ADVANCED BRANCHING PROCESSES

FLORIAN SIMATOS
Eindhoven University of Technology, Eindhoven, Netherlands

INTRODUCTION

Branching processes have their roots in the study of the so-called family name extinction problem (see ***Introduction to Branching Processes***) and find their most natural and important applications in biology, especially in the study of population dynamics. They were also motivated by the study of nuclear fission reactions and underwent rapid development during the Manhattan project under the impulse of Szilárd and Ulam. To date, they continue to be very important in reactor physics. They also play a major role in (applied) probability at large, and appear in a wide variety of problems in queuing theory, percolation theory, random graphs, statistical mechanics, the analysis of algorithms, and bins and balls problems, to name a few.

The appearance of branching processes in so many contexts has triggered the need for extensions and variations around the classical Galton—Watson branching process. For instance, their application in particle physics provided an impetus to study them in continuous time. The possible extensions are almost endless, and indeed new models of processes exhibiting a branching structure are frequently proposed and studied. Such models allow for instance time and/or space to be continuous, individuals to have one of several types, immigration to take place, catastrophes to happen, individuals to move in space, each individual's dynamic to depend on time, space, the state of the process itself or some exogenous resources, a combination of all these ingredients, and many more.

In this article, we focus more specifically on three advanced models of branching processes: (i) branching processes in random environment, which are examples of branching processes, where the dynamic evolves (randomly) over time; (ii) branching random walks that exhibit a spatial feature; and (iii) continuous state branching processes (CSBPs) that can be seen as continuous approximations of Galton—Watson processes where both time and space are continuous. The presentation of CSBPs will also be a good place to briefly discuss superprocesses. We each time focus on the most basic properties of these processes, such as the extinction probability or the behavior of extremal particles.

This choice of topics does not aim to be exhaustive and reflects a personal selection of exciting and recent research on branching processes. It leaves out certain important classes of models, which include multitype branching processes, branching processes with immigration, and population-size-dependent branching processes. For multitype branching processes, the book by Mode [1] offers a good starting point. Branching processes with immigration were initially proposed by Heathcote [2,3] as branching models that could have a non-trivial stationary distribution. Lyons *et al*. [4] showed, via change of measure arguments, that they played a crucial role in the study of Galton—Watson processes. Finally, population-size-dependent branching processes, originally proposed by Labkovskii [5], find their motivations in population dynamics: they are elegant models that introduce dependency between individuals and can account for the important biological notion of carrying capacity, see for instance [6–8]. The interested reader can find more results in the extensive survey by Vatutin and Zubkov [9,10] that gathers results up to 1993 as well as in the recent books by Haccou *et al*. [11] and by Kimmel and Axelrod [12].

Before going on, recall (see ***Introduction to Branching Processes***) that a Galton—Watson branching process $(Z_n, n \geq 0)$ is an $\mathbb{N}$-valued Markov chain obeying to the following recursion:

Wiley Encyclopedia of Operations Research and Management Science, edited by James J. Cochran

$$Z_{n+1} = \sum_{k=1}^{Z_n} X_{ni}, \quad n = 0, 1, 2, \ldots, \qquad (1)$$

where the $(X_{ni}, n, i = 0, 1, 2, \ldots)$ are independent and identically distributed (i.i.d.) random variables following the so-called *offspring distribution*. A Galton—Watson process is classified according to the value of the mean $m = \mathbb{E}(X_{ni})$ of its offspring distribution. If $m < 1$, the process is subcritical: it dies out almost surely, the survival probability $\mathbb{P}(Z_n > 0)$ decays exponentially fast at speed m^n, and Z_n conditioned on being non-zero converges weakly. If $m = 1$, the process is critical: it dies out almost surely, the survival probability $\mathbb{P}(Z_n > 0)$ decays polynomially fast, and Z_n conditioned on being non-zero grows polynomially fast. Finally, if $m > 1$, the process is supercritical: it may survive forever, and grows exponentially fast in the event $\{\forall n \geq 0 : Z_n > 0\}$ of survival.

BRANCHING PROCESSES IN RANDOM ENVIRONMENT

A first possible generalization of the Galton—Watson model allows for the offspring distribution to vary over time: then, the recursion (1) still holds, the X_{ni} s are still independent but the law of X_{ni} may depend on n. If π_{n+1} is the offspring distribution in generation n and $\Pi = (\pi_n)$, that is, π_{n+1} is the common law of the $(X_{ni}, i = 0, 1, 2, \ldots)$ and Π is the environmental process, then this model defines a branching process in varying environment Π. We talk about branching process in random environment when the sequence Π is itself random and independent from Z_0. Note that in this case, π_n is a random probability distribution on $\mathbb{N}$.

As always in the case of stochastic processes in random environment, one may follow two approaches for their study: (i) the quenched approach, which fixes a realization of the environment and studies the process in it; it is most natural from the point of view of the applications and (ii) the annealed approach, where the various characteristics of interest are calculated by averaging over the environment. For instance, the extinction probability is a random variable in the quenched approach, and a deterministic number in the annealed approach.

When the environmental process is assumed to be stationary and ergodic, which includes for instance the case of i.i.d. environment or the case where the environment is a stationary Markov chain, it is known since the pioneering works of Smith [13], Smith and Wilkinson [14], and Athreya and Karlin [15,16] that the extinction problem and the description of the asymptotic growth have fairly general solutions. Although in the classical Galton—Watson case, the classification of Z_n is in terms of the mean of the offspring distribution, it is not difficult to see that in the case of random (stationary and ergodic) environment the mean of the logarithm of the mean is the meaningful quantity to look at. More precisely, if π is a probability distribution on $\mathbb{N}$, let $m(\pi) = \sum_y y\pi(\{y\})$ be its mean. Then, by definition (1), we have

$$\mathbb{E}\,(Z_n \mid \Pi) = Z_0 m(\pi_1) \cdots m(\pi_n) = Z_0 \mathrm{e}^{S_n},$$

where we have defined

$$S_n = \log m(\pi_1) + \cdots + \log m(\pi_n).$$

By the ergodic theorem, we have $S_n/n \to \mathbb{E}(\log m(\pi_1))$ as $n \to +\infty$, which implies that $\left[\mathbb{E}(Z_n \mid \Pi)\right]^{1/n} \to \exp[\mathbb{E}(\log m(\pi_1))]$. In particular, conditionally on the environment, the mean of Z_n goes to 0 if $\mathbb{E}(\log m(\pi_1)) < 0$ and to $+\infty$ if $\mathbb{E}(\log m(\pi_1)) > 0$. This suggests to classify the behavior of Z_n in terms of $\mathbb{E}(\log m(\pi_1))$, and under some mild technical assumptions it holds indeed that Z_n dies out almost surely if $\mathbb{E}(\log m(\pi_1)) \leq 0$ (subcritical and critical cases) and has a positive chance of surviving if $\mathbb{E}(\log m(\pi_1)) > 0$ (supercritical case). More precisely, we have the following quenched result: if $q(\Pi)$ is the (random) extinction probability of Z_n given Π, then $\mathbb{P}(q(\Pi) = 1) = 1$ in the former case and $\mathbb{P}(q(\Pi) < 1) = 1$ in the latter case.

In the supercritical case $\mathbb{E}(\log m(\pi_1)) > 0$, there is an interesting technical condition that is both necessary and sufficient to allow the process to survive with positive probability: namely, in addition to $\mathbb{E}(\log m(\pi_1)) > 0$ one also needs to assume

$\mathbb{E}(-\log(1-\pi_1(\{0\})) < +\infty$. This condition shows the interesting interplay that arises between Z_n and the environment: even though $\mathbb{E}(\log m(\pi_1)) > 0$ is sufficient to make the conditional mean of Z_n diverge, if $\mathbb{E}(-\log(1-\pi_1(\{0\})) = +\infty$ then the process almost surely dies out because the probability of having an unfavorable environment is large, where by unfavorable environment we mean an environment π where the (random) probability $\pi(\{0\})$ of having no offspring is close to 1. In other words, if $\mathbb{E}(-\log(1-\pi_1(\{0\})) = +\infty$ then the process gets almost surely extinct because of the wide fluctuation of the environment.

The classification of Z_n into the subcritical, critical, and supercritical cases also corresponds to different asymptotic behaviors of Z_n conditioned on non-extinction (here again, we have the quenched results of Athreya and Karlin [15] in mind). In that respect, Z_n shares many similarities with a Galton—Watson process, although there are some subtle differences as we see at the end of this section. In the supercritical case, Z_n grows exponentially fast in the event of non-extinction, whereas in the subcritical case, Z_n conditioned on being non-zero converges weakly to a non-degenerate random variable. In the critical case, Z_n conditioned on being non-zero converges weakly to $+\infty$, a result that can be refined in the case of i.i.d. environment.

Indeed, the case where the (π_i) are i.i.d. has been extensively studied. In this case, S_n is a random walk and recent works have highlighted the intimate relation between Z_n and S_n. In particular, the classification of Z_n can be generalized as follows. It is known from random walk theory that, when one excludes the trivial case where $S_n = S_0$ for every n, there are only three possibilities concerning the almost sure behavior of (S_n): either it drifts to $-\infty$, or it oscillates with $\liminf_n S_n = -\infty$ and $\limsup_n S_n = +\infty$, or it drifts to $+\infty$. Then, without assuming that the mean $\mathbb{E}(\log m(\pi_1))$ exists, Z_n can be said to be subcritical if $S_n \to -\infty$, critical if S_n oscillates, and supercritical if $S_n \to +\infty$. Within this terminology, Afanasyev *et al.* [17] studied the critical case and were able to obtain striking results linking the behavior of Z_n to the behavior of its associated random walk. In particular, this work emphasized the major role played by fluctuation theory of random walks in the study of branching processes in random (i.i.d.) environment, a line of thought that has been very active since then.

Let us illustrate this idea with some of the results of Afanasyev *et al.* [17], so consider Z_n a critical branching process in random environment. As Z_n is absorbed at 0, we have $\mathbb{P}(Z_n > 0 \mid \Pi) \le \mathbb{P}(Z_m > 0 \mid \Pi)$ for any $m \le n$ and as Z_n is integer-valued, we obtain $\mathbb{P}(Z_n > 0 \mid \Pi) \le \mathbb{E}(Z_m \mid \Pi)$. It follows that

$$\mathbb{P}(Z_n > 0 \mid \Pi) \le Z_0 \exp\left(\min_{0 \le m \le n} S_m\right),$$

which gives an upper bound, in term of the infimum process of the random walk S_n, on the decay rate of the extinction probability in the quenched approach. It turns out that this upper bound is essentially correct, and that the infimum also leads to the correct decay rate of the extinction probability in the annealed approach, although in a different form. Indeed, it can be shown under fairly general assumptions that

$$\mathbb{P}(Z_n > 0) \sim \theta \mathbb{P}(\min(S_1, \dots, S_n) > 0) \qquad (2)$$

for some $\theta \in (0, \infty)$. Moreover, conditionally on $\{Z_n > 0\}$, Z_n / e^{S_n} converges weakly to a random variable W, almost surely finite and strictly positive, showing that S_n essentially governs the growth rate of Z_n. Finally, although it is natural to consider the growth rate and extinction probability of the process Z_n, one can also reverse the viewpoint and study the kind of environment that makes the process survive for a long time. And actually, the conditioning $\{Z_n > 0\}$ has a strong impact on the environment: although S_n oscillates, conditionally on $\{Z_n > 0\}$ the process $(S_k, 0 \le k \le n)$ suitably rescaled can be shown to converge to the meander of a Lévy process, informally, a Lévy process conditioned on staying positive. This provides another illustration of the richness of this class of models, where the interplay between the environment and the process leads to very interesting behavior.

These various results concern the annealed approach: Equation (2) is for instance obtained by averaging over the environment. However, the connection between Z_n and S_n continues to hold in the quenched approach. In Ref. 18, it is for instance shown that Z_n passes through a number of bottlenecks at the moments close to the sequential points of minima in the associated random walk. More precisely, if $\tau(n) = \min\{k \geq 0 : S_j \geq S_k, \quad j = 0, \ldots, n\}$ is the leftmost point of the interval $[0, n]$ at which the minimal value of $(S_j,\ j = 0, \ldots, n)$ is attained, $Z_{\tau(n)}$ conditionally on the environment and on $\{Z_n > 0\}$ converges weakly to a finite random variable. For further reading on this topic, the reader is referred to Refs 19 and 20.

Let us conclude this section by completing the classification of branching processes in random environment. We have mentioned that similarly as Galton—Watson processes, branching processes in random environment could be classified as subcritical, critical, or supercritical according to whether $\mathbb{E}(Y) < 0$, $\mathbb{E}(Y) = 0$, or $\mathbb{E}(Y) > 0$, with $Y = \log m(\pi_1)$ (in the "simple" case where Y is indeed integrable). Interestingly, assuming that $\mathbb{E}(\mathrm{e}^{tY})$ is finite for every $t \geq 0$, the subcritical phase can be further subdivided, according to whether $\mathbb{E}(Y\mathrm{e}^Y) > 0$, $\mathbb{E}(Y\mathrm{e}^Y) = 0$, or $\mathbb{E}(Y\mathrm{e}^Y) < 0$ corresponding respectively, in the terminology of Birkner *et al*. [21], to the weakly subcritical, intermediate subcritical, and strongly subcritical cases. These three cases correspond to different speeds of extinction: in the weakly subcritical case, there exists $\beta \in (0, 1)$ such that $\mathbb{E}(Y\mathrm{e}^{\beta Y}) = 0$ and $\mathbb{P}(Z_n > 0)$ decays like $n^{-3/2}[\mathbb{E}(\mathrm{e}^{\beta Y})]^n$; in the intermediate subcritical case, $\mathbb{P}(Z_n > 0)$ decays like $n^{-1/2}[\mathbb{E}(\mathrm{e}^Y)]^n$; finally, in the strongly subcritical case, $\mathbb{P}(Z_n > 0)$ decays like $[\mathbb{E}(\mathrm{e}^Y)]^n$. These decay rates are to be compared to the classical Galton—Watson process, wherein the subcritical case $\mathbb{P}(Z_n > 0)$ decays like m^n, corresponding to the strongly subcritical case, because when Y is determinist we have the relation $m = \mathbb{E}\mathrm{e}^Y$.

Further reading on branching processes in random environment includes, for example, Refs 22 and 23 for the study of the subcritical case using the annealed approach, while the trajectories of Z_n under various conditionings, namely dying at a distant given moment and attaining a high level, have been studied in Refs 24,25 and 26,27, respectively. In Ref. 28, the survival probability of the critical multitype branching process in a Markovian random environment is investigated.

BRANCHING RANDOM WALKS

Branching random walks are extension of Galton—Watson processes that, in addition to the genealogical structure, add a spatial component to the model. Each individual has a location, say for simplicity on the real line. Typically, each individual begets a random number of offspring, as in a regular branching process, and the positions of these children form a point process centered around the location of the parent. For instance, if $B = \sum_{x \in B} \delta_x$ is the law of the point process governing the locations of the offspring of a given individual, with δ_x the Dirac measure at $x \in \mathbb{R}$, the locations of the children of an individual located at y are given by the atoms of the measure $\sum_{x \in B} \delta_{y+x}$. Branching random walks can therefore naturally be seen as measure-valued Markov processes, which will turn out to be the right point of view when discussing superprocesses. Another viewpoint is to see branching random walks as random labeled trees: the tree itself represents the genealogical structure, whereas the label on an edge represents the displacement of a child with respect to its parent. Nodes of the tree then naturally inherit labels recursively, where the root is assigned any label, and the label of a node that is not the root is given by the label of its parent plus the label on the corresponding edge.

There is an interesting connection between branching random walks and (general) branching processes. In the case where the atoms of B are in $(0, \infty)$, particles of the branching random walk live on the positive half-line and their positions can therefore be interpreted as the time at which the corresponding particle is born. Keeping track of the filiation between particles, we see that within this interpretation, each particle gives birth at times given by the atoms of the random measure B. This is exactly the model of general, or Crump—Mode—Jagers, branching processes (see ***Introduction to Branching Processes***).

One of the most studied questions related to branching random walks concerns the long-term behavior of extremal particles. Of course, as the branching random walk is absorbed when there are no more particles, this question only makes sense when the underlying Galton—Watson process is supercritical and conditioned on surviving. Let for instance M_n be the location of the leftmost particle in the nth generation, that is, the smallest label among the labels of all nodes at depth n in the tree. Assume for simplicity that each individual has two children with i.i.d. displacements, say with distribution D. Then by construction, a typical line of descent (i.e., the labels on the successive nodes on an infinite path from the root) is equal in distribution to a random walk with step distribution D, thus drifting to $+\infty$ if $\mathbb{E}D > 0$. However, M_n is then equal in distribution to the minimum between 2^n random walks, and although a typical line of descent goes to $+\infty$, the exponential explosion in the number of particles makes it possible for the minimal displacement M_n to follow an atypical trajectory and, say, diverge to $-\infty$. Finer results are even available, and the speed at which M_n diverges has been initiated in a classical work by Hammersley [29] (who was interested in general branching processes), and later extended by Kingman [30] and Biggins [31] leading to what is now commonly referred to as the *Hammersley—Kingman—Biggins theorem*. For instance, in the simple case with binary offspring and i.i.d. displacements, it can be shown that $M_n \to -\infty$ if $inf_{\theta\geq 0}\mathbb{E}(e^{-\theta D}) > 1/2$ and simple computations even give a precise idea of the speed at which this happens. Indeed, if $S_n^{(k)}$ for $k = 1, \dots, 2^n$ are the 2^n labels of the nodes at depth n in the tree, we have by definition for any $a \in \mathbb{R}$

$$\begin{aligned}\mathbb{P}(M_n \leq an) &= \mathbb{P}\left(S_n^{(k)} \leq an \text{ for some } k \leq 2^n\right)\\ &\leq \sum_{1\leq k\leq 2^n} \mathbb{P}\left(S_n^{(k)} \leq an\right)\end{aligned}$$

using the union bound for the last inequality. As the $S_n^{(k)}$'s are identically distributed, say with common distribution S_n equal to the value at time n of a random walk with step distribution D, we obtain for any $\theta \geq 0$

$$\begin{aligned}\mathbb{P}(M_n \leq an) &\leq 2^n \mathbb{P}(S_n \leq an)\\ &\leq 2^n \left[e^{\theta an}\mathbb{E}(e^{-\theta S_n})\right] \leq \left[2\mu(a)\right]^n\end{aligned}$$

using Markov inequality for the second inequality, and defining $\mu(a)$ as $\mu(a) = inf_{\theta\geq 0}\left(e^{\theta a}\mathbb{E}(e^{-\theta D})\right)$ in the last term. In particular, $\mathbb{P}(M_n/n \leq a) \to 0$ if a is such that $\mu(a) < 1/2$, which makes $M_n/n \to \gamma$, with $\gamma = inf\{a : \mu(a) > 1/2\}$, the best we could hope for. It is quite surprising that these simple computations lead to the right answer, but the almost sure convergence $M_n/n \to \gamma$ is indeed the content of the aforementioned Hammersley—Kingman—Biggins theorem.

The case $\gamma = 0$ can be seen as a critical case, where the speed is sublinear; for a large class of branching random walks, this case also corresponds, after some renormalization (typically, centering the branching random walk), to study the second-order asymptotic behavior of M_n for a general γ. In the case $\gamma = 0$, several asymptotic behaviors are possible and the reader can for instance consult Addario-Berry and Reed [32] for more details. Bramson [33] proved that if every particle gives rise to exactly two particles and the displacement takes the value 0 or 1 with equal probability, then $M_n - \log\log n/\log 2$ converges almost surely. Recently, Aïdékon [34] proved in the so-called boundary case that $M_n - (3/2)\log n$ converges weakly.

These results concern the behavior of the extremal particle, and there has recently been an intense activity to describe the asymptotic behavior of all extremal particles, that is, the largest one, together with the second largest one, and third largest one. Informally, one is interested in the behavior of the branching random walk "seen from its tip," which technically amounts to consider the point process recording the distances from every particle to the extremal one. This question was recently solved by Madaule [35], building on previous results by different authors, in particular the aforementioned work by Aïdékon [34]. The limiting point process can be seen as a "colored" Poisson process, informally obtained by attaching to each atom of some Poisson process independent copies of some other point process.

Initially, one of the main motivations for studying the extremal particle of a branching random walk comes from a connection with the theory of partial differential equations. Namely, one can consider a variation of the branching random walk model, called the *branching Brownian motion*. In this model, time and space are continuous; each particle lives for a random duration, exponentially distributed, during which it performs a Brownian motion, and is replaced on death by k particles with probability p_k. Then, McKean [36] and later Bramson [37] observed that if $\sum_k kp_k = 2$ and $\sum_k k^2 p_k < +\infty$, then the function $u(t,x) = \mathbb{P}(M(t) > x)$, with $M(t)$ now the maximal displacement of the branching Brownian motion at time t, that is, the location of the rightmost particle, is a solution to the so-called Kolmogorov—Petrovskii—Piskunov (KPP) equation, which reads

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + \sum_{k\geq 1} p_k u^k - u$$

with the initial condition $u(0,x) = 1$ if $x \geq 0$ and $u(0,x) = 0$ if $x < 0$. One of the key properties of the KPP equation is that it admits traveling waves: there exists a unique solution satisfying

$$u\left(t, m(t) + x\right) \to w(x) \text{uniformly in } x \text{ as } t \to +\infty.$$

Using the connection with the branching Brownian motion, Bramson [37] was able to derive extremely precise results on the position of the traveling wave, and essentially proved that $m(t) = \sqrt{2}t - (3/2^{3/2}) \log t$. In probabilistic terms, this means that $M(t) - \sqrt{2}t + (3/2^{3/2}) \log t$ converges weakly. Similarly as for the branching random walk, there has recently been an intense activity to describe the branching Brownian motion seen from its tip, which culminated in the recent works by Arguin *et al*. [38] and Aïdékon [34].

Beyond the behavior of extremal particles, the dependency of the branching random walk on the space dimension has been investigated in Refs 39–45. There are also a number of articles for the so-called catalytic random walk when the particles performing random walk on $\mathbb{Z}^d$ reproduce at the origin only [42,46,47] for which a wide range of phenomena has been investigated. For this and more, the reader can for instance consult the recent survey by Bertacchi and Zuccha [48].

CONTINUOUS STATE BRANCHING PROCESSES

From a modeling standpoint, it is natural in the context of large populations to wonder about branching processes in continuous time and with a continuous state space. In the same vein, Brownian motion (or, more generally, a Lévy process) approximates a random walk evolving on a long time scale. The definition (1) does not easily lend itself to such a generalization.

An alternative and, from this perspective, more suitable characterization of Galton—Watson processes is through the branching property. If Z^y denotes a Galton—Watson process started with $y \in \mathbb{N}$ individuals, the family of processes $(Z^y, y \in \mathbb{N})$ is such that

$$Z^{y+z} \stackrel{(d)}{=} Z^y + \widetilde{Z}^z, y, z \in \mathbb{N}, \tag{3}$$

where $\stackrel{(d)}{=}$ means equality in distribution and $\widetilde{Z}^z$ is a copy of Z^z, independent from Z^y. In words, a Galton—Watson process with offspring distribution X started with $y + z$ individuals is stochastically equivalent to the sum of two independent Galton—Watson processes, both with offspring distribution X, and where one starts with y individuals and the other with z. It can actually be shown that this property uniquely characterizes Galton—Watson processes, and Lamperti [49] uses this characterization as the definition of a continuous state branching process (CSBP). Formally, CSBPs are the only time-homogeneous Markov processes (in particular, in continuous time) with state space $[0, \infty]$ that satisfy the branching property, see also Ikeda *et al*. [50] for more general state spaces. Note that even in the case of real-valued branching processes, the state space includes $+\infty$: in contrast with Galton—Watson branching processes, CSBP can in principle explode in finite time.

One of the achievements of the theory is the complete classification of CSBPs, the main result being a one-to-one correspondence between CSBPs and Lévy processes with no negative jumps and, in full generality, possibly killed after an exponential time. This result has a long history dating back from Lamperti [49], for which the reader can find more details in the introduction of Caballero *et al*. [51] (note that CSBPs were first considered by Jiřina [52]). There are two classical ways to see this result. Until further notice, let $Z = (Z(t), t \geq 0)$ be a CSBP started at $Z(0) = 1$.

The first one is through random time-change manipulations, and more specifically through the Lamperti transformation $\mathcal{L}$ that acts on positive functions as follows. If $f : [0, \infty) \to [0, \infty)$, then $\mathcal{L}(f)$ is defined implicitly by $\mathcal{L}(f)(\int_0^t f) = f(t)$ for $t \geq 0$, and explicitly by $\mathcal{L}(f) = f \circ \kappa$ with $\kappa(t) = inf\{u \geq 0 : \int_0^u f > t\}$. Then, it can be proved that $\mathcal{L}(Z)$ is a Lévy process with no negative jumps, stopped at 0 and possibly killed after an exponential time; conversely, if Y is such a Lévy process, then $\mathcal{L}^{-1}(Y)$ is (well defined and is) a CSBP.

The second way is more analytical. Let $u(t, \lambda) = -\log \mathbb{E}(\mathrm{e}^{-\lambda Z(t)})$: then u satisfies the semigroup property $u(s + t, \lambda) = u(s, u(t, \lambda))$, which leads, for small $h > 0$, to the approximation

$$\begin{aligned} u(t+h,\lambda) - u(t,\lambda) &= u(h, u(t,\lambda)) - u(0, u(t,\lambda)) \\ &\approx h\frac{\partial u}{\partial t}(0, u(t,\lambda)) = -h\Psi(u(t,\lambda)) \end{aligned}$$

once one defines $\Psi(\lambda) = -\frac{\partial u}{\partial t}(0, \lambda)$. It can indeed be shown that u satisfies the so-called branching equation

$$\frac{\partial u}{\partial t} = -\Psi(u), \tag{4}$$

with boundary condition $u(0, \lambda) = \lambda$. In particular, Ψ uniquely characterizes u, and thus Z, and it is called the *branching mechanism of* Z. Moreover, it can be proved that Ψ is a Lévy exponent, that is, there exists a Lévy process Y with no negative jumps such that $\Psi(\lambda) = -\log \mathbb{E}(\mathrm{e}^{-\lambda Y(1)})$ or equivalently, in view of the Lévy—Khintchine formula, Ψ is of the form

$$\begin{aligned} \Psi(\lambda) = \varepsilon + \alpha\lambda - \frac{1}{2}\beta\lambda^2 \\ + \int_{(0,\infty)} \left(1 - \mathrm{e}^{-\lambda x} - \lambda x 1_{\{x<1\}}\right) \pi(\mathrm{d}x) \end{aligned}$$

for some measure π satisfying $\int_{(0,\infty)} (1 \wedge x^2)\pi(\mathrm{d}x) < +\infty$ and some parameters $\varepsilon, \beta \geq 0$, and $\alpha \in \mathbb{R}$. Conversely, any Ψ of this form is the branching mechanism of a CSBP, and as Lévy processes are characterized by their Lévy exponent, this provides an alternative view on the one-to-one correspondence between CSBPs and Lévy processes.

The long-term behavior of CSBPs was one of the earliest questions studied by Grey [53], and they exhibit a richer range of behavior than Galton—Watson processes. First, Z is conservative, that is, $\mathbb{P}(Z(t) < +\infty) = 1$ for every $t \geq 0$, if and only if $1/\Psi$ is integrable at $0+$, a sufficient condition for this being that $\Psi(0) = 0$ and $\Psi'(0) > -\infty$. Further, we observe by differentiating the branching equation (4) with respect to λ and using $\mathbb{E}(Z(t)) = \frac{\partial u}{\partial \lambda}(t, 0)$ that $\mathbb{E}(Z(t)) = \exp(-\Psi'(0)t)$, which suggests to classify a (conservative) CSBP as subcritical, critical, or supercritical according to whether $\Psi'(0) > 0$, $\Psi'(0) = 0$, or $\Psi'(0) < 0$, respectively. This classification turns out to be essentially correct for conservative processes, under the additional requirement $\beta > 0$: in this case, supercritical processes may survive forever, with probability $\mathrm{e}^{-\lambda_0}$ where λ_0 is the largest root of the equation $\Psi(\lambda) = 0$, while critical and subcritical processes die out almost surely, that is, the time $inf\{t \geq 0 : Z(t) = 0\}$ is almost surely finite.

When $\beta = 0$, the situation may be slightly different. indeed, for any CSBP, the extinction probability $\mathbb{P}(\exists t \geq 0 : Z(t) = 0)$ is strictly positive if and only if $1/\Psi$ is integrable at $+\infty$ and $\Psi(\lambda) > 0$ for λ large enough; in this case, the extinction probability is equal to $\mathrm{e}^{-\lambda_0}$ with λ_0 as discussed earlier. In particular, we may have a subcritical CSBP (with $\Psi'(0) > 0$) satisfying both $\beta = 0$ and $\int^\infty (1/\Psi) = +\infty$, in which case $Z(t) \to 0$ but $Z(t) > 0$ for every $t \geq 0$. In other words, although Z vanishes, in the absence of the stochastic fluctuations induced by β, it never hits 0. This behavior is to some extent quite natural, because the α term corresponds to a deterministic exponential decay (for $\Psi(\lambda) = \alpha\lambda$ we have $Z(t) = \mathrm{e}^{-\alpha t}$) and the jumps of Z are only positive, so one needs stochastic fluctuations in order to make Z hit 0.

We have mentioned in the beginning of this section the motivation for studying CSBPs as continuous approximations of Galton—Watson processes. This line of thought is actually present in one of the earliest articles by Lamperti [54] on the subject. In particular, CSBPs are the only possible scaling limits of Galton—Watson processes, that is, if $(Z^{(n)}, n \geq 1)$ is a sequence of Galton—Watson processes with $Z_0^{(n)} = n$ such that the sequence of rescaled processes $(\overline{Z}^{(n)}, n \geq 1)$, where $\overline{Z}^{(n)}(t) = Z^{(n)}_{\lfloor a_n t \rfloor}/n$ for some normalizing sequence a_n, converges weakly to some limiting process $\overline{Z}$, then $\overline{Z}$ must be a CSBP. And conversely, any CSBP can be realized in this way.

There is, at least informally, an easy way to see this result, by extending the Lamperti transformation at the discrete level of Galton—Watson processes. Indeed, consider $(S(k), k \geq 0)$ a random walk with step distribution $X' = X - 1$ for some integer-valued random variable X, and define recursively $Z_0 = S(0)$ and $Z_{n+1} = S(0) + S(Z_1 + \cdots + Z_n)$ for $n \geq 0$. Then, writing $S(k) = S(0) + X'_1 + \cdots + X'_k$ with (X'_k) i.i.d. copies of X', we have

$$Z_{n+1} - Z_n = X'_{Z_1+\cdots+Z_{n-1}+1} + \cdots + X'_{Z_1+\cdots+Z_{n-1}+Z_n}$$

and so Z_{n+1} is the sum of Z_n i.i.d. copies of $X' + 1$, that is, Z_n is a branching process with offspring distribution X. This realizes Z as the time-change of a random walk, and leveraging on classical results on the convergence of random walks toward Lévy processes and continuity properties of the time-change involved [55], one can prove that the limit of any sequence of suitably renormalized Galton—Watson processes must be the time-change of a Lévy process, that is, a CSBP. This approach is for instance carried on by Ethier and Kurtz [56].

If a CSBP can be viewed as a continuous approximation of a Galton—Watson process, it is natural to ask about the existence of a corresponding genealogical structure. This question was answered by Duquesne and Le Gall [57], who for each CSBP Z exhibited a process H, which they call height process, such that Z is the local time process of H. This question is intrinsically linked to the study of continuum random trees initiated by Aldous [58,59]. As a side remark, note that this genealogical construction plays a key role in the construction of the Brownian snake [60]. There has also been considerable interest in CSBP allowing immigration of new individuals: these processes were defined by Kawazu and Watanabe [61], their genealogy studied by Lambert [62] and the corresponding continuum random trees by Duquesne [63].

Finally, let us conclude this section on CSBPs by mentioning superprocesses. Superprocesses are the continuous approximations of branching random walks, in the same vein as CSBPs are continuous approximations of Galton—Watson processes. They were constructed by Watanabe [64], and can technically be described as measure-valued Markov processes. Similarly as for the branching Brownian motion, Dynkin [65] showed that superprocesses are deeply connected to partial differential equations. The recent book by Li [66] offers a nice account on this topic.

REFERENCES

1. Mode CJ. Multitype branching processes. Theory and applications, Modern Analytic and Computational Methods in Science and Mathematics, No. 34. New York: American Elsevier Publishing Co., Inc.; 1971.
2. Heathcote CR. A branching process allowing immigration. J R Stat Soc Ser B 1965;27:138–143.
3. Heathcote CR. Corrections and comments on the paper "A branching process allowing immigration". J R Stat Soc Ser B 1966;28:213–217.
4. Lyons R, Pemantle R, Peres Y. Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. Ann Probab 1995;23(3):1125–1138.

5. Labkovskii V. A limit theorem for generalized random branching processes depending on the size of the population. Theory Probab Appl 1972;17(1): 72–85.
6. Jagers P. Population-size-dependent branching processes. J Appl Math Stochast Anal 1996;9(4): 449–457.
7. Jagers P, Klebaner FC. Population-size-dependent, age-structured branching processes linger around their carrying capacity. J Appl Probab 2011;48A (New frontiers in applied probability: a Festschrift for Soren Asmussen): 249–260.
8. Klebaner FC. On population-size-dependent branching processes. Adv Appl Probab 1984;16(1): 30–55.
9. Vatutin VA, Zubkov AM. Branching processes. I. J Math Sci 1987;39:2431–2475. DOI: 10.1007/BF01086176.
10. Vatutin VA, Zubkov AM. Branching processes. II. J Math Sci 1993;67:3407–3485. DOI: 10.1007/BF01096272.
11. Haccou P, Jagers P, Vatutin VA. Branching processes: variation, growth, and extinction of populations, Cambridge Studies in Adaptive Dynamics. Cambridge: Cambridge University Press; 2007.
12. Kimmel M, Axelrod DE. Branching processes in biology. Volume 19, Interdisciplinary Applied Mathematics. New York: Springer-Verlag; 2002.
13. Smith WL. Necessary conditions for almost sure extinction of a branching process with random environment. Ann Math Stat 1968;39:2136–2140.
14. Smith WL, Wilkinson WE. On branching processes in random environments. Ann Math Stat 1969;40:814–827.
15. Athreya KB, Karlin S. Branching processes with random environments. II. Limit theorems. Ann Math Stat 1971;42: 1843–1858.
16. Athreya KB, Karlin S. On branching processes with random environments. I. Extinction probabilities. Ann Math Stat 1971;42:1499–1520.
17. Afanasyev VI, Geiger J, Kersting G, *et al*. Criticality for branching processes in random environment. Ann Probab 2005;33(2): 645–673.
18. Vatutin V, Dyakonova E. Branching processes in a random environment and bottlenecks in the evolution of populations. Theory Probab Appl 2007;51(1): 189–210.
19. Vatutin V, Dyakonova E. Galton-Watson branching processes in a random environment I: limit theorems. Theory Probab Appl 2004;48(2): 314–336.
20. Vatutin V, Dyakonova E. Galton–Watson branching processes in a random environment. II: Finite-dimensional distributions. Theory Probab Appl 2005;49(2): 275–309.
21. Birkner M, Geiger J, Kersting G. Branching processes in random environment—a view on critical and subcritical cases. Interacting stochastic systems. Berlin: Springer; 2005. p 269–291.
22. Afanasyev VI, Böinghoff C, Kersting G, *et al*. Limit theorems for weakly subcritical branching processes in random environment. J Theor Probab 2012;25(3): 703–732.
23. Afanasyev V, Böinghoff C, Kersting G, *et al*. Conditional limit theorems for intermediately subcritical branching processes in random environment. To appear in *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques*.
24. Böinghoff C, Dyakonova EE, Kersting G, *et al*. Branching processes in random environment which extinct at a given moment. Markov Process Relat Fields 2010;16(2): 329–350.
25. Vatutin V, Wachtel V. Sudden extinction of a critical branching process in a random environment. Theory Probab Appl 2010;54(3): 466–484.
26. Afanasyev V. Brownian high jump. Theory Probab Appl 2011;55(2): 183–197.
27. Afanasyev V. Invariance principle for a critical Galton-Watson process attaining a high level. Theory Probab Appl 2011;55(4): 559–574.
28. Dyakonova E. Multitype Galton–Watson branching processes in Markovian random environment. Theory Probab Appl 2012;56(3): 508–517.
29. Hammersley JM. Postulates for subadditive processes. Ann Probab 1974;2:652–680.
30. Kingman JFC. The first birth problem for an age-dependent branching process. Ann Probab 1975;3(5): 790–801.
31. Biggins JD. The first- and last-birth problems for a multitype age-dependent branching process. Adv Appl Probab 1976;8(3): 446–459.
32. Addario-Berry L, Reed B. Minima in branching random walks. Ann Probab 2009;37(3): 1044–1079.

33. Bramson MD. Minimal displacement of branching random walk. Z Wahrsch Verw Gebiete 1978;45(2): 89–108.
34. Aïdékon E. Convergence in law of the minimum of a branching random walk. Annals of Probability. 2013;41(3A):1362–1426.
35. Madaule T. Convergence in law for the branching random walk seen from its tip 2011. arXiv:1107.2543.
36. McKean HP. Application of Brownian motion to the equation of Kolmogorov–Petrovskii–Piskunov. Commun Pure Appl Math 1975;28(3): 323–331.
37. Bramson MD. Maximal displacement of branching Brownian motion. Commun Pure Appl Math 1978;31(5): 531–581.
38. Arguin L-P, Bovier A, Kistler N. The extremal process of branching Brownian motion. Probab Theory Relat Fields 2012. DOI: 10.1007/s00440-012-0464-x.
39. Bramson M, Cox JT, Greven A. Ergodicity of critical spatial branching processes in low dimensions. Ann Probab 1993;21(4): 1946–1957.
40. Bramson M, Cox JT, Greven A. Invariant measures of critical spatial branching processes in high dimensions. Ann Probab 1997;25(1): 56–70.
41. Cox JT, Greven A. On the long term behavior of some finite particle systems. Probab Theory Relat Fields 1990;85(2): 195–237.
42. Fleischmann K, Vatutin V, Wakolbinger A. Branching systems with long-living particles at the critical dimension. Theory Probab Appl 2003;47(3): 429–454.
43. Fleischmann K, Vatutin VA. An integral test for a critical multitype spatially homogeneous branching particle process and a related reaction-diffusion system. Probab Theory Relat Fields 2000;116(4): 545–572.
44. Klenke A. Different clustering regimes in systems of hierarchically interacting diffusions. Ann Probab 1996;24(2): 660–697.
45. Klenke A. Clustering and invariant measures for spatial branching models with infinite variance. Ann Probab 1998;26(3): 1057–1087.
46. Albeverio S, Bogachev LV. Branching random walk in a catalytic medium. I. Basic equations. Positivity 2000;4(1): 41–100.
47. Vatutin V, Topchii V. Limit theorem for critical catalytic branching random walks. Theory Probab Appl 2005;49(3): 498–518.
48. Bertacchi D, Zucca F. Statistical mechanics and random walks: principles, processes and applications. Recent results on branching random walks. Nova Science Publishers Inc.; 2012.
49. Lamperti J. Continuous state branching processes. Bull Am Math Soc 1967;73:382–386.
50. Ikeda N, Nagasawa M, Watanabe S. Branching Markov processes. I. J Math Kyoto Univ 1968;8:233–278.
51. Caballero ME, Lambert A, Uribe Bravo G. Proof(s) of the Lamperti representation of continuous-state branching processes. Probab Surv 2009;6:62–89.
52. Jiřina M. Stochastic branching processes with continuous state space. Czech Math J 1958;8(83):292–313.
53. Grey DR. Asymptotic behaviour of continuous time, continuous state-space branching processes. J Appl Probab 1974;11(4): 669–677.
54. Lamperti J. The limit of a sequence of branching processes. Z Wahrsch Verw Gebiete 1967;7:271–288.
55. Helland IS. Continuity of a class of random time transformations. Stoch Process Appl 1978;7(1): 79–99.
56. Ethier SN, Kurtz TG. Markov Processes, Characterization and Convergence. Wiley Series in Probability and Mathematical Statistics. New-York: John Wiley & Sons Inc.; 1986.
57. Duquesne T, Le Gall J-F. Random trees, Lévy processes and spatial branching processes. Astérisque 2002;281:vi+147.
58. Aldous D. The continuum random tree. I. Ann Probab 1991;19(1): 1–28.
59. Aldous D. The continuum random tree. III. Ann Probab 1993;21(1): 248–289.
60. Le Gall J-F. Spatial branching processes, random snakes and partial differential equations. Lectures in Mathematics ETH Zürich. Basel: Birkhäuser Verlag; 1999.
61. Kawazu K, Watanabe S. Branching processes with immigration and related limit theorems. Theory Probab Appl 1971;16(1): 36–54.
62. Lambert A. The genealogy of continuous-state branching processes with immigration. Probab Theory Relat Fields 2002;122(1): 42–70.
63. Duquesne T. Continuum random trees and branching processes with immigration. Stoch Process Appl 2009;119(1): 99–129.

64. Watanabe S. A limit theorem of branching processes and continuous state branching processes. J Math Kyoto Univ 1968;8: 141–167.
65. Dynkin EB. Superprocesses and partial differential equations. Ann Probab 1993; 21(3):1185–1262.
66. Li Z. Measure-valued branching Markov processes. Probability and its Applications. Springer: Berlin Heidelberg; 2011.