalized TL on carabid beetles abundances. The column  $k/j$  gives the asymptotic model prediction for the exponent  $b_{jk}$ . The point estimate is computed as the average  $b_{jk}$  across years and site type, not by pooling all the data from different years and site types to calculate means and variances. The confidence intervals are obtained via bootstrapping with  $10^6$  bootstrap samples from the set of  $b_{jk}$ .

$j, k$	$k/j$	$b_{jk}$ point estimate	2.5% percentile	97.5% percentile
1,2	2	2.005	1.984	2.025
1,3	3	3.005	2.961	3.042
1,4	4	3.994	3.936	4.057

### 4.3 Discussion

The multiplicative growth model is one of numerous demographic models that predict TL. The exponent  $b = 2$  for the scaling of the variance versus the mean is typical of deterministic dynamics. For example, an exponential model of clonal growth (Cohen, 2013), where clones grow exponentially with different growth rates (variability enters here only through the different growth rates and initial densities), and the above symmetric model for  $\lambda = 0$  or  $\lambda = 1$  both predict TL with exponent  $b = 2$ . Although found in deterministic models, the exponent  $b = 2$  is also observed in stochastic models such as the continuous-time birth-death process and the Galton-Watson branching process (Cohen, 2014a). Such models yield population exponents  $b = 2$  and  $b = 1$  respectively for asymptotically growing and decaying populations (Cohen, 2014a).

The theoretical investigation of multiplicative population processes showed that the generalized TL sample exponents  $b_{jk}$  satisfy  $b_{jk} \simeq k/j$  asymptotically for large  $t$  for a broad ensemble of transition matrices  $\Pi$  and sets of positive multiplicative factors. Additionally, the large-deviation approach and the small-sample argument adopted here suggest that the entropic term in Eq. 4.11 dominates over the other terms that contain the specifications of the demographic process. Thus, the result might be more general than the class of multiplicative population growth models. It is shown in the Appendix that  $b_{jk} = k/j$  holds also for the population exponents of other population growth processes, such as the birth-death process in the case of expanding populations. The empirical confirmation and the novel finding that other demographic models predict the generalized TL with  $b_{jk} = k/j$  (Appendix) indicate that these predictions are probably insensitive to the details of the dynamics, just as the original TL is quite robust (Fronczak and Fronczak, 2010, Kendal and Jørgensen, 2011, Xiao et al., 2014).

In conclusion, a general mechanism that yields TL with the widely observed sample exponent  $b \simeq 2$  was uncovered. For a broad range of parameters within the class of multiplicative models, and other demographic processes, the generalized TL describes the scaling of moments and cumulants with the sample exponent  $b_{jk}$  asymptotically equal to  $k/j$ . This phenomenon may be attributable to the finite size of both ecosystems and sampling efforts. TL may not reflect (or depend on) the underlying population dynamics. The theoretical prediction is supported by two empirical examples and invites further testing. Notably, this study suggests that limited sampling efforts might hinder the observation of abrupt transitions in population exponents that were recently discovered for theoretical multiplicative growth processes. Because fluctuations in population abundances strongly affect ecological dynamics, particularly

## **Chapter 4. Taylor's law of fluctuation scaling**

---

extinction risk, comparable real-world abrupt transitions could harm fish populations, forests, and public health, and could alter agricultural pest dynamics. This study shows that limited sampling hinders the detection and anticipation of such abrupt transitions. The calculation of the minimum number of samples required to reveal such transitions provided here may help to identify early-warning signals of abrupt biotic change following smooth changes in the environment.

## 4.4 Appendix

### 4.4.1 Analysis of the discontinuity in $b$ as a function of $r$ and $s$

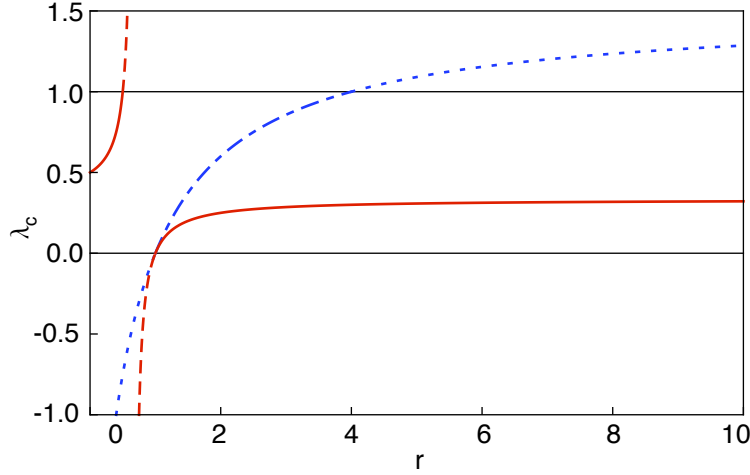
A discontinuity in the population TL exponent  $b$  (Fig. 1, Eq. 4.7) is present when the limiting growth rate of the mean abundance is zero, i.e.,  $\lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}[N(t)] = 0$ . Consider Fig. 4.2 and fix  $r$  and  $s$  with  $r \neq s$ . The value of  $\lambda$  shapes the form of  $I_{\Pi}(x)$  (black curve in Fig. 4.2); in particular, the second derivative can be easily calculated from Eq. 4.4 and shown to increase for larger  $\lambda$ . A discontinuity may eventually appear for the value  $\lambda = \lambda_c$  such that the curve  $I_{\Pi}(x)$  and the line  $G(x)$  (blue line in Fig. 4.2) are tangent. In other words,  $\lim_{t \rightarrow \infty} t^{-1} \log \mathbb{E}[N(t)] = \sup_{x \in [0,1]} [G(x) - I_{\Pi}(x)] = 0$  for  $\lambda = \lambda_c$  such that:

$$\log \frac{1}{2} \left[ (1 - \lambda_c)(r + s) + \sqrt{4(2\lambda_c - 1)rs + (\lambda_c - 1)^2(r + s)^2} \right] = 0, \quad (4.26)$$

with constraints  $r, s > 0$  and  $0 < \lambda_c < 1$ .  $\lambda_c$  exists only for certain values of  $r$  and  $s$ , thus a discontinuity in the population TL exponent  $b$  is not always possible. Solving Eq. 4.26 with respect to  $\lambda_c$  gives  $\lambda_c = \frac{1-r-s+rs}{-r-s+2rs}$ ; thus, for any given  $s$ ,  $\lambda_c = 0$  for  $r = 1$  and  $\lambda_c = 1$  for  $r = 1/s$ . For fixed  $s \neq 1$  one has  $d\lambda_c/dr > 0$  (except for  $r = s$  where  $d\lambda_c/dr|_{r=s} = 0$ ); thus,  $\lambda_c$  exists for  $0 < r \leq 1/s$  and  $r \geq 1$  if  $s > 1$  and for  $1 \leq r \leq 1/s$  if  $s < 1$  (see Fig. 4.7). Fig. 4.8 schematically illustrates the behavior of  $b(\lambda)$  for different pairs  $\{r, s\}$  of multiplicative factors. Discontinuities analogous to that of  $b(\lambda)$  appear for certain values of  $r$ ,  $s$  and  $\lambda$  in the population exponents  $b_{jk}$  (Eq. 4.8), when  $\lim_{t \rightarrow \infty} t^{-1} \log \mathbb{E}[N(t)^j] = \sup_{x \in [0,1]} [jG(x) - I_{\Pi}(x)] = 0$ .

### 4.4.2 Compatibility of Eq. 4.7 here and Eq. 8 in Cohen (2014a)

It is shown here that Eq. 4.7 coincides with Eq. 8 in Cohen (2014a), under the assumption (stronger than in Cohen (2014a)) that the transition matrix  $\Pi$  is positive and  $r \neq s$ . The rate function Eq. 4.3 can be written as (Section 4.3 of Touchette (2009) or Theorem 3.1.7 of Dembo and Zeitouni (2009))  $I_{\Pi}(x) = \sup_q \{qx - \log \zeta(\Pi_q)\}$ , where  $\Pi_q$  is the matrix with elements  $\Pi_q(i, j) = \Pi(i, j) \exp(q\delta_{j,1})$ , and  $\zeta(\cdot)$  indicates the spectral radius (i.e., the Perron-Frobenius eigenvalue).  $\zeta(\Pi_q)$  is unique and analytic in  $q$ ; thus,  $\xi(q) \equiv \log \zeta(\Pi_q)$  is differentiable and the rate function can be expressed as  $I_{\Pi}(x) = q(x)x - \xi(q(x))$ , where  $q(x)$  is the unique solution of  $\xi'(q) = x$ . Eq. 4.6 for the  $k$ th moment of  $N(t)$  then reads  $\lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}[N(t)^k] = \sup_{x \in [0,1]} [kG(x) - q(x)x + \xi(q(x))]$ . The argument of the supremum is maximum at  $x^*$  such that  $k \log(r/s) - q(x^*) = 0$ , that is,  $x^* = \xi'(k \log(r/s))$ . Thus, evaluating the supremum one has  $\lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}[N(t)^k] = k \log s + \xi(k \log(r/s)) = \log [s^k \zeta(\Pi_{k \log(r/s)})] = \log \zeta(\Pi \text{diag}(r, s)^k)$ , which coincides with



**Figure 4.7:** The critical transition probability  $\lambda_c$  as a function of  $r$  (with  $s$  fixed). Below the black horizontal line at  $\lambda_c = 0$  and above the black horizontal line at  $\lambda_c = 1$ ,  $\lambda_c$  does not exist. The red (solid for  $0 \leq \lambda_c \leq 1$  and dashed otherwise) and blue (dash-dotted for  $0 \leq \lambda_c \leq 1$  and dotted otherwise) lines  $\lambda_c = \frac{1-r-s+rs}{-r-s+2rs}$  were calculated by solving Eq. 4.26 with respect to  $\lambda_c$  with, respectively,  $s = 2$  and  $s = 1/4$ . For any given  $s$ ,  $\lambda_c = 0$  for  $r = 1$  and  $\lambda_c = 1$  for  $r = 1/s$ .

Eqs. 13, 14 of Cohen (2014a) (Eqs. 13, 14 in Cohen (2014a)) are expressed in terms of the column-stochastic matrix  $\Pi^\top$  that corresponds to the row-stochastic matrix  $\Pi$ ; because  $\zeta(\text{diag}(r, s)^k \Pi^\top) = \zeta(\Pi \text{diag}(r, s)^k)$ , the equations coincide).

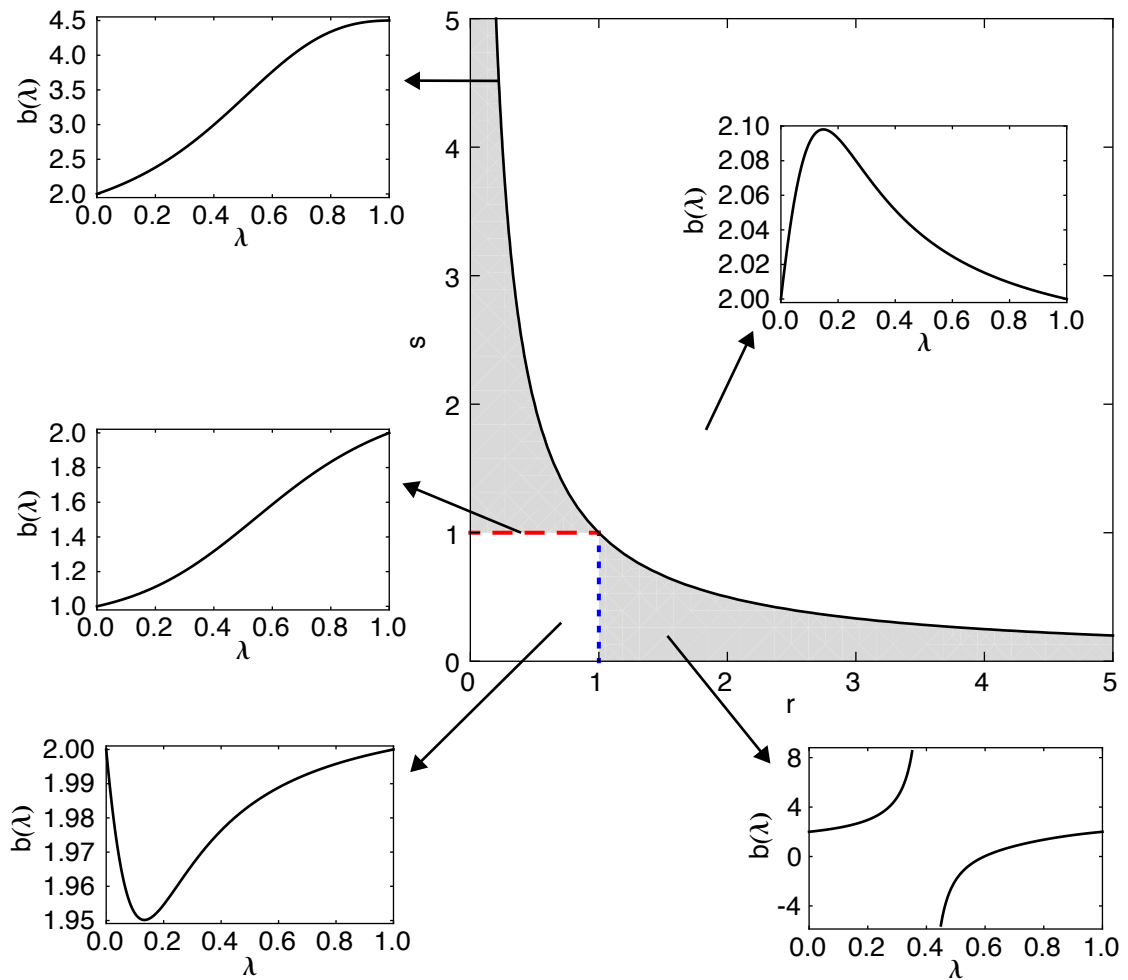
### 4.4.3 Comparison with other demographic models

It is shown here that  $b_{jk} = k/j$  holds for the population exponents of the birth-death process in the case of expanding populations. The moments of the birth-death process with constant birth rate  $\lambda$  and constant death rate  $\mu$  can be computed via the associated moment generating function  $M$ , which is equal to (Bailey, 1964):

$$M(\theta, t) = \left( \frac{\mu v(\theta, t) - 1}{\lambda v(\theta, t) - 1} \right)^{N_0}, \quad (4.27)$$

where  $v(\theta, t) = \frac{(e^\theta - 1)e^{(\lambda - \mu)t}}{\lambda e^\theta - \mu}$  and  $N_0$  is the initial population size. The  $k$ -th moment of population size can be computed as  $\langle N^k \rangle = \frac{\partial^k M(\theta, t)}{\partial \theta^k} |_{\theta=0}$ . Here,  $N_0 = 1$  is assumed (but the result holds for any  $N_0$ ) and an expanding population, i.e.,  $\lambda - \mu > 0$ . Because  $v(0, t) = 0$ ,  $\frac{\partial v}{\partial \theta}(\theta, t) = (\lambda - \mu)e^{(\lambda - \mu)t} \frac{e^h}{(-e^h \lambda + \mu)^2} \propto e^{(\lambda - \mu)t}$  and  $\frac{\partial^k v}{\partial \theta^k}(\theta, t) \propto e^{(\lambda - \mu)t}$ , the leading term in the partial derivatives of  $M(\theta, t)$  with respect to  $\theta$ , evaluated in  $\theta = 0$ , can





**Figure 4.8:** Existence of a critical transition probability  $\lambda_c$ . Smaller panels show the population exponent  $b(\lambda)$  (Eq. 4.7) for various choices of the multiplicative factors in different regions of the plane  $(r, s)$  (larger panel). Only in the interior of the gray region of the plane  $(r, s)$ ,  $\lambda_c$  exists. The solid black line represents the curve  $rs = 1$ .

be written as:

$$\begin{aligned} \left. \frac{\partial^k M}{\partial \theta^k}(\theta, t) \right|_{\theta=0} &= (-1)^{k+1} (\lambda - \mu) \lambda^{k-1} \frac{\left( \frac{\partial v}{\partial \theta} \right)^k}{(-1 + \lambda v)^{k+1}} \Big|_{\theta=0} + o \left[ \left( \frac{\partial v}{\partial \theta} \right)^k \Big|_{\theta=0} \right] \\ &= (\lambda - \mu)^{1-k} \lambda^{k-1} e^{k(\lambda - \mu)t} + o \left[ e^{k(\lambda - \mu)t} \right], \end{aligned} \quad (4.28)$$

where the little- $o$  notation indicates that the remaining terms are negligible in the limit  $t \rightarrow \infty$ . Derivation of the equation for  $\left. \frac{\partial^k M}{\partial \theta^k}(\theta, t) \right|_{\theta=0}$  (first line of Eq. 4.28) shows that the leading term in  $\left. \frac{\partial^{k+1} M}{\partial \theta^{k+1}}(\theta, t) \right|_{\theta=0}$  is equal to  $(\lambda - \mu)^k \lambda^k e^{(k+1)(\lambda - \mu)t} + o \left[ e^{(k+1)(\lambda - \mu)t} \right]$ , which coincides with replacing  $k$  by  $k + 1$  in Eq. 4.28. Eq. 4.28 can be obtained by considering that, because  $\frac{\partial^k v}{\partial \theta^k} \propto e^{(\lambda - \mu)t}$  and  $v(0, t) = 0$ , the leading term in  $\left. \frac{\partial M}{\partial \theta}(\theta, t) \right|_{\theta=0} = (\lambda - \mu) \frac{\partial v / \partial \theta}{(-1 + \lambda v)^2}$  evaluated at  $\theta = 0$  is the second term in the quotient rule  $(f/g)' = (f'g - fg')/g^2$ , that is, the term that raises the exponent of  $\frac{\partial v}{\partial \theta}$  by one unit. For subsequent derivatives, the quotient rule is applied to the leading term. All other terms in  $\left. \frac{\partial^k M}{\partial \theta^k}(\theta, t) \right|_{\theta=0}$  contain products of partial derivatives, i.e.,  $\prod_{j=1}^k \left( \frac{\partial^j v}{\partial \theta^j} \right)^{q_j}$ , with  $\sum_{j=1}^k q_j < k$  (with  $q_j \in \mathbb{N}$ ) and are thus negligible in the limit  $t \rightarrow \infty$ . From Eq. 4.28 it follows that  $\lim_{t \rightarrow \infty} \frac{1}{t} \log \langle N^k \rangle = k(\lambda - \mu)$ ; thus, the generalized TL holds with  $b_{jk} = k/j$ .

The asymptotic behavior of exponents, i.e.,  $\lim_{t \rightarrow \infty} \frac{1}{t} \log \langle N^k \rangle = k(\lambda - \mu)$ , can also be computed via the continuous approximation of the birth-death process. Although such calculations do not provide further understanding of the birth-death process, the fact that the continuous approximation of the birth-death process coincides with that of the Galton-Watson branching process (Feller, 1957, Harris, 1963, Rubin et al., 2014) suggests an even broader validity for the generalized TL result  $b_{jk} = k/j$ . The detailed calculation of exponents in the continuous approximation of the birth-death process and the Galton-Watson branching process is provided in the following section.

### **Moments of population density in the continuous approximation of the birth-death process and the Galton-Watson branching process**

The forward Kolmogorov equation for the continuous approximation of the birth-death process reads (Feller, 1957, Harris, 1963, Rubin et al., 2014):

$$\frac{\partial p(x, t)}{\partial t} = -\alpha \frac{\partial [xp(x, t)]}{\partial x} + \frac{\beta}{2} \frac{\partial^2 [xp(x, t)]}{\partial x^2}, \quad (4.29)$$

where  $p(x, t)$  is the probability density function for the population density  $x$  at time  $t$  (here,  $x \in \mathbb{R}$  is the population density and should not be confused with the frequency of multiplicative factors used in the rest of the chapter). Eq. 4.29 is the continuous approximation of a birth-death process with birth rate  $\lambda$  and death rate  $\mu$  such that  $\alpha = \lambda - \mu$  and  $\beta = \lambda + \mu$ . Eq. 4.29 also arises as the continuous approximation of the

Galton-Watson branching process for large populations (Feller, 1957, Harris, 1963, Rubin et al., 2014). The solution of Eq. 4.29 with initial condition  $x(0) = x_0$  is known (Bailey, 1964) and is equal to:

$$p(x, t) = \frac{2\alpha}{\beta(e^{\alpha t} - 1)} \left( \frac{x_0 e^{\alpha t}}{x} \right)^{\frac{1}{2}} \exp \left[ \frac{-2\alpha(x_0 e^{\alpha t} + x)}{\beta(e^{\alpha t} - 1)} \right] I_1 \left[ \frac{4\alpha(x_0 x e^{\alpha t})^{\frac{1}{2}}}{\beta(e^{\alpha t} - 1)} \right], \quad (4.30)$$

where  $I_1$  is the modified Bessel function of the first kind. Differentiation with respect to  $\gamma$  of the identity  $\int_0^\infty dx I_1(x) e^{-\gamma x^2} = e^{1/(4\gamma)} - 1$  gives the following equation:

$$C \int_0^\infty dx x^k x^{-\frac{1}{2}} I_1(x^{\frac{1}{2}} A) e^{Bx} = 2CA^{-(2k+1)} \left( -\frac{d}{d\gamma} \right)^k \Big|_{\gamma=-\frac{B}{A^2}} \left( e^{\frac{1}{4\gamma}} - 1 \right), \quad (4.31)$$

which allows calculating the moments of Eq. 4.30 with  $A = \frac{4\alpha(x_0 e^{\alpha t})^{\frac{1}{2}}}{\beta(e^{\alpha t} - 1)}$ ,  $B = \frac{2\alpha}{\beta(e^{\alpha t} - 1)}$  and  $C = \frac{2\alpha(x_0 e^{\alpha t})^{\frac{1}{2}}}{\beta(e^{\alpha t} - 1)} \exp \left[ -\frac{2\alpha x_0 e^{\alpha t}}{\beta(e^{\alpha t} - 1)} \right]$ . For an expanding population,  $\alpha > 0$ ; thus asymptotically for large  $t$ :

$$\begin{aligned} A &\propto e^{-\frac{\alpha t}{2}}, \\ B &\propto e^{-\alpha t}, \\ C &\propto e^{-\frac{\alpha t}{2}}. \end{aligned} \quad (4.32)$$

Therefore,  $\gamma = -\frac{B}{A^2}$  tends to a constant and one has:

$$\langle x^k \rangle \propto CA^{-2k+1} \propto (e^{\alpha t})^k, \quad (4.33)$$

which implies that, asymptotically, the generalized TL holds with exponent  $b_{jk} = k/j$ .



## 5 Scaling body size fluctuations

### **Abstract**

The size of an organism matters for its metabolic, growth, mortality and other vital rates. Scale-free community size spectra (i.e., size distributions regardless of species) are routinely observed in natural ecosystems and are the product of intra- and inter-species regulation of the relative abundance of organisms of different sizes. Intra- and inter-species distributions of body sizes are thus major determinants of ecosystems' structure and function. Here, it is shown experimentally that single-species mass distributions of unicellular eukaryotes covering different phyla exhibit both characteristic sizes and universal features over more than four orders of magnitude in mass. Remarkably, it is found that the mean size of a species is sufficient to fully characterize its size distribution and that the latter has a universal form across all species. An analytical physiological model accounts for the observed universality, which can be synthesized in a log-normal form for the intra-species size distributions. The characterization of scaling intra-specific body size distributions sets constraints for the interaction of ecological and physiological processes that give rise to scale-invariant community size spectra.

### 5.1 Introduction

Why should a continuous, gap-free spectrum of organismic sizes emerge from the ecological and evolutionary processes that shape ecosystems? The origins and the implications of the absence of preferential body sizes, which is routinely observed across a variety of ecosystems regardless of broad differences in climatic and environmental conditions (Cavender-Bares et al., 2001, Cohen et al., 2003, Huete-Ortega et al., 2012, Sheldon et al., 1972), have been attracting much interest from field and theoretical ecologists (Bascompte and Solé, 1995, Brown et al., 2004, Brown and West, 2000, Chisholm, 1992, Levin, 1992, Marquet et al., 2005, Solé et al., 1999). Scale invariance, epitomized by power-law probability distributions (Bak, 1997, Bak et al., 1987, Camacho and Solé, 2001, Mandelbrot, 1973, Marquet, 2000, Marquet et al., 2005, Rodriguez-Iturbe and Rinaldo, 1997, Solé et al., 1999), requires regularities of the component parts (the species' size distributions) making up the whole (the community size spectra, i.e., the probability distributions of size regardless of species). In particular, a necessary condition for scaling community size spectra is the lack of peaks that pinpoint frequent occurrences and therefore excess abundance (and vice-versa) within any given range of sizes. Such features are particularly interesting if robust to environmental fluctuations as their dynamic origin could lie in the self-organization of complex adaptive systems (Bak et al., 1987, Levin, 1992, Solé et al., 1999).

Body size distributions in natural ecosystems are strongly related to the life history of the organisms and to the dynamics of their living communities (Marquet, 2000); thereby, they modulate the structure and function of the ecosystem at any scale. Size spectra, which display the relative abundance of organisms of different sizes within or across species, convey a synoptic, possibly taxon-independent image of ecological communities (Cohen et al., 2003, Holling, 1992, Sheldon et al., 1972, White et al., 2007). As such, they have long been attracting much interest in ecology as they hold important predictive power, e.g., fish stock projections from planktonic size spectra (Sheldon et al., 1972). Because examples and counterexamples of scaling spectra abound (Cavender-Bares et al., 2001, Chisholm, 1992, Holling, 1992, Lampert and Tlustý, 2013, Rinaldo et al., 2002, Rodriguez and Mullin, 1986, Sheldon et al., 1972), it is an unsettled issue whether scaling size spectra represent some central tendency of statistically stationary states of natural ecosystems. For instance, the operational computation of mean phytoplankton size was shown to typically depend on the sample size (Chisholm, 1992) and scaling relationships were documented for interspecific plant biomass (Niklas and Enquist, 2001, Simini et al., 2010, White et al., 2007), whereas some terrestrial ecosystems exhibit ubiquitous gaps in size and uneven relative abundances of organisms (Cohen et al., 2003, Holling, 1992).

Single species inhabiting communities, however, do exhibit species-specific mean and variance of their sizes, as even common sense suggests. There naturally exists, therefore, the mean size of a particular species, as usually implied by most biological scaling laws (Brown and West, 2000) where one typical mass subsumes a whole distribution of sizes. One wonders, therefore, how evolutionary and ecological processes interact to modulate species' abundances, the range of sizes proper to each functional group and number of species existing within a given niche or range of sizes to concoct regular, taxon-independent continuous size spectra. Moreover, one expects that the existence of a range of possible sizes for a species (and how such range varies for different mean sizes) has to be taken into account when addressing scaling laws in biology, e.g., allometric ones (Brown and West, 2000, Cohen et al., 2012, Damuth, 1981, May, 1988).

This chapter is organized as follows. First, the experiment where species size distributions were measured in standardized environmental conditions is described and a finite-size scaling framework is proposed to prove the universality of body size distributions. Second, body size distributions are measured in different environmental conditions or in the presence of other species. Such distributions are shown to adhere to the same finite-size scaling framework. Third, a mathematical model of cell growth and division is shown to produce body size distribution with the functional form found in the data.

## 5.2 Results

Here, the intra-species size distributions of 13 species of protists were measured in isolation or in competition, covering a relatively broad set of field conditions (see Methods). Examples of such distributions as functions of the linear size in standard environmental conditions are shown in Fig. 5.1. The corresponding transformed distributions as functions of volume span over four orders of magnitude and are shown in Fig. 5.2A. Let  $p_k(m)$  denote the measured size spectrum of the  $k$ th species:  $p_k(m)$  measures the relative proportion of individuals of a given species  $k$  with mass belonging to  $(m, m + dm)$ , assuming a continuous distribution of sizes. Herein, it was tested whether  $p_k(m)$  exhibits a finite-size scaling form (Banavar et al., 2007, 1999a,b, Fisher and Barber, 1972, Rinaldo et al., 2002) obtained by the product of two terms, an algebraic power of size multiplied by a suitable scaling function  $F$ , i.e.,

$$p_k(m) = \frac{1}{m^\Delta} F\left(\frac{m}{\langle m \rangle_k^\phi}\right), \quad (5.1)$$

## Chapter 5. Scaling body size fluctuations

---

where  $\langle m \rangle_k$  is the mean mass of the  $k$ th species and  $F(x)$ , critically, is the same scaling function for all species (see Methods for dimensional analysis and for normalization conditions that  $F$  must satisfy). Eq. 5.1 implies that the only species-dependence of the size distribution occurs through the average mass  $\langle m \rangle_k$  of species  $k$ . Note that the two exponents in Eq. 5.1,  $\Delta$  and  $\phi$ , are not independent. This follows from imposing  $\int_{\mathcal{R}} dm m p_k(m) \propto \langle m \rangle_k$  (where  $\mathcal{R}$  is the suitable range of sizes), in fact  $\int_{\mathcal{R}} dm m p_k(m) = \int_{\mathcal{R}} dm m \frac{1}{m^\Delta} F\left(\frac{m}{\langle m \rangle_k^\phi}\right) \propto \langle m \rangle_k^{(2-\Delta)\phi}$  is proportional to  $\langle m \rangle_k$  only if the two exponents satisfy  $(2 - \Delta)\phi = 1$  (see Methods for details concerning constraints on the exponents). To verify the hypothesis,  $m^\Delta p_k(m)$  is plotted versus  $m/\langle m \rangle_k^\phi$  for all 13 protist species (Fig. 5.2B) and  $\Delta$  and  $\phi$  are varied until a satisfactory data collapse (Barenblatt, 1983) is observed. The best collapse is found for  $\Delta = 1.0$  (and therefore  $\phi = 1.0$ , see Fig. 5.2B). A quantitative method (Bhattacharjee and Seno, 2001) to produce the best collapse yields  $\Delta = 1.01 \pm 0.05$  (see inset in Fig. 5.2B).

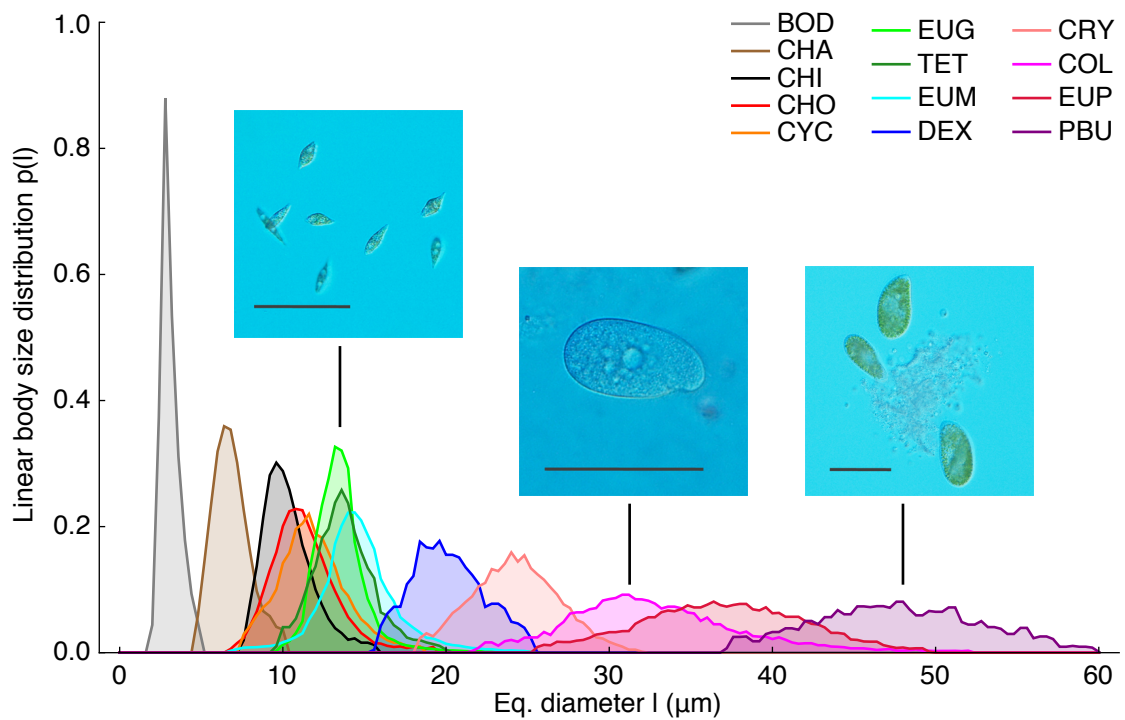
A relevant consequence of Eq. 5.1, where  $\Delta = \phi = 1$ , is that the  $j$ -th moment  $\langle m^j \rangle_k = \int_0^\infty m^j p_k(m) dm$  is proportional to  $(\langle m \rangle_k)^j$  (where  $j = 1, 2, 3, \dots$ ). In particular, the variance of the species' sizes does increase with the mean size. The proportionality of successive moments ratios to  $\langle m \rangle_k$  provides a test which further corroborates the validity of Eq. 5.1 (see Fig. 5.2C). Thus, it is found that a single parameter, the average mass of a species, is sufficient to fully characterize its size distribution. This is far from trivial because, in general, a probability distribution is determined by all its moments (Van Kampen, 2007).

Environmental factors are capable of affecting the size distribution of any given species (Bradshaw, 1965, Forster and Hirst, 2012). To test the effects of environmental conditions on Eq. 5.1, it is questioned whether the measured size distributions might still be described by the universal functional form:

$$p(m) = \frac{1}{m} F\left(\frac{m}{\langle m \rangle}\right), \quad (5.2)$$

where  $\langle m \rangle$  is the mean mass, determined either i) by the species or ii) by phenotypic plasticity due to environmental factors. To that end, a set of manipulated field conditions were investigated. Specifically, some of the above protist species were grown at different temperatures, or in pairwise competition with each other, over more than 15 generations to achieve relevant ecological timescales. Although obviously far from exhausting field-like scenarios, a sizeable plasticity was observed (see Fig. 5.3A). Crucially, once rescaled by the actual mean body size of the sample, whether constrained by temperature or by competition, all distributions collapse again and the scaling exponent estimate proves unaffected ( $\Delta = 1.01 \pm 0.10$ , see inset in Fig. 5.3B).





**Figure 5.1:** Size distributions of 13 protist species as functions of the equivalent diameter (i.e., the diameter a cell would have if it was spherical). The pictures show, from left to right, individuals of the species *Euglena gracilis*, *Colpidium* sp. and *Paramecium bursaria* (scale bar 100  $\mu\text{m}$ ). The legend links each color to the corresponding species (abbreviations as in Methods. Protist pictures by Regula Illi and Florian Altermatt).

## Chapter 5. Scaling body size fluctuations

---

Therefore, environmental factors are capable of affecting the size distribution of any given species, although they do not alter its scaling nature as both the exponents and the scaling function  $F$  are unchanged.

The observed regularities are compelling given the profound diversity of the species (belonging to the phyla/divisions of Ciliophora, Euglenozoa, Chlorophyta and Cryptophyta) considered in this study (Adl et al., 2012). Besides, protists and unicellular algae are of key ecological significance. In fact, they are the basic food source of almost all aquatic foodwebs, and unicellular algae are responsible for almost 50% of the world-wide biomass production (Field et al., 1998). Additionally, the observed universality of eukaryotes intra-species size distributions holds in a range of more than four orders of magnitude in mass. This suggests the existence of a simple underlying mechanism responsible for the empirical patterns observed.

One possible explanation for the reported universality may lie in the physiological processes that determine the size of unicellular species, namely cellular growth and cell division. Here, it is found that a simple mathematical model of these processes (Diekmann et al., 1983, Rading et al., 2011, Tyson and Diekmann, 1986) can justify the scaling form of unicellular eukaryotes' size distributions without the need to specify further biological details. Let  $N(m, t)$  be the number of organisms of mass  $m$  at time  $t$ : a cell's mass grows exponentially in time with rate  $\mu$  (i.e.,  $\dot{m} = \mu m$ ) and cell division occurs in time at a rate  $b(m)$ . Therefore, the fission rate depends on the mass of the cell (a mechanism known as sloppy size control, Tyson and Diekmann (1986)) and  $b(m)dt$  is the probability that a cell of mass  $m$  divides in a time  $dt$ . A maximum possible size  $M$  for a cell is introduced, i.e.  $N(m, t) = 0 \forall m > M$ , which requires (Diekmann et al., 1983):

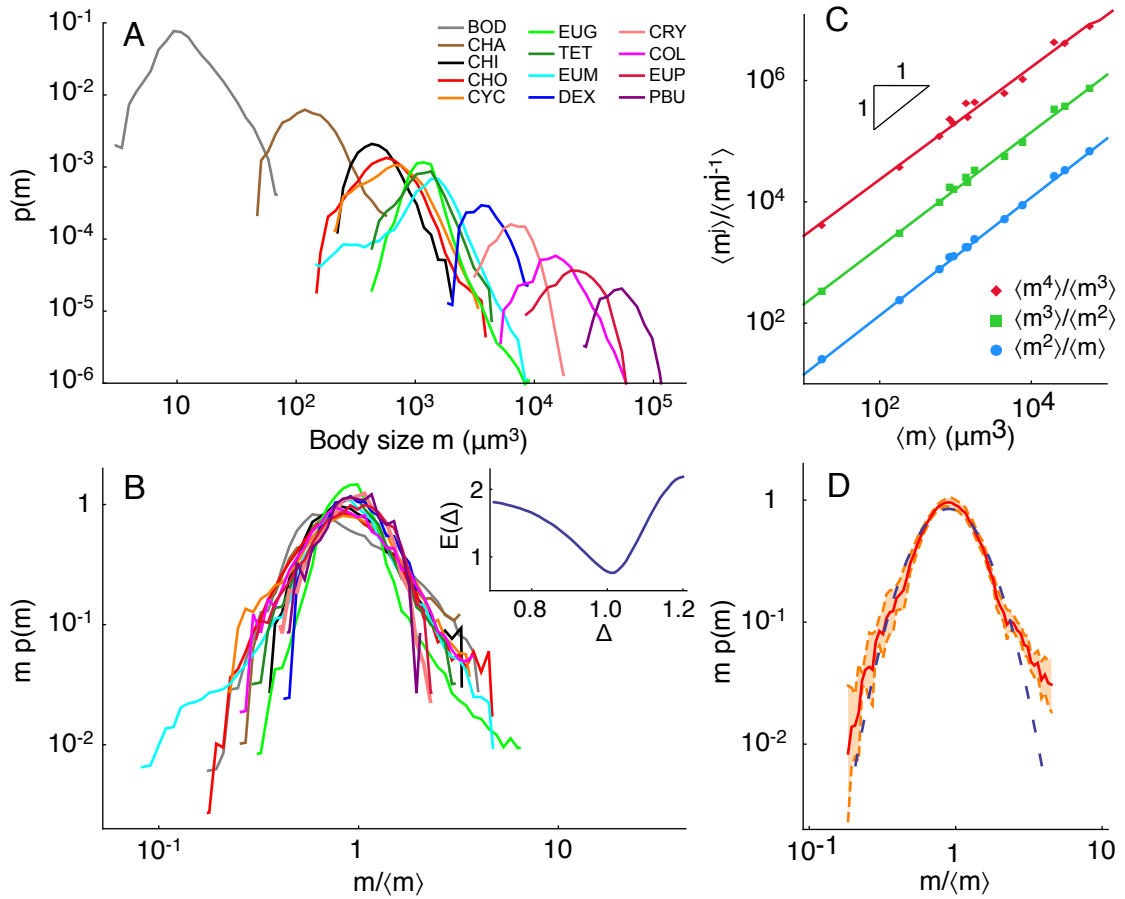
$$\int_0^M dm b(m) = \infty. \quad (5.3)$$

Considering the balance of growth and division in an infinitesimal time interval  $dt$  and in the size interval  $[m_1, m_2]$ , expanding at first order in  $dt$ , one has:

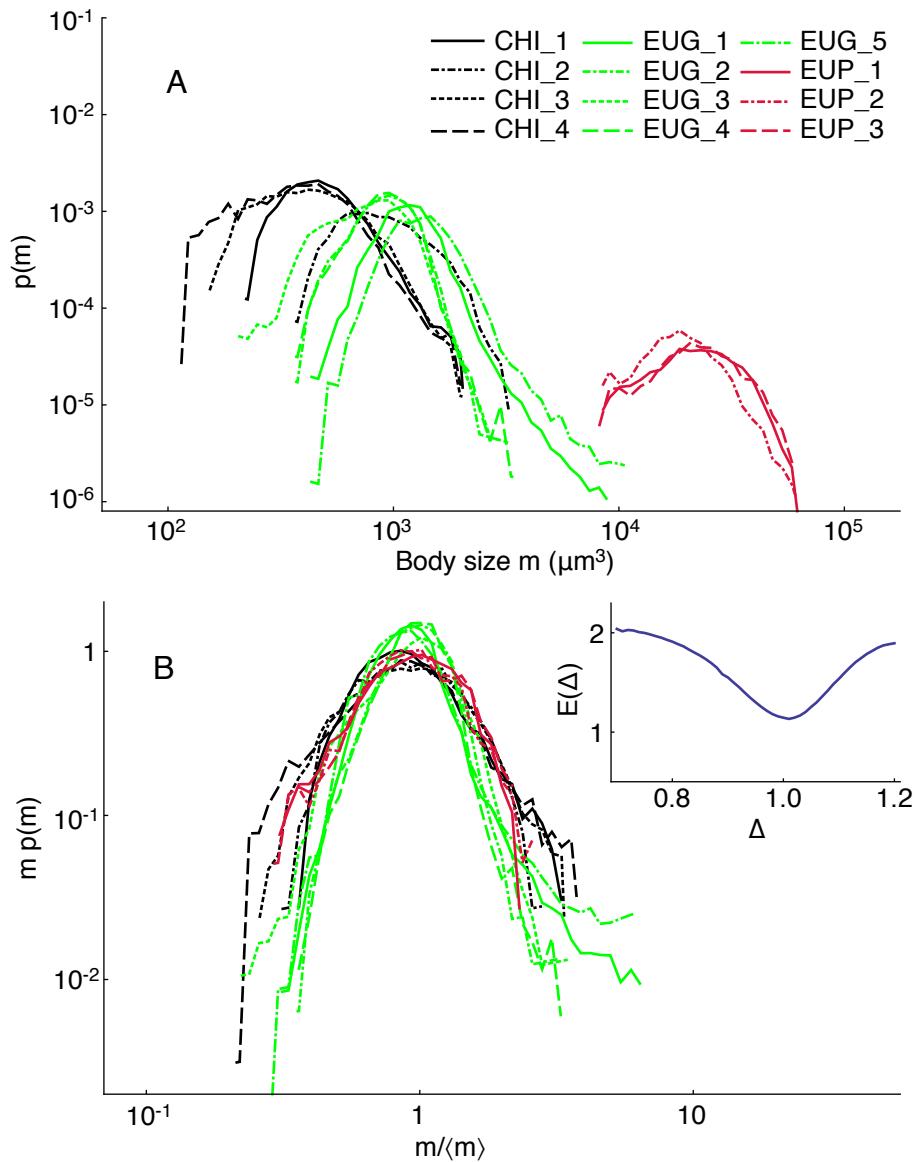
$$\int_{m_1}^{m_2} dm \left( \frac{\partial N}{\partial t}(m, t) + \frac{\partial[\mu m N(m, t)]}{\partial m} + b(m)N(m, t) - 4b(2m)N(2m, t) \right) = 0. \quad (5.4)$$

The equation governing the balance of growth and cell division is then:

$$\frac{\partial N}{\partial t}(m, t) + \frac{\partial[\mu m N(m, t)]}{\partial m} + b(m)N(m, t) - 4b(2m)N(2m, t) = 0. \quad (5.5)$$



**Figure 5.2:** Evidence for a universal single-species size distribution. (A) Volume probability distributions of 13 protist species, spanning four orders of magnitude in mass. (B) Data collapse of  $m^\Delta p(m)$  versus  $m/\langle m \rangle^\phi$ : the best collapse is observed for  $\Delta = \phi = 1.0$ . Inset: the minimum of the functional  $E(\Delta)$  provides the best estimate for the exponent and the associated error (Bhattacharjee and Seno, 2001). (C) The proportionality of successive moments of  $m$  to  $\langle m \rangle$  is an independent verification of the hypothesis in Eq. 5.1. (D) The fit of a Gaussian scaling function  $F$  as function of  $\log m/\langle m \rangle$  (dashed blue line) contrasting the ensemble average size distribution (red line), the orange region is the 99.7% confidence interval around the average. The scaling function yields a log-normal form for  $p_k(m)$ .



**Figure 5.3:** The universal form of the single-species size distribution is robust to biotic and abiotic forcings. (A) Volume probability distributions of three protist species in various environmental conditions and competition scenarios, different colors identify different species (abbreviations as in Methods). (B) The best data collapse of  $m^\Delta p(m)$  versus  $m/\langle m \rangle^\phi$  is observed for  $\Delta = \phi = 1.0$ . Inset: the minimum of the functional  $E(\Delta)$  provides the best estimate for the exponent and the associated error (Bhattacharjee and Seno, 2001).

It is assumed that at stationarity  $N(m, t) = \lambda(m)e^{kt}$ , where  $\lambda(m)$  is proportional to the stationary cell size distribution. Introducing  $N(m, t) = \lambda(m)e^{kt}$  in Eq. 5.5 one finds for  $\lambda(m)$ :

$$\mu \frac{d[m\lambda(m)]}{dm} = -[k + b(m)]\lambda(m) + 4b(2m)\lambda(2m). \quad (5.6)$$

One then has, integrating Eq. 5.6 in  $[0, M]$ :

$$k \int_0^M dm \lambda(m) + \mu \int_0^M dm \frac{d[m\lambda(m)]}{dm} = \int_0^M dm b(m)\lambda(m). \quad (5.7)$$

Eq. 5.7 imposes  $\int_0^M dm b(m)\lambda(m) < \infty$  and therefore (see Eq. 5.3) one has  $\lim_{m \rightarrow M} \lambda(m) = 0$ , which implies  $\int_0^M dm \frac{d[m\lambda(m)]}{dm} = 0$  and as a result:

$$k = \frac{\int_0^M dm b(m)\lambda(m)}{\int_0^M dm \lambda(m)}. \quad (5.8)$$

The scale invariance of  $\lambda(m)$  can be deduced directly from Eq. 5.6 as follows.  $\lambda$  depends on  $m$  and  $M$ , i.e.,  $\lambda = \lambda(m, M)$ . It is assumed that the total mass present at  $t = 0$  is equal to 1 (Eq. 5.6 is linear in  $\lambda$  and therefore if  $\lambda$  is a solution, so is  $C\lambda$  with  $C$  an arbitrary constant). As  $M$  is the only scale in the problem, it is assumed that  $b(m, M) = \hat{b}\left(\frac{m}{M}\right)$  and Eq. 5.6 is rewritten with  $x = m/M$  as:

$$\mu \frac{d[x\lambda(Mx, M)]}{dx} = -[k + \hat{b}(x)]\lambda(Mx, M) + 4\hat{b}(2x)\lambda(2Mx, M), \quad (5.9)$$

where  $x \in [0, 1]$ . Therefore one has the solution  $\lambda(Mx, M) = \hat{\lambda}(x)/M$ , i.e.,  $\lambda(m, M) = \frac{1}{M}\hat{\lambda}\left(\frac{m}{M}\right)$  which satisfies:

$$\mu \frac{d[x\hat{\lambda}(x)]}{dx} = -[k + \hat{b}(x)]\hat{\lambda}(x) + 4\hat{b}(2x)\hat{\lambda}(2x) \quad (5.10)$$

in  $x \in [0, 1]$ , with  $\int_0^1 dx \hat{\lambda}(x) = 1$  and  $\hat{\lambda}(x) = 0 \forall x > 1$ . In particular, the size distribution  $p(m)$  can be written as:

$$p(m) = \frac{1}{m} \frac{m}{M} \hat{\lambda}\left(\frac{m}{M}\right) = \frac{1}{m} G\left(\frac{m}{M}\right). \quad (5.11)$$

Computing the average mass  $\langle m \rangle$  one finds that it is proportional to  $M$ :

$$\langle m \rangle = \int_0^M dm m p(m) = M \int_0^1 dy G(y) = cM, \quad (5.12)$$

## Chapter 5. Scaling body size fluctuations

---

so that the stationary size distribution is of the form:

$$p(m) = \frac{1}{m} F\left(\frac{m}{\langle m \rangle}\right), \quad (5.13)$$

which is precisely the scaling ansatz proposed and observed in the data. The solution to Eq. 5.10 can be written as  $\lambda(x) = \lambda_n(x)$  with  $2^{-n} \leq x \leq 2^{-n+1}$  for  $n = 1, 2, \dots$  and:

$$\lambda_n(x) = e^{-\int_{2^{-n}}^x dy h(y)} \left[ C_n + \frac{4}{\mu} \int_{2^{-n+1}}^{2x} dy \frac{\hat{b}(y)}{y} \lambda_{n-1}(y) \right], \quad (5.14)$$

where  $\lambda_0(y) = 1$ ,  $x \in [2^{-n}, 2^{-n+1}]$ ,  $h(y) = [k + \hat{b}(y) + \mu] / y$  and  $C_0$  depends on the normalization condition, i.e.,  $\int_0^1 dy \hat{\lambda}(y) = 1$ . The  $C_n$ s ( $n \geq 1$ ) are determined recursively imposing the continuity  $\hat{\lambda}_n(2^{-n+1}) = \hat{\lambda}_{n-1}(2^{-n+1})$  for  $n \geq 2$ . For instance, if  $n = 2$ :

$$\hat{\lambda}_1(x) = C_0 e^{-\int_{1/2}^x dy h(y)} \Rightarrow \hat{\lambda}_1\left(\frac{1}{2}\right) = C_0 \quad (5.15)$$

(note that  $\hat{\lambda}_1(1) = 0$  due to the singularity in Eq. 5.3),

$$\begin{aligned} \hat{\lambda}_2(x) = e^{-\int_{1/4}^x dy h(y)} & \left[ C_1 + \right. \\ & \left. + \frac{4}{\mu} C_0 \int_{1/2}^{2x} dy \frac{\hat{b}(y)}{y} e^{-\int_{1/2}^y dz h(z)} \right] \end{aligned} \quad (5.16)$$

and  $\hat{\lambda}_2(1/2) = \hat{\lambda}_1(1/2) = C_0$  allows to compute  $C_1/C_0$ . Iterating, one can compute  $C_n/C_0 \forall n \geq 1$ .

## 5.3 Methods

### 5.3.1 Protist cultures

Replicated single-species cultures of 13 different protists and unicellular algae (in this chapter all called ‘protists’) were initialized with three species of freshwater bacteria (*Serratia fonticola*, *Breviactillus brevis* and *Bacillus subtilis*) as a food resource in a climatized room at 20 ° C under constant fluorescent light three weeks before the measurements. Previous studies (Altermatt et al., 2011, 2014, Carrara et al., 2012) support that the composition and size spectra of these communities are rather stable

over this time period, thus cultures were assumed to be at their carrying capacity while the measurements were performed. Single-species cultures were grown in 500 mL Schott flasks containing a nutrient medium made of sterilized local spring water and Protozoan Pellets (Carolina Biological Supply, Burlington, NC, USA) in a density of  $0.45 \text{ g L}^{-1}$ . Here, these are referred to as 'standard conditions'. The 13 protist species were *Bodo saltans* (BOD), *Chilomonas* sp. (CHI), *Chlamydomonas* sp. (CHA), *Chlorogonium euchlorum* (CHO), *Colpidium* sp. (COL), *Cryptomonas curvata* (CRY), *Cyclidium glaucoma* (CYC), *Dexiostoma campylum* (DEX), *Euglena gracilis* (EUG), *Euglena mutabilis* (EUM), *Euplotes aediculatus* (EUP), *Paramecium bursaria* (PBU) and *Tetrahymena* sp. (TET). CHI and TET were purchased at Carolina Biological Supply, BOD, CYC, CHO, CRY, EUM and DEX were supplied by SAMS Research Services Ltd. (Oban, Scotland, UK), whereas all other species were isolated from a natural pond and used in previous studies (Altermatt et al., 2014). All species are heterotrophs and feed on bacteria, whereas CHA, CHO, EUG, EUM, EUP and PBU can also photosynthesize. The rotifer *Cephalodella* sp., also isolated from a natural pond (Altermatt et al., 2014), was employed both in competition (Fig. 5.3) and in isolation (Fig. 5.4).

Besides the single species cultures grown in standard conditions as described in the previous paragraph (which are referred to in Fig. 5.3 as CHI\_1, EUG\_1 and EUP\_1), CHI, EUG and EUP were also grown at additional temperatures and nutrients conditions, or in competition with each other. In the latter case, the two competing species were always well separated in their size spectrum, so that their two distributions did not overlap. The following conditions were studied with at least three replicates each:

- i) single species at  $15 \text{ }^{\circ}\text{C}$  (EUG\_2, EUP\_2);
- ii) single species at  $25 \text{ }^{\circ}\text{C}$  (CHI\_2, EUG\_3, EUP\_3);
- iii) competition at  $20 \text{ }^{\circ}\text{C}$ , with two species competing for resources, initialized at half of their carrying capacity in 10 mL-well plates (*Chilomonas* sp. with *Dexiostoma campylum* - CHI\_3, *Chilomonas* sp. with *Colpidium* sp. - CHI\_4, *Euglena gracilis* with the rotifer *Cephalodella* sp. - EUG\_5);
- iv) *Euglena gracilis* at a low protist medium concentration ( $0.045 \text{ g L}^{-1}$ , i.e., one tenth of the 'standard conditions' concentration - EUG\_4).

#### 5.3.2 Size distributions

Size distribution measurements were performed with a Cell Counter and Analyzer System (CASY) model TTC, Roche Applied Science. Size measurements were performed by suspending a sample taken from a protist culture in a buffer solution (CASYton) which is developed specifically to aspirate cells through a precision hole in the instrument at constant speed. To perform size measurements, capillaries with diameters

60, 150  $\mu\text{m}$  and 200  $\mu\text{m}$  were employed depending on the size of the protists under investigation. Smaller capillaries resolve better size distributions at low scale (5 – 20  $\mu\text{m}$ ), but can be blocked if larger particles pass through (it is therefore necessary to use larger capillaries to measure larger species). As a general rule, for each species the smallest capillary that enabled to unequivocally separate the protist peak from the debris in the instrument output was used. The size spectra of a sample of living cells is returned by the instrument as function of the equivalent diameter  $l$  of each cell, assuming cells to be spherical. From the definition of size distribution,  $p(l)dl$  is the fraction of individuals with equivalent diameter in  $(l, l + dl)$  and  $p(m)dm$  is the fraction of individuals with mass in  $(m, m + dm)$ .  $p(m)$  can then be calculated via the variable transformation  $p(l)dl = p(m)dm$  ( $m = \pi/6 d^3$ ). Constant density equal to the density of water is assumed (Chisholm, 1992, Fenchel, 1974) and therefore volume and mass are used without distinction. It is also assumed that size distributions do not depend on time, i.e., the cultures are in a steady state characterized by size distributions of constant shape. A peak at small sizes exists due to debris in the culture. Peaks at larger sizes are due to protists. To deconvolve the two peaks, the debris peak is fitted with an exponential decay (in a region adjacent to the peak, where data lie on a straight line in a log-linear plot) and the resulting curve is subtracted from the overall spectrum. On the right hand side of the protist peak the data were truncated when the measured frequency of a size channel was below 20 occurrences, to separate it from the noise. Noise was uniformly distributed on all size channels with frequencies of approximately 10 – 20 counts per channel, as demonstrated by measuring pure buffer solution only. For each species, several measurements of different cultures (grown in the same conditions) were collected and summed to get an ensemble average representative of the species.

### 5.3.3 Finite-size scaling distributions

For a general account of finite-size scaling (Fisher and Barber, 1972) in ecology see (Banavar et al., 1999a). General properties of finite-size scaling distributions (as in Eq. 5.1) are detailed in Banavar et al. (2007). A brief account of the derivation of the most relevant results, adapted to the case at hand, is given here. This section is dedicated to the study of the normalization conditions for size distributions of the form:

$$p_k(m) = \frac{1}{m^\Delta} F\left(\frac{m}{\langle m \rangle_k^\phi}\right), \quad (5.17)$$



where  $\phi > 0$  and, for dimensional reasons:

$$F\left(\frac{m}{\langle m \rangle_k^\phi}\right) = \frac{1}{m_0^{1-\Delta}} \hat{F}\left(\frac{1}{m_0^{1-\phi}} \frac{m}{\langle m_k \rangle^\phi}\right), \quad (5.18)$$

where  $m_0$  is the minimum mass of an organism or a cutoff in the system. From this point onwards all masses are measured in units of  $m_0$ , so  $F$  and  $\hat{F}$  coincide and  $\frac{\langle m \rangle_k}{m_0} \rightarrow \langle m \rangle_k$  is arbitrarily large. In order for the distribution in Eq. 5.17 to be normalized (i.e.,  $\int dm p_k(m) = 1$ ) one needs to make the following assumptions: i)  $F(x)$  approaches a constant when  $x \ll 1$ , ii)  $F(x)$  goes to zero sufficiently fast when  $x \gg 1$ . With these conditions one has:

$$\begin{aligned} 1 &= \int_1^\infty dm p_k(m) = \int_1^\infty dm \frac{1}{m^\Delta} F\left(\frac{m}{\langle m \rangle_k^\phi}\right) = \langle m \rangle_k^{\phi(1-\Delta)} \int_{\langle m \rangle_k^{-\phi}}^\infty dx \frac{1}{x^\Delta} F(x) = \\ &= \langle m \rangle_k^{(1-\Delta)\phi} \left[ \int_{\langle m \rangle_k^{-\phi}}^1 dx x^{-\Delta} + \int_1^\infty dx x^{-\Delta} F(x) \right] = a + b \langle m \rangle_k^{(1-\Delta)\phi}, \end{aligned} \quad (5.19)$$

where  $a$  and  $b$  are constants. Now, save corrections to the scaling (i.e., from Eq. 5.19:  $p_k(m) = m^{-\Delta} F\left(\frac{m}{\langle m \rangle_k^\phi}\right) (a + b \langle m \rangle_k^{(1-\Delta)\phi})^{-1} = m^{-\Delta} \left[ \frac{1}{a} F\left(\frac{m}{\langle m \rangle_k^\phi}\right) - \frac{b}{a^2} \langle m \rangle_k^{(1-\Delta)\phi} F\left(\frac{m}{\langle m \rangle_k^\phi}\right) + \dots \right]$ , which adds an additional term to Eq. 5.17), everything is consistent if  $a = 1$ ,  $(1-\Delta)\phi < 0$  and  $\Delta > 1$  ( $\phi > 0$ ) or if  $\Delta = 1$ , in which case one has:

$$1 = \int_1^\infty dm p_k(m) = \int_1^\infty dm \frac{1}{m} F\left(\frac{m}{\langle m \rangle_k^\phi}\right) = \int_{\langle m \rangle_k^{-\phi}}^\infty dx \frac{1}{x} F(x) \quad (5.20)$$

and two possibilities arise: *i*)  $\int_0^\infty dx \frac{F(x)}{x} = 1$ , with  $F(x) \rightarrow 0$  sufficiently fast for  $x \rightarrow 0$  (which is consistent with the data), *ii*)  $F(x) \sim_{x \rightarrow 0} (-\ln x)^{-\alpha}$ , so that  $\int_{\langle m \rangle_k^{-\phi}}^\infty dx \frac{1}{x} F(x) \sim (\ln \langle m \rangle_k)^{\max(0, 1-\alpha)}$  and, if  $\alpha > 1$  one is back to case *i*), while if  $\alpha < 1$  one has logarithmic corrections to the scaling. In fact, if  $\alpha < 1$ , one finds  $p_k(m) = \frac{1}{m} (\ln \langle m \rangle_k)^{1-\alpha} F\left(\frac{m}{\langle m \rangle_k^\phi}\right)$ .

A test for the validity of a scaling size distribution of the form  $p_k(m) = \frac{1}{m} F\left(\frac{m}{\langle m \rangle_k}\right)$  is the proportionality of successive moments ratios  $\langle m^j \rangle_k / \langle m^{j-1} \rangle_k$  ( $j > 1$ ) to the first moment  $\langle m \rangle_k$ . In fact, if  $p_k(m) = \frac{1}{m} F\left(\frac{m}{\langle m \rangle_k}\right)$ , one has:

$$\frac{\langle m^j \rangle_k}{\langle m^{j-1} \rangle_k} = \frac{\int dm m^j \frac{1}{m} F\left(\frac{m}{\langle m \rangle_k}\right)}{\int dm m^{j-1} \frac{1}{m} F\left(\frac{m}{\langle m \rangle_k}\right)} = \frac{\langle m \rangle_k^j \int dx x^j \frac{1}{x} F(x)}{\langle m \rangle_k^{j-1} \int dx x^{j-1} \frac{1}{x} F(x)} \propto \langle m \rangle_k, \quad (5.21)$$

## Chapter 5. Scaling body size fluctuations

where  $x = \frac{m}{\langle m \rangle_k}$  and  $j > 1$ . Fig. 5.2C shows successive moments ratios calculated from the data (cultures in standard conditions) and linear regressions on these data. The slopes of the linear regressions are compatible with the value of 1 (linear regressions on log-transformed data). Coefficient of determination  $R^2$  values for the regressions are:  $R^2_{\langle m^2 \rangle / \langle m \rangle} = 0.999$ ,  $R^2_{\langle m^3 \rangle / \langle m^2 \rangle} = 0.996$ ,  $R^2_{\langle m^4 \rangle / \langle m^3 \rangle} = 0.988$ . The same holds for the body size distributions of CHI, EUG and EUP in different environmental conditions. In this case, the slopes of the linear regressions are compatible with the value of 1 (linear regressions on log-transformed data). Coefficient of determination  $R^2$  values for the regressions are:  $R^2_{\langle m^2 \rangle / \langle m \rangle} = 0.998$ ,  $R^2_{\langle m^3 \rangle / \langle m^2 \rangle} = 0.987$ ,  $R^2_{\langle m^4 \rangle / \langle m^3 \rangle} = 0.962$ .

A fitting procedure suggests the viability of an analytical log-normal form for the universal size distribution, i.e.:

$$p(m) = \frac{1}{m\sqrt{2\pi\sigma^2}} e^{-\frac{\left(\ln \frac{m}{\langle m \rangle} - \mu\right)^2}{2\sigma^2}}, \quad (5.22)$$

where  $\sigma^2$  and  $\mu$  are constants,  $\langle m \rangle$  depends on the species and  $\ln$  is the natural logarithm, i.e., the logarithm to the base  $e$ . In order for the distribution Eq. 5.22 to have the scaling form  $p(m) = 1/mF(m/\langle m \rangle)$  one has to impose that  $\langle m \rangle = \int_0^\infty dm mp(m)$ , which implies  $\mu = -\sigma^2/2$ , i.e.,  $\mu$  and  $\sigma$  are not independent. Thus, the following analytical form for the universal size distribution, which depends on only one parameter,  $\sigma^2$ , is proposed:

$$p(m) = \frac{1}{m\sqrt{2\pi\sigma^2}} e^{-\frac{\left(\ln \frac{m}{\langle m \rangle} + \frac{\sigma^2}{2}\right)^2}{2\sigma^2}}. \quad (5.23)$$

The scaling function  $F(x)$  is therefore of the form:

$$F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\left(\ln x + \frac{\sigma^2}{2}\right)^2}{2\sigma^2}}, \quad (5.24)$$

as suggested by the fact that a parabola fits well the log-transformed data  $mp(m)$  versus  $m/\langle m \rangle$  (least-squares fit on log-transformed data). In order to have a good estimate of the mean of  $F$ , the fit to Eq. 5.24 was performed in the common support of at least half of the protist species, with only one parameter  $\sigma^2$ . The best estimate for the parameter is  $\sigma^2 = 0.222 \pm 0.003$  and the coefficient of determination is  $R^2 = 0.92$ . The fit of the scaling function is shown in Fig. 5.2D superimposed to the ensemble average of the experimental size distributions, showing a remarkable overlap.

The postulated universal scaling form for  $p(m)$  may arise from general dynamical considerations (Banavar et al., 2007, Clauset and Erwin, 2008, Lande, 1976). Consider ecosystem dynamics over ecological timescales. Ecological processes governing the abundances and niche occupancy of species are expected to change their characteristic size. One would expect, however, that offsprings would have a mass proportional to the mass of the parent organism (Clauset and Erwin, 2008). Thus, fluctuations in size within same species ought to be measured in the order of percent variations and the natural variable is  $x = \log(m/\bar{m})$  (Clauset and Erwin, 2008, Lande, 1976),  $\bar{m}$  being the characteristic mass of the reference species  $k$ , e.g., proportional to the mean  $\bar{m} = \alpha \langle m \rangle$  (Banavar et al., 2007, Rinaldo et al., 2002). In this framework, the results of ecological processes can be represented by a random walk in the variable  $x$ , as a fixed percent increase (decrease) of the mass corresponds to a shift to the right (left) of the variable  $x$  by a constant amount. In the simplest model, the results of ecological processes could be represented by an Ornstein-Uhlenbeck process (Uhlenbeck and Ornstein, 1930, Van Kampen, 2007). This process is a modification of a Wiener process, where the walk tends to move towards a central location, which, in the context of phenotypic evolution (Lande, 1976), has been identified as the optimum in the adaptive zone for the phenotype. A phenotypic character like body size, therefore, is expected to be distributed around a fitness optimum in this framework. The physical analogy of this process is a noisy relaxation process, e.g., a spring fluctuating around its rest length in the presence of disturbances. The fraction of organisms  $q(x, t)$  with a (log-)body mass  $x$  at time  $t$  is governed by the dynamical equation:

$$\frac{\partial q(x, t)}{\partial t} = D \frac{\partial^2 q(x, t)}{\partial x^2} + \frac{\partial [kxq(x, t)]}{\partial x}, \quad (5.25)$$

where  $k$  is a constant,  $k > 0$ . The stationary solution is obtained by setting the rate of change of  $q(x, t)$  to zero and is known to be gaussian (Uhlenbeck and Ornstein, 1930, Van Kampen, 2007):

$$q(x) = \sqrt{\frac{k}{\pi D}} e^{-\frac{kx^2}{2D}} \quad (5.26)$$

One obtains then, for the mass distribution  $p(m)$ :

$$p(m) = q(x) \frac{dx}{dm} = \frac{1}{m} \sqrt{\frac{k}{2\pi D}} e^{-\frac{k}{2D} \left( \ln \frac{m}{\langle m \rangle} - \ln \alpha \right)^2}, \quad (5.27)$$

i.e., a log-normal distribution of mass. Imposing Eq. 5.27 to have mean  $\langle m \rangle$  one finds

$\alpha = \exp[-D/(2k)]$  and the distribution of size is therefore:

$$p(m) = q(x) \frac{dx}{dm} = \frac{1}{m} \sqrt{\frac{k}{2\pi D}} e^{-\frac{k}{2D} \left( \ln \frac{m}{\langle m \rangle} + \frac{D}{2k} \right)^2}, \quad (5.28)$$

i.e., a log-normal distribution of mass with mean  $\langle m \rangle$ , as in Eq. 5.23. Therefore the scaling function  $F$  in Fig. 5.2 is  $F(x) = \sqrt{\frac{k}{2\pi D}} \exp \left[ -\frac{k}{2D} \left( \ln x + \frac{D}{2k} \right)^2 \right]$ .

One might wonder whether the size distribution obtained in Eq. 5.28 is in agreement with the model of cellular growth and division investigated above. The cellular growth and division model, as treated in the previous section, assumes the existence of a maximum mass  $M$  and allows to study the scaling properties of the stationary size distribution. It is possible, however, to relax this hypothesis allowing the cells to assume all masses in the range  $[0, \infty]$  and obtain an implicit relation for the stationary size distribution  $p(m)$  (Rading et al., 2011), which allows to compute the asymptotic behavior of the distribution for large mass, i.e.,  $m \rightarrow \infty$ . In the notation of the previous section the size distribution, for large  $m$ , satisfies the relation:

$$p(m) \rightarrow \frac{1}{m} \exp \left( - \int_{\alpha}^{m/\langle m \rangle} dy \frac{k + d(y)}{\mu y} \right), \quad (5.29)$$

which behaves as a log-normal if one assumes a division rate  $d(y)$  increasing logarithmically with size, i.e.,  $d(y) \propto \ln y$ .

### 5.3.4 Data collapses

Data collapse is a tool widely used in statistical physics to establish scaling laws and extract information on their exponents (Barenblatt, 1983). Traditionally, the procedure to produce a data collapse is to rely on the direct visualization of it and eyeballing the exponent which gives the best collapse. A less subjective method was proposed (Bhattacharjee and Seno, 2001), which introduces a measure (error functional  $E$ , insets of Figs. 5.2, 5.3) to quantify the goodness of a collapse. Let  $\Delta$  be the exponent that is tuned to find the best collapse:  $E(\Delta)$  is the cumulative area enclosed between all pairs of curves that one tries to collapse, within their common support, for the value  $\Delta$  of the exponent. The value  $\Delta^*$  of the exponent which minimizes  $E(\Delta)$  is taken as the best estimate for the exponent, i.e., the smaller the area, the best the collapse. Errors are associated to the determination of  $\Delta^*$  and are obtained from the width of the minimum. Further details can be found in Bhattacharjee and Seno (2001).

### 5.3.5 Community size spectra

In this section it is shown how a power law community size spectrum arises as a sum of single species size distributions of finite-size scaling form (Rinaldo et al., 2002). Assume that the size distribution of species  $k$  is of the form:

$$p_k(m) = \frac{1}{m} F\left(\frac{m}{\langle m \rangle_k}\right). \quad (5.30)$$

Let  $N_k$  be the stationary abundance of species  $k$  in an ecosystem (i.e.,  $N_k = N_k(t \rightarrow \infty)$ ) and  $S$  is the total number of species. The community size spectrum is defined as:

$$f(m) = \sum_{k=1}^S N_k p_k(m) / \sum_{k=1}^S N_k. \quad (5.31)$$

It is assumed, supported by a number of observations (Damuth, 1981), that the population abundance of the  $k$ th species scales as:

$$N_k \propto \langle m \rangle_k^\alpha, \quad (5.32)$$

where  $\alpha < 0$  implies that the total number of organisms decreases with increasing typical size. From Eq. 5.30 and Eq. 5.32 one has that:

$$f(m) \propto \sum_{k=1}^S N_k p_k(m) \propto \sum_{k=1}^S \langle m \rangle_k^\alpha m^{-1} F\left(\frac{m}{\langle m \rangle_k}\right). \quad (5.33)$$

Let  $g(\bar{m})$  be the fraction of species of typical size  $\bar{m}$ . The above equation can be rewritten, treating  $\langle m \rangle_k$  as a continuous variable for easiness of computation, as:

$$f(m) \propto \frac{1}{m} \int d\bar{m} g(\bar{m}) \bar{m}^\alpha F\left(\frac{m}{\bar{m}}\right) \propto \int dx g(xm) x^\alpha m^\alpha F\left(\frac{1}{x}\right). \quad (5.34)$$

Theoretical predictions from a scaling macroecological framework (Banavar et al., 2007, Rinaldo et al., 2002) and data (Fenchel, 1993, Marquet et al., 2005, May, 1988) suggest a pure power-law behavior for  $g(\bar{m})$ :

$$g(\bar{m}) \propto \frac{1}{\bar{m}^\beta}, \quad (5.35)$$

which, because of normalization, is assumed to hold between an upper and lower cutoff. One then has for the size spectrum  $f(m)$ :

$$f(m) \propto m^{\alpha-\beta} \int dx x^{\alpha-\beta} F\left(\frac{1}{x}\right) \propto m^{\alpha-\beta}, \quad (5.36)$$

which has the form of a power law. In the case of a limited range of sizes one might argue that the number of species  $S$  within the range of sizes investigated could be assumed as constant to first order. This, of course, is the particular case for which  $\beta = 0$ .

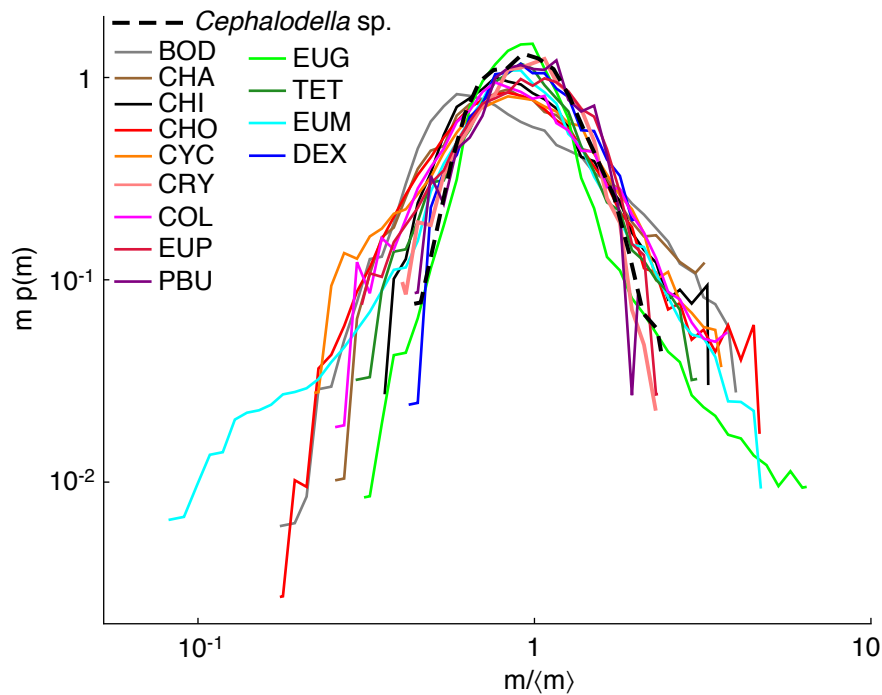
Overall, it is clear that to obtain a scaling community size spectrum (Eq. 5.31), a necessary condition is an adaptive fine tuning of the specific abundances. This is epitomized by the relation Eq. 5.32, which in turn implies thinning relations that are recurrent in the literature of macroecological empirical laws (Banavar et al., 2007, Southwood et al., 2006).

### 5.4 Discussion

The detailed identification of the scaling function  $F$  in Eq. 5.2 is interesting but inessential for the tenet of this investigation as the collapse of the distributions suffices in documenting the universality sought after. However, a log-normal functional form for  $p(m)$  proves admissible and rooted in a theoretical framework for the time-evolution of the distribution of body sizes in ecological timescales (see inset in Fig. 5.2D). In this context, the size distribution of organisms of a given species is the stationary distribution of an Ornstein-Uhlenbeck process (Uhlenbeck and Ornstein, 1930) in the variable  $x = \log(m/\langle m \rangle)$  (Clauset and Erwin, 2008, Lande, 1976). A log-normal form for  $p(m)$  can also be recovered as a particular case of the physiological model cited above, so the two models are not mutually exclusive (Methods). Recently, models of cell growth and division with different assumptions on the timings of cell duplication have also been shown to predict the log-normality of cell distributions in bacteria (Amir, 2014) and the scaling behavior of other relevant properties of cell division such as the initial size at division and doubling times (Kennard et al., 2014) have been substantiated experimentally. These works give further support to the claim that the physiological processes of growth and reproduction are responsible for the observed scaling behavior of body size distributions.

A yet unproven but reasonable ansatz would posit that this behavior might apply to multicellular or arbitrarily complex organisms as well, resulting in even broader validity. An indication supporting this statement is the experimental size distribution of a multicellular organism that was measured with the same methods in the same laboratory conditions. In fact, it was found that the size distribution of a multicellular species (*Cephalodella* sp.) showed a very good collapse with the protist size distributions once rescaled according to Eq. 5.2, see Fig.5.4.

The observation of more than 20 orders of magnitude in organismic sizes in natural



**Figure 5.4:** Collapse of rescaled size distributions as in Fig. 5.2B, with superimposed the size distribution of the multicellular organism *Cephalodella* sp. (black dashed line), whose size distribution collapses well with those of unicellular protists. Colors as in Fig. 5.2A.

ecosystems (Kleiber, 1947, Sheldon et al., 1972) leads to the conclusion that there has been little long-term impediment to the development, on evolutionary timescales, of any particular size. On ecological timescales, however, a characteristic size emerges as a fundamental property of a species, determined by biological constraints and by biotic and abiotic interactions. Such characteristic size, in turn, modulates the entire size distribution of the species.

The results reported here decouple the effects of biotic and abiotic interactions, which regulate abundances and can affect a species' mean size, from the individuals' physiology, which shapes intra-species size distributions. The replicated, controlled experiments performed here corroborate and extend comparative field findings of marine microbial size spectra to broader size and taxonomic diversity (Rinaldo et al., 2002). It is speculated here that such behavior may extend over broader domains. Then, theoretical linkages of diverse empirical macroecological relationships, traditionally treated as independent (Banavar et al., 2007, Southwood et al., 2006), would be substantiated. In fact, because single-species size distributions would be characterized by specific mean values and variances, a precise requirement would be cast on the number of species existing at stationarity within a niche of size and on the related abundances, in order to produce community size spectra that lack characteristic

## Chapter 5. Scaling body size fluctuations

---

scales. Ecological interactions among species would consequently adapt to produce thinning laws, i.e., to control the relative abundance of species given their characteristic sizes (Damuth, 1981, May, 1988). The fact that these thinning laws have been shown to be robust to perturbations further emphasizes their universal character (Marquet et al., 1990).

Finally, a distribution of the form in Eq. 5.2 implies that the variance of the species' sizes increases quadratically with the mean size. As the characteristic mass of a species is frequently adopted as the independent variable in allometric scaling laws (Banavar et al., 2002, 2010, 1999b, Brown et al., 2004, Brown and West, 2000, Cohen et al., 2012, Damuth, 1981, Fenchel, 1993, Kleiber, 1947, May, 1988, McMahon and Bonner, 1983, West et al., 1997), its increasing variance must impact the scaling of the dependent one, such as metabolic rates in Kleiber's law. This investigation might thus pose the basis for a re-examination of allometric relations, by considering appropriate fluctuations in both the dependent (Dodds et al., 2001, Labra et al., 2007) (metabolic rate) and the independent (mass) variables.



## Conclusions and perspectives

Fluctuations are ubiquitous in nature and any investigation of their role in ecological patterns and processes is by necessity limited in its scope. This thesis focuses on the characterization and impact of the main sources of fluctuations on relevant ecological patterns and processes. The main conclusions are summarized here and perspectives for future research are outlined.

The role of demographic and environmental stochasticity on biological invasions was studied in Chapters 1, 2 and 3, through laboratory experiments with *Tetrahymena* sp. and *E. gracilis* and theoretically in the framework of stochastic partial differential equations. The study of front propagation in uniform environments (Chapter 1) allowed establishing experimentally a quantitative link between the processes acting at the local scale, movement and growth, and the colonization dynamics at the larger scale. Demographic stochasticity acting the front of the invasion wave, where population densities are low, was argued to induce the observed fluctuations of the front across replicated invasions in a uniform landscape. Accordingly, the variability in the front position measured experimentally was quantitatively comparable with the prediction of the Fisher-Kolmogorov equation modified to include demographic stochasticity, whose strength was estimated locally via the measured growth curves.

In Chapter 2, the phototactic movement of *E. gracilis* was characterized experimentally and a Keller-Segel framework was derived and shown to reproduce both the stationary accumulation patterns in the presence of light gradients and the temporal dynamics of accumulation and relaxation following the removal of light. The phototactic response of *E. gracilis* was shown to be described by a generalized receptor law that accounts for both positive and negative phototaxis. Whereas other forms of taxis received much attention in the literature (Tindall et al., 2008), Chapter 2 appears as the first quantitative investigation and modeling of the phototactic response function capable of reconciling both positive and negative phototaxis within the same mathematical framework.

The generalized Keller-Segel model derived in Chapter 2 and the stochastic Fisher-

## Conclusions and perspectives

---

Kolmogorov equation investigated in Chapter 1 were combined in Chapter 3 to study the propagation of invasion fronts in linear landscapes endowed with heterogeneous distributions of resources. The joint effect of demographic stochasticity and environmental heterogeneity was investigated numerically across different autocorrelation lengths of the resource distribution. Demographic stochasticity was shown to be critical in slowing-down the propagating front for increasing resource autocorrelation length. The directed movement of individuals towards favorable regions of the landscape was shown to contribute to such slowing-down. Experiments with *E. gracilis* in landscapes with manipulated light intensity profiles confirmed the prediction that larger resource autocorrelation lengths induce a slowing-down of the propagating front. The results of Chapter 3 thus suggest that the local extinctions caused by demographic stochasticity at the front of the invasion and the heterogeneity of the resource distribution, characterized in terms of its autocorrelation length, critically affect the propagation of biological invasions.

Chapters 1 and 3 extend the current understanding of biological invasions by recognizing and investigating experimentally the role of fluctuations on the dynamics of propagating fronts, a subject that suffers an acknowledged lack of experimentation. Demographic stochasticity is found to induce fluctuations of the front position across replicated invasions in uniform landscapes (Chapter 1). The role of the temporal autocorrelation of environmental fluctuations in shaping ecological dynamics is widely recognized in the literature. This study provides evidence for the role of the spatial autocorrelation of environmental heterogeneity on spatial ecological processes such as biological dispersal.

In Chapter 4, population fluctuations induced by temporal environmental stochasticity were studied in the framework of Taylor's law of fluctuation scaling. Results from large deviations theory were used to show that, for any ecosystem of finite size, the spatial sample variance of population abundance scales with the mean sample abundance with the exponent  $b = 2$  irrespectively of the details of the population growth model. The temporal autocorrelation of environmental stochasticity only affects the scaling exponent of Taylor's law expressed in terms of the population cumulants. Understanding whether widely reported macroecological patterns such as the species-area relationship are statistical artifacts rather than the outcome of ecological processes was listed as one of the 100 fundamental ecological questions cited in the introduction to this thesis. The results outlined in Chapter 5 suggest that the widespread observation of the TL scaling exponent  $b = 2$  may be a sampling artifact rather than an emergent property of ecosystems.

Finally, Chapter 5 investigated the heterogeneity of body sizes in the search for univer-

sal properties across levels of biological organization. Laboratory experiments with protists suggest that intra-specific body size distributions may be controlled by fundamental physiological processes, causing such distributions to be identical across very different species covering over four orders of magnitude in body mass and belonging to four different taxa. The observation that the variance of intra-specific body size distributions scales with the mean body size highlights another occurrence of Taylor's law. Because body size is one of the most relevant ecological traits, the characterization of intra-specific body size distributions provided here may have implications for the several scaling relationships where body size appears as the independent variable. In particular, our study invites further investigation of the joint scaling of body size and metabolic rates accounting for the scaling fluctuations of both the dependent and the independent variables.

Further developments will deal with theoretical and experimental studies of geometrical heterogeneities affecting the environmental matrix, like for example studies of biological invasions in network substrates. Theoretical predictions for the speed of fronts propagating in dendritic networks are available and will be tested experimentally. Furthermore, the experimental system employed in Chapters 2 and 3 could be readily used to impose temporal fluctuations of the environment during the invasion process, possibly autocorrelated both in time and space. A comprehensive understanding of the effect of heterogeneities and fluctuations on biological invasions will have important practical implications, because they influence predictions on the speed of spread of invasive species or disease propagules, especially along dendritic habitats such as rivers, waterways or mountain-ridges. The results derived in Chapter 4 may be investigated in protist microcosms, for example generating Markovian environments with different first-order temporal autocorrelations by controlling the temperature experienced by the microcosms. The emergence of the Taylor's law exponent  $b = 2$  as a sampling artifact depending on the number of replicates could therefore be explored experimentally. Finally, the characterization of body size fluctuations (Chapter 5) will be complemented with the study of metabolic rate fluctuations by measuring the joint distribution of body size and metabolic rates. The implications of scaling metabolic rate and body size fluctuations on linked macroecological laws will then be explored within the framework of finite-size scaling.



## Bibliography

- Adl, S. M. et al. (2012). The revised classification of eukaryotes. *Journal of Eukaryotic Microbiology*, 59:429–493.
- Adler, J., Hazelbauer, G. L., and Dahl, M. M. (1973). Chemotaxis toward sugars in *Escherichia coli*. *Journal of Bacteriology*, 115:824–847.
- Altermatt, F., Bieger, A., Carrara, F., Rinaldo, A., and Holyoak, M. (2011). Effects of connectivity and recurrent local disturbances on community structure and population density in experimental metacommunities. *PLoS ONE*, 6:e19525.
- Altermatt, F. et al. (2014). Big answers from small worlds: a user's guide for protist microcosms as a model system in ecology and evolution. *Methods in Ecology and Evolution*, published ahead of print, doi:10.1111/2041-210x.12312.
- Amir, A. (2014). Cell size regulation in bacteria. *Physical Review Letters*, 112:208102.
- Anderson, R. M., Gordon, D. M., Crawley, M. J., and Hassell, M. P. (1982). Variability in the abundance of animal and plant species. *Nature*, 296:245–248.
- Andow, D. A., Kareiva, P. M., Levin, S. A., and Okubo, A. (1990). Spread of invading organisms. *Landscape Ecology*, 4:177–188.
- ASTM International (2008). Standard tables for reference solar spectral irradiances, direct normal and hemispherical on 37° tilted surface. ASTM Standard G173 (ASTM International, West Conshohocken, PA).
- Azam, F. (1998). Microbial control of oceanic carbon flux: the plot thickens. *Science*, 280:694–696.
- Bailey, D. J., Otten, W., and Gilligan, C. A. (2000). Saprotrophic invasion by the soil-borne fungal plant pathogen *Rhizoctonia solani* and percolation thresholds. *New Phytologist*, 146:535–544.

## Bibliography

---

- Bailey, N. T. J. (1964). *The Elements of Stochastic Processes With Applications to the Natural Sciences*. Wiley, New York.
- Bak, P. (1997). *How Nature works. The science of self-organized criticality*. Copernicus-Springer, Berlin.
- Bak, P., Tang, C., and Wiesenfeld, K. (1987). Self-organized criticality: An explanation of  $1/f$  noise. *Physical Review Letters*, 59:381–384.
- Banavar, J. R., Damuth, J., Maritan, A., and Rinaldo, A. (2002). Supply-demand balance and metabolic scaling. *Proceedings of the National Academy of Sciences of the United States of America*, 99:10506–10509.
- Banavar, J. R., Damuth, J., Maritan, A., and Rinaldo, A. (2007). Scaling in ecosystems and the linkage of macroecological laws. *Physical Review Letters*, 98:068104.
- Banavar, J. R. et al. (2010). A general basis for quarter-power scaling in animals. *Proceedings of the National Academy of Sciences of the United States of America*, 107:15816–15820.
- Banavar, J. R., Green, J. L., Harte, J., and Maritan, A. (1999a). Finite size scaling in ecology. *Physical Review Letters*, 83:4212–4214.
- Banavar, J. R., Maritan, A., and Rinaldo, A. (1999b). Size and form in efficient transportation networks. *Nature*, 399:130–132.
- Barbara, G. M. and Mitchell, J. G. (2003). Bacterial tracking of motile algae. *FEMS Microbiology Ecology*, 44:79–87.
- Barenblatt, G. I. (1983). *Introduction to Phase Transitions and Critical Phenomena*. Cambridge University Press, Cambridge.
- Bascompte, J. and Solé, R. V. (1995). Rethinking complexity - modeling spatiotemporal dynamics in ecology. *Trends in Ecology & Evolution*, 10:361–366.
- Behrenfeld, M. J. et al. (2006). Climate-driven trends in contemporary ocean productivity. *Nature*, 444:752–755.
- Benincà, E., Dakos, V., Van Nes, E. H., Huisman, J., and Scheffer, M. (2011). Resonance of plankton communities with temperature fluctuations. *The American Naturalist*, 178:E85–E95.
- Benton, T. G., Lapsley, C. T., and Beckerman, A. P. (2001). Population synchrony and environmental variation: an experimental demonstration. *Ecology Letters*, 4:236–243.

- Benton, T. G., Solan, M., Travis, J. M. J., and Sait, S. M. (2007). Microcosm experiments can inform global ecological problems. *Trends in Ecology & Evolution*, 22:516–521.
- Berg, H. C. (1993). *Random Walks in Biology*. Princeton University Press, Princeton.
- Bergelson, J., Newman, J. A., and Floresroux, E. M. (1994). Rates of weed spread in spatially heterogeneous environments. *Ecology*, 74:999–1011.
- Bertuzzo, E., Casagrande, R., Gatto, M., Rodriguez-Iturbe, I., and Rinaldo, A. (2010). On spatially explicit models of cholera epidemics. *Journal of The Royal Society Interface*, 7:321–333.
- Bertuzzo, E. et al. (2011). Spatial effects on species persistence and implications for biodiversity. *Proceedings of the National Academy of Sciences of the United States of America*, 108:4346–4351.
- Bertuzzo, E., Maritan, A., Gatto, M., Rodriguez-Iturbe, I., and Rinaldo, A. (2007). River networks and ecological corridors: Reactive transport on fractals, migration fronts, hydrochory. *Water Resources Research*, 43:W04419.
- Bhattacharjee, S. M. and Seno, F. (2001). A measure of data collapse for scaling. *Journal of Physics A-Mathematical and General*, 34:6375–6380.
- Bhaya, D. (2004). Light matters: phototaxis and signal transduction in unicellular cyanobacteria. *Molecular Microbiology*, 53:745–754.
- Bonachela, J., Muñoz, M. A., and Levin, S. A. (2012). Patchiness and demographic noise in three ecological examples. *Journal of Statistical Physics*, 148:723–739.
- Botet, R. et al. (2001). Universal fluctuations in heavy-ion collisions in the Fermi energy domain. *Physical Review Letters*, 86:3514–3517.
- Boyce, D. G., Lewis, M. R., and Worm, B. (2010). Global phytoplankton decline over the past century. *Nature*, 466:591–596.
- Bradshaw, A. D. (1965). Evolutionary significance of phenotypic plasticity in plants. *Advances in Genetics*, 13:115–155.
- Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., and West, G. B. (2004). Toward a metabolic theory of ecology. *Ecology*, 85:1771–1789.
- Brown, J. H. and West, G. B. (2000). *Scaling in Biology*. Oxford University Press, New York.

## Bibliography

---

- Burkart, U. and Häder, D. P. (1980). Phototactic attraction in light trap experiments: A mathematical model. *Journal of Mathematical Biology*, 10:257–269.
- Burnham, K. P. and Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer-Verlag, New York, 2 edition.
- Cadotte, M. W., Drake, J. A., and Fukami, T. (2005). Constructing nature: Laboratory models as necessary tools for investigating complex ecological communities. *Advances in Ecological Research*, 37:333–353.
- Camacho, J. and Solé, R. V. (2001). Scaling in ecological size spectra. *Europhysics Letters*, 55:774–780.
- Campos, D., Fort, J., and Méndez, V. (2006). Transport on fractal river networks: application to migration fronts. *Theoretical Population Biology*, 69:88–93.
- Campos, D. and Méndez, V. (2005). Reaction-diffusion wave fronts on comblike structures. *Physical Review E*, 71:051104.
- Carrara, F., Altermatt, F., Rodriguez-Iturbe, I., and Rinaldo, A. (2012). Dendritic connectivity controls biodiversity patterns in experimental metacommunities. *Proceedings of the National Academy of Sciences of the United States of America*, 109:5761–5766.
- Cavender-Bares, K. K., Rinaldo, A., and Chisholm, S. W. (2001). Microbial size spectra from natural and nutrient enriched ecosystems. *Limnology and Oceanography*, 46:778–789.
- Celani, A. and Vergassola, M. (2010). Bacterial strategies for chemotaxis response. *Proceedings of the National Academy of Sciences of the United States of America*, 107:1391–1396.
- Chassot, E. et al. (2010). Global marine primary production constrains fisheries catches. *Ecology Letters*, 13:495–505.
- Chesson, P. and Huntly, N. (1997). The roles of harsh and fluctuating conditions in the dynamics of ecological communities. *The American Naturalist*, 150:519–553.
- Chisholm, S. W. (1992). Phytoplankton size. In Falkowski, A. and Woodhead, A., editors, *Primary Productivity and Biogeochemical Cycles in the Sea*, volume 43, pages 213–237, New York.
- Clauset, A. and Erwin, D. H. (2008). The evolution and distribution of species body size. *Science*, 321:399–401.



- Cohen, J. E. (2013). Taylor's power law of fluctuation scaling and the growth-rate theorem. *Theoretical Population Biology*, 88:94–100.
- Cohen, J. E. (2014a). Stochastic population dynamics in a markovian environment implies Taylor's power law of fluctuation scaling. *Theoretical Population Biology*, 93:30–37.
- Cohen, J. E. (2014b). Taylor's law and abrupt biotic change in a smoothly changing environment. *Theoretical Ecology*, 7:77–86.
- Cohen, J. E., Jonsson, T., and Carpenter, S. R. (2003). Ecological community description using the food web, species abundance, and body size. *Proceedings of the National Academy of Sciences of the United States of America*, 100:1781–1786.
- Cohen, J. E., Plank, M. J., and Law, R. (2012). Taylor's law and body size in exploited marine ecosystems. *Ecology and evolution*, 2:3168–3178.
- Cohen, J. E., Xu, M., and Schuster, W. S. F. (2013). Stochastic multiplicative population growth predicts and interprets Taylor's power law of fluctuation scaling. *Proceedings of the Royal Society B-Biological Sciences*, 280:20122955.
- Crawley, M. J. (2007). *The R Book*. Wiley, Chichester, 1st edition.
- Dahlstedt, J. and Jensen, H. J. (2005). Fluctuation spectrum and size scaling of river flow and level. *Physica A: Statistical Mechanics and its Applications*, 348:596–610.
- Damuth, J. (1981). Population-density and body size in mammals. *Nature*, 290:699–700.
- Dembo, A. and Zeitouni, O. (2009). *Large Deviations Techniques and Applications*. Stochastic Modelling and Applied Probability. Springer, New York.
- Den Boer, P. J. (1977). *Dispersal Power and Survival*. Miscellaneous Papers 14. Landbouwhogeschool Wageningen.
- den Hollander, F. (2008). *Large Deviations*. Fields Institute monographs. American Mathematical Society, Providence.
- Diehl, S. (2002). Phytoplankton, light, and nutrients in a gradient of mixing depths: theory. *Ecology*, 83:386–398.
- Diekmann, O., Lauwerier, H. A., Aldenberg, T., and Metz, J. A. J. (1983). Growth, fission and the stable size distribution. *Journal of Mathematical Biology*, 18:135–148.

## Bibliography

---

- Dodds, P. S., Rothman, D. H., and Weitz, J. S. (2001). Re-examination of the "3/4-law" of metabolism. *Journal of Theoretical Biology*, 209:9–27.
- Doering, C. R., Sargsyan, K. V., and Smereka, P. (2005). A numerical method for some stochastic differential equations with multiplicative noise. *Physics Letters A*, 344:149–155.
- Dornic, I., Chaté, H., and Muñoz, M. A. (2005). Integration of Langevin equations with multiplicative noise and the viability of field theories for absorbing phase transitions. *Physical Review Letters*, 94:100601.
- Duncan, A. B., Gonzalez, A., and Kaltz, O. (2013). Stochastic environmental fluctuations drive epidemiology in experimental host-parasite metapopulations. *Proceedings of the Royal Society B-Biological Sciences*, 280:20131747.
- Eisler, Z. (2008). Fluctuation scaling in complex systems: Taylor's law and beyond. *Advances in Physics*, 57:89–142.
- Eisler, Z. and Kertész, J. (2005). Random walks on complex networks with inhomogeneous impact. *Physical Review E*, 71:057104.
- Eisler, Z. and Kertész, J. (2006). Scaling theory of temporal correlations and size-dependent fluctuations in the traded value of stocks. *Physical Review E*, 73:046109.
- Elton, C. S. (1958). *The Ecology of Invasions by Animals and Plants*. Methuen, London.
- Falkowski, P. G. (1998). Biogeochemical controls and feedbacks on ocean primary production. *Science*, 281:200–206.
- Feller, W. (1957). *An Introduction to Probability Theory and its Applications, Vol. 1*. Wiley, New York.
- Fenchel, T. (1974). Intrinsic rate of natural increase - relationship with body size. *Oecologia*, 14:317–326.
- Fenchel, T. (1993). There are more small than large species? *Oikos*, 68:375–378.
- Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*, 281:237–240.
- Fisher, M. E. and Barber, M. N. (1972). Scaling theory for finite-size effects in critical region. *Physical Review Letters*, 28:1516–1519.

- Fisher, R. A. (1937). The wave of advance of advantageous genes. *Annals of Eugenics*, 7:355–369.
- Fontaine, C. and Gonzalez, A. (2005). Population synchrony induced by resource fluctuations and dispersal in an aquatic microcosm. *Ecology*, 86:1463–1471.
- Forster, J. and Hirst, A. G. (2012). Warming-induced reductions in body size are greater in aquatic than terrestrial species. *Proceedings of the National Academy of Sciences of the United States of America*, 109:19310–19314.
- Fort, J. (2012). Synthesis between demic and cultural diffusion in the neolithic transition in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, 109:18669–18673.
- Fort, J. and Solé, R. V. (2013). Accelerated tumor invasion under non-isotropic cell dispersal in glioblastomas. *New Journal of Physics*, 15:055001.
- Fox, J. W., Vasseur, D. A., Hausch, S., and Roberts, J. (2011). Phase locking, the Moran effect and distance decay of synchrony: experimental tests in a model system. *Ecology Letters*, 14:163–168.
- Fronczak, A. and Fronczak, P. (2010). Origins of Taylor's power law for fluctuation scaling in complex systems. *Physical Review E*, 81:066112.
- Fürth, R. (1920). Die Brownsche Bewegung bei Berücksichtigung einer Persistenz der Bewegungsrichtung. Mit Anwendungen auf die Bewegung lebender Infusorien. *Zeitschrift für Physik*, 2:244–256.
- Garcia, X., Rifaï, S., and Peyla, P. (2013). Light Control of the Flow of Phototactic Microswimmer Suspensions. *Physical Review Letters*, 110:138106.
- García-Carreras, B. and Reuman, D. C. (2011). An empirical link between the spectral colour of climate and the spectral colour of field populations in the context of climate change. *Journal of Animal Ecology*, 80:1042–1048.
- Gardiner, C. (2006). *Stochastic Methods*. Springer, Berlin, 4 edition.
- Gardner, R. H., Kemp, W. M., Kennedy, V. S., and Petersen, J. E. (2001). *Scaling Relations in Experimental Ecology*. Columbia University Press, New York.
- Gause, G. F. (1934). *The Struggle for Existence*. Williams and Wilkins, Baltimore.
- Gillespie, D. T. (2000). The chemical Langevin equation. *The Journal of Chemical Physics*, 113:297–306.

## Bibliography

---

- Giometto, A., Altermatt, F., Carrara, F., Maritan, A., and Rinaldo, A. (2013). Scaling body size fluctuations. *Proceedings of the National Academy of Sciences of the United States of America*, 110:4646–4650.
- Giometto, A., Rinaldo, A., Carrara, F., and Altermatt, F. (2014). Emerging predictable features of replicated biological invasion fronts. *Proceedings of the National Academy of Sciences of the United States of America*, 111:297–301.
- Gonzalez, A. and Holt, R. D. (2002). The inflationary effects of environmental fluctuations in source-sink systems. *Proceedings of the National Academy of Sciences of the United States of America*, 99:14872–14877.
- Greenwood, J. and Craig, P. (2014). A simple numerical model for predicting vertical distribution of phytoplankton on the continental shelf. *Ecological Modelling*, 273:165–172.
- Grosholz, E. D. (1996). Contrasting rates of spread for introduced species in terrestrial and marine systems. *Ecology*, 77:1680–1686.
- Häder, D.-P. and Griebenow, K. (1988). Orientation of the green flagellate, *Euglena gracilis*, in a vertical column of water. *FEMS Microbiology Ecology*, 53:159–167.
- Häder, D.-P. and Lebert, M. (1998). The photoreceptor for phototaxis in the photosynthetic flagellate *Euglena gracilis*. *Photochemistry and Photobiology*, 8:260–265.
- Hallatschek, O., Hersen, O., Ramanathan, S., and Nelson, D. R. (2007). Genetic drift at expanding frontiers promotes gene segregation. *Proceedings of the National Academy of Sciences of the United States of America*, 104:19926–19930.
- Hallatschek, O. and Korolev, K. (2009). Fisher waves in the strong noise limit. *Physical Review Letters*, 103:108103.
- Hanski, I. (1982). On patterns of temporal and spatial variation in animal populations. *Annales Zoologici Fennici*, 19:21–37.
- Hanski, I. (1999). *Metapopulation Ecology*. Oxford Series in Ecology and Evolution. Oxford University Press, Oxford.
- Hanski, I. and Ovaskainen, O. (2000). The metapopulation capacity of a fragmented landscape. *Nature*, 404:755–758.
- Harris, T. (1963). *The Theory of Branching Processes*. Springer, Berlin.
- Hastings, A. et al. (2005). The spatial spread of invasions: new developments in theory and evidence. *Ecology Letters*, 8:91–101.

- Hays, G. C. (2003). A review of the adaptive significance and ecosystem consequences of zooplankton diel vertical migrations. *Hydrobiologia*, 503:163–170.
- Hazelbauer, G. L., Berg, H. C., and Matsumura, P. (1993). Bacterial motility and signal transduction. *Cell*, 73:15–22.
- Hidalgo, J. et al. (2014). Information-based fitness and the emergence of criticality in living systems. *Proceedings of the National Academy of Sciences of the United States of America*, 111:10095–10100.
- Holling, C. S. (1992). Cross-scale morphology, geometry and dynamics of ecosystems. *Ecological Monographs*, 62:447–502.
- Holmes, E. E. (1993). Are diffusion models too simple? A comparison with telegraph models of invasion. *The American Naturalist*, 142:779–795.
- Holyoak, M. and Lawler, S. P. (2005). The contribution of laboratory experiments on protists to understanding population and metapopulation dynamics. In *Advances in ecological research 37: Population dynamics and laboratory ecology*, volume 37, pages 245–271. Academic Press, London.
- Holyoak, M., Leibold, M. A., and Holt, R. D. (2005). *Metacommunities: Spatial Dynamics and Ecological Communities*. The University of Chicago Press, Chicago.
- Hubbell, S. P., editor (2001). *A Unified Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton.
- Huete-Ortega, M., Cermeño, P., Calvo-Díaz, A., and Maraño, E. (2012). Isometric size-scaling of metabolic rate and the size abundance distribution of phytoplankton. *Proceedings of the Royal Society B-Biological Sciences*, 279:1815–1823.
- Hurn, A. S., Jeisman, J. I., and Lindsday, K. A. (2007). Transitional densities of diffusion processes: a new approach to solving the Fokker-Planck equation. *Journal of Derivatives*, 14:86–94.
- Itoh, A. (2004). *Euglena* motion control by local illumination. In Kato N., Ayers J., M. H., editor, *Bio-mechanisms of Swimming and Flying*, chapter 2, pages 13–26. Springer, Tokio.
- Itoh, A. and Tamura, W. (2008). Object manipulation by a formation-controlled *Euglena* group. In Kato, N. and Kamimura, S., editors, *Bio-mechanisms of Swimming and Flying*, chapter 4, pages 41–52. Springer, Tokio.

## Bibliography

---

- Jeffrey, S. W. (1984). Responses of unicellular marine plants to natural blue-green light environments. In Senger, H., editor, *Blue Light Effects in Biological Systems*, Proceedings in Life Sciences, pages 497–508, Berlin. Springer.
- Jékely, G. (2009). Evolution of phototaxis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364:2795–2808.
- Jékely, G. et al. (2008). Mechanism of phototaxis in marine zooplankton. *Nature*, 456:395–399.
- Jessup, C. M., Kassen, R., Forde, S. E., Kerr, B., Buckling, A., Rainey, P. B., and Bohannan, B. J. M. (2004). Big questions, small worlds: microbial model systems in ecology. *Trends in Ecology & Evolution*, 19:189–197.
- Jiang, J., DeAngelis, D. L., Zhang, B., and Cohen, J. E. (2014). 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.
- Jørgensen, B. (1987). Exponential dispersion models. *Journal of the Royal Statistical Society: Series B*, 49:127–162.
- Kalyuzhny, M. et al. (2014). Temporal fluctuation scaling in populations and communities. *Ecology*, 95:1701–1709.
- Keeling, M. J. (2000). Simple stochastic models and their power-law type behaviour. *Theoretical Population Biology*, 58:21–31.
- Keller, E. and Segel, L. (1970a). Initiation of slime mold aggregation viewed as an instability. *Journal of Theoretical Biology*, 26:399–415.
- Keller, E. and Segel, L. (1970b). Travelling bands of chemotactic bacteria: A theoretical analysis. *Journal of Theoretical Biology*, 30:235–248.
- Keller, E. and Segel, L. (1971). Model for chemotaxis. *Journal of Theoretical Biology*, 30:225–234.
- Kendal, W. S. and Jørgensen, B. (2011). Taylor’s power law and fluctuation scaling explained by a central-limit-like convergence. *Physical Review E*, 83:066115.
- Kendall, B. E. and Fox, G. A. (2003). Unstructured individual variation and demographic stochasticity. *Conservation Biology*, 17:1170–1172.
- Kennard, A. S. et al. (2014). Individuality and universality in the growth-division laws of single *E. coli* cells. *arXiv:1411.4321v1*.

- Kilpatrick, A. M. and Ives, A. R. (2003). Species interactions can explain Taylor's power law for ecological time series. *Nature*, 422:65–68.
- Kinezaki, N., Kawasaki, K., Takasu, F., and Shigesada, N. (2003). Modeling biological invasions into periodically fragmented environments. *Theoretical Population Biology*, 64:291–302.
- Kingston, M. B. (1999). Effect of light on vertical migration and photosynthesis of *Euglena proxima* (Euglenophyta). *Journal of Phycology*, 35:245–253.
- Klausmeier, C. A. and Litchman, E. (2001). Algal games: The vertical distribution of phytoplankton in poorly mixed water columns. *Limnology and Oceanography*, 46:1998–2007.
- Kleiber, M. (1947). Body size and metabolic rate. *Physiological reviews*, 27:511–541.
- Kolmogorov, A. N., Petrovskii, I. G., and Piskunov, N. S. (1937). A study of the diffusion equation with increase in the amount of substance, and its application to a biological problem. *Moscow University Mathematics Mechanics Bulletin*, 1:1–25.
- Kussell, E. and Leibler, S. (2005). Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, 309:2075–2078.
- Labra, F. A., Marquet, P. A., and Bozinovic, F. (2007). Scaling metabolic rate fluctuations. *Proceedings of the National Academy of Sciences of the United States of America*, 104:10900–10903.
- Lampert, A. and Tlustý, T. (2013). Resonance-induced multimodal body-size distributions in ecosystems. *Proceedings of the National Academy of Sciences of the United States of America*, 110:205–209.
- Lampert, W., McCauley, E., and Manly, B. F. J. (2003). Trade-offs in the vertical distribution of zooplankton: ideal free distribution with costs? *Proceedings of the Royal Society B-Biological Sciences*, 270:765–773.
- Lande, R. (1976). Natural selection and random genetic drift in phenotypic evolution. *Evolution*, 30:314–334.
- Lapidus, I. R. and Schiller, R. (1976). Model for the chemotactic response of a bacterial population. *Biophysical Journal*, 16:779–789.
- Laughlin, S. (1981). A simple coding procedure enhances a neuron's information capacity. *Zeitschrift für Naturforschung C*, 36:910–912.

## Bibliography

---

- Lazova, M. D., Ahmed, T., Bellomo, D., Stocker, R., and Shimizu, T. S. (2011). Response rescaling in bacterial chemotaxis. *Proceedings of the National Academy of Sciences of the United States of America*, 108:13780–13875.
- Lebert, M., Porst, M., Richter, P., and Häder, D.-P. (1999). Physical characterization of gravitaxis in *Euglena gracilis*. *Journal of Plant Physiology*, 155:338–343.
- Levin, S. A. (1992). The problem of pattern and scale in ecology. *Ecology*, 73:1943–1967.
- Levin, S. A. (1994). Patchiness in marine and terrestrial systems: from individuals to populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 343:99–103.
- Lewontin, R. C. and Cohen, D. (1969). On population growth in a randomly varying environment. *Proceedings of the National Academy of Sciences of the United States of America*, 62:1056–1060.
- Lubina, J. A. and Levin, S. A. (1988). The spread of a reinvading species: Range expansion in the California sea otter. *The American Naturalist*, 131:526–543.
- Mandelbrot, B. B. (1973). *The fractal geometry of nature*. Freeman, San Francisco.
- Marquet, P. A. (2000). Ecology - invariants, scaling laws, and ecological complexity. *Science*, 289:1487–1488.
- Marquet, P. A. et al. (2005). Scaling and power-laws in ecological systems. *Journal of Experimental Biology*, 208:1749–1769.
- Marquet, P. A., Navarrete, S. A., and Castilla, J. C. (1990). Scaling population-density to body size in rocky intertidal communities. *Science*, 250:1125–1127.
- May, R. M. (1988). How many species are there on Earth? *Science*, 241:1441–1449.
- McMahon, T. A. and Bonner, J. T. (1983). *On Size and Life*. Scientific American Library, New York.
- Melbourne, B. A. et al. (2007). Invasion in a heterogeneous world: resistance, coexistence or hostile takeover? *Ecology letters*, 10:77–94.
- Melbourne, B. A. and Hastings, A. (2008). Extinction risk depends strongly on factors contributing to stochasticity. *Nature*, 454:100–103.
- Melbourne, B. A. and Hastings, A. (2009). Highly variable spread rates in replicated biological invasions: Fundamental limits to predictability. *Science*, 325:1536–1539.



- Mellard, J. P., Yoshiyama, K., Litchman, E., and Klausmeier, C. A. (2011). The vertical distribution of phytoplankton in stratified water columns. *Journal of Theoretical Biology*, 269:16–30.
- Méndez, V. and Camacho, J. (1997). Dynamics and thermodynamics of delayed population growth. *Physical Review E*, 55:6476–6482.
- Méndez, V., Campos, D., and Bartumeus, F. (2014). *Stochastic Foundations in Movement Ecology*. Springer, Berlin.
- Méndez, V., Fedotov, S., and Horsthemke, W. (2010). *Reaction-Transport Systems*. Springer, Berlin.
- Méndez, V., Fort, J., Rotstein, H., and Fedotov, S. (2003). Speed of reaction-diffusion fronts in spatially heterogeneous media. *Physical Review E*, 68:041105.
- Méndez, V., Llopis, I., Campos, D., and Horsthemke, W. (2011). Effect of environmental fluctuations on invasion fronts. *Journal of Theoretical Biology*, 281:31–38.
- Mesibov, R., Ordal, G. W., and Adler, J. (1973). The range of attractant concentrations for bacterial chemotaxis and the threshold and size of response over this range Weber law and related phenomena. *The Journal of General Physiology*, 62:203–223.
- Mora, T. and Bialek, W. (2011). Are biological systems poised at criticality? *Journal of Statistical Physics*, 144:268–302.
- Moran, P. A. P. (1953). The statistical analysis of the Canadian lynx cycle. II. Synchronization and meteorology. *Australian Journal of Zoology*, 1:291–298.
- Moro, E. (2004). Numerical schemes for continuum models of reaction-diffusion systems subject to internal noise. *Physical Review E*, 70:045102.
- Murray, J. D. (2004). *Mathematical Biology I: An Introduction*. Springer, Berlin, 3 edition.
- Nelson, D. R. (1999). Population dynamics and Burgers' equation. *Physica A: Statistical Mechanics and its Applications*, 274:85–90.
- Nelson, D. R. (2012). Biophysical dynamics in disorderly environments. *Annual Review of Biophysics*, 41:371–402.
- Nelson, D. R. and Schnerb, N. M. (1998). Non-Hermitian localization and population biology. *Physical Review E*, 58:1383–1403.

## Bibliography

---

- Niklas, K. J. and Enquist, B. J. (2001). Invariant scaling relationships for interspecific plant biomass production rates and body size. *Proceedings of the National Academy of Sciences of the United States of America*, 98:2922–2927.
- Novikov, E. A. (1965). Functionals and the random-force method in turbulence theory. *Soviet Physics, Journal of Experimental and Theoretical Physics*, 20:1290–1294.
- Okubo, A. and Levin, S. A. (2002). *Diffusion and Ecological Problems: Modern Perspectives*. Springer, Berlin, 2 edition.
- Ooka, H., Ishii, T., Hashimoto, K., and Nakamura, R. (2014). Light-induced cell aggregation of *Euglena gracilis* towards economically feasible biofuel production. *RSC Advances*, 4:20693.
- Parmesan, C. et al. (1999). Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature*, 399:579–583.
- Peters, R. H. (1983). *The ecological implications of body size*. Cambridge University Press, Cambridge.
- Post, E. and Forchhammer, M. C. (2002). Synchronization of animal population dynamics by large-scale climate. *Nature*, 420:168–171.
- Rading, M. M., Engel, T. A., Lipowsky, R., and Valleriani, A. (2011). Stationary size distributions of growing cells with binary and multiple cell division. *Journal of Statistical Physics*, 145:1–22.
- Ragni, M. and Ribera D'Alcalà, M. (2004). Light as an information carrier underwater. *Journal of Plankton Research*, 26:433–443.
- Ramsayer, J., Fellous, S., Cohen, J. E., and Hochberg, M. E. (2012). Taylor's law holds in experimental bacterial populations but competition does not influence the slope. *Biology letters*, 8:316–319.
- Redner, S. (1990). Random multiplicative processes: An elementary tutorial. *American Journal of Physics*, 58:267–273.
- Rinaldo, A., Maritan, A., Cavender-Bares, K. K., and Chisholm, S. W. (2002). Cross-scale ecological dynamics and microbial size spectra in marine ecosystems. *Proceedings of the Royal Society B-Biological Sciences*, 269:2051–2059.
- Ringelberg, J. and Flik, B. J. G. (1994). Increased phototaxis in the field leads to enhanced diel vertical migration. *Limnology and Oceanography*, 39:1855–1864.

- Rodriguez, J. and Mullin, M. M. (1986). Relation between biomass and body-weight of plankton in a steady-state oceanic ecosystem. *Limnology and Oceanography*, 31:361–370.
- Rodriguez-Iturbe, I., Muneeppeerakul, R., Bertuzzo, E., Levin, S. A., and Rinaldo, A. (2009). River networks as ecological corridors: A complex systems perspective for integrating hydrologic, geomorphologic, and ecologic dynamics. *Water Resources Research*, 45:W01413.
- Rodriguez-Iturbe, I. and Rinaldo, A. (1997). *Fractal River Basins: Chance and Self-Organization*. Cambridge University Press, New York.
- Rosen, G. (1976). Existence and nature of band solutions to generic chemotactic transport equations. *Journal of Theoretical Biology*, 59:243–246.
- Rosindell, J., Hubbell, S. P., He, F., Harmon, L. J., and Etienne, R. S. (2012). The case for ecological neutral theory. *Trends in Ecology & Evolution*, 27:203–208.
- Roy, M., Holt, R. D., and Barfield, M. (2005). Temporal autocorrelation can enhance the persistence and abundance of metapopulations comprised of coupled sinks. *The American Naturalist*, 166:246–261.
- Rubin, K. J., Pruessner, G., and Pavliotis, G. A. (2014). Mapping multiplicative to additive noise. *Journal of Physics A: Mathematical and Theoretical*, 47:195001.
- Sbalzarini, I. F. and Koumoutsakos, P. (2005). Feature point tracking and trajectory analysis for video imaging in cell biology. *Journal of Structural Biology*, 151:182–195.
- Schiesser, W. E. (1991). *The Numerical Method of Lines Integration of Partial Differential Equations*. Academic Press, San Diego.
- Scott, D. W. (1979). On optimal and data-based histograms. *Biometrika*, 66:605–610.
- Sheldon, R. W., Prakash, A., and Sutcliffe, W. H. (1972). The size distribution of particles in the ocean. *Limnology and Oceanography*, 17:327–340.
- Shigesada, N., Kawasaki, K., and Ei, T. (1986). Traveling periodic waves in heterogeneous environments. *Theoretical Population Biology*, 30:143–160.
- Shoval, O., Goentoro, L., Hart, Y., Mayo, A., Sontag, E., and Alon, U. (2010). Fold-change detection and scalar symmetry of sensory input fields. *Proceedings of the National Academy of Sciences of the United States of America*, 107:15995–16000.

## Bibliography

---

- Simini, F., Anfodillo, T., Carrer, M., Banavar, J., and Maritan, A. (2010). Self-similarity and scaling in forest communities. *Proceedings of the National Academy of Sciences of the United States of America*, 107:7658–7662.
- Skellam, J. G. (1951). Random dispersal in theoretical populations. *Biometrika*, 38:196–218.
- Smayda, T. J. (1997). Harmful algal blooms: Their ecophysiology and general relevance to phytoplankton blooms in the sea. *Limnology and Oceanography*, 42:1137–1153.
- Solé, R. and Bascompte, J. (2006). *Self-Organization in Complex Ecosystems*. Princeton University Press, Princeton.
- Solé, R., Manrubia, S. C., Benton, M., Kauffman, S., and Bak, P. (1999). Criticality and scaling in evolutionary ecology. *Trends in Ecology & Evolution*, 14:156–160.
- Southwood, T. R. E., May, R. M., and Sugihara, G. (2006). Observations on related ecological exponents. *Proceedings of the National Academy of Sciences of the United States of America*, 103:6931–6933.
- Srivastava, D. S. et al. (2004). Are natural microcosms useful model systems for ecology? *Trends in Ecology & Evolution*, 19:379–384.
- Stanley, H. E. et al. (2000). Scale invariance and universality: organizing principles in complex systems. *Physica A: Statistical Mechanics and its Applications*, 281:60–68.
- Steele, J. H. and Yentsch, C. S. (1960). The vertical distribution of chlorophyll. *Journal of the Marine Biological Association of the United Kingdom*, 39:217–226.
- Stocker, R. (2012). Marine microbes see a sea of gradients. *Science*, 338:628–633.
- Strother, G. K. and Wolken, J. J. (1960). Microspectrophotometry of *Euglena* chloroplast and eyespot. *Nature*, 188:601–602.
- Sutherland, W. J. et al. (2013). Identification of 100 fundamental ecological questions. *Journal of Ecology*, 101:58–67.
- Taylor, L. R. (1961). Aggregation, variance and the mean. *Nature*, 189:732–735.
- Taylor, L. R. and Woiwod, I. P. (1982). Comparative synoptic dynamics. I. Relationships between inter- and intra-specific spatial and temporal variance/mean population parameters. *Journal of Animal Ecology*, 51:879–906.
- Tindall, M. J., Maini, P. K., Porter, S. L., and Armitage, J. P. (2008). Overview of mathematical approaches used to model bacterial chemotaxis II: bacterial populations. *Bulletin of Mathematical Biology*, 70:1570–1607.

- Torney, C. and Neufeld, Z. (2008). Phototactic clustering of swimming microorganisms in a turbulent velocity field. *Physical Review Letters*, 101:078105.
- Touchette, H. (2009). The large deviation approach to statistical mechanics. *Physics Report*, 478:1–69.
- Turing, A. M. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 237:37–72.
- Turner, M. G. (2005). Landscape ecology: what is the state of the science? *Annual Review of Ecology, Evolution, and Systematics*, 36:319–344.
- Tweedie, M. C. K. (1946). The regression of the sample variance on the sample mean. *Journal of the London Mathematical Society*, 21:22–28.
- Tyson, J. J. and Diekmann, O. (1986). Sloppy size control of the cell-division cycle. *Journal of Theoretical Biology*, 118:405–426.
- Uhlenbeck, G. E. and Ornstein, L. S. (1930). On the theory of the brownian motion. *Physical Review*, 36:0823–0841.
- Uttley, P. and McHardy, I. M. (2001). The flux-dependent amplitude of broadband noise variability in X-ray binaries and active galaxies. *Monthly Notices of the Royal Astronomical Society*, 30:L26–L30.
- Vallee, B. L. and Falchuk, K. H. (1993). The biochemical basis of zinc physiology. *Physiological Reviews*, 73:79–118.
- Van Kampen, N. G. (2007). *Stochastic Processes in Physics and Chemistry*. North Holland, Amsterdam.
- Vasseur, D. A. (2007). Environmental colour intensifies the Moran effect when population dynamics are spatially heterogeneous. *Oikos*, 116:1726–1736.
- Vasseur, D. A. and Fox, J. W. (2009). Phase-locking and environmental fluctuations generate synchrony in a predator-prey community. *Nature*, 460:1007–1010.
- Vellend, M. (2010). Conceptual synthesis in community ecology. *The Quarterly Review of Biology*, 85:183–206.
- Vincent, R. V. and Hill, N. A. (1996). Bioconvection in a suspension of phototactic algae. *Journal of Fluid Mechanics*, 327:343–371.
- Volpert, V. and Petrovskii, S. (2009). Reaction-diffusion waves in biology. *Physics of Life Reviews*, 6:267–310.

## Bibliography

---

- West, G. B., Brown, J. H., and Enquist, B. J. (1997). A general model for the origin of allometric scaling laws in biology. *Science*, 276:122–126.
- White, E. P., Ernest, S. K. M., Kerkhoff, A. J., and Enquist, B. J. (2007). Relationships between body size and abundance in ecology. *Trends in Ecology & Evolution*, 22:323–330.
- Williams, C. R. and Bees, M. A. (2011a). Photo-gyrotactic bioconvection. *Journal of Fluid Mechanics*, 678:41–86.
- Williams, C. R. and Bees, M. A. (2011b). A tale of three taxes: photo-gyro-gravitactic bioconvection. *The Journal of Experimental Biology*, 214:2398–2408.
- Williamson, J. and Harrison, S. (2002). Biotic and abiotic limits to the spread of exotic revegetation species. *Ecological Applications*, 12:40–51.
- Wolken, J. J. (1961). *Euglena: An experimental organism for biochemical and biophysical studies*. Rutgers University Press, New Brunswick.
- Xiao, X., Locey, K. J., and White, E. P. (2014). A process-independent explanation for the general form of Taylor's law. *arXiv:1410.7283*.
- Yoshiyama, K., Mellard, J. P., Litchman, E., and Klausmeier, C. A. (2009). Phytoplankton competition for nutrients and light in a stratified water column. *The American Naturalist*, 174:190–203.
- Zauderer, E. (1989). *Partial Differential Equations of Applied Mathematics*. Wiley, New York, 2 edition.

## Curriculum Vitae

February 2015

### Personal Information

*Name:* Andrea

*Surname:* Giometto

*Date of Birth:* October 10<sup>th</sup>, 1987

*Nationality:* Italian

*E-mail:* andrea.giometto@epfl.ch; andrea.giometto@gmail.com

### Research Interests

Complex systems, Population dynamics, Scaling in ecology, Biological invasions, Stochastic modeling of ecosystems dynamics, Motility of microorganisms.

### Education

- **February 2015:** PhD, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland;
- **June 2012:** Diploma Galileiano (70/70 cum laude), Scuola Galileiana di Studi Superiori, Università degli Studi di Padova, Padova, Padova, Italy;
- **July 2011:** M. Sc., Physics (110/110 cum laude), Università degli Studi di Padova, Padova, Italy;
- **July 2009:** B. Sc., Physics (110/110 cum laude), Università degli Studi di Padova, Padova, Italy.

### Certifications

- **July 2011:** IELTS (International English Language Testing System), British Council. Overall band score: 8.0 (CEFR<sup>1</sup> level: C2);
- **May 2005:** FCE (First Certificate in English), University of Cambridge;
- **April 2004:** European Computer Driving License (ECDL).

---

<sup>1</sup>Common European Framework of Reference

### Workshop, Schools and International Experiences

- **June 2014:** Workshop “Discounting and evaluation of environmental policies”, Istituto Veneto di Scienze, Lettere ed Arti, Venice, Italy;
- **November 2013:** “Movement and Dispersal Workshop”, University of Aberdeen, United Kingdom;
- **April 2013:** Joint Centre de Recerca Matemàtica (CRM) and Imperial College “School and Workshop in Complex Systems”, CRM Barcelona, Spain;
- **February 2013:** Winter Graduate School “Ecological Theory”, University of Zürich and ETH Zürich, Centro Stefano Franscini, Ascona, Switzerland;
- **2010-2011:** M. Sc. thesis research conducted at the Institute for Mathematical Sciences, Imperial College, London, United Kingdom;
- **August 2010:** Summer School “Physics of the Climate System”, Utrecht University, The Netherlands;
- **August 2008:** Summer School “Progetto Lauree Scientifiche”, Italian Physical Society (SIF), Perugia, Italy;
- **June 2008:** BESS Summer School “Climate Forcing and Global Patterns”, Istituto Veneto di Scienze, Lettere ed Arti, Venice, Italy;
- **August 2007:** Summer School “Progetto Lauree Scientifiche”, Italian Physical Society (SIF), Perugia, Italy.

### Professional Experiences

- **2011-2015:** Visiting scientist at Eawag (Dübendorf, Zürich), Department of Aquatic Ecology.
- **2013-2014:** Silvia Zaoli’s M. Sc. thesis co-supervisor (Physics), Università degli Studi di Padova, Italy;
- **2013-2014:** Teaching assistant for the course “Limnoecology”, ETH Zürich, Switzerland;
- **2012-2013:** Teaching assistant for the course “Limnoecology”, ETH Zürich, Switzerland;
- Reviewer for *Proceedings of the Royal Society B-Biological Sciences*.



## Awards

- **August 2013:** “Best poster award”, Institute of Environmental Engineering (IIE) Green Days, EPFL;
- **May 2013:** “Best student talk award”, Eawag PhD Symposium;
- **September 2012:** “Best student talk award”, Eawag Symposium;
- **April 2010:** “Erasmus Mundus” scholarship, European Commission;
- **2006-2009:** “Progetto Lauree Scientifiche” scholarship, Italian Society of Physics;
- **September 2005:** “Best Student” scholarship, Liceo Brocchi, Bassano del Grappa, Italy;
- **September 2004:** “International Mobility” scholarship, Regione Veneto, Italy.

## Peer-reviewed journal articles

- Giometto, A. and Jensen, H. J. (2012). Connecting the microdynamics to the emergent macrovariables: Self-organized criticality and absorbing phase transitions in the deterministic lattice gas. *Physical Review E* 85:011128;
- Giometto, A., Altermatt, F., Carrara, F., Maritan, A. and Rinaldo, A. (2013). Scaling body size fluctuations. *Proceedings of the National Academy of Sciences of the United States of America* 110:4646–4650;
- Giometto, A., Rinaldo, A., Carrara, F. and Altermatt, F. (2014). Emerging predictable features of replicated biological invasion fronts. *Proceedings of the National Academy of Sciences of the United States of America* 111:297–301;
- Carrara, F., Rinaldo, A., Giometto, A. and Altermatt, F. (2014). Complex interaction of dendritic connectivity and hierarchical patch size on biodiversity in river-like landscapes. *The American Naturalist* 183:13–25;
- Altermatt, F., Fronhofer, E., Garnier, A., Giometto, A., Hammes, F., Klecka, J., Legrand, D., Mächler, E., Thomas, M. M., Pennekamp, F., Plebani, M., Pontarp, M., Schtickzelle, N., Virginie, T. and Petchey, O. L. (2014). Big answers from small worlds: a user’s guide for protist microcosms as a model system in ecology and evolution. *Methods in Ecology and Evolution*. In press, doi: 10.1111/2041-210x.12312;

## Curriculum Vitae

---

- Carrara, F., Giometto, A., Seymour, A., Rinaldo, A. and Altermatt, F. (2014). Experimental evidence for strong stabilizing forces at high functional diversity of aquatic microbial communities. *Ecology. In press.*

## Pre-prints

- Evans, T. S. and Giometto, A. (2011). Turnover rate of popularity charts in neutral models. *arXiv:1105.4044* [physics.soc-ph];
- Zaoli, S., Giometto, A., Formentin, M., Azaele, S., Rinaldo, A. and Maritan, A. (2014). Phenomenological modeling of the motility of self-propelled microorganisms. *arXiv:1407.1762* [physics.bio-ph];
- Giometto, A., Formentin, M., Rinaldo, A., Cohen, J. E. and Maritan, A. (2014). Sample and population exponents of generalized Taylor's law. *arXiv:1412.5026* [q-bio.PE].

## Conference abstracts

- Giometto, A., Altermatt, F., Carrara, F., and Rinaldo, A. Mean, fluctuations and predictability in biological dispersal (Poster). AGU Fall Meeting 2013, San Francisco, United States;
- Giometto, A., Rinaldo, A., Carrara, F. and Altermatt, F. Mean, variance and predictability in biological dispersal (Oral Presentation). Movement and Dispersal conference 2013, Aberdeen, United Kingdom;
- Giometto, A., Rinaldo, A., Carrara, F. and Altermatt, F. Mean, variance and predictability in biological dispersal (Poster). INTECOL 2013, London, United Kingdom;
- Giometto, A., Altermatt, F., Carrara, F., and Rinaldo, A. Scaling body size fluctuations (Oral Presentation). Meeting of the German Society for Protozoology 2013, Kartause Ittingen, Switzerland;
- Giometto, A., Rinaldo, A., Carrara, F. and Altermatt, F. Mean, variance and predictability in biological dispersal (Oral Presentation). IOP Physics of Emergent Behaviour 2013, Brighton, United Kingdom;
- Giometto, A., Altermatt, F., Carrara, F., Maritan, A. and Rinaldo, A. Scaling and universality in microbial size spectra (Poster). AGU Fall Meeting 2013, San Francisco, United States;

- Giometto, A., Carrara, F., Altermatt, F. and Rinaldo, A. Microcosms metacommunities in river networks: niche effects and biodiversity (Poster). EGU General Assembly 2012, Vienna, Austria.

### Personal Skills

#### Languages:

- **Italian:** Mother tongue;
- **English:** Proficient User (CEFR<sup>2</sup> level: C2);
- **French:** Independent User (CEFR level: B1).

**Computer Skills:** Mathematica, Matlab, Office, Adobe Illustrator, Adobe Photoshop, LaTeX, C++, High Performance Computing (HPC).

**Laboratory Skills:** Experiments with protist microcosms: cell tracking, body size and abundance measurements with cell counters and flow cytometers. Working knowledge of Arduino electronics platform.

---

<sup>2</sup>Common European Framework of Reference

### References

To get additional information please contact:

- Prof. Andrea Rinaldo ([andrea.rinaldo@epfl.ch](mailto:andrea.rinaldo@epfl.ch)), Laboratory of Ecohydrology (ECHO), Environmental Engineering Institute (IIE), School of Architecture, Civil and Environmental Engineering (ENAC), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland;
- Prof. Florian Altermatt ([florian.altermatt@eawag.ch](mailto:florian.altermatt@eawag.ch)), Eawag: Swiss Federal Institute of Aquatic Science and Technology, Department of Aquatic Ecology, Überlandstrasse 133, Dübendorf, Switzerland;
- Prof. Amos Maritan ([amos.maritan@pd.infn.it](mailto:amos.maritan@pd.infn.it)), Dipartimento di Fisica, Università degli Studi di Padova, Padova, Italy.