BRANCHING RANDOM WALKS IN RANDOM ENVIRONMENT: A SURVEY

By Wolfgang König[1]

WIAS Berlin and TU Berlin

(14 October, 2020)

We consider branching particle processes on discrete structures like the hypercube in a random fitness landscape (i.e., random branching/killing rates). The main question is about the location where the main part of the population sits at a late time, if the state space is large. For answering this, we take the expectation with respect to the migration (mutation) and the branching/killing (selection) mechanisms, for fixed rates. This is intimately connected with the parabolic Anderson model, the heat equation with random potential, a model that is of interest in mathematical physics because of the observed prominent effect of intermittency (local concentration of the mass of the solution in small islands). We present several advances in the investigation of this effect, also related to questions inspired from biology.

Keywords. multitype branching random walk; random potential; parabolic Anderson model; Feynman-Kac-type formula; annealed moments; large deviations.

MSC2020. 60J80; 60J55; 60F10; 60K37.

1. Introduction

In this chapter we study a topic that is of interest both in mathematics and in mathematical population biology: branching random walks on a graph that either models the space in which we live or a space of genotypes. That is, the particles are under the influence of three random mechanisms: movement, branching and killing. In the case of genotypes, the movement can be understood as mutation, and the killing as selection, hence such models belong to the cornerstones of the mathematical description of random population processes. We introduce this mathematical model in Section 2 and explain its biological interpretation in Section 2.2.

The main point that we are interested in here is the situation where both the branching rates and the killing rates may depend on the state that they are attached to, and that these rates are taken as random, typically independent and identically distributed. In this case, the expected number of particles in all the sites at a given time is the solution to the heat equation with

[1]Technische Universität Berlin, Str. des 17. Juni 136, 10623 Berlin, and Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39, 10117 Berlin, Germany, koenig@wias-berlin.de. This text will appear as a chapter in the proceedings volume of the DFG Priority Programme 1590 *Probabilistic Structures in Evolution*

potential, which is now random. This model is called the *parabolic Anderson model* (PAM) and has high physical significance also in other parts of applied mathematics because of the interesting effect of a high concentration property called *intermittency*. All this is explained in Section 2.

In Section 3 we introduce a time-discrete version of the model that has two state spaces: one describes the location, the other describes the type of the particles. We develop a version of a Feynman–Kac formula in this setting and derive the large-time asymptotics for the expected total mass of the branching process for a particular disctribution of the fitness variables. The formulas that describe these asymptotics contain interesting information about the counterplay of the two state spaces; mostly we are interested in the description of the limiting structures that the main part of the particles is sitting in.

A slightly different question is investigated in Section 4: For a particularly simple fitness distribution on a finite but large graph, how much time is needed for the main part of the population to move to the fittest site?

In Section 5, we extend the setting of the parabolic Anderson model (which describes the first moment of the number of particles) to a Feynman–Kac type formula for all the moments of the number of particles, taken over migration, branching and killing, valid for all choices of the branching/killing rates. This formula is employed afterwards to derive a large-time asymptotics for the high moments, for a particular potential distribution.

In Section 6 we mention and motivate future work on extensions of the study of branching random walks in random environment, like the study of the high-moment asymptotics, the parabolic Anderson model on random tree-like graphs and the effect coming from inserting a mutually repellent force between the particles.

2. Spatial branching random walks in random environment

We introduce the basic model in Section 2.1, explain its interpretation for evolutionary applications in Section 2.2, its relation with the parabolic Anderson model in Section 2.3 and give a brief account on the types of limiting statements available in the literature in Section 2.4.

2.1. The model.

Let $\mathcal{X}$ be some graph and $X = (X_t)_{t\in[0,\infty)}$ be a Markov process in continuous time on $\mathcal{X}$, which is most often taken as the simple random walk. The main examples for the state space $\mathcal{X}$ are the infinite space $\mathbb{Z}^d$ or the finite hypercube $\mathcal{X} = \{1,\dots,K\}^n$ for some $K, n \in \mathbb{N}$. Let $\xi_+ = (\xi_+(x))_{x\in\mathcal{X}}$ and $\xi_- = (\xi_-(x))_{x\in\mathcal{X}}$ be two collections of nonnegative numbers, attached to the sites of the state space. Now the spatial branching random walk with potential (ξ_+, ξ_-) is introduced as follows. Starting with a single particle at time zero at the origin $\mathcal{O} \in \mathcal{X}$, the particle runs like a copy of X through $\mathcal{X}$. Furthermore, when located at x, the particle splits into two particles with rate $\xi_+(x)$ and dies (is removed from the system) with rate $\xi_-(x)$. Hence $\xi_+(x)$ is the *branching rate* at x, and $\xi_-(x)$ is the *killing rate*. The two newborn copies proceed independently in the same way as the initial particle.

This is one of the basic models for the time-evolution of a population in a space with several random mechanisms: *migration*, *branching* and *killing*. The fields of rates, ξ_+ and ξ_-, introduce disorder in the latter two mechanisms; indeed we will consider them here often as being random and independent and identically distributed. Areas with high values of ξ_+ will produce many offspring over large time, and areas with high values of ξ_- will contain fewer particles

over long time. Since the $\xi_+(x)$ are rates, the number of particles at x at time t is likely to grow exponentially fast with rate $\xi_+(x)$ (neglecting the killing mechanism), but the migration mechanism diminishes this effect, since the many newborn particles have the tendency to move away from this site (but may return later). See [6] for more explanations about random walks with branching and killing in a random medium, in particular in connection with equations of the type that we will be examining in Sections 2.2 and 2.3.

The space $\mathcal{X}$ is often taken as (a discrete version of) the real space in which we live, i.e., the Euclidean space $\mathbb{Z}^3$, and many investigations of the model concentrate on this interpretation. However, from the viewpoint of evolution of a population in a biological sense, it is also highly interesting to view $\mathcal{X}$ as a space of genotypes, i.e., as a space of more abstract, biological properties of the particles. One particularly descriptive example is the choice of $\mathcal{X}$ as the set of *gene sequences*, representing each particle by its genome, which is a sequence composed of four or, for simplicity, two alleles. Here the hypercube $\mathcal{X} = \{-1, 1\}^N$ is the most canonical choice. In Section 2.2 we explain the interpretation of our model from the viewpoint of biological evolution.

The standard reflex of a probabilist is to choose the migration mechanism as the nearest-neighbour simple random walk, where two gene sequences are considered neighbours if they differ in precisely one allele. We will consider this in Section 6.2, but let us note that there are more biologically inspired migration mechanisms that respect the way in which mutation and selection of gene sequences really happen [3]. A mathematical investigation of the above branching model with this type of random motion is widely open and would be rather interesting and well-motivated.

A spatial branching process model (called the *logistic branching random walk*) without randomness in the birth rates, but a kind of birth control in the case of a local overpopulation is considered in [5, Section 3] in this volume. Here the birth rate in a given site is bounded, and is small (possibly zero) if the current number of individuals in a neighbourhood is large. The main results reported there are criteria for survival and for long-time distributional convergence to equilibrium, conditioned on survival. Proofs are based on a comparison technique using coupling.

2.2. **The heat equation and a mutation-selection population model.** The model of branching random walks introduced in Section 2 on the hypercube $\mathcal{X} = \{-1, 1\}^N$ with random branching rates has a profound interpretation in terms of a mutation-selection population model in a random fitness landscape. (See [14] for theoretical background about the notion of fitness in biological context and about the numbers and lengths of paths between sites of different fitnesses.) In [2], this is described as follows.

The mutation-selection model is given by the solution $v_N(t, \cdot, y)$ of the partial differential equation (PDE)

$$v_N(t, x, y) = \frac{1}{N}\Delta v_N(t, x, y) + [\xi(x) - \overline{\xi}(t)]\, v_N(t, x, y), \qquad t \in [0, \infty), x \in \{-1, 1\}^N, \tag{2.1}$$

with the localised initial condition $v_N(t, \cdot, y) = \delta_y(\cdot)$. The potential $\xi = \xi_+ - \xi_-$ is interpreted as a *fitness landscape*, and the mean fitness is given by

$$\overline{\xi}(t) = \sum_{x \in \{-1,1\}^N} v_N(t, x, y)\xi(x). \tag{2.2}$$

Here Δ is the *Laplace operator* defined by

$$\Delta f(x) = \sum_{y\in\mathcal{X}\colon y\sim x} \big(f(y) - f(x)\big)\,. \tag{2.3}$$

Now, (2.1) is not a particle model, but a PDE; however, let us already note that such types of PDEs are very suitable for describing the expectation of the number of particles in a spatial branching process for fixed branching rates. This will be explained in a broader perspective in Section 2.5.

Let us briefly explain the biological meaning of (2.1). Haploid genotypes are identified with linear arrangement of N sites $x = (x(1), \dots, x(N))$ with each site taking values -1 or $+1$. In the multilocus context, sites correspond to loci and the variables $x(i)$ to alleles. In the context of molecular evolution, x corresponds to a DNA (or RNA) sequence, where the nucleotides are lumped into purines (say, $+1$) and pyrimidines (say,-1). In the biology literature, the hypercube is usually called the sequence space. Then the mutation-selection model in (2.1) describes the evolution of an infinite population of haploids that experience only mutation and selection. The population evolves in continuous time (non-overlapping generations) with mutation and selection occurring independently (in parallel). $\xi(x)$ is the Malthusian fitness of type x and the $\xi(x)$ form a fitness landscape, which in our case is random. Site mutations happen with rate $1/N$ (hence, at a total rate of one). From (2.2) it follows that $\sum_x v_N(t,x,y) = 1$, and $v_N(t,x,y)$ corresponds to the frequency of type x under this evolution. Finally, note that the localised initial condition means that initially the population consists of only type y.

The competition between diffusion and potential in the PAM translates into a competition between mutation and selection, two driving forces of Darwinian evolution. We refer to the classical book [7] for an introduction to population genetics and to [9] for an excellent survey that involves the statistical physics methods used to solve mutation-selection models for a wide range of landscapes.

In the following, we will concentrate on the situation where the fitness landscape is random. The motivation for this is the following. Realistic landscapes are expected to be complex with structures such as valleys and hills. Random fitness landscapes naturally form a class of complex landscapes. The first obvious choice, that is, an i.i.d. landscape, is also known as the *house of cards* model.

Let us also mention that it is well-known [15] that

$$v_N(t,x,y) = \frac{u_N(t,x,y)}{u_N(t,x)}, \tag{2.4}$$

where u_N is the solution to (2.1) with $\overline{\xi}$ replaced by zero, which is called the *heat equation* that we are going to introduce in Section 2.3. In a way, $u_N(t,x,y)$ can be thought of as an absolute frequency.

2.3. **The parabolic Anderson model.** Let us return to the branching random walk model introduced in Section 2.1 with fixed branching rates $\xi_+(x)$ and killing rates $\xi_-(x)$. We want to explain this model from the viewpoint of PDEs, like in (2.1). For the time being, it is inessential that the rates are random. By $\eta(t,x)$ we denote the number of particles at time t at site x, and by $u(t,x)$ its expectation with respect to migration, branching and killing. Note that the expectation of $\eta(t,x)$ does not depend on ξ_+ and ξ_-, but only on the difference field

$\xi = \xi_+ - \xi_-$, since the branching and killing balances out in the expectation. We can therefore view $u(t,\cdot)$ as a time-dependent field that depends on $\xi = (\xi(x))_{x\in\mathcal{X}}$.

As a matter of fact, the time-dependent field $u(t,\cdot)$ satisfies the *heat equation with potential* $\xi = \xi_+ - \xi_-$, which is the equation

$$\begin{aligned} \partial_t u(t,x) &= \Delta u(t,x) + \xi(x)u(t,x), \qquad && x\in\mathcal{X},\, t\in(0,\infty),\\ u(0,x) &= \mathbb{1}_0(x), && x\in\mathcal{X}. \end{aligned} \tag{2.5}$$

Here Δ is the generator of the Markov process X, i.e., in case of a simple random walk on the graph $\mathcal{X}$, it is the Laplace operator defined in (2.3).

The solution theory for the heat equation is rather rich and explicit. Indeed, the solution u to (2.5) can be represented in terms of the *Feynman–Kac formula*

$$u(t,x) = \mathbb{E}_{\mathcal{O}}\Big[\mathrm{e}^{\int_0^t \xi(X_s)\,\mathrm{d}s}\mathbb{1}\{X_t = x\}\Big], \tag{2.6}$$

and in terms of an eigenvalue expansion for the eigenvalues $\lambda_1 > \lambda_2 \geq \lambda_3 \geq \dots$ of the operator $\Delta + \xi$,

$$u(t,x) = \sum_k \mathrm{e}^{t\lambda_k}\phi_k(\mathcal{O})\phi_k(x), \tag{2.7}$$

where $(\phi_k)_k$ is an orthonormal sequence of corresponding eigenfunctions.

So far, the field ξ of differences of branching and killing rates collects just the coefficients of the PDE in (2.5), and their randomness is inessential for the above. But let us turn now to the case that ξ is a *random* collection of numbers, which we then call a *random potential*. In this case, the operator $\Delta + \xi$ and its spectrum are also random; it is called the *Anderson operator* and plays an important role in mathematical physics; in particular the famous phenomenon of Anderson localisation has kept the interest in the description of its spectrum awake for decades. The heat equation with random potential in (2.5) is then called the *parabolic Anderson model*.

As we are interested in the behaviour of $u(t,\cdot)$ for large t, we will be concerned only with the boundary part of the spectrum, mostly only with the principal eigenvalue λ_1 (this is easily seen from (2.7)). Anderson localisation is, roughly speaking, the phenomenon that all the eigenfunctions of the Anderson operator $\Delta + \xi$, at least the ones corresponding to eigenvalues close to the boundary of the spectrum, are highly concentrated in small areas that are randomly located. The principal eigenvalue is included in this statement. Again looking at (2.7), we can guess that this concentration property is inherited by the random function $u(t,\cdot)$. This has been shown to be true in some generality, and this is the basis of a detailed understanding of the large-time behaviour of the entire branching process in a random potential of branching/killing rates.

The main assumption on the random potential ξ is that it is a collection of independent, identically distributed (i.i.d.) random variables. Many distributions can be considered, and for many applications there is no canonical choice. Not even ubiquitous distributions like the standard normal distribution or the exponential distribution can be considered canonical from the viewpoint of applications nor from the viewpoint of interesting emerging effects in this model. Rather, it turned out in the study of the moments of u that the most natural choice in this respect is perhaps the *double-exponential distribution*,

$$\mathrm{Prob}\big(\xi(0) > u\big) = \mathrm{e}^{-\mathrm{e}^{u/\varrho}}, \qquad u\in\mathbb{R}, \tag{2.8}$$

since it turned out in several investigations that the emerging picture in the long-time behaviour of the process has a non-trivial and not too complicated structure.

Our main interest here is in the description of the long-time (i.e., large-t) behaviour of the random field $u(t,\cdot)$, in particular of its *total mass*

$$U(t) = \sum_{x\in\mathcal{X}} u(t,x). \tag{2.9}$$

Of particular interest is the question whether $u(t,\cdot)$ develops some (randomly located) areas with particularly high values, i.e., particularly high accumulation of particles. Such areas are often called *intermittent islands*. Further questions concern the size and the location of these islands and characterisations of the potential $\xi(\cdot)$ and the solution $u(t,\cdot)$ inside them.

2.4. **Asymptotics.** Let us explain some of the fundamental results on the large-t behaviour of the parabolic Anderson model introduced in Section 2.5. See [12] for a comprehensive survey on the research until 2016 on the PAM, which was almost exclusively done on the Euclidean space $\mathbb{R}^d$ or $\mathbb{Z}^d$. We abstain in this section from giving explicit references to the original literature and refer to [12] across-the-board.

Here is a fundamental assertion that has been proved about the large-t behaviour of the PAM: the asymptotics of the moments of the total mass of the solution, $\langle U(t)\rangle$ (we write $\langle\cdot\rangle$ for expectation with respect to the potential ξ). We present this here only for the double-exponential distribution, as this is the one that we will consider further. The following is also the historically first result of this type; together with its proof, it gave guidance to the analysis of the PAM for many other potential distributions over the last 20 years. Introducing $H(t) = \log\langle \mathrm{e}^{t\xi(0)}\rangle$, the cumulant generating function of one of the potential variables, we have the following (see Theorem 3.13 and Remark 3.17 in [12]).

Theorem 2.1 (Moment asymptotics of the PAM)**.** *We consider the PAM as in (2.5) on $\mathcal{X} = \mathbb{Z}^d$ and assume that the random potential $\xi = (\xi(x))_{x\in\mathbb{Z}^d}$ is i.i.d. and doubly-exponentially distributed as in (2.8) with parameter $\rho\in(0,\infty)$. Then there is a number $\chi = \chi(\rho)\in[0,2d]$ such that*

$$\frac{1}{t}\log\langle U(t)\rangle = \frac{H(t)}{t} - \chi + o(1), \qquad t\to\infty. \tag{2.10}$$

Note that $H(t) = \rho t\log t + \rho t + o(t)$ for $t\to\infty$, hence the leading term $H(t)/t$ tends to infinity, while the second term χ is a constant.

Indeed, the characteristic quantity χ is represented by a certain variational formula that is interpretable and can be investigated more deeply. The interpretation is that the best contribution to the expected total mass comes from a localised peak in a certain island in $\mathbb{Z}^d$, called the *intermittent island*. More precisely, it comes from a pretty small area in which the potential ξ assumes particularly high values. This area is essentially a centered ball with a deterministic, fixed radius (not depending on t), which is a special feature that only the double-exponential distribution shows. We can also argue in terms of the Feynman–Kac formula in (2.6) that the random walk $(X_s)_{s\in[0,t]}$ spends practically all the time in this island in order to benefit as much as possible from the high potential values. In terms of the eigenvalue expansion (2.7), we can argue that the local principal eigenvalue (say, with zero boundary condition) in that island is enormously large. The leading term $H(t)/t$ expresses the high potential value, and the second term χ expresses the optimal compromise between the long stay of the random

walk in the island and the size of the island. (In order to see that the large-t asymptotics may have anything to do with any kind of optimisation, note that both the Feynman–Kac formula and the eigenvalue expansion reveal that $U(t)$ is something like a t-th exponential moment, and recall that, for any random variable Y, the moment $\mathbb{E}[\mathrm{e}^{tY}]$ runs like $\mathrm{e}^{t\,\mathrm{esssup}Y}$.)

Another fundamental result is on the almost sure behaviour of the total mass, which reads as follows in the above special case (see Theorem 5.1 and Remark 3.17 in [12]).

Theorem 2.2 (Almost-sure asympotics of the PAM)**.** *Under the same assumptions as in Theorem 2.1, there is a number $\widetilde{\chi} \in [0, 2d]$ such that*

$$\frac{1}{t} \log U(t) = h_t - \widetilde{\chi} + o(1), \qquad t \to \infty, \ \textit{almost surely}, \tag{2.11}$$

where $h_t = \varrho \log\log |B_t|$, and $|B_t|$ is the cardinality of the box $B_t = [-t, t]^d \cap \mathbb{Z}^d$.

Indeed, h_t is the asymptotics of the maximum of all the potential values in B_t, and $\widetilde{\chi}$ is another variational formula that describes the optimal principal eigenvalue of the operator $\Delta + q$ under the assumption that the potential $q\colon \mathbb{Z}^d \to \mathbb{R}$ is not too improbable to be realised under ξ. The explanation is that, in a box of some radius r_t around the starting site 0, one searches for a local region in which the local principal eigenvalue of $\Delta + \xi$ is as large as possible. Then the path in the Feynman–Kac formula runs quickly to that place (which is typically $\approx r_t$ away) and stays in that island for the remaining time until t. Here the radius r_t is chosen in such a way that an optimal compromise is reached between the probabilistic costs to reach that remote place in a short time and to have enough time left to benefit from that favourable potential. As a deeper investigation in [4] has shown, the optimal choice is $r_t = \varrho t / \log t \log_3 t$ (with $\log_k t$ the k-fold iteration of log), and the time that is used for getting there is of the order $t / \log t \log_2 t \log_3 t$.

Like for the moments, it is characteristic for the double-exponential distribution in (2.8) that the size of the intermittent island does not depend on t. Other potential distributions lead to different sizes of the intermittent islands. As a rule, the heavier the upper tails of the potential variable is, the smaller the island is. E.g., for standard normal, the Weibull or the Pareto distribution, one obtains single-site islands, while for bounded random variables the islands depend on t and grow unboundedly.

3. Multitype branching random walk in random potential

In this section, we review the results of [11] about a version of the spatial branching walk model defined in Section 2.1 with the additional feature that we give *two* categories of properties to each particle, a spatial information and information about its type. We now interpret $\mathcal{X}$ as the location space and take some Markov chain on it as a basis, and concerning the type, we introduce an additional discrete set $\mathcal{T}$, the *type space*. We will let the rates depend on the types of the particle and of the offspring. Mathematically, there are certainly possibilities to conceive $\mathcal{X} \times \mathcal{T}$ as one state space, but we are interested in the effect coming from the difference between the two components, since there are different modeling ideas behind $\mathcal{X}$ and $\mathcal{T}$.

We consider a natural and flexible discrete-time version. Denote the transition matrix of the Markov chain on $\mathcal{X}$ by $P = (P_{xy})_{x,y\in\mathcal{X}}$. We equip $\mathcal{T}$ with a set $\mathcal{A}$ of directed edges $(i, j) \in \mathcal{T} \times \mathcal{T}$ and obtain a directed finite graph $\mathcal{G} = (\mathcal{T}, \mathcal{A})$. We assume that each directed edge appears at

most once in $\mathcal{A}$, and for each $i \in \mathcal{T}$, there is at least one $j \in \mathcal{T}$ such that $(i,j) \in \mathcal{A}$. Self-edges (i,i) may appear in $\mathcal{A}$. To each $y \in \mathcal{X}$ we attach a matrix $F_y = (F_y^{(i,j)})_{(i,j)\in\mathcal{A}}$ of probability distributions on $\mathbb{N}_0$, the environment. Given $F = (F_y)_{y\in\mathcal{X}}$, we define a discrete-time Markov process $(\eta_n)_{n\in\mathbb{N}_0}$ on $\mathbb{N}_0^{\mathcal{T}\times\mathcal{X}}$, where $\eta_n(i,x)$ is the number of particles of type i at site x at time n. The environment F does not depend on time and is fixed throughout the evolution of particles. We specify the transition mechanism of $(\eta_n)_{n\in\mathbb{N}_0}$ as follows: for a given state η, during the time interval $(n, n+1)$, given that the configuration is equal to η at time n,

(1) A particle of type i located at site $y \in \mathcal{X}$ produces, independently for $j \in \mathcal{T}$ such that $(i,j) \in \mathcal{A}$, precisely k particles of type j at the same site y with probability $F_y^{(i,j)}(k)$, for any $k \in \mathbb{N}_0$. This offspring production is independent over all the particles in $\mathcal{X}$ and over the time $n \in \mathbb{N}_0$.
(2) Immediately after creation, each new particle at x independently chooses a site y with probability P_{xy} and moves there.
(3) The resulting particle configuration is η_{n+1}.

Fix a site $y \in \mathcal{X}$ and type $j \in \mathcal{T}$ as starting sites. We start the Markov chain $(\eta_n)_n$ with the initial configuration $\eta_0(i,x) = \delta_j(i)\delta_y(x)$, and by $\mathtt{P}_{j,y}$ and $\mathtt{E}_{j,y}$ we denote its distribution and expectation, respectively. Note that the Markov chain depends on the realisation of the environment F. We are interested in the expectation of the global population size, $|\eta_n| := \sum_{i\in\mathcal{T},x\in\mathcal{X}} \eta_n(i,x)$,

$$u_n(i,x) := \mathtt{E}_{i,x}[|\eta_n|], \qquad n \in \mathbb{N}_0, x \in \mathcal{X}, i \in \mathcal{T}. \tag{3.1}$$

As the first main result, we develop two formulas for u_n: as the n-th power of a certain $\mathcal{X}\times\mathcal{T}$-matrix, and as an expectation of a multiplicative functional over n steps of a particular Markov chain on that set. These two formulas are extensions of well-known representations of the expected number of particles of multitype branching processes without space; we like the fact that one can see these formulas from the viewpoint of a Feynman–Kac formula.

To formulate this, we need to introduce a Markov chain $T = (T_n)_{n\in\mathbb{N}_0}$ on the type space $\mathcal{T}$ with transition probabilities

$$p_{ij} = \frac{\mathbb{1}\{(i,j) \in \mathcal{A}\}}{\deg(i)}, \qquad i,j \in \mathcal{T}, \tag{3.2}$$

where $\deg(i) = |\{k \in \mathcal{T}\colon (i,k) \in \mathcal{A}\}|$ is the outdegree of i. We define T and X independently on a common probability space and write $\mathbb{P}_{i,x}^{(T,X)}$ and $\mathbb{E}_{i,x}^{(T,X)}$ for probability and expectation, respectively, where T starts from i and X from x. We denote by $m_{ij}(y) = \sum_{k\in\mathbb{N}_0} kF_y^{(i,j)}(k)$ the expectation of $F_y^{(i,j)}$ (the offspring expectation) and collect these random numbers in the matrix $M(y) := (m_{ij}(y))_{i,j\in\mathcal{T}}$, where we put $m_{ij}(y) = 0$ if $(i,j) \notin \mathcal{A}$. The following is [11, Proposition 1].

Proposition 3.1 (Representations for u_n)**.** *For any $i \in \mathcal{T}$ and any $x \in \mathcal{X}$ and any $n \in \mathbb{N}_0$,*

$$u_n(i,x) = \mathbb{E}_{i,x}^{(T,X)}\Big[\prod_{l=1}^{n}\big(m_{T_{l-1}T_l}(X_{l-1})\deg(T_{l-1})\big)\Big], \tag{3.3}$$

$$u_n(i,x) = B^n\mathbb{1}(i,x) = \sum_{j\in\mathcal{T}}, y \in \mathcal{X}B^n_{(i,x),(j,y)}, \tag{3.4}$$

where B is the $(\mathcal{T}\times\mathcal{X})\times(\mathcal{T}\times\mathcal{X})$ matrix with coefficients

$$B_{(i,x),(j,y)} = m_{ij}(x)P_{xy}\mathbb{1}\{(i,j)\in\mathcal{A}\}. \tag{3.5}$$

The equation in (3.4) easily follows from (3.3) by writing out explicitly the expectation over the Markov chain and the n-fold matrix product. The interpretation of (3.3) is as follows. Every n-step path $((X_0,T_0),\dots,(X_n,T_n))$, together with the product over the $m_{T_{l-1}T_l}(X_{l-1})$, stands for the union of all n-step branching subtrees that produce, in the l-th step, for any $l\in\{1,\dots,n\}$, from a particle of type T_{l-1} located at X_{l-1} a particle of type T_l that makes a step to X_l immediately after creation. The product over the $P_{X_{l-1}X_l}p_{T_{l-1}T_l}$, together will the product over the $\deg(T_{l-1})$ (note the partial canceling of terms), summarises the probabilities of all the jumps in the state space. In this way, we encounter a discrete-time version of a Feynman–Kac formula for the Markov chain (T,X) on $\mathcal{T}\times\mathcal{X}$, however with an interesting difference: the potential $\log m_{i,j}(x)$ depends on the *vertices* of the space $\mathcal{X}$ and on the *edges* of the type space $\mathcal{T}$.

So far, the environment F was fixed, but let us now turn to the case of a *random* environment, and let us describe our assumptions. We assume that the collection of all distributions $F_y^{(i,j)}$ with $y\in\mathcal{X}$ and $(i,j)\in\mathcal{A}$ is independent. Their distribution depends on (i,j), but not on y. We call $F=(F_y)_{y\in\mathcal{X}}$ the *random environment* and denote by Prob and $\langle\cdot\rangle$ probability and expectation with respect to F, respectively. Since we are only interested in the expectation of the global number of particles here, we will make our assumptions on the environment only in terms of the quantities $m_{ij}(y)=\sum_{k\in\mathbb{N}_0}kF_y^{(i,j)}(k)$. In particular, we assume that the collection of the $m_{ij}(y)$ is stochastically independent over $y\in\mathcal{X}$ and $i,j\in\mathcal{T}$ (but of course not identically distributed).

We will study the case where the upper tails of $m_{ij}(y)$ lie in the vicinity of the *Weibull distribution*, with parameter $1/\rho_{ij}\in(0,\infty)$, i.e.,

$$\mathrm{Prob}(m_{ij}(y)>r)\approx\exp\{-r^{1/\rho_{ij}}\},\qquad r\to\infty. \tag{3.6}$$

Hence, $\log m_{ij}(y)$ lies in the vicinity of the double-exponential distribution defined in (2.8) with parameter ρ_{ij}. Let $H_{ij}(t):=\log\langle m_{ij}(y)^t\rangle$ denote the logarithm of the moment generating function. For $(i,j)\in\mathcal{T}^2\setminus\mathcal{A}$, we put $\rho_{ij}=0$. Hence, our environment distribution is characterised by the matrix-valued parameter $\rho=(\rho_{ij})_{i,j\in\mathcal{T}}$. The larger ρ_{ij} is, the thicker are the tails of $m_{ij}(y)$, i.e., the easier it is for $m_{ij}(y)$ to achieve extremely high values.

The second main result of [11] is a formula for the large-n asymptotics for the expectation of the Feynman–Kac formula from (3.3). This formula is in the spirit of (2.10), however, it has some pecularities and some novelties. Let us first mention that there are basically two possible lines of proof for deriving the result: one about the large deviations for the empirical pair measure of the underlying Markov chain, and one using Frobenius eigenvalue theory. We decided to carry out only the former one. For any discrete set S, we denote by $\mathcal{M}_1(S)$ the set of probability measures on S and by $\mathcal{M}_1^{(\mathrm{s})}(S^2)$ the set of probability measures on S^2 with equal marginals. The first quantity of interest is

$$\lambda(\rho)=\sup\big\{\langle\mu,\rho\rangle\colon\mu\in\mathcal{M}_1^{(\mathrm{s})}(\mathcal{T}^2)\big\},\qquad\text{where }\langle\mu,\rho\rangle=\sum_{(i,j)\in\mathcal{A}}\mu(i,j)\rho_{ij}, \tag{3.7}$$

and the set of the corresponding maximisers:

$$\Lambda(\rho):=\big\{\mu\in\mathcal{M}_1^{(\mathrm{s})}(\mathcal{T}^2)\colon\langle\mu,\rho\rangle=\lambda(\rho)\big\}. \tag{3.8}$$

We introduce some notation. Each measure $\nu \in \mathcal{M}_1^{(s)}((\mathcal{T}\times\mathcal{X})^2)$ has a number of marginal measures that are defined on different spaces, but in order to keep the notation simple, we denote by $\overline{\nu}$ all these marginals, namely,

$$\begin{aligned} \overline{\nu}(i,j,x) &= \sum_{y\in\mathcal{X}} \nu((i,x),(j,y)), \qquad \overline{\nu}(i,x) = \sum_{j\in\mathcal{T}} \overline{\nu}(i,j,x), \\ \overline{\nu}(i,j) &= \sum_{x\in\mathcal{X}} \overline{\nu}(i,j,x), \qquad \overline{\nu}(i) = \sum_{j\in\mathcal{T}} \overline{\nu}(i,j). \end{aligned} \tag{3.9}$$

To describe the second term in the asymptotics, we need to introduce two functionals on probability measures $\nu \in \mathcal{M}_1^{(s)}((\mathcal{T}\times\mathcal{X})^2)$, an energy functional $\mathcal{S}$ and an entropy functional $\mathcal{I}$. Indeed, define

$$\begin{aligned} \mathcal{S}(\nu) &:= \sum_{(i,j)\in\mathcal{A}} \rho_{ij} \sum_{x\in\mathcal{X}} \overline{\nu}(i,j,x) \log \overline{\nu}(i,j,x) + \sum_{(i,j)\in\mathcal{A}} \overline{\nu}(i,j)\rho_{ij} \log \rho_{ij}, \\ \mathcal{I}(\nu) &:= \sum_{(i,j)\in\mathcal{A}} \sum_{x,y\in\mathcal{X}} \nu\big((i,x),(j,y)\big) \log \frac{\nu\big((i,x),(j,y)\big)}{\overline{\nu}(i,x)P_{xy}}. \end{aligned}$$

We set $\mathcal{I}(\nu)=\infty$ if ν is not absolutely continuous with respect to the measure $((i,x),(j,y)) \mapsto \overline{\nu}(i,x)P_{xy}1\{(i,j)\in\mathcal{A}\}$. Then $\mathcal{I}(\nu)$ is equal to the relative entropy of ν with respect to this measure, up to the missing normalisation; note that the reference measure is not normalised, but has mass equal to $\sum_{i\in\mathcal{T}} \overline{\nu}(i)\deg(i)$.

Now we can state our main result, [11, Theorem 3]:

Theorem 3.2. *For any $i\in\mathcal{T}$ and $x\in\mathcal{X}$, as $n\to\infty$,*

$$\langle u_n(i,x)\rangle = (n!)^{\lambda(\rho)} \mathrm{e}^{-n\chi(\rho)} \mathrm{e}^{o(n)} = \exp\Big(\lambda(\rho) n \log\frac{n}{\mathrm{e}} - n\chi(\rho) + o(n)\Big), \tag{3.10}$$

where

$$\chi(\rho) = \inf\big\{\mathcal{I}(\nu) - \mathcal{S}(\nu) \colon \nu\in\mathcal{M}_1^{(s)}\big((\mathcal{T}\times\mathcal{X})^2\big), \overline{\nu}\in\Lambda(\rho)\big\}. \tag{3.11}$$

The central object in the proof and in the understanding of this result is the *empirical pair measure*

$$\nu_n = \frac{1}{n}\sum_{l=1}^{n} \delta_{((T_{l-1},X_{l-1}),(T_l,X_l))}. \tag{3.12}$$

In terms of the space-time random walk (X,T), the number $n\nu_n((i,x),(j,y))$ is equal to the number of j-type offspring of any i-type particle located at x by time n that makes a step to y right after creation. Hence, ν_n stands for the union of all n-step paths $((X_0,T_0),\dots,(X_n,T_n))$ that make precisely $n\nu_n((i,x),(j,y))$ steps $(i,x)\to(j,y)$ for every $i,j\in\mathcal{T}$ and every $x,y\in\mathcal{X}$. The term $\mathcal{I}(\nu)$ is the negative exponential rate of the probability of this union under the Markov chain X, together with the combinatorial complexity of the trajectories of types, and $\mathcal{S}(\nu)$, together with the leading term $\lambda(\rho)$, is the one under the expectation w.r.t. the random environment.

Theorem 3.2 in particular shows that the main contribution to the annealed moments of the numbers of particles, $\lambda(\rho)$, comes from those n-step branching process subtrees that produce, for some $\mu\in\Lambda(\rho)$, at approximately $n\mu(i,j)$ of the n steps a number of j-type particles from one or more i-type particles, for any $i,j\in\mathcal{T}$. Then the value $\langle\mu,\rho\rangle$ gives the leading contribution on the scale $n\log\frac{n}{\mathrm{e}}$. It is interesting to note that the optimality of the leading term has nothing

to do with the spatial part of the branching process, but only with the creation of particles. The reason is that all the probabilities of spatial actions, i.e., of the random walk X, are on the scale n, but the values of the offspring expectations $m_{ij}(x)$ are typically on the scale $\mathrm{e}^{O(\log n)}$.

In this light, let us analyse the leading term $\lambda(\rho)$ a bit more closely. A *simple cycle* on $\mathcal{G}$ is a path $\gamma = (i_1, \dots, i_l)$ in $\mathcal{T}$, with steps (i_m, i_{m+1}) in $\mathcal{A}$, that begins and ends at the same vertex $i_1 = i_{m+1}$, but otherwise has no repeated vertices or edges. We write $(i,j) \in \gamma$ if the directed edge (i,j) belongs to γ, that is, if $(i,j) = (i_m, i_{m+1})$ for some $m \in \{1, \dots, l\}$. We call $|\gamma| = l$ its *length*. We denote by Γ_l the set of all simple cycles of length l and by Γ the set of all simple cycles. We define the uniform measure on the edges of a simple cycle γ,

$$\mu_\gamma(i,j) = \begin{cases} 1/|\gamma| & \text{if } (i,j) \in \gamma, \\ 0 & \text{otherwise.} \end{cases} \tag{3.13}$$

It is clear that $\mu_\gamma \in \mathcal{M}_1^{(\mathrm{s})}(\mathcal{T}^2)$ for any $\gamma \in \Gamma$. Simple cycles are important for the asymptotics of the annealed moments because the set of extreme points of $\mathcal{M}_1^{(\mathrm{s})}(\mathcal{T}^2)$ consists exactly of the μ_γ's with γ the simple cycles of the graph $\mathcal{G}$. The following is [11, Lemma 1].

Lemma 3.3. *The set of extreme points of the convex set $\mathcal{M}_1^{(\mathrm{s})}(\mathcal{T}^2)$ is equal to $\{\mu_\gamma \colon \gamma \in \Gamma\}$.*

Since the optimisation problem in (3.7) is a linear optimisation problem on the convex, compact set $\mathcal{M}_1^{(\mathrm{s})}(\mathcal{T}^2)$, the Krein-Milman theorem and Lemma 3.3 imply the following characterisation of the leading term in (3.10). The following is [11, Lemma 2].

Lemma 3.4.

$$\lambda(\rho) = \max\Big\{\langle \nu_\gamma^P, \rho\rangle : \gamma \in \Gamma\Big\} = \max\Big\{\frac{1}{|\gamma|}\sum_{m=1}^{|\gamma|} \rho_{i_{m-1}i_m} \colon (i_1, \dots, i_{|\gamma|}) \in \Gamma\Big\}.$$

The interpretation of Lemma 3.4 is that the leading contribution to the expected population size comes from optimal cycles $(i_1, \dots, i_{|\gamma|}) \in \Gamma_{|\gamma|}$. In terms of branching process trees, they are considered optimal if they produce only i_{m+1}-type particles from i_m-type particles for any $m \in \{1, \dots, |\gamma|\}$ (with $i_{|\gamma|+1} = i_1$), but no other offspring.

4. The PAM on finite graphs

Another, more biologically inspired, direction in the investigation of the PAM, is the PAM on some finite graph (like the full graph or the hypercube) and the concentration on the question about the amount of time that one has to give to the branching process such that the overwhelming part of the population has found the way to the "fittest" site. Here we rely on the interpretation that we explained in Section 2.2, and the question is investigated in the limit of a large (but finite) graph and late times, and the relation of the growths of time and space is crucial. In order to make a mathematical treatment feasable without too many technicalities, one typically assumes the random potential (the "fitness landscape") to be i.i.d. exponentially distributed times a growth parameter, such that one is in the regime where the intermittent islands (in the understanding of Section 2.4) are singletons.

The first work on the PAM on a finite graph $\mathcal{X}$ was — to the best of our knowledge — [8], which considered the complete graph with N nodes and an exponentially distributed i.i.d. random potential ξ. However, instead of Δ in (2.5), the *rescaled* Laplace operator $\frac{1}{N}\Delta$ is

considered. This is mathematically equivalent to (2.5) with graph-size dependent potential $N\xi$; the large prefactor N supports a strong concentration of the mass of branching particles in small intermittent islands; actually here we are concerned with single sites.

In that work, the initial condition was taken to be localised in the site of the k-th largest of the potential values, and the question was raised, for what choices of $t = t_N$, in the large-N limit, the main mass of the particle system travels to the site of the largest potential value, and for what choices it stays at the initial site by time t. The authors found the leading scale and the criterion for the answer, and they derived the first term in the asymptotics for the expectation of the total mass with this initial condition. We do not go deeper into these details.

In [2], the same question was raised for the hypercupe $\{-1,1\}^N$ with an i.i.d. potential, the assumptions of which are formulated by requiring that 2^N i.i.d. copies of the potential variables leave gaps of order one between their consecutive leading order statistics, asymptotically as $N \to \infty$. According to standard extreme-value statistics, this includes the case of centred Gaussians with variance N. The main result is the following (see [2, Theorem 1.2]).

Theorem 4.1. *Assume that $U_N(t)$ is the total mass of the solution to the PAM in (2.5) on $\mathcal{X} = \{-1,1\}^N$ with Δ replaced by $\frac{1}{N}\Delta$ and with an i.i.d. random potential ξ on $\mathcal{X}$ such that their leading order statistics $\max_{\mathcal{X}} \xi = \xi_{1,2^N} > \xi_{2,2^N} > \xi_{3,2^N} > \dots$ leaves gaps of order one between its consecutive values, asymptotically as $N \to \infty$. We assume that the initial condition in (2.5) is the delta-measure in the site of $\mathcal{X}$ in which $\xi_{k,2^N}$ sits with a fixed integer $k > 1$. Fix $\varepsilon > 0$. If $t_N \geq (1+\varepsilon)(N \log N)/2(\xi_{1,2^N} - \xi_{k,2^N})$, then, almost surely,*

$$U_N(t) = \exp\Big\{(\xi_{1,2^N} - 1)t_N + \frac{1}{2}(1+o(1))N\log N\Big\}, \qquad N \to \infty.$$

In contrast, if $t_N \leq (1-\varepsilon)(N \log N)/2(\xi_{1,2^N} - \xi_{k,2^N})$, then, almost surely,

$$U_N(t) = \exp\{(\xi_{k,2^N} - 1)t_N\}\big(1+o(1)\big), \qquad N \to \infty.$$

The interpretation is that the main mass has, in the first case of a large time horizon, enough time to move to the site of the maximal potential value, $\xi_{1,2^N}$, while in the second case (small time horizon), it stays in the initial site.

5. Higher moments of the numbers of particles

As we mentioned in Section 2.3, the first moment (the expectation) $u(t,x)$ of the number of particles in the branching random walk system at time t in the site x, satisfies the heat equation with random potential in (2.5), which is rather amenable to a deeper analysis, due to the Feynman–Kac formula in (2.6) and the eigenvalue expansion in (2.7). However, this concerns only the *first* moment, and this is quite short of a deeper understanding of the entire particle system. Substantially more information is contained in the n-th moments of the particle number at time t at site x for $n \in \mathbb{N}$, i.e., in the expectation of $\eta(t,x)^n$, again taken only over the migration, branching and killing mechanisms with fixed branching/killing rates. In this section, we are going to report on the work [10] on this aspect. To keep things simple in this survey, we only consider here $U_n(t)$, the expectation of $(\sum_{x\in\mathcal{X}} \eta(t,x))^n$. Hence, the case $n = 1$ is the one that we handled before. Again, like in Section 3, we will present two main results: An explicit formula of Feynman–Kac type for $U_n(t)$ for fixed branching/killing rates, and the identification of the large-t asymptotics of the expectation of $U_n(t)$, taken with respect to a particular choice of the distribution of these random rates.

One of our motivations to study $U_n(t)$ is that asymptotic knowledge on the behaviour of $\langle U_n(t)\rangle$ might be able to tell something about the distributional behaviour of the total number of particles, as in the general theorem where (under some technical conditions) the convergence of all the n-th moments of a random variable implies its weak convergence towards a random variable whose n-th moments are equal to the limits.

Meanwhile, also deep investigations have been carried out directly on the branching particle system in the spirit of the PAM and using a great deal of the results and methods specific for the PAM [16], but this concerned only the simplest potential distribution (the Pareto distribution) and was technically enormously cumbersome.

The method of choice to analyse $U_n(t)$ is to derive a similar equation as the heat equation with random potential and to try to employ the techniques that were helpful in the analysis of the PAM. In earlier work [1], a recursive equation was derived for $u_n(t,x)$ (defined as the expectation of $\eta(t,x)^n$) in terms of all the functions $u_1,\dots,u_{n-1}$. This formula was explicit enough to derive, for a particular potential distribution that leads to single-site islands, the first term in the large-t asymptotics of the moments of $u_n(t,\cdot)$. However, this is obviously unsatisfactory for several reasons. Instead, in [10], a direct formula for the expectation of u_n (in fact, for its p-th moment for any $p\in\mathbb{N}$) is derived, and this is so explicit that, for the double-exponential distribution in (2.8), the two main terms could be derived in terms of a variational formula that admits an interpretation and further analysis. The main tool for deriving the direct formula is the many-to-one formula from the theory of branching processes, derived via spine techniques.

Unlike in the case $n=1$ of the first moment, the n-th moment is not just a function of the difference $\xi=\xi_+-\xi_-$, but we need to keep track of both the splitting rate $\xi_+=\xi_2$ and the killing rate $\xi_-=\xi_0$.

Let us state the formula from [10, Theorem 2.1] in the special case $n=2$.

Theorem 5.1 (Feynman–Kac-like formula for U_2)**.** *Let $U_n(t)$ denote the n-th moment of the total number of particles, for any $n\in\mathbb{N}$. Then, in the case $n=2$, we have $U_2=U_1+\widetilde{U}_2$, where*

$$\widetilde{U}_2(t)=\int_0^t \mathbb{E}_{\mathcal{O}}\Big[\mathrm{e}^{\int_0^s \xi(X_r)\,\mathrm{d}r}\mathrm{e}^{\int_s^t \xi(X'_r)\,\mathrm{d}r}\mathrm{e}^{\int_s^t \xi(X''_r)\,\mathrm{d}r}2\xi_2(X_s)\Big]\,\mathrm{d}s, \tag{5.1}$$

where $(X_r)_{r\in[0,s]}$, $(X'_r)_{r\in[s,t]}$ and $(X''_r)_{r\in[s,t]}$ are three independent simple random walks, conditioned on $X_0=\mathcal{O}$ and $X'_s=X''_s=X_s$.

In other words, the right-hand side of (5.1) is the expectation over a branching random walk with precisely one splitting event in the time interval $[0,t]$, namely at time s. The first term in the representation $U_2=U_1+\widetilde{U}_2$ corresponds to absence of splitting, the second to exactly one splitting. For general n, the formula for U_n is similar (one has to sum over all numbers of splitting events in $\{0,\dots,n-1\}$), but much more cumbersome, since combinatorial prefactors are involved and multiple powers of ξ_2-values at the splitting locations. The formula in (5.1) is the main result of [10] about a representation of the n-th moments of the number of particles, for fixed branching/killing rates, in the special case $n=2$.

Now we turn to the second question. For the analysis of the large-t asymptotics of the expectation of $\widetilde{U}_2(t)$ and for its intuitive understanding, the formula in (5.1) is well suitable, at least for the case that both random fields ξ_2 and ξ_0 are i.i.d. sequences of double-exponentially distributed variables as in (2.8) and are independent. The first observation is that the term $2\xi_2(X_s)$

should hardly have any influence, and this is indeed true under some technical assumption that forbids too thick upper tails of the random variable $\xi_2(0)$.

The second observation is that, as we are considering the case that the potential can assume very large (positive) values, it should be giving particularly large values if the splitting time s is as early as possible, since it is then *two* independent random walks that can contribute to the random-potential expectation: it is indeed the expectation of a *square* of the Feynman–Kac formula, since each of the two copies contributes independently like one total mass. This reasoning can be done already when considering the leading terms of the expectation: Since we have three random walks with running time s, $t-s$ and $t-s$, one should expect that the leading logarithmic term is roughly equal to $H(s+(t-s)+(t-s)) = H(2t-s)$, which is maximal if s is minimal, since we assume that $\frac{1}{t}H(t) \to \infty$ (as ξ is unbounded). Hence, the first term in the asymptotics for $\frac{1}{t}\log\langle \widetilde{U}_2(t)\rangle$ should be $H(2t)$, and the second one should be $-2\chi(p)$ in the notation of (2.10). The same picture is true for the higher moments, i.e., for the moment asymptotics of $U_n(t)$, with 2 replaced by n at both occurrences. This has indeed been proved as the second main result of [10] in Theorem 1.3 there:

Theorem 5.2 (Moment asymptotics for the branching random walk in random environment)**.** *Let the random potentials $\xi_2 = (\xi_2(x))_{x\in\mathbb{Z}^d}$ and $\xi_0 = (\xi_0(x))_{x\in\mathbb{Z}^d}$ be independent and i.i.d. and double-exponentially distributed as in (2.8) with parameter $\rho \in (0,\infty)$. Denote by $U_n(t)$ the n-th moment of the total number of particles in the branching random walk in the random environment (ξ_0,ξ_2). Then, for any $n,p\in\mathbb{N}$,*

$$\langle U_n(t)^p\rangle = \exp\big\{H(npt) - npt\chi + o(t)\big\}, \qquad t\to\infty,$$

where $\chi\in[0,2d]$ is the number appearing in Theorem 2.1.

The most important conclusion is that, at least for the potential double-exponentially distributed, the main contribution to the expected p-th power of the total mass of the n-th moment comes from early splittings in the moment formula. This implies that

$$\langle U_n(t)^p\rangle = \langle U(t)^{np}\rangle \mathrm{e}^{o(t)} = \langle U(tnp)\rangle \mathrm{e}^{o(t)}, \qquad t\to\infty.$$

Recall that we believe that this phenomenon comes from the unboundedness of the branching rates $\xi(0)$, more precisely from the super-linear behaviour of the leading term, the logarithmic moment generating function $H(t)$. If the potential random variable $\xi(0)$ is not positive, but attains only strictly negative values, then we expect that the opposite behaviour is crucial, i.e., a very late splitting, and the result should be that $\langle U_n(t)\rangle = \langle U_1(t)\rangle \mathrm{e}^{o(t)}$ as $t\to\infty$. If the essential supremum of the potential is zero, then we expect that deeper investigations are necessary and that, in some cases, richer pictures may appear.

6. Further perspectives

6.1. **High-moment asymptotics.** In ongoing work, we are currently deriving the large-n asymptotics of the n-th moment (taken over all randomnesses: migration, branching, killing, and all the rates) of the number of particles in the branching random walk model of Section 2.1 on a finite time interval in $\mathbb{Z}^d$, i.e., the asymptotics of $\langle U_n(1)\rangle$. This will say something about the most probable way of the branching particle system to produce as much offspring as possible over a finite time interval, i.e., about the questions how high the potential values should be, how often and how quickly after each other the splitting of the particles occurs, and

how many steps all the many paths make. For deriving this, we are exploiting the moment formula that we explained in Section 5.

6.2. **Branching random walks on (random) graphs.** As we explained in Section 4, an investigation of the PAM on (possibly finite) graphs is biologically sound, but has not been done yet on a broad front; we are actually aware only of the two works [8] and [2]. In another ongoing investigation, the state space $\mathcal{X}$ is taken as a *random* graph that is locally tree-like. The main examples are a Galton–Watson tree with bounded degrees and the configuration model. The main goal is — for the random potential double-exponentially distributed as in (2.8) — to find the large-t asymptotics for the total mass $U(t)$ of the solution with high probability and to describe the structure of the parts of the random graphs that give the main contribution, i.e., of the intermittent islands. The two main interests here are to understand in general the influence (1) of the exponential structure of the tree, i.e., the fact that the volume of a ball with diameter r runs exponentially fast in r, and (2) of the randomness of the graph structure.

In the long-term, we hope to answer also the questions that are answered in the case of the state space $\mathcal{X} = \{-1, 1\}^N$ in Theorem 4.1, but the infiniteness of the state space, its randomness and the fact that the intermittent islands are not single sites, but have some structure make the question about the location and the shape of the intermittent islands a big enterprise.

6.3. **Self-repellent random walk in random potential.** As described in Section 2.4, the Feynman–Kac formula that represents the total mass of the solution to the PAM actually displays a random walk in random potential, and the large-time asymptotics are carried by those random walk paths that find their way to an optimal local region in the potential. The random walk in random potential feels an attractive force towards the extremal regions of the potential.

In another ongoing work, we are investigating the effect of an additional counter force: some additional self-interaction that suppresses self-intersections of the random walk until time t. In other words, we replace the free walk by the well-known *weakly self-avoiding* or *self-repellent walk*, which is given by an exponential weight of the form

$$\frac{1}{Z_{T,\beta}} \exp\Big\{ -\beta \int_0^T \int_0^T \mathbb{1}\{X_s = X_t\} \, \mathrm{d}s\mathrm{d}t \Big\},$$

where $\beta \in (0, \infty)$ is a parameter, and $Z_{T,\beta}$ is the normalising constant. In this model, the path in the Feynman–Kac formula cannot spend too much time anymore in single sites, the intermittent islands. The goal is to describe what it does instead. If the underlying state space is the lattice $\mathbb{Z}^d$, then the most obvious conjecture is that it will visit not only one of the intermittent islands, but several ones after each other, even though they are far away from each other. This is a random strategy, depending via an opimisation problem on the limiting spatial extreme-value order statistics of the potential. We think that this strategy is indeed optimal for most values of the thickness parameter of the potential distribution, but for some values this does not seem to be true, and we currently have no clue how to decribe the typical behaviour of the path properly.

References

[1] S. Albeverio, L.V. Bogachev, S.A. Molchanov, and E.B. Yarovaya, Annealed moment Lyapunov exponents for a branching random walk in a homogeneous random branching environment. *Markov Processes Relat. Fields* **6** (2000), 473–516. 13

[2] L. Avena, O. Gün, and M. Hesse, The parabolic Anderson model on the hypercube, *Stoch. Process. Appl.* **130** (2020), 3369–3393. 3, 12, 15

[3] E. Baake and H.-O. Georgii, Mutation, selection, and ancestry in branching models: a variational approach. *J. Math. Biol.* **54** (2007), 257–303. 3

[4] M. Biskup, W. König, and R. Soares dos Santos, Mass concentration and aging in the parabolic Anderson model with doubly-exponential tails, *Probab. Theory Relat. Fields* **171** (2018), 251–331. 7

[5] M. Birkner and N. Gantert, Ancestral lineages in spatial population models with local regulation, *this volume*. 3

[6] R. Carmona and S.A. Molchanov, *Parabolic Anderson problem and intermittency*, AMS, Providence, RI, 1994. 3

[7] J.F. Crow and M. Kimura, *An Introduction to Population Genetics Theory.* Harper & Row, New York, 1970. 4

[8] K. Fleischmann and S.A. Molchanov, Exact asymptotics in a mean field model with random potential, *Probab. Theory Relat. Fields* **86** (1990), 239–251. 11, 15

[9] W. Gabriel and E. Baake, Biological evolution through mutation, selection, and drift: An introductory review. In: *Annual Reviews of Computational Physics*, D. Stauffer (ed.), vol. VII, World Scientific, Singapore, 2000, pp. 203–264. 4

[10] O. Gün, W. König, and O. Sekulovic, Moment asymptotics for branching random walks in random environment, *Electron. J. Probab.* **18** (2013), 1–18. 12, 13, 14

[11] O. Gün, W. König, and O. Sekulovic, Moment asymptotics for multitype branching random walks in random environment, *J. Theoret. Probab.* **28** (2015), 1726–1742. 7, 8, 9, 10, 11

[12] W. König, *The Parabolic Anderson Model. Random Walk in Random Potential*, Birkhäuser, Springer Berlin, 2016. 6, 7

[13] S.M. Krone and C. Neuhauser, Ancestral processes with selection, *Theor. Popul. Biol.* **51** (1997), 210–237.

[14] J. Krug, Accessibility percolation in random fitness landscapes, *this volume*. 3

[15] P.A.P. Moran, Global stability of genetic systems governed by mutation and selection. *Math. Proc. Camb. Philos. Soc.* **80** (1976), 331–336. 4

[16] M. Ortgiese and M. Roberts, Scaling limit and ageing for branching random walk in Pareto environment. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** (2018), 1291–1313. 13