FLUCTUATIONS, SCALING, AND UNIVERSALITY IN FIRST-PASSAGE PROCESSES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

David Hathcock

July 2022

FLUCTUATIONS, SCALING, AND
UNIVERSALITY IN FIRST-PASSAGE PROCESSES

David Hathcock, Ph.D.
Cornell University 2022

Here we study fluctuations, scaling, and universality in a variety of first-passage scenarios. First, we explore fluctuations in fixation times in evolutionary dynamics. We compute the fixation-time distribution for several models of evolution and determine how the shape of the distribution depends on the fitness advantage provided by a genetic mutation. Our results reveal an interesting dichotomy: for neutral mutations the distribution is highly-skewed, while for non-neutral mutations, two particular distributions arise. In the latter case, depending on population structure the fixation-time distribution is either a Gaussian or the (moderately skewed) Gumbel distribution. Next, we show that the Gaussian and Gumbel distributions are universal; they arise generically across a variety of stochastic models of evolution, ecology, epidemiology, and chemical reactions. The distinguishing feature is the decay of the stochastic transition rates near the absorbing state: lack of decay leads to Gaussian distributions, while linear decay leads to Gumbel distributions. Distributions resulting from other power-law decays in the transition rates are also classified. Finally, we formulate a renormalization group approach and scaling theory for barrier crossing phenomena near a noisy saddle-node bifurcation, where the barrier vanishes. We derive the universal scaling behavior and corrections to scaling for the mean barrier escape time in overdamped systems with arbitrary barrier height. We also develop an accurate approximation for the fluctuations in es-

cape times, capturing the full distribution of barrier escape times at any barrier height. This critical theory draws links between barrier crossing in chemistry, the renormalization group, and bifurcation theory.

BIOGRAPHICAL SKETCH

David Hathcock grew up in Longmont, Colorado and graduated from Niwot High School in 2013. He then attended Case Western Reserve University, where he initially planned on studying economics. But his first-year mechanics course, which (somewhat unconventionally) covered topics from chaos theory, inspired a change of direction. David eventually changed his major when the Physics department chair emailed him wondering why an economics major was taking so many advanced physics courses. He received a B.S. in Mathematics and Physics in 2017.

David's foray into physics research at CWRU began in 2015 with studying cellular signaling and other areas of theoretical biophysics under Michael Hinczewski. He also spent the summer of 2016 at the University of Maryland, working with Ed Ott and Michelle Girvan, where he first learned about networks, oscillators, and nonlinear dynamics. These experiences motivated David to pursue graduate studies at Cornell, where he has been working on his PhD in Physics under Steve Strogatz and has also been collaborating regularly with Jim Sethna. At the time of writing, David's research interests include statistical physics and nonlinear dynamics, particularly applications to population dynamics and cellular and molecular biology. He has published work on cell signaling, molecular motors, evolutionary dynamics, universal fluctuations in absorption processes, and on a scaling and renormalization group framework for chemical reactions (the last three of which make up this thesis).

This document is dedicated to Louie, Haggis, and Madeleine.

ACKNOWLEDGEMENTS

First, a huge thanks to Michael Hinczewski, whose mentorship during my undergraduate years fostered my love for statistical physics. Mike's early influence on my development into a physicist set the foundation for much of the work in this thesis.

Thanks to Bertrand Ottino-Löffler, who took me under his wing during my first year and introduced me to the world of evolutionary dynamics. Bertrand's own research and mentorship inspired Chapter 2 of this thesis.

Thanks to Alen Senanian for the constant friendship and entertainment during my time at Cornell.

Thanks to Itay Griniasty and Jaron Kent-Dobias who have each been valuable collaborators and with whom I have shared countless memorable conversations about various areas of physics and life over the years.

Thanks to the various fellowships that generously supported my graduate work: the Cornell University Fellowship, Cornell Presidential Life Science Fellowship, and National Science Foundation Graduate Research Fellowship. These enabled broad exploration and collaboration that would not have otherwise been possible.

Thanks to my partner Madeleine Dempsey and our two dogs, Haggis and Louie, for the constant love and support throughout this journey.

Thanks to my parents whose own advanced degrees inspired and enabled me to pursue this dream.

Finally, thanks to my advisors Steve Strogatz and Jim Sethna for their invaluable mentorship, inspiration, and collaboration. Thanks to Steve for teaching me enough about oscillators to last a lifetime, helping me improve my writing by a thousandfold, and believing in my work (even when referees were skep-

tical). Thanks to Jim for his constant energy and excitement, teaching me the magic and beauty of the renormalization group, and for the honor of teaching statistical mechanics alongside him. It has truly been a pleasure learning from these two scientists.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The concept of first passage underlies many physical, chemical and biological processes and even arises commonly in daily life. Many readers might recall asking the question "are we there yet?" on long car rides as a young child. During the incubation of a disease, symptoms onset when the invasive bacteria or virus first reaches a certain population level. Chemical reactions and magnetic spin avalanches each involve first-passage escape over an energy barrier separating metastable and stable states. A molecular motor, walking along actin filaments or microtubules to drag cargo across the cell, diffusively searches for binding sites and binds to those it first encounters. Finally, the extinction of a species and the eradication of an epidemic both occur when the population (of animals or of sick individuals) reaches zero.

For each of these examples (many of which we will revisit in this thesis), reaching a threshold *for the first time* either triggers a response (e.g., onset of disease) or terminates the dynamics entirely (e.g., extinction). The key quantity of interest for studying these processes is the first-passage time T, how long does it take to first reach the threshold?[1] When the underlying dynamics of the system are stochastic, T is itself a random quantity, with a probability distribution $f_T(t)$.

The broad goal of this thesis is to determine how the underlying physics of a system, which determines the space of possible states and the rates of transition between these states, influences the statistics of the first-passage time. For example, the state of the system might be the number of sick individuals

[1] In the following chapters we will refer to fixation time in the context of evolutionary dynamics, absorption time for Markov chains with an absorbing boundary state, and barrier escape time for chemical reactions. These are all different examples of first-passage times.

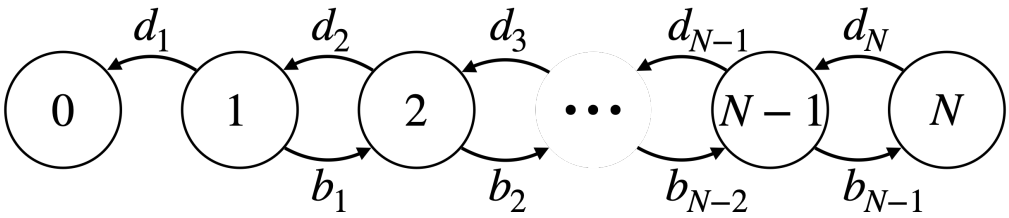

Figure 1.1: A birth-death Markov chain. The system is composed of a linear chain of N states, with b_m and d_m designating the transition rates between neighboring states. In this example, the upper boundary $m = N$ is reflecting and the lower boundary $m = 0$ is absorbing.

in a population or the positions of all the atoms in a molecule undergoing a chemical reaction. The transitions between states are respectively determined by the contact network of the population and the intramolecular forces holding the molecule together. How do the details of the contact network or molecular forces influence the shape of the first-passage time distribution $f_T(t)$?

Figure 1.1 shows one of the simplest possible systems: a birth-death Markov chain, which is composed of a discrete one-dimensional chain of N states with transitions only between nearest neighbor states occurring at rates b_m and d_m. Models of this form (and those with absorbing boundaries on both ends of the chain) will be the focus of Chapters 2 and 3. In Chapter 4, we consider the analogous continuous system, where the state undergoes diffusive motion.

While these one-dimensional models are simplistic, they often provide effective coarse-grained descriptions of higher-dimensional phenomena. For example, averaging over the configurations of infected individuals in network models of epidemiology produces effective transition rates b_m and d_m that accurately reproduce the dynamics of the infected population [3, 4]. Even when the state space is genuinely high-dimensional, the first-passage behavior may still be well described by the birth-death Markov chain if the dynamics collapse onto

a one-dimensional slow manifold near the absorbing state (which is often the case in dynamical systems). Similarly, the theory of chemical reactions employs a coarse-grained one-dimensional reaction coordinate, which parametrizes the path of least resistance between reactant and product chemical species [5] (also see Chapter 4 for further discussion).

Much effort has gone toward computing mean first-passage times,

$$\langle T \rangle = \int t f_T(t)\, dt \tag{1.1}$$

for simple stochastic models that fall into the birth-death Markov chain class shown in Figure 1.1, both for specific models and more generally. In particular, Doering and colleagues obtained asymptotic ($N \gg 1$) expressions for the mean time to hit the absorbing state at 0 for several broad classes of Markov chains based on generic features of the transition rates b_m and d_m [6, 7].

In this thesis we go beyond the mean to provide an understanding of the higher-order fluctuations in first-passage times. We develop new analytical approaches and approximations that enable characterization of the cumulants of first-passage times, κ_n, defined via the generating function

$$\log \tilde{f}_T(\omega) = \sum_{n=1}^{\infty} \kappa_n (i\omega)^n / n!, \tag{1.2}$$

where $\tilde{f}_T(\omega)$ is the Fourier transform of the first-passage time distribution. The first two cumulants are familiar: κ_1 is the mean, while κ_2 is the variance. The higher-order cumulants characterize the shape of the distribution, for instance $\kappa_3/\kappa_2^{3/2}$ is the skew. We will answer the following questions in the chapters below: what characteristics of the stochastic transition rates in a system (b_m and d_m for the birth-death Markov chain) determine the cumulants and hence the shape of the distribution $f_T(t)$? How do the mean and higher-order cumulants

scale, for large systems or near a critical transition? Finally, are there universal features of first-passage processes? Do the same fluctuations and scaling appear across models and applications?

Chapter 2 focuses on a particular application: evolutionary dynamics. There we investigate how the fixation time, the time for a genetic mutation to spread to an entire population via competitive reproductive dynamics and natural selection, depends on the fitness advantage (or lack thereof) conferred by the mutation. We find a striking transition: for neutral fitness the fixation-time distributions are highly skewed, while for non-neutral evolution the distribution is either Gaussian or Gumbel (depending on the population structure).

The appearance of Gaussian and Gumbel distributions for broad ranges of fitness and across different models of evolution suggests a degree of universality. In Chapter 3 we classify the birth-death Markov chains (Figure 1.1) that have Gaussian or Gumbel absorption-time distributions, connecting the resulting distribution to basic features of the underlying dynamics. We also identify a broader family of universal skewed distributions that interpolates between the Gaussian and Gumbel.

Finally, Chapter 4 analyzes barrier-crossing phenomena in continuous systems, applicable to chemical reactions and other noise-driven systems with metastable states. Here we study the scaling of the barrier escape time near a critical transition: the saddle-node bifurcation where the barrier vanishes. Using a renormalization group approach inspired by Feigenbaum's renormalization group for iterated maps [8], we derive the universal scaling near the saddle-node transition for both the mean and full distribution of barrier crossing times. The results capture the crossover from high-barrier to barrier-less systems, uni-

fying classical reaction rate theory with dynamical systems theory for the noisy saddle-node bifurcation.

To conclude, Chapter 5 outlines a number of experimental systems which may be used to test our predictions of universality and scaling for first-passage times and their fluctuations. Real-world epidemiological processes as well as bacterial evolution, single-cell aging, and optoelectronic laboratory experiments each have underlying first-passage processes, whose measurement provides a direct test of our results. When the experimental measurements agree with the predicted first-passage-time distributions or scaling, the theory provides a clear picture of the underlying physics governing the system.

CHAPTER 2

FITNESS DEPENDENCE OF THE FIXATION-TIME DISTRIBUTION FOR EVOLUTIONARY DYNAMICS ON GRAPHS

2.1 Introduction[1]

Reproducing populations undergo evolutionary dynamics. Mutations can endow individuals with a fitness advantage, allowing them to reproduce more quickly and outcompete non-mutant individuals [9]. Two natural questions arise: If a single mutant individual is introduced into a population, what is the *probability* that the mutant lineage will spread and ultimately take over the population (an outcome known as fixation)? And if fixation occurs, how much *time* does it take?

These questions have been addressed, in part, by evolutionary graph theory, which studies evolutionary dynamics in structured populations. Thanks to this approach, fixation probabilities are now well understood for various models on various networks [10–20]. Less is known about fixation times. Given a model of evolutionary dynamics, one would like to predict the mean, variance, and ideally the full distribution of its fixation times.

Of these quantities, the mean is the best understood. Numerical and analytical results exist for mean fixation times on both deterministic [12, 14, 19–25] and random [4, 24–26] networks. Yet although mean fixation times are important to study, the information they provide can be misleading, because fixation-time

[1]This chapter is reproduced from: David Hathcock and Steven H. Strogatz, "Fitness dependence of the fixation-time distribution for evolutionary dynamics on graphs." Physical Review E **100**, 012498 (2019)

distributions tend to be broad and skewed and hence are not well characterized by their means alone [19, 27–30]. Initial analytical results have determined the asymptotic fixation-time distribution for several simple networks, but only when the relative fitness of the mutants is infinite [31–33]. For other values of the relative fitness, almost nothing is known. Preliminary results suggest that at neutral fitness (when mutants and non-mutants are equally fit), the fixation-time distribution becomes highly right-skewed [33].

In this chapter we investigate the full fitness dependence of fixation-time distributions for the Moran process [34, 35], a simple model of evolutionary dynamics. In the limit of large network size, we derive asymptotically exact results for the fixation-time distribution and its skew for two network structures at opposite ends of the connectivity spectrum: the complete graph, in which every individual interacts with every other individual; and the one-dimensional ring lattice, in which each individual interacts only with its nearest neighbors on a ring.

The specific model we consider is the Moran Birth-death (Bd) process[2], defined as follows. On each node of the network there is an individual, either mutant or non-mutant. The mutants have a fitness level r, which designates their relative reproduction rate compared to non-mutants. When $r > 1$, the mutants have a fitness advantage, whereas when $r = 1$ they have neutral fitness. At each time step we choose a node at random, with probability proportional to its fitness, and choose one of its neighbors with uniform probability. The first individual gives birth to an offspring of the same type. That offspring replaces the neighbor, which dies. The model population is updated until either the mutant

[2]We use the convention that capital letters designate a fitness dependent step in the Moran process (e.g., for the Bd process nodes give birth at a rate proportional to their fitness, but die with uniform probability). See Ref. [33, Box 2] for a detailed explanation of this nomenclature.

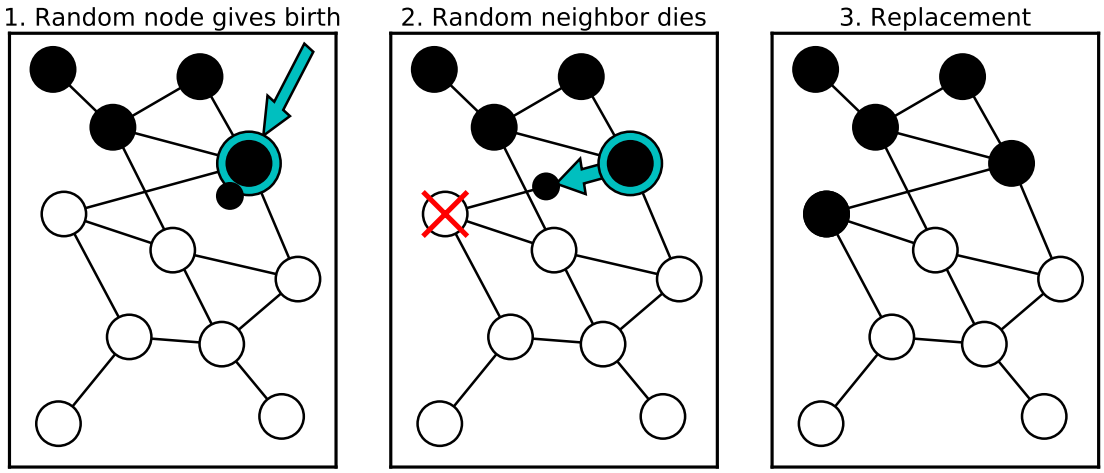

Figure 2.1: Dynamics of the Moran Birth-death (Bd) process. At each time step (1) a random node is chosen to give birth, (2) one of its neighbors is chosen to die, and (3) the new offspring replaces the dying node.

lineage takes over (in which case fixation occurs) or the mutant lineage goes extinct (a case not considered here). Figure 2.1 illustrates a single step of this update procedure.

As mentioned above, the distribution of fixation times is often skewed. The skew emerges from the stochastic competition between mutants and non-mutants through multiple mechanisms. For instance, when the mutants have neutral fitness the process resembles an unbiased random walk. We find that the asymptotic fixation-time distribution for a simple random walk is only skewed when the walk is unbiased. The lack of bias allows for occasional long recurrent excursions (that are suppressed in biased walks) during successful runs to fixation. The fixation-time distribution is strongly skewed because there are many ways to execute such walks that are much longer than usual, but comparably few ways for mutants to sweep through the population much faster than usual.

Depending on network structure, the fixation-time skew can also come from a second, completely separate mechanism, which involves characteristic slow-

downs that arise because individuals do not discriminate between mutants and non-mutants during the replacement step of the Moran process. For example, when very few non-mutants remain, the mutants can waste time replacing each other. These slowdowns are reminiscent of those seen in a classic problem from probability theory, the coupon collector's problem, which asks: How long does it take to complete a collection of N distinct coupons if a random coupon is received at each time step? The intuition for the long slowdowns is clear: when nearly all the coupons have been collected, it can take an exasperatingly long time to collect the final few, because one keeps acquiring coupons that one already has. The problem was first solved by Erdős and Rényi, who proved that for large N, the time to complete the collection has a Gumbel distribution [36]. In fact, for evolutionary processes with infinite fitness there exists an exact mapping onto coupon collection [32, 33]. Remarkably, while this correspondence breaks down for finite fitness, the coupon collection heuristic still allows us to predict correct asymptotic fixation-time distributions for non-neutral fitness.

In the following sections we show that for $N \gg 1$, the neutral-fitness Moran process on the complete graph and the one-dimensional ring lattice has highly skewed fixation-time distributions, and we solve for their cumulants exactly. For non-neutral fitness the fixation-time distribution is normal on the lattice and a weighted convolution of Gumbel distributions on the complete graph. These results are novel; apart from the infinite fitness limit and some partial results at neutral fitness (noted below), the fitness dependence of these distributions was previously unknown.

We begin by developing a general framework for computing fixation-time distributions and cumulants of birth-death Markov chains, and then apply it to

the Moran process to prove the results above. We also consider the effects of truncation on the process and examine how long it takes to reach partial, rather than complete, fixation. The fixation-time distributions have rich dependence on the fitness level and the degree of truncation, with both discontinuities and regions of universality. To conclude, we discuss extensions of our results to two-fitness Moran models and to more complicated network topologies.

2.2 General Theory for Birth-Death Markov Processes

For simplicity, we restrict attention to network topologies and initial mutant populations for which the probability of adding or removing a mutant in a given time step depends only on the number of existing mutants, not on where the mutants are located on the network. The state of the system can therefore be defined in terms of the number of mutants, $m = 0, 1, \ldots, N$, where N is the total number of nodes on the network. The Moran process is then a birth-death Markov chain with $N + 1$ states, transition probabilities b_m and d_m determined by the network structure, and absorbing boundaries at $m = 0$ and $m = N$. In this section we review several general analytical results for absorbing birth-death Markov chains, explaining how they apply to fixation times in evolutionary dynamics. We also develop an approach, which we call *visit statistics*, that enables analytical estimation of the asymptotic fixation time cumulants.

On more complicated networks, the probability of adding or removing a mutant depends on the configuration of existing mutants. For some of these networks, however, the transition probabilities can be accurately estimated using a mean-field approximation [4, 30, 32, 33]. Then, to a good approximation,

the results below apply to such networks as well.

2.2.1 Eigendecomposition of the birth-death process

Assuming a continuous-time process, the state of the Markov chain described above evolves according to the master equation,

$$\dot{\mathbf{p}}(t) = \Omega \cdot \mathbf{p}(t), \tag{2.1}$$

where $\mathbf{p}(t)$ is the probability of occupying each state of the system at time t and Ω is the transition rate matrix, with columns summing to zero. In terms of the transition probabilities b_m and d_m, the entries of Ω are

$$\Omega_{mn} = b_n\delta_{m,n+1} + d_n\delta_{m,n-1} - (b_n + d_n)\delta_{m,n}, \tag{2.2}$$

where m and n run from 0 to N, $\delta_{m,n}$ is the Kronecker delta, and $b_0 = d_0 = b_N = d_N = 0$. The final condition guarantees the system has absorbing boundaries with stationary states $p_m = \delta_{m,0}$ and $p_m = \delta_{m,N}$ when the population is homogeneous. Thus we can decompose the transition matrix into stationary and transient parts, defining the transient part Ω_{tr} as in Eq. (2.2), but with $m, n = 1, \ldots, N-1$. The transient transition matrix acts on the transient states of the system, denoted $\mathbf{p}_{\text{tr}}(t)$. The eigenvalues of Ω_{tr} are real and strictly negative, since probability flows away from these states toward the absorbing boundaries. To ease notation in the following discussion and later applications, we shall refer to the positive eigenvalues of $-\Omega_{\text{tr}}$ as the eigenvalues of the transition matrix, denoted λ_m, where $m = 1, \ldots, N-1$.

From the perspective of Markov chains, the fixation time T is the time required for first passage to state $m = N$, given m_0 initial mutants, $p_m(0) = \delta_{m,m_0}$. At

time t, the probability that state N has been reached (i.e., the cumulative distribution function for the first-passage times) is simply $\varphi_{m_0}^{-1} p_N(t)$, where φ_{m_0} is the fixation probability given m_0 initial mutants. The distribution of first-passage times is therefore $\varphi_{m_0}^{-1} \dot{p}_N(t) = \varphi_{m_0}^{-1} b_{N-1} p_{N-1}(t)$. Since we normalize by the fixation probability, this is precisely the fixation-time distribution conditioned on reaching N.

The solution to the transient master equation is the matrix exponential $\mathbf{p}_{\mathrm{tr}}(t) = \exp(\Omega_{\mathrm{tr}} t) \cdot \mathbf{p}_{\mathrm{tr}}(0)$, yielding a fixation-time distribution $\varphi_{m_0}^{-1} b_{N-1} [\exp(\Omega_{\mathrm{tr}} t) \cdot \mathbf{p}_{\mathrm{tr}}(0)]_{N-1}$ [37]. If we assume one initial mutant $m_0 = 1$ this becomes $\varphi_1^{-1} b_{N-1} [\exp(\Omega_{\mathrm{tr}} t)]_{N-1,1}$. The matrix exponential can be evaluated in terms of the eigenvalues λ_m by taking a Fourier (or Laplace) transform (for details, see Ref. [29]). For a single initial mutant, the result is that the fixation time T has a distribution $f_T(t)$ given by

$$f_T(t) = \sum_{j=1}^{N-1} \left(\prod_{k=1, k \neq j}^{N-1} \frac{\lambda_k}{\lambda_k - \lambda_j} \right) \lambda_j e^{-\lambda_j t}. \tag{2.3}$$

This formula holds as long as the eigenvalues λ_m are distinct, which for birth-death Markov chains occurs when b_m and d_m are non-zero (except at the absorbing boundaries) [38]. Generalizations of this result for arbitrarily many initial mutants have also recently been derived, in terms of eigenvalues of the transition matrix and certain sub-matrices [29].

The distribution in Eq. (2.3) is exactly that corresponding to a sum of exponential random variables with rate parameters λ_m. The corresponding cumulants equal $(n-1)! \sum_{m=1}^{N-1} (\lambda_m)^{-n}$. As our primary interest is the asymptotic shape of the distribution, we normalize T to zero mean and unit variance and study $(T - \mu)/\sigma$, where μ and σ denote the mean and standard deviation of T. The standardized distribution is then given by $\sigma f_T(\sigma t + \mu)$. The rescaled fixation

time has cumulants

$$\kappa_n(N) = (n-1)!\left(\sum_{m=1}^{N-1}\frac{1}{\lambda_m^n}\right)\Bigg/\left(\sum_{m=1}^{N-1}\frac{1}{\lambda_m^2}\right)^{n/2}, \tag{2.4}$$

which, for many systems including those considered below, are finite as $N \to \infty$. When the limit exists, we define the asymptotic cumulants by $\kappa_n = \lim_{N\to\infty}\kappa_n(N)$. In particular, because we have standardized our distribution, the third cumulant κ_3 is the skew. In practice the limit $N \to \infty$ is taken by computing the leading asymptotic behavior of both the numerator and denominator in Eq. (2.4). As we will see below the scaling of these terms with N depends on both the population network structure and the mutant fitness (see also asymptotic analysis in the Appendix, Sections 2.9.4 & 2.9.5). This approach allows us to characterize the asymptotic shape of the fixation-time distribution in terms of the constants κ_n. Since $\lambda_m > 0$, it is clear from this expression that, for finite N, the skew and all higher order cumulants must be positive, in agreement with results for random walks with non-uniform bias [39]. As $N \to \infty$ this is not necessarily true; in some cases the cumulants vanish.

The eigendecomposition gives the fixation-time distribution and cumulants in terms of the non-zero eigenvalues of the transition matrix. In general the eigenvalues must be found numerically, but in cases where they have a closed form expression the asymptotic form of the cumulants and distribution can often be obtained exactly.

2.2.2 Analytical cumulant calculation: Visit statistics

In this section we develop machinery to compute the cumulants of the fixation time analytically without relying on matrix eigenvalues. For this analysis, we

specialize to cases where $b_m/d_m = r$ for all m, relevant for the Moran processes considered below. These processes can be thought of as biased random walks overlaid with non-constant waiting times at each state.

It is helpful to consider the Markov chain conditioned on hitting N, with new transition probabilities $\tilde{b}_m$ and $\tilde{d}_m$ so that the fixation probability $\varphi_{m_0} = 1$. If X_t is the state of the system at time t, then $\tilde{b}_m = \mathcal{P}(X_t = m \to X_{t+1} = m + 1 | X_\infty = N)$ with $\tilde{d}_m$ defined analogously. We derive explicit expressions for $\tilde{b}_m$ and $\tilde{d}_m$ in the Appendix, Section 2.9.2. Conditioning is equivalent to a similarity transformation on the transient part of the transition matrix: $\tilde{\Omega}_{\text{tr}} = S\,\Omega_{\text{tr}}\,S^{-1}$, where S is diagonal with $S_{mm} = 1 - 1/r^m$. Furthermore, since $b_m/d_m = r$, we can decompose $\Omega_{\text{tr}} = \Omega_{\text{RW}} D$, where D is a diagonal matrix, $D_{mm} = b_m + d_m$, that encodes the time spent in each state and Ω_{RW} is the transition matrix for a random walk with uniform bias,

$$[\Omega_{\text{RW}}]_{nm} = \frac{r}{1+r}\delta_{m,n+1} + \frac{1}{1+r}\delta_{m,n-1} - \delta_{m,n}. \tag{2.5}$$

Applying the results of the previous section and using the fact that the columns of Ω sum to zero, we can write there fixation-time distribution of the conditioned Markov chain as $f_T(t) = -\mathbf{1}\tilde{\Omega}_{\text{tr}} \exp(\tilde{\Omega}_{\text{tr}} t)\mathbf{p}_{\text{tr}}(0)$, where $\mathbf{1}$ is the row vector containing all ones. This distribution has characteristic function [37]

$$\phi(\omega) \coloneqq E[\exp(i\omega T)] = \mathbf{1}\tilde{\Omega}_{\text{tr}}(i\omega + \tilde{\Omega}_{\text{tr}})^{-1}\mathbf{p}_{\text{tr}}(0). \tag{2.6}$$

and the derivatives $(-i)^n \phi^{(n)}(0)$ give the moments of T

$$E[T^n] = (-1)^n n! \mathbf{1}\tilde{\Omega}_{\text{tr}}^{-n}\mathbf{p}_{\text{tr}}(0), \tag{2.7}$$

in terms of $\tilde{\Omega}_{\text{tr}}^{-1} = D^{-1} S \Omega_{\text{RW}}^{-1} S^{-1}$. This inverse has a nice analytical form because S and D are diagonal and Ω_{RW} is tridiagonal Toeplitz. We call this approach *visit statistics* because the elements V_{ij} of $V = -S\Omega_{\text{RW}}^{-1}S^{-1}$ encode the average number of visits to state i starting from state j.

Each power of $\tilde{\Omega}_{\rm tr}$ in Eq. (2.7) produces products of $(b_i + d_i)$ that arise in linear combinations determined by the visit numbers V_{ij}. Therefore, the cumulants of the fixation time have the general form

$$\kappa_n(N) = \frac{\displaystyle\sum_{i_1,i_2,\ldots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r, N|m_0)}{(b_{i_1} + d_{i_1})(b_{i_2} + d_{i_2}) \cdots (b_{i_n} + d_{i_n})}}{\left(\displaystyle\sum_{i,j=1}^{N-1} \frac{w^2_{ij}(r, N|m_0)}{(b_i + d_i)(b_j + d_j)} \right)^{n/2}}, \tag{2.8}$$

where $w^n_{i_1 i_2 \cdots i_n}(r, N|m_0)$ are weighting factors based on the visit statistics of the biased random walk, given the initial number of mutants m_0. In what follows, we always assume $m_0 = 1$ and suppress the dependence of the weighting factors on initial condition, writing $w^n_{i_1 i_2 \cdots i_n}(r, N)$ instead. A detailed derivation of Eq. (2.8) and explicit expressions for $w^2_{ij}(r, N)$ and $w^3_{ijk}(r, N)$ are given in the Appendix, Section 2.9.1 below.

To the best of our knowledge this representation of the fixation-time cumulants has not been previously derived, although a similar approach was recently used to compute mean fixation times for evolutionary dynamics on complex networks [4]. This expression is equivalent to the well-known recurrence relations for absorption-time moments of birth-death processes [28, 40] but is easier to handle asymptotically, and can be useful even without explicit expressions for $w^n_{i_1 i_2 \cdots i_n}(r, N)$. Estimating the sums in Eq. (2.8) allows us to compute the asymptotic fixation time cumulants exactly.

2.2.3 Recurrence relation for fixation-time moments

Evaluation of the eigenvalues of the transition matrix for large systems can be computationally expensive, with the best algorithms having run times

quadratic in matrix size. Numerical evaluation of the expression given in Eq. (2.8) is even worse, as it requires summing $O(N^n)$ elements. If only a finite number of fixation time cumulants (and not the full distribution) are desired, there are better numerical approaches. Using standard methods from probability theory [41], we derive a recurrence relation that allows numerical moment computation with run time linear in system size N. For completeness we provide the full derivation of the reccurence for the fixation-time skew in the Appendix, Section 2.9.3.

2.2.4 Equivalence between advantageous and disadvantageous mutations

In the following applications, we will generally speak of the mutants as having a fitness advantage, designated by the parameter $r > 1$. Our results, however, can be immediately extended to disadvantageous mutations. In particular, the fixation-time distributions (conditioned on fixation occurring) for mutants of fitness r and $1/r$ are identical. When a mutant with fitness $1/r$ is introduced into the population (and eventually reaches fixation), the non-mutants are r times as fit as the mutants. Therefore, this system is equivalent to another system that starts with $N - 1$ fitness r mutants which eventually die out (the mutants in the former system are the non-mutants in the latter). It has been shown that the times to go from one initial mutant to fixation ($m = 1 \rightarrow m = N$) and from $N - 1$ initial mutants to extinction ($m = N - 1 \rightarrow m = 0$) have identical distribution [29]. Thus indeed, the conditioned fixation-time distributions are identical for mutants of fitness r and $1/r$. Of course the fixation probability is very different

in the two cases: for the disadvantageous mutations it approaches 0 for large N [13].

2.3 One-Dimensional Lattice

We now specialize to Moran Birth-death (Bd) processes, starting with the one-dimensional (1D) lattice. We assume periodic boundary conditions, so that the N nodes form a ring. The mutants have relative fitness r, meaning they give birth r times faster, on average, than non-mutants do.

Starting from one mutant, suppose that at some later time m of the N nodes are mutants. On the 1D lattice, the population of mutants always forms a connected arc, with two mutants at the endpoints of the arc. Therefore, the probability b_m of increasing the mutant population by one in the next time step is the probability of choosing a mutant node at an endpoint to give birth, namely $2r/(rm+N-m)$, times the probability $1/2$ that the neighboring node to be replaced is not itself a mutant. (The latter probability equals $1/2$ because there are two neighbors to choose for replacement: a mutant neighbor on the interior of the arc and a non-mutant neighbor on the exterior. Only the second of these choices produces an increase in the number of mutants.) Multiplying these probabilities together we obtain

$$b_m = \frac{r}{rm + N - m}, \qquad d_m = \frac{1}{rm + N - m}, \tag{2.9}$$

where the probability d_m of decreasing the mutant population by one is found by similar reasoning. Note that this derivation fails for $m = 1$ ($m = N - 1$) when the arc of mutants (non-mutants) contains only one node, but one can check Eq. (2.9) still holds for these cases. These quantities play the role of transition

probabilities in a Markov transition matrix. The next step is to find the eigenvalues of that matrix.

2.3.1 Neutral fitness

First we work out the eigenvalues for the case of neutral fitness, $r = 1$. In this case, the transition probabilities are equal, $b_m = d_m = 1/N$, and independent of m. Therefore, the Moran process is simply a random walk, with events occurring at a rate of $2/N$ per time step. The associated transition matrix is tridiagonal Toeplitz, which has eigenvalues given by

$$\lambda_m = \frac{2}{N} - \frac{2}{N}\cos\left(\frac{m\pi}{N}\right), \quad m = 1, 2, \ldots, N-1. \tag{2.10}$$

Applying Eq. (2.4) and computing the leading asymptotic form of the sums $S_n = \sum_{m=1}^{N-1}(\lambda_m)^{-n}$ (see Appendix, Section 2.9.4), we find that as $N \to \infty$, the fixation-time distribution has cumulants

$$\kappa_n = (n-1)!\,\frac{\zeta(2n)}{\zeta(4)^{n/2}}, \tag{2.11}$$

where ζ denotes the Riemann zeta function. In particular, the skew $\kappa_3 = 4\sqrt{10}/7 \approx 1.807$, as previously calculated by Ottino-Löffler et al. [33] via martingale methods. The other cumulants (and characteristic function below) haven't previously been computed for the Bd process on the 1D lattice. The largeness of the skew stems from the recurrent property of the random walk. As $N \to \infty$, long walks with large fixation times become common and the system revisits each state infinitely often [42].

Knowledge of the cumulants allows us to obtain the exact characteristic

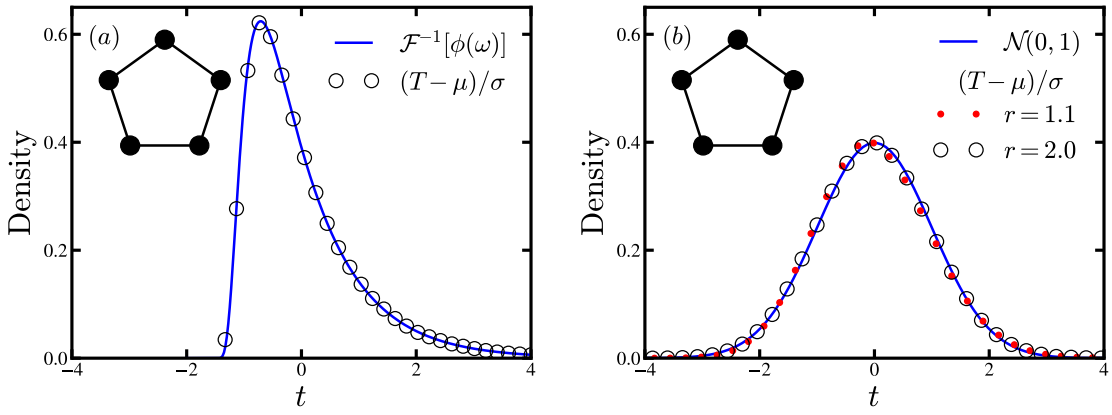

Figure 2.2: Fixation-time distributions on the 1D lattice obtained from 10^6 simulation runs. All distributions are standardized to zero mean and unit variance. Solid curves are the theoretical predictions. Shown are the fixation-time distributions for (a) a 1D lattice of $N = 100$ nodes with neutral fitness $r = 1$ and (b) a 1D lattice of $N = 5000$ nodes with mutant fitnesses $r = 1.1$ and $r = 2.0$. For the neutral fitness case, the theoretical distribution was generated by numerical inverse Fourier transform of the characteristic function (Eq. (2.12)). The $r = 1.1$ distribution is slightly but visibly skewed due to finite network size.

function of the fixation-time distribution:

$$\phi(\omega) = e^{-\sqrt{\frac{5}{2}\omega}}\,\Gamma\left(1 - \frac{90^{1/4}\sqrt{\omega}}{\pi}\right)\Gamma\left(1 + \frac{90^{1/4}\sqrt{\omega}}{\pi}\right). \tag{2.12}$$

Although we cannot find a simple expression for the distribution itself, we can efficiently evaluate it by taking the inverse Fourier transform of the characteristic function numerically. Figure 2.2(a) shows that the predicted fixation-time distribution agrees well with simulations.

2.3.2 Non-neutral fitness

Next, consider $r \neq 1$ with the transition probabilities given by Eq. (2.9). Then the eigenvalues of the transition matrix are no longer expressible in closed form. If r is not too large, however, the probabilities b_m and d_m do not vary dramatically

with m, the number of mutants. In particular, $b_m \sim 1/N$ for all m when N is large. Therefore, as a first approximation we treat the Bd process on a 1D lattice as a biased random walk with $b_m = r/(1+r)$ and $d_m = 1/(1+r)$. The eigenvalues of the corresponding transition matrix are

$$\lambda_m = 1 - \frac{2\sqrt{r}}{1+r}\cos\left(\frac{m\pi}{N}\right), \quad m = 1, 2, \ldots, N-1. \tag{2.13}$$

The cumulants again involve sums $S_n = \sum_{m=1}^{N-1}(\lambda_m)^{-n}$, which can be approximated in the limit $N \to \infty$ by,

$$S_n \approx \frac{N}{\pi}\int_0^{\pi} \frac{1}{(1 - 2\sqrt{r}/(1+r)\cos x)^n}\,\mathrm{d}x. \tag{2.14}$$

Since the integral is independent of N and converges for $r \neq 1$, each of the sums scales linearly: $S_n \sim N$. Thus, using Eq. (2.4), we see that all cumulants past second order approach 0,

$$\kappa_n \sim \frac{1}{N^{(n-2)/2}} \xrightarrow{N\to\infty} 0, \quad n \geq 3. \tag{2.15}$$

Hence the fixation-time distribution is asymptotically normal, independent of fitness level.

By evaluating the integrals in Eq. (2.14), we can more precisely compute the scaling of the cumulants. For the skew we find

$$\kappa_3 \approx \frac{2 + 2r(r+4)}{(r+1)\sqrt{(r^2-1)}}\frac{1}{\sqrt{N}}. \tag{2.16}$$

The integral approximation becomes accurate when the first term in the sums S_n becomes close to the value of the integrand evaluated at the lower bound ($x = 0$). The fractional difference between these quantities is

$$\begin{aligned}\Delta &= \left|\frac{(1 - 2\sqrt{r}/(1+r))^n}{(1 - 2\sqrt{r}/(1+r)\cos(\pi/N))^n} - 1\right| \\ &= \frac{\sqrt{r}n\pi^2}{(\sqrt{r}-1)^2 N^2} + O(1/N^4).\end{aligned} \tag{2.17}$$

Then we have $\Delta \ll 1$ when $N \gg N_c$ where $N_c \approx 2\pi\sqrt{n}/(r-1)$ (assuming r is near 1). For the skew, we require the sums with $n = 2$ and 3, giving $N_c \approx 10/(r-1)$.

The above calculation fails for $r \gg 1$, because when $r = \infty$ the transition probabilities $b_m = 1/m$ have different asymptotic behavior as $N \to \infty$. In particular, more time is spent waiting at states with large m. The process still has normally distributed fixation times [33], but the skew becomes

$$\kappa_3^\infty = 2\left(\sum_{m=1}^{N-1} m^3\right)\Big/\left(\sum_{m=1}^{N-1} m^2\right)^{3/2} \approx \frac{3\sqrt{3}}{2}\frac{1}{\sqrt{N}}, \tag{2.18}$$

for large N. Notice that the coefficient is different from that given by the infinite-r limit of Eq. (2.16), $\kappa_3 \approx 2/\sqrt{N}$. We conjecture that there is a smooth crossover between these two scaling laws with the true skew given approximately by

$$\tilde{\kappa}_3 = \kappa_3\left[r^{-q} + \frac{3\sqrt{3}}{4}(1 - r^{-q})\right] \tag{2.19}$$

for some exponent q, where κ_3 is the skew given in Eq. (2.16). For small r this ansatz has skew similar to that of a random walk, but captures the correct large-r limit. We do not have precise theoretical motivation for this ansatz, but as discussed below, it works quite well.

Numerical calculation of the skew for the 1D lattice was performed using the recurrence relation method discussed in Section 2.2.3. The results are shown in Figure 2.3 for a few values of r. This calculation confirms our initial hypothesis, near neutral fitness the waiting times are uniform enough that the process is well approximated by a biased random walk and the skew approaches 0, scaling in excellent agreement with Eq. (2.16). When $N \ll N_c$, the bias is not sufficient to give the mutants a substantial advantage: the process is dominated by drift and the fixation-time distribution has large skew $\kappa_3 \approx 1.807$, as found in the preceding section. For $N \gg N_c$, selection takes over, the cumulants approach 0,

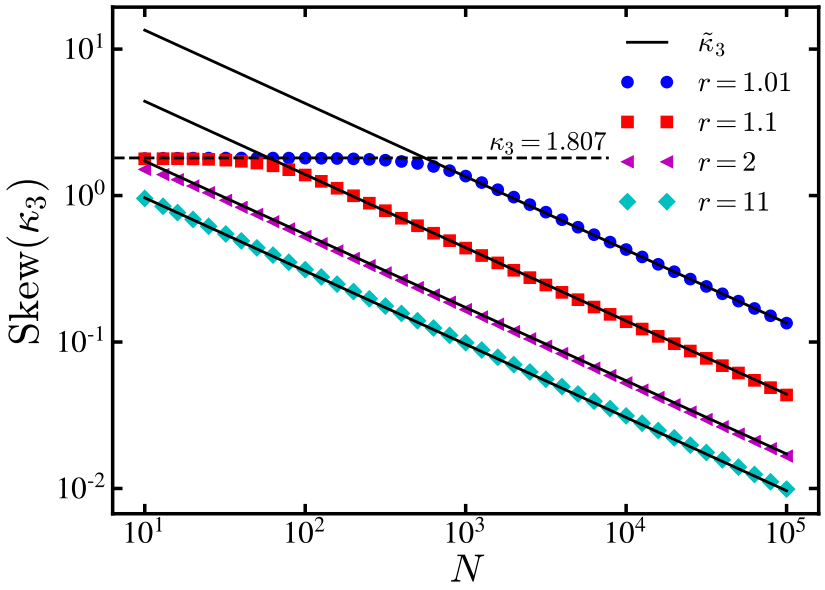

Figure 2.3: Scaling of the skew of the fixation-time distribution on the 1D lattice with non-neutral fitness. Data points show numerical calculation of the skew for various fitness levels. The solid lines are the predicted scaling given in Eq. (2.19) with exponent $q = 1/2$ for each value of fitness r. For small N (and small enough r), the skew is that of a random walk, namely $\kappa_3 = 1.807$, as shown by the dashed line. For large N, the skew $\kappa_3 \sim 1/\sqrt{N}$ with an r-dependent coefficient.

and the distribution becomes normal. A similar crossover appears in the study of the fixation probability, where a transition from $\varphi_1 \sim 1/N$ to $\varphi_1 \sim 1 - 1/r$ is seen when N passes a critical system size (that is slightly different than N_c). For large fitness $r \gg 1$, the ansatz Eq. (2.19) captures the scaling behavior if we use an exponent $q = 1/2$. Direct numerical simulations of the process confirm that, for any $r > 1$, the fixation time on the 1D lattice has an asymptotically normal distribution [Figure 2.2(b)].

The random walk approximation allows us to find the asymptotic scaling of the fixation-time cumulants, but ignores the heterogeneity of waiting times present in the Moran process. Using visit statistics we can compute the cumulants exactly with Eq. (2.8) and rigorously prove they vanish as $N \to \infty$, ver-

ifying that the waiting times have no influence on the asymptotic form of the distribution. Details are provided in the Appendix, Section 2.9.4.

Our analysis of the 1D lattice reveals an intriguing universality property of its fixation-time distribution. For any value of relative fitness r other than $r = 1$, the fixation-time distribution approaches a normal distribution as $N \to \infty$. Thus, for $r \neq 1$ the asymptotic shape of the distribution is universal and independent of r (though bear in mind, its mean and variance do depend on r).

When $r = 1$, corresponding to precisely neutral fitness, the unbiased random walk yields a qualitatively different distribution with considerably larger skew. This qualitative change as r passes through unity leads to a discontinuous jump in the skew at $r = 1$.

As one might expect, the discontinuity stems from passage to the infinite-N limit. For finite but large N, the distribution varies continuously with r, though our numerical results indicate that the sharp increase in skew still occurs very close to $r = 1$. We will see in the next section that the discontinuity and highly skewed distribution at neutral fitness persist when we alter the network structure from a locally connected 1D lattice to a fully connected complete graph.

2.4 Complete Graph

Next we consider the Moran process on a complete graph, useful for modeling well-mixed populations in which all individuals interact. By similar reasoning to above, given m mutants the probability of adding a mutant in the next time

step is

$$b_m = \frac{rm}{rm + N - m} \cdot \frac{N - m}{N - 1}, \tag{2.20}$$

while the probability of subtracting a mutant is

$$d_m = \frac{N - m}{rm + N - m} \cdot \frac{m}{N - 1}. \tag{2.21}$$

Interestingly, as we will see in this section, these transition probabilities give rise to a fitness dependent fixation-time distribution, in stark contrast to the universality of the normal distribution observed on the 1D lattice.

2.4.1 Neutral Fitness

Again we begin with neutral fitness $r = 1$. Now $b_m = d_m = (Nm - m^2)/(N^2 - N)$. The eigenvalues of this transition matrix also have a nice analytical form:

$$\lambda_m = \frac{m(m+1)}{N(N-1)}, \quad m = 1, 2, \ldots N - 1. \tag{2.22}$$

The asymptotic form of the sums $S_n = \sum_{m=1}^{N-1} (\lambda_m)^{-n}$, can be found by taking the partial fraction decomposition of $(\lambda_m)^{-n}$ and evaluating each term individually. The resulting cumulants are

$$\begin{aligned} \kappa_n &= (n-1)! \frac{3^{n/2}}{(\pi^2 - 9)^{n/2}} \\ &\quad \times (-1)^n \sum_{k=1}^{n} \binom{2n-k-1}{n-1} \left[\zeta(k) \left(1 + (-1)^k \right) - 1 \right]. \end{aligned} \tag{2.23}$$

Our knowledge of the eigenvalues also allows us to obtain a series expression for the asymptotic distribution using Eq. (2.3). For $N \to \infty$ the standardized distribution is,

$$\begin{aligned} \sigma f_T(\sigma t + \mu) = c_\sigma \sum_{j=1}^{\infty} (-1)^{j+1} j(j+1)(2j+1) \\ \times \exp\left[j(j+1)(c_\sigma t + 1) \right], \end{aligned} \tag{2.24}$$

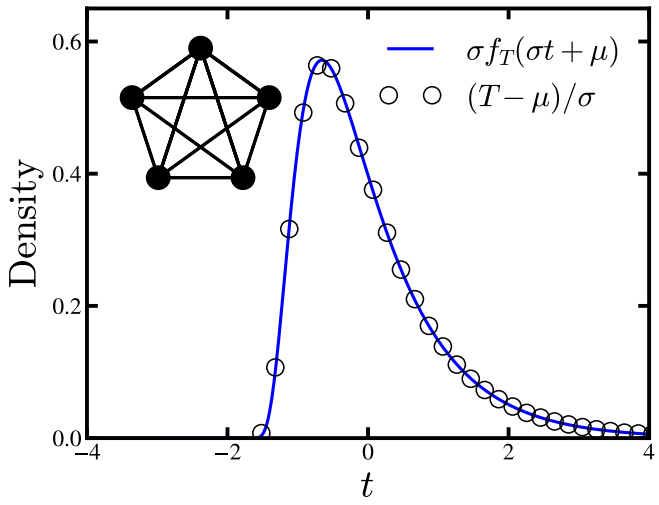

Figure 2.4: Fixation-time distributions on the complete graph with $N = 100$ nodes and neutral fitness ($r = 1$) obtained from 10^6 simulation runs. The distribution is standardized to zero mean and unit variance. The solid curve is the theoretical distribution obtained by numerically evaluating the infinite series in Eq. (2.24) for each value of t.

where to leading order in N the mean and standard deviation are $\mu = N^2$ and $\sigma = c_\sigma N^2$, with $c_\sigma = \sqrt{\pi^2/3 - 3}$. This distribution was previously found using a different approach by Kimura, who also computed the first few fixation-time moments [43]. We have extended these results, obtaining the cumulants to all orders. Figure 2.4 shows that the predicted asymptotic distribution agrees well with numerical experiments.

The numerical value of the fixation-time skew for the Birth-death process on the complete graph is $\kappa_3 = 6\sqrt{3}(10-\pi^2)/(\pi^2-9)^{3/2} \approx 1.6711$, slightly less than that for the 1D lattice. This decrease is the result of two competing effects contributing to the skew. First, since the birth and death transition probabilities are the same, the process is a random walk, which has a highly skewed fixation-time distribution, as shown above. The average time spent in each state, however, varies with m. For instance, when $m = 1$ or $N - 1$, $b_m \to 0$ for large N. But if

$m = \alpha N$ for some constant $0 < \alpha < 1$ independent of N, then b_m approaches a constant.

Intuitively, the beginning and end of the mutation-spreading process are very slow because the transition probabilities are exceedingly small. To start, the single mutant must be selected by chance to give birth from the N available nodes, a selection problem like finding a needle in a haystack. Similarly, near fixation the reproducing mutant must find and replace one of the few remaining non-mutants, again choosing it by chance from an enormous population.

The characteristic slowing down at certain states is reminiscent of "coupon collection", as discussed earlier. Erdős and Rényi proved that for large N, the normalized time to complete the coupon collection follows a Gumbel distribution [36], which we denote by Gumbel(α, β) with density

$$f(t) = \beta^{-1} e^{-(t-\alpha)/\beta} \exp(-e^{-(t-\alpha)/\beta}). \tag{2.25}$$

For the Moran process, each slow region is produced by long waits for the random selection of rare types of individuals: either mutants near the beginning of the process or non-mutants near the end. In the next section we show that the two coupon collection regions of the Bd process on a complete graph lead to fixation-time distributions that are convolutions of two Gumbel distributions. In the case of neutral fitness, these Gumbel distributions combine with the random walk to produce a new highly skewed distribution with cumulants given by Eq. (2.23).

2.4.2 Non-neutral fitness

We saw in Section 2.3.2 that when the average time spent in each state is constant or slowly varying the fixation-time distribution is asymptotically normal. Birth-death dynamics on the complete graph, however, exhibit coupon collection regions at the beginning and end of the process, where the transition probabilities vanish. We begin this section with a heuristic argument that correctly gives the asymptotic fixation-time distribution in terms of independent iterations of coupon collection.

Differentiating b_m with respect to m, we find the slope near $m = 0$ is $(r + 1)/N$, while the slope near $m = N$ has magnitude $(r + 1)/(rN)$ for $N \gg 1$. The transition rates approach zero at each of these points, so we expect behavior similar to coupon collection giving rise to two Gumbel distributions. Since the slope is greater for m near 0 than for m near N, the Moran process completes its coupon collection faster near the beginning of the process than near fixation.

This heuristic suggests that the asymptotic fixation time should be equal in distribution to the sum of two Gumbel random variables, one weighted by r, which is the ratio of the slopes in the coupon collection regions. Specifically, if T is the fixation time with mean μ and variance σ^2, we expect

$$\frac{T - \mu}{\sigma} \xrightarrow{d} \frac{G + rG'}{\sqrt{1 + r^2}}, \tag{2.26}$$

where $\xrightarrow{d}$ means convergence in distribution for large N. Here G and G' denote independent and identically distributed Gumbel random variables with zero mean and unit variance. It is easy to check that the correct distribution is Gumbel$(-\gamma\sqrt{6}/\pi, \sqrt{6}/\pi)$, where $\gamma \approx 0.5772$ is the Euler-Mascheroni constant.

Let us make this argument more rigorous. Previous theoretical analysis

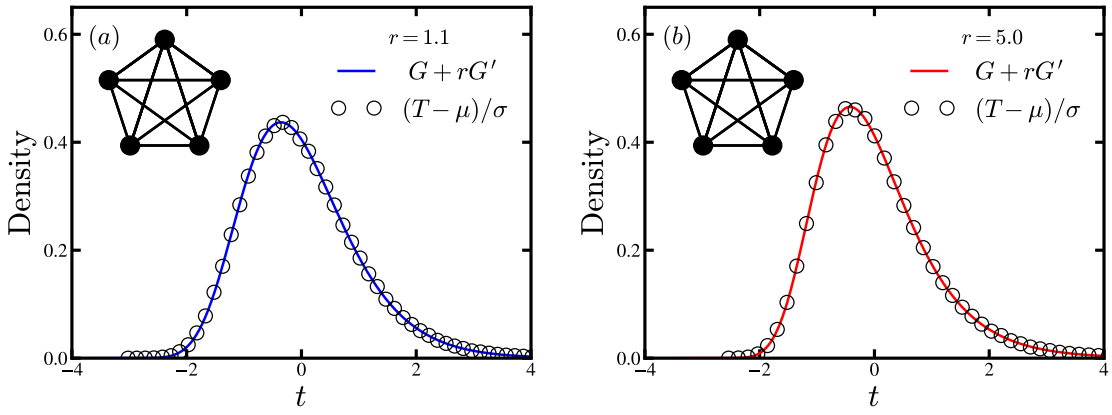

Figure 2.5: Fixation-time distributions on the complete graph with $N = 5000$ nodes and non-neutral fitness ($r > 1$) obtained from 10^6 simulation runs. All distributions are standardized to zero mean and unit variance. Solid curves are the theoretical predictions obtained by numerical convolution of two Gumbel distributions, one weighted by r. Distributions are shown for (a) $r = 1.1$ and (b) $r = 5.0$. For larger r, the distribution has larger skew and a slightly sharper peak.

showed that in the infinite fitness limit, the fixation time has an asymptotically Gumbel distribution [33]. This result can be recovered within our framework, since when $r = \infty$ it follows that $d_m = 0$, so the eigenvalues of the transition matrix are just $\lambda_m = b_m = (N - m)/(N - 1)$ and the cumulants can be directly calculated using Eq. (2.4).

For large (but not infinite) fitness, the number of mutants is monotonically increasing, to good approximation, since the probability that the next change in state increases the mutant population is $r/(1 + r) \approx 1$. The time spent waiting in each state, however, changes dramatically, especially near $m = 1$. Here, $b_1 \to 0$ for large N, in stark contrast to the infinite fitness system where $b_1 \to 1$. The time spent at each state, t_m is an exponential random variable, $\mathcal{E}(b_m + d_m)$. In this approximation each state is visited exactly once, so the total fixation time is a

sum of these waiting times:

$$T \approx \sum_{m=1}^{N-1} \mathcal{E}(b_m + d_m). \tag{2.27}$$

But this sum of exponential random variables has density given by Eq. (2.3), with the substitution $\lambda_m \to b_m + d_m$. Thus, the cumulants of $(T - \mu)/\sigma$ are

$$\kappa_n = \frac{1 + r^n}{(1 + r^2)^{n/2}} \times \frac{(n-1)!\zeta(n)}{\zeta(2)^{n/2}}, \tag{2.28}$$

which are exactly the cumulants corresponding to the sum of Gumbel random variables given in Eq. (2.26). In the limit $r \to \infty$, the first term in Eq. (2.28) becomes 1, and the cumulants are those for a single Gumbel distribution, in agreement with previous results [33].

Remarkably, these cumulants are exact for any $r > 1$, not just in the large-r limit. We can see this directly for the skew κ_3 using the visit statistics approach, computing the asymptotic form of Eq. (2.8) with the complete graph transition probabilities, Eqs. (2.20) and (2.21). Details of the asymptotic analysis are provided in the Appendix, Section 2.9.5. Numerical simulations of the Moran process corroborate our theoretical results. As shown in Figure 2.5, for $r = 1.1$ and $r = 5$ the agreement between simulated fixation times and the predicted convolution of Gumbel distributions is excellent, at least when N is sufficiently large. Again, our calculations show a discontinuity in the fixation-time distribution at $r = 1$. In particular, the $r \to 1$ limit of the cumulants for non-neutral fitness in Eq. (2.28) is not the same as the cumulants for neutral fitness found in the preceding section [Eq. (2.23)].

For smaller networks, it is fascinating to see how the results converge to the asymptotic predictions as N grows. Figure 2.6 shows how the skew of the fixation-time distribution depends on r and N for the complete graph. As dis-

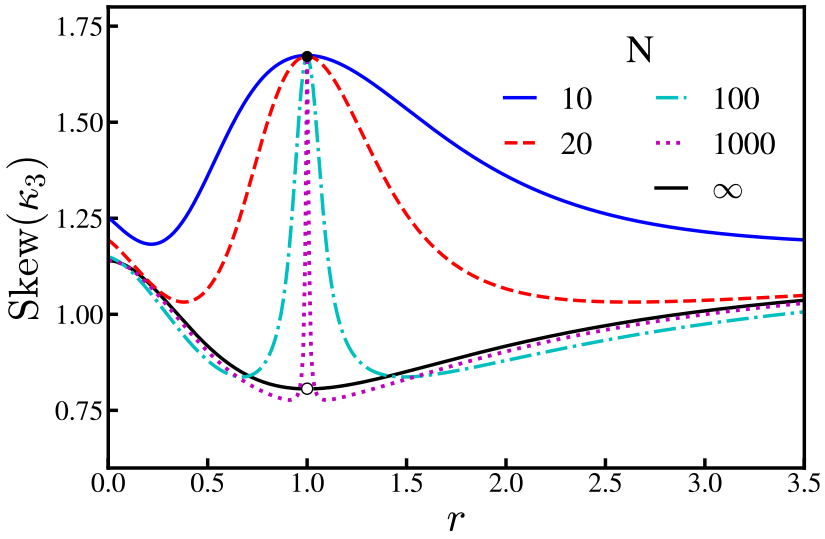

Figure 2.6: Fitness dependence of fixation-time skew for the Moran Birth-death process on the complete graph. The skew is shown for $r \geq 0$ and is invariant under $r \to 1/r$. For finite N, the skew does not have a discontinuity, but does show non-monotonic dependence on fitness r. In particular, for a given N, there is a certain fitness level with minimum skew. As $N \to \infty$, we see non-uniform convergence to the predicted skew given by κ_3 in Eq. (2.28), leading to the discontinuity at $r = 1$. Moreover, for fixed r, the convergence to the $N = \infty$ skew is non-monotonic.

cussed in Section 2.2.4, the fixation-time distributions for these systems are invariant under $r \to 1/r$. Therefore we show the skew for all $r > 0$, to emphasize the intriguing behavior near neutral fitness, where $r = 1$. We find that non-uniform convergence of the fixation-time skew leads to the discontinuity predicted at $r = 1$. For finite N, the skew is a non-monotonic function of r and has a minimum value at some fitness $r_{\min}(N)$. Furthermore, at fixed r, the convergence to the $N = \infty$ limit is itself non-monotone. Though beyond the scope of the current study, further investigation of this finite-N behavior would be worth pursuing.

2.5 Partial fixation times

In many applications, we may be interested in the time to partial fixation of the network. For instance, considering cancer progression [44–46] or the incubation of infectious diseases [33], symptoms can appear in a patient even when a relatively small proportion of cells are malignant or infected. We therefore consider T_α, the total time to first reach αN mutants on the network, where $0 < \alpha < 1$. The methods developed in Section 2.2 apply to these processes as well. For the eigendecomposition approach we instead use the sub-matrix of Ω_{tr} containing the first αN rows and columns. In calculations involving the numerical recurrence relations or visit statistics, we simply cut the sums off at αN instead of N and for the latter, replace $w_{i_1 i_2 \cdots i_n}(r, N)$ with $w_{i_1 i_2 \cdots i_n}(r, \alpha N)$.

2.5.1 One-dimensional lattice

Truncating the Moran Bd process on the 1D lattice by a factor α has no effect on the asymptotic shape of the fixation-time distributions. In both the neutral fitness system and the random walk approximation to the non-neutral fitness system, the transition matrix has no explicit dependence on the state or system size [aside from proportionality factors that cancel in Eq. (2.4)]. Thus, the eigenvalues are identical to those calculated previously, but correspond to a smaller effective system size αN. Taking the limit $N \to \infty$ therefore yields the same asymptotic distributions found in Section 2.3.

2.5.2 Complete graph: truncating coupon collection

The complete graph exhibits more interesting dependence on truncation. Since the transition probabilities have state dependence, the eigenvalues change with truncation (they don't correspond to the same system with smaller effective N). Our intuition from coupon collection, however, lets us predict the resulting distribution.

First consider non-neutral fitness. Then there are two coupon collection stages, one near the beginning and another near the end of the process, and together they generate a fixation-time distribution that is a weighted convolution of two Gumbel distributions. The effect of truncating the process near its end should now become clear: it simply removes the second coupon collection. The truncated process stops before the mutants have to laboriously find and replace the last remaining non-mutants. Therefore, we intuitively expect the fixation time for non-neutral fitness to be distributed according to a single Gumbel distribution, regardless of fitness level.

The only exception occurs if $r = \infty$; then no coupon collection occurs at the beginning of the process either, as the lone mutant is guaranteed to be selected to give birth in the first time step, thanks to its infinite fitness advantage. Thus, when fitness is infinite and the process is truncated at the end, both coupon collection phases are removed and the fixation times are normally distributed.

Similar reasoning applies to the Birth-death process with neutral fitness. It also has two coupon collection regions, one of which is removed by truncation. In this case, however, the random walk mechanism contributes to the skew of the overall fixation-time distribution, combining non-trivially with the coupon

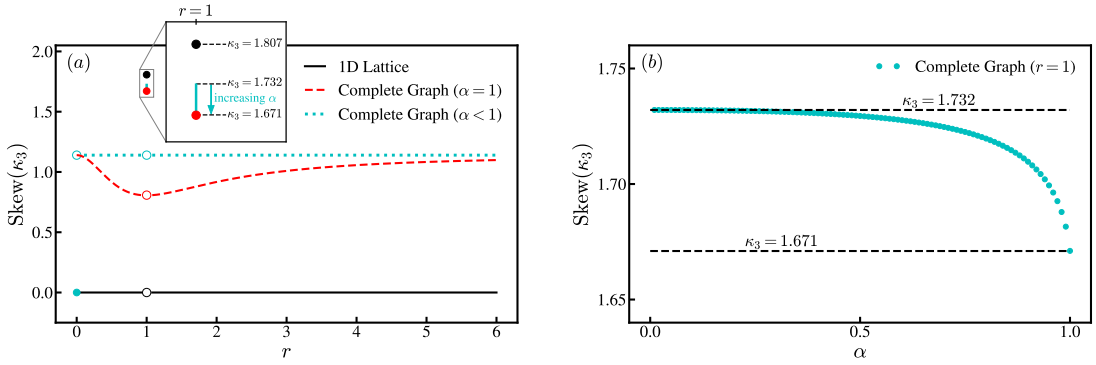

Figure 2.7: Variation of fixation-time skew κ_3 with fitness level r and truncation factor α for different network structures. (a) The skew of the fixation-time distribution is plotted versus fitness for the 1D lattice (black solid line), complete graph (red dashed line), and complete graph with truncation (green dotted line). The skew is shown for all $r \geq 0$ and is invariant under $r \to 1/r$. When $r \neq 1$ and $r < \infty$, the fixation-time distribution is normal for the 1D lattice, and hence has zero skew ($\kappa_3 = 0$). The distribution becomes a fitness-weighted convolution of Gumbel distributions for the complete graph, and a single Gumbel distribution for the complete graph with truncation (for any $\alpha < 1$). Each curve jumps discontinuously at $r = 1$, where the distributions become highly skewed with $\kappa_3 > 1.5$. The inset shows a blow-up of the neutral fitness results, specifying the skew for each case. On the complete graph with truncation, the skew is continuously variable at $r = 1$, taking on an interval of values between $6\sqrt{3}(10 - \pi^2)/(\pi^2 - 9)^{3/2} \approx 1.671$ when $\alpha = 1$, and $\sqrt{3} \approx 1.732$ when $\alpha = 0$. This range is indicated by the green vertical line. The truncated fixation time on the complete graph has a second discontinuity at $r = \infty$ (shown here at $r = 0$, by exploiting the $r \to 1/r$ invariance). At this discontinuity the functional form of the distribution jumps from Gumbel to normal. (b) The fixation-time skew for the complete graph with neutral fitness, plotted versus the truncation factor α. These points correspond to the green vertical line in panel (a) at $r = 1$.

collection-like process. We find that the skew of the fixation time depends on the truncation factor α, varying between $6\sqrt{3}(10 - \pi^2)/(\pi^2 - 9)^{3/2} \approx 1.6711$ when $\alpha = 1$, and $\sqrt{3} \approx 1.732$ when $\alpha = 0$. A derivation of this $\alpha \to 0$ limit is given in the Appendix, Section 2.9.5.

2.5.3 Summary of main results

The main results from Sections 2.3–2.5 are summarized in Figure 2.7, which shows the asymptotic fitness dependence of fixation-time skew for each network considered in this chapter. We again show the skew for all $r > 0$ (not just $r > 1$) to emphasize the discontinuities at zero, neutral, and infinite fitness. On the 1D lattice, independent of the truncation factor α, the Bd process has normally distributed fixation times, except at neutral fitness where the distribution is highly skewed. The complete graph fixation-time distributions are the weighted convolution of two Gumbel distributions for $r \neq 1$, again with a highly skewed distribution at $r = 1$. With truncation by a factor $\alpha < 1$, the distribution for the complete graph is Gumbel for $1 < r < \infty$, and normal for $r = \infty$. With neutral fitness the fixation distribution is again highly skewed, with skew dependent on the truncation factor α.

2.6 Extensions

It is natural to ask whether our results are generic; do the same fixation-time distributions appear in other models of evolutionary dynamics? Here we explore the robustness of our results to various changes in the model update dynamics and the network topology. The main finding is that our results are insensitive to these changes, at least qualitatively. The distributions typically remain right-skewed and even follow the same functional forms derived above.

2.6.1 Other update dynamics

Two-fitness Moran process

The Moran Bd processes considered above require a single fitness level, designating the relative reproduction rates between mutants and non-mutants. Another common model is the Moran Birth-Death (BD) process, which has a second fitness level $\tilde{r}$ measuring the resilience of mutants versus non-mutants during the replacement step [17]. Taking this into account, when a mutant or non-mutant is trying to replace its neighbors, mutants are replaced with probability proportional to $1/\tilde{r}$. Taking $\tilde{r} = 1$ returns to the model used throughout the preceding sections. The two-fitness model may better capture the complexity of real-world evolutionary systems but does not generally give rise to qualitatively different fixation-time distributions. For brevity, we simply discuss the resulting fixation-time distributions for the BD model. Details supporting the results quoted below are provided in the Appendix, Section 2.9.6.

Writing down the transition probabilities for the Moran BD process, we find that $b_m/d_m \to r\tilde{r}$ as $N \to \infty$. This motivates the definition of an effective fitness level, $r_{\text{eff}} = r\tilde{r}$. When $r_{\text{eff}} \neq 1$ our results from above translate to this model. On the 1D lattice the fixation times are normally distributed, while on the complete graph the fixation time distribution is a weighted convolution of Gumbel distributions $G + (r/\tilde{r})G'$, with relative weighting $r/\tilde{r}$ (instead of r). When $r_{\text{eff}} = 1$, the process is asymptotically unbiased and we expect a highly skewed fixation-time distribution. This is indeed the case, although numerical calculations indicate there is an entire family of distributions, dependent on $r = 1/\tilde{r}$.

It is interesting to contrast the above observations with a result in evolu-

tionary dynamics known as the isothermal theorem. The theorem states that for $\tilde{r} = 1$, the Moran process on a large class of networks, known as isothermal graphs, has fixation probability identical to the complete graph [13]. Recent work has shown that this breaks down if $\tilde{r} \neq 1$; the fixation probability develops new network dependence [17]. In contrast, even isothermal graphs (including the complete graph and 1D lattice) have fixation-time distributions that depend on network structure. The two-fitness BD model breaks the universality in fixation probabilities predicted by the isothermal theorem, but leads to the same family of fixation distributions that arise due to network structure.

The Death-Birth Moran process

A two-fitness Death-Birth (DB) Moran process is also frequently used to study evolutionary dynamics. In this model, the birth and death events are reversed in order. At each time step a node is chosen at random, with probability proportional to $1/\tilde{r}$, and one of its neighbors is chosen with probability proportional to r. The first individual dies and is replaced by an offspring of the same type as the neighbor. The process continues until the mutation either reaches fixation or goes extinct.

The BD and DB processes obey a duality property [17]. Starting from the BD transition probabilities, if we swap the two fitness levels $r \leftrightarrow \tilde{r}$ and substitute $m \to N - m$ (which swaps mutants and non mutants), we obtain the DB transition probabilities. Therefore, the transition matrix for the DB model is identical to that for the corresponding dual BD process, but has the main-, super-, and sub-diagonal entries reversed in order. This leaves the matrix eigenvalues unchanged, so that the DB process has identical fixation-time distributions to those

given in the preceding section for the dual BD process.

In principle, the correspondence between DB and BD fixation times could break down for the truncated process considered in Section 2.5. In practice, however, the results are again generally identical. For the truncated DB process, the fixation times on the 1D lattice remain normally distributed. On the complete graph, one of two coupon collection regions is removed by truncation leading to fixation-times following a single Gumbel distribution.

One exception, where the dual models yield different results under truncation, is at infinite fitness. As in Section 2.5, at infinite fitness ($r \to \infty$) the BD model performs a single coupon collection near fixation, which is cut off by truncation, leading to a normal fixation-time distribution. In contrast, in the dual infinite-fitness DB model ($\tilde{r} \to \infty$) the coupon collection occurs at the beginning of the process and even under truncation the Gumbel fixation-time distribution is preserved. This effect was previously observed by Ottino-Löffler et al. [33].

2.6.2 Other networks: Approximate results via mean-field transition probabilities

While the 1D lattice and complete graph provide illustrative exactly solvable models of the fitness dependence of fixation-time distributions, other networks may be more realistic. On more complicated networks the analytical tools developed here fail because the transition probabilities (the probability of adding or subtracting a mutant given the current state) depend on the full configu-

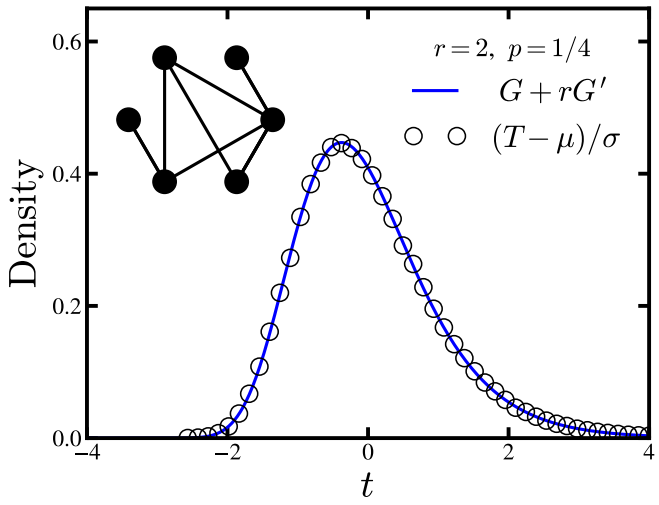

Figure 2.8: Fixation-time distribution on an Erdős-Rényi random graph with $N = 100$ nodes, edge probability $p = 1/4$, and fitness $r = 2$, obtained from 10^6 simulation runs (the same graph is used for each run). The distribution is standardized to zero mean and unit variance. The solid curve is the theoretical prediction for the complete graph, obtained by numerical convolution of two Gumbel distributions, one weighted by r. For these parameters, the random graph fixation time is captured by the mean field approximation.

ration of mutants, not just the number of mutants. Such systems can still be modeled as a Markov process, but the state space becomes prohibitively large. Fortunately, for certain networks the effect of different configurations can be averaged over, giving a mean-field approximation to the transition probabilities. This approach has been used on a variety of networks to calculate fixation times [4, 30, 32, 33]. In this section we discuss how such mean-field approaches can be used to calculate fixation-time distributions for evolution on several different networks.

Erdős-Rényi random graph

We start with the Erdős-Rényi random graph, for which the mean-field transition probabilities were recently estimated [4]. The result is identical to the

complete graph probabilities [Eqs. (2.20)-(2.21)] up to a constant factor $1 - 2/Np$, which depends on the edge probability p for the network. This correction is important for computing the mean fixation time, but does not affect the shape of the fixation-time distribution, since proportionality factors cancel in Eq. (2.4). Therefore we expect the asymptotic fixation-time distribution will be a weighted sum of two Gumbel distributions. This prediction holds for infinite fitness, where the fixation time on an Erdős-Rényi network has a Gumbel distribution [33].

Preliminary simulations show that the Erdős-Rényi network has the expected fixation-time distributions for $p = 1/4$ and $r = 2$ (see Figure 2.8). Further investigation is required to determine the range of fitness and edge probabilities for which this result holds asymptotically (as $N \to \infty$). For constant p, the average degree is proportional to the system size $\langle k \rangle = pN$, similar to the complete graph. It may be, however, that for some p and r the mean-field approximation is not sufficient to capture the higher-order moments determining the shape of the distribution. It is also traditional to consider N-dependent edge probabilities with $p(N)$ chosen, for example, to fix $\langle k \rangle$. It is unclear whether such graphs will behave like the ring (due to their sparsity), like the complete graph (due to their short average path length), or somewhere in between these extreme cases. In the same vein, which other networks admit accurate mean-field approximations to the transition probabilities? Do many complex networks have fixation-time distributions identical to the complete graph?

Stars and superstars: evolutionary amplifiers

Another nice approximation maps the Moran process on a star graph, a simple amplifier of selection, onto a birth-death Markov chain [23]. The resulting transition probabilities exhibit coupon collection regions, similar to the complete graph. The ratio of slopes near these regions (few mutants or non-mutants), however, is r^2. Our heuristic predicts the fixation-time distribution on the star is $G + r^2G'$. In addition to amplifying fixation probability, the star increases fixation-time skew. This raises a broader question: do evolutionary amplifiers also amplify fixation-time skew? Computing fixation times for evolutionary dynamics on superstars (which more strongly amplify selection [13]) remains an open problem.

Growth of cancerous tumors: evolutionary dynamics on d-dimensional lattices

Mean-field arguments have also been applied to d-dimensional lattices in the infinite-fitness limit [32, 33]. In this limit the mutant population grows in an approximately spherical shape near the beginning of the process and the population of non-mutants is approximately spherical near fixation. The surface area to volume ratio of the d-dimensional sphere gives the probability of adding a mutant. With finite fitness, non-mutants can now replace their counterparts and the surface of the sphere of growing mutants roughens [44]. For near-neutral fitness, the configuration of mutants resembles the shape of real cancerous tumors. Perhaps mean-field approaches can draw connections between the fitness-dependent roughness of growing mutant populations and fixation-time distributions for evolution on lattices.

Table 2.1: Asymptotic fixation-time statistics for the Moran Birth-death and Death-birth processes on the complete graph and the 1D lattice. Together with the mean and variance, the standardized distributions give a complete statistical description of the fixation time. The mean and variance given are to leading order in N for each case.

		Asymptotic Fixation-Time Statistics		
Network	Fitness	Mean	Variance	Distribution
1D Lattice	$r = 1$	$\frac{1}{6}N^3$	$\frac{1}{90}N^6$	Highly Skewed [Eqs. (2.11) & (2.12)]
	$r > 1$	$\frac{r+1}{2(r-1)}N^2$	$\frac{(r+1)(r^2+r+1)}{3(r-1)^3}N^3$	$\mathcal{N}(0,1)$
Complete Graph	$r = 1$	N^2	$\left(\frac{\pi^2}{3} - 3\right)N^4$	Highly Skewed [Eqs. (2.23) & (2.24)]
	$r > 1$	$\frac{r+1}{r-1}N \log N$	$\frac{\pi^2(r+1)^2}{6(r-1)^2}N^2$	$G + rG'$

2.7 Summary

In this chapter we have obtained the first closed-form solutions for the fitness dependence of fixation-time distributions of the Moran Birth-death process on the 1D lattice and complete graph. Previous analyses were restricted to the limit of infinite fitness, with some partial results for neutral fitness. To reiterate our new results: There is a dichotomy between neutral and non-neutral fitness. When fitness is neutral, the distribution always exhibits a discontinuity; whether the graph is complete or a 1D lattice, the skew jumps up discontinuously in either case. On the other hand, when fitness is non-neutral but otherwise arbitrary, the results depend strongly on network topology. Specifically, on the complete graph the fixation-time distribution is a fitness-weighted convolution of Gumbel distributions and hence is always skewed, whereas on the 1D lattice the distribution is normal and hence is never skewed.

Together with the mean and variance, the distributions derived here give a complete statistical description of the asymptotic fixation time (see Table 2.1). Our analysis revealed that these results are robust in the sense that similar distributions arise under truncation, in some other models, and in some other network structures, including the Erdős-Rényi random graph.

2.8 Future Directions

Though the model we have focused on here (the Moran Birth-death model) is deliberately simplified, we expect our results will be useful in applications. For instance, the theory should allow a more refined analysis of the rate of evolution, by extending the seminal work by Kimura, whose neutral theory of evolution predicted a molecular clock [47]. In his model, neutral mutations become fixed at a constant rate, independent of population size. This result, with some refinements, is now used widely in estimating evolutionary time scales [48]. The fixation-time distributions discussed here should allow one to go beyond Kimura's classic analysis to capture the full range of evolutionary outcomes, by providing information about the expected deviations from the constant-rate molecular clock, as well as how this prediction is affected by population structure. More generally, it would be interesting to study the implications of these distributions for rates of evolution at various fitness levels.

Furthermore, our results provide concrete predictions that are testable via bacterial evolution experiments. Does the same fitness and network structure dependence of fixation-time distributions arise in real systems?

Future theoretical studies could analyze random networks and lattices more

deeply, as well as stars and superstars, the prototypical evolutionary amplifiers [13]. More sophisticated models involving evolutionary games are also of interest. These have skewed fixation-time distributions [29] whose asymptotic form remains unknown. Finally, we hope that methods developed here will prove useful in other areas, such as epidemiology [7], ecology [6], and protein folding [49], where stochastic dynamics may similarly give rise to skewed first-passage times.

2.9 Appendix

2.9.1 Visit Statistics

In this Appendix we formulate the visit statistics approach. We first provide further details in the derivation of the series expression for the fixation-time cumulants given in Eq. (2.8), and then explicitly compute the weighting factors that appear in this expression to third order. This result requires constant selection, $b_m/d_m = r$, as is the case for the Moran process. Under constant selection the transient transition matrix can be written as $\Omega_{\rm tr} = \Omega_{\rm RW} D$, where D is diagonal with elements $D_{mm} = b_m + d_m$ and $\Omega_{\rm RW}$ is the transition matrix for a random walk,

$$[\Omega_{\rm RW}]_{nm} = \frac{r}{1+r}\delta_{m,n+1} + \frac{1}{1+r}\delta_{m,n-1} - \delta_{m,n}. \tag{2.29}$$

Since we are interested in the fixation-time distribution, we condition on fixation occurring. As discussed in Section 2.2.2 (see also Section 2.9.2 below), the conditioned transition matrix $\tilde{\Omega}_{\rm tr} = S\,\Omega_{\rm tr}\,S^{-1}$, where S is diagonal with $S_{mm} = 1 - 1/r^m$.

Combining these results, we have that

$$\tilde{\Omega}_{\text{tr}} = S\ \Omega_{\text{RW}}\, S^{-1} D, \tag{2.30}$$

where we have used the fact that both D and S are diagonal matrices, and therefore commute.

We found in Section 2.2.2 that the moments of the fixation time T can be expressed as,

$$\mu_n \coloneqq E[T^n] = (-1)^n n!\, \mathbf{1}\, \tilde{\Omega}_{\text{tr}}^{-n}\, \mathbf{p}_{\text{tr}}(0), \tag{2.31}$$

where $\mathbf{1}$ is a row vector of ones and $\mathbf{p}_{\text{tr}}(0)$ is the initial state of the system, with $[p_{\text{tr}}(0)]_m = \delta_{m,m_0}$ for m_0 initial mutants. To compute these moments, we need the inverse $\tilde{\Omega}_{\text{tr}}^{-1} = D^{-1} S\ \Omega_{RW}^{-1} S^{-1}$. Since Ω_{RW} is a tridiagonal Toeplitz matrix, its inverse has a well-known form [50]:

$$(-\Omega_{RW})^{-1}_{ij} = \begin{cases} \dfrac{(r+1)(r^i-1)(r^N-r^j)}{r^j\,(r-1)(r^N-1)} & \text{if } i \le j, \\ \dfrac{(r+1)(r^j-1)(r^N-r^i)}{r^j\,(r-1)(r^N-1)} & \text{if } i > j. \end{cases} \tag{2.32}$$

Hence the matrix $V = -S\ \Omega_{\text{RW}}^{-1} S^{-1}$ has elements

$$V_{ij} = \begin{cases} \dfrac{(r+1)(r^i-1)^2(r^N-r^j)}{r^i\,(r-1)(r^j-1)(r^N-1)} & \text{if } i \le j, \\ \dfrac{(r+1)(r^i-1)(r^N-r^i)}{r^i\,(r-1)(r^N-1)} & \text{if } i > j. \end{cases} \tag{2.33}$$

The matrix V, sometimes called the fundamental matrix, encodes the visit statistics of the conditioned random walk: V_{ij} is the mean number of visits to state i from state j before hitting the absorbing state N [51]. The Moran process has the same visit statistics, but on average spends a different amount of time, designated by $(b_i + d_i)^{-1}$, waiting in each state.

While one could now compute the moments μ_n in Eq. (2.31) directly, we find that the cumulants yield nicer expressions. Furthermore, the normal and Gum-

bel fixation-time distributions, predicted by our simulations and approximate calculations, are more simply described in terms of their cumulants. The non-standardized cumulants κ'_n are linear combinations involving products of moments whose orders sum to n. Thus each term in the cumulants has n powers of D producing n factors of $(b_i+d_i)^{-1}$ with a weight designated by the visit statistics. With this observation, it is clear the standardized cumulants $\kappa_n = \kappa'_n/(\kappa'_2)^{n/2}$ have the form given in Eq. (2.8),

$$\kappa_n(N) = \frac{\displaystyle\sum_{i_1,i_2,\ldots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r, N|m_0)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})}}{\left(\displaystyle\sum_{i,j=1}^{N-1} \frac{w^2_{ij}(r, N|m_0)}{(b_i+d_i)(b_j+d_j)}\right)^{n/2}}, \tag{2.34}$$

where $w^n_{i_1 i_2 \cdots i_n}(r, N|m_0)$ are the weighting factors coming entirely from the visit statistics of a biased random walk (starting from m_0 initial mutants). As in the sections above, we take the initial state to be a single mutant $m_0 = 1$, and will suppress the dependence of the weighting factors on initial condition, writing $w^n_{i_1 i_2 \cdots i_n}(r, N)$ instead. Generalizations to other cases are straightforward and are discussed briefly below.

We emphasize that even without explicit knowledge of the factors $w^n_{i_1 i_2 \cdots i_n}(r, N)$, this formulation can be extremely useful. For instance when $b_i + d_i$ is constant, these are just the cumulants for the (possibly biased) random walk, which were computed in Section 2.3 to approximate the Moran process on the 1D lattice. In particular, the sums over weighting factors obtained from setting $b_i + d_i = 1$ in Eq. (2.34) have leading asymptotic form given by Eq. (2.14). This fact can be used to bound the cumulants even when $b_i + d_i \neq 1$, which in some cases is sufficient to determine the leading asymptotic behavior. When this is not possible, the weighting factors must be computed explicitly. We now turn our focus to derviing $w^2_{ij}(r, N)$ and $w^3_{ijk}(r, N)$.

We can compute the weighting factors by writing out the matrix multiplication of $\tilde{\Omega}_{\text{tr}}^{-1}$. First note that

$$[-\tilde{\Omega}_{\text{tr}}^{-1}]_{ij} = \frac{V_{ij}}{b_i + d_i}. \tag{2.35}$$

Then the first three moments of the fixation time are,

$$\begin{aligned}
\mu_1 &= \sum_{i=1}^{N} \frac{V_{i1}}{b_i + d_i}, \\
\mu_2 &= 2 \sum_{i,j=1}^{N} \frac{V_{ij} V_{j1}}{(b_i + d_i)(b_j + d_j)}, \\
\mu_3 &= 6 \sum_{i,j,k=1}^{N} \frac{V_{ij} V_{jk} V_{k1}}{(b_i + d_i)(b_j + d_j)(b_k + d_k)}.
\end{aligned} \tag{2.36}$$

The corresponding non-standardized cumulants are given by the usual formulas, $\kappa_2' = \mu_2 - \mu_1^2$ and $\kappa_3' = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3$. In terms of the visit numbers the non-standardized cumulants become

$$\begin{aligned}
\kappa_2' &= \sum_{i,j=1}^{N} \frac{2\, V_{ij} V_{j1} - V_{i1} V_{j1}}{(b_i + d_i)(b_j + d_j)}, \\
\kappa_3' &= \sum_{i,j=1}^{N} \frac{6\, V_{ij} V_{jk} V_{k1} - 6\, V_{ij} V_{j1} V_{k1} + 2\, V_{i1} V_{j1} V_{k1}}{(b_i + d_i)(b_j + d_j)(b_k + d_k)}.
\end{aligned} \tag{2.37}$$

From here we can read off the weighting factors accordingly. For convenience, we can choose $w_{ij}^2(r, N)$ and $w_{ijk}^3(r, N)$ to be symmetric by averaging the numerators in Eq. (2.37) over the permutations of the indices. Then,

$$\begin{aligned}
w_{ij}^2(r, N) &= \frac{1}{2} \sum_{\sigma \in \Pi_2} 2\, V_{\sigma(1)\sigma(2)} V_{\sigma(2)1} - V_{\sigma(1)1} V_{\sigma(2)1}, \\
w_{ijk}^3(r, N) &= \frac{1}{6} \sum_{\sigma \in \Pi_3} 6\, V_{\sigma(1)\sigma(2)} V_{\sigma(2)\sigma(3)} V_{\sigma(3)1} - 6\, V_{\sigma(1)\sigma(2)} V_{\sigma(2)1} V_{\sigma(3)1} + 2\, V_{\sigma(1)1} V_{\sigma(2)1} V_{\sigma(3)1},
\end{aligned} \tag{2.38}$$

where Π_2 is the set of permutations of $\{i, j\}$ and Π_3 are the permutations of $\{i, j, k\}$. We note that these expressions also hold for general initial condition by replacing the subscript 1 with m_0. Plugging Eq. (2.33) into this expression for w_{ij}^2 we

obtain, after some algebra,

$$w_{ij}^2(r,N) = \frac{(r+1)^2(r^j-1)^2(r^N-r^i)^2}{r^{i+j}\,(r-1)^2(r^N-1)^2}, \tag{2.39}$$

for $i \geq j$. Since we have constructed $w_{ij}^2(r,N)$ to be symmetric, when $j > i$ the formula is identical with i and j exchanged. Similarly, using Eq. (2.33) together with the expression for w_{ijk}^3 in Eq. (2.38) leads to

$$w_{ijk}^3(r,N) = 2\frac{(r+1)^3(r^k-1)^2(r^j-1)(r^N-r^i)^2(r^N-r^j)}{r^{i+j+k}\,(r-1)^3(r^N-1)^3}, \tag{2.40}$$

for $i \geq j \geq k$. Again, the formula for different orderings of the indices i, j, k is the same with the indices permuted appropriately, so that w_{ijk}^3 is perfectly symmetric.

This completes the derivation of the visit statistics expression for the fixation-time cumulants. Together, Eqs. (2.34), (2.39) and (2.40) give a closed form expression for the fixation-time skew which is manageable for the purpose of asymptotic approximations. The diagonal terms in the higher-order weighting factors are also particularly simple, $w_{ii\cdots i}^n(r,N) = (n-1)!V_{ii}^n$. While we will not explicitly compute them, the off diagonal weights $w_{i_1 i_2 \cdots i_n}^n(r,N)$ can be found by a straightforward generalization of the above procedure. Example applications of this approach are given in Sections 2.9.4 and 2.9.5, where we show that all cumulants of the fixation time vanish for the Moran process on the 1D lattice and compute the asymptotic skew for the Moran process on the complete graph.

2.9.2 Birth-death Markov chain conditioned on fixation

For both the numerical recurrence relation and the visit statistics approach, it is useful to consider the birth-death Markov chain conditioned on hitting N, which

has an identical fixation-time distribution to the unconditioned process. This Markov chain has new conditioned transition probabilities denoted $\tilde{b}_m$ and $\tilde{d}_m$. If X_t is the state of the system at time t, then $\tilde{b}_m \coloneqq \mathcal{P}(X_t = m \to X_{t+1} = m+1 | X_\infty = N)$ and $\tilde{d}_m$ is defined similarly. Applying the laws of conditional probability, we find that

$$\begin{aligned}
\tilde{b}_m &= \frac{\mathcal{P}(X_{t+1} = m+1 \text{ AND } X_t = m \text{ AND } X_\infty = N)}{\mathcal{P}(X_t = m \text{ AND } X_\infty = N)} \\
&= \frac{\mathcal{P}(X_\infty = N | X_t = m+1)}{\mathcal{P}(X_\infty = N | X_t = m)} \mathcal{P}(X_{t+1} = m+1 | X_t = m) \\
&= \frac{\mathcal{P}(X_\infty = N | X_t = m+1)}{\mathcal{P}(X_\infty = N | X_t = m)} b_m,
\end{aligned} \tag{2.41}$$

where b_m is the transition rate in the original Markov chain. Following the same procedure, we find the backward transition probabilities are related by

$$\tilde{d}_m = \frac{\mathcal{P}(X_\infty = N | X_t = m-1)}{\mathcal{P}(X_\infty = N | X_t = m)} d_m. \tag{2.42}$$

The conditioned Markov chain has a few nice properties. First, the fixation probability in the conditioned system is one, by construction. This is particularly helpful for accelerating simulations of the Moran process. Conditioning the transition probabilities also accounts for the normalization of the fixation-time distribution. Furthermore, this operation only changes the relative probability of adding versus subtracting a mutant. The probability that the system leaves a given state is unchanged:

$$\begin{aligned}
\tilde{b}_m + \tilde{d}_m &= \frac{\mathcal{P}(X_{t+1} = m+1 \text{ AND } X_t = m \text{ AND } X_\infty = N)}{\mathcal{P}(X_t = m \text{ AND } X_\infty = N)} \\
&\quad + \frac{\mathcal{P}(X_{t+1} = m-1 \text{ AND } X_t = m \text{ AND } X_\infty = N)}{\mathcal{P}(X_t = m \text{ AND } X_\infty = N)} \\
&= 1 - \frac{\mathcal{P}(X_{t+1} = m \text{ AND } X_t = m \text{ AND } X_\infty = N)}{\mathcal{P}(X_t = m \text{ AND } X_\infty = N)} \\
&= 1 - \mathcal{P}(X_{t+1} = m | X_t = m) \\
&= b_m + d_m.
\end{aligned} \tag{2.43}$$

This invariance, along with Eqs. (2.41) and (2.42), shows that conditioning the Markov chain is equivalent to a similarity transformation on the transient transition matrix with a diagonal change of basis:

$$\tilde{\Omega}_{\text{tr}} = S\,\Omega_{\text{tr}}\,S^{-1} \qquad S_{mn} = \mathcal{P}(X_\infty = N|X_t = m)\delta_{m,n}, \tag{2.44}$$

where Ω_{tr} is the birth-death transition matrix with absorbing states removed as defined in Section 2.2.

For the Moran Birth-death process considered in this chapter, $b_m/d_m = r$. In this case, by setting up a linear recurrence it is easy to show that the probability of fixation, starting from m mutants, is

$$\mathcal{P}(X_\infty = N|X_t = m) = \frac{1 - 1/r^m}{1 - 1/r^N}, \tag{2.45}$$

so that

$$\tilde{b}_m = \frac{r^{m+1} - 1}{r^{m+1} - r} b_m, \qquad \tilde{d}_m = \frac{r^m - r}{r^m - 1} d_m. \tag{2.46}$$

Note that we can scale the similarity matrix S by an overall constant, so it is convenient to choose $S_{mn} = (1 - 1/r^m)\delta_{m,n}$. For the two-fitness Moran Birth-Death model discussed in Section 2.6.1 fixation probabilities derived by Kaveh et al. [17] can be used together with Eq. (2.44) to condition the Markov chain on fixation.

2.9.3 Recurrence relation for fixation-time skew

With the conditioned transition probabilities derived in Section 2.9.2, there is a reflecting boundary at $m = 1$, which lets us set up a recurrence relation for the fixation-time moments. This derivation follows the method described by

Keilson in Ref. [41]. Let $S_m(t)$ be the first-passage time densities from state m to state $m+1$. Clearly, $S_1(t)$ has an exponential distribution,

$$S_1(t) = \tilde{b}_1 e^{-\tilde{b}_1 t}. \tag{2.47}$$

From $m > 1$, the state $m+1$ can be reached either directly, with exponentially distributed times, or indirectly by first stepping backwards to $m-1$, returning to m, and then reaching $m+1$ at a latter time. Thus, the densities $S_m(t)$ satisfy

$$S_m(t) = \tilde{b}_m e^{-(\tilde{b}_m+\tilde{d}_m)t} + \tilde{d}_m e^{-(\tilde{b}_m+\tilde{d}_m)t} * S_{m-1}(t) * S_m(t), \tag{2.48}$$

where the symbol $*$ denotes a convolution. This equation can be solved by Fourier transform to obtain

$$S_m(\omega) = \frac{\tilde{b}_m}{\tilde{b}_m + \tilde{d}_m - \tilde{d}_m S_{m-1}(\omega) - i\omega}. \tag{2.49}$$

We can compute a recurrence relation for the moments of the first-passage time densities $S_m(t)$ by differentiating Eq. (2.49). Let μ_m, ν_m and γ_m to be the first, second, and third moments of $S_m(t)$ respectively. Using the relations $\nu_m = -iS'(\omega=0)$, $\xi_m = (-i)^2 S''(\omega=0)$, and $\zeta_m = (-i)^3 S'''(\omega=0)$, we find that

$$\begin{aligned}
\nu_m &= \tilde{b}_m^{-1}(1 + \tilde{d}_m \nu_{m-1}), \\
\xi_m &= \tilde{b}_m^{-2}[\tilde{b}_m \tilde{d}_m \xi_{m-1} + 2(1 + \tilde{d}_m \nu_{m-1})^2], \\
\zeta_m &= \tilde{b}_m^{-3}[\tilde{b}_m^2 \tilde{d}_m \zeta_{m-1} + 6\tilde{b}_m \tilde{d}_m \xi_{m-1}(1 + \tilde{d}_m \nu_{m-1}) + 6(1 + \tilde{d}_m \nu_{m-1})^3],
\end{aligned} \tag{2.50}$$

with boundary conditions $\nu_0 = \xi_0 = \zeta_0 = 0$. The recurrence relations in Eq. (2.50) give the moments of incremental first-passage time distributions $S_m(t)$. The total fixation time, T is the sum of these incremental first-passage times. Thus, the cumulants of T are the sum of the cumulants of the incremental times and the skew of T can be expressed as,

$$\kappa_3(N) = \left(\sum_{m=1}^{N-1} \zeta_m - 3\xi_m \nu_m + 2\nu_m^3\right) \Bigg/ \left(\sum_{m=1}^{N-1} \xi_m - \nu_m^2\right)^{3/2}. \tag{2.51}$$

Numerical computation of for $\kappa_3(N)$ requires calculating the $3N$ moments and carrying out the two sums in Eq. (2.51). By bottom-up tabulation of the incremental moments, this procedure can be completed in $O(N)$ time, asymptotically faster than the eigenvalue decomposition and the exact series solution from visit statistics.

2.9.4 Asymptotic Analysis for the 1D Lattice

Neutral Fitness

For our detailed asymptotic analysis, we will begin with the neutral fitness Moran process on a 1D lattice with periodic boundary conditions. In this case, the eigenvalues of the transition matrix describing the system are,

$$\lambda_m = \frac{2}{N} - \frac{2}{N}\cos\left(\frac{m\pi}{N}\right), \quad m = 1, 2, \ldots, N-1. \tag{2.52}$$

From the eigen-decomposition of the Markov birth-death process described Section 2.2.1, the standardized fixation-time cumulants are given by

$$\kappa_n(N) = (n-1)!\left(\sum_{m=1}^{N-1}\frac{1}{\lambda_m^n}\right)\Bigg/\left(\sum_{m=1}^{N-1}\frac{1}{\lambda_m^2}\right)^{n/2}. \tag{2.53}$$

Note that the constant factor $2/N$ cancels in Eq. (2.53), so we may equivalently consider rescaled eigenvalues $\lambda_m = 1 - \cos(m\pi/N)$. To derive the asymptotic cumulants, we compute the leading asymptotic behavior of sums

$$S_n = \sum_{m=1}^{N-1}\frac{1}{[1-\cos(m\pi/N)]^n}. \tag{2.54}$$

The function $(1-\cos x)^{-n}$ can be expanded as a Laurent series $\sum_{k=0}^{\infty} c_k(n)x^{2(k-n)}$, which is absolutely convergent for $x \neq 0$ in the interval $(-2\pi, 2\pi)$. So the sum S_n

can then be expressed as

$$\begin{aligned} S_n &= \sum_{m=1}^{N-1} \sum_{k=0}^{\infty} c_k(n) \left(\frac{\pi m}{N}\right)^{2(k-n)} \\ &= \sum_{k=0}^{\infty} c_k(n) (N/\pi)^{2(n-k)} H_{N-1,2(n-k)} \\ &= \frac{c_0(n)\zeta(2n)}{\pi^{2n}} N^{2n} + O(N^{2(n-1)}) \end{aligned} \tag{2.55}$$

where $H_{N,q} = \sum_{m=1}^{N} m^{-q}$ is the generalized harmonic number and in the last line we used the asymptotic approximation

$$H_{N,2q} = \begin{cases} \zeta(2q) + O(N^{1-2q}) & q > 0, \\ \dfrac{N^{1-2q}}{2q+1} + O(N^{-2q}) & q \leq 0. \end{cases} \tag{2.56}$$

It is easy to check that $c_0(n) = 2^n$. Now the cumulants are $\kappa_n(N) = (n-1)! S_n / S_2^{n/2}$, which for $N \to \infty$ are

$$\begin{aligned} \kappa_n &= (n-1)! \left(\frac{2^n \zeta(2n)}{\pi^{2n}}\right) \Big/ \left(\frac{2^2 \zeta(4)}{\pi^4}\right)^{n/2} \\ &= (n-1)! \, \frac{\zeta(2n)}{\zeta(4)^{n/2}}, \end{aligned} \tag{2.57}$$

as reported in Section 2.3.1.

Non-neutral fitness

For non-neutral fitness, we showed in Section 2.3.2 that in the random walk approximation the fixation-time distribution is asymptotically normal. Here we use the visit statistics approach to prove this holds even when the variation in time spent in each state is accounted for. From the visit statistics formulation, the standardized cumulants of the fixation time (starting from a single initial

mutant) can be written as,

$$\kappa_n(N) = \left(\sum_{i_1,i_2,\ldots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})} \right) \Bigg/ \left(\sum_{i,j=1}^{N-1} \frac{w^2_{ij}(r,N)}{(b_i+d_i)(b_j+d_j)} \right)^{n/2}, \tag{2.58}$$

where $w^n_{i_1 i_2 \cdots i_n}(r,N)$ are the weighting factors that depend on the visit statistics of a biased random walk. To prove the fixation-time distribution is normal, we derive bounds on the sums

$$S_n = \sum_{i_1,i_2,\ldots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})} \tag{2.59}$$

that appear in Eq. (2.58) and show that $\kappa_n(N) \to 0$ as $N \to \infty$. First, note that $w^n_{i_1 i_2 \cdots i_n}(r,N) = (n-1)!\, V^n_{ii} \geq 1$ if $i_1 = i_2 = \cdots = i_n \equiv i$. Furthermore, we claim that $w^n_{i_1 i_2 \cdots i_n}(r,N) \geq 0$ for all $i_1, i_2, \ldots, i_n$. If this were not the case, one could construct a birth-death process with negative fixation-time cumulants by choosing $b_i + d_i$ appropriately. But we know the fixation-time cumulants are positive from the eigen-decomposition, Eq (2.53). With these observations, we can bound S_n from below by the sum over unweighted diagonal elements. Similarly, the sums are bounded from above by the maximum value of $(b_i+d_i)^{-n}$ times the sum over the weighting factors. Putting these together, we obtain

$$\sum_{i=1}^{N-1} \frac{1}{(b_i+d_i)^n} \leq S_n \leq \left(\max_{1 \leq i < N} \frac{1}{b_i+d_i} \right)^n \times \sum_{i_1,i_2,\ldots,i_n=1}^{N-1} w^n_{i_1 i_2 \cdots i_n}(r,N). \tag{2.60}$$

The Moran process on the 1D lattice has transition probabilities $b_i + d_i = (1+r)/(rm+N-m)$. Then, as $N \to \infty$, the lower bound is

$$\sum_{i=1}^{N-1} \frac{1}{(b_i+d_i)^n} = \frac{1}{(r+1)^n} \sum_{m=1}^{N-1} (rm+N-m)^n = \frac{1+r+r^2+\cdots+r^n}{(n+1)(1+r)^n} N^{n+1} + O(N^n). \tag{2.61}$$

For the upper bound, first note that

$$\left(\max_{1 \leq i < N} \frac{1}{b_i+d_i} \right)^n = [r(N-1)+1]^n = r^n N^n + O(N^{n-1}). \tag{2.62}$$

The sums over the weighting factors give the (non-standardized) fixation-time cumulants corresponding to a process with $b_i + d_i = 1$ and uniform bias r. This is exactly the biased random walk model used to approximate the Moran process in Section 2.3.2. It follows that as $N \to \infty$,

$$\sum_{i_1,i_2,\ldots,i_n=1}^{N-1} w^n_{i_1 i_2 \cdots i_n}(r,N) = (n-1)! \sum_{i=1}^{N-1} \left(\frac{1}{1 - 2\sqrt{r}/(r+1)\cos(m\pi/N)} \right)^n, \tag{2.63}$$

where the denominators in the second sum are the eigenvalues of the transition matrix for the biased random walk, $\lambda_m = 1 - 2\sqrt{r}/(r+1)\cos(m\pi/N)$. As in Section 2.3.2, we can estimate the leading asymptotics of this sum by converting to an integral,

$$\sum_{i_1,i_2,\ldots,i_n=1}^{N-1} w^n_{i_1 i_2 \cdots i_n}(r,N) = \frac{N}{\pi} \int_0^\pi \frac{(n-1)!}{(1 - 2\sqrt{r}/(1+r)\cos x)^n} \, \mathrm{d}x + O(1). \tag{2.64}$$

Combining the results from Eqs. (2.60)–(2.62) and (2.64) we arrive at

$$\frac{1 + r + r^2 + \cdots + r^n}{(n+1)(1+r)^n} N^{n+1} + O(N^n) \leq S_n \leq \frac{N^{n+1}}{\pi} \int_0^\pi \frac{(n-1)!}{(1 - 2\sqrt{r}/(1+r)\cos x)^n} \, \mathrm{d}x + O(N^n). \tag{2.65}$$

For each n, our upper and lower bounds have the same asymptotic scaling as a power of N, with different r-dependent coefficients. Using these results together in Eq. (2.58), it follows that for $N \gg 1$, the cumulants to leading order are

$$\kappa_n(N) = C_n(r) \frac{1}{N^{(n-2)/2}} + O(N^{-n/2}), \tag{2.66}$$

where $C_n(r)$ is a fitness-dependent constant. Thus, indeed $\kappa_n(N) \to 0$ as $N \to \infty$.

This result confirms the claim made in Section 2.3.2. Even with heterogeneity in the time spent in each state, the skew and higher-order cumulants of the fixation time vanish asymptotically. Therefore, the Moran Birth-death process on the 1D lattice with non-neutral fitness $r > 1$ has an asymptotically normal fixation-time distribution. The normal distribution is universal, independent of fitness level for this population structure.

2.9.5 Asymptotic Analysis for the Complete Graph

Non-neutral fitness

In Section 2.4.2 we predicted that the asymptotic fixation-time distribution for the Moran Birth-death process on the complete graph is a convolution of two Gumbel distributions by applying our intuition from coupon collection. Furthermore, our calculation of the fixation-time cumulants in the large (but finite) fitness limit agrees with this prediction. Surprisingly, numerical calculations using the recurrence relation formulated above and direct simulations of the Moran process indicate that this result holds for all $r > 1$. In this section we prove, using the visit statistics formulation, that the asymptotic skew of the fixation time for $r > 1$ is identical to that of a convolution of Gumbel distributions. Based on our numerical evidence, we conjecture that an analogous calculation holds to all orders. The below calculation also shows why the coupon collection heuristic works: the asymptotically dominant terms come exclusively from the regions near fixation ($m = N - 1$) and near the beginning of the process when a single mutant is introduced into the system ($m = 1$).

As for the 1D lattice, we want to derive the asymptotic behavior of the sums

$$S_n = \sum_{i_1, i_2, \ldots, i_n = 1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r, N)}{(b_{i_1} + d_{i_1})(b_{i_2} + d_{i_2}) \cdots (b_{i_n} + d_{i_n})}, \tag{2.67}$$

where the transition probabilities b_i and d_i are those for the Moran process on the complete graph,

$$b_i + d_i = \frac{(1 + r) i (N - i)}{(N - 1)(ri + N - i)} \tag{2.68}$$

and the weights $w^2_{ij}(r, N)$ and $w^3_{ijk}(r, N)$ are respectively given by

$$w^2_{ij}(r, N) = \frac{(r + 1)^2 (r^j - 1)^2 (r^N - r^i)^2}{r^{i+j} (r - 1)^2 (r^N - 1)^2} \tag{2.69}$$

for $i > j$ and

$$w^3_{ijk}(r,N) = 2\frac{(r+1)^3(r^k-1)^2(r^j-1)(r^N-r^i)^2(r^N-r^j)}{r^{i+j+k}\,(r-1)^3(r^N-1)^3}, \tag{2.70}$$

for $i > j > k$. The expressions for different orderings of indices are the same but with the indices permute appropriately so that w^2_{ij} and w^3_{ijk} are perfectly symmetric.

To start, consider the sums Eq. (2.67), but with two indices i_1 and i_2 constrained to integers from αN to $(1-\alpha)N$ for $1/2 > \alpha > 0$. This sum may be written as

$$S^{\alpha}_n = \sum_{i_1,i_2=\alpha N}^{(1-\alpha)N} \sum_{i_3,\cdots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2\cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})}. \tag{2.71}$$

Now we may apply the upper bound in Eq. (2.60), but for the sums restricted to $\alpha N < i, j < (1-\alpha)N$, the maximum of $(b_i+d_i)^{-1}$ can also be restricted to this range,

$$\begin{aligned} S^{\alpha}_n &\le \left(\max_{1<i<N}\frac{1}{b_i+d_i}\right)^{n-2} \times \left(\max_{\alpha N<i<(1-\alpha)N}\frac{1}{b_i+d_i}\right)^2 \times \sum_{i_1,i_2,\cdots,i_n=1}^{N} w^n_{i_1 i_2\cdots i_n}(r,N) \\ &= N^{n-1}\left\{\left(\frac{r}{1+r}\right)^{n-2}\left(\frac{r(1-\alpha)+\alpha}{(1+r)(1-\alpha)\alpha}\right)^2 \times \frac{1}{\pi}\int_0^{\pi}\frac{(n-1)!}{(1-2\sqrt{r}/(1+r)\cos x)^n}\,\mathrm{d}x\right\} + O(N^{n-2}). \end{aligned} \tag{2.72}$$

In the second line we used the integral approximation from Eq. (2.64) and evaluated the maximum of $(b_i+d_i)^{-1}$ over the indicated intervals. Since we constructed $w^n_{i_1 i_2\cdots i_n}(r,N)$ to be symmetric, this upper bound holds for any permutation of the indices in Eq. (2.71).

We now consider the same sums but with $1 < i_1 < \alpha N$ or $(1-\alpha)N < i_1 < N-1$,

$$S^{\alpha,1}_n = \sum_{i_1=1}^{\alpha N} \sum_{i_2=\alpha N}^{(1-\alpha)N} \sum_{i_3,\cdots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2\cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})}. \tag{2.73}$$

and

$$S^{\alpha,2}_n = \sum_{i_1=(1-\alpha)N}^{N-1} \sum_{i_2=\alpha N}^{(1-\alpha)N} \sum_{i_3,\cdots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2\cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})}. \tag{2.74}$$

These sums can be estimated using the same upper bound, but without extending the sum on $w^n_{i_1 i_2 \cdots i_n}(r, N)$ to the entire domain. Specifically,

$$S_n^{\alpha,1} \leq N^{n-1} \left\{ \left(\frac{r}{1+r} \right)^{n-1} \left(\frac{r(1-\alpha)+\alpha}{(1+r)(1-\alpha)\alpha} \right) \right\} \times \sum_{i_1=1}^{\alpha N} \sum_{i_2=\alpha N}^{(1-\alpha)N} \sum_{i_3,\cdots,i_n=1}^{N-1} w^n_{i_1 i_2 \cdots i_n}(r, N) + O(N^{n-2}). \tag{2.75}$$

Note that the weighting factors fall off exponentially away from the diagonal elements. This is because the visit numbers in the biased random walk become only very weakly correlated if the states are far away from each other. Thus, the sum in Eq. (2.75) over terms away from the diagonal elements converges to a constant as $N \to \infty$. We have verified this explicitly for $w^2_{ij}(r, N)$ and $w^3_{ijk}(r, n)$. The series $S_n^{\alpha,2}$ is similarly bounded, as are all sums of the form Eq. (2.73) or (2.74) with the indices permuted.

The remaining terms in S_n involve all indices in either $[1, \alpha N]$ or $[(1-\alpha)N, N-1]$. If not all indices are in the same interval, the weighting factors are exponentially small: the visit numbers near $m = 1$ are uncorrelated with those near $m = N - 1$. Thus each term in the sum is exponentially suppressed and doesn't contribute to S_n asymptotically. With this observation only two parts of the sum remain: those with bounds $1 \leq i_1, i_2 \ldots i_n \leq \alpha N$ or $(1-\alpha)N \leq i_1, i_2 \ldots i_n \leq N-1$. We call the sums with these bounds S_n^{c1} and S_n^{c2} respectively. As we will see below, the sums over these regions have leading order $O(N^n)$. Since all the above terms are order $O(N^{n-1})$ or smaller, the asymptotic behavior of the cumulants is entirely determined by these regions near the beginning and end of the process, i.e. the coupon collection regions. The fact that we can restrict the sums to this region allows us to make approximations that do not change the leading asymptotics, but make the sums easier to carry out. For instance, in S_2^{c1}, we can

set $r^N - r^i \to r^N$ and $(N-i) \to N$, since the indices run only up to αN. This gives

$$\begin{aligned} S_2^{c1} &= \frac{N^2}{(r-1)^2}\left\{\sum_{i=1}^{\alpha N}\frac{(r^i-1)^2}{i^2 r^{2i}} + 2\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1}\frac{(r^j-1)^2}{ij\,r^{i+j}}\right\} + O(N) \\ &= \frac{N^2\zeta(2)}{(r-1)^2} + O(N), \end{aligned} \tag{2.76}$$

for $N \gg 1$. A similar calculation shows $S_2^{c2} = r^2N^2\zeta(2)/(r-1)^2$. For the third order sums, we find

$$\begin{aligned} S_3^{c1} = 2\frac{N^3}{(r-1)^3}\Bigg\{ &\sum_{i=1}^{\alpha N}\frac{(r^i-1)^3}{i^3 r^{3i}} + 3\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1}\frac{(r^j-1)^3}{ij^2 r^{i+2j}} \\ &+ 3\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1}\frac{(r^i-1)(r^j-1)^2}{i^2 j r^{2i+j}} + 6\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}\frac{(r^j-1)(r^k-1)^2}{ijk\,r^{i+j+k}}\Bigg\} + O(N^2) \\ = 2\frac{N^3\zeta(3)}{(r-1)^3} + O(N^2), \end{aligned} \tag{2.77}$$

for $N \gg 1$. Again the other sum, with indices near $N-1$, is identical up to a factor of r^3, $S_3^{c2} = 2r^3N^3\zeta(3)/(r-1)^3$. Overall, we have that

$$S_2 = \frac{N^2(1+r^2)\zeta(2)}{(r-1)^2} + O(N) \qquad \text{and} \qquad S_3 = \frac{2N^3(1+r^3)\zeta(3)}{(r-1)^3} + O(N^2). \tag{2.78}$$

The asymptotic skew is given by

$$\kappa_3 = \frac{2(1+r^3)\zeta(3)}{(r-1)^3}\Bigg/\left(\frac{(1+r^2)\zeta(2)}{(r-1)^2}\right)^{3/2} = \frac{1+r^3}{(1+r^2)^{3/2}} \times \frac{2\zeta(3)}{\zeta(2)^{3/2}}, \tag{2.79}$$

which is exactly the skew corresponding to the convolution of Gumbel distributions with relative weighting given by the fitness, $G + rG'$. While evaluating the series to higher orders is increasingly difficult, our simulations and the large-fitness approximation suggest this result holds to all orders and that indeed, the asymptotic fixation-time distribution is a weighted convolution of Gumbel distributions.

Neutral fitness with truncation

As discussed in Section 2.5.2, the neutral fitness Moran process on the complete graph has a fixation-time skew that depends on the level of truncation. That is, the time T_α it takes for the process to reach αN mutants, where $0 \leq \alpha \leq 1$, has a distribution whose skew depends on α. Here we show that the $\alpha \to 0$ limit of the fixation-time skew equals $\sqrt{3}$.

To start, we take the neutral fitness limit of the weighting factors to obtain

$$w_{ij}^2(1, \alpha N) = \frac{4j^2(\alpha N - i)^2}{\alpha^2 N^2} \tag{2.80}$$

for $i \geq j$ and

$$w_{ijk}^3(1, \alpha N) = \frac{16jk^2(\alpha N - i)^2(\alpha N - j)}{\alpha^3 N^3} \tag{2.81}$$

for $i \geq j \geq k$, again with the expressions for other orderings obtained by permuting the indices accordingly. The neutral fitness Moran process on the complete graph has transition probabilities $b_i + d_i = 2(Ni - i^2)/(N^2 - N)$. Since we are computing the truncated fixation-time skew, we use Eq. (2.58), but cut the sums off at αN. In this case, these sums are dominated by the off-diagonal terms, so that

$$\begin{aligned} S_2 = \sum_{i,j=1}^{\alpha N} \frac{w_{ij}^2(1, \alpha N)}{(b_i + d_i)(b_j + d_j)} &= 2 \sum_{i=1}^{\alpha N} \sum_{j=1}^{i-1} \frac{j(\alpha N - i)^2(N-1)^2}{\alpha^2 i(N-i)(N-j)} + O(N^3) \\ &= 2 \sum_{i=1}^{\alpha N} \sum_{j=1}^{i-1} \frac{j(\alpha N - i)^2}{\alpha^2 i} + O(N^3) \\ &= \frac{\alpha^2 N^4}{12} + O(N^3), \end{aligned} \tag{2.82}$$

where in the second line we approximated $N - i$ and $N - j$ by N. This approximation is exact in the limit $\alpha \to 0$ since the upper limit on the sum, αN, is much

smaller than N. Using analogous approximations, we find

$$\begin{aligned} S_3 &= \sum_{i,j,k=1}^{\alpha N} \frac{w_{ijk}^2(1,\alpha N)}{(b_i+d_i)(b_j+d_j)(b_k+d_k)} \\ &= 6\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} \frac{2k(\alpha N-i)^2(\alpha N-j)(N-1)^3}{\alpha^3 i(N-i)(N-j)(N-k)} + O(N^5) \\ &= 12\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} \frac{k(\alpha N-i)^2(\alpha N-j)}{\alpha^2 i} + O(N^5) \\ &= \frac{\alpha^3 N^6}{24} + O(N^5). \end{aligned} \tag{2.83}$$

The asymptotic fixation-time skew as $\alpha \to 0$ is therefore

$$\kappa_3 = \frac{\alpha^3 N^6/24}{(\alpha^2 N^4/12)^{3/2}} = \sqrt{3}, \tag{2.84}$$

as claimed above. This value agrees perfectly with our numerical calculations, which show the above approximation breaks down when $\alpha \approx 1/2$. Above this threshold, the random walk causes mixing between the two coupon collection regions, thereby lowering the overall skew of the fixation-time distribution toward the $\alpha = 1$ value of $\kappa_3 = 6\sqrt{3}(10-\pi^2)/(\pi^2-9)^{3/2} \approx 1.6711$.

2.9.6 Fixation-time distributions in the two-fitness Moran Process

As noted in Section 2.6.1, the two-fitness Birth-Death (BD) Moran process has the same family of fixation-time distributions as the Birth-death (Bd) process with only one fitness level. Here we provide further details leading to this conclusion. In particular, we give the transition probabilities for the two-fitness model and describe how the calculations from the sections above generalize to this system. Here r is the fitness level during the birth step, while $\tilde{r}$ is the fitness level during the death step in the Moran process.

One-dimensional lattice

On the 1D lattice, the Moran process with fitness at both steps (birth and death), has new transition probabilities

$$b_m = \frac{r}{rm + N - m}\frac{\tilde{r}}{1+\tilde{r}}, \qquad d_m = \frac{1}{r\tilde{r}}b_m, \tag{2.85}$$

for $1 < m < N - 1$. The probabilities are different when there is only one mutant or non-mutant ($m = 1$ or $m = N - 1$ respectively). In these cases the nodes on the population boundary don't have one mutant and one non-mutant as neighbors, as is the case for all other m. In the limit $N \gg 1$, however, changing these two probabilities does not affect the fixation-time distribution and we can use the probabilities given in Eq. (2.85).

The two-fitness Moran BD model on the 1D lattice differs from the previously considered Bd process in two ways. First, the transition probabilities have the same functional form as before, but are scaled by a factor $\tilde{r}(1+\tilde{r})^{-1}$. This factor determines the time-scale of the process but does not alter the shape of the fixation-time distribution because it drops out of the expressions for the cumulants, Eqs. (2.53) and (2.58). Second, the ratio $b_m/d_m = r\tilde{r}$ shows that the process is still a random walk, but with new bias corresponding to an effective fitness level $r_{\text{eff}} = r\tilde{r}$. With these observations, when $r_{\text{eff}} \neq 1$, our preceding analysis applies and we predict normally distributed fixation times. If $r_{\text{eff}} = 1$, the random walk is unbiased, and we expect highly skewed fixation-time distributions.

Complete Graph

On the complete graph, considering fitness during the replacement step leads to transition probabilities

$$b_m = \frac{rm}{rm + N - m} \cdot \frac{\tilde{r}(N-m)}{\tilde{r}(N-m) + m - 1}, \qquad d_m = \frac{N-m}{rm + N - m} \cdot \frac{m}{\tilde{r}(N-m-1) + m}. \tag{2.86}$$

In this case, the ratio of transition probabilities is m-dependent, but $b_m/d_m \to r\tilde{r}$ as $N \to \infty$, again motivating the definition of the effective fitness level $r_{\text{eff}} = r\tilde{r}$. If we take the large (but not infinite) fitness limit $r_{\text{eff}} \gg 1$, so that the mutant population is monotonically increasing to good approximation, then the fixation time cumulants are again given by Eq. (2.53) with $\lambda_m \to b_m + d_m$. As $N \to \infty$, the cumulants become

$$\kappa_n = \frac{1 + r^n/\tilde{r}^n}{(1 + r^2/\tilde{r}^2)^{n/2}} \cdot \frac{(n-1)!\zeta(n)}{\zeta(2)^{n/2}}, \tag{2.87}$$

identical to the Moran Bd process on the complete graph, with $r \to r/\tilde{r}$. Numerical calculations (using the moment recurrence relation derived in Section 2.9.3) again indicate this expression for the cumulants holds for all r, not just in the large fitness limit. When $r_{\text{eff}} = 1$, we expect highly skewed fixation distributions arising from the unbiased random walk underlying the dynamics. This is indeed the case, though numerics indicate there is an entire family of distributions dependent on $r = 1/\tilde{r}$.

CHAPTER 3

ASYMPTOTIC ABSORPTION-TIME DISTRIBUTIONS IN EXTINCTION-PRONE MARKOV PROCESSES

3.1 Introduction[1]

Modeling extinction-prone dynamics is essential to our understanding of epidemics, disease incubation, and evolution. For example, a key goal in epidemiology is to implement control measures (such as social distancing or vaccination) that push the dynamics toward a state where the disease is eradicated on a reasonable timescale [2, 52, 53]. Similarly, disease incubation [33, 54] and evolution [9, 13] involve highly fit infectious cells or mutant species outcompeting their less fit counterparts.

In these fields the distribution of extinction times, rather than just the mean, is crucial. For example, how long must a patient wait after exposure to a disease to be sure they are not infected? In the best and worst case scenarios, how long must epidemiological control measures be imposed to stop an outbreak? Knowledge of the extinction-time distribution provides an answer to these questions. Incubation period distributions have long been measured empirically to inform treatment regimens or public health initiatives [54]. Similarly, a recent study used a data-driven model of African sleeping sickness in the Democratic Republic of Congo to predict the distribution of times until the disease is eradicated [2].

[1]This chapter is reproduced from: David Hathcock and Steven H. Strogatz, "Asymptotic absorption-time distributions in extinction-prone Markov processes." *Physical Review Letters* **128**, 218301 (2022).

In this chapter, we show that two particular extinction-time distributions—Gaussian and Gumbel distributions—arise generically from basic features of the stochastic dynamics driving the system. These distributions were found previously in several models of evolutionary dynamics [1, 32, 33]. We show now that these same distributions appear in much more general classes of birth-death Markov chains, along with a family of skewed distributions that include the Gumbel. Extending the approach introduced in Ref. [1], we provide analytical criteria that predict when the asymptotic absorption-time distribution is normal, Gumbel, or a member of the family of skewed distributions. We apply our results to models of epidemiology [55–57], ecology [58–60], stochastic chemical reactions [7, 61], and evolutionary games [62], for which the predicted distributions agree with those measured via simulation. To our knowledge, this is the first calculation of the asymptotic absorption-time distributions for these models. As an application, we show that the Gumbel distribution closely resembles eradication-time distributions for African sleeping sickness.

We analyze birth-death Markov processes with a linear chain of states $m = 0, 1, \ldots, N$. For example, m might represent the number of infected individuals in an epidemic. The system has an absorbing state at $m = 0$ (where nobody is infected) and a reflecting state at $m = N$ (the maximum allowed infected population). Transitions occur only between neighboring states, i.e., the population can only increment by 1 in either direction. The dynamics of $p_m(t)$, the probability of occupying state m at time t, obey the master equation,

$$\dot{p}_m(t) = b_{m-1}p_{m-1}(t) + d_{m+1}p_{m+1}(t) - (b_m + d_m)p_m(t), \tag{3.1}$$

where b_m and d_m are respectively the birth and death rates at which the state increases or decreases from state m. The master equation can also be expressed as $\dot{\mathbf{p}}(t) = \Omega \cdot \mathbf{p}(t)$, where Ω is the transition matrix containing the birth and death

rates. Since the state at $m = 0$ is absorbing and the state $m = N$ is reflecting, we have $b_0 = b_N = 0$. For simplicity we assume the system starts in an initial state $m = N$, i.e. $p_m(0) = \delta_{m,N}$, but our results apply more broadly (see Section 3.6.3). The quantity we are interested in is the first-passage time T to the absorbing state $m = 0$; here we focus on obtaining the probability distribution about the mean.

3.2 Exact expression for the absorption-time cumulants

Building on our results in Chapter 2 and Ref. [1], we develop an approach to determine the absorption-time distributions for general classes of birth-death Markov chains in the limit of large system size. The key insight is to introduce a change of variables, $D_m = b_m + d_m$ and $r_m = b_m/d_m$. If the system is in state m, it waits on average a time D_m^{-1} before increasing or decreasing. The probabilities of the next step being forward or backward are $r_m/(1 + r_m)$ and $1/(1 + r_m)$ respectively; r_m is the ratio of these probabilities. Thus, our coordinate change separates the random-walk portion of the Markov process, which describes the relative probabilities of stepping forward or backward at each state, from the times spent waiting in each state. This change of variables leads to a transition matrix decomposition, $\Omega = \Omega_{RW} D$, where D is diagonal with elements D_m and Ω_{RW} is the transition matrix for a biased random walk. The number of times the system visits each state depends only on the random-walk portion of the process. The elements V_{ij} of $V = -\Omega_{RW}^{-1}$ encode the average number of visits to state i before absorption, starting from state j.

To characterize the asymptotic distributions, we compute the cumulants

$\kappa_n(N)$ of the absorption time T, which describe the shape of the distribution. For instance, κ_1 is the mean, κ_2 is the variance, and $\kappa_3/\kappa_2^{3/2}$ is the skew. Following Ref. [1] we use the matrix decomposition above to derive the cumulants (generalizing the previous result to non-constant r_j):

$$\kappa_n(\{r_j\}, N) = \sum_{1 \le i_1 \le i_2 \le \cdots \le i_n \le N} \frac{w^n_{i_1 i_2 \cdots i_n}(\{r_j\})}{(b_{i_1} + d_{i_1}) \cdots (b_{i_n} + d_{i_n})}. \tag{3.2}$$

Here $w^n_{i_1 i_2 \cdots i_n}(\{r_j\})$ are weighting factors that depend only on the visit statistics of the random walk; for example, $w^1_i(\{r_j\}) = V_{ii}$. Section 3.2.1 below provides a derivation of this formula and explicit expressions for the first few weighting factors, each of which are polynomials of the visit numbers V_{ij}. Equation (3.2) is equivalent to well known recursive relations for absorption time moments [63], but this form enables the asymptotic analysis leading to the results below.

3.2.1 Derivation of the absorption-time cumulants

In this section we derive the general formula for the absorption-time cumulants, Eq. 3.2. This derivation follows Ref. [1], but we generalize to Markov chains where the ratio $r_m = b_m/d_m$ is non-constant. We start from the master equation, Eq. (3.1) above, and restrict our attention to the transient (non-absorbing) states, $m > 0$, since these determine the time it takes to reach absorption. The master equation for these states can be expressed as $\dot{\mathbf{p}}(t) = \Omega \cdot \mathbf{p}(t)$, where Ω is the transient transition matrix with elements

$$\Omega_{mn} = b_n \delta_{m,n+1} + d_n \delta_{m,n-1} - (b_n + d_n)\delta_{m,n} \tag{3.3}$$

for $m, n = 1, \ldots, N$ and $\mathbf{p}(t)$ is the vector of transient state occupancy probabilities.

The entire first-passage process can be characterized in terms of the transition matrix Ω. In fact, the first-passage distribution $p(t)$ can be written in terms

of an element of the matrix exponential, $p(t) = d_1[\exp(\Omega t)]_{1,N}$ and the moments of T are

$$\mu_n \coloneqq E[T^n] = (-1)^n n! \mathbf{1}\Omega^{-n}\mathbf{p}(0), \tag{3.4}$$

where $\mathbf{1}$ is a row vector containing all 1's and E denotes expected value.

As noted above, to proceed it is useful to introduce the following decomposition of the transition matrix: $\Omega = \Omega_{RW}D$, where D is a diagonal matrix $D_{mm} = b_m + d_m$ and

$$[\Omega_{RW}]_{mn} = \frac{r_n}{1+r_n}\delta_{m,n+1} + \frac{1}{1+r_n}\delta_{m,n-1} - \delta_{m,n}, \tag{3.5}$$

with $r_n = b_n/d_n$. The rates D_{mm} determine how long the system waits in state m before taking a step and r_m is the relative probability of stepping forward versus backward along the chain. Defining $V = -\Omega_{RW}^{-1}$, the elements V_{ij} are the average number of visits to state i before absorption starting from an initial state j.

With the above decomposition we can easily invert the transition matrix,

$$[-\Omega^{-1}]_{ij} = \frac{V_{ij}}{b_i + d_i}, \tag{3.6}$$

where visit numbers V_{ij} are given by

$$V_{ij} = (1 + r_i) \sum_{n=1}^{\min(i,j)} \prod_{m=n}^{i-1} r_m. \tag{3.7}$$

Then, using Eq. (3.4) the moments can be expressed as

$$\mu_n = n! \sum_{i_1, i_2, \ldots i_n = 1}^{N} \frac{V_{i_1 i_2} V_{i_2 i_3} \cdots V_{i_{n-1} i_n} V_{i_n N}}{(b_{i_1} + d_{i_1})(b_{i_2} + d_{i_2}) \cdots (b_{i_n} + d_{i_n})}. \tag{3.8}$$

To compute the cumulants, we use the standard conversion formulas: $\kappa_1 = \mu_1$, $\kappa_2 = \mu_2 - \mu_1^2$, $\kappa_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3$, and so on. Since the relation between cumulants and moments is polynomial, if we collect terms with common denominators it

follows that the cumulants have the form quoted above,

$$\kappa_n(\{r_j\}, N) = \sum_{1 \le i_1 \le i_2 \le \cdots \le i_n \le N} \frac{w^n_{i_1 i_2 \cdots i_n}(\{r_j\})}{(b_{i_1} + d_{i_1}) \cdots (b_{i_n} + d_{i_n})}. \tag{3.9}$$

where the weights w^n depend on the visit numbers V_{ij} (and hence are functions of only the ratios $\{r_j\}$). Note that we sum over $i_1 \le i_2 \le \cdots \le i_n$, so that each product in the denominator of Eq. (3.9) appears exactly once. The weights are determined using Eq. (3.8) and the moment-cumulant conversion formulas. For example, the second and third cumulants are

$$\kappa_2 = \sum_{i,j=1}^{N} \frac{2V_{ij}V_{jN} - V_{iN}V_{jN}}{(b_i + d_i)(b_j + d_j)} \tag{3.10}$$

$$\kappa_3 = \sum_{i,j,k=1}^{N} \frac{6V_{ij}V_{jk}V_{kN} - 6V_{ij}V_{jN}V_{kN} + 2V_{iN}V_{jN}V_{kN}}{(b_i + d_i)(b_j + d_j)(b_k + d_k)}. \tag{3.11}$$

From here we can read off the weights w^n: they are simply the numerators in the above expressions, summed over distinct permutations of the indices (since these terms have the same denominators). Carrying out the sum we obtain,

$$w^2_{ij} = \sum_{\sigma \in \Pi_2} 2V_{\sigma_1\sigma_2}V_{\sigma_2 N} - V_{\sigma_1 N}V_{\sigma_2 N} \tag{3.12}$$

$$w^3_{ijk} = \sum_{\sigma \in \Pi_3} 6V_{\sigma_1\sigma_2}V_{\sigma_2\sigma_3}V_{\sigma_3 N} - 6V_{\sigma_1\sigma_2}V_{\sigma_2 N}V_{\sigma_3 N} + 2V_{\sigma_1 N}V_{\sigma_2 N}V_{\sigma_3 N}, \tag{3.13}$$

where Π_2 is the set of distinct permutations of indices $\{i, j\}$ and Π_3 is the set of distinct permutations of $\{i, j, k\}$.

3.2.2 Properties of the weighting factors

The weighting factors have some convenient properties. First, they are non-negative: $w^n_{i_1 i_2 \cdots i_n}(\{r_j\}) \ge 0$ and *increasing* functions of each r_j. Second, the weighting factors appear to fall off exponentially away from the diagonal. For constant

$r_j = r$, this exponential decay can be shown explicitly [1]. We conjecture that the same decay holds for arbitrary transition probabilities $\{r_j\}$. The intuition is that the visits to state i are uncorrelated with those to state j (for $N \gg 1$ and $i - j = O(N)$), due to the Markov property.

Here we will show the non-negativity and monotonicity explicitly up to order $n = 4$ and conjecture these properties hold for all orders. To proceed, we use the fact that $V_{ii} = V_{ij}$ for any $i < j$. This is easy to see from Eq. (3.7), but also has an intuitive physical interpretation. Since the system is eventually absorbed at the boundary state 0, if it starts from a state $j > i$ it must visit i before absorption. After the first visit, the statistics of the random walk are identical to a walk initialized in state i. Using this property the sum over permutations above dramatically simplifies. For $i < j$ we have

$$\begin{aligned} w_{ij}^2 &= (2V_{ij}V_{jN} - V_{iN}V_{jN} + 2V_{ji}V_{iN} - V_{jN}V_{iN}) \\ &= 2V_{ji}V_{iN}. \end{aligned} \tag{3.14}$$

Similarly, after simplification we find

$$w_{ijk}^3 = 3!\,(V_{kj}V_{ji}V_{iN} + V_{ki}V_{iN}V_{jN}) \tag{3.15}$$

$$\begin{aligned} w_{ijkl}^4 = 4!\,(&V_{lk}V_{kj}V_{ji}V_{iN} + V_{lk}V_{ki}V_{iN}V_{jN} + V_{lj}V_{ji}V_{iN}V_{kN} \\ &+ V_{lj}V_{jN}V_{ki}V_{iN} + V_{li}V_{iN}V_{kj}V_{jN} + V_{li}V_{iN}V_{kN}V_{iN}) \end{aligned} \tag{3.16}$$

when $i < j < k < l$. When some indices are identical, these results still hold, but they must be divided by the number of permutations of the identical indices, e.g., $w_{ii}^2 = V_{ii}V_{iN}$ (notice this differs from Eq. (3.14) by a factor of 2). The important feature of these expressions is that they are *positive* sums of products of the visit numbers V_{ij}. We conjecture that the weights at every order can also be written as a positive sums of products of the visit numbers (though we omit

the expressions here, we have checked this is true up to order $n = 6$). If this is the case, it immediately follows that the weights w^n are positive and increasing functions of each r_j because the visit numbers, Eq. (3.7), themselves also have these properties.

3.3 Gaussian Universality Class

The first universality class of birth-death Markov chains we consider have normally distributed absorption times. As an instructive special case, consider the process $b_m = 0$, $d_m = d$, which visits each state exactly once before absorption, waiting a time d^{-1} on average at each. The time to absorption is simply $T = \sum_m \mathcal{E}_m(d)$ where $\mathcal{E}_m(d)$ is an exponential random variable. Since T is a sum of identical random variables we expect it to be normally distributed by the central limit theorem. Alternatively, the cumulants of T are $\kappa_n = N/d^n$. In units of the standard deviation the higher order cumulants vanish: $\kappa_n/\kappa_2^{n/2} = N^{1-n/2} \to 0$ as $N \to \infty$. Hence the distribution is asymptotically normal.

We might also expect this asymptotic normality to hold for transition rates with mild state dependence: if $b_m + d_m$ does not vary too much (we will give a precise condition below), the absorption time is a sum of nearly identical exponential random times. Similarly, for $r_m = b_m/d_m > 0$, the system randomly walks back and forth, but as long as $r_m < 1$ the average number of visits to each state is finite. Under either of these generalizations the distribution is asymptotically normal.

To characterize more precisely which Markov chains lead to normally distributed absorption times, we compute the asymptotic form of the cumulants in

Eq. (3.2) by introducing two auxiliary Markov chains. These have the same b_i+d_i as the original system, but b_i and d_i are adjusted so that the ratios are $r_j = r_{\max}$ or $r_j = r_{\min}$, where $r_{\max} = \lim_{N\to\infty} \max_{1<j<N} r_j$ and $r_{\min} = \lim_{N\to\infty} \min_{1<j<N} r_j$. In other words, we construct two Markov chains where the time spent waiting in each state is identical to that for the original system, but the odds of moving toward the absorbing state are increased or decreased to be uniform.

Above we noted that the weighting factors w^n in Eq. (3.2) are increasing functions of r_j. Thus, we can bound the cumulants in our system by those for the auxiliary Markov chains, $\kappa_n(r_{\min}, N) \leq \kappa_n(\{r_j\}, N) \leq \kappa_n(r_{\max}, N)$. The asymptotic form of $\kappa_n(r, N)$ (where r is constant across states) was computed in Ref. [1]; we utilize this result in the asymptotic analysis given in the Appendix, Section 3.9.1. To nail down the asymptotics of $\kappa_n(r, N)$ we require the waiting times to be 'flat' in the following sense:

$$\frac{1}{N}\sum_{m=1}^{N} t_m \sim c \max_{1\leq m\leq N} t_m, \tag{3.17}$$

where $t_m = (b_m+d_m)^{-1}$ is the mean waiting time at state m and c is a constant independent of N. In other words, the mean waiting time $\langle t_m \rangle$ across all states is the same asymptotic order as the maximum waiting time: the process fluctuates at an approximately uniform rate across the entire Markov chain, without spending a disproportionate amount of time in any one state. Gaussian absorption times have also been found in the continuum limit via the linear-noise approximation, which removes state dependence from the noise [64]. This approximation is similar to the condition (3.17), which requires the noise amplitude $b_m + d_m$ to vary only mildly across states.

If Eq. (3.17) holds, then $\kappa_n(r, N) \sim c_n(r) f(N)^n N$, where $f(N) \sim \max_{1\leq i\leq N}(b_i+d_i)^{-1}$. Since these asymptotics hold for $r = r_{\min}$ and $r = r_{\max}$, it follows that $\kappa_n(\{r_j\}, N) \sim$

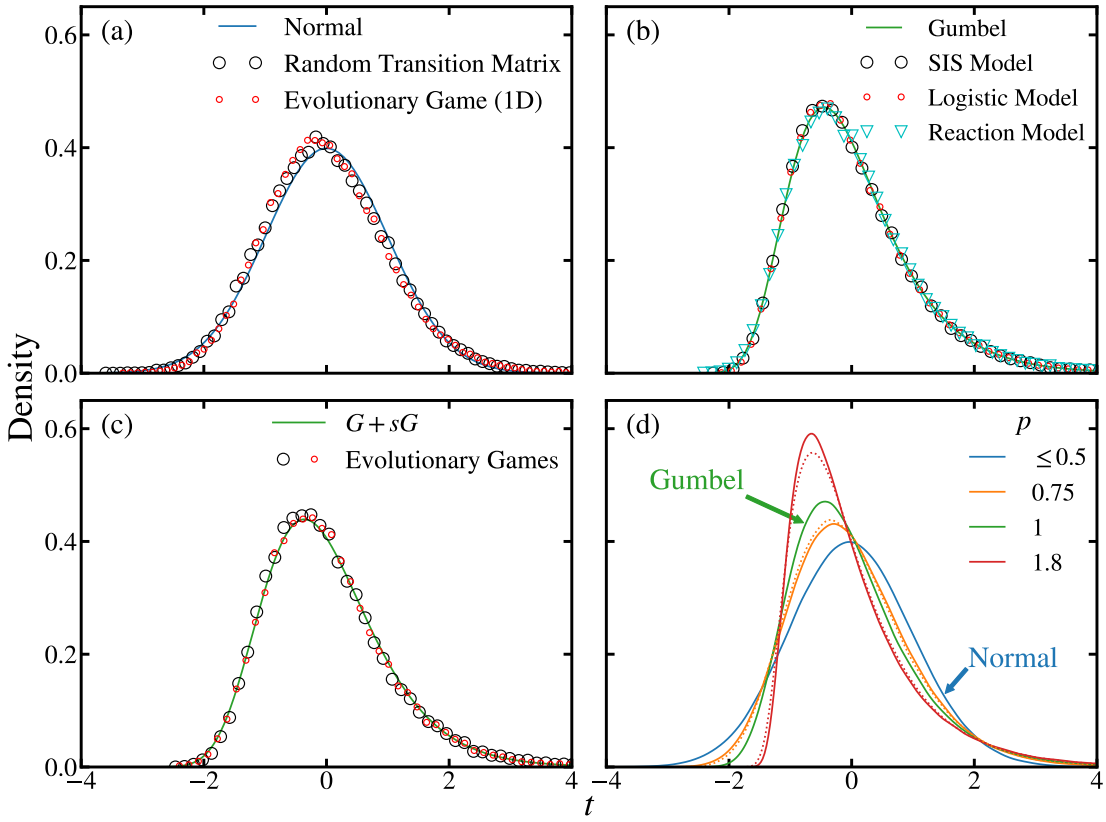

Figure 3.1: Absorption-time distributions for (a) the random transition matrix model (large black circles) and the evolutionary game on a ring (small red circles), (b) SIS model (large black circles), logistic model (small red circles), and autocatalytic chemical reaction model (cyan triangles), (c) the well-mixed evolutionary game, and (d) the process $b_m = rd_m = rm^p$, for $r = 0$ and $p = 0.3$ (blue), $p = 0.75$ (orange), $p = 1$ (green), and $p = 1.8$ (red). The $r = 0.8$ distributions are indicated by dotted lines (when they differ from the $r = 0$ counterparts). See Section 3.8 for models and parameters. We used system sizes (a-b) $N = 500$ and (c-d) $N = 1000$ and simulated (a) 5×10^4, (b-c) 10^5, and (d) 10^6 trials to measure the distributions, which have been standardized to have zero mean and unit variance. In (c) the distributions are a convolution of Gumbel distributions with relative weighting $s \approx 0.73$. Deviations from predicted normal and Gumbel distributions in (a-c) are due to finite system size.

$c_n(\{r_j\})f(N)^n N$ as well.

With the asymptotic form of the cumulants established, we analyze the shape of the distribution using the standardized cumulants $\tilde{\kappa}_n = \kappa_n/\kappa_2^{n/2}$ for $n \geq 2$ (which are rescaled so that the variance $\tilde{\kappa}_2 = 1$). Using the asymptotic form ob-

tained above, we find $\tilde{\kappa}_n \sim \tilde{c}_n N^{1-n/2}$. In particular, $\tilde{\kappa}_n \to 0$ as $N \to \infty$ for $n > 2$, so that the distribution becomes Gaussian for large N (the cumulants past second order vanish for normal distributions).

For finite N, the dominant correction to the normal distribution comes from the non-zero skew $\tilde{\kappa}_3 \sim \tilde{c}_3/\sqrt{N}$. The coefficient in this scaling depends on the ratios r_j; in the Appendix, Section 3.9.1 we compute a bound on this coefficient, which is useful for estimating the rate of convergence in applications. The ratio of the standard deviation $\kappa_2^{1/2}$ to the mean κ_1 also scales like $\kappa_2^{1/2}/\kappa_1 \sim \tilde{c}_1/\sqrt{N}$, similar to the skew. As the distribution converges to the Gaussian, the relative width of the distribution narrows at the same rate. To summarize, any birth-death Markov chain that satisfies the 'flatness' condition, Eq. (3.17), and has an absorbing state toward which the system flows on average ($r_j < 1$) will have asymptotically Gaussian distributed absorption times.

Our first example of a Markov chain with normally distributed absorption times is a toy model with *random* transition probabilities. Here we select $b_m + d_m$ uniformly at random between 0.1 and 2 and r_m uniformly at random between 0 and 0.9, which satisfies the conditions described above. This example shows that the transition rates need not be smooth in m; systems with disordered transition rates still belong to this universality class.

Next we study evolutionary game dynamics on a one-dimensional ring [19, 65]. Mutant and wild-type individuals compete via the following dynamics: an individual is chosen randomly, proportional to its (frequency dependent) fitness. The selected individual gives birth to an offspring of the same type, which in turn replaces a random neighbor. The model runs until the mutation spreads to the entire population.

Figure 3.1(a) shows simulation results for the random transition system and the evolutionary game. Both display the expected normal distribution. Interestingly, for the evolutionary game, the normal distribution appears for a wide range of parameters, while the mean absorption time and absorption probability depend more intricately on parameters [19, 65].

3.4 Gumbel Universality Class

Gumbel distributions, known for their role in extreme value theory [66], also arise generically in absorption processes. This second universality class is closely related to the 'coupon collector' problem in probability theory, which asks the following: if there are N distinct coupons and we are given a random one (with replacement) at each time step, how long does it take to collect all N coupons? The collection process displays a characteristic slowdown: when nearly all coupons have been collected, it takes a long time to acquire the final few because duplicates keep getting selected. Erdős and Rényi showed that for large N the time to complete the collection follows a Gumbel distribution [36].

The coupon collector problem can be modeled using Markov chains. Let m be the number of coupons missing from the collection of N total coupons. The probability of obtaining a new coupon (thereby decreasing m) is m/N and the number of missing coupons never increases. Thus, the coupon collection process is described by birth-death dynamics with $b_m = 0$ and $d_m = m/N$. The linear decay of the transition probability d_m near the absorbing boundary is the key feature that gives rise to the characteristic slowdown. For this case the cumulants can be computed exactly, $\tilde{\kappa}_n = (n-1)!\zeta(n)/\zeta(2n)^{n/2}$, and match those for a Gumbel

distribution. Similar to the Gaussian class above, we find that the Gumbel distribution is preserved for non-zero $r_m < 1$ and nonlinear transition rates as long as the linear decay is dominant near 0. Specifically, if $b_m + d_m = f(N)m[1 + O(m/N)]$, with $b_{\alpha N} + d_{\alpha N}$ of order at least $O(Nf(N))$ for any $0 < \alpha \leq 1$, and if $r_m = r + O(m/N)$ for large N, then the absorption-time distribution is asymptotically Gumbel [2].

By bounding the cumulants (3.2), we show (Section 3.9.2) their leading order behavior for $N \gg 1$ is dominated by the states near 0, where the approximations $b_m + d_m \approx f(N)m$ and $r_m \approx r$ become asymptotically exact, so that

$$\kappa_n(\{r_j\}, N) \sim \frac{1}{f(N)^n} \sum_{1 \leq i_1 \leq i_2 \leq \cdots \leq i_n \leq N} \frac{w^n_{i_1 i_2 \cdots i_n}(r)}{i_1 i_2 \cdots i_n}. \tag{3.18}$$

The factors $f(N)^n$ set the timescale of the process but do not affect the shape of the distribution (they cancel in $\tilde{\kappa}_n = \kappa_n / \kappa_2^{n/2}$). Thus, we have shown that the cumulants are asymptotic to those for a process with $b_m + d_m = m$ and $b_m / d_m = r$. The absorption-time distribution for this process can be computed exactly (see Ref. [59, Appendix B]) and approaches a Gumbel distribution as $N \to \infty$ (see Section 3.9.2). Therefore, any system with transition rates vanishing linearly and ratios r_j that approach a constant near the absorbing boundary will fall into the Gumbel universality class.

As in the Gaussian class, the relative width of the Gumbel distributions becomes small for $N \gg 1$. In this case, however, the standard deviation-to-mean ratio scales like $\kappa_2^{1/2} / \kappa_1 \sim C_1 / \ln N$. On the other hand, the deviations from the Gumbel cumulants decay like $\delta\tilde{\kappa}_n = \tilde{\kappa}_n - \tilde{\kappa}_n^{\text{Gumbel}} \sim C_n N^{-1} \ln N$ (see Appendix, Section 3.9.2). Thus the distribution narrows very slowly compared to the convergence to the Gumbel shape. Therefore, in applications we expect to see the

[2] More generally it is sufficient to have $b_m + d_m = f(N)m[1 + O(m/g(N)]$ for any function $g(N)$ that diverges for $N \to \infty$. If $g(N)$ grows linearly or sublinearly, deviations in the cumulants scale like $\delta\tilde{\kappa}_n \sim C_n[\ln g(N)]/g(N)$. Otherwise, $\delta\tilde{\kappa}_n \sim C_n/N$

Gumbel distribution appear before the fluctuations become negligible.

Finally, if the transition rates vanish near the initial condition N, scaling like $b_m + d_m = \tilde{f}(N)(N-m) + O((N-m)^2)$, there will be another coupon-collection slowdown at the beginning of the process. An identical analysis to that above shows that the contributions from the two coupon collection regions simply add together to give the cumulants. The resulting absorption-time distribution is therefore a convolution of two Gumbels, with one weighted by $s = \lim_{N\to\infty} f(N)/\tilde{f}(N)$.

To illustrate the Gumbel universality class we use the susceptible-infected-susceptible (SIS) model of epidemiology [57], the logistic model from ecology [58], and an autocatalytic chemical reaction model [7, 61] (model details in Section 3.8). In each case the transition rates decrease linearly near the absorbing state. For example, in the SIS model, $b_m = \Lambda m(1-m/N)$ and $d_m = m$, where Λ is the infection rate.

Our simulations show that these models each have the expected Gumbel distribution [Figure 3.1(b)]. The distribution is also insensitive to parameter choices (e.g., a Gumbel appears in the SIS model for any $\Lambda < 1$).

If we study the aforementioned evolutionary game in a well-mixed population, the transition rates vanish linearly as $m \to 0$ and $m \to N$ (see Section 3.8.1 and Ref. [67]) . As discussed above, we expect a convolution of Gumbel distributions with relative weighting s given by the ratio of the linear coefficients at these two boundaries. Figure 3.1(c) shows that this prediction is borne out in simulations.

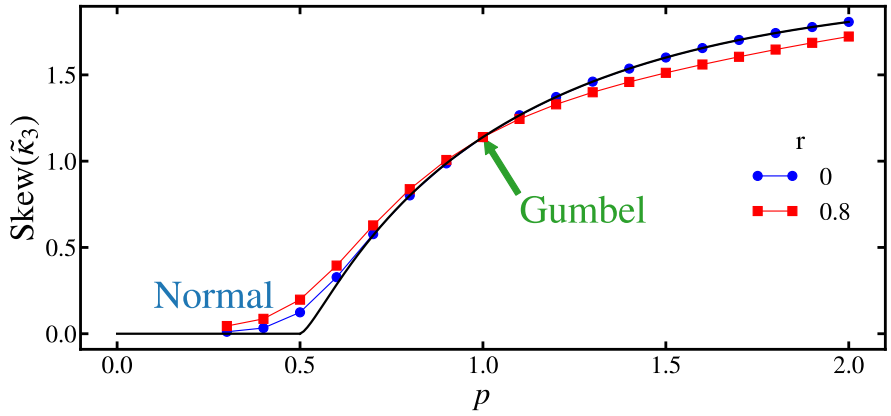

Figure 3.2: Absorption-time skew for the process $b_m = rd_m = rm^p$ with $r = 0$ (blue circles) and $r = 0.8$ (red squares), plotted as a function of the power-law exponent p. Skews were numerically computed for $N = 10^5$ using the recurrence relation approach described in Ref. [1]. The black line shows the asymptotic skew $2\zeta(3p)/\zeta(2p)^{3/2}$ for $r = 0$. The curves cross at $p = 1$ where the distribution is Gumbel, independent of r. For $p \leq 0.5$ the skew approaches zero and the distribution is Gaussian. The numerical skew is slightly larger than expected for $p \lesssim 0.6$ due to finite size effects.

3.5 Absorption-time distributions for power-law processes

In addition to Gumbel and Gaussian classes, other absorption-time distributions arise if the transition rates have power-law decay: $b_m + d_m = f(N)m^p[1 + O(m/N)]$. For $p < 1/2$, the decay is sufficiently slow that the normal distribution is maintained: the system still fluctuates at an approximately uniform rate across states. On the other hand, if $p > 1/2$ we find a generalized coupon collection phenomenon giving rise to a family of skewed distributions. Slowdown near the boundary dominates the absorption process and the distribution is asymptotic to that for the minimal model $b_m = rd_m = rm^p$ (rigorous asymptotics given in Section 3.9.3). When $r = 0$ the cumulants can be computed analytically: $\tilde{\kappa}_n = (n-1)!\zeta(np)/\zeta(2p)^{n/2}$ [32, 33]. Figure 3.1(d) shows the resulting distributions for a few values of p. Interestingly for $p \neq 1$, the shape of the distribution depends subtly on r. Figure 3.2 shows the skew of these distributions as a func-

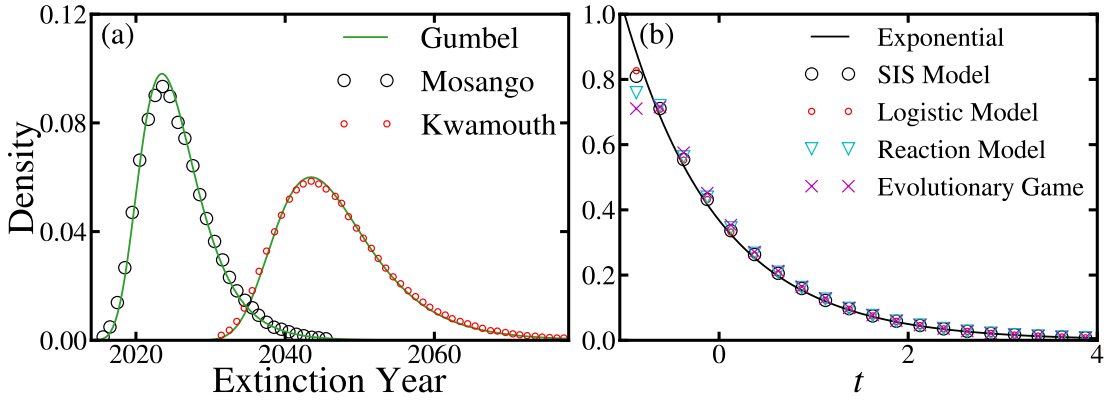

Figure 3.3: Generalizations to high-dimensional models and Markov chains with internal sinks. (a) Extinction-time distributions for sleeping sickness predicted using a 17-dimensional compartmental model that was fit to case data from the Mosango (large black circles) and Kwamouth (small red circles) regions of the Democratic Republic of Congo (data from Ref. [2]). Mean extinction times (measured from 2016) are approximately 9.5 and 31 years for the Mosango and Kwamouth regions respectively, with standard deviations of 4.8 and 7.9 years. Disease eradication times approximately follow a Gumbel distribution (fit using the mean and variance). (b) Simulations of the SIS, logistic, reaction, and well-mixed evolutionary game models have exponential absorption-time distributions (standardized to zero mean and unit variance) if parameters are chosen so that the dynamics have an internal sink state. For each case, we used $N = 50$ and simulated 10^6 trials. See Section 3.8 for model details and parameters.

tion of p, elucidating the transition from normal distributions to the skewed family.

3.6 Extensions

3.6.1 High-dimensional models

Beyond simple one-dimensional Markov processes, the eradication-time distributions for African sleeping sickness predicted by a 17-dimensional data-driven model [2] closely resemble the Gumbel [Figure 3.3(a)]. This result suggests that

the Gumbel distribution is also generic in higher dimensions if the dynamics collapse onto a one-dimensional slow manifold near absorption. Crucially, although the distributions have converged to the Gumbel shape, the fluctuations still matter: the probable extinction times span years. The ratio between the standard deviation and the mean is approximately 0.5 and 0.25 for the Mosango and Kwamouth regions respectively. Similar results hold for a variety of high-dimensional systems. Their dynamics are accurately approximated by birth-death processes with transition rates that vanish as a power-law m^p near the boundary. Examples include evolutionary dynamics on D-dimensional lattices ($p = 1 - 1/D$) and complex networks [4, 32, 33] as well as epidemics on networks [3].

3.6.2 Transition matrix spectrum

So far, we have characterized universality classes for absorption times in birth-death Markov chains. While our results are formulated in terms of the transition rates b_i and d_i, we can also connect the shape of the absorption-time distribution to the spectrum of the transition matrix. In the following sections, we discuss the classes of transition matrix spectra that give rise to either Gaussian or Gumbel absorption-time distributions.

Eigenvalue spectrum for the Gaussian class

If Eq. (3.17) is satisfied with $b_i + d_i$ replaced by the eigenvalues λ_i of the negative transition matrix $-\Omega$, the absorption-time distribution will be Gaussian. To show this, we use the spectral representation of the absorption-time cumulants

[1, 67],

$$\kappa_n = (n-1)! \sum_{i=1}^{N} \lambda_i^{-n}. \tag{3.19}$$

If Eq. (3.17) is satisfied for the eigenvalues, we have

$$\sum_{i=1}^{N} \lambda_i^{-1} \sim c_n \left(\max_{1 \le i \le N} \lambda_i^{-1} \right) N, \tag{3.20}$$

Since the left-hand side of this expression is exactly the cumulant κ_n (up to the constant $(n-1)!$), it immediately follows that $\kappa_n \sim c_n g(N)^n N$ where $g(N) \sim \max_{1 \le i \le N} \lambda_i^{-1}$ as $N \to \infty$. This scaling implies that the standardized cumulants vanish for large N: $\tilde{\kappa}_n \sim \tilde{c}_n N^{1-n/2}$ and the distribution is asymptotically Gaussian.

More generally, the distribution approaches a Gaussian as long as $\tilde{\kappa}_n \to 0$ as $N \to \infty$. This condition with Eq. (3.19) describes a broader class of eigenvalue spectra that give rise to Gaussian absorption-time distributions. Specifically, we need

$$\left(\sum_{i=1}^{N} \lambda_i^{-n} \right) \Big/ \left(\sum_{i=1}^{N} \lambda_i^{-2} \right)^{n/2} \xrightarrow{N \to \infty} 0. \tag{3.21}$$

While this condition is difficult to interpret, we consider two examples that illustrate the type of spectra that can give rise to Gaussian absorption-time distributions. First, if $\lambda_m = m^p$, the above condition is satisfied for $p \le 1/2$. This result is related to the emergence of Gaussian distributions for the systems considered in Section 3.5, which have transition rates that decay as a power-law with $p \le 1/2$. Also, if $\lambda_m = P(m)/Q(m)$ for some polynomials P and Q, the condition is satisfied when the degree of Q is greater than that of P.

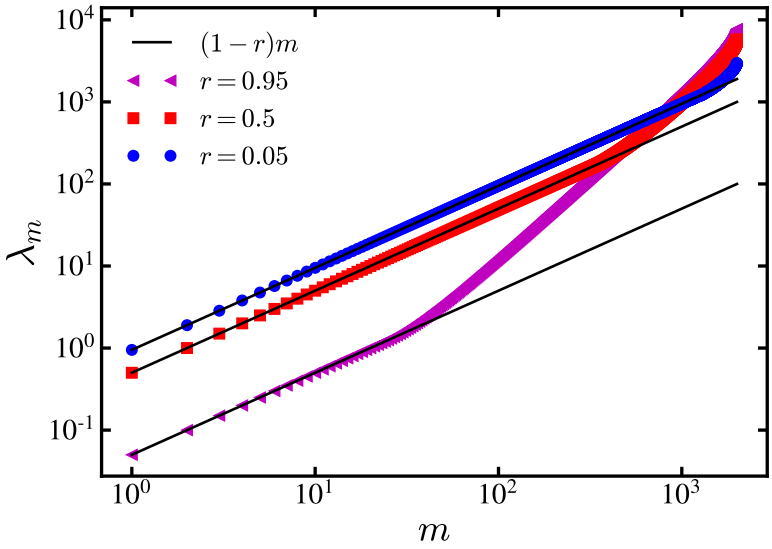

Figure 3.4: The eigenvalues of the transition matrix for the canonical model $b_m = rm$, $d_m = m$ with $N = 2000$ and $r = 0.05$, 0.5, and 0.95 plotted on a log-log scale. The black lines show $(1 - r)m$ for each value of r. The eigenvalues closely follow this linear relation up to a cut-off m_c that is dependent on r. Since the leading eigenvalues are linear the absorption-time distribution is Gumbel.

Eigenvalue spectrum for the Gumbel class

The Gumbel distribution arises if the transition matrix eigenvalues decay linearly. For instance, suppose $\lambda_m = bm$. Then, using Eq. (3.19) and taking $N \to \infty$, we have that

$$\tilde{\kappa}_n = \left((n-1)! \sum_{m=1}^{\infty} (bm)^{-n}\right) \Big/ \left(\sum_{m=1}^{\infty} (bm)^{-2}\right)^{n/2} = (n-1)! \frac{\zeta(n)}{\zeta(2)^{n/2}}, \tag{3.22}$$

which are precisely the cumulants for a standardized Gumbel distribution. The result is unchanged if the dominant eigenvalues are approximately linear, i.e. $\lambda_m \approx bm$ for $m < \alpha N$ where α is a constant $0 < \alpha < 1$. In this case, the standardized cumulants are still $\tilde{\kappa}_n = (n-1)!\zeta(n)/\zeta(2)^{n/2}$ with the larger eigenvalues contributing $O(1/N)$ corrections that vanish asymptotically.

This second case appears to be what happens in practice: for $N \gg 1$ the eigenvalues become linear up to a cutoff. We have carried out numerical cal-

culations of the spectrum for the canonical model, $b_m = rm$, $d_m = m$ for a few values of $r < 1$. As shown in Figure 3.4, the leading eigenvalues in the spectrum become equally spaced: $\lambda_m \approx (1 - r)m$ for indices below a cutoff m_c. Numerical tests indicate m_c is approximately a constant proportion of N, i.e. $m_c \approx \alpha(r)N$, where $0 < \alpha(r) < 1$. Above this cutoff the eigenvalues grow super-linearly. The above calculation illustrates how the Gumbel absorption-time distribution arises in this model from the perspective of the eigenvalue spectrum.

3.6.3 Distributions are robust to changes in initial and boundary conditions

In this chapter we specialize to Markov chains with a finite state space of size N, a reflecting upper boundary, and initial condition at the maximal state $p_m(0) = \delta_{m,N}$. Our asymptotic absorption-time distributions, however, should be robust to changes in initial and boundary conditions. Because the dynamics are extinction-prone, the system quickly progresses toward the absorbing state, spending negligible time near the reflecting boundary. Therefore, if the initial condition m_0 is sufficiently large ($m_0 \sim N$ for large N), corrections due to variation in the initial condition will be sub-dominant as $N \to \infty$. By the same argument, we expect the same asymptotic distributions to occur for infinite systems with free boundary conditions and no maximal state N, but large initial condition. On the other hand, if the upper boundary is absorbing, our result describes the absorption-time distribution, given that the absorbing state at 0 is reached (i.e. if we ignore all trajectories that are absorbed at the upper boundary) [1]. Finally, our results can also be used to determine the first-passage-time

distribution to an arbitrary state m, since the first-passage problem can be solved by making the target state absorbing [61].

3.7 Future Directions

Future work might focus on characterizing additional universality classes beyond those studied here. For example, simulations [Figure 3.3(b)] show that exponential absorption-time distributions arise frequently in systems with an internal sink state, toward which transitions are more likely [68]. The emergence of the exponential distribution makes sense intuitively: the system quickly settles into a quasiequilibrium mode around the sink, whose slow exponential decay dominates the absorption process [69]. To our knowledge, however, there is no rigorous classification of this case. It would also be fascinating to investigate whether there is a universal crossover between different members of our family of absorption-time distributions. For example, how do the distributions change if the transition rates have mixed decay $m^p + \epsilon m^q$? Understanding the crossover scaling between these cases will enable the classification for an even broader class of extinction-prone Markov chains.

3.8 Example models

To conclude this chapter, we provide details of the evolutionary game, SIS, logistic, and autocatalytic chemical reaction models, each of which exhibit Gaussian, Gumbel, or exponential absorption-time distributions in different parameter regimes. Parameters used for the simulations presented above are provided

Table 3.1: Parameter choices for the simulations used to measure absorption-time distributions shown in Figures 3.1(a)-(c) and 3.3(b). See Section 3.8 for model and parameter definitions. Evolutionary games use well-mixed population structure except in Figure 3.1(a). In Figure 3.1(c) the relative weighting of the convolution of Gumbel distributions is $s = (1 + e^{\beta(c-a)})/(1 + e^{\beta(b-d)}) \approx 0.73$ for both sets of parameters.

Figure	Model	Parameters
3.1(a)	1D Evolutionary Game	$\beta = 1, a = 2, b = 4, c = 1, d = 0.1$
3.1(b)	SIS Model	$\Lambda = 0.5$
	Logistic Model	$B = 0.5, K = 1$
	Chemical Reaction Model	$k_1 = 1, k_2 = 0.75, k_3 = 1.25$
3.1(c)	Evolutionary Game (black)	$\beta = 1, a = 1, b = 0.5, c = 0.8, d = 0.1$
	Evolutionary Game (red)	$\beta = 2, a = 0.3, b = 1.3, c = 0.06, d = 1.2$
3.3(b)	SIS Model	$\Lambda = 1.4$
	Logistic Model	$B = 1.4, K = 1$
	Chemical Reaction Model	$k_1 = 1, k_2 = 1.35, k_3 = 0.14$
	Evolutionary Game	$\beta = 1, a = 1, b = 1.5, c = 1.2, d = 1$

in Table 3.1.

3.8.1 Evolutionary games

In the preceding sections, we present absorption-time distributions measured via simulations of a two-strategy evolutionary game. In this game, two types of individuals, mutants (M) and wild-types (W), compete and have *frequency dependent* fitness, which means that an individual's fitness depends on the identity of its neighbors. This dependence is encoded by the payoff matrix,

	M	W
M	a	b
W	c	d

(3.23)

For example, a mutant (M) with 2 mutant neighbors and 3 wild-type neighbors will have payoff $\pi = 2a + 3b$. The fitness is then $\exp(\beta\pi)$, where the parameter

β, the selection intensity, controls how strongly payoff influences fitness. This choice is known as the exponential fitness mapping [67]; we note that other fitness mappings do not change the qualitative behavior discussed below. The dynamics of the model are as follows: an individual is chosen randomly, proportional to their fitness. The selected individual gives birth to an offspring of the same time (M or W) which replaces a random neighbor (selected uniformly). We will let m denote the number of wild-types in the population. Thus, when $m = 0$, the mutants have taken over the population (in the jargon, the mutation becomes fixed). We focus on cases in which the mutation becomes fixed, ignoring those when the mutation dies out (which have infinite absorption time). We consider evolution in two types of network populations: a one-dimensional (1D) ring of individuals and a well-mixed (complete graph) population. Each exhibits different absorption-time behavior.

1D ring population structure

First we consider individuals connected in a 1D periodic ring [19]. Assuming a single initial mutant, the mutant population grows as a connected chain. Any changes in the population must occur at the boundary between mutants and wild-types. The boundary mutants and wild-types have payoff $a + b$ and $c + d$ respectively (they have one of each type as a neighbor). Thus the probability b_m of removing a mutant, and the probability d_m of adding a mutant, are given by

$$b_m = e^{\beta(c+d)}/F_m\,, \qquad d_m = e^{\beta(a+b)}/F_m\,, \qquad \text{for } 1 < n < N-1, \tag{3.24}$$

where F_m is the average fitness:

$$F_m = 2e^{\beta(a+b)} + (N-m-2)e^{\beta 2a} + 2e^{\beta(c+d)} + (m-2)e^{\beta 2d}. \tag{3.25}$$

The rates are slightly different for $m = 1$ and $m = N - 1$ [19]. For example, when $m = 1$ there is a single wild-type with two mutant neighbors. These transition rates are:

$$b_1 = \frac{e^{\beta 2c}}{2e^{\beta(a+b)} + (N-3)e^{\beta 2a} + e^{\beta 2c}} \qquad d_1 = \frac{e^{\beta(a+b)}}{2e^{\beta(a+b)} + (N-3)e^{\beta 2a} + e^{\beta 2c}} \quad (3.26)$$

$$b_{N-1} = \frac{e^{\beta(c+d)}}{e^{\beta 2b} + 2e^{\beta(c+d)} + (N-3)e^{\beta 2d}} \qquad d_{N-1} = \frac{e^{\beta 2b}}{e^{\beta 2b} + 2e^{\beta(c+d)} + (N-3)e^{\beta 2d}}. \quad (3.27)$$

For large N, however, these changes to the transition rates do not affect the absorption-time distribution. One can check that these transition rates satisfy the requirements of the Gaussian universality class if $(a + b) > (c + d)$.

Well-mixed population

If the population is well-mixed, every individual has contact with every other, and hence their fitness depends simply on the fraction of mutants in the population. The payoffs (per contact) for mutants and wild-types respectively are $\pi_M = a(N-m-1)/(N-1) + bm/(N-1)$ and $\pi_W = c(N-m)/(N-1) + d(m-1)/(N-1)$, where again a, b, c, and d are elements of the payoff matrix Eq. (3.23) and m is the number of wild-types in the population. The rates at which the wild-type population increases or decreases are [67]

$$b_m = \frac{m\, e^{\beta\pi_W}}{m\, e^{\beta\pi_W} + (N-m)\, e^{\beta\pi_M}} \frac{(N-m)}{N-1}, \qquad d_m = \frac{(N-m)\, e^{\beta\pi_M}}{m\, e^{\beta\pi_W} + (N-m)\, e^{\beta\pi_M}} \frac{m}{N-1}. \quad (3.28)$$

For the birth (death) rate the first fraction represents the probability of choosing a wild-type (mutant) to give birth, while the second fraction is the probability of the offspring replacing a mutant (wild-type) in the populations.

Probability flows toward the absorbing state ($r_m < 1$) if $b > d$ and $a > c$. From the transition probabilities it is clear $b_m + d_m$ decays linearly near both $m = 0$

and $m = N$. Expanding around these points, $b_m + d_m = f(N)(N - m) + O(m)$ and $b_m + d_m = \tilde{f}(N)(N - m) + O((N - m)^2)$. Our theory predicts the distribution will be a convolution of Gumbel distributions with relative weighting $s = \lim_{N\to\infty} f(N)/\tilde{f}(N)$. Taking this limit for the transition probabilities in Eq. (3.28) we find

$$s = \frac{1 + e^{\beta(c-a)}}{1 + e^{\beta(b-d)}}. \tag{3.29}$$

3.8.2 SIS model

The stochastic susceptible-infected-susceptible (SIS) model of epidemiology [57] describes the spread of an infectious disease through a population. The population is broken into two groups, those susceptible to the disease and those currently infected. This model describes diseases that do not confer immunity following recovery (or the immunity lasts only for a short time compared to the time scale on which the disease spreads). The rate per contact at which the disease is transmitted between individuals is Λ/N, and we set the time-scale so that the recovery rate is 1. Letting m represent the number of infected individuals, there are $m(N - m)$ contacts between infected and susceptible people in a well-mixed population. The m infected individuals each recover at rate 1. Thus the rates at which the infected population increases and decreases are respectively

$$b_m = \Lambda m(1 - m/N), \qquad d_m = m. \tag{3.30}$$

This system has the vanishing transition probabilities near $m = 0$, indicating coupon collection behavior (it is also straightforward to explicitly check it satisfies our requirements for the universality class as long as $\Lambda < 1$). Our simulations show that it has the expected Gumbel distribution of times for the infection to die out (Figure 3.1(b) above).

3.8.3 Logistic model

The stochastic logistic model describes the dynamics and fluctuations of an ecological population [58]. The model assumes a constant birth rate B per individual as well as a constant death rate (which we set to 1 by choosing the appropriate time scale) when the population is sparse. For higher populations, competition between individuals increases the death rate quadratically. The transition rates are

$$b_m = Bm, \qquad d_m = m + Km^2/N, \tag{3.31}$$

where the parameter K controls how strongly competition influences death rates (this parameter is related to the carrying capacity of the ecosystem). Again the transition rates vanish linearly near $m = 0$ and this model belongs to the Gumbel universality class as long as $B < 1$.

3.8.4 Autocatalytic chemical reaction model

Our final example model describes a stochastic autocatalytic chemical reaction [7],

$$2X + A \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 3X, \qquad X \xrightarrow{k_3} B, \tag{3.32}$$

where k_i are the reaction rates. The concentrations of species A and B are fixed at saturation levels and we want to describe the dynamics and fluctuations of m, the number of particles of species X. This is a variation of the Schlögl model where the reaction $X \to B$ is irreversible and the reactions cease when no particles of X remain. Applying our results to this model we will classify the distribution of reaction times: how long does the reaction proceed before the supply of X is exhausted.

The birth-death transition rates for the reaction given above are

$$b_m = \frac{k_1}{N} m(m-1), \qquad d_m = \frac{k_2}{N^2} m(m-1)(m-2) + k_3 m. \tag{3.33}$$

Again, the transition rates decay linearly near the absorbing boundary at $m = 0$, indicating the Gumbel universality class; it is straightforward to check that the required expansions hold. The conditions that guarantee $r_m = b_m/d_m < 1$ are more intricate. In particular, if $k_3 > k_2$, then $r_m < 1$ as long as $k_1 < k_2 + k_3$. On the other hand, if $k_3 < k_2$ we need $k_2 < 2\sqrt{k_2 k_3}$. With either of these conditions satisfied the autocatalytic reaction times will be Gumbel distributed. Note that this model has an infinite state space: the number of X particles can be any positive integer. We expect our results to apply to this class of models as long as the initial condition is large. The simulation shown in Figure 3.1(b) indicates this expectation is indeed borne out.

3.9 Appendix

3.9.1 Asymptotic analysis for the Gaussian Universality Class

Cumulant bounds

To estimate the asymptotics of the cumulants we start from Eq. (3.2) derived above. Since the weights w^n are increasing functions of the r_j, we argued in Section 3.3 that

$$\kappa_n(r_{\min}, N) \leq \kappa_n(\{r_j\}, N) \leq \kappa_n(r_{\max}, N), \tag{3.34}$$

where $r_{\max} = \lim_{N\to\infty} \max_{1<j<N} r_j$ and $r_{\min} = \lim_{N\to\infty} \min_{1<j<N} r_j$. The cumulants $\kappa_n(r_{\max}, N)$ and $\kappa_n(r_{\min}, N)$ correspond to auxiliary Markov chains where $b_j + d_j$ is

unchanged, but $r_j = r_{\max}$ or $r_j = r_{\min}$ respectively.

Following Ref. [1], we provide asymptotic bounds on $\kappa(r, N)$ that lead to an analytic criterion for the Gaussian universality class. Since the diagonal elements of the weights w^n are greater than 1, we can bound the cumulant κ_n from below by a sum of the unweighted diagonal elements $(b_i + d_i)^{-n}$. To bound from above we can take the maximum value of $(b_i + d_i)^{-n}$ times the sum of the weighting factors. The sum over weighting factors $w^n_{i_1 i_2 \cdots i_n}(r)$ is precisely the n^{th} cumulant for a biased random walk (with $b_i + d_i = 1$ and uniform r). This sum can be computed exactly using eigenvalues of the transition matrix [1]. In particular, the sum is $O(N)$ for any n as long as $r < 1$ and asymptotically can be represented in the integral form given below. Note that $r_{\max} < 1$ as long as $r_j < 1 - \epsilon$ for all j and some $\epsilon > 0$: this condition was the first requirement for the Gaussian universality class quoted in Section 3.3. Altogether we have,

$$\begin{aligned} \sum_{n=1}^{N} \frac{1}{(b_i + d_i)^n} \leq \kappa_n(r, N) &\leq \left(\max_{1 \leq i \leq N} \frac{1}{b_i + d_i} \right)^n \times \frac{N}{\pi} \int_0^{\pi} \frac{(n-1)!}{(1 - 2\sqrt{r}/(1+r)\cos x)^n} \mathrm{d}x \\ &= \left(\max_{1 \leq i \leq N} \frac{1}{b_i + d_i} \right)^n \times O(N). \end{aligned} \tag{3.35}$$

We can now read off the second condition for the Gaussian universality class. To nail down the asymptotics of $\kappa_n(r, N)$ we want the upper and lower bounds in Eq. (3.35) to have the same scaling for large N. Specifically, we require

$$\frac{1}{N} \sum_{i=1}^{N} \frac{1}{(b_i + d_i)^n} \sim c_n \left(\max_{1 \leq i \leq N} \frac{1}{b_i + d_i} \right)^n, \tag{3.36}$$

for some N-independent constant c_n. Setting $n = 1$ in this equation leads to the condition Eq. (3.17). We can make this simplification because when Eq. (3.36) is satisfied for $n = 1$, it is also satisfied for $n > 1$. To see this fact, first note that $\langle (b_i + d_i)^{-n} \rangle < \max_i (b_i + d_i)^{-n}$ trivially. Furthermore, we can write the left hand side of Eq. (3.36) as $N^{-1} \| (b + d)^{-1} \|_p^p$, where $\| \cdot \|_p$ is the p-norm and $(b + d)^{-1}$ is

the vector containing elements $(b_i + d_i)^{-1}$. Using p-norm inequalities, we have $N^{-1}\|(b+d)^{-1}\|_1 < N^{-1/n}\|(b+d)^{-1}\|_n$ for $n > 1$. Then if $c \cdot \max_i(b_i + d_i)^{-1} < \langle (b_i + d_i)^{-1} \rangle$ as $N \to \infty$ for some constant c, it follows that $c^n \max_i(b_i + d_i)^{-n} < \langle (b_i + d_i)^{-n} \rangle$ in this limit as well. Thus, it is sufficient to check Eq. (3.36) holds for $n = 1$, since this implies the condition holds for all $n > 1$.

As discussed previously, the condition Eq. (3.36) can be interpreted as the waiting times being 'flat' in the following sense: all (or at least a significant fraction) of the $(b_i + d_i)^{-1}$ are the same order as their maximum value. If this condition holds, then for large N we have that $\kappa_n(r, N) \sim c_n(r) f(N)^n N$ where $f(N) \sim \max_{1 \le i \le N}(b_i + d_i)^{-1}$ as $N \to \infty$. Since these asymptotics hold for both $r = r_{\min}$ and $r = r_{\max}$, it follows from Eq. (3.34) that $\kappa_n(\{r_j\}, N) \sim c_n(\{r_j\}) f(N)^n N$ as well, possibly with a different constant $c_n(r_{\min}) < c_n(\{r_j\}) < c_n(r_{\max})$. These asymptotics imply that the higher-order cumulants are dominated by the variance and hence the distribution looks normal for large N, i.e. the standardized cumulants $\tilde{\kappa}_n = \kappa_n / \kappa_2^{n/2} \to 0$ as $N \to \infty$.

Leading correction to the Gaussian

The leading correction to the Gaussian distribution for finite N comes from the skew, $\tilde{\kappa}_3 = \tilde{c}_3 / \sqrt{N}$. Here we will give a bound on the magnitude of the skew, that can be used to predict when finite systems will have a nearly Gaussian absorption-time distribution. First, define

$$K_2 = \lim_{N \to \infty} \frac{1}{N f(N)^2} \sum_{i=1}^{N} \frac{1}{(b_i + d_i)^2}, \tag{3.37}$$

where $f(N) \sim \max_{1 \le i \le N}(b_i + d_i)^{-1}$ as $N \to \infty$ as above. Then from Eq. (3.35) that $\kappa_2 \ge K_2 N f(N)^2$. Evaluating the integral in Eq. (3.35) we have $\kappa_3 \le 2 f(N)^3 N (r_{\max} +$

$1)^3(r_{\max}^2 + 4r_{\max} + 1)/(1 - r_{\max})^5$. Putting these together,

$$\tilde{\kappa}_3 \leq \frac{2(r_{\max} + 1)^3(r_{\max}^2 + 4r_{\max} + 1)}{(1 - r_{\max})^5 K_2^{3/2}} \frac{1}{\sqrt{N}}. \tag{3.38}$$

The convergence is slowest (i.e. the coefficient of $1/\sqrt{N}$ is large), when the conditions for the universality class are pushed to their limits: if the system is barely extinction-prone, $r_{\max} \approx 1$, or the waiting times are not very uniform, $K_2 \ll 1$ (the sum in Eq. (3.36) is nearly dominated by the maximal term). Finally, we note that this is a rough upper bound; in many cases the convergence is much faster, e.g., if only a few $r_j \approx 1$ but the rest are very small. Replacing $r_{\max}$ with the average r_j in Eq. (3.38) may often give a better estimate of the actual skew for a given system, even if it does not give a strict upper bound.

3.9.2 Asymptotic analysis for the Gumbel Universality Class

Cumulant bounds

For the Gumbel universality class we require $b_m + d_m = f(N)m[1 + O(m/N)]$, $b_{\alpha N} + d_{\alpha N}$ be of order at least $O(Nf(N))$ for any $0 < \alpha < 1$, and $r_m = r + O(m/N)$ for large N. These properties are sufficient to guarantee that the absorption-time cumulants are asymptotic to those for an exactly solvable canonical model (for which the above equalities hold exactly, not just to leading order). Following Ref. [1], we restrict two of the indices in Eq. (3.2) to be $O(N)$ away from the absorbing state, $\alpha N \leq i_{n-1} \leq i_n \leq N$. With this restriction we can bound the sums,

$$\sum_{\substack{1 \leq i_1 \leq i_2 \leq \cdots \leq i_{n-1} \\ \alpha N \leq i_{n-1} \leq i_n \leq N}} \frac{w^n_{i_1 i_2 \cdots i_n}(\{r_j\})}{(b_{i_1} + d_{i_1}) \cdots (b_{i_n} + d_{i_n})} \leq \frac{1}{f(N)^n N^2} \sum_{1 \leq i_1 \leq i_2 \leq \cdots \leq i_n \leq N} w^n_{i_1 i_2 \cdots i_n}(r). \tag{3.39}$$

In the previous section, we established that the sum over the weighting factors is $O(N)$, so this portion of the sum is $O(f(N)^{-n} N^{-1})$. We now consider indices

$1 < i_1 < \alpha N$ and $\alpha N < i_n < N$,

$$\sum_{\substack{i_1 \le i_2 \le \cdots \le i_n \\ 1 \le i_1 \le \alpha N \le i_n \le N}} \frac{w^n_{i_1 i_2 \cdots i_n}(\{r_j\})}{(b_{i_1} + d_{i_1}) \cdots (b_{i_n} + d_{i_n})} \le \frac{1}{f(N)^n N} \sum_{\substack{i_1 \le i_2 \le \cdots \le i_n \\ 1 \le i_1 \le \alpha N \le i_n \le N}} w^n_{i_1 i_2 \cdots i_n}(r). \tag{3.40}$$

Since the weighting factors decay exponentially away from the diagonal elements, the sum on the right hand side of Eq. (3.40) is $O(1)$ and this portion of the sum is also $O(f(N)^{-n} N^{-1})$.

Since the same bounds also apply for any other pair of the indices, the only remaining portion of the cumulant sum Eq. (3.2) is that where all indices are near 0. Here the approximations that $b_m + d_m$ is linear and r_m is constant become asymptotically exact so that,

$$\sum_{1 \le i_1 \le i_2 \le \cdots \le i_n \le \alpha N} \frac{w^n_{i_1 i_2 \cdots i_n}(\{r_j\})}{(b_{i_1} + d_{i_1}) \cdots (b_{i_n} + d_{i_n})} \sim \frac{1}{f(N)^n} \sum_{1 \le i_1 \le i_2 \le \cdots \le i_n \le \alpha N} \frac{w^n_{i_1 i_2 \cdots i_n}(r)}{i_1 i_2 \cdots i_n}. \tag{3.41}$$

The right hand side of Eq. (3.41) is at least $O(f(N)^{-n})$ and therefore this region of the cumulant sum dominates asymptotically compared to the $O(f(N)^{-n} N^{-1})$ terms estimated above. In other words, the absorption process is entirely dominated by the coupon collection behavior near the absorbing state. Furthermore, we can freely extend the upper limit of the sum to N (instead of αN) since this will only add subdominant terms. Finally, we obtain the result quoted in Eq. (3.18),

$$\kappa_n(\{r_j\}, N) \sim \frac{1}{f(N)^n} \sum_{1 \le i_1 \le i_2 \le \cdots \le i_n \le N} \frac{w^n_{i_1 i_2 \cdots i_n}(r)}{i_1 i_2 \cdots i_n}. \tag{3.42}$$

Thus, for any Markov chain satisfying the conditions at the beginning of this section, the cumulants are asymptotic to those for the "canonical model" with $b_m + d_m = f(N)m$ and $r_m = r$ exactly. In Section S4.C we show this model has an asymptotically Gumbel absorption-time distribution.

Leading correction to the Gumbel

The leading correction $\delta\kappa_2$ to the standard deviation comes from the quadratic term in the transition rates, $b_m+d_m \approx f(N)m(1+m/N)$. Plugging this into Eq. (3.42) for one set of rates in the denominator and using the partial fraction decomposition $1/i(j+j^2/N) = 1/ij + 1/i(j+N)$ leads to

$$\begin{aligned}\delta\kappa_2 &\sim \frac{1}{f(N)^2}\sum_{1\le i\le j\le \alpha N}\frac{w_{i,j}^2(r)}{i(j+N)} \\ &= \frac{1}{f(N)^2}\sum_{i=1}^{\alpha N}\frac{(1+r)^2(1-r^i)^2}{(1-r)^2 i(i+N)} + 2\sum_{i=1}^{j-1}\sum_{j=1}^{\alpha N}\frac{r^{j-i}(1+r)^2(1-r^i)^2}{(1-r)^2 i(j+N)}\end{aligned} \tag{3.43}$$

where in the second line we make use of the fact that $r_m \approx r$ is approximately constant to write the explicit expression for $w_{i,j}^2$ (obtained from Eqs. (3.7) and (3.14)). The sums in the second line can be evaluated explicitly in terms of special functions, including harmonic numbers and the Lerch transcendent. The first sum is asymptotically dominant, leading to $\delta\kappa_2 \sim f(N)^{-2}N^{-1}\ln(N)$. More generally, the asymptotics above hold if $b_m + d_m \approx f(N)m(1 + m/g(N))$ as long as the function $g(N) \to \infty$ as $N \to \infty$. An analogous calculation shows that for this case $\delta\kappa_2 \sim [\ln g(N)]/g(N)f(N)^2$. If $g(N)$ grows superlinearly, however, $[\ln g(N)]/g(N)f(N)^2$ is dominated by the corrections due to Eq. (3.39) and (3.40) computed above, leading to $\delta\kappa_2 \sim f(N)^{-2}N^{-1}$.

The higher-order cumulants can be analyzed in similar fashion. Since the weights decay exponentially away from the diagonal, the terms with $i_1 = i_2 = \cdots = i_n \equiv i$ is asymptotically dominant. For these elements $w^n_{i_1 i_2\cdots i_n}(\{r_j\}) = (n-1)!V_{ii}^n$ and it is straightforward to show that $\delta\kappa_n \sim N^{-1}f(N)^{-n}$. for the standardized cumulants $\tilde{\kappa}_n = \kappa_n/\kappa_2^{3/2}$, the factors of $f(N)$ in the asymptotics cancel and we are left with $O(N^{-1}\ln N)$ corrections from the standard deviation. In other words, the deviations from the Gumbel cumulants scale like $\delta\tilde{\kappa}_n = \tilde{\kappa}_n - \tilde{\kappa}_n^{\text{Gumbel}} \sim$

$C_n N^{-1} \ln N$ for large N. For the more general case, where the quadratic term in the rates is suppressed by $g(N)$, the scaling is $\delta\tilde{\kappa}_n \sim C_n[\ln g(N)]/g(N)$ for sublinear $g(N)$ and $\kappa_n \sim C_n/N$ otherwise.

Large-N limit for the canonical model

The canonical Markov model with coupon-collection behavior has $b_m + d_m = f(N)m$ and $r_m = r$. Note that $f(N)$ simply sets the time scale for the process and does not affect the shape of the absorption-time distribution. Therefore, for convenience, we will rescale time $t \to t(r+1)/f(N)$ so that $b_m = rm$ and $d_m = m$. For this system, the absorption-time distribution $p(t)$ has been computed exactly using generating functions [59, Appendix B],

$$p(t) = \frac{N e^{\nu t} \nu^2}{(e^{\nu t} - 1)^2 (1 + \frac{\nu}{e^{\nu t} - 1})^{N+1}} \tag{3.44}$$

where $\nu = 1 - r$. To derive the asymptotic form of the distribution we standardize to zero mean and unit variance. The standardized distribution is simply $\sigma p(\sigma t + \mu)$, where $\mu \sim (\ln N + \ln \nu + \gamma)/\nu$ and $\sigma \sim \pi/\nu\sqrt{6}$ are the mean and standard deviation of the absorption time. Here $\gamma \approx 0.5772$ is the Euler-Mascheroni constant. Plugging in this transformation and taking $N \to \infty$, we find

$$\sigma p(\sigma t + \mu) \xrightarrow{N\to\infty} \frac{\pi}{\sqrt{6}} \exp\left(-\gamma - \pi t/\sqrt{6} - e^{-\gamma - \pi t/\sqrt{6}}\right), \tag{3.45}$$

which is precisely the standardized Gumbel distribution.

3.9.3 Asymptotic Analysis for the power-law processes

In this section we generalize the Gumbel criteria discussed above. Consider Markov processes with transition rates that satisfy $b_m + d_m = f(N)m^p[1 + O(m/N)]$.

Moreover, suppose that $b_{\alpha N} + d_{\alpha N}$ is of order at least $O(N^p f(N))$ for any $0 < \alpha < 1$, and $r_m = r + O(m/N)$ for large N. In other words, this process has transition rates that vanish as a power-law m^p near the boundary. In Section 3.5 we claimed that $p \leq 1/2$ gives rise to Gaussian absorption times, while $p > 1/2$ leads to a skewed family of distributions (whose shape depends on the parameters p and r). We rigorously justify these claims in the following subsections.

Skewed family for $p > 1/2$

When $p > 1/2$ the transition rates decay quickly enough that the process is dominated by slowdown near the boundary (similar to coupon collection), giving rise to skewed distributions. To analyze this case, we can apply similar asymptotic analysis to that given in Section S4.A for the Gumbel class. Repeating the bounds in Eqs. (3.39) and (3.40) for the power-law process, we find that

$$\sum_{\substack{1 \leq i_1 \leq i_2 \leq \cdots \leq i_{n-1} \\ \alpha N \leq i_{n-1} \leq i_n \leq N}} \frac{w^n_{i_1 i_2 \cdots i_n}(\{r_j\})}{(b_{i_1} + d_{i_1}) \cdots (b_{i_n} + d_{i_n})} = O(f(N)^{-n} N^{1-2p}) \tag{3.46}$$

$$\sum_{\substack{i_1 \leq i_2 \leq \cdots \leq i_n \\ 1 \leq i_1 \leq \alpha N \leq i_n \leq N}} \frac{w^n_{i_1 i_2 \cdots i_n}(\{r_j\})}{(b_{i_1} + d_{i_1}) \cdots (b_{i_n} + d_{i_n})} = O(f(N)^{-n} N^{-p}). \tag{3.47}$$

As long as $p > 1/2$, these terms are each asymptotically dominated by the indices near 0,

$$\sum_{1 \leq i_1 \leq i_2 \leq \cdots \leq i_n \leq \alpha N} \frac{w^n_{i_1 i_2 \cdots i_n}(\{r_j\})}{(b_{i_1} + d_{i_1}) \cdots (b_{i_n} + d_{i_n})} \sim \frac{1}{f(N)^n} \sum_{1 \leq i_1 \leq i_2 \leq \cdots \leq i_n \leq \alpha N} \frac{w^n_{i_1 i_2 \cdots i_n}(r)}{i_1^p i_2^p \cdots i_n^p}, \tag{3.48}$$

which are at least of order $O(f(N)^{-n})$. Similar to the Gumbel class, the absorption process is dominated by the slow behavior near the absorbing state, where the transition rates decay. Extending the upper limit on the sum from αN to N (which only adds subdominant terms), we find that the cumulants κ_n satisfy

$$\kappa_n(\{r_j\}, N) \sim \frac{1}{f(N)^n} \sum_{1 \leq i_1 \leq i_2 \leq \cdots \leq i_n \leq N} \frac{w^n_{i_1 i_2 \cdots i_n}(r)}{i_1^p i_2^p \cdots i_n^p}. \tag{3.49}$$

Notice that this asymptotic formula for the cumulants is identical to Eq. (3.42), but with the denominator $(i_1 i_2 \cdots i_m)$ raised to the power p. Thus, we have shown the absorption-time cumulants for a general Markov process, with $b_m + d_m = f(N)m^p[1 + O(m/N)]$, are asymptotic to those for the minimal model $b_m = rd_m = rm^p$ (after rescaling time so that $f(N) = r + 1$). The absorption-time distributions for the minimal model were explored numerically in Figures 3.1(d) and 3.2. For $p > 1/2$, we find a family of distributions that become more skewed as p increases. The shape of the distributions depends subtly on r except when $p = 1$, where the distribution is Gumbel, as revealed by our analysis above.

Gaussian distributions for $p \leq 1/2$

To show the normality of the absorption-time distribution for $p \leq 1/2$, we show that the variance κ_2 diverges at least as fast as the higher-order cumulants. Using the asymptotic estimate from the previous section, we have that

$$\kappa_n(\{r_j\}, N) \sim \frac{1}{f(N)^n} \sum_{1 \leq i_1 \leq i_2 \leq \cdots \leq i_n \leq N} \frac{w^n_{i_1 i_2 \cdots i_n}(r)}{i_1^p i_2^p \cdots i_n^p} + O(f(N)^{-n} N^{1-2p}). \tag{3.50}$$

As noted above, when $p < 1/2$, we can not guarantee that the first term is dominant. If the second term is dominant than the cumulants scale like $\kappa_n \sim c_n f(N)^{-n} N^{1-2p}$. Then the standardized cumulants asymptotically vanish, $\tilde{\kappa}_n = \kappa_n / \kappa_2^{n/2} \propto N^{(2-n)(1-2p)/2} \to 0$ as $N \to \infty$. Hence, the absorption-time distribution is asymptotically normal. On the other hand, if the first term in Eq. (3.50) is dominant, we can show the distribution is still Gaussian. Since the weight factors fall off exponentially away from the diagonal, the diagonal terms are asymptotically dominant. Using the fact that $w^n_{i_1 i_2 \cdots i_n}(\{r_j\}) = (n-1)! V^n_{ii} \geq 1$, when

$i_1 = i_2 = \cdots = i_n \equiv i$ together with Eq. (3.7), we have

$$\kappa_n(\{r_j\}, N) \sim \frac{(1+r)^n}{(1-r)^n f(N)^n} \sum_{i=1}^{N} \frac{(1-r^i)^n}{i^{np}}. \tag{3.51}$$

Notice that when $p = 1/2$, the sum in this expression diverges as $N \to \infty$ for the variance ($n = 2$), but converges for the higher-order cumulants ($n > 2$). More generally, for any $p \leq 1/2$, it is straightforward to show that the sum in Eq. (3.51) always diverges faster with N for the variance than for the higher-order cumulants. As above, this scaling leads to $\kappa_n/\kappa_2^{n/2} \to 0$ for large N, so that the distribution asymptotically approaches a Gaussian. Figures 3.1(d) and 3.2 show this result is confirmed in numerical simulations: the distribution for $p \leq 1/2$ looks approximately normal and the skew approaches 0.

CHAPTER 4

REACTION RATES AND THE NOISY SADDLE-NODE BIFURCATION: RENORMALIZATION GROUP FOR BARRIER CROSSING

4.1 Introduction[1]

In this chapter, we investigate deep connections between barrier crossing, the renormalization group, and the noisy saddle-node bifurcation. In particular, we show that Kramers' reaction rates can be understood as an asymptotic limit of the universal scaling near the continuous transition between high-barrier and barrier-less regimes. Applying methods from stochastic processes theory we derive an analytical expression for the universal scaling function for the mean barrier escape time near the critical point, giving the crossover between high and low barrier limits. The renormalization group provides a framework within which this result can be understood and systematically improved by perturbative calculations of corrections to scaling, some of which we give explicitly.

Barrier crossing arises in applications across physics, chemistry, and biology. In 1940, Kramers computed the barrier crossing rate for particles in both overdamped and underdamped regimes [70]. This result and others [71–73] provided the theoretical explanation for the Arrhenius equation describing chemical rate coefficients $k \sim \exp(-E_b/k_BT)$, where E_b is the energy barrier for activation [74]. More recent efforts have established the escape rate at arbitrary damping, giving the crossover between the low- and high-damping limits [75, 76], and have accounted for the effects of state-dependent [77, 78], non-Gaussian

[1] This chapter is reproduced from: David Hathcock and James P. Sethna, "Reaction rates and the noisy saddle-node bifurcation: renormalization group for barrier crossing." *Physical Review Research* **3**, 013156 (2021).

[79–81], and colored [81–83] noise, anharmonic corrections [84, 85], and fluctuating barriers [83, 86].

Most transition-state calculations assume a large barrier limit. This means the barrier escape is a rare event, with a separation of time scales between the relaxation into a quasi-equilibrium state and the escape from that state [5]. In the limit of vanishing barrier, however, there is a qualitative change in behavior. Particles instead slide down a monotonic potential, spending the most time near its inflection point. To capture the low barrier escape rate, extensions to Kramers' theory have been developed (e.g., incorporating anharmonic corrections), but these have significant errors when the barrier and thermal energy are comparable ($E_b \approx k_B T$) [84].

Finite barrier escape problems have garnered increasing theoretical interest over the past decade, with several studies contributing further low barrier refinements of existing theories [87, 88, 88–92] or focusing directly on the saddle-node bifurcation where the barrier vanishes [93, 94]. Such escape processes are relevant to certain high precision measurements. For instance, force spectroscopy experiments apply a force on a single bond in a biomolecule until it breaks [88, 95]. For typical molecules, the critical force, at which the energy barrier for breaking vanishes and Kramers' theory breaks down, is now well within the reach of atomic force microscopy and optical tweezers [95]. Another exciting application is in micro- and nano-electromechanical devices, which sensitively switch oscillation amplitude in response to an input signal by operating near the barrier-less critical point [93, 96]. Here, an analytical theory of low barrier crossing would help to distinguish between noise and signal activated switching.

We develop a critical theory for barrier crossing with a renormalization group approach that gives a complete scaling description of the noisy saddle-node bifurcation. We are inspired by previous work on the 'intermittency' [2] route to chaos [97–99], where the renormalization group coarse-grains in time, then rescales the system to fix a certain term in the potential. In chaos theory, this procedure involves iterating and rescaling a discrete map [98, 99], leading to a different fixed point for the same renormalization group equations used by Feigenbaum to study period doubling [8]. We take the continuous time limit, reducing the renormalization group to a series of elementary rescalings and yielding a simplified description applicable to barrier escape problems. Our procedure organizes what amounts to dimensional analysis, providing an elegant renormalization-group framework that unifies Kramers' theory for Arrhenius barrier crossing with the dynamical systems theory of a noisy saddle-node bifurcation.

Why do we frame our analysis in terms of the renormalization group, if the scaling form can be justified using dimensional analysis and the analytical methods we use are drawn from more traditional stochastic analysis? On the one hand, this forms a wonderful case study, unifying and illuminating bifurcation theory, the renormalization group, and chemical reaction theory. Second, barrier crossing forms the solvable limiting case of much more complex phenomena: coupling to colored-noise heat baths, nucleation of abrupt phase transitions, and depinning transitions in disordered systems (see Section 4.5). The scale invariance of random walks does not demand a renormalization group proof of the central limit theorem, and the one-dimensional Ising model can be

[2] In this chapter 'intermittency' refers to the chaotic dynamical intermittency studied in Refs. [97–99] that emerges in discrete maps via a tangency bifurcation. We use this term sparingly to avoid confusion with intermittency in fluid dynamics and other areas.

solved without the machinery of flows in Hamiltonian space. But framing the problems in terms of the renormalization group provide excellent pedagogical exercises, and the natural framework for extensions to self-avoiding random walks and Ising models in higher dimensions.

Using our scaling theory as an organizing framework, we derive an analytical expression for the scaling form of the mean escape time near the saddle-node bifurcation. We also compute corrections to scaling due to anharmonicity in the potential and finite initial or final position. Going beyond the mean, we develop an accurate approximation to the scaling form for the full distribution of escape times by assuming the eigenvalues of the Fokker-Plank operator are equally spaced.

As a starting point, we consider the equation of motion for an overdamped particle in a general potential $V(x)$ and driven by spatially dependent white noise,

$$\dot{x} = f(x) + g(x)\,\xi(t). \tag{4.1}$$

Here $f(x) = -\eta^{-1}\,dV/dx$ is the force exerted on the particle (divided by the damping coefficient η) and $g(x)$ is the spatially varying noise amplitude (with the damping absorbed). The noise $\xi(t)$ has zero mean, $\langle \xi(t) \rangle = 0$ and is uncorrelated in time, $\langle \xi(t)\xi(t') \rangle = \delta(t - t')$. With barrier crossing phenomena in mind, we consider potentials with boundary conditions $V(x) \to \infty$ as $x \to -\infty$ and $V(x) \to -\infty$ as $x \to \infty$. The potential either has a single barrier or is monotonically decreasing (e.g., Figure 4.1). The quantity of interest is the mean barrier crossing time τ, defined as the time particles take to reach $+\infty$ from an initial position at $-\infty$.

Besides the experimental systems discussed above, this model also serves as

the natural description for a general chemical reaction, involving the transition between metastable species A and B. These species are points in a $3N$ dimensional configuration space defined by the locations of N reaction constituents. As derived by Hänggi *et al.* this system can be reduced to the one-dimensional model we study [5]. They coarse-grain to a one dimensional reaction coordinate, which parametrizes the minimal gradient path between the states A and B, neglecting effects of memory friction and noise correlations, and taking the overdamped limit produces Eq. (4.1). The effective potential along the reaction coordinate has a barrier separating species A and B.

4.2 Renormalization Group and Scaling Theory

To begin, we parameterize Eq. (4.1) by the Taylor coefficients of $g(x)$ and $f(x)$,

$$\frac{dx}{dt} = \sum_{n=0}^{\infty} \epsilon_n x^n + \xi(t) \sum_{n=0}^{\infty} g_n x^n. \tag{4.2}$$

The renormalization group defines a flow in this space of systems described by a single reaction coordinate x. Near the renormalization group fixed point, the behavior is most effectively described by a single Taylor expansion at the origin. In contrast, for large barriers in Kramers' theory, the escape time is characterized by two expansions, capturing the harmonic oscillations in the potential well and at the top of the barrier. These two equivalent schemes are shown in Figure 4.1. Given the later expansion at the two extrema, the expansion at the origin can be reconstructed via a two-point Padé approximation [100].

As discussed above, the discrete renormalization group coarse-grains by iterating a map, evolving the equations forward in time. Ignoring the noise for the moment, we consider a discrete approximation to Eq. (4.1), $x_{n+1} = x_n +$

$\delta t f(x_n) \equiv h(x_n)$. References [97–99] study this discrete equation using the Feigenbaum renormalization group transformation. This transformation iterates the map and rescales space, inducing a flow in function space, $T[h](x) = ah(h(x/a))$, where a is the rescaling factor. The renormalization fixed point is the function $h(x)$ that obeys $T[h] = h$ for a particular rescaling a. We simplify this calculation by taking a continuum limit. Expanding the renormalization group to first order in δt, $T[h](x) = x + 2\delta t(af(x/a))$. Thus, in the continuous time limit, the RG iteration becomes a simple rescaling of time and space. We will use this below to derive our scaling theory for barrier crossing near a saddle-node bifurcation.

Within the context of the renormalization group for singular perturbations developed by Goldenfeld, Oono, and others [101, 102], our problem can be understood as having zero anomalous dimension. For our calculations in the following sections we do not need to use the traditional renormalization-group machinery; instead we assemble a variety of tools from probability theory and Markov processes, to express the mean and distributions of escape times as universal scaling functions near the transition where the barrier vanishes. The scaling theory provides a powerful and elegant structure which organizes our understanding barrier crossing.

Following the above expansion of the discrete renormalization group, we 'coarse grain' the system in time by scaling, $\hat{t} = t/b$. As the time-scale shrinks, the noise is amplified, $\hat{\xi}(\hat{t}) = b^{1/2}\xi(t)$ (the exponent $1/2$ follows from the units of the correlation function). Our goal is to understand the scaling properties near the critical point, where a qualitative change in behavior occurs. For a generic analytic potential this happens when the barrier vanishes and $V(x) = -x^3$ is locally a perfect cubic. Therefore, we rescale our system to fix the coefficient ϵ_2,

corresponding to the cubic term in the potential. The correct rescaling defines a new spatial coordinate $\hat{x} = bx$. After both coarse graining and rescaling, we arrive at

$$\frac{d\hat{x}}{d\hat{t}} = \sum_{n=0}^{\infty} \epsilon_n b^{2-n} \hat{x}^n + \hat{\xi}(\hat{t}) \sum_{n=0}^{\infty} b^{3/2-n} g_n \hat{x}^n. \tag{4.3}$$

We can then read off how the parameters flow under the renormalization group, $\hat{\epsilon}_n = b^{2-n}\epsilon_n$ and $\hat{g}_n = b^{3/2-n} g_n$. These flows and exponents exactly match those found under the discrete-time renormalization group [98, 99], indicating that the scaling of the 'intermittency route to chaos' is also non-anomalous [102]. Taking the coarse graining factor to be close to 1, $b = (1 + d\ell)$, we obtain continuous flow equations,

$$\frac{d\epsilon_n}{d\ell} = (2 - n)\epsilon_n, \quad \frac{dg_n}{d\ell} = (3/2 - n) g_n. \tag{4.4}$$

The eigenfunctions of the renormalization group in our continuum theory are the monomials x^n and noisy monomials $\xi(t)x^n$. If the right hand side of Eq. (4.2) is an eigenfunction, it is scaled by a constant factor under the action of the renormalization group. These eigenfunctions are the much simpler continuous time limit of those for the discrete-time renormalization group [99]. In particular, the cubic potential $V(x) \propto -x^3$ (without noise) is the fixed point. At the fixed point, particle trajectories $x(t) \sim 1/t$ exhibit scale invariance in time as they approach the cubic inflection point at $x = 0$. Perturbations away from the fixed point lead to dynamics with non-power law decay to a locally stable state or over the inflection point.

The mean barrier crossing time is a function of the potential shape and the noise correlation, encoded through the expansion coefficients ϵ_n and g_n. Thus, the escape time can be expressed as $\tau(\{\epsilon_n\}, \{g_n\})$, where $n \geq 0$. If we coarse-grain

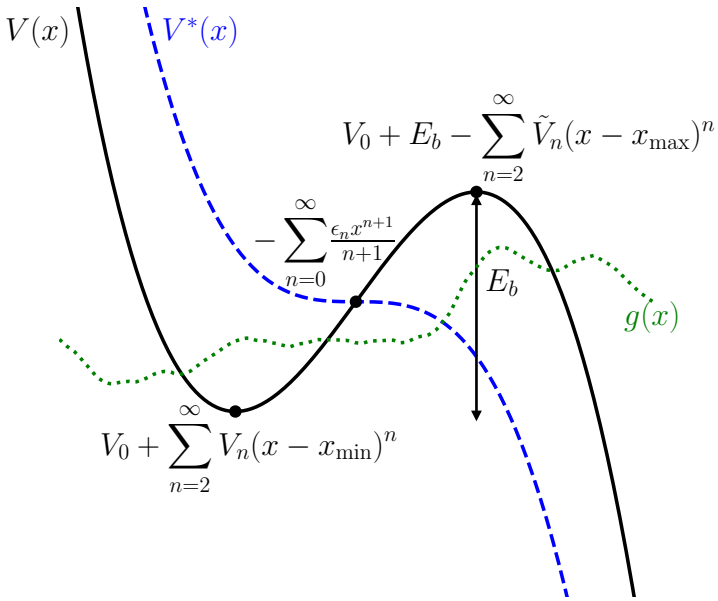

Figure 4.1: Typical potentials in the high barrier Arrhenius limit (solid curve) and at the renormalization group fixed point (dashed curve). Kramers' theory utilizes a two point series expansion at $x_{\min}$ in the potential well and at $x_{\max}$, the top of the barrier. For our renormalization group approach the natural description is in terms of a single expansion at the origin parameterizing perturbations away from the fixed point potential $V^*(x) \propto -x^3$. Also shown is the noise amplitude g(x), which generically has spatial dependence (dotted curve).

until $g_0(\ell^*) = 1$, we find that the escape time has the form

$$\tau = g_0^{-2/3} \mathcal{T} \left(\{\epsilon_n / g_0^{2(2-n)/3}\}, \{g_n / g_0^{1-2n/3}\} \right), \tag{4.5}$$

where $\mathcal{T}$ is a universal scaling function, with $n \geq 1$ for the second term in brackets.

While the scaling form Eq. (4.5) could have been written down using dimensional analysis, the renormalization group approach provides the natural structure and motivation for our approach. The parameter space flows indicate that, with a fixed quadratic force, the constant and linear force and noise terms $\{\epsilon_0, \epsilon_1, g_0, g_1\}$ are relevant, growing under coarse graining and dominant on long time scales. Other variables are irrelevant and can be incorporated perturba-

tively. Of the relevant variables, the linear force coefficient ϵ_1 can be set to zero by placing the origin at the inflection point of the potential. The spatial dependence of the noise (including the relevant linear term g_1) can also be removed by a change of coordinates $x \to \tilde{x}$ with $\tilde{x}$ defined by [103]

$$x = \int^{\tilde{x}} \frac{g_0}{g(y)}\, dy, \tag{4.6}$$

producing a system with constant noise $\tilde{g}(\tilde{x}) = g_0$ and force $\tilde{f}(\tilde{x}) = f(\tilde{x})/g(\tilde{x})$ (hence g_1 was relevant before we removed it, because it contributes to the linear term in the expansion of $\tilde{f}$).

Systems near enough to the critical point therefore can be modeled as a cubic potential with a linear perturbation $V(x) = -x^3/3 - \epsilon_0 x$ and constant noise g_0. This is the 'normal form' used in bifurcation theory for the saddle-node transition, and might have been anticipated from Taylor's theorem. The escape time scaling form becomes,

$$\tau = g_0^{-2/3} \mathcal{T}(\epsilon_0 / g_0^{4/3}). \tag{4.7}$$

Thus, the problem asymptotically reduces to finding the universal function of a single variable $\mathcal{T}(\alpha)$, where $\alpha = \epsilon_0/g_0^{4/3}$. The limiting form of the scaling function $\mathcal{T}(\alpha)$ must give the known solutions. In the limit $\alpha \to -\infty$ the barrier is large compared to the noise, so the Kramers approximation [70] applies, so

$$\mathcal{T}(\alpha) \sim \frac{\pi}{|\alpha|^{1/2}} e^{\frac{8}{3}|\alpha|^{3/2}}, \quad \alpha \to -\infty. \tag{4.8}$$

For our choice of parameters, the energy barrier is given by $E_b/k_BT = 8/3|\alpha|^{3/2}$. In the opposite limit $\alpha \to \infty$, the potential is downward sloping with gradient much larger than the noise level. The passage of particles over the inflection point occurs even in the absence of noise (in contrast to the Kramers limit, which requires noise for barrier escape). Therefore, the crossing time approaches that

for a deterministic particle in the cubic potential. One can easily show that the limiting scaling form is

$$\mathcal{T}(\alpha) \sim \frac{\pi}{\alpha^{1/2}}, \quad \alpha \to \infty. \tag{4.9}$$

4.3 Mean Escape Time

4.3.1 Analytical escape time for relevant variables

We now turn our focus to obtaining an exact analytical expression for $\mathcal{T}(\alpha)$ that is valid for all α. To this end, we study the trajectories of particles injected at position x_i and time t_i into a general potential $V(x)$ with noise g_0 and compute the mean first-passage time to x_f, following the standard approach [5, 97, 104]. Let $\mathcal{P}(x,t)$ be the distribution of particles over positions x at time t, with $\mathcal{P}(x,t_i) = \delta(x - x_i)$. The probability that a particle has not reached x_f at time t is

$$\mathcal{P}(t) = \int_{-\infty}^{x_f} \mathcal{P}(x,t)\, dx. \tag{4.10}$$

Note that $\mathcal{P}(0) = 1$ and $\mathcal{P}(t) \to 0$ as $t \to \infty$ as long as there is noise driving the system, which guarantees particles reach x_f. The distribution of first-passage times is $p(t) = -d\mathcal{P}/dt$ so that the mean first-passage time is

$$\tau(x_i|x_f) = \int_0^\infty t\, p(t)\, dt = \int_0^\infty \mathcal{P}(t)\, dt, \tag{4.11}$$

where we integrate by parts for the second equality. To derive a differential equation for $\tau(x_i|x_f)$, we start from the Kolmogorov backward equation for distribution $\mathcal{P}(x,t)$ with initial condition x_i [105],

$$-\frac{d\mathcal{P}(x,t)}{dt_i} = -V'(x_i)\frac{d\mathcal{P}(x,t)}{dx_i} + \frac{1}{2}g_0^2\frac{d^2\mathcal{P}(x,t)}{dx_i^2}. \tag{4.12}$$

To write this equation in terms of the mean first-passage time τ, we multiply both sides by t and integrate over x and t. Using the relations in Eqs. (4.10) and (4.11) and the identity $d\mathcal{P}(x,t)/dt_i = -d\mathcal{P}(x,t)/dt$, we arrive at

$$\frac{1}{2}g_0^2\tau''(x_i|x_f) - V'(x_i)\tau'(x_i|x_f) = -1. \tag{4.13}$$

This gives an ordinary differential equation for the first-passage time from x_i to x_f of particles in potential $V(x)$ and constant noise with amplitude g_0. The boundary conditions are $\tau(x_f|x_f) = 0$ and $\tau'(-\infty|x_f) = 0$, which encode absorbing and reflecting boundaries respectively. Writing the solution to Eq. (4.13) in integral form, we arrive at the result obtained in Refs. [5, 97, 104],

$$\tau(x_i|x_f) = \frac{2}{g_0^2}\int_{x_i}^{x_f} dy \int_{-\infty}^{y} dz\, e^{-\frac{2}{g_0^2}[V(z)-V(y)]}, \tag{4.14}$$

which satisfies the boundary conditions as long as $V'(x) \to \infty$ as $x \to -\infty$. For large barriers, it is known that Eq. (4.14) reproduces Kramers escape rate formula via a saddle point approximation that expands the potential around the maximum and the minimum (as shown in Figure 4.1) to second order [5].

Our renormalization group analysis allows us to restrict our focus to the relevant variables. For the cubic potential (systems on the unstable manifold of the renormalization group fixed point), the escape time can be computed analytically using Eq. (4.14) in the limit $x_f = -x_i \to \infty$. We find that $\tau = g_0^{-2/3}\mathcal{T}(\alpha)$ with the universal scaling function given by

$$\mathcal{T}(\alpha) = 2^{1/3}\pi^2\left[\mathrm{Ai}^2(-2^{2/3}\alpha) + \mathrm{Bi}^2(-2^{2/3}\alpha)\right], \tag{4.15}$$

where $\mathrm{Ai}(x)$ and $\mathrm{Bi}(x)$ are the first and second Airy functions and $\alpha = \epsilon_0/g_0^{4/3}$ as above. This solution is shown in Figure 4.2, along with the Arrhenius and deterministic limits given in Eqs. (4.8) and (4.9) respectively and the mean barrier crossing times from direct simulations of the Langevin process [Eq. (4.1)]. The

universal scaling function $\mathcal{T}(\alpha)$ reproduces the two known limits when the barrier is large or the potential is strongly downward sloping and agrees excellently with the numerical results.

Kramers' escape rate for the cubic potential follows from Eq. (4.15) and the asymptotic form of the second Airy function. As $\alpha \to 0$, however, contributions from the first Airy function become important so that Kramers' theory and extensions involving anharmonic corrections break down. The difference between Eqs. (4.8) and (4.15) is also related to the narrowing of the spectral gap of the barrier crossing Fokker-Plank operator (which has been measured numerically [106] and is discussed below in Section 4.4).

4.3.2 Corrections to scaling

Finite launching and absorbing positions

One limitation of our result Eq. (4.15) is that we assume initial and final states at infinity. In a real chemical or mechanical system, the transition of interest generally occurs between states with finite coordinates. For systems with large barriers, the escape time is exponentially large compared to the time it takes to settle into a metastable state in the well. Thus, the scaling function Eq. (4.15) is universal: independent of initial and final conditions. When the potential is downward sloping, the barrier crossing time is still dominated by the time spent near the inflection point and we can systematically compute corrections to scaling due to the finite initial and final positions.

Working in terms of scaling variables $\alpha = \epsilon_/ g_0^{4/3}$, $\chi_- = x_i/g_0^{2/3}$, and $\chi_+ = x_f/g_0^{2/3}$,

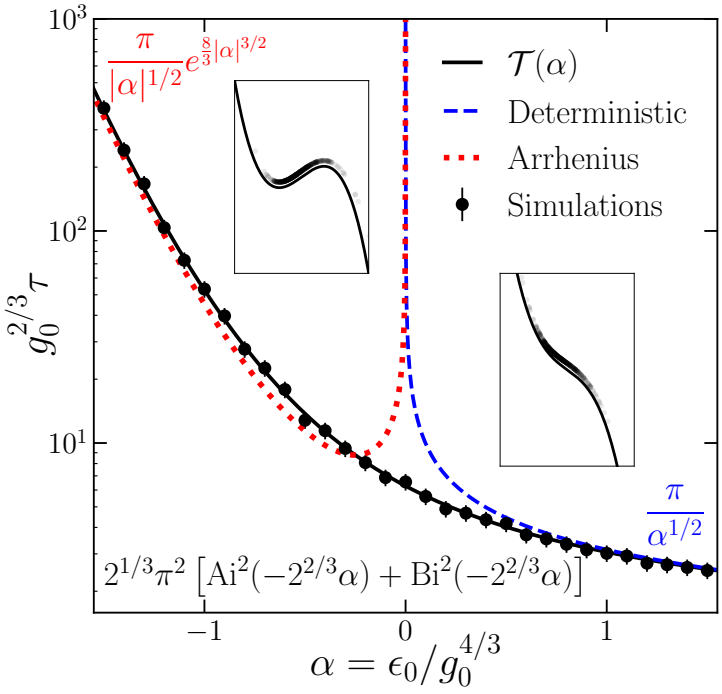

Figure 4.2: Comparison of the universal scaling function $\mathcal{T}(\alpha)$ (solid curve) to the Arrhenius (dotted curve) and deterministic (dashed curve) limits. Also shown are the mean escape times for 500 simulations of the barrier escape process. For the simulations we fixed $g_0 = 1$ while varying ϵ_0 and used boundary conditions $x_f = -x_i = 25$. Agreement with our analytic expression for $\mathcal{T}(\alpha)$ is excellent. The insets show snapshots of the barrier crossing simulations for $\epsilon_0 = \pm 1$.

we can write the barrier crossing time for the cubic potential with arbitrary initial and final conditions as $\tau = g_0^{-2/3}\mathcal{T}(\alpha, \chi_-, \chi_+)$ with the scaling function,

$$\mathcal{T}(\alpha, \chi_-, \chi_+) = \mathcal{T}(\alpha) - \mathcal{T}_-(\alpha, \chi_-) - \mathcal{T}_+(\alpha, \chi_+). \tag{4.16}$$

Here $\mathcal{T}_\pm(\alpha, \chi_\pm)$ are the universal corrections for finite final and initial conditions respectively. These have integral representations,

$$\mathcal{T}_-(\alpha, \chi_-) = 2\int_{-\infty}^{\chi_-} dy \int_{-\infty}^{y} dz\, e^{-2\left(y^3/3+\alpha y - z^3/3 - \alpha z\right)} \tag{4.17}$$

and

$$\mathcal{T}_+(\alpha, \chi_+) = 2\int_{\chi_+}^{\infty} dy \int_{-\infty}^{y} dz\, e^{-2\left(y^3/3+\alpha y - z^3/3 - \alpha z\right)}. \tag{4.18}$$

Assuming $\chi_- < 0$ and $\chi_+ > 0$ these can be expanded in powers of $1/\chi_\pm$,

$$\begin{aligned}\mathcal{T}_\pm(\alpha,\chi_\pm) =& |\alpha|^{-1/2}\tan^{-1}(|\alpha|^{1/2}/|\chi_\pm|) \\ &+ \frac{1}{\chi_\pm^4}\left(\frac{1}{4} - \frac{\alpha}{2\chi_\pm^2} + \frac{3\alpha^2}{4\chi_\pm^4}\right) \\ &\pm \frac{1}{\chi_\pm^7}\left(\frac{5}{14} - \frac{13\alpha}{9\chi_\pm^2}\right) + O(\chi_\pm^{-10}).\end{aligned} \tag{4.19}$$

The first term, which we can compute exactly, is simply the correction due to the deterministic trajectory in the cubic potential between $\chi_\pm$ and $\pm\infty$. The higher order terms describe the influence of noise on the corrections due to finite initial conditions. These terms appear to have the form $\chi_\pm^{-1-3j} f_j(\alpha/\chi_\pm^2)$ for integers j and some functions f_j. The expansions of f_1 and f_2 are given in parenthesis in the expression above.

Our corrections to scaling for finite launching and absorbing positions are universal if the initial conditions are sufficiently close to the cubic inflection point so that anharmonic corrections are small. Thus, for an arbitrary potential, the universal escape time near a saddle-node bifurcation is the given by Eq. (4.15) corrected using Eq. (4.19). In the following section we will separately treat the anharmonic corrections to scaling. The interplay between corrections to scaling due to anharmonicity and initial conditions will be an interesting subject for future studies.

Anharmonic corrections

The scaling function Eq. (4.15) also serves as a starting point from which the theory can be systematically improved by computing anharmonic corrections to scaling. The higher order terms in the potential are irrelevant variables under the renormalization group flows and hence can be treated perturbatively. For

instance, consider a quartic perturbation $\delta V(x) = -\epsilon_3 x^4/4$ and let $\beta = \epsilon_3 g_0^{2/3}$. In the Kramers regime, $\alpha \to -\infty$, we have that $\mathcal{T}(\alpha,\beta) \approx \mathcal{T}(\alpha) + \beta^2 \mathcal{T}_3(\alpha)$ to leading order, where

$$\mathcal{T}_3(\alpha) \xrightarrow{\alpha \ll 0} \pi \sqrt{|\alpha|} e^{\frac{8}{3}|\alpha|^{3/2}} (8|\alpha|^{3/2} + 11)/8. \tag{4.20}$$

In the deterministic regime, $\alpha \to \infty$, we also add a quintic term as a regulator on the boundary conditions of the potential, $\delta V(x) = -\epsilon_3 x^4/4 - \epsilon_4 x^5/5$, with $\epsilon_4 > 0$ and sufficiently large so that the potential remains monotonically decreasing. To quadratic order in β and $\gamma = \epsilon_4 g_0^{4/3}$, the universal scaling function is [3]

$$\begin{aligned} \mathcal{T}(\alpha,\beta,\gamma) \quad &\xrightarrow{\alpha \gg 0} \frac{\pi}{\sqrt{\alpha}} - \beta^2 \left(\frac{15}{8} \pi \sqrt{\alpha} - \frac{3\pi}{4\sqrt{\gamma}} \right) \\ &-\pi \sqrt{\gamma} + \frac{3}{2} \pi \sqrt{\alpha} \gamma - \frac{5}{2} \pi \alpha \gamma^{3/2} + \frac{35}{8} \pi \alpha^{3/2} \gamma^2. \end{aligned} \tag{4.21}$$

The term $\pi/\sqrt{\alpha}$ is just the deterministic limit of the scaling form for the cubic potential and $\beta^2 \mathcal{T}_3(\alpha) = -15\pi \sqrt{\alpha} \beta^2/8$ comes from the quartic perturbation to the inflection point. Other terms arise from quintic corrections or global changes in the potential. Here γ is a dangerous irrelevant variable [107, Sections 3.6, 5.4, & 5.6], which has a pole $3\pi\beta^2/4\sqrt{\gamma}$ in the expansion about 0, because it is needed to keep the potential monotonic (for $\beta \neq 0$).

4.4 Approximating the distribution of escape times

To completely characterize the escape times, we require their distribution, which captures the full range of outcomes we might expect from the stochastic dynamics of the Langevin equation. For high barriers (the Arrhenius limit) the escape is dominated by the decay of a single (quasi-equilibrium) mode in

[3] Here and in Eq. 4.20 we could have avoided branch cuts in our scaling functions by using the variables $\sqrt{\alpha}$, $\sqrt{\beta}$, and $\sqrt{\gamma}$. We choose our convention to avoid complex valued scaling variables and imaginary numbers in Eq. 4.20.

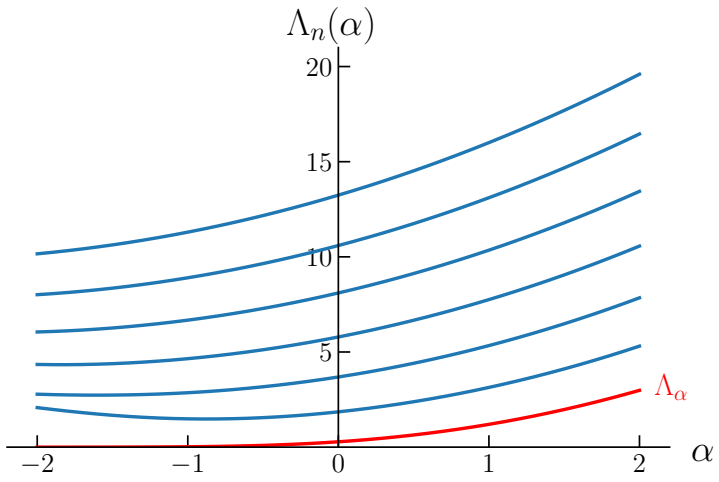

Figure 4.3: The scaling forms $\Lambda_n(\alpha)$ for the first seven eigenvalues. For large positive α the eigenvalues are approximately evenly spaced. For large negative α the leading eigenvalue approaches 0 and the gap to the second eigenvalue grows. The scaling form for the leading eigenvalue $\Lambda_0(\alpha) \equiv \Lambda_\alpha$ used in our approximation to the distribution of escape times is shown in red.

the bottom of the potential well, so the distribution is exponential with rate parameter given by $1/\tau$. On the other hand, for small barriers or sloping potentials many modes contribute and the mean may not be representative of the escape times in general. In this section, we develop an approximation to the distribution of escape times which is accurate for all α (i.e. any cubic potential). The approximate distribution is given as an analytical scaling form parameterized by two variables which are computed numerically for a given α.

4.4.1 First-passage distributions in Markov processes

To study the distribution, we borrow a result from the theory of Markov processes. Consider a birth-death process, i.e. a Markov chain with a one dimensional state space and hopping only between nearest neighbor states, that has one reflecting boundary and one absorbing boundary. This is precisely what we

would obtain by discretizing space in the Langevin equation Eq. (4.1), which has a one dimensional state space with a reflecting boundary for large negative x values and an absorbing boundary for large positive x values. For Markov systems initialized near the reflecting boundary, the distribution of times to reach the absorbing state can be characterized completely in terms of the eigenvalues of the transition matrix. In particular the distribution, $p(t) = -d\mathcal{P}(t)/dt$ (using the notation from the preceding section) can be shown to be a convolution of exponentials [29],

$$p(t) = \mathcal{E}(\lambda_0) * \mathcal{E}(\lambda_1) * \cdots * \mathcal{E}(\lambda_N), \tag{4.22}$$

where λ_n are the negative eigenvalues of the Markov transition matrix, $\mathcal{E}(\lambda_n)$ are exponential distributions, and $*$ denotes a convolution. As we will see below, it is useful consider the Fourier transform $\tilde{p}(\omega)$ of the distribution $p(t)$ and the cumulant generating function

$$\log \tilde{p}(\omega) = \sum_{m=1}^{\infty} \kappa_m (i\omega)^m / m! \tag{4.23}$$

since the cumulants κ_m are easily expressed in terms of the eigenvalues of the Markov matrix,

$$\kappa_m = (m-1)! \sum_{n=0}^{N} \frac{1}{\lambda_n^m}. \tag{4.24}$$

Note that the first cumulant κ_1 is just the mean barrier crossing time τ, which was the focus of the previous section.

4.4.2 Spectra of the Fokker-Planck operator

When the state space becomes continuous, the process is generated by the Fokker-Planck operator rather than a finite transition matrix. Thus we want

to understand the spectra of the right hand side of the Fokker-Planck equation,

$$\frac{\partial \mathcal{P}(x,t)}{\partial t} = \frac{\partial}{\partial x}\left(V'(x)\mathcal{P}(x,t)\right) + \frac{1}{2}g_0^2\frac{\partial^2 \mathcal{P}(x,t)}{\partial x}. \tag{4.25}$$

For numerical evaluation of the eigenvalues is convenient to change variables so that the differential operator is Hermitian. To do this we rescale $\mathcal{P}(x,t)$ by the square root of the Boltzmann factor, defining $\sigma(x,t) = \mathcal{P}(x,t)/\exp(-V(x)/g_0^2)$. Then $\sigma(x,t)$ satisfies

$$\frac{\partial \sigma(x,t)}{\partial t} = \left\{\frac{g_0^2}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}V''(x) - \frac{1}{2g_0^2}\Big(V'(x)\Big)^2\right\}\sigma(x,t), \tag{4.26}$$

which has the form of a Schrödinger equation with imaginary time. Importantly, even in the limit of continuous space, the spectrum of the Fokker-Plank operator is discrete. This is easy to see from Eq. (4.26), because the 'effective quantum potential' $(V'(x))^2/2g_0^2 - V''(x)/2$ is bounded from below for the cubic potential as well as any polynomial potential that diverges super-linearly as $x \to \pm\infty$.

Within the scaling theory developed in Section 4.2, if we specialize to the cubic potential (i.e. relevant variables of the renormalization group) the eigenvalues have scaling forms

$$\lambda_n = g_0^{2/3}\Lambda_n(\alpha), \tag{4.27}$$

written in terms of the RG-invariant quantity $\alpha = \epsilon/g_0^{4/3}$. This also implies a scaling form for the cumulants $\kappa_m = g_0^{-2m/3}\mathcal{K}_m(\alpha)$. The scaling forms $\Lambda_n(\alpha)$ are plotted in Figure 4.3 for the first several eigenvalues. These curves describing the α-dependence of the Fokker-Planck spectra are universal for systems near a saddle-node bifurcation (where α is the only relevant variable). For large positive α, the eigenvalues are approximately evenly spaced, while for large negative α, a single slowly decaying mode with eigenvalue very close to 0 dominates the behavior.

4.4.3 Evenly spaced eigenvalue approximation

To develop an approximation to the distribution of barrier escape times for the cubic potential, we assume the eigenvalues are equally spaced, $\Lambda_n(\alpha) = \Lambda_\alpha + n\Delta_\alpha$, where $\Lambda_\alpha \equiv \Lambda_0(\alpha)$ is the decay rate of the slowest decaying eigenmode. This approximation becomes exact in the limit $\alpha \to \infty$, for which the effective potential in Eq. (4.26) becomes harmonic. Though we can see in Figure 4.3 that this approximation is clearly not correct for $\alpha \to -\infty$, the behavior in this limit is dominated by the vanishing leading eigenvalue Λ_α. For instance, the cumulants in Eq. (4.24) are insensitive to Δ_α when $\Delta_\alpha \gg \Lambda_\alpha$ and $\Lambda_\alpha \to 0$. Thus, this approximation will produce a family of distributions which correctly captures the behavior for both large negative and positive α. As we will see below, the approximation is also quite accurate over the full range of α.

One caveat of the equal spacing approximation is that the mean escape time $\mathcal{T} = \mathcal{K}_0 = \sum_{n=1}^{\infty} \Lambda_n^{-1}$ always diverges. Therefore, we must fix the mean to the value derived in Section 4.3.1. On the other hand, the higher order cumulants, which determine the shape of the distribution, always converge.

To proceed, we sum Eq. (4.24) using our eigenvalue ansatz. For $m \geq 2$ the result is,

$$\mathcal{K}_m(\alpha) = (-1)^m \frac{\psi^{(m-1)}(\Lambda_\alpha/\Delta_\alpha)}{\Delta_\alpha^m}, \tag{4.28}$$

where $\psi^{(m)}(x)$ is the polygamma function of order m. Neglecting the mean, we can sum the series Eq. (4.23), then exponentiate and inverse Fourier transform to obtain the distribution $p(t)$. Writing the distribution in terms of its universal scaling function, $p(t, \epsilon, g_0) = g_0^{2/3}\rho(s, \alpha)$ with $s = (t - \tau)/g_0^{2/3}$ and $\alpha = \epsilon/g_0^{4/3}$, the

final result is

$$\rho(s,\alpha) = \frac{\Delta_\alpha}{\Gamma(\Lambda_\alpha/\Delta_\alpha)} \exp\Big(e^{\psi^{(0)}(\Lambda_\alpha/\Delta_\alpha)-\Delta_\alpha s} - \Lambda_\alpha s + \frac{\Lambda_\alpha}{\Delta_\alpha}\psi^{(0)}(\Lambda_\alpha/\Delta_\alpha)\Big). \tag{4.29}$$

To use this family of distributions, we compute the leading eigenvalue Λ_α using a shooting method on the right hand side of Eq. (4.26). The effective Δ_α is chosen so that the second cumulant $\mathcal{K}_2$ (i.e. the variance) is exact. To do this, we follow the same approach used in Section 4.3.1 to write an integral expression for the variance. The result is $\mathcal{K}_2 = \tau_2 - \tau^2$ with

$$\tau_2(x_i|x_f) = \frac{2}{g_0^2}\int_{x_i}^{x_f} dy \int_{-\infty}^{y} dz\, \tau(z|x_f)\, e^{-\frac{2}{g_0^2}[V(z)-V(y)]}, \tag{4.30}$$

where $\tau(z|x_f)$ is the mean given in Eq. (4.14). After evaluating these integrals numerically to obtain the variance, we solve $\mathcal{K}_2 = \psi^{(1)}(\Lambda_\alpha)/\Delta_\alpha^2$ to fix the eigenvalue spacing Δ_α.

In Figure 4.4 we compare the approximate distribution Eq. (4.29) to those obtained by direct numerical simulation of the Fokker-Plank equation. Our approximation captures the shape of the distributions remarkably well for the full range of α.

We can also show analytically that this distribution reproduces the correct large α limits. For large negative α, taking the $\Delta_\alpha \to \infty$ limit of Eq. (4.29), gives the expected exponential distribution, $\rho(s,\alpha) \to \Lambda_\alpha e^{-s\Lambda_\alpha - 1}$ (shifted to have zero mean). For large positive we can estimate Λ_α and Δ_α using the harmonic approximation to the potential in Eq. (4.26). We find $\Lambda_\alpha \sim \alpha^2$ and $\Delta_\alpha \sim \sqrt{\alpha}$ so that the cumulants scale like $\mathcal{K}_m(\alpha) \sim \alpha^{3(1-m)/2}$ for $\alpha \to \infty$. In particular, the ratio $\mathcal{K}_m/\mathcal{K}_2^{m/2} \to 0$ for $m > 2$, i.e. the higher order cumulants are small compared to the variance and the distribution approaches a Gaussian for large α. This

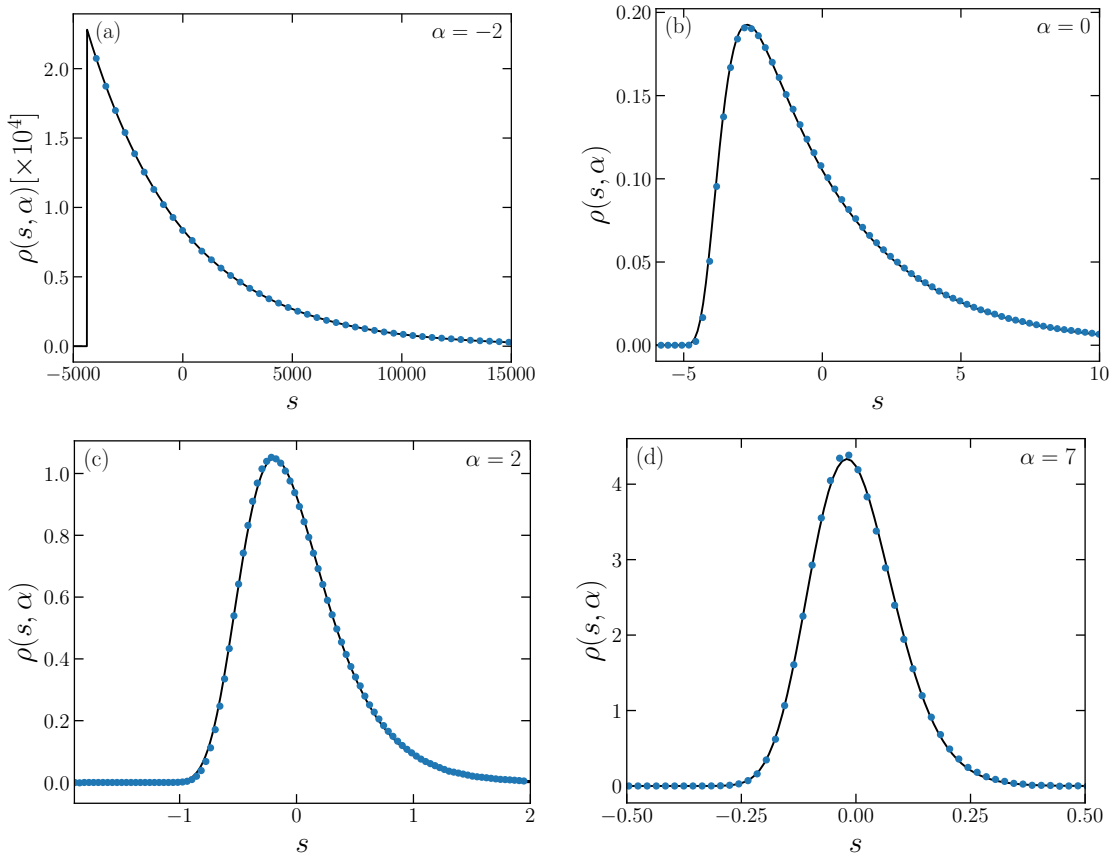

Figure 4.4: The barrier crossing time distributions obtained using our evenly spaced eigenvalue approximation Eq. (4.29) (lines) and from direct simulation of the Fokker-Planck equation (symbols) for (a) $\alpha = -2$, (b) $\alpha = 0$, (c) $\alpha = 2$, and (d) $\alpha = 7$. In all cases agreement between the theory and simulations is excellent. In the large barrier limit (a) the distribution is approximately exponential and in the strongly sloped potential (d) it is nearly Gaussian.

prediction is born out in direct simulation of the Fokker-Plank equation (see Figure 4.4d).

The distribution Eq. (4.29) combined with our analytical understanding of the mean escape time provides a complete description of the barrier crossing process. Our procedure allows for accurate approximations to the distribution by evaluating just the two universal scaling forms Λ_α and Δ_α. Knowledge of these quantities allows for evaluation of the distribution of barrier crossing

times for any system near a saddle-node bifurcation.

4.5 Discussion

We expect our results will be directly applicable to barrier crossing processes in which thermal fluctuations are comparable to the energy barrier including the aforementioned experimental systems, narrow escape problems in cellular biology [108], and downhill protein folding scenarios [109, 110]. Our approximation to the distribution of barrier escape times, combined with our analytical results for the mean, provides an accurate and complete characterization of the barrier crossing process for systems near a saddle-node bifurcation.

A more thorough analysis of incorporating perturbative corrections from irrelevant variables into Eq. (4.14) would be both theoretically interesting and useful in applications. The interplay between anharmonic corrections and finite initial conditions is also important. Computing these corrections will extend the applicability of our theory to systems in which the boundary conditions correspond to positions with non-negligible anharmonicity. It will be interesting to test the accuracy of our evenly spaced eigenvalue approximation to the barrier escape time distribution when irrelevant variables are incorporated. We conjecture that, at least perturbatively, Eq. (4.29) will still accurately parameterize the distributions if Λ_α and Δ_α are corrected to account for the irrelevant variables.

Our analysis directly translates to higher order cuspoid catastrophes [111, Section 36.2], which form their own universality classes with different exponents (in fact, these have already been analyzed for the discrete iterated map [98, 99]). For these bifurcations, the fixed point potential will be a higher order

monomial and our analysis can be used to identify the relevant variables and develop a scaling theory for quantities like the barrier crossing time mean and distribution.

More generally, it would be useful to study the applicability of our renormalization group and scaling analysis to systems with colored noise, multiple dimensions, or in other damping regimes. The effects of colored noise are encoded in the correlation function $\langle \xi(t)\xi(t') \rangle = G(x, t - t')$. The renormalization group transformation can be adapted to act on the Fourier transform of this quantity $\tilde{G}(x, \omega)$, giving flows of the colored noise under coarse-graining (for example, barrier crossing between two symmetric wells – a noisy pitchfork bifurcation – coupled to an Ohmic heat bath leads in the quantum limit to a critical point in the same universality class as the Kondo problem [112]). We expect short-range correlations will be irrelevant under coarse-graining, while those with power-law decay will give rise to new anomalous scaling. For some reactions, an underdamped model or multi-dimensional reaction coordinate may be required for an accurate description. Renormalization group scaling will provide a natural organizing framework for these studies.

Nucleation of abrupt phase transitions (e.g., raindrop formation) is also described by Arrhenius rates. Here the noiseless bifurcation underlying critical droplet theory is the *spinodal line*. This line – a discredited mean-field boundary between nucleation and spontaneous phase separation – could play the role of our renormalization-group fixed point in a future generalization of this work to higher spatial dimensions.

Finally, the saddle-node bifurcation is the simplest example of a depinning transition. One anticipates studying how adding noise would affect depinning

of earthquakes, vortices in superconductors, plastic flow in crystals, raindrops on windshields, coffee soaking into napkins, and other depinning phenomena. Each of these has anomalous exponents and RG treatments [113] even without added noise. Our work provides a stepping stone toward understanding the universal scaling near noisy depinning transitions in these more sophisticated systems as well.

CHAPTER 5
CONCLUSIONS

The results of the preceding chapters lay out a series of universality classes and scaling laws that arise in a broad range of first-passage processes. Chapters 2 and 3 found that a few common absorption-time distributions arise in discrete stochastic systems and connected these resulting distributions to basic features of the underlying dynamics. Chapter 4 characterized the universal scaling in barrier escape times and distributions near the critical saddle-node transition.

While these results are quite compelling, it remains to be seen whether and to what extent they arise in practice: do observed first-passage times in real-world systems follow a Gumbel distribution, Gaussian distribution or another member of the skewed family of distributions classified in Chapter 3? Are the metastable lifetimes for systems operating near a saddle-node bifurcation captured by the scaling law derived in Chapter 4? To conclude this thesis, this chapter will discuss a number of experimental systems in which our predictions can be tested to answer these questions. We also mention a few broad future theoretical directions that will likely need to be explored to capture the behavior of experiments that break the universality or scaling predicted by our work.

Since first-passage processes are so ubiquitous, there is a broad range of contexts in which our theoretical predictions can be tested: from epidemiological field data to bacterial evolution, single-cell aging, and optical laboratory experiments. In each of these examples (discussed further below), if the theoretically predicted scaling or fluctuations are observed, we can use this agreement to extract information about the physics underlying the system. For instance, the

shape of an absorption-time distribution might imply linear or power-law decay of transition rates. Similarly the distribution and mean lifetime of a metastable state can be used to infer energy barrier heights and the proximity to a saddle-node bifurcation. On the other hand, experimental results that break our universality classes or scaling laws will help inform future theoretical work outlined in the following section.

The first potential experimental test of our results was briefly mentioned in Chapter 3: eradication of African sleeping sickness. There we found that the eradication times measured in simulations of a high-dimensional model [2] of the disease that was fit to case data from different regions of the Democratic Republic of Congo (DRC) followed the Gumbel distribution predicted by our one-dimensional theory. In the coming years, tracking the time to eradicate sleeping sickness in the 168 endemic health-zones in the DRC will enable a more direct comparison between the predicted Gumbel distribution and the epidemiological data.

A similar study might be possible for COVID-19, for which an abundance of case data is available on county, state, and country scales [114]. In this case, mutation complicates the disease dynamics: rather than the epidemic dying out entirely, new outbreaks involving mutated variants of the disease emerge [115]. Perhaps, however, the outbreak duration (the time between a peak and trough in case loads before emergence of a new variant) follows similar statistics to the epidemic duration in the absence of mutations. Does the Gumbel distribution explain the variability in outbreak durations between different locations?

Our results can also be tested in a variety of laboratory experiments. For example, large scale bacteria evolution experiments enable tracking of the abun-

dance of genetic mutations and associated fitness levels over many generations [116, 117]. Starting with a genetically homogeneous population and measuring the time for new mutations to become fixed would directly test the predictions made in Chapter 2. For example, we expect a fixation-time distribution that becomes more skewed for mutations that provide larger fitness advantages. Longer term evolution experiments [117] can also be used to test whether our results are robust to heterogeneity in the evolving population.

Aging and death can be thought of as a first-passage process. Both mammalian species [118, 119] as well as yeast and other single-cell organisms [118] are known to follow the Gompert-Makeham mortality law, which describes how the mortality rate increases as a function of age. It turns out the Gompertz distribution is closely related to the Gumbel distribution, which arose as a universal absorption-time distribution in our theory. It would be exciting to test whether the distribution of lifetimes in single-cell aging experiments follows a Gumbel distribution. Perhaps the slowdown in reproduction rates near death in yeast [120] is analogous to the coupon-collection slowdowns that lead to Gumbel absorption-times.

Finally, optical systems provide a particularly effective medium for testing the scaling-laws near the noisy saddle-node bifurcation obtained in Chapter 4. Optoelectronic systems have previously been used to observe complex dynamical states, including chimeras [121] and laminar chaos [122, 123]. Laser optics enable nonlinear dynamics, which are necessary to undergo a saddle-node bifurcation, and noise can be introduced to the system in a well controlled way.

Beyond tests of our results, the scaling laws in Chapter 4 may prove useful for design of signal detectors [93, 96] and controllable magneto-elastic machines

[124], which both operate in the vicinity of a saddle node bifurcation. The lifetime of a metastable state determines the detector's sensitivity and the machines stability to perturbations.

Inevitably, many experimental systems will go beyond the universality classes discussed in the chapters of this thesis. Such experiments might display new and interesting behaviors that arise, for example, due to higher dimensionality, finite size effects, corrections to scaling, or non-Markovian noise. Building on the results of this thesis and modeling each of these effects will further expand our understanding of fluctuations and scaling in first-passage processes.

BIBLIOGRAPHY

[1] David Hathcock and Steven H. Strogatz. Fitness dependence of the fixation-time distribution for evolutionary dynamics on graphs. *Phys. Rev. E*, 100:012408, 2019.

[2] Maryam Aliee, Kat S. Rock, and Matt J. Keeling. Estimating the distribution of time to extinction of infectious diseases in mean-field approaches. *Journal of The Royal Society Interface*, 17(173):20200540, 2020.

[3] F. Di Lauro, J.-C. Croix, M. Dashti, L. Berthouze, and I. Z. Kiss. Network inference from population-level observation of epidemics. *Scientific Reports*, 10(1):18779, 2020.

[4] Mahdi Hajihashemi and Keivan Aghababaei Samani. Fixation time in evolutionary graphs: A mean-field approach. *Phys. Rev. E*, 99:042304, 2019.

[5] Peter Hänggi, Peter Talkner, and Michal Borkovec. Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.*, 62:251–341, 1990.

[6] Charles R. Doering, Khachik V. Sargsyan, and Leonard M. Sander. Extinction times for birth-death processes: Exact results, continuum asymptotics, and the failure of the Fokker–Planck approximation. *Multiscale Modeling & Simulation*, 3(2):283–299, 2005.

[7] Charles R. Doering, Khachik V. Sargsyan, Leonard M. Sander, and Eric Vanden-Eijnden. Asymptotics of rare events in birth–death processes bypassing the exact solutions. *Journal of Physics: Condensed Matter*, 19(6):065145, 2007.

[8] Mitchell J. Feigenbaum. Quantitative universality for a class of nonlinear transformations. *Journal of Statistical Physics*, 19(1):25–52, 1978.

[9] Martin A. Nowak. *Evolutionary Dynamics*. Harvard University Press, 2006.

[10] Takeo Maruyama. A markov process of gene frequency change in a geographically structured population. *Genetics*, 76(2):367–377, 1974.

[11] Takeo Maruyama. A simple proof that certain quantities are independent of the geographical structure of population. *Theoretical Population Biology*, 5(2):148 – 154, 1974.

[12] Montgomery Slatkin. Fixation probabilities and fixation times in a subdivided population. *Evolution*, 35(3):477–488, 1981.

[13] Erez Lieberman, Christoph Hauert, and Martin A. Nowak. Evolutionary dynamics on graphs. *Nature*, 433:312–, 2005.

[14] Tibor Antal and István Scheuring. Fixation of strategies for an evolutionary game in finite populations. *Bulletin of Mathematical Biology*, 68(8):1923–1944, 2006.

[15] Bahram Houchmandzadeh and Marcel Vallade. The fixation probability of a beneficial mutation in a geographically structured population. *New Journal of Physics*, 13(7):073020, 2011.

[16] Josep Díaz, Leslie Ann Goldberg, George B. Mertzios, David Richerby, Maria Serna, and Paul G. Spirakis. Approximating fixation probabilities in the generalized moran process. *Algorithmica*, 69(1):78–91, 2014.

[17] Kamran Kaveh, Natalia L. Komarova, and Mohammad Kohandel. The

duality of spatial death–birth and birth–death processes and limitations of the isothermal theorem. *Royal Society Open Science*, 2(4):140465, 2015.

[18] Alastair Jamieson-Lane and Christoph Hauert. Fixation probabilities on superstars, revisited and revised. *Journal of Theoretical Biology*, 382:44 – 56, 2015.

[19] P. M. Altrock, A. Traulsen, and M. A. Nowak. Evolutionary games on cycles with strong selection. *Phys. Rev. E*, 95:022407, 2017.

[20] Josef Tkadlec, Andreas Pavlogiannis, Krishnendu Chatterjee, and Martin A. Nowak. Population structure determines the tradeoff between fixation probability and fixation time. *Communications Biology*, 2(1):138, 2019.

[21] Motoo Kimura. Average time until fixation of a mutant allele in a finite population under continued mutation pressure: Studies by analytical, numerical, and pseudo-sampling methods. *Proceedings of the National Academy of Sciences*, 77(1):522–526, 1980.

[22] Philipp M. Altrock and Arne Traulsen. Fixation times in evolutionary games under weak selection. *New Journal of Physics*, 11(1):013012, 2009.

[23] Marcus Frean, Paul B. Rainey, and Arne Traulsen. The effect of population structure on the rate of evolution. *Proceedings of the Royal Society B: Biological Sciences*, 280(1762):20130211, 2013.

[24] Marziyeh Askari and Keivan Aghababaei Samani. Analytical calculation of average fixation time in evolutionary graphs. *Phys. Rev. E*, 92:042707, 2015.

[25] Marziyeh Askari, Zeinab Moradi Miraghaei, and Keivan Aghababaei Samani. The effect of hubs and shortcuts on fixation time in evolu-

tionary graphs. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(7):073501, 2017.

[26] Suzan Farhang-Sardroodi, Amir H. Darooneh, Moladad Nikbakht, Natalia L. Komarova, and Mohammad Kohandel. The effect of spatial randomness on the average fixation time of mutants. *PLOS Computational Biology*, 13(11):1–22, 2017.

[27] David Dingli, Arne Traulsen, and Jorge M. Pacheco. Stochastic dynamics of hematopoietic tumor stem cells. *Cell Cycle*, 6(4):461–466, 2007.

[28] Philipp M. Altrock, Arne Traulsen, and Floyd A. Reed. Stability properties of underdominance in finite subdivided populations. *PLOS Computational Biology*, 7(11):1–10, 2011.

[29] Peter Ashcroft, Arne Traulsen, and Tobias Galla. When the mean is not enough: Calculating fixation time distributions in birth-death processes. *Phys. Rev. E*, 92:042154, 2015.

[30] Li-Min Ying, Jie Zhou, Ming Tang, Shu-Guang Guan, and Yong Zou. Mean-field approximations of fixation time distributions of evolutionary game dynamics on graphs. *Frontiers of Physics*, 13(1):130201, 2018.

[31] David Aldous. Interacting particle systems as stochastic social dynamics. *Bernoulli*, 19(4):1122–1149, 2013.

[32] Bertrand Ottino-Löffler, Jacob G. Scott, and Steven H. Strogatz. Takeover times for a simple model of network infection. *Phys. Rev. E*, 96:012313, 2017.

[33] Bertrand Ottino-Löffler, Jacob G Scott, and Steven H Strogatz. Evolutionary dynamics of incubation periods. *eLife*, 6:e30212, 2017.

[34] P. A. P. Moran. Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 54(1):60–71, 1958.

[35] P. A. P. Moran. The effect of selection in a haploid genetic population. *Mathematical Proceedings of the Cambridge Philosophical Society*, 54(4):463–467, 1958.

[36] Paul Erdős and Alfréd Rényi. On a classical problem of probability theory. *Publ. Math. Inst. Hung. Acad. Sci.*, 6:215–220, 1961.

[37] Søren Asmussen. *Applied Probability and Queues*. Springer–Verlag, 2003.

[38] Julian Keilson. *Markov Chain Models – Rarity and Exponentiality*. Springer–Verlag, 1979.

[39] Yuri Bakhtin. Universal statistics of incubation periods and other detection times via diffusion models. *Bulletin of Mathematical Biology*, 2018.

[40] Narendra S Goel and Nira Richter-Dyn. *Stochastic models in biology*. Academic Press, 1974.

[41] J. Keilson. A review of transient behavior in regular diffusion and birth–death processes. Part II. *Journal of Applied Probability*, 2(2):405–428, 1965.

[42] Rick Durrett. *Probability: Theory and Examples*. Cambridge University Press, 2010.

[43] Motoo Kimura. The length of time required for a selectively neutral mutant to reach fixation through random frequency drift in a finite population. *Genetical Research*, 15(1):131–133, 1970.

[44] Trevor Williams and Rolf Bjerknes. Stochastic model for abnormal clone spread through epithelial basal layer. *Nature*, 236(5340):19–21, 1972.

[45] Andrea Sottoriva, Inmaculada Spiteri, Sara G. M. Piccirillo, Anestis Touloumis, V. Peter Collins, John C. Marioni, Christina Curtis, Colin Watts, and Simon Tavaré. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proceedings of the National Academy of Sciences*, 110(10):4009–4014, 2013.

[46] Ivana Bozic, Johannes G. Reiter, Benjamin Allen, Tibor Antal, Krishnendu Chatterjee, Preya Shah, Yo Sup Moon, Amin Yaqubie, Nicole Kelly, Dung T. Le, Evan J. Lipson, Paul B. Chapman, Jr Diaz, Luis A., Bert Vogelstein, and Martin A. Nowak. Evolutionary dynamics of cancer in response to targeted combination therapy. *eLife*, 2:e00747, 2013.

[47] Motoo Kimura. Evolutionary rate at the molecular level. *Nature*, 217(5129):624–626, 1968.

[48] Sudhir Kumar. Molecular clocks: four decades of evolution. *Nature Reviews Genetics*, 6:654, 2005. Perspective.

[49] Robert Zwanzig, Attila Szabo, and Biman Bagchi. Levinthal's paradox. *Proceedings of the National Academy of Sciences*, 89(1):20–22, 1992.

[50] C.M. da Fonseca and J. Petronilho. Explicit inverses of some tridiagonal matrices. *Linear Algebra and its Applications*, 325(1):7 – 21, 2001.

[51] John G. Kemeny and J. Laurie Snell. *Finite Markov Chains*. Springer, 1983.

[52] Alan R. Hinman. Global progress in infectious disease control. *Vaccine*, 16(11):1116–1121, 1998.

[53] Donald R. Hopkins. Disease eradication. *New England Journal of Medicine*, 368(1):54–63, 2013.

[54] Philip E. Sartwell. The distribution of incubation periods of infectious disease. *American Journal of Epidemiology*, 51(3):310–318, 1950.

[55] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. Epidemic processes in complex networks. *Rev. Mod. Phys.*, 87:925–979, 2015.

[56] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. Critical phenomena in complex networks. *Rev. Mod. Phys.*, 80:1275–1335, 2008.

[57] John A. Jacquez and Carl P. Simon. The stochastic SI model with recruitment and deaths I. Comparison with the closed SIS model. *Mathematical Biosciences*, 117(1):77 – 125, 1993.

[58] Johan Grasman and Reinier HilleRisLambers. On local extinction in a metapopulation. *Ecological Modelling*, 103(1):71 – 80, 1997.

[59] Sandro Azaele, Samir Suweis, Jacopo Grilli, Igor Volkov, Jayanth R. Banavar, and Amos Maritan. Statistical mechanics of ecological systems: Neutral theory and beyond. *Rev. Mod. Phys.*, 88:035003, 2016.

[60] Amar Gandhi, Simon Levin, and Steven Orszag. "Critical slowing down" in time-to-extinction: an example of critical phenomena in ecology. *Journal of Theoretical Biology*, 192(3):363–376, 1998.

[61] Nicolaas Godfried Van Kampen. *Stochastic Processes in Physics and Chemistry*, volume 1. Elsevier, 1992.

[62] Immanuel Meyer and Nadav M. Shnerb. Evolutionary dynamics in fluctuating environment. *Phys. Rev. Research*, 2:023308, 2020.

[63] Narendra S Goel and Nira Richter-Dyn. *Stochastic models in biology*. Elsevier, 2016.

[64] Peter G. Hufton, Elizabeth Buckingham-Jeffery, and Tobias Galla. First-passage times and normal tissue complication probabilities in the limit of large populations. *Scientific Reports*, 10(1):8786, 2020.

[65] Hisashi Ohtsuki and Martin A Nowak. Evolutionary games on cycles. *Proceedings of the Royal Society B: Biological Sciences*, 273(1598):2249–2256, 2006.

[66] Emil Julius Gumbel. Statistical theory of extreme values and some practical applications. *NBS Applied Mathematics Series*, 33, 1954.

[67] Peter Ashcroft. *The Statistical Physics of Fixation and Equilibration in Individual-Based Models*. Springer, 2016.

[68] Yitzhak Yahalom, Bnaya Steinmetz, and Nadav M. Shnerb. Comprehensive phase diagram for logistic populations in fluctuating environment. *Phys. Rev. E*, 99:062417, 2019.

[69] Pierre Collet, Servet Martínez, and Jaime San Martín. *Quasi-stationary distributions: Markov chains, diffusions and dynamical systems*, volume 1. Springer, 2013.

[70] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284 – 304, 1940.

[71] L Farkas. Keimbildungsgeschwindigkeit in übersättigten dämpfen. *Z. Phys. Chem.*, 125(1):236–242, 1927.

[72] R. Becker and W. Döring. Kinetische behandlung der keimbildung in übersättigten dämpfen. *Annalen der Physik*, 416(8):719–752, 1935.

[73] Henry Eyring. The activated complex in chemical reactions. *The Journal of Chemical Physics*, 3(2):107–115, 1935.

[74] S. Arrhenius. Über die reaktionsgeschwindigkeit bei der inversion von rohrzucker durch säuren. *Z. Phys. Chem.*, 4:226, 1889.

[75] V. I. Mel'nikov and S. V. Meshkov. Theory of activated rate processes: Exact solution of the Kramers problem. *The Journal of Chemical Physics*, 85(2):1018–1027, 1986.

[76] Bo Cartling. Kinetics of activated processes from nonstationary solutions of the Fokker-Planck equation for a bistable potential. *The Journal of Chemical Physics*, 87(5):2638–2648, 1987.

[77] Donald L. Ermak and J. A. McCammon. Brownian dynamics with hydrodynamic interactions. *The Journal of Chemical Physics*, 69(4):1352–1360, 1978.

[78] A. W. C. Lau and T. C. Lubensky. State-dependent diffusion: Thermodynamic consistency and its path integral formulation. *Phys. Rev. E*, 76:011123, 2007.

[79] Jing-Dong Bao, Hai-Yan Wang, Ying Jia, and Yi-Zhong Zhuo. Cancellation phenomenon of barrier escape driven by a non-gaussian noise. *Phys. Rev. E*, 72:051105, 2005.

[80] B. Dybiec, E. Gudowska-Nowak, and P. Hänggi. Escape driven by α-stable white noises. *Phys. Rev. E*, 75:021109, 2007.

[81] Alendu Baura, Monoj Kumar Sen, Gurupada Goswami, and Bidhan Chandra Bag. Colored non-gaussian noise driven open systems: Generalization of Kramers' theory with a unified approach. *The Journal of Chemical Physics*, 134(4):044126, 2011.

[82] George P. Tsironis and Paolo Grigolini. Escape over a potential barrier in the presence of colored noise: Predictions of a local-linearization theory. *Phys. Rev. A*, 38:3749–3757, 1988.

[83] Peter Hänggi. Escape over fluctuating barriers driven by colored noise. *Chemical Physics*, 180(2):157 – 166, 1994.

[84] O. Edholm and O. Leimar. The accuracy of Kramers' theory of chemical kinetics. *Physica A: Statistical Mechanics and its Applications*, 98(1):313 – 324, 1979.

[85] W. Bez and P. Talkner. A new variational method to calculate escape rates in bistable systems. *Physics Letters A*, 82(7):313 – 316, 1981.

[86] Martin Bier and R. Dean Astumian. Matching a diffusive and a kinetic approach for escape over a fluctuating barrier. *Phys. Rev. Lett.*, 71:1649–1652, 1993.

[87] Vladimir V Palyulin and Ralf Metzler. How a finite potential barrier decreases the mean first-passage time. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(03):L03001, 2012.

[88] J. J. Mazo, O. Y. Fajardo, and D. Zueco. Thermal activation at moderate-to-high and high damping: Finite barrier effects and force spectroscopy. *The Journal of Chemical Physics*, 138(10):104105, 2013.

[89] N E Aktaev. Theoretical approach to modelling the low-barrier chemical reactions initiated by pulsed electron beam. *Journal of Physics: Conference Series*, 552:012033, 2014.

[90] Eli Pollak and Reuven Ianconescu. Finite barrier corrections to the PGH solution of Kramers' turnover theory. *The Journal of Chemical Physics*, 140(15):154108, 2014.

[91] Eli Pollak and Reuven Ianconescu. Kramers' turnover theory: Improvement and extension to low barriers. *The Journal of Physical Chemistry A*, 120(19):3155–3164, 2016.

[92] Zhan-Wu Bai. Simple generalization of Kramers theory to finite barrier height in spatial diffusion regime. *Physics Letters A*, 382(32):2103 – 2107, 2018.

[93] Nicholas J. Miller and Steven W. Shaw. Escape statistics for parameter sweeps through bifurcations. *Phys. Rev. E*, 85:046202, 2012.

[94] Corentin Herbert and Freddy Bouchet. Predictability of escape for a stochastic saddle-node bifurcation: When rare events are typical. *Phys. Rev. E*, 96:030201(R), 2017.

[95] J. Husson, M. Dogterom, and F. Pincet. Force spectroscopy of a single artificial biomolecule bond: The Kramers' high-barrier limit holds close to the critical force. *The Journal of Chemical Physics*, 130(5):051103, 2009.

[96] Yukihiro Tadokoro, Hiroya Tanaka, and M. I. Dykman. Driven nonlinear nanomechanical resonators as digital signal detectors. *Scientific Reports*, 8(1):11284, 2018.

[97] J. E. Hirsch, B. A. Huberman, and D. J. Scalapino. Theory of intermittency. *Phys. Rev. A*, 25:519–532, 1982.

[98] J.E. Hirsch, M. Nauenberg, and D.J. Scalapino. Intermittency in the presence of noise: A renormalization group formulation. *Physics Letters A*, 87(8):391 – 393, 1982.

[99] Bambi Hu and Joseph Rudnick. Exact solutions to the Feigenbaum renormalization-group equations for intermittency. *Phys. Rev. Lett.*, 48:1645–1648, 1982.

[100] Carl M Bender and Steven A Orszag. *Advanced Mathematical Methods for Scientists and Engineers: Asymptotic Methods and Perturbation Theory*. Springer-Verlag, 1999.

[101] Nigel Goldenfeld, Olivier Martin, and Y. Oono. Intermediate asymptotics and renormalization group theory. *Journal of Scientific Computing*, 4(4):355–372, 1989.

[102] Nigel Goldenfeld, Olivier Martin, Y. Oono, and Fong Liu. Anomalous dimensions and the renormalization group in a nonlinear diffusion process. *Phys. Rev. Lett.*, 64:1361–1364, 1990.

[103] Rouslan L Stratonovich. *Topics in the theory of random noise*, volume 2. Gordon and Breach, 1967.

[104] Askold N. Malakhov and Andrey L. Pankratov. *Evolution Times of Probability Distributions and Averages – Exact Solutions of the Kramers' Problem*, pages 357–438. John Wiley & Sons, Ltd, 2002.

[105] Rouslan L Stratonovich. *Topics in the theory of random noise*, volume 1. Gordon and Breach, 1963.

[106] Yahui Zhan and Bernie D. Shizgal. Diffusion in a bistable system: The eigenvalue spectrum of the Fokker-Planck operator and Kramers' reaction rate theory. *Phys. Rev. E*, 99:042101, 2019.

[107] John Cardy. *Scaling and Renormalization in Statistical Physics*. Cambridge University Press, 1996.

[108] Z. Schuss, A. Singer, and D. Holcman. The narrow escape problem for diffusion in cellular microdomains. *Proceedings of the National Academy of Sciences*, 104(41):16098–16103, 2007.

[109] J. Sabelko, J. Ervin, and M. Gruebele. Observation of strange kinetics in protein folding. *Proceedings of the National Academy of Sciences*, 96(11):6031–6036, 1999.

[110] Robert B. Best and Gerhard Hummer. Diffusive model of protein folding dynamics with Kramers turnover in rate. *Phys. Rev. Lett.*, 96:228104, 2006.

[111] *NIST Digital Library of Mathematical Functions*. Release 1.0.28 of 2020-09-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.

[112] Sudip Chakravarty. Quantum fluctuations in the tunneling between superconductors. *Phys. Rev. Lett.*, 49:681–684, 1982.

[113] Daniel S. Fisher. Collective transport in random media: from superconductors to earthquakes. *Physics Reports*, 301(1):113 – 150, 1998.

[114] Centers for Disease Control and Prevention. Covid Data Tracker. Atlanta, GA: US Department of Health and Human Services, CDC;

2022, August 15. https://covid.cdc.gov/covid-data-tracker/#datatracker-home.

[115] Nicola Bellomo, Diletta Burini, and Nisrine Outada. Multiscale models of covid-19 with mutations and variants. *Networks and Heterogeneous Media*, 17(3):293–310, 2022.

[116] Thomas Scheuerl, Meirion Hopkins, Reuben W. Nowell, Damian W. Rivett, Timothy G. Barraclough, and Thomas Bell. Bacterial adaptation is constrained in complex communities. *Nature Communications*, 11(1):754, 2020.

[117] Richard E Lenski. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. *The ISME Journal*, 11(10):2181–2194, 2017.

[118] James W. Vaupel, James R. Carey, Kaare Christensen, Thomas E. Johnson, Anatoli I. Yashin, Niels V. Holm, Ivan A. Iachine, Väinö Kannisto, Aziz A. Khazaeli, Pablo Liedo, Valter D. Longo, Yi Zeng, Kenneth G. Manton, and James W. Curtsinger. Biodemographic trajectories of longevity. *Science*, 280(5365):855–860, 1998.

[119] Omer Karin, Amit Agrawal, Ziv Porat, Valery Krizhanovsky, and Uri Alon. Senescent cell turnover slows with age providing an explanation for the gompertz law. *Nature Communications*, 10(1):5495, 2019.

[120] Jing Yang, Ziwei Wang, Xili Liu, Hao Li, and Qi Ouyang. Yeast replicative aging leads to permanent cell cycle arrest in g1 effectuated by the start repressor whi5. *bioRxiv*, 2018.

[121] Aaron M. Hagerstrom, Thomas E. Murphy, Rajarshi Roy, Philipp Hövel, Iryna Omelchenko, and Eckehard Schöll. Experimental observation of chimeras in coupled-map lattices. *Nature Physics*, 8(9):658–661, 2012.

[122] David Müller-Bender, Andreas Otto, Günter Radons, Joseph D. Hart, and Rajarshi Roy. Laminar chaos in experiments and nonlinear delayed langevin equations: A time series analysis toolbox for the detection of laminar chaos. *Phys. Rev. E*, 101:032213, Mar 2020.

[123] Joseph D. Hart, Rajarshi Roy, David Müller-Bender, Andreas Otto, and Günter Radons. Laminar chaos in experiments: Nonlinear systems with time-varying delays and noise. *Phys. Rev. Lett.*, 123:154101, Oct 2019.

[124] Teaya Yang, David Hathcock, Yuchao Chen, James P. Sethna Sethna, Itai Cohen, and Itay Griniasty. Bifurcation instructed design of magneto elastic machines. *Unpublished*, 2022.