PHYSICAL REVIEW E **108**, 044406 (2023)

Extinction time distributions of populations and genotypes

David A. Kessler and Nadav M. Shnerb
Department of Physics, Bar-Ilan University, Ramat-Gan 52900, Israel

(Received 18 July 2023; accepted 29 September 2023; published 20 October 2023)

Ultimately, the eventual extinction of any biological population is an inevitable outcome. While extensive research has focused on the average time it takes for a population to go extinct under various circumstances, there has been limited exploration of the distributions of extinction times and the likelihood of significant fluctuations. Recently, Hathcock and Strogatz [D. Hathcock and S. H. Strogatz, Phys. Rev. Lett. **128**, 218301 (2022)] identified Gumbel statistics as a universal asymptotic distribution for extinction-prone dynamics in a stable environment. In this study we aim to provide a comprehensive survey of this problem by examining a range of plausible scenarios, including extinction-prone, marginal (neutral), and stable dynamics. We consider the influence of demographic stochasticity, which arises from the inherent randomness of the birth-death process, as well as cases where stochasticity originates from the more pronounced effect of random environmental variations. Our work proposes several generic criteria that can be used for the classification of experimental and empirical systems, thereby enhancing our ability to discern the mechanisms governing extinction dynamics. Employing these criteria can help clarify the underlying mechanisms driving extinction processes.

DOI: 10.1103/PhysRevE.108.044406

I. INTRODUCTION

Biological populations are inevitably destined for extinction. Over 99% of all known species that have ever existed on earth have already become extinct and the others are awaiting their inevitable turn. The concern over the anthropogenic acceleration of extinction rates has sparked heated debates in the past decade regarding whether such acceleration is indeed observed in local populations [1,2] and, if so, what are the global implications of this change. Understanding the likelihood of extinction under specific conditions and the distribution of extinction times is crucial for predicting future extinction events and assessing the threat to biodiversity. The same questions also arise when the objective is to eliminate a particular biological entity, such as in the case of pest control, pathogen eradication, or combating genetic diseases.

The dynamics of biological populations is influenced by deterministic and stochastic factors. At the deterministic level, the dynamics can be classified into two main types: those attracted to a manifold (such as a fixed point) with finite population and those attracted to an extinction point. In the latter case the population decays over time towards zero. Populations of the first type would be expected to persist indefinitely, while populations of the second type disappear. In the common case of exponential decline, the extinction time is logarithmic in the size of the original population.

Stochasticity makes this picture much more subtle. Since the state of zero population is an absorbing state, the ultimate fate of *any* stochastic dynamics is extinction. The sharp distinction between extinction-prone and stable populations thus blurs, and the focus must switch to the characteristics of the extinction process and in particular to the statistical properties of extinction times.

Typically, stochasticity in biological systems is quite strong, even under extremely stable experimental conditions [3]. Stochastic fluctuations are usually classified into two categories, demographic stochasticity (or genetic drift or internal noise) and temporal environmental stochasticity (extrinsic noise) [4]. Demographic noise reflects the inherent randomness of the birth-death process caused by small-scale random events that affect the reproductive success of individuals in an uncorrelated manner. Temporal environmental stochasticity (TES) is associated with large-scale events that affect entire populations. Mathematically speaking, this implies that the parameters of a given model (usually, the growth rates) vary in time, where the amplitude and correlation times of the fluctuations characterize the environment. Abundance variations induced by TES are usually proportional to the population size, whereas those induced by demographic stochasticity scale with the square root of population size. Therefore, demographic stochasticity is typically negligible when population size is large [5–10], but it becomes important at the brink of extinction or during invasion [11–14].

This brings us to a third type of systems: those in which the deterministic dynamics is weak or negligible and stochasticity is the main, or only, driver of fluctuations. In this case, we are talking about neutral dynamics, a topic of great importance in population genetics and community ecology [9,15–17]. In sum, our classification consists of six types of systems: persistent, extinction prone, and neutral, each of which can be analyzed under pure demographic noise or under a combination of demographic and environmental stochasticity.

In recent works, Strogatz and Hathcock [18,19] analyzed the distribution of extinction times for an extinction-prone (negative growth rate, exponentially decaying) population with pure demographic stochasticity. These authors found a universal asymptotic behavior, i.e., that the fluctuations around the expected extinction time obey a Gumbel distribution. Furthermore, the width of this distribution is extremely narrow: While the deterministic mean time to extinction scales with the logarithm of the initial population size N_0, the width is N_0 independent. Therefore, relative fluctuations around the mean vanish as $N \to \infty$.

Here we would like to extend the work of Hathcock and Strogatz [19] and consider the statistics of extinction times in all six typical scenarios mentioned above. Some of these cases have already been discussed in the literature and we provide the relevant details below. Still, we believe that there is great benefit in presenting a complete picture, including both original and former works, so that a researcher interested in this topic can compare and contrast the different alternatives. In what follows we devote a single section to each of the six scenarios and in the last section we will discuss the results and provide a general outlook.

II. EXTINCTION-PRONE DYNAMICS IN A FIXED ENVIRONMENT

In this section we first revisit the class of systems considered by Hathcock and Strogatz [19], presenting in Sec. II A an alternate derivation of the fact that the Gumbel distribution is a universal limit. In Sec. II B we provide an example of a non-Gumbel scenario and analyze some of its features, from which a few aspects of the general picture emerge.

Special attention is directed to the relationship between the average lifespan of a population and the width of the distribution (variance-mean ratio, also known as the index of dispersion or Fano factor). In the Gumbel cases the mean time to extinction diverges (albeit logarithmically) at the large-N limit, while the width of the distribution remains constant. Therefore, fluctuations become negligible in large systems. This characteristic reflects the negligible effect of demographic noise when the abundance is large.

In certain systems, as we will explore, the average time to extinction is unaffected by the initial population size. In such cases, even in the limit of large initial size, the mean-variance ratio is $O(1)$, indicating significant fluctuations. Specifically, we examine a population-genetics model for diploids with dominance and offer insights into the broader scenario.

A. Density-independent dynamics and Gumbel statistics

We begin with a simple example in which the general answer is attainable and suggest an argument for the general case.

Let us consider a system with no density-dependent effects. In that case, for any single individual the birth and death rates, per unit time, are fixed, i.e., are independent of the state of other individuals. The death rate is taken to be μ and the birth rate is λ. If the population is extinction prone, $\mu > \lambda$.

The chance $P_n(t)$ of having n individuals at time t satisfies the differential equation

$$\frac{dP_n(t)}{dt} = \mu(n+1)P_{n+1} + \lambda(n-1)P_{n-1} - (\mu+\lambda)nP_n. \quad (1)$$

We would like to solve this equation and find P_0, the chance of extinction, given that $P_n(t=0) = \delta_{n,N_0}$. To do that we introduce the generating function

$$G(x,t) = \sum_{n=0}^{\infty} P_n x^n, \quad (2)$$

obeying

$$\dot{G} = \mu \sum_n x^n(n+1)P_{n+1} + \lambda \sum_n x^n(n-1)P_{n-1} - (\mu+\lambda)\sum_n x^n nP_n. \quad (3)$$

Redefinition of indices yields a first-order differential equation for G,

$$\dot{G} = \mu G' + \lambda x^2 G' - (\mu+\lambda)xG' = [\mu + \lambda x^2 - (\mu+\lambda)x]G' = \tilde{Q}(x)G'. \quad (4)$$

Equation (4) is a first-order equation that may be solver using characteristics [20]. Every function of the form $G(F(x)+t)$ will solve Eq. (4) if $dF/dx = 1/\tilde{Q}(x)$. For Eq. (4) the desired F is

$$F(x) = \frac{\ln\left(\frac{x-1}{\lambda x-\mu}\right)}{\lambda-\mu}. \quad (5)$$

What is left is to determine the functional form of $G[F(x)+t]$, and this has to do with the initial condition. Suppose at $t=0$ we have only one individual. In that case, by definition $G(x, t=0) = x$ and

$$G^{-1}(F(x)) = x. \quad (6)$$

The solution for Eq. (6) is

$$G(F, t=0) = \frac{\mu e^{(\lambda-\mu)F} - 1}{\lambda e^{(\lambda-\mu)F} - 1}. \quad (7)$$

Therefore the generating function at any time t is

$$G(F,t) = \frac{\mu e^{(\lambda-\mu)(F+t)} - 1}{\lambda e^{(\lambda-\mu)(F+t)} - 1} = \frac{\mu e^{(\lambda-\mu)t}\left(\frac{x-1}{\lambda x-\mu}\right) - 1}{\lambda e^{(\lambda-\mu)t}\left(\frac{x-1}{\lambda x-\mu}\right) - 1}. \quad (8)$$

Hence, the chance $Q(t)$ that at time t the lineage of a given individual has already gone extinct is

$$Q_{1\to 0}(t) = 1 - \frac{\mu-\lambda}{\mu e^{t(\mu-\lambda)} - \lambda}. \quad (9)$$

Since the dynamics of the lineages of all individuals are statistically identical (no density-dependent effects), if the population at $t=0$ has N_0 individuals,

$$Q_{N_0\to 0}(t) = \left(1 - \frac{\mu-\lambda}{\mu e^{(\mu-\lambda)t} - \lambda}\right)^{N_0}. \quad (10)$$

To see the connection between the distribution (10) and the Gumbel distribution, let us measure time in units of μ and define a decay parameter $\kappa = 1 - \lambda/\mu$. When $N_0 \to \infty$, the time t in which all individuals goes extinct is large and therefore

$$Q_{N_0\to 0}(t) \approx e^{-\kappa N_0 e^{-\kappa t}}. \quad (11)$$

The chance of extinction at t is $P(t) = dQ/dt$.

Now let us define $t = (s+\nu)/\kappa$, where $\nu = \ln(\beta N_0)$ is the point at which the large-N_0 distribution of extinction times,

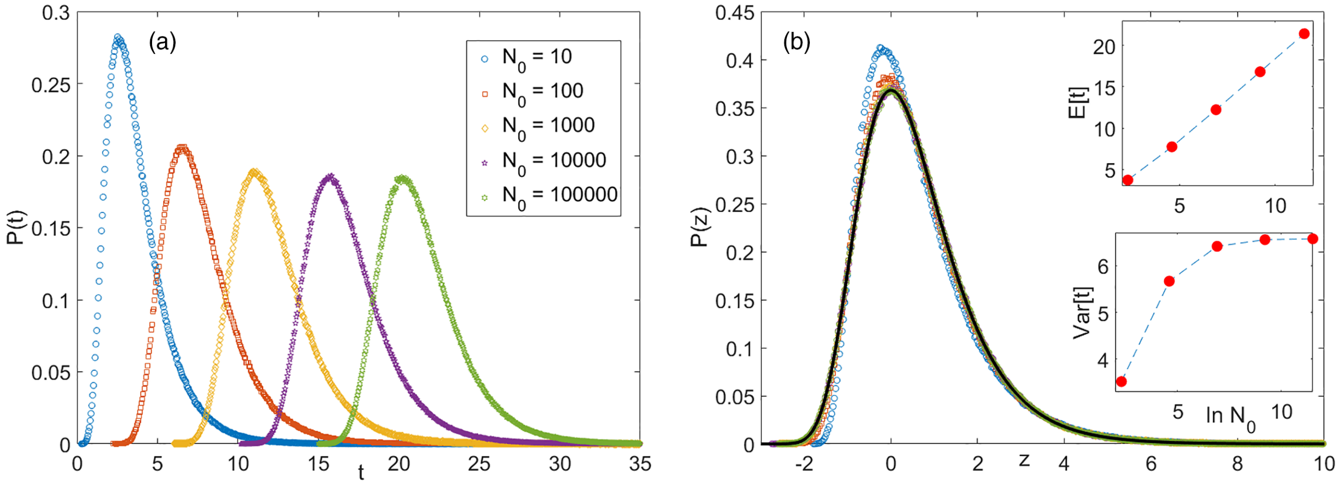

FIG. 1. (a) Distribution of extinction times $P(t)$ for extinction-prone populations when stochasticity is purely demographic. The dynamics is logistic, with finite carrying capacity N_0. In the presence of N individuals the total death rate is N and the total birth rate is $0.5N(1 - N/N_0)$; the initial population was taken to be N_0. The Gumbel distribution parameters $\beta = \sqrt{6\,\mathrm{Var}[t]/\pi^2}$ and $\mu = \mathbb{E}[t] - \gamma_E \beta$ (γ_E is Euler constant) were extracted for each N_0. (b) Plot of a histogram of the adjusted variable $z = (t - \mu)/\beta$, showing that all data collapse and fit the Gumbel distribution $\exp\{-[z + \exp(-z)]\}$ (black solid line). Small deviations are observed for $N_0 = 10$ and 100, but above these numbers there is perfect agreement between the predicted and the observed distribution. The mean and the variance for each N_0 are shown in the insets in (b). While the mean grows linearly with $\ln N_0$ (top inset) the variance saturates (bottom inset) to its predicted value for $\kappa = \frac{1}{2}$, namely, $2\pi^2/3 \approx 6.58$.

$P(t)$, is peaked, so the second derivative of the cumulative distribution $Q(t)$ vanishes. With that definition,

$$P(s) = e^{-(s+e^{-s})}, \tag{12}$$

which is the cumulative distribution function of the Gumbel distribution whose scale parameter is $\beta = 1/\kappa$ and its mode is $\mu = \nu/\kappa$. The standard deviation of this distribution is $\pi/\sqrt{6\kappa^2}$, an $O(1)$, N_0-independent number.

The variance-mean ratio (VMR) is then

$$\mathrm{VMR} = \frac{\pi^2}{6\kappa(\nu + \gamma_E)} = \frac{\pi^2}{6\kappa(\ln N_0 + \gamma_E - \ln\kappa)}, \tag{13}$$

where γ_E is Euler's number. Importantly, this ratio decays like $1/\ln N_0$ in the large-N_0 limit. As explained in the Appendix, the effect of demographic stochasticity is negligible out of the extinction zone in which $n < n_c$. In the region dominated by demographic noise, the dynamics is more or less neutral (discussed in Sec. IV) and hence the variance of extinction-time distribution is proportional to n_c. For extinction-prone systems with no density dependence, n_c is N_0 independent (see the Appendix). This feature may change in other scenarios, as demonstrated in Sec. II B.

The general result of Hathcock and Strogatz [19] may be interpreted as follows. Once the population is in decline, the intraspecific interactions are usually negligible. The question of extinction time of N_0 individuals is thus governed by the chance of the last lineage to go extinct. In the large-N_0 limit this becomes the classical extreme-event problem, so as long as the chance of a single lineage to persist decays exponentially at long times, the limit distribution is Gumbel [21]. The same answer holds for any other single-lineage extinction-time distribution which is neither compact nor fat tailed. The Gumbel statistics is demonstrated, in Fig. 1, for logistic dynamics with negative growth rate.

B. Density-dependent dynamics: Non-Gumbel scenarios

As pointed out by Hathcock and Strogatz [19], the Gumbel distribution is a universal asymptotic limit of many extinction-time statistics provided the rates of demographic events (transition rates) decrease linearly towards zero in the vicinity of the extinction point. This property reflects the weakening of the interactions between individuals in the extinction zone, so the rate of events is linearly proportional to the number of individuals. When this condition is not fulfilled, the distribution is not Gumbel. In this section we consider a specific example and then we derive a few formulas that provide insights into the more general cases.

As a realistic example, let us consider a population-genetics model for diploids with dominance [22–24]. This model describes the dynamics of two alleles A and a in a randomly mating diploid population. The allele A is always dominant to a, so the phenotype of an aA heterozygote is the same as the phenotype of AA. If the fraction of a alleles in the gamete pool is x and the fraction of A is $1 - x$, then, after random mating, the zygote genotypes follow classic Hardy-Weinberg proportions, with AA:Aa:aa as $(1 - x)^2$:$2x(1 - x)$:x^2.

Setting the fitness of AA and Aa phenotypes to unity and the fitness of aa to $f < 1$, one expects the a allele to disappear from a well-mixed fixed-size population. This purifying selection process is however very slow for small x, because an individual will only suffer from low fitness when both of its alleles are of type a. Since the number of a homozygotes is proportional to x^2, the process is always density dependent and one expects a non-Gumbel skewed distribution.

Figure 2 shows results from a simulation of this process. In each time step one individual is chosen to die, so with probability x^2 two a alleles are lost, with probability $2x(1 - x)$ only one a is lost, and the chance of zero a loss is $(1 - x)^2$.

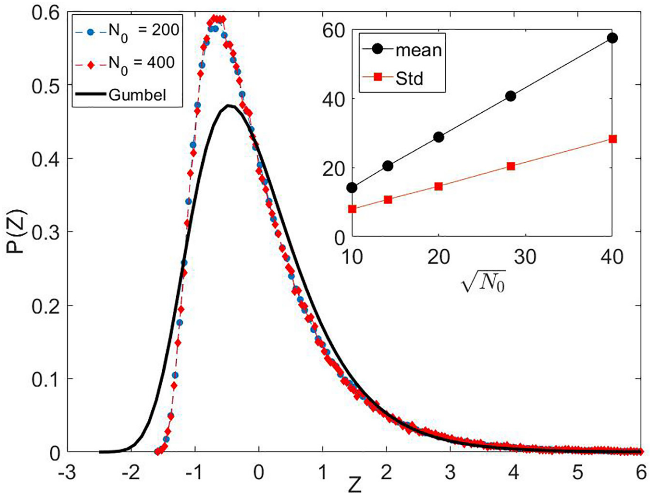

FIG. 2. Diploid with dominance: distribution of normalized extinction times $P(z)$ vs z (using the adjusted variable $z = (t - \mu)/\beta$, where $\beta = \sqrt{6\,\mathrm{Var}[t]/\pi^2}$ and $\mu = \mathbb{E}[t] - \gamma_E \beta$), where t is the time to extinction of the a allele whose fitness is $f = \frac{1}{2}$. Here N is the a allele initial frequency, out of total population of $2N_0$ alleles (N_0 diploid individuals). Results are shown for $N_0 = 200$ and 400 (each statistic reflects 10^5 numerical experiments). Both distributions are almost identical and differ substantially from the Gumbel curve (black line). The inset shows the mean (black circles) and the standard deviation (red squares) for the same system, plotted vs $\sqrt{N_0}$ for $N_0 = 100, 200, 400, 800, 1600$. Both quantities scale linearly with $\sqrt{N_0}$, so the variance-mean ratio is finite even in the thermodynamic limit.

Then a new individual is introduced, whose two alleles are chosen at random from the gamete pool in which the fraction of a is now

$$\frac{fx^2 + x(1-x)}{fx^2 + 2x(1-x) + (1-x)^2}. \tag{14}$$

Although the distribution of the standardized variables is again narrow and appears to be N independent, it does not satisfy Gumbel statistics, as demonstrated in Fig. 2. More importantly, as demonstrated in the inset of Fig. 2, both the mean and the standard deviation scale with the square root of N_0 and therefore the width of the distribution is proportional to its mean even in the large-N_0 limit.

These examples suggest a general insight as to the N_0 scaling of the width of the distribution and its mean. As explained in the Appendix, the width reflects the effect of demographic stochasticity, which is relatively weak and becomes prominent only when the deterministic forces are tiny. A population undergoing demographic stochasticity and decline can be described by the Langevin equation

$$dn = -\kappa n^p dt - \sigma_d \sqrt{n} dW, \tag{15}$$

where κ is the decay coefficient (related to κ and f in the above examples), p is the power that characterizes the interaction between individuals in the dilute limit ($p = 1$ for exponential decay with no interactions and $p = 2$ for diploids with dominance), and σ_d is the amplitude of demographic variations. The last term in Eq. (15) becomes important only below $n < n_c$. In the Appendix we show that $n_c \sim N^{(p-1)/p}$ as long as $p > 1$, and n_c is $O(1)$ for $p \leqslant 1$. For $n < n_c$ the dynamics is neutral (discussed in Sec. V), so the contribution of this extinction zone (both to the mean time to extinction and to its standard deviation) is proportional to n_c. The regime $n < n_c$ is the only place in which demographic fluctuations are important, so the variance of the extinction-time distribution is n_c^2.

The mean time to extinction, on the other hand, is the sum of the deterministic timescale, i.e., the time required to decline from N_0 to n_c, and the stochastic period that scales with n_c. The deterministic timescale for the dynamics described by Eq. (15) is N_0^{p-1} for $p < 1$, $\log N_0$ for $p = 1$, and $O(1)$ for $p > 1$. Accordingly, the variance-mean ratio goes to zero if $p \leqslant 1$ (assuming both N_0 and N diverge together). For $p > 1$, the mean and the standard deviation both have the same scaling with $n_c \sim N^{(p-1)/p}$, so the variance-mean ratio diverges as N and N_0 go to infinity.

III. EXTINCTION-PRONE DYNAMICS IN A STOCHASTIC ENVIRONMENT

In this section we consider the scenario of an extinction-prone population influenced by environmental stochasticity. In that case, Lande and Orzack [25] and Dennis *et al.* [26] suggested an inverse Gaussian statistics of extinction times. Here we recapitulate their argument while detailing the differences between this system and the systems we surveyed in the preceding section and clarifying when and how to use it.

In what follows, the term “environment” encompasses any external factor that impacts the demographic rates of an entire population, including factors such as competition and/or predation pressure from other species. When the environment undergoes stochastic variations, the birth and death rates of the population also fluctuate. Consequently, the overall growth rate (birth rate minus death rate) experiences corresponding variations, leading the population to exhibit either growth or decay. The population is prone to extinction if its mean growth rate is negative [27,28].

Let us reemphasize the distinction between demographic and environmental stochasticity. Stochasticity, in both of its forms, can be attributed to fluctuations in the environment. The distinction between these two forms of stochasticity has to do with their range. When the mean demographic rates remain constant over time and the fluctuations affect individuals in an uncorrelated manner, it is considered demographic noise. On the other hand, if an entire population is affected by the stochasticity, it is classified as environmental stochasticity. Demographic noise is commonly characterized as white noise, where different birth or death events are uncorrelated in time. In contrast, the correlation time becomes a significant characteristic of environmental variations.

Specifically, let us consider a simple, purely environmental, two-state system (telegraphic noise). We assume that the environment may be in either of two states, say, state 1 and state 2. The environment remains in a particular state for a certain duration (referred to as the dwell time, which is considered the unit time of the process) before switching to the alternative state with a probability of $\frac{1}{2}$. In each of these states of the environment, the population either grows exponentially or decreases exponentially, so if the number of individuals is n, $\ln n$ increases or decreases linearly with time between switches.

When the number of individuals is large, demographic stochasticity is negligible with respect to environmental stochasticity [4]. Therefore, in many studies the effect of demographic stochasticity is taken into account only by introducing a threshold at a given density, below which the population is considered extinct. Recent analyses suggest that this threshold has to be taken to be the value of N at which the strength of demographic stochasticity is equal to the strength of environmental stochasticity [14,29].

Once demographic stochasticity is neglected, the dynamics of n is simply $n(t+\tau) = n(t)\exp(\zeta\tau)$, where τ is the dwell time and ζ is the (time-dependent) growth exponent (if the environment admits two states, ζ is either $\zeta_1 > 0$ or $\zeta_2 < 0$). Taking τ as the unit time, one arrives at

$$x_{t+1} = x_t + \zeta_t, \tag{16}$$

where $x = \ln n$. The random walk in x space is characterized by the mean and the variance of ζ, namely, $\kappa = \overline{\zeta} = (\zeta_1 + \zeta_2)/2$ and $\sigma^2 = \text{Var}[\zeta]$, where κ, the decline rate, is assumed to be negative.

When the initial population N_0 is large, the problem is mapped to the classical first-passage time for a biased random walker. Accordingly, if $x_0 = \ln N_0$ is the initial location of the random walker and $x_1 = \ln(N_{\text{th}})$ is the threshold density below which the population is considered extinct, the probability distribution function for the time required to cross the logarithmic-space distance $\Delta x = x_0 - x_1 = \ln(N_0/N_{\text{th}})$ is given by the inverse Gaussian distribution [25,26]

$$P(t) = \frac{\Delta x}{\sigma\sqrt{2\pi t^3}} e^{-(\Delta x - \kappa t)^2/2\sigma^2 t}. \tag{17}$$

The mean of this distribution is $\mathbb{E}[t] = \Delta x/\kappa$ and its variance $\text{Var}[t] = \mathbb{E}[t]\sigma^2/\kappa^2$. Therefore, the variance-mean ratio in that case is N_0 independent, $\text{VMR} = \sigma^2/\kappa^2$.

The chance of the system to survive until t (i.e., the cumulative distribution function) is given by

$$Q(t) = \frac{1}{2}\left[1 - \text{erf}\left(\frac{\kappa t - \Delta x}{\sqrt{2t\sigma^2}}\right) - e^{2\kappa\Delta x/\sigma^2}\text{erfc}\left(\frac{\kappa t + \Delta x}{\sqrt{2t\sigma^2}}\right)\right]. \tag{18}$$

In the case of exponential decay ($p = 1$) with pure demographic noise considered in Sec. II, the mean time to extinction is also logarithmic in the initial population size, but the variance and the higher cumulants are $O(1)$. Here both the mean and variance are linear in $\ln N$, so the distribution is much wider than the one that characterizes the purely demographic case. When the noise is demographic, its effect becomes non-negligible only when the number of individuals n is $O(1)$ (i.e., smaller than N_{th}), while for systems with environmental stochasticity the noise affects the system all the way down from N_0 to extinction, no matter how large N_0 is.

The given example focuses on a specific example, namely, telegraphic noise. However, at its core, the analysis considers the dynamics of a random walker (in the logarithmic-abundance space) with a bias. It can be shown (see, e.g., [28], Appendix) that, as long as the logarithmic-abundance steps are not excessively large, the diffusion approximation is applicable and the long-term characteristics of the dynamics are solely influenced by the mean and the variance of the $\zeta(t)$ process. Therefore, the results presented above remain valid.

IV. MARGINAL DYNAMICS WITH PURE DEMOGRAPHIC STOCHASTICITY: THE NEUTRAL MODEL

In Secs. II and III we focused on the persistence time statistics of populations prone to extinction. In the next two sections our aim is to examine the same question but with a focus on marginal populations. These populations are characterized by deterministic dynamics that support a marginally stable manifold, which includes the extinction state. A classic example is the case of competition between two populations or two types that possess identical fitness. For instance, consider two genotypes that differ only by a synonymous mutation, resulting in the same phenotype. In that case, and in general on a marginally stable manifold, stochasticity is the only driver of abundance variations.

The famous neutral models proposed by Kimura [15,30] in population genetics and Hubbell [16,17,31] in community ecology address such systems, where the main driver of dynamics (up to relatively rare mutation or speciation events) is demographic stochasticity. Under neutral dynamics, species identity is irrelevant. One can consider a single species as a focal species and pool over the effect of all other species together as a single entity (an effective rival species). Therefore, in what follows we derive the appropriate formulas for a single species within a community of N individuals, whose dynamics ends at one of the two absorbing states, i.e., the zero abundance state (extinction) or at abundance N (fixation).

The systems considered in previous sections admit deterministic decline dynamics, so ultimately the overall population never grows beyond its initial value N_0 and hence extinction times and extinction statistics are governed by N_0, the initial abundance, and not by N, the maximum carrying capacity. Under neutral dynamics, a population may either decline to extinction or grow to fixation and therefore N sets the relevant timescales. When $N \to \infty$ (no fixation), the time to extinction for N_0 individuals is given by the limit $\mu \to \lambda$ of Eq. (10) above. For finite N the value of N_0 affects the statistics only through its relationship with N, as explained below.

A. Case I: A macroscopic population

In a neutral model with pure demographic stochasticity, one considers the dynamics of $x = n/N$, where n is the number of individuals of a given focal species and N is the total number of individuals. In this section we assume that the initial frequency N_0/N is $O(1)$. We would like to obtain the statistics of absorption (either fixation or extinction) times where the dynamics of $P(x,t)$ is given by

$$\frac{\partial P(x,\tau)}{\partial \tau} = \frac{\partial^2}{\partial x^2}[x(1-x)P(x,t)],$$
$$P(0,\tau) = P(1,\tau) = 0, \; P(x,\tau = 0) = \delta\left(x - \frac{1}{2}\right). \tag{19}$$

Here τ is the dimensionless timescale t/N.

Defining

$$W(x,\tau)=x(1-x)P(x,\tau), \tag{20}$$

W satisfies

$$\frac{\partial W(x,\tau)}{\partial \tau}=x(1-x)\frac{\partial^2 W(x,t)}{\partial t^2}. \tag{21}$$

Taking $W(x,\tau)=W_m(x)e^{\lambda_m \tau}$, the equation for the eigenfunctions $W_m(x)$ and the eigenvalues λ_m is

$$\frac{\partial^2 W_m(x)}{\partial t^2}-\lambda_m\frac{W_m(x)}{x(1-x)}=0. \tag{22}$$

The general solution of (22) is a linear combination of two independent functions. One is a Meijer G function that diverges at the origin, so its contribution must vanish [since $P(x)$ vanishes at $x=0$ and at $x=1$, so does W]. Thus the solution, up to a constant, is given by the other solution, which vanishes at $x=0$,

$$W_m(x)=x\,{}_2F_1\big(\tfrac{1}{2}(1-\sqrt{1-4\lambda_m}),\tfrac{1}{2}(1+\sqrt{1-4\lambda_m});2;x\big), \tag{23}$$

where ${}_2F_1$ is the hypergeometric function.

The λ_m are determined by the condition $W(x=1)=0$, which yields

$$W_m(x)=\frac{\cos\left(\frac{1}{2}\pi\sqrt{1-4\lambda_m}\right)}{\pi\lambda_m}. \tag{24}$$

Therefore,

$$\lambda_m=-m(m+1). \tag{25}$$

Since m is an integer, the corresponding eigenfunction simplifies to

$$W_m(x)=x\,{}_2F_1(-m,m+1;2;x)=xP_m^{(1,-1)}\frac{1-2x}{m+1}, \tag{26}$$

where the $P_m^{(\alpha,\beta)}(x)$ are Jacobi polynomials. Accordingly, the general solution to Eq. (22) takes the form

$$W(x,\tau)=\sum_{m=1}^{\infty}A_mW_m(x)e^{-m(m+1)\tau}. \tag{27}$$

The $m=0$ (time-independent) term yields a non-normalizable probability function and therefore it has been excluded.

The constants A_m are determined by the initial condition. The orthogonality relationships of the Jacobi polynomials, when translated to functions of $1-2x$, are

$$\int_0^1\frac{x}{1-x}P_m^{(1,-1)}(1-2x)P_n^{(1,-1)}(1-2x)=\delta_{n,m}\frac{m+1}{m(2m+1)}. \tag{28}$$

To find A_m from $W(x,0)=x(1-x)\delta(x-\frac{1}{2})$ one multiplies both the left- and the right-hand side of this equation by $P_n^{(1,-1)}(1-2x)$, integrates over x from zero to one, and applies the relationship (28). That yields

$$A_m\frac{1}{m(2m+1)}=\begin{cases}0 & \text{if } m \text{ even}\\ \frac{P_m^{(1,-1)}(0)}{m+1}=\frac{(-1)^{m_1}C_{m_1}}{2^m} & \text{if } m \text{ odd,}\end{cases} \tag{29}$$

where $m=2m_1+1$ and $C_{m_1}=(2m_1)!/m_1!(m_1+1)!$ are the Catalan numbers.

The chance to survive until t, $Q(t)$, is given by the integral of $P(x)$ over x from zero to one. Using Eq. (27), the definition (20), the relationships between Jacobi polynomials and W, and the integral

$$\int_0^1 dx\,\frac{xP_m^{(1,-1)}(1-2x)}{x(1-x)}=\frac{2}{m}, \tag{30}$$

one finds

$$\begin{aligned}Q(t)&=\int_0^1 dx\sum_{m=0}^{\infty}A_{2m+1}\frac{W_{2m+1}(x)}{x(1-x)}e^{-(2m+1)(2m+2)\tau}\\&=\sum_{m=0}^{\infty}\frac{(-1)^{m+1}C_m}{2^{2m+1}}(4m+3)e^{-(2m+1)(2m+2)\tau}.\end{aligned} \tag{31}$$

Accordingly, the chance of extinction at τ, $\mathcal{P}(t)$, is

$$\begin{aligned}\mathcal{P}(t)=-\frac{dQ(t)}{dt}&=\frac{1}{N}\sum_{m=0}^{\infty}\frac{(-1)^mC_m}{2^{2m+1}}(2m+1)(2m+2)\\&\times(4m+3)e^{-(2m+1)(2m+2)t/N}.\end{aligned} \tag{32}$$

Figure 3 shows the correspondence between the predicted and the measured $\mathcal{P}(t)$.

Following the calculation that leads to Eq. (49) below, one obtains an expression for the asymptotic behavior of the nth moment of the extinction-time distribution

$$\overline{t^n}=B_nN^nn!\ln(2). \tag{33}$$

The numbers B_n are given by a complicated set of hypergeometric functions; however, $B_1=1$ and in general $B_n\approx\exp[-0.68(n-1)]$ provides an excellent approximation for the first ten moments, as demonstrated in Fig. 3. The mean and the variance are then

$$\bar{t}=N\ln 2,\qquad \mathrm{Var[t]}=(2B_2-1)(N\ln 2)^2, \tag{34}$$

so the VMR scales like N.

B. Case II: A single neutral mutant

Let us now consider the case of other initial conditions $W(x,0)=x(1-x)\delta(x-x_0)$ and in particular the survival-time distribution of a single mutant $x_0=1/N$. Now the general solution for $W(x,\tau)$ takes the form

$$\begin{aligned}W(x,\tau)=&\sum_{m=1}^{\infty}\frac{m(2m+1)}{m+1}x_0P_m^{(1,-1)}(1-2x_0)xP_m(1-2x)\\&\times e^{-m(m+1)\tau}.\end{aligned} \tag{35}$$

Dividing by $x(1-x)$ and integrating over x,

$$Q(x,\tau)=\sum_{m=1}^{\infty}\frac{(2m+1)}{m+1}x_0P_m^{(1,-1)}(1-2x_0)e^{-m(m+1)\tau}. \tag{36}$$

If N is large, for the dynamics of a single mutant ($x_0=1/N$) one may use the Mehler-Heine formula for the Jacobi polynomials [32],

$$P_m^{(1,-1)}(1-2/N)\approx\sqrt{N}J_1\left(\frac{2m}{\sqrt{N}}\right), \tag{37}$$

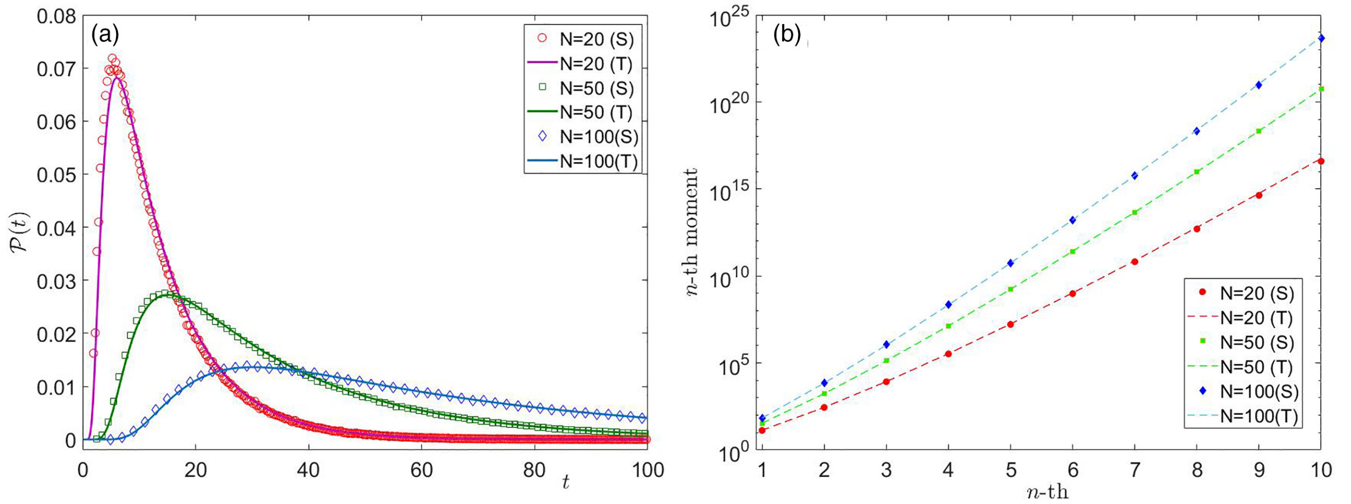

FIG. 3. (a) Extinction probability at t, $\mathcal{P}(t)$, plotted against t for populations of $N = 20$, 50, and 100. The initial condition is $n(t = 0) = N/2$, namely, $x = \frac{1}{2}$. The result of Eq. (32) (solid curves) are compared with the normalized distribution obtained numerically (open circles). In the numerical experiment, the chance of the focal population to increase, or to decrease, by one unit in each elementary step is $\frac{1}{2}$, and in each elementary step time is incremented by $1/2x(1-x)$. (b) Theoretical prediction for the moments [Eq. (33), dashed lines] compared with the moments of these distributions (closed markers).

where J_1 is the first Bessel function. Accordingly,

$$Q(x,\tau) \approx \frac{1}{N}\sum_{m=1}^{\infty}\frac{(2m+1)}{m+1}J_1\left(\frac{2m}{\sqrt{N}}\right)e^{-m(m+1)\tau}. \tag{38}$$

Since J_1 vanishes at zero, the small-m behavior yields a negligible contribution to the sum. This facilitates the approximation

$$Q(x,\tau) \approx \frac{2}{N}\int_{x=1}^{\infty}J_1\left(\frac{2x}{\sqrt{N}}\right)e^{-x^2t/N} = 1 - e^{-1/t} - \frac{1-e^{-t/N}}{t}, \tag{39}$$

so the chance of the lineage of a single mutant to reach extinction at t is

$$\mathcal{P}(t) = -\frac{dQ(t)}{dt} = \frac{e^{-1/t}}{t^2} - \frac{1-e^{-t/N}(1+t/N)}{t^2}. \tag{40}$$

The first moment may be obtained from this expression, and one gets $\bar{t} = \ln N + 1 - 2\gamma_E$, where γ_E is Euler's gamma. To get the higher moments we implement the procedure described above, $\overline{t^n} = -\int t^n (dQ/dt)dt$,

$$\overline{t^n} = n!N^{n-1}\sum_{m=1}^{\infty}\frac{2m+1}{m^n(m+1)^{n+1}}P_m^{(1,-1)}(1-2/N). \tag{41}$$

Since the main contribution comes from the small-m region, we can approximate $P_m^{(1,-1)}(1-2/N) \approx (m+1)$ and therefore

$$\overline{t^n} = n!N^{n-1}\sum_{m=1}^{\infty}\frac{2m+1}{m^n(m+1)^n}. \tag{42}$$

Figure 4 demonstrates the validity of these results.

Note that the time required for a single mutant to be absorbed follows a logarithmic scaling of $\ln N$, whereas the time for a macroscopic population scales linearly with N. Additionally, the variance of extinction times for a single mutant is $O(N)$, while for a macroscopic population, it scales with N^2. In general, the ratio between the moments described in Eq. (42) and the corresponding moments in Eq. (33) is the factor N. This characteristic highlights the fact that an individual either goes extinct within a timescale of $O(1)$ or, with a probability that scales like $1/N$, avoids extinction and achieves macroscopic population sizes.

V. MARGINAL DYNAMICS WITH ENVIRONMENTAL STOCHASTICITY: THE TIME-AVERAGED NEUTRAL MODEL

The neutral model, for which we provided the extinction analysis in the preceding section, was initially introduced by Kimura as a model describing competition between two alleles with equal fitness and later (with certain modifications) was implemented by Hubbell to describe the dynamics of an ecological community in which all species have equal fitness. Both variations of the model gained immense popularity. In particular, its community ecology version successfully explained the distribution of species abundance in high-diversity assemblages using a small number of parameters [15–17,31].

However, it seems that the neutral model fails to capture the dynamics of ecological communities. According to the neutral model, which contains only demographic stochasticity and so generates binomial noise, one expects the per-generation changes in abundance to be proportional to the square root of population size. In practice, changes in abundance are usually much larger [33] and usually scale with population size as expected in systems where stochasticity is environmental [5,6], not demographic. Similarly, the times to the most recent common ancestor proposed by the neutral model are way too long [34,35], also reflecting the unrealistic slowness of the demographically driven neutral dynamics.

To address these issues, the time-averaged neutral model of biodiversity was proposed [9,36,37]. This is essentially a neutral model with temporal environmental stochasticity. Like the original neutral model, the dynamics is purely stochastic, but in this model, the stochasticity has two sources,

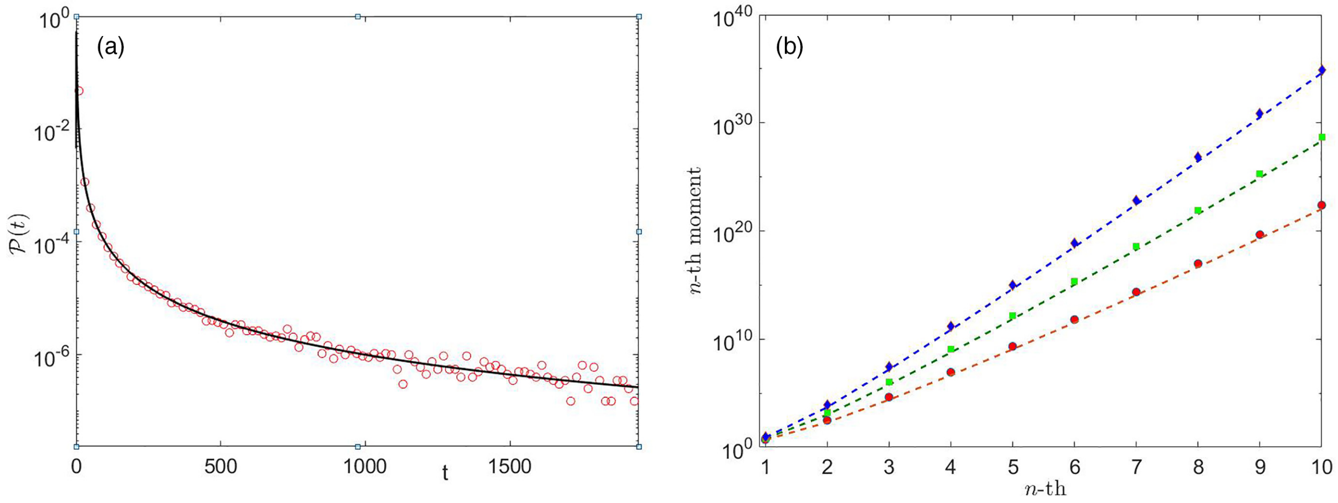

FIG. 4. (a) Extinction probability of a single mutant $[n(t=0)=1]$ at t, $\mathcal{P}(t)$, plotted against t for a population of $N = 100\,000$. The result of Eq. (40) (solid curve) is compared with the normalized distribution obtained numerically (open circles). (b) Theoretical prediction for the moments [Eq. (42), dashed lines] compared with the moments of these distributions (closed markers).

both demographic and environmental. All species have the same time-averaged fitness, but at any given moment, there are lower-fitness and higher-fitness species. This immediately leads to abundance variations that scale with population size, as expected, and the theory accounts for both static and dynamic patterns of community assembly [9].

In the following treatment, we consider a focal species representing a fraction x of the community, competing with all other species representing a fraction $1-x$ of the same community. Once again, we address the question of the distribution of times until the focal species reaches either extinction or fixation, this time under environmental noise. If we allow ourselves to neglect the demographic noise, by replacing it with an absorbing boundary condition for populations below a certain threshold, what we obtain is an unbiased random walk in logit ($z = \ln[x/(1-x)]$) space. Therefore, the problem reduces to the distribution of times for a one-dimensional simple random walk with absorbing boundary conditions. This problem was solved using a Laplace transform (see, e.g., [38] and references therein). Here we provide an alternative eigenfunction-based solution.

Mathematically speaking, we consider the dynamics of a population whose fraction $x = n/N$ satisfies $\dot{x} = \zeta(t)x(1-x)$, where $\zeta(t)$ is a zero-mean stochastic process whose variance is σ^2. Therefore, the logit variable $z \equiv \ln[x/(1-x)]$ is an unbiased random walk $z(t) = z_0 \int^t \zeta(t')dt'$. If the threshold fraction $x_{\text{th}} = N_{\text{th}}/N \ll 1$, the boundary conditions are, to the left, $z_{\text{th},L} \approx \ln N_{\text{th}}/N$ and, to the right, $z_{\text{th},R} \approx \ln N/N_{\text{th}}$. Since there is no bias, the specific values is not important and so we focus on the corresponding diffusion equation

$$\frac{\partial P(z,t)}{\partial t} = D\frac{\partial^2 P(z,t)}{\partial t^2},$$

$$P(0,t) = P(L,t) = 0, \;\; P(x,0) = \delta(z - L/2), \tag{43}$$

where $L = z_{\text{th},R} - z_{\text{th},L}$.

The problem is thus equivalent to the heat equation in a one-dimensional slab. The general form of the solution is

$$P(z,t) = \sum_{m=1}^{\infty} A_m \sin\left(\frac{m\pi z}{L}\right)e^{-\lambda_m t}, \tag{44}$$

where

$$\lambda_m = \frac{Dm^2\pi^2}{L^2}. \tag{45}$$

Thus, the solution that satisfies both the boundary and initial conditions is

$$P(z,t) = \sqrt{\frac{2}{L}}\sum_m (-1)^m \sin\left(\frac{(2m+1)\pi z}{L}\right)e^{-(2m+1)^2\tau}, \tag{46}$$

where $\tau \equiv \pi^2 Dt/L^2$.

The chance of the random walker to survive to time t, $Q(t)$, is

$$Q(t) = \int_0^L P(z,t)dx = \frac{4}{\pi}\sum_m \frac{(-1)^m}{2m+1}e^{-(2m+1)^2\tau}. \tag{47}$$

The chance of extinction at a given time t is $-dQ/dt$ and therefore the nth moment of t is given by

$$\overline{t^n} = -\int_0^\infty dt\, t^n \frac{dQ}{dt} = n\int_0^\infty dt\, t^{n-1}Q(t)$$

$$= n\left(\frac{L^2}{\pi^2 D}\right)^n \int_0^\infty d\tau\, \tau^{n-1}Q(\tau). \tag{48}$$

Evaluating the integral, one finds

$$\overline{t^n} = \frac{n!}{4^{2n}\pi}\left(\frac{L^2}{\pi^2 D}\right)^n\left[\zeta\left(2n+1,\frac{1}{4}\right) - \zeta\left(2n+1,\frac{3}{4}\right)\right], \tag{49}$$

where ζ is the Riemann zeta function. The agreement between these theoretical predictions and the outcomes of a standard Monte Carlo simulation is demonstrated in Fig. 5.

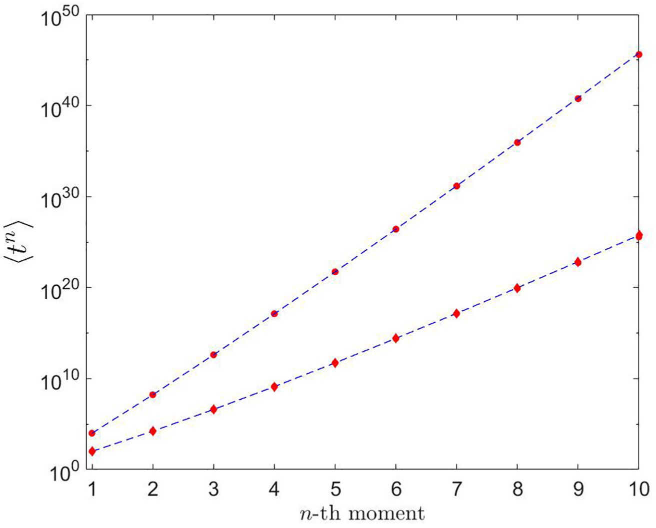

FIG. 5. The nth moment of the extinction time, for a random walker that started at $z = L/2$. In each step the random walker jumps to the left or to the right with probability $\frac{1}{2}$, and time is incremented by one unit. Moments were calculated for extinction times evaluated in 10^5 numerical experiments for $L = 200$ (circles) and $L = 20$ (diamonds). Dashed lines are the corresponding predictions from Eq. (49) with $D = \frac{1}{2}$.

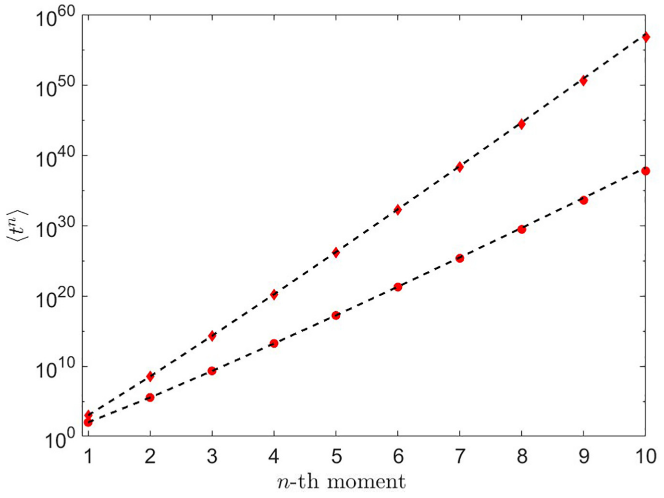

FIG. 6. The nth moment of the extinction time, for a random walker that started at $z_0 = 1$. In each step the random walker jumps to the left or to the right with probability $\frac{1}{2}$, and time is incremented by one unit. Moments were calculated for extinction times evaluated in 10^5 numerical experiments for $L = 100$ (circles) and $L = 1000$ (diamonds). Dashed lines are the corresponding predictions from Eq. (51) with $D = \frac{1}{2}$.

Here the general scaling of the nth cumulant is L^{2n}, so the mean time to absorption scales like $\ln^2 N$ and the variance like $\ln^4 N$. As in the case of neutral dynamics with pure demographic stochasticity, the VMR diverges as $N \to \infty$.

For generic initial conditions $P(z, 0) = \delta(z - z_0)$, Eq. (47) is replaced by

$$Q(t) = \int_0^L P(z,t)dz = \frac{4}{\pi}\sum_{m=0}^{\infty} \frac{\sin[(2m+1)\pi z_0/L}{2m+1} e^{-(2m+1)^2\tau}, \tag{50}$$

so

$$\overline{t^n} = \frac{n!}{4^{2n}\pi}\left(\frac{L^2}{\pi^2 D}\right)^n \sum_{m=0}^{\infty} \frac{\sin[(2m+1)\pi z_0/L]}{(2m+1)^{2n+1}}. \tag{51}$$

The main contribution to this sum, even for $n = 1$, comes from the small-m regime (see Fig. 6). When $z_0 \to 0$ (close to the absorbing boundaries) the argument of the sine function is negligibly small. Therefore, when $z_0 = \epsilon$ the moments are

$$\overline{t^n} \approx \frac{4n!\epsilon}{L}\left(\frac{L^2}{\pi^2 D}\right)^n (1 - 4^{-n})\zeta(2n). \tag{52}$$

Again there is a factor of $1/L$ between the single-mutant case and the macroscopic population case, because the chance of a single mutant to avoid extinctions on timescales that are $O(1)$ and reach macroscopic abundances is proportional to $1/L$.

VI. STABLE POPULATIONS

Now let us discuss systems that exhibit deterministic dynamics with an attractive fixed point capable of supporting large population. One example is the logistic system described by the equation $dn/dt = rn(1 - n/K)$, where $r > 0$ (throughout this section we refer to K as the number of individuals in the equilibrium state). In such cases, the occurrence of extinctions, even in the presence of stochastic fluctuations, is relatively rare. We can think of the stochastic process as a random walk biased towards the equilibrium state. For extinction (or approaching the zero-population point) to happen, the random walker would need to take numerous steps against the current, an event with an extremely low probability.

The stochastic dynamics of a stable system is some sort of a random walk biased away from the extinction point. Given that, the path to extinction consists of a series of implausible steps, where any plausible step leads to an increase in the population size. Therefore, the most probable decay path is composed of a consecutive sequence of these implausible steps. Under pure demographic stochasticity, this series requires K consecutive death events without any birth event, and the likelihood of this decreases exponentially as $\exp(-c_1 K)$, where c_1 is some coefficient. In cases where environmental variations allow for periods of negative growth rate, the most probable path to extinction involves a long period T of adverse conditions. The duration T scales logarithmically with K, resulting in the frequency of extinctions, which is exponentially rare in T, decaying as a power-law function of K. These arguments were extensively discussed and presented in detail in [27,28].

However, beyond the differences in the scaling of the average extinction time with K, stable systems have a common characteristic that determines the distribution around that mean. As mentioned, the extinction event is a rare fluctuation and the typical timescale associated with the decline, $T_d \sim \ln K$, is much shorter than the persistence time of a system in the asymptotic limit of large K. This separation of timescales, between the decline time and the persistence time, allows us to treat this stochastic process as a binomial process in which, during each increment T_d, an extinction

TABLE I. Summary of the main results.

Scenario	Mean	Variance	Distribution
extinction prone, demographic, density independent ($p = 1$)	$\ln N_0$	$O(1)$	Gumbel [Eq. (12)]
extinction prone, demographic, $p > 1$	$N^{(p-1)/p}$	$N^{2(p-1)/p}$	skewed (Fig. 2)
extinction prone, environmental	$\ln N_0$	$\ln N_0$	inverse Gaussian [Eq. (17)]
neutral, demographic, macroscopic population	N	N^2	Eq. (32)
neutral, demographic, single mutant	$\ln N$	N	Eq. (40)
neutral and environmental stochasticity, macroscopic population	$\ln^2 N$	$\ln^4 N$	Eq. (46)
neutral and environmental stochasticity, single mutant	$\ln N$	$\ln^3 N$	Eq. (51)
stable population, demographic stochasticity	$\exp(K)$	$\exp(2K)$	exponential[a]
stable population, environmental stochasticity	K^α (power law)	$K^{2\alpha}$	exponential[b]

[a] From Refs. [39–41].

[b] From Refs. [27,28].

event occurs with a tiny probability. If extinction does not happen, even if it almost happens (the population declines to small abundance), the system recovers and returns to its equilibrium state. Therefore, the lifetime distribution of stable systems is simply an exponential distribution with an average equal to the average persistence time, as shown in [28].

Mathematically, extensive efforts have been made to calculate the mean time to extinction and determine its numerical value, including the coefficient c_1 mentioned earlier or the prefactor of the exponential term [39–42]. These studies have revealed that the spectrum of the Markov matrix governing such a stochastic process exhibits several interesting properties. First, it supports an extinction state whose decay rate (logarithm of its eigenvalue) is zero, indicating the absorbing nature of the extinction state. Second, there exists a single quasistationary state whose decay rate decreases to zero as K increases. Finally, the decay rates of all other eigenstates are $O(1)$, independent of K. These results are in agreement with the qualitative picture illustrated above: starting from an arbitrary initial state, which is a linear combination of many eigenstates of the corresponding Markov matrix, the system converges to the quasistationary state on timescales that are $O(1)$ and then the survival probability decays exponentially.

VII. SUMMARY AND DISCUSSION

In this paper we have discussed the extinction-time statistics in various generic scenarios. The main results we derived or quoted are summarized in Table I. In the mean and in the variance columns of this table we provided only the dependences of the times on the relevant large parameter, be it the initial population size N_0, the total population N, or the population at the attractive fixed point K.

It is worth dwelling on this point: the determining factor, be it K, N_0, or N. In a stable system, this factor is K, the number of focal species individuals in the stable state. It is independent of the initial population size N_0, because the system usually flows towards the stable state. Similarly, it has nothing to do with the total carrying capacity N (how many total individuals, regardless of species, are allowed in the system).

In marginal and neutral systems there is no specific abundance for a particular species. Accordingly, the determining factor is the total carrying capacity of the system, N, because every species has a non-negligible chance of reaching it regardless of its initial size. In contrast, in an exponentially decaying system, the initial condition N_0 is the only important factor since the population does not generally increase in size.

An exceptional case is when a population undergoes density-dependent extinction dynamics, as demonstrated in the diploid with dominance dynamics. In this case, the deterministic extinction time depends only weakly on the initial population size. Therefore, the factor that governs extinction times is the width of the fluctuation-dominated region, where the system exhibits neutral behavior. Consequently, in these cases ($p > 1$) the important quantity is again N, since it determines the width of the stochastically dominated zone.

The width of the distribution and the variance-mean ratio are governed by the stochastic part of the dynamics. When the origin of these fluctuations is demographic and the deterministic forces take the system to extinction, these fluctuations are important only in a narrow region around zero ($n < n_c$). In the Gumbel case, or by and large when $p \leqslant 1$, this implies that the variance-mean ratio goes to zero in the thermodynamic limit. When $p > 1$ two things happen. First, n_c is proportional to N, and second, the time required to reach n_c, starting from N_0, is $O(1)$. Therefore, the properties of the distribution of extinction times when $p < 1$ are more or less identical to the corresponding properties of a neutral system with $N \sim n_c$.

A significant number of experimental [43–45] and empirical [46–49] studies have been dedicated to investigating the distribution of extinction times. However, in order to interpret these results in the context of the archetypal models discussed in this paper, further analysis is required. Nevertheless, we believe that this article can serve as a point of reference for future analyses of extinction statistics. The key characteristics observed in each study of extinction times, such as their dependence on initial conditions or carrying capacity, first moments, and variance-mean ratio, can provide valuable insights for classifying the basic dynamics of the system. This classification can then facilitate more detailed examinations, revealing other, system-specific features. Together, these valuable insights possess the potential to significantly enhance our comprehension of the underlying mechanisms that drive extinctions. Such knowledge can play a pivotal role in bolstering conservation efforts and guiding strategic approaches aimed at safeguarding biodiversity and promoting ecosystem stability.

ACKNOWLEDGMENTS

We would like to express our gratitude to S. P. Ellner for bringing to our attention the experimental works of Drake and Griffen.

APPENDIX: WIDTH OF THE EXTINCTION ZONE

In Sec. II we considered the distribution of extinction times for extinction-prone population with demographic stochasticity. In this Appendix we present a general argument that allows us to estimate the variance of this distribution.

During the process of extinction, the population is influenced by deterministic forces that drive it towards zero, as well as demographic stochasticity. We can define a critical population size, denoted by n_c, above which the deterministic forces dominate, rendering stochasticity negligible. Below n_c the population dynamics is essentially neutral, but the population cannot escape to $n > n_c$ due to the dominant deterministic forces.

In the stochastic regime, once the system reaches n_c, both the time to extinction (measured in generations) and its variance scale with n_c. Consequently, the time to extinction can be divided into two components: the deterministic time required for the population to transition from its initial state to n_c, which produces little to no variance, and the stochastic time with a mean and variance proportional to n_c.

Therefore, the crucial step in getting semiquantitative insight regarding the gross features of the extinction-time distribution is to estimate n_c. This may be done in several ways. Here we implement a dominant balance approach to the backward Kolmogorov equation (BKE).

Let us begin with the simplest case of an exponentially decaying population. The BKE, as derived in [50], for example, is

$$T''(x) - \kappa N T' = -\frac{N}{x}. \tag{A1}$$

This equation was derived for a two-species competition in a community of N individuals, when $x \ll 1$ is the fraction of focal species and T is the mean time to extinction. Here κ is the selection parameter, and when $\kappa < 0$ the focal species population declines exponentially.

Clearly, the T' term corresponds to the deterministic decline and the T'' term represents stochasticity. If we neglect the stochastic term, $T' = 1/\kappa x$ and therefore $T'' = -1/\kappa x^2$. The stochastic term thus dominates when

$$\frac{1}{\kappa x^2} > \kappa N T' = \frac{N}{x}, \tag{A2}$$

i.e., the stochastic regime is below $x_c = n_c/N = 1/\kappa N$, or $n_c = 1/\kappa$. Therefore, for large N_0 the mean time to extinction scales like $\ln N_0/\kappa$ (this is the timescale required for a population that satisfies $\dot{N} = -\kappa N$ to decline below a certain small value) plus an extinction time that scales like $1/\kappa$ and therefore is negligible when N_0 is large. On the other hand, the contribution for the variance comes only from the stochastic regime and, following Sec. V, must scale like $1/\kappa$. These two predictions yield the correct scaling for the parameters ν and β in Sec. II A.

For the diploid with dominance problem of Sec. II B the relevant BKE is

$$T''(x) - \kappa N x T' = -\frac{N}{2x}. \tag{A3}$$

Now $T' \sim 1/\kappa x^2$ and therefore $T'' \sim 1/\kappa x^3$. The first term thus becomes equal to the second term at $x_c = 1/\sqrt{\kappa N}$, so $n_c = \sqrt{N/\kappa}$.

Note that the deterministic time in that case is $O(1)$ (N_0 independent) and therefore both the mean and the variance scale, in the thermodynamic limit, like $\sqrt{N}$. Extending this argument, one finds that for deterministic dynamics that satisfies $\dot{x} = -\kappa x^p$, the variance scales like $N^{(p-1)/p}$. The deterministic timescale is $O(1)$ if $p > 1$ and scales like N_0^{p-1} for $p < 1$. Note that at $p \to \infty$ the stochastic timescale approaches N, since the dynamics becomes neutral.

[1] M. Dornelas, N. J. Gotelli, B. McGill, H. Shimadzu, F. Moyes, C. Sievers, and A. E. Magurran, Science **344**, 296 (2014).

[2] A. Gonzalez, B. J. Cardinale, G. R. Allington, J. Byrnes, K. Arthur Endsley, D. G. Brown, D. U. Hooper, F. Isbell, M. I. O'Connor, and M. Loreau, Ecology **97**, 1949 (2016).

[3] D. R. Hekstra and S. Leibler, Cell **149**, 1164 (2012).

[4] R. Lande, S. Engen, and B.-E. Saether, *Stochastic Population Dynamics in Ecology and Conservation* (Oxford University Press, Oxford, 2003).

[5] M. Kalyuzhny, E. Seri, R. Chocron, C. H. Flather, R. Kadmon, and N. M. Shnerb, Am. Nat. **184**, 439 (2014).

[6] M. Kalyuzhny, Y. Schreiber, R. Chocron, C. H. Flather, R. Kadmon, D. A. Kessler, and N. M. Shnerb, Ecology **95**, 1701 (2014).

[7] R. A. Chisholm, R. Condit, K. A. Rahman, P. J. Baker, S. Bunyavejchewin, Y.-Y. Chen, G. Chuyong, H. Dattaraja, S. Davies, C. E. Ewango *et al.*, Ecol. Lett. **17**, 855 (2014).

[8] A. O. Bergland, E. L. Behrman, K. R. O'Brien, P. S. Schmidt, and D. A. Petrov, PLoS Genet. **10**, e1004775 (2014).

[9] M. Kalyuzhny, R. Kadmon, and N. M. Shnerb, Ecol. Lett. **18**, 572 (2015).

[10] J. Grilli, Nat. Commun. **11**, 4743 (2020).

[11] L. Pechenik and H. Levine, Phys. Rev. E **59**, 3893 (1999).

[12] I. Dornic, H. Chaté, and M. A. Muñoz, Phys. Rev. Lett. **94**, 100601 (2005).

[13] M. Assaf and B. Meerson, J. Phys. A: Math. Theor. **50**, 263001 (2017).

[14] J. Pande, Y. Tsubery, and N. M. Shnerb, Ecol. Lett. **25**, 1783 (2022).

[15] M. Kimura, *The Neutral Theory of Molecular Evolution* (Cambridge University Press, Cambridge, 1985).
[16] S. P. Hubbell, *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton University Press, Princeton, 2001).
[17] I. Volkov, J. R. Banavar, S. P. Hubbell, and A. Maritan, Nature (London) **424**, 1035 (2003).
[18] D. Hathcock and S. H. Strogatz, Phys. Rev. E **100**, 012408 (2019).
[19] D. Hathcock and S. H. Strogatz, Phys. Rev. Lett. **128**, 218301 (2022).
[20] D. G. Kendall, Ann. Math. Stat. **19**, 1 (1948).
[21] R. A. Fisher and L. H. C. Tippett, Proc. Cambridge Philos. Soc. **24**, 180 (1928).
[22] J. Haldane and S. Jayakar, J. Genet. **58**, 237 (1963).
[23] A. Dean and N. M. Shnerb, Ecology **101**, e03098 (2020).
[24] S. Karlin and B. Levikson, Theor. Popul. Biol. **6**, 383 (1974).
[25] R. Lande and S. H. Orzack, Proc. Natl. Acad. Sci. USA **85**, 7418 (1988).
[26] B. Dennis, P. L. Munholland, and J. M. Scott, Ecol. Monogr. **61**, 115 (1991).
[27] Y. Yahalom and N. M. Shnerb, Phys. Rev. Lett. **122**, 108102 (2019).
[28] Y. Yahalom, B. Steinmetz, and N. M. Shnerb, Phys. Rev. E **99**, 062417 (2019).
[29] Axel G. Rossberg *et al.*, arXiv:2111.03867.
[30] W. J. Ewens, *Mathematical Population Genetics 1: Theoretical Introduction* (Springer Science + Business Media, New York, 2012), Vol. 27.
[31] S. Azaele, A. Maritan, S. J. Cornell, S. Suweis, J. R. Banavar, D. Gabriel, and W. E. Kunin, Methods Ecol. Evol. **6**, 324 (2015).
[32] M. Abramowitz, I. A. Stegun, and R. H. Romer, Am. J. Phys. **56**, 958 (1988).
[33] E. G. Leigh, J. Evol. Biol. **20**, 2075 (2007).
[34] S. Nee, Funct. Ecol. **19**, 173 (2005).
[35] R. E. Ricklefs, Ecology **87**, 1424 (2006).
[36] M. Danino and N. M. Shnerb, Phys. Rev. E **97**, 042406 (2018).
[37] J. Pande and N. M. Shnerb, J. Theor. Biol. **539**, 111053 (2022).
[38] S. Redner, *A Guide to First-Passage Processes* (Cambridge University Press, Cambridge, 2001), p. 49.
[39] V. Elgart and A. Kamenev, Phys. Rev. E **70**, 041106 (2004).
[40] M. Assaf and B. Meerson, Phys. Rev. Lett. **97**, 200602 (2006).
[41] D. A. Kessler and N. M. Shnerb, J. Stat. Phys. **127**, 861 (2007).
[42] A. Kamenev, B. Meerson, and B. Shklovskii, Phys. Rev. Lett. **101**, 268103 (2008).
[43] J. M. Drake, Ecology **87**, 2215 (2006).
[44] B. D. Griffen and J. M. Drake, Proc. R. Soc. B **275**, 2251 (2008).
[45] J. M. Drake and B. D. Griffen, Nature (London) **467**, 456 (2010).
[46] H. L. Jones and J. M. Diamond, Condor **78**, 526 (1976).
[47] G. Ferraz, G. J. Russell, P. C. Stouffer, R. O. Bierregaard Jr, S. L. Pimm, and T. E. Lovejoy, Proc. Natl. Acad. Sci. USA **100**, 14069 (2003).
[48] D. Matthies, I. Bräuer, W. Maibom, and T. Tscharntke, Oikos **105**, 481 (2004).
[49] E. Bertuzzo, S. Suweis, L. Mari, A. Maritan, I. Rodríguez-Iturbe, and A. Rinaldo, Proc. Natl. Acad. Sci. USA **108**, 4346 (2011).
[50] M. Danino, D. A. Kessler, and N. M. Shnerb, Theor. Popul. Biol. **119**, 57 (2018).