CrossMark

*Adv. Appl. Prob.* **53,** 251–278 (2021)
*doi:*10.1017/apr.2020.59

# BRANCHING PROCESSES WITH INTERACTIONS: SUBCRITICAL COOPERATIVE REGIME

ADRIÁN GONZÁLEZ CASANOVA,* *Universidad Nacional Autónoma de México*
JUAN CARLOS PARDO** *** AND
JOSÉ LUIS PÉREZ,** **** *Centro de Investigación en Matemáticas*

**Abstract**

In this paper, we introduce a family of processes with values on the nonnegative integers that describes the dynamics of populations where individuals are allowed to have different types of interactions. The types of interactions that we consider include pairwise interactions, such as competition, annihilation, and cooperation; and interactions among several individuals that can be viewed as catastrophes. We call such families of processes *branching processes with interactions*. Our aim is to study their long-term behaviour under a specific regime of the pairwise interaction parameters that we introduce as the *subcritical cooperative regime*. Under such a regime, we prove that a process in this class comes down from infinity and has a moment dual which turns out to be a jump-diffusion that can be thought as the evolution of the frequency of a trait or phenotype, and whose parameters have a classical interpretation in terms of population genetics. The moment dual is an important tool for characterizing the stationary distribution of branching processes with interactions whenever such a distribution exists; it is also an interesting object in its own right.

*Keywords:* Branching coalescing processes; stability; moment duality; population genetics

2010 Mathematics Subject Classification: Primary 60K35
Secondary 60J80

## 1. Introduction and main results

Branching processes are one of the most important families of probabilistic models that describe the dynamics of a given population. The simplest branching model is the so-called Bienaymé–Galton–Watson (BGW) process, which is a Markov chain whose time steps are non-overlapping generations, with individuals reproducing independently and giving birth to a (random) number of offspring in the next generation. These random offspring have all the same probability distribution.

Received 18 July 2018; revision received 28 July 2020.
* Postal address: Instituto de Matemáticas, Universidad Nacional Autónoma de México (UNAM), Área de la Investigación Científica, Circuito Exterior, Ciudad Universitaria, 04510, México, D.F. Email address: adriangcs@matem.unam.mx
** Postal address: Centro de Investigación en Matemáticas, A.C., Calle Jalisco s/n, Col. Valenciana CP 36023 Guanajuato, Gto., México, Apartado Postal 402, CP 36000.
*** Email address: jcpardo@cimat.mx
**** Email address: jluis.garmendia@cimat.mx

© The Author(s), 2021. Published by Cambridge University Press on behalf of Applied Probability Trust.



https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

A similar model can also be introduced in the continuous-time setting. In this case, each individual possesses an exponential clock such that, when it rings, the individual dies and is replaced by a random number of offspring. The number of offspring and the exponential clock associated to each individual are independent and identically distributed. This model possesses overlapping generations and is known as a BGW process in continuous time. Since we are only interested in the continuous-time setting, we will refer to these processes as BGW processes, omitting the phrase 'continuous time'.

In order to make this probabilistic model more realistic, many authors have introduced different types of density-dependence to branching processes (see for instance Jagers [18], Lambert [22], and the references therein). One approach consists in generalizing the birth and death rates of continuous-time branching processes by considering polynomial rates as functions of the population size. This way of modelling density-dependence seems to be popular in the biology community (see for instance Mattis and Kiffe [27] and Nåsell [31]).

In this manuscript, we follow this approach by considering polynomial rates as functions of the population size that can be interpreted as different types of interactions between individuals. To be more precise, we are interested in a model that considers several specific phenomena, such as (pairwise) *competition pressure*, *annihilation*, and *cooperation*, as well as interactions among several individuals that can be viewed as *catastrophes*. We refer to this family of processes as *branching processes with interactions.*

Before we provide a formal definition of branching processes with interactions, we fix some notation and recall some examples that have already appeared in the literature. We define $\mathbb{N} := \{1, 2, \dots\}$, the set of strictly positive integers; $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, the set of nonnegative integers; $\overline{\mathbb{N}} := \mathbb{N}_0 \cup \{\infty\}$, the compactification of $\mathbb{N}_0$ induced by the metric $d(n, m) = |1/n - 1/m|$; $\mathbb{R}_+ := (0, \infty)$, the set of strictly positive real numbers; and $\mathbb{R}_{+,0} := [0, \infty)$.

Our first example is the so-called logistic branching process, which was studied in depth by Lambert in [22]. In this model, each individual produces a random number of offspring independently of the others, as in BGW processes, but also experiences competition pressure; in other words, each pair of individuals interacts at a fixed rate and one of them is killed as result of this interaction. The logistic branching process $L = (L_t, t \in \mathbb{R}_{+,0})$, with positive (i.e. in $\mathbb{R}_{+,0}$) parameters $c$ and $(\pi_i, i \in \mathbb{N})$ such that $\sum_{i\in\mathbb{N}} \pi_i = \rho > 0$, is a continuous-time Markov chain with values in $\overline{\mathbb{N}} := \mathbb{N} \cup \{0, \infty\}$, with infinitesimal generator $Q = (q_{i,j})_{i,j\in\overline{\mathbb{N}}}$, where

$$q_{i,j} = \begin{cases} i\pi_{j-i} & \text{if } i \in \mathbb{N} \text{ and } j > i, \\ di + ci(i-1) & \text{if } i \in \mathbb{N} \text{ and } j = i-1, \\ -i(d + \rho + c(i-1)) & \text{if } i \in \mathbb{N} \text{ and } j = i, \\ 0 & \text{otherwise.} \end{cases}$$

Observe that $\pi_i/\rho$ represents the probability of having $i$ new individuals born at each reproduction event. It is important to note that the states $\{0, \infty\}$ are absorbing states. Moreover, from the main results in Lambert [22] (see Theorems 2.2 and 2.3 in [22]), we deduce that the log-moment condition

$$\sum_{i\in\mathbb{N}} \log(i)\pi_i < \infty$$

is sufficient to ensure that the logistic branching process $L$ does not explode in finite time, almost surely (a.s.).

It turns out that the logistic branching process is also useful in the field of population genetics. Indeed, it appears in a *duality* relationship with the frequency of a phenotype

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

with selective disadvantage in a given population, which can be modelled by the following stochastic differential equation (SDE) with values in [0, 1]:

$$X_t = x - \rho \int_0^t X_s(1 - X_s)\mathrm{d}s + \int_0^t \sqrt{X_s(1 - X_s)}\mathrm{d}B_s,$$

where $x \in [0, 1]$, $\rho > 0$, and $B = (B_t, t \in \mathbb{R}_{+,0})$ is a standard Brownian motion (see Krone and Neuhauser [20, 21]). Krone and Neuhauser observed that one can study the above SDE using the block counting process of the ancestral selection graph, which turns out to be a particular case of the logistic branching process. Specifically, if we take $c = 1$, $d = 0$, $\pi_1 = \rho$, and $\pi_i = 0$ for all $i \in \mathbb{N} \setminus \{1\}$, in the logistic branching process $L$ defined before, the moments of $X = (X_t, t \in \mathbb{R}_{+,0})$ can be written in terms of the moment-generating function of $L$ as follows:

$$\mathbb{E}_x[X_t^n] = \mathbf{E}_n[x^{L_t}], \qquad x \in [0, 1], \quad n \in \mathbb{N}, \quad t \in \mathbb{R}_{+,0},$$

where $\mathbb{E}_x$ and $\mathbf{E}_n$ denote the expectations of $X$ starting from $x$ and $L$ starting from $n$ individuals, respectively.

The above relationship is known as *moment duality* and appears between many interesting branching processes with interactions and frequency processes that arise in population genetics. For instance, Athreya and Swart [2] considered the following moment duality. Let $\rho, c, d \in \mathbb{R}_{+,0}$, and denote by $C = (C_t, t \in \mathbb{R}_{+,0})$ the process that counts the number of particles of the branching–coalescing process defined by the initial value $C_0 = n$ and the following dynamics: each particle splits into two particles at rate $\rho$, each particle dies at rate $d$, and each ordered pair of particles coalesces into one particle at rate $c$. All these events occur independently of each other. The authors of [2] called this process the $(1, \rho, c, d)$-braco-process. Note that the braco-process is also a particular case of the logistic branching process with parameters $c, d, \pi_1 = \rho$, and $\pi_i = 0$ for all $i \in \mathbb{N} \setminus \{1\}$. Its dual process $X = (X_t, t \in \mathbb{R}_{+,0})$ is the unique strong solution taking values in [0, 1] of the SDE

$$X_t = x_0 + \int_0^t (\rho - d)X_s\mathrm{d}s - \int_0^t \rho X_s^2\mathrm{d}s + \int_0^t \sqrt{2cX_s(1 - X_s)}\mathrm{d}B_s, \qquad t \in \mathbb{R}_{+,0},$$

where $x_0 \in [0, 1]$ and $B$ is a standard Brownian motion. Athreya and Swart called this process the resampling–selection process with selection rate $\rho$, resampling rate $c$, and mutation rate $d$, or, for short, the $(1, \rho, c, d)$-resem-process. In particular, they observed the following moment duality:

$$\mathbb{E}_x[(1 - X_t)^n] = \mathbf{E}_n\big[(1 - x)^{C_t}\big], \qquad x \in [0, 1], \quad n \in \mathbb{N}, \quad t \in \mathbb{R}_{+,0},$$

where $\mathbb{E}_x$ and $\mathbf{E}_n$ denote the expectations of $X$ starting from $x$ and $C$ starting from $n$ individuals, respectively. Recently, Alkemper and Hutzenthaler [1] derived a unified stochastic picture for the moment duality between the resampling–selection model and the branching–coalescing process of Athreya and Swart. It is important to note that the previous duality relationships include the moment duality between the Wright–Fisher diffusion and the so-called Kingman's coalescent.

Other type of duality relationships have been considered for haploid population models and two-sex population models by Möhle [30] and for the Wright–Fisher diffusions with $d$-types and the Moran model (in both the presence and the absence of mutation) by Carinci *et al.* [7].

Because of the power of this relationship, the question of which models admit a moment duality is interesting in its own right. In Section 2.1, we provide conditions for this moment duality to hold for a large family of branching processes with interactions that include existing examples in the literature. However, the focus of this manuscript is the long-term behaviour of

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

branching processes with interactions, and we use the moment duality technique as a tool to determine the invariant distribution whenever it exists. In particular, we consider interactions that have appeared independently in the literature, and that have previously been studied using moment duality.

As in the examples of above, we are interested in branching processes where individuals reproduce and die, as in the BGW process, but also are allowed to have different types of interactions. To be more precise, our model has the following dynamics:

(i) *death:* each individual in the population dies at rate $d \in \mathbb{R}_{+,0}$;
(ii) *reproduction:* each individual produces $i$ new individuals at rate $\pi_i \in \mathbb{R}_{+,0}$, for $i \in \mathbb{N}$;
(iii) *competition pressure:* each pair of individuals interact at a fixed rate $c \in \mathbb{R}_{+,0}$, and one of them is killed as result of this interaction (see for instance Athreya and Swart [2] and Lambert [22]);
(iv) *annihilation:* each pair of individuals interact at a fixed rate $a \in \mathbb{R}_{+,0}$, and both of them are killed as result of this interaction (see for instance Athreya and Swart [3] and Blath and Kurt [5]);
(v) *cooperation:* each pair of individuals interact and produce $i$ new individuals at rate $b_i \in \mathbb{R}_{+,0}$, for $i \in \mathbb{N}$ (see for instance Sturm and Swart [36]).

Finally, we consider interactions among several individuals in the sense of $\Lambda$-coalescent events as in Foucart [12] and Griffiths [16], which we call *catastrophes*. Let $\Lambda$ be a finite measure on [0, 1]. If $n$ individuals are present in the population, each $(k-1)$-tuple die simultaneously at rate

$$\lambda_{n,k} := \int_0^1 y^k (1-y)^{n-k} \frac{\Lambda(\mathrm{d}y)}{y^2}, \qquad \text{for} \quad k \in \{2, 3, \ldots, n\}.$$

We refer the reader to Pitman [33] and Sagitov [35] for a formal definition of $\Lambda$-coalescent events and processes.

Throughout the paper, we assume that all the parameters $d$, $c$, $a$, $b_i$, $\pi_i$ for $i \in \mathbb{N}$ are positive (i.e. belong to $\mathbb{R}_{+,0}$), and $\Lambda$ is a finite measure on [0, 1], such that

$$0 \le \rho := \sum_{i\in\mathbb{N}} \pi_i < \infty, \qquad 0 \le \lambda_i := \sum_{k=2}^{i} \binom{i}{k} \lambda_{i,k}, \quad \text{and} \quad 0 \le b := \sum_{i\in\mathbb{N}} b_i < \infty.$$

Note that when $\rho = 0$ (similarly when $b = 0$), we have $\pi_i = 0$ (and respectively $b_i = 0$) for all $i \in \mathbb{N}$.

The branching process with interactions $Z = (Z_t, t \in \mathbb{R}_{+,0})$ is a continuous-time Markov chain with values in $\overline{\mathbb{N}}$ whose extended generator $Q = (q_{i,j})_{i,j\in\overline{\mathbb{N}}}$ is given by

$$q_{i,j} = \begin{cases} i\pi_{j-i} + i(i-1)b_{j-i} & \text{if } i \in \mathbb{N} \text{ and } j > i, \\ di + ci(i-1) + \binom{i}{2}\lambda_{i,2} & \text{if } i \in \mathbb{N}\setminus\{1\} \text{ and } j = i-1, \\ ai(i-1) + \binom{i}{3}\lambda_{i,3} & \text{if } i \in \{3, 4, \ldots\} \text{ and } j = i-2, \\ \binom{i}{k}\lambda_{i,k} & \text{if } i \in \{4, \ldots\},\ k \in \{4, \ldots i\} \text{ and } j = i-k+1, \\ -i(d+\rho) - i(i-1)(b+c+a) - \lambda_i & \text{if } i \in \mathbb{N}\setminus\{1\} \text{ and } j = i, \\ 2a & \text{if } i = 2 \text{ and } j = 0, \\ d & \text{if } i = 1 \text{ and } j = 0, \\ -d-\rho & \text{if } i = 1 \text{ and } j = 1, \\ 0 & \text{otherwise.} \end{cases}$$

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

Equivalently, the extended generator $Q$ of $Z$ acts as follows. For $f \in C_0(\overline{\mathbb{N}}, \mathbb{R})$, the set of continuous functions from $\overline{\mathbb{N}}$ to $\mathbb{R}$ which vanish at infinity, we have

$$Qf(0) = Qf(\infty) = 0, \qquad Qf(1) = d(f(0) - f(1)) + \sum_{i\in\mathbb{N}} \pi_i(f(1+i) - f(1)),$$

and for all $n \in \mathbb{N} \setminus \{1\}$,

$$\begin{aligned} Qf(n) =& dn(f(n-1) - f(n)) + \sum_{i\in\mathbb{N}} n\pi_i(f(n+i) - f(n)) \\ &+ n(n-1)a(f(n-2) - f(n)) + n(n-1)c(f(n-1) - f(n)) \\ &+ n(n-1)\sum_{i\in\mathbb{N}} b_i(f(n+i) - f(n)) + \sum_{k=2}^{n} \binom{n}{k} \lambda_{n,k}(f(n-k+1) - f(n)), \end{aligned} \tag{1.1}$$

whenever it is well-defined. For instance the domain of the extended generator $Q$, here denoted by $\mathcal{D}(Q)$, includes the set of linear combinations of exponential functions of the form

$$f(n) = \sum_{\ell=1}^{k} a_\ell e^{-m_\ell n} \tag{1.2}$$

for $a_1, a_2, \ldots, a_k \in \mathbb{R}$, $m_1, m_2, \ldots, m_k \in \mathbb{R}_+$, and $k \in \mathbb{N}$, which is dense in $C_0(\overline{\mathbb{N}}, \mathbb{R})$ by the Stone–Weierstrass theorem for locally compact sets. Indeed, the latter claim follows from the construction in Section 4.2 and Problem 15(a) in Section 4.11 of [11] (we leave the details to the interested reader).

For our purposes, we introduce the parameters

$$\mathbf{m} := -d + \sum_{i\in\mathbb{N}} i\pi_i \qquad \text{and} \qquad \varsigma := -c - 2a + \sum_{i\in\mathbb{N}} ib_i,$$

which we assume to be finite. As we will see below, we can characterize the long-term behaviour of branching processes with interactions depending on the value of what we call the *cooperative parameter* $\varsigma \in \mathbb{R}$. We say that a branching process with interactions is *supercritical*, *critical*, or *subcritical cooperative* accordingly as $\varsigma > 0$, $\varsigma = 0$, or $\varsigma < 0$.

In order to understand the long-term behaviour of this family of processes, we use stability theory for continuous-time Markov chains (see for instance Tweedie [37], Chen [8], Meyn and Tweedie [28], and the notes of Hairer [17]). Our first main result says that a branching process with interactions $Z$ which is subcritical or critical cooperative, i.e. $\varsigma \le 0$, is conservative, or, in other words, that it does not explode in finite time. In the particular case when there are no catastrophes and $\varsigma > 0$, the process $Z$ explodes in finite time with positive probability.

In the sequel, we denote by $\mathbf{P}_n$ the law of the process $Z$ starting from $n \in \mathbb{N}_0$. Since $\{0\}$ is an absorbing state, we consider only the case where $Z_0$ takes values in $\mathbb{N}$, a.s.

**Theorem 1.** *Let $Z$ be a branching process with interactions such that $|\mathbf{m}| < \infty$. If the process is subcritical or critical cooperative, i.e. $\varsigma \le 0$, then the process $Z$ is conservative, i.e.*

$$\mathbf{P}_n(Z_t < \infty) = 1, \qquad \text{for any} \quad n \in \mathbb{N}, \ \ t \in \mathbb{R}_{+,0}.$$

*Moreover, if there are no catastrophes, i.e. $\Lambda \equiv 0$, and $Z$ is supercritical cooperative (i.e. $\varsigma > 0$) starting at $Z_0 \in \{2, 3, ...\}$ a.s., then the process $Z$ explodes in finite time with positive probability.*

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

The proof of the first part of our previous result follows from a Lyapunov-type condition for non-explosion found in Chen [8], and for the supercritical case we use a coupling argument. The supercritical cooperative regime seems to be more involved in the event of catastrophes. Nonetheless, we believe that there must be cases when the process may explode in the presence of catastrophes.

In the sequel, we assume that the process satisfies $|\mathbf{m}| < \infty$ and that it is subcritical cooperative, i.e $\varsigma < 0$, unless we specifically state otherwise. It is important to note that when there is no annihilation, i.e. $a = 0$, the previous assumption implies that $c > 0$.

Our next result characterizes the long-term behaviour of $Z$. For simplicity, we only deal with the case with no annihilation; the annihilation case is more involved and will be studied at the end of this section.

**Proposition 1.** *Assume that $|\mathbf{m}| < \infty$ and that $Z$ is a subcritical cooperative branching process with interactions, i.e. $\varsigma < 0$, such that $\mathbb{P}(Z_0 \in \mathbb{N}) = 1$. If there is no annihilation, i.e. $a = 0$, we have the following:*

(i) *If $d = 0$ and $\rho = 0$, then $Z$ is absorbed in the state $\{1\}$, and $\{0\}$ is not accessible.*
(ii) *If $d = 0$ and $\rho > 0$, then $\{0\}$ is not accessible, and $Z$ is positive recurrent in $\mathbb{N}$.*
(iii) *If $d > 0$, then $Z$ is absorbed in the state $\{0\}$.*

The proof of Part (ii) of this result follows from the Foster–Lyapunov conditions for positive recurrence, and the remaining cases follow from a coupling argument. Recurrence in the critical cooperative case does not seem easy to handle. Actually a different approach from the one we present here for the subcritical cooperative case is needed, and we conjecture that the criteria for recurrence depend not only on the cooperative parameters but also on the branching parameters. Further developments on the critical case appear in González Casanova *et al.* [14], where a particular example is treated.

When $a = 0 = d$ and $\rho > 0$, the state $\{0\}$ is not accessible, and since $Z$ is irreducible (as a process with values in $\mathbb{N}$) and positive recurrent, there exists a unique stationary distribution (see for instance Theorem 21.14 in [23]). Thus a natural question arises: *can we determine the invariant distribution of $Z$?* In order to provide a positive answer to this question, we first introduce the moment dual of $Z$, which also exists in the critical cooperative case.

The unique moment dual of the branching process with interactions $Z$ is a jump-diffusion taking values in $[0, 1]$ that can be defined as the unique strong solution of the following SDE:

$$\mathrm{d}X_t = \mu(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t + \int_{(0,1]}\int_{(0,1]} z\big(\mathbf{1}_{\{u\le X_{t-}\}} - X_{t-}\big)\widetilde{N}(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u), \tag{1.3}$$

with initial condition $X_0 = x \in [0, 1]$. Here $\widetilde{N}$ is a compensated Poisson random measure on $\mathbb{R}_{+,0} \times (0, 1]^2$ with intensity $\mathrm{d}t z^{-2}\Lambda(\mathrm{d}z)\mathrm{d}u$, and the functions $\mu : [0, 1] \to \mathbb{R}$ and $\sigma : [0, 1] \to \mathbb{R}_{0,+}$ are continuous and satisfy

$$\mu(z) := d(1 - z) + \sum_{i\in\mathbb{N}} \pi_i(z^{i+1} - z) \quad \text{and} \quad \sigma(z) := \sqrt{2c(z - z^2) + 2\sum_{i\in\mathbb{N}} b_i(z^{i+2} - z^2)}.$$

We denote by $\mathbb{P}_x$ the law of the process $X = (X_t, t \in \mathbb{R}_{0,+})$ starting from $x \in [0, 1]$.

**Theorem 2.** *Assume that $|\mathbf{m}| < \infty$, $a = 0$, and $\varsigma \le 0$. Then the SDE (1.3) with starting point $X_0 \in [0, 1]$ has a unique strong solution, which we denote by $X$, taking values on $[0, 1]$.*

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

*Moreover, $X$ is the unique (in distribution) moment dual of $Z$, the (sub)critical cooperative branching process with interactions having the same parameters as $X$. More precisely, for $x \in [0, 1]$ and $n \in \mathbb{N}_0$, we have*

$$\mathbb{E}_x[X_t^n] = \mathbf{E}_n\left[x^{Z_t}\right] \qquad \text{for} \quad t \in \mathbb{R}_{+,0}.$$

**Remark 1.** Recall that $a = 0$. The assumption $\varsigma \leq 0$ guarantees that

$$c(x - x^2) + \sum_{i \in \mathbb{N}} b_i\left(x^{i+2} - x^2\right) \geq 0 \qquad \text{for all} \quad x \in [0, 1],$$

and thus $\sigma : [0, 1] \mapsto \mathbb{R}_{+,0}$. Indeed, the latter follows since

$$c(x - x^2) + \sum_{i \in \mathbb{N}} b_i(x^{i+2} - x^2) = x(1 - x)\left(c - \sum_{i \in \mathbb{N}} b_i \sum_{j=0}^{i-1} x^{j+1}\right),$$

implying that the previous quantity is always positive for any $x \in [0, 1]$, if the mapping

$$x \in [0, 1] \mapsto g(x) := c - \sum_{i \in \mathbb{N}} b_i \sum_{j=0}^{i-1} x^{j+1}$$

is positive. On the other hand, the function $g$ is decreasing on $[0, 1]$ and satisfies $g(0) = c$. Under the assumption that $\varsigma < 0$, we have that $c > 0$ and $g(1) = c - \sum_{i \in \mathbb{N}} i b_i = -\varsigma > 0$. If $\varsigma = 0$, we have that $c \geq 0$ (if $c = 0$ then $b = 0$) and $g(1) = c - \sum_{i \in \mathbb{N}} i b_i = 0$. Moreover, the assumption that $X_0 \in [0, 1]$ guarantees that the SDE (1.3) takes values in $[0, 1]$.

The process $X$ can be thought as the evolution of the frequency of a trait or phenotype, and some of the parameters have a classical interpretation in terms of population genetics. For instance, $d$ represents the rate at which a mutation affects an individual, $\pi_1$ has been interpreted as the weak selection parameter, $c$ is also known as the strength of the random genetic drift, and the Poisson random measure $N$ may model the occurrence of reproduction events that affect large fractions of the population. Recently, the parameters $(\pi_i, i \geq 2)$ have been interpreted in terms of frequency-dependent selection by Gonzalez Casanova and Spanò [15].

The interpretation of the parameters $(b_i, i \in \mathbb{N})$ is not so classical. In Gonzalez Casanova *et al.* [14], a biological interpretation of the parameter $b_1$ is studied which can be related to the efficiency of individuals. To be more precise, let us imagine a Wright–Fisher model where the population size is coupled with the frequency of individuals of a given type, in such a way that its frequency process $X^{(N)} = (X_n^{(N)}, n \in \mathbb{N}_0)$ is a Markov chain taking values in $[0, 1]$ such that conditionally on $X_{n-1}^{(N)} = x \in [0, 1]$, the random variable $\lfloor N_x \rfloor X_n^N$ is distributed as a binomial random variable with parameters $\lfloor N_x \rfloor$ and $x$, where

$$N_x = \frac{N}{1 - b_1 x},$$

and $x \mapsto \lfloor x \rfloor$ denotes the floor function. The interpretation of $N_x$ is related to efficiency; indeed, assume that at each generation there are $N$ units of resources available, and suppose that there are two types of individuals: the *inefficient* population, consisting of individuals that need 1 unit of resources to be created, and the *efficient* population, consisting of individuals that require

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

$1 - b_1$ units of resources. Since, in a given generation, the frequency of efficient individuals is $x$, the number of individuals that the next generation will be able to sustain is approximately $N_x$. In other words, we can interpret the process $X^{(N)}$ as the frequency of efficient individuals. Now observe that for all $x \in [0, 1]$, we have

$$\mathbb{E}\left[X_n^{(N)} \middle| X_{n-1}^{(N)} = x\right] = x, \qquad \mathbb{V}\mathrm{ar}\left[X_n^{(N)} \middle| X_{n-1}^{(N)} = x\right] = \frac{x(1-x)}{\lfloor N_x \rfloor} \sim \frac{x(1-x)(1-b_1 x)}{N},$$

as $N$ increases. From the previous observation, it is not so difficult to guess that the scaling limit of $(X^{(N)}_{\lfloor Nt \rfloor}, t \geq 0)$, as $N$ goes to infinity, must be the unique strong solution of the SDE

$$\mathrm{d}X_t = \sqrt{2X_t(1-X_t)(1-b_1 X_t)}\mathrm{d}B_t,$$

which belongs to the family of diffusions described by the SDE (1.3) when $c = 1$, $b_1 \in (0, 1)$, and $\Lambda \equiv 0$. In other words, the parameters $(b_i, i \in \mathbb{N})$ can then be understood as frequency-dependent effective population size, in the sense of Gonzalez Casanova and Spanò [15].

It is important to note that recently, Foucart [12] and Griffiths [16] have studied asymptotic properties of the so-called $\Lambda$-Wright–Fisher process with selection using moment duality. The latter process can be defined as the unique strong solution of the SDE (1.3) with $\mu(x) = \rho(x^2 - x)$ and $\sigma(x) = 0$, and its unique moment dual is a binary branching process with $\Lambda$-catastrophes. Foucart and Griffiths were interested in understanding the conditions under which the solution of such an SDE eventually goes to zero with probability one. This question can be interpreted in a biological sense as follows: *under what conditions is fixation of the fittest phenotype certain, in a population with skewed reproduction?*

As we have mentioned before, the moment duality property has recently been used by Gonzalez Casanova and Spanò [15] for different purposes. The authors of [15] studied a model related to selection which happens to be moment dual to the solution of the SDE (1.3) with $c, d = 0$ and $b_i = 0$ for $i \in \mathbb{N}$; they used moment duality to understand the conditions under which fixation of the fittest phenotype is certain. It is important to note that the results that we present here are complementary to those obtained in Foucart [12], Griffiths [16], and Gonzalez Casanova and Spanò [15]. As we will see below, fixation at 1 of the process $X$ is related to determining the invariant distribution of $Z$.

One of our main results shows that when there is no annihilation, $d = 0$, and $\rho > 0$, the invariant distribution of $Z$ can be determined by the fixation at 1 of its moment dual $X$.

**Theorem 3.** *Assume that there is no annihilation, i.e. $a = 0 = d$, $|\mathbf{m}| < \infty$, and $\rho > 0$. Then a subcritical cooperative branching process with interactions Z, i.e. $\varsigma < 0$, has a unique stationary distribution, here denoted by $\nu$. Moreover,*

$$\lim_{t\to\infty} \sup_{n\in\mathbb{N}} ||\mathbf{P}_n(Z_t = \cdot) - \nu(\cdot)|| = 0, \tag{1.4}$$

*and if X is the unique moment dual of Z, then $\mathbb{P}_x(X_\infty \in \{0, 1\}) = 1$ for $x \in (0, 1)$, where $X_\infty := \lim_{t\to\infty} X_t$ and the generating function associated to $\nu$ satisfies*

$$f(x) := \sum_{n\in\mathbb{N}} x^n \nu(n) = \mathbb{P}_x(X_\infty = 1).$$

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

It is important to note that when there are no catastrophes, i.e. $\Lambda \equiv 0$, the invariant distribution of Z can be determined explicitly in terms of the scale function of the diffusion described by the SDE (1.3), as explained below.

**Corollary 1.** *Assume that* $|\mathbf{m}| < \infty$ *and* $\varsigma < 0$*, there is no annihilation (i.e.* $a = 0$*), and there are no catastrophes (i.e.* $\Lambda \equiv 0$*). If* $d = 0$ *and* $\rho > 0$*, we have*

$$\sum_{n\in\mathbb{N}} x^n \nu(n) = \mathbb{P}_x(X_{T_{0,1}} = 1) = \frac{S(x) - S(0)}{S(1) - S(0)}, \qquad \text{for} \quad x \in [0, 1],$$

*where* $T_{0,1} = \inf\{t \geq 0 : X_t \in \{0, 1\}\}$*, and S denotes the so-called scale function of X, which satisfies*

$$S(x) = \int_{\theta_1}^{x} \exp\left\{-\int_{\theta_2}^{y} \frac{2\mu(z)}{\sigma^2(z)} \mathrm{d}z\right\} \mathrm{d}y,$$

*where* $\theta_1$ *and* $\theta_2$ *are arbitrary numbers in* [0, 1].

The previous corollary follows from Theorem 3, the theory of exit problems for diffusions (see for instance Revuz and Yor [34]), and Feller's boundary test (see for instance Proposition 2.4 in [29]). Indeed, from Feller's boundary test, the process is absorbed at 1 (respectively, at 0) at finite time with positive probability if $S(1) < \infty$ ($|S(0)| < \infty$) and

$$\int^{1} \frac{S(1) - S(u)}{\sigma^2(u)S'(u)} \mathrm{d}u < \infty \qquad \left(\int_0 \frac{S(u) - S(0)}{\sigma^2(u)S'(u)} \mathrm{d}u < \infty\right).$$

Both conditions at 1 and at 0 are fulfilled in our case since

$$\frac{2\mu(z)}{\sigma^2(z)} \sim -\frac{\rho}{c} \quad \text{as} \quad z \to 0, \qquad \frac{2\mu(z)}{\sigma^2(z)} \sim \frac{\mathbf{m}}{\varsigma} \quad \text{as} \quad z \to 1,$$

and

$$\frac{S(1) - S(u)}{\sigma^2(u)S'(u)} \sim -\frac{1}{\varsigma} \quad \text{as} \quad u \to 1, \qquad \frac{S(u) - S(0)}{\sigma^2(u)S'(u)} \sim \frac{1}{c} \quad \text{as} \quad u \to 0.$$

Before we continue with our exposition, we study an interesting example, which is related to the model of efficient individuals studied in [14] and where the invariant distribution of Z can be computed in closed form. Assume that there are no catastrophes, and let $d = 0$, $\pi_1 = \rho > 0$, $b_1 = b > 0$, $c > b$, and $\pi_i = 0 = b_i$ for $i \geq 2$. In this particular case, the unique strong solution X of the SDE (1.3) is a diffusion whose parameters are given by

$$\mu(x) := \rho(x^2 - x) \qquad \text{and} \qquad \sigma^2(x) := 2x(1 - x)(c - bx).$$

If $2\rho \neq b$, the scale function associated to the diffusion X satisfies

$$S(x) = \int_0^x \exp\left\{\frac{\rho}{b}\int_0^y \frac{\mathrm{d}v}{\frac{c}{b} - v}\right\} \mathrm{d}y = K\left(1 - \left(1 - \frac{b}{c}x\right)^{1-\frac{\rho}{b}}\right), \quad \text{for} \quad x \in [0, 1],$$

where K is a constant that depends on $(c, b, \rho)$ and is positive or negative accordingly as $\rho < b$ or $b < \rho$. If $\rho = b$, then

$$S(x) = K \ln\left(\frac{1}{1 - \frac{b}{c}x}\right), \quad \text{for} \quad x \in [0, 1],$$

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

with K positive. Therefore, if $b < \rho$, we have

$$\sum_{n\geq 0} x^n \nu(n) = \mathbb{P}_x(X_{T_{0,1}} = 1) = \frac{S(x) - S(0)}{S(1) - S(0)} = C_{c,b,\rho}\left(\left(1 - \frac{b}{c}x\right)^{1-\frac{\rho}{b}} - 1\right),$$

where

$$C_{c,b,\rho} = \left(\left(1 - \frac{b}{c}\right)^{1-\frac{\rho}{b}} - 1\right)^{-1}.$$

By the binomial theorem, we deduce

$$\nu(k) = C_{c,b,\rho}\frac{\Gamma\left(\frac{\rho}{b} - 1 + k\right)}{\Gamma\left(\frac{\rho}{b} - 1\right)k!}\left(\frac{b}{c}\right)^k, \qquad k \geq 1,$$

where $\Gamma$ denotes the so-called gamma function. The invariant distribution for the case $b > \rho$ is the same, but instead of using the binomial theorem we need to use the generating function of a beta-geometric distribution with parameters $(1 - \frac{\rho}{b}, \frac{\rho}{b})$.

In the case $\rho = b$, the representation of the function $f(z) = -\ln(1 - z)$, for $z \in [0, 1)$, as a hypergeometric function leads to

$$\nu(k) = \widetilde{C}_{c,b}\left(\frac{b}{c}\right)^k \frac{1}{k}, \qquad k \geq 1,$$

with

$$\widetilde{C}_{c,b} = \left(\ln\left(\frac{1}{1 - \frac{b}{c}}\right)\right)^{-1}.$$

The moment duality is also useful for showing that a subcritical cooperative branching process with interactions and no annihilation comes down from infinity. Formally, we define the law $\mathbf{P}_\infty$ starting from infinity with values in $\overline{\mathbb{N}}$ as the limit of the laws $\mathbf{P}_n$ of the process issued from $n$. When the limiting process is nondegenerate, it hits finite values in finite time with positive probability. Moreover, we say that the process $Z$ *comes down from infinity* if $\mathbf{P}_\infty(Z_t < \infty) = 1$ for all $t > 0$.

**Theorem 4.** *Assume* $|\mathbf{m}| < \infty$ *and* $a = 0$. *Then a subcritical cooperative (i.e.* $\varsigma < 0$*) branching process with interactions Z comes down from infinity. Furthermore, we have that*

$$\mathbf{E}_\infty[\tau_1] < \infty,$$

*where* $\tau_1 = \inf\{t > 0 : Z_t = 1\}$.

Another interesting consequence of the moment duality and Proposition 1 is the following result, which describes the asymptotic behaviour of jump-diffusions of the form (1.3). In particular, we can determine the probability of fixation at 1 of the jump-diffusion $X$.

**Corollary 2.** *Let X be the unique strong solution of* (1.3) *which is the unique moment dual of a subcritical cooperative branching process with interactions Z and no annihilation (i.e.* $\varsigma < 0$ *and* $a = 0$*) and* $X_\infty := \lim_{t\to\infty} X_t$. *Then, for all* $x \in [0, 1]$,

*(i) if* $d = 0$ *and* $\rho = 0$,

$$\mathbb{P}_x(X_\infty = 1) = x = 1 - \mathbb{P}_x(X_\infty = 0);$$

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

*(ii) if $d = 0$ and $\rho > 0$,*

$$\mathbb{P}_x(X_\infty = 1) = \sum_{i \in \mathbb{N}} x^i \nu(i),$$

*where $\nu$ is the unique stationary distribution of $Z$;*

*(iii) if $d > 0$,*

$$\mathbb{P}_x(X_\infty = 1) = 1.$$

Finally, we study the case $a > 0$. This case has been studied before in different contexts; see for instance [3, 4, 5, 6] and the references therein. It is important to note that in this case monotonicity is lost, in the sense that it is not true that a bigger population has more probability of survival. For instance, if there are two individuals (or particles) the probability of extinction is higher than if there is only one individual (or one particle). Monotonicity is a very important and useful property; thus the case $a > 0$ seems to be technically more involved, and many properties of processes with annihilation events remain unknown.

We point out that in this case there is also a moment duality relationship between a branching process with interactions and a jump-diffusion similar to (1.3). In this case, the existence of a unique strong solution is more complicated than the case $a = 0$. We will not prove this claim, since in this case there is no relevant interpretation for the moment dual of $Z$, and we can study its long-term behaviour via a coupling argument. Moreover, it seems more involved to determine the invariant distribution of $Z$ (whenever it exists) in terms of the probability of fixation at, for example, $-1$ and 1.

Finally, our last result provides a detailed description of the long-term behaviour of $Z$ in the case when $a > 0$. As stated before, the behaviour of $Z$ is more involved than in the case with no annihilation.

**Proposition 2.** *Let $Z$ be a subcritical cooperative branching process with interactions (i.e. $\varsigma < 0$) with $a > 0$, and let*

$$\tau_{0,1} = \inf\{t > 0 : Z_t \in \{0, 1\}\} \qquad \text{and} \qquad \tau_i = \inf\{t > 0 : Z_t = i\},$$

*for $i \in \{0, 1\}$.*

*(i) Suppose $d = 0$.*

*(a) If either $\Lambda(0, 1] > 0$, $c > 0$, or $\sum_{i \in \mathbb{N}} \pi_{2i-1} + \sum_{i \in \mathbb{N}} b_{2i-1} > 0$, then we have the following:*

*(a.1) If $\rho = 0$, then $Z$ gets absorbed in $\{0, 1\}$, and $\sup_{n \in \mathbb{N}} \mathbf{E}_n[\tau_{0,1}] < \infty$.*

*(a.2) If $\rho > 0$, then $Z$ gets absorbed in $\{0\}$, and $\sup_{n \in \mathbb{N}} \mathbf{E}_n[\tau_0] < \infty$.*

*(b) If $\Lambda(0, 1] = 0$, $c = 0$, and $\sum_{i \in \mathbb{N}} \pi_{2i-1} + \sum_{i \in \mathbb{N}} b_{2i-1} = 0$, then we have the following:*

*(b.1) If $Z_0 = 2n$ for some $n \in \mathbb{N}$, then $Z$ takes values in the even numbers, $Z$ gets absorbed in the state $\{0\}$, and $\sup_{n \in \mathbb{N}} \mathbf{E}_n[\tau_0] < \infty$.*

*(b.2) If $Z_0 = 2n + 1$ for some $n \in \mathbb{N}_0$, then $Z$ takes values in the odd numbers. Furthermore, we have the following:*

*(b.2.1) If $\rho = 0$, then $Z$ gets absorbed in the state $\{1\}$, and $\sup_{\in \mathbb{N}} \mathbf{E}_n[\tau_1] < \infty$.*

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

*(b.2.2) If $\rho > 0$, then $Z$ has a unique stationary distribution $\nu$ satisfying*

$$\lim_{t\to\infty} \sup_{n\in\mathbb{N}} ||\mathbf{P}_n(Z_t = \cdot) - \nu(\cdot)|| = 0.$$

*(ii) If $d > 0$, then $Z$ gets absorbed in the state $\{0\}$, and $\sup_{n\in\mathbb{N}} \mathbf{E}_n[\tau_0] < \infty$.*

Figure 1 depicts the different long-term behaviours depending on the choice of parameters.

## 2. Proofs

### 2.1. Proof of Theorem 1

*Proof.* For the first part of the statement, we apply Theorem 1.11 in Chen [8] (see also Theorem 2.2 in Tweedie [37] for the case $\mathbf{m} \le 0$ and Theorem 2.1 in Meyn and Tweedie [28] for the case $\mathbf{m} > 0$). The proof of the aforementioned result follows from a localization argument, and its conditions can be written in terms of the localized transition rates without changing the conclusion. More precisely, we consider the particular case where the state space is $\mathbb{N}_0$ and introduce the sequence of sets $E_n = \{0, 1, 2, \dots, n\}$ for $n \in \mathbb{N}$, which increases towards $\mathbb{N}_0$. We also introduce the localized transition rates $(q_{i,j}^{(n)})_{i,j\in\mathbb{N}_0}$ as follows:

$$q_{i,j}^{(n)} = \mathbf{1}_{E_n}(i) q_{i,j} \qquad \text{with} \qquad -q_{i,i}^{(n)} = \sum_{j\neq i} q_{i,j}^{(n)} \quad \text{for } i, j \in \mathbb{N}_0.$$

If a nonnegative function $\varphi = (\varphi(i), i \in \mathbb{N}_0)$ satisfies

$$\lim_{n\to\infty} \inf_{i\notin E_n} \varphi(i) = \infty, \qquad \text{and} \qquad \sum_{j\in\mathbb{N}_0} q_{i,j}^{(n)} (\varphi(j) - \varphi(i)) \le c\varphi(i) \qquad \text{for } i \in E_n,$$

where $c \in \mathbb{R}$, then the associated process to $(q_{i,j})_{i,j\in\mathbb{N}_0}$ does not explode, since the arguments from Parts (ii), (iii) and (iv) of the proof of Theorem 1.11 in Chen [8] still hold with the above assumptions. In our particular case, we take

$$\varphi(i) = \begin{cases} i & \text{if } i \in E_n, \\ n & \text{otherwise.} \end{cases}$$

Since $\inf_{i\notin E_n} \varphi(i) = n$ and $\varsigma \le 0$, we deduce

$$\sum_{j\in\mathbb{N}_0} q_{i,j}^{(n)} \Big(\varphi(j) - \varphi(i)\Big) \le i\mathbf{m} + i(i-1)\varsigma - \sum_{k=2}^{i} (k-1) \binom{i}{k} \lambda_{i,k} \le \varphi(i)\mathbf{m} \qquad \text{for } i \in E_n,$$

which implies from above that the process $Z$ does not explode (or it is unique or regular, in the terminology of [8]).

For the second statement, we assume that the measure $\Lambda$ is identically zero and that $\varsigma > 0$. We first treat the case when $d = \rho = 0$. In order to do so, we introduce $A = (A_t, t \in \mathbb{R}_{+,0})$, a compound Poisson process with parameter $a + b + c$ and jump distribution $\eta$ satisfying

$$\eta(i) = \begin{cases} \dfrac{a}{a+b+c} \text{ if } i = -2, \\ \dfrac{c}{a+b+c} \text{ if } i = -1, \\ \dfrac{b_i}{a+b+c} \text{ if } i \in \mathbb{N}. \end{cases}$$

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

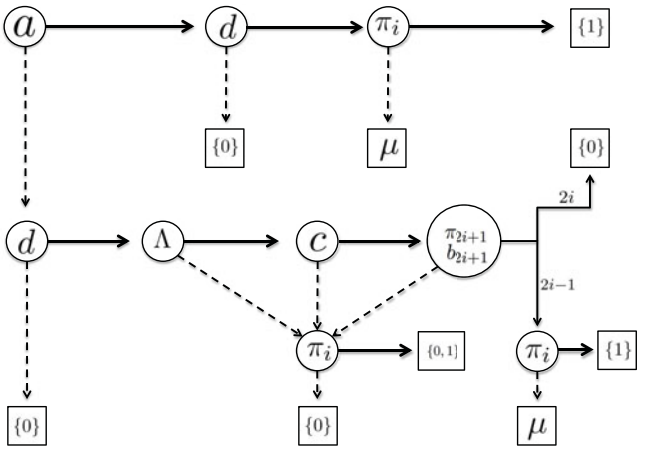



FIGURE 1: This figure exemplifies the long-term behaviour of branching processes with interactions in the subcritical cooperative regime, depending on its parameters. Dotted lines mean that the parameter is different from zero, while continuous lines mean that the parameter is zero. For example, if $a = 0$, $d = 0$, and $\pi_i = 0$ for all $i \in \mathbb{N}$, we take the right edge three times. In this case the process gets absorbed at $\{1\}$. The symbol $\mu$ means that the process has a nondegenerate stationary distribution.

We also consider the function $g(i) = i(i-1)$, which is nonnegative for $i \in \mathbb{N}_0$, with $g(i) = 0$ for $i = 0, 1$. Hence, from Theorems 6.1.1 and 6.1.3 in Ethier and Kurtz [11], the random time-changed process

$$Z_t = A_{\int_0^t g(Z_s)\mathrm{d}s}, \qquad t \le \theta_{0,1} := \inf\{s : A_s = 0 \text{ or } 1\},$$

defines a Markov process with absorbing states $\{0, 1\}$ whose infinitesimal generator is given by

$$Qf(n) = n(n-1)\mathcal{B}f(n),$$

where, for $f \in C_0(\mathbb{Z} \cup \{\infty\}, \mathbb{R})$ with $f(i) = 0$ for $i \in \{\ldots, -2, -1, 0, 1\}$, the infinitesimal generator $\mathcal{B}$ acts as follows:

$$\mathcal{B}f(n) = a[f(n-2) - f(n)] + c[f(n-1) - f(n)] + \sum_{i\in\mathbb{N}} b_i[f(n+i) - f(n)], \quad n \in \mathbb{N} \setminus \{1\}.$$

In other words, the process $Z$ has the same law as a branching process with interactions with parameters $a$, $c$, and $b_i$ for $i \in \mathbb{N}$.

In the same probability space, we also consider a Poisson process $\widetilde{A}$ with parameter $\delta \in (0, \varsigma)$; then, by using a similar random time change as above, we introduce a branching process with interactions with parameters $c = a = 0$, $b_1 = \delta > 0$, and $b_i = 0$ for $i \in \mathbb{N} \setminus \{1\}$. We observe that the process $\widetilde{Z}$, which is the process with these parameters in this probability space, explodes a.s. Indeed, let us define $\widetilde{\tau}_i^+ = \inf\{t > 0 : \widetilde{Z}_t \ge i\}$ and observe that its lifetime, here denoted by $\widetilde{\tau}_\infty$, satisfies $\widetilde{\tau}_\infty = \sup_{i\in\mathbb{N}} \widetilde{\tau}_i^+$. Thus, we see that

$$\widetilde{\mathbf{E}}_n[\widetilde{\tau}_\infty] \le \widetilde{\mathbf{E}}_2[\widetilde{\tau}_\infty] = \sum_{i\in\mathbb{N}} \widetilde{\mathbf{E}}_i[\widetilde{\tau}_{i+1}^+ - \widetilde{\tau}_i^+] = \frac{1}{\delta} \sum_{i\in\mathbb{N}\setminus\{1\}} \frac{1}{i(i-1)} < \infty, \tag{2.5}$$

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

where $\widetilde{\mathbf{E}}_n$ denotes the expectation of $\widetilde{Z}$ starting from $n \in \mathbb{N} \setminus \{1\}$. The previous identity implies that $\widetilde{\tau}_\infty$ is finite a.s., or equivalently that $\widetilde{Z}$ explodes a.s.

Next, we assume that $A_0 = n > m = \widetilde{A}_0$ and define the event $E = \{A_t > \widetilde{A}_t, \text{ for all } t > 0\}$. It is straightforward to deduce that the process $S_t = A_t - \widetilde{A}_t$, for $t \geq 0$, is a compound Poisson process starting from $n - m > 0$ and drifting to $\infty$, since $\varsigma > \delta$. Hence with positive probability the process $S$ stays strictly positive, implying that $\mathbb{P}(E|S_0 = n - m) > 0$. In other words, conditionally on $E$ and given that $S_0 = n - m > 0$, we have that the process $\widetilde{Z}$ is stochastically dominated by $Z$, and since the former explodes, $Z$ also explodes. This proves our statement for the case $d = \rho = 0$.

For the case $d + \rho > 0$, our analysis will be based on the following coupling argument. Let us consider two branching processes with interactions, $Z = (Z_t, t \in \mathbb{R}_{+,0})$ and $\overline{Z} = (\overline{Z}_t, t \in \mathbb{R}_{+,0})$, with respective parameters $d$, $a$, $c$, $b_i$, and $\pi_i$, for $i \in \mathbb{N}$, such that $d + \rho > 0$; and $\overline{d} = \sum_{i\in\mathbb{N}} \overline{\pi_i} = 0$, $\overline{c} = c + \epsilon d$, $\overline{a} = a$, and $\overline{b_i} = b_i$ for all $i \in \mathbb{N}$, where $\epsilon$ is chosen positive and such that

$$-2\overline{a} - \overline{c} + \sum_{i\in\mathbb{N}} i\overline{b_i} = -2a - c - \epsilon d + \sum_{i\in\mathbb{N}} i b_i > 0. \tag{2.6}$$

We denote by $\overline{\mathbf{P}}_n$ the law of $\overline{Z}$ starting from $\overline{Z}_0 = n$. We also observe, from above, that $\overline{Z}$ explodes with positive probability. Moreover, let $n_0 = \inf\{n \in \mathbb{N} : 1 < (n-1)\epsilon\}$ and note that for any starting condition $Z_0 = \overline{Z}_0 = n > n_0$, we necessarily have

$$nd + n(n-1)c < n(n-1)\overline{c} = \epsilon n(n-1)d + n(n-1)c.$$

Next, we consider the following coupling between the Markov chains $Z$ and $\overline{Z}$. Let us introduce the process $\{(U_t, \overline{U}_t), t \in \mathbb{R}_{+,0}\}$ with the following dynamics. Assume that $U_0 = \overline{U}_0 > n_0$. For all $t < \gamma_{n_0} := \inf\{t > 0 : \overline{U}_t \leq n_0\}$, the couple $(U_t, \overline{U}_t)$ has the following transitions rates, for $m \geq n$:

$$(U_t, \overline{U}_t) \text{ goes from } (n, m) \text{ to } \begin{cases} (n+i, m+i), & \text{with rate } m(m-1)b_i, \\ (n+i, m), & \text{with rate } n\pi_i + [n(n-1) - m(m-1)]b_i, \\ (n-1, m-1), & \text{with rate } \underline{\Theta}(n, m), \\ (n, m-1), & \text{with rate } m(m-1)(\epsilon d + c) - \underline{\Theta}(n, m), \\ (n-1, m), & \text{with rate } n(n-1)c + nd - \underline{\Theta}(n, m), \\ (n-2, m-2), & \text{with rate } n(n-1)a, \end{cases} \tag{2.7}$$

where $\underline{\Theta}(n, m) = \min(n(n-1)c + nd, m(m-1)(\epsilon d + c))$. For $t \geq \gamma_{n_0}$, the Markov chains $U_t$ and $\overline{U}_t$ evolve independently with the same transition rates as $Z$ and $\overline{Z}$, respectively. We denote by $\mathbf{P}_{(n,m)}$ the law of $\{(U_t, \overline{U}_t), t \in \mathbb{R}_{+,0}\}$ starting from $(n,m)$. Note from our construction that the Markov chains $U$ and $\overline{U}$ are equal in distribution to $Z$ and $\overline{Z}$, respectively. Furthermore, we necessarily have

$$\mathbf{P}_{(n,n)}(U_t \geq \overline{U}_t, \text{ for } t < \gamma_{n_0}) = 1.$$

We deduce that for $n > n_0$, $\mathbf{P}_{(n,n)}(U_t = \infty) > \mathbf{P}_{(n,n)}(\overline{U}_t = \infty, \gamma_{n_0} > t)$, where the right-hand side of the inequality is positive for $t$ sufficiently large. The latter implies that $Z$ is not conservative. □

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

For the proof of Proposition 1, we verify the Foster–Lyapunov conditions for positive recurrence. In order to do so, we consider the jump Markov chain associated to $Z$, which we denote by $Y = (Y_n, n \in \mathbb{N}_0)$, and whose jump matrix is given by $P = (p_{i,j})_{i,j\in\mathbb{N}_0}$, where

$$p_{i,j} = \begin{cases} -q_{i,j}/q_{i,i} & \text{if } j \neq i \\ 0 & \text{if } j = i, \end{cases}$$

for $i \in \mathbb{N}$ and $p_{0,0} = 1$. On the other hand, we recall that if a state is recurrent or transient for the jump chain $Y$, then the state is also recurrent or transient for the chain $Z$. (See for instance the monograph of Norris [32] for further details on Markov chains and their associated jump chains.)

*Proof of Proposition 1.* We first prove Part (ii). Recall that $d = 0$ and $\rho > 0$ and that under our assumption $c > 0$. From the form of the jump matrix $P$ defined above, we observe that the jump chain $Y$ is irreducible in $\mathbb{N}$ and that the state $\{0\}$ is not accessible from any state $i \in \mathbb{N}$.

In order to analyse the class properties of the state space of the jump chain $Y$, and implicitly for $Z$, we apply the Foster–Lyapunov criteria (see for instance Proposition 1.3 in Hairer [17] or Theorem 4.2 in Meyn and Tweedie [28]). Let us consider the discrete generator associated to $Y$, which is defined as follows:

$$\mathcal{L}f(i) = Pf(i) - f(i) = -\frac{1}{q_{i,i}} \sum_{j\in\mathbb{N}_0} q_{i,j} f(j),$$

for $f : \mathbb{N}_0 \to \mathbb{R}$. We choose the Lyapunov function to be $f(i) = Ci$ for $i \in \mathbb{N}_0$ and $C$ a positive constant that we will specify later. According to the Foster–Lyapunov criteria we know that if $\mathcal{L}f(i) \leq -1$ for all but finitely many values of $i$, then $Y$ is positive recurrent. Hence straightforward computations lead to

$$\begin{aligned} \mathcal{L}f(i) &= -\frac{C}{q_{i,i}} \left( i\mathbf{m} + i(i-1)\varsigma - \sum_{k=2}^{i} (k-1) \binom{i}{k} \lambda_{i,k} \right) \\ &\leq C \left( \frac{m}{(i-1)(b+c+a)} + \frac{\varsigma}{(b+c+a)} - \frac{\sum_{k=2}^{i} (k-1)\binom{i}{k}\lambda_{i,k}}{i(i-1)(b+c+a)} \right). \end{aligned} \tag{2.8}$$

For $i$ sufficiently large, we can clearly choose $C$ in such a way that $\mathcal{L}f(i) \leq -1$, implying that the Foster–Lyapunov criteria hold. In other words, $Z$ is positive recurrent in $\mathbb{N}$.

Now we prove Part (i). Recall that $\rho = 0$ and $d = 0$ and that under our assumption $c > 0$. Again from the form of the jump matrix $P$, we observe that the state $\{1\}$ is absorbing and that $\{0\}$ is not accessible. Moreover, if $b > 0$ the class $\{2, 3, \ldots\}$ is irreducible for $Z$, and if $b = 0$, the process $Z$ is decreasing. In any case, we can always stochastically dominate $Z$ with another branching process with interactions, here denoted by $\widetilde{Z}$, with the same parameters except $\pi_1 > 0$. We refer the reader to Section 2 of López *et al.* [26] for the conditions on the intensities $(q_{i,j})_{i,j\geq 0}$ that guarantee the existence of an order-preserving coupling. Since $\widetilde{Z}$ is recurrent and reaches the state $\{1\}$ in finite time a.s., it follows that $Z$ is absorbed at 1 in finite time a.s.

Finally, we deduce Part (iii). Since in this case $d > 0$, from the form of the jump matrix $P$ we observe that the state $\{0\}$ is accessible. If $\rho = 0$ and $b = 0$, the process $Z$ is decreasing, and if $\rho > 0$ or $b > 0$, then $\mathbb{N}$ is irreducible for the $\bar{Z}$ constucted in Part (ii), and again an order-preserving coupling argument guarantees that $Z$ is absorbed at $\{0\}$ in finite time a.s. □

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

### 2.2. Moment duality property and proof of Theorem 2

The aim of this section is to prove Theorem 2, which characterizes a class of jump-diffusions that fulfils the moment duality property with respect to (sub)critical branching processes with interactions and without annihilation, i.e. $a = 0$. This moment duality will be very helpful for our purposes. The proof of Theorem 2 follows Lemmas 1 and 2 below.

In order to prove Theorem 2 we first introduce the class of jump-diffusions via the following SDE:

$$\mathrm{d}X_t = \mu(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t + \int_{(0,1]}\int_{(0,1]} zg(X_{t-}, u)\widetilde{N}(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u), \tag{2.9}$$

where $\widetilde{N}$ is a compensated Poisson random measure on $(0, \infty) \times (0, 1]^2$ with intensity $\mathrm{d}tz^{-2}\Lambda(\mathrm{d}z)\mathrm{d}u$, the function $g(x,u)$ is defined by

$$g(x, u) = \mathbf{1}_{\{u \le 1 \wedge x\}} - (1 \wedge x)\mathbf{1}_{\{x \ge 0\}},$$

and the functions $\mu : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}$ are continuous and satisfy

$$\mu(x) = d\mathbf{1}_{\{x<0\}} + \left(d(1-x) + \sum_{i\in\mathbb{N}} \pi_i(x^{i+1} - x)\right)\mathbf{1}_{\{x\in[0,1]\}}$$

and

$$\sigma^2(x) = \left(c(x - x^2) + \sum_{i\in\mathbb{N}} b_i(x^{i+2} - x^2)\right)\mathbf{1}_{\{x\in[0,1]\}},$$

for some positive constants $d$, $c$, $b_i$, and $\pi_i$ with $i \in \mathbb{N}$.

**Lemma 1.** *Fix $d$, $c$, $b_i$, $\pi_i$ for $i \in \mathbb{N}$, where all terms are nonnegative. If such coefficients satisfy $|\mathbf{m}| < \infty$ and $\varsigma \le 0$, then the SDE (2.9) with starting point $X_0 \in [0, 1]$ has a unique strong solution taking values on* $[0, 1]$.

*Proof.* We first observe that if $X$ satisfies (2.9) with $X_0 \in [0, 1]$, a.s., then $X_t \in [0, 1]$ for all $t \in \mathbb{R}_+$, a.s. Suppose that there exists $\epsilon > 0$ such that $\tau := \inf\{t \ge 0 : X_t \le -\epsilon\}$ is finite with positive probability. Then on the event $\{\tau < \infty\}$, we have $X_\tau = -\epsilon$ and

$$\tau > \tau_0 = \inf\{s < \tau : X_t \le 0, \forall s \le t \le \tau\}.$$

So we take $r \ge 0$ such that $\{\tau_0 \le r < \tau\}$ occurs with strictly positive probability, and on that event, we observe

$$X_{t\wedge\tau} = X_{r\wedge\tau} + d(t \wedge \tau - r \wedge \tau), \qquad t \ge r,$$

implying that $t \mapsto X_{t\wedge\tau}$ is nondecreasing on $[r, \infty)$. Since $X_r > -\epsilon$ on $\{\tau_0 \le r < \tau\}$, we get a contradiction. The previous argument implies that $X_t \ge 0$ a.s. A similar argument proves that $X_t \le 1$ a.s. We leave the details to the reader.

Now, we show that $X$ has a unique strong solution on $[0, 1]$. In order to do so, we first observe that for $x, y \in [0, 1]$,

$$|\mu(x) - \mu(y)| \le \left(d + \sum_{i\in\mathbb{N}} (i+1)\pi_i\right)|x - y|.$$

The latter inequality implies that $\mu$ is Lipschitz continuous on $[0, 1]$.

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

On the other hand, we also observe that $\sigma$ satisfies, for all $x, y \in [0, 1]$,

$$|\sigma^2(x) - \sigma^2(y)| \le \left| c(x-y) + c(y^2 - x^2) + \sum_{i\in\mathbb{N}} b_i \left( x^{2+i} - y^{i+2} - (x^2 - y^2) \right) \right|$$
$$\le \left( 3c + \sum_{i\in\mathbb{N}} b_i (i+2) + 2 \sum_{i\in\mathbb{N}} b_i \right) |x-y|,$$

implying from our assumptions that $\sigma$ is Hölder continuous on $[0, 1]$ with exponent 1/2.

Finally, we observe from Corollary 6.2 in Li and Pu [24] that $x \mapsto x + g(x, u)$ is nondecreasing, and for any $0 \le x, y \le 1$,

$$\int_{(0,1]} \frac{\Lambda(\mathrm{d}z)}{z^2} \int_0^1 \mathrm{d}r\, z^2 |g(x, r) - g(y, r)| \le |x - y| \int_{(0,1]} \Lambda(\mathrm{d}z).$$

Hence the conditions (3.a), (3.b), and (5.a) in Li and Pu [24] are satisfied, implying that there is a unique strong solution to the SDE (2.9). □

Let $X = (X_t, t \in \mathbb{R}_{+,0})$ be the unique strong solution of the SDE (2.9). We denote by $\mathcal{A}$ the infinitesimal generator of the process X, which satisfies

$$\mathcal{A}f(x) := \lim_{t\downarrow 0} \frac{\mathbb{E}_x[f(X_t)] - f(x)}{t}$$

for any $f : [0, 1] \mapsto \mathbb{R}$ continuous such that the limit exists uniformly on $x \in [0, 1]$. If this limit exists, then we say that $f \in \mathcal{D}_{\mathcal{A}}$, the domain of the infinitesimal generator $\mathcal{A}$. Observe from Itô's formula that for $f \in C_b^2([0, 1], \mathbb{R})$ (the set of twice continuously differentiable bounded functions) and $x \in [0, 1]$, the infinitesimal generator $\mathcal{A}$ satisfies

$$\mathcal{A}f(x) = \mu(x) f'(x) + \frac{\sigma^2(x)}{2} f''(x)$$
$$+ \int_{[0,1]} (xf(x(1-z)+z) + (1-x)f(x(1-z)) - f(x)) \frac{\Lambda(\mathrm{d}z)}{z^2},$$

where $\mu, \sigma : [0, 1] \to \mathbb{R}$ and $\Lambda$ are defined as above. In other words, for $f \in C_b^2([0, 1], \mathbb{R})$, the process

$$f(X_t) - \int_0^t \mathcal{A}f(X_s)\mathrm{d}s$$

is a martingale. On the other hand, we also observe that $C_b^2([0, 1], \mathbb{R})$ is dense in $C([0, 1], \mathbb{R})$, the set of continuous functions from $[0, 1]$ to $\mathbb{R}$. It is not so difficult to see that the generator $\mathcal{A}$ satisfies the positive maximum principle. Since $[0, 1]$ is a compact space and the space of all polynomials on $[0, 1]$ is dense on $C([0, 1], \mathbb{R})$, the Hille–Yosida theorem implies that the closure of $\mathcal{A}$ restricted to $C_b^2([0, 1], \mathbb{R})$ generates a strongly continuous, time-homogeneous, positive contraction Feller semigroup on $C([0, 1], \mathbb{R})$; see for instance Theorem 4.2.2 and Lemma 4.2.3 in [11]. Theorem 4.2.7 and Corollary 4.2.8 in [11] imply that the process $X$ is Feller.

Recall that $\mathbb{P}_x$ and $\mathbf{P}_n$ denote the laws of $X$ starting from $x \in [0, 1]$ and $Z$ starting from $n \ge 0$, respectively. Moreover, from our assumptions we have that $\mathbf{P}_n(Z_t < \infty) = 1$ for any $n \in \mathbb{N}_0$ and $t \in \mathbb{R}_{+,0}$.

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

In practice, an effective methodology to show duality between two Markov processes is via their infinitesimal generators. Indeed, the following result is a special case of Theorem 4.11 and Corollary 4.15 in Ethier and Kurtz [11] and can also be seen as a small modification of Proposition 1.2 of Jansen and Kurt [19].

**Proposition 3.** *Let $(Y_t^{(1)}, t \in \mathbb{R}_{+,0})$ and $(Y_t^{(2)}, t \in \mathbb{R}_{+,0})$ be two Markov processes taking values on $E_1$ and $E_2$, respectively. Let $H : E_1 \times E_2 \to \mathbb{R}$ be a bounded and continuous function, and assume that there exist functions $g_i : E_1 \times E_2 \mapsto \mathbb{R}$, for $i = 1, 2$, such that for every $n \in E_1$, $x \in E_2$ and every $T > 0$, the processes $(M_t^{(1)}, 0 \le t \le T)$ and $(M_t^{(2)}, 0 \le t \le T)$, defined as*

$$M_t^{(1)} := H\big(Y_t^{(1)}, x\big) - \int_0^t g_1\big(Y_s^{(1)}, x\big)\mathrm{d}s, \tag{2.10}$$

$$M_t^{(2)} := H\big(n, Y_t^{(2)}\big) - \int_0^t g_2\big(n, Y_s^{(2)}\big)\mathrm{d}s, \tag{2.11}$$

*are martingales with respect to the natural filtrations of $Y_t^{(1)}$ and $Y_t^{(2)}$ respectively. Then, if*

$$g_1(n, x) = g_2(n, x) \qquad \text{for all} \qquad n \in E_1, x \in E_2,$$

*the processes $\big(Y_t^{(1)}, t \in \mathbb{R}_{+,0}\big)$ and $\big(Y^{(2)}, t \in \mathbb{R}_{+,0}\big)$ are dual with respect to $H$.*

This proposition is a direct consequence of Theorem 4.11 in Ethier and Kurtz [11], with H taken to be bounded and continuous and $\alpha = \beta = 0$.

If the function $H$ is such that $H(n, x) = x^n$ for $x \in [0, 1]$ and $n \in \mathbb{N}_0$, the above duality relationship is known as *moment duality*. The moment duality between $X$ and $Z$ reads as follows.

**Lemma 2.** *Assume $|\mathbf{m}| < \infty$, $\varsigma \le 0$, and $a = 0$. Let $X$ be the unique strong solution of the SDE (2.9); then it is the unique (in distribution) moment dual of $Z$ a (sub)critical branching process with interactions and no annihilation. More precisely, for $x \in [0, 1]$ and $n \in \mathbb{N}_0$, we have*

$$\mathbb{E}_x[X_t^n] = \mathbf{E}_n\big[x^{Z_t}\big] \qquad \text{for} \quad t \in \mathbb{R}_{+,0}.$$

*Proof.* We take $H(n, x) = x^n$ for $x \in [0, 1]$ and $n \in \mathbb{N}$, and when $n = 0$, $H(0, x) := 1$ for $x \in [0, 1]$; these are continuous and bounded.

Now, we claim that for fixed $x \in [0, 1]$, $H(n, x)$ is in the domain of $Q$, and for fixed $n \in \mathbb{N}_0$, it is in the domain of $\mathcal{A}$. The latter follows from the fact that $H(n, \cdot)$ is a polynomial on $[0, 1]$, which clearly belongs to $C_b^2([0, 1], \mathbb{R})$. Implicitly, we have that

$$H(n, X_t) - \int_0^t \mathcal{A}H(n, X_s)\mathrm{d}s$$

is a martingale. The first claim follows from recalling that linear combinations of exponential functions as in (1.2) belong to the domain of $Q$ (see Section 1), and the fact that $H(n, x) = e^{n \log x}$ implies that the process

$$H(Z_t, x) - \int_0^t QH(Z_s, x)\mathrm{d}s$$

is also a martingale.

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

Next, we observe that for $x \in (0, 1)$ and $n \in \mathbb{N} \setminus \{1\}$,

$$\begin{aligned}\mathcal{A}H(n, x) &= nd(x^{n-1} - x^n) + n \sum_{i \in \mathbb{N}} \pi_i (x^{n+i} - x^n) \\ &\quad + n(n-1) \sum_{i \in \mathbb{N}} b_i (x^{n+i} - x^n) + n(n-1)c(x^{n-1} - x^n) \\ &\quad + \int_{[0,1]} (x(x(1-z)+z)^n + (1-x)(x(1-z))^n - x^n) \frac{\Lambda(\mathrm{d}z)}{z^2}.\end{aligned}$$

On the other hand, let $B$ be a Bernoulli random variable taking the value 1 with probability $x$ and 0 with probability $1 - x$. Then, for $n \in \mathbb{N} \setminus \{1\}$, we observe that

$$\begin{aligned}x(x(1-z)+z)^n + (1-x)(x(1-z))^n - x^n &= \mathbb{E}\big[(x(1-z) + Bz)^n\big] - x^n \\ &= \mathbb{E}\left[\sum_{k=0}^{n} \binom{n}{k} x^{n-k} (1-z)^{n-k} z^k B^k\right] - x^n \\ &= \sum_{k=2}^{n} \binom{n}{k} x^{n-k+1} (1-z)^{n-k} z^k \\ &\quad + x^n (1-z)^n + n x^n (1-z)^{n-1} z - x^n \\ &= \sum_{k=2}^{n} \binom{n}{k} x^{n-k+1} (1-z)^{n-k} z^k \\ &\quad - x^n \big(1 - (1-z)^n - n(1-z)^{n-1} z\big) \\ &= \sum_{k=2}^{n} \binom{n}{k} (1-z)^{n-k} z^k \big(x^{n-k+1} - x^n\big).\end{aligned}$$

Hence, for $n \in \mathbb{N} \setminus \{1\}$, we deduce that

$$\int_{[0,1]} \big(x(x(1-z)+z)^n + (1-x)(x(1-z))^n - x^n\big) \frac{\Lambda(\mathrm{d}z)}{z^2} = \sum_{k=2}^{n} \binom{n}{k} \lambda_{n,k} [x^{n-k+1} - x^n],$$

implying that $\mathcal{A}H(n, x) = QH(n, x)$ for $x \in (0, 1)$ and $n \in \mathbb{N} \setminus \{1\}$. For the case $n = 1$ and $x \in (0, 1)$, we have

$$\mathcal{A}H(1, x) = d(1-x) + \sum_{i \in \mathbb{N}} \pi_i (x^{1+i} - x) = QH(1, x).$$

Finally, we consider the remaining cases. For $n \in \mathbb{N}$ and $x = 0$ or $x = 1$, it is clear that $\mathcal{A}H(n, x) = 0 = QH(n, x)$. Moreover, when $n = 0$, we have $QH(0, x) = 0 = \mathcal{A}H(0, x)$ for $x \in [0, 1]$.

Putting all the pieces together, we deduce that $\mathcal{A}H(n, x) = QH(n, x)$ for $x \in [0, 1]$ and $n \in \mathbb{N}_0$. Hence we can apply Proposition 3 to deduce the moment duality between $X$ and $Z$. □

We finish this section with the following useful lemma, which provides an upper bound for the expected fixation time of the process $X$ defined by (2.9) with $\mu \equiv 0$ and $\Lambda \equiv 0$. We will use

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

this lemma for the proofs of Theorems 3 and 4. In particular, we have that the fixation time for $X$ with $\mu \equiv 0$, $\Lambda \equiv 0$ and starting at $x \in [0, 1]$ is finite $\mathbb{P}_x$-a.s.

**Lemma 3.** *Let $X$ be the process defined by the SDE (2.9) with $\varsigma < 0$, $\Lambda \equiv 0$, $d = 0$, $\pi_i = 0$ for all $i \in \mathbb{N}$, and recall that $T_{0,1} = \inf\{t > 0 : X_t \in \{0, 1\}\}$. Then for any $x \in [0, 1]$,*

$$\mathbb{E}_x[T_{0,1}] \leq -\frac{2}{c - \sum_{i=1}^{\infty} ib_i}\big((1-x)\ln(1-x) + x\ln x\big).$$

*Moreover, we have $\mathbb{P}_x(X_{T_{0,1}} = 1) = x$.*

*Proof.* Since $\Lambda \equiv 0$, $d = 0$, and $\pi_i = 0$ for all $i \in \mathbb{N}$, we observe that the process $X$, under $\mathbb{P}_x$, satisfies the following SDE:

$$X_t = x + \int_0^t \sigma(X_s)\mathrm{d}B_s, \tag{2.12}$$

where $\sigma$ is defined as above. The process $X$ is a diffusion taking values on $[0, 1]$ with $\{0, 1\}$ as absorbing states. Observe that its associated scale function is given by $S(y) = y + \mathbf{c}$, where $\mathbf{c}$ is a constant, and its associated speed measure satisfies

$$m(\mathrm{d}y) = \frac{2}{\sigma^2(y)}\mathrm{d}y.$$

Hence, from Corollary VII 3.8 in [34], we have

$$\mathbb{E}_x[T_{0,1}] = 2\int_0^1 G(x, y)\frac{1}{\sigma^2(y)}\mathrm{d}y,$$

where

$$G(x, y) = \begin{cases} x(1-y) & \text{if } 0 \leq x \leq y \leq 1, \\ y(1-x) & \text{if } 0 \leq y \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Now, we observe that for $y \in [0, 1]$, the following inequality holds:

$$\sigma^2(y) \geq \Big(c - \sum_{i\in\mathbb{N}} ib_i\Big)y(1-y).$$

This implies

$$\begin{aligned}\mathbb{E}_x[T_{0,1}] &\leq \frac{2}{\Big(c - \sum_{i\in\mathbb{N}}^{\infty} ib_i\Big)}\left(\int_0^x \frac{y(1-x)}{y(1-y)}\mathrm{d}y + \int_x^1 \frac{x(1-y)}{y(1-y)}\mathrm{d}y\right) \\ &= -\frac{2}{\Big(c - \sum_{i\in\mathbb{N}} ib_i\Big)}\Big((1-x)\ln(1-x) + x\ln x\Big),\end{aligned}$$

which provides the desired inequality.

Finally, since $X$ is a local martingale taking values on $[0, 1]$, it is a uniformly integrable martingale. Hence from the optional stopping theorem, we deduce that

$$x = \mathbb{E}_1[X_0] = \mathbb{E}_x[X_{T_{0,1}}] = \mathbb{P}_x(X_{T_{0,1}} = 1).$$

This completes the proof. □

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

### 2.3. Proof of Theorem 3 & Corollary 2

For the proof of Theorem 3, we need the following lemma.

**Lemma 4.** *Assume that there is no annihilation, no death, and no branching; i.e.* $\varsigma < 0$, $a = 0$, $d = 0$, *and* $\rho = 0$. *Then*

$$\sup_{n\in\mathbb{N}} \mathbf{E}_n\big[\tau_1\big] < \infty,$$

*where* $\tau_1 = \inf\{t > 0 : Z_t = 1\}$.

*Proof.* To prove the statement, it is enough to prove it in the case when there are no catastrophes, since an order-preserving coupling can be established. Indeed, we can stochastically dominate the case with catastrophes with another branching process with interactions having the same parameters but no catastrophes. We refer the reader to Section 2 of López *et al.* [26] for the conditions on the intensities $(q_{i,j})_{i,j\in\mathbb{N}_0}$ that ensure the existence of an order-preserving coupling.

From the moment duality, i.e. Lemma 2, we get

$$\mathbf{E}_n\big[x^{Z_t}\big] = \mathbb{E}_x[X_t^n] = \mathbb{E}_x\Big[X_t^n \mathbf{1}_{\{T_{0,1}\geq t\}}\Big] + \mathbb{E}_x\Big[X_{T_{0,1}}^n \mathbf{1}_{\{T_{0,1}<t\}}\Big].$$

Thus, by letting $t$ go to $\infty$, applying the dominated convergence theorem, and recalling that $T_{0,1}$ is finite $\mathbb{P}_x$-a.s. (see Lemma 2), we obtain

$$\lim_{t\to\infty} \mathbf{E}_n\big[x^{Z_t}\big] = \mathbb{P}_x(X_{T_{0,1}} = 1) = x. \tag{2.13}$$

On the other hand, again from the moment duality between $X$ and $Z$, we deduce the following inequality:

$$\mathbf{P}_n(Z_t = 1) = \frac{\mathbf{E}_n\big[x^{Z_t}\big] - \sum_{i=2}^{\infty} x^i \mathbf{P}_n(Z_t = i)}{x} \geq \frac{\mathbf{E}_n\big[x^{Z_t}\big]}{x} - x.$$

Therefore, from the identity (2.13), we get

$$\lim_{t\to\infty} \mathbf{P}_n(Z_t = 1) \geq 1 - x.$$

Implicitly, there are $t_1 > 0$ and $\epsilon > 0$ such that for every $n \in \mathbb{N}$,

$$\mathbf{P}_n(Z_{t_1} = 1) > \epsilon.$$

Next we observe, using the Markov property of $Z$ starting from $n \in \mathbb{N}$ at times $(jt_1)_{j\in\mathbb{N}}$, that $\sigma_1$ is stochastically dominated by a geometric random variable with parameter $\epsilon$, here denoted by $\Upsilon_\epsilon$, that bounds the number of steps taken by the discretized Markov chain $(Z_{jt_1}, j \geq 0)$ to reach the state 1. In other words, for any $n, m \in \mathbb{N}$, we have

$$\mathbf{P}_n(\tau_1 > mt_1) \leq \mathbb{P}(\Upsilon_\epsilon > mt_1).$$

The latter inequality implies, in particular, that

$$\sup_{n\in\mathbb{N}} \mathbf{E}_n\big[\tau_1\big] < \frac{t_1}{\epsilon},$$

which completes the proof. □

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

*Proof of Theorem 3.* We first deal with the case with no catastrophes, i.e. $\Lambda \equiv 0$. From our assumptions, $Z$ has parameters $d = 0$, $c$, $\pi_i$, and $b_i$, for $i \in \mathbb{N}$, such that $\rho = \sum_{i\in\mathbb{N}} \pi_i > 0$ and $\varsigma < 0$. Our arguments will be based on a coupling argument. More precisely, let us consider two branching processes with interactions, $Z = (Z_t, t \in \mathbb{R}_{+,0})$ and $\overline{Z} = (\overline{Z}_t, t \in \mathbb{R}_{+,0})$, with respective parameters $a = d = 0$, $c$, $b_i$, and $\pi_i$, for $i \in \mathbb{N}$, such that $\rho > 0$; and $\overline{d} = \overline{a} = \sum_{i\in\mathbb{N}} \overline{\pi_i} = 0$, $\overline{c} = c$, and $\overline{b_i} = b_i + \epsilon\pi_i$ for all $i \in \mathbb{N}$, where $\epsilon$ is chosen positive and such that

$$-\overline{c} + \sum_{i\in\mathbb{N}} i\overline{b_i} = -c + \sum_{i\in\mathbb{N}} ib_i + \epsilon \sum_{i\in\mathbb{N}} i\pi_i < 0. \tag{2.14}$$

We denote by $\overline{\mathbf{P}}_n$ the law of $\overline{Z}$ starting from $\overline{Z}_0 = n$ and observe, from the first part of the proof and under our assumptions, that $\overline{Z}$ does not explode and gets absorbed in the state $\{1\}$. Moreover, let $n_0 = \inf\{n \in \mathbb{N} : 2 < (n-1)\epsilon\}$ and observe that for any starting condition $Z_0 = \overline{Z}_0 = n > n_0$, we necessarily have

$$n\pi_i + n(n-1)b_i < n(n-1)\overline{b}_i = \epsilon n(n-1)\pi_i + n(n-1)b_i.$$

Next, we consider the following coupling between the Markov chains $Z$ and $\overline{Z}$. Let us introduce the process $\{(U_t, \overline{U}_t), t \in \mathbb{R}_{+,0}\}$ with the following dynamics. For all $t < \gamma_{n_0} := \inf\{t > 0 : U_t \le n_0\}$, the couple $(U_t, \overline{U}_t)$ has the following transition rates, for $m \ge n$:

$$(U_t, \overline{U}_t) \text{ goes from } (n, m) \text{ to } \begin{cases} (n+i, m+i), & \text{with rate } n\pi_i + n(n-1)b_i, \\ (n, m+i), & \text{with rate } m(m-1)\overline{b}_i - n\pi_i + n(n-1)b_i, \\ (n-1, m-1), & \text{with rate } n(n-1)c, \\ (n, m-1), & \text{with rate } m(m-1)c - n(n-1)c. \end{cases} \tag{2.15}$$

For $t \ge \gamma_{n_0}$, the Markov chains $U_t$ and $\overline{U}_t$ evolve independently with the same transition rates as $Z$ and $\overline{Z}$, respectively. We denote by $\mathbf{P}_{(n,m)}$ the law of $\{(U_t, \overline{U}_t), t \in \mathbb{R}_{+,0}\}$ starting from $(n, m)$. Note that from our construction, the Markov chains $U$ and $\overline{U}$ are equal in distribution to $Z$ and $\overline{Z}$, respectively. (We refer the reader to Chapters IV and V in Lindvall [25] for further details on this type of coupling and stochastic dominance.)

Furthermore, we necessarily have, for all $n > n_0$,

$$\mathbf{P}_{(n,n)}\big(U_t \le \overline{U}_t, \text{ for } t < \gamma_{n_0}\big) = 1,$$

where $\mathbf{P}_{(n,n)}$ denotes the law of $\{(U_t, \overline{U}_t), t \in \mathbb{R}_{+,0}\}$ starting from $(n, n)$. Hence, from Lemma 4 and the fact that $Z$ and $\overline{Z}$ are skip-down-free (or skip-free to the left), we deduce that

$$\mathbf{E}_n[\tau_{n_0}] = \mathbb{E}[\gamma_{n_0} | U_0 = n] \le \mathbb{E}[\overline{\gamma}_{n_0} | \overline{U}_0 = n] = \overline{\mathbf{E}}_n[\overline{\tau}_{n_0}] < \overline{\mathbf{E}}_n[\overline{\tau}_1] < K, \tag{2.16}$$

where $\gamma_{n_0} := \inf\{t > 0 : U_t \le n_0\}$, $\overline{\gamma}_{n_0} := \inf\{t > 0 : \overline{U}_t \le n_0\}$, $\tau_i := \inf\{t > 0 : Z_t = i\}$, $\overline{\tau}_i := \inf\{t > 0 : \overline{Z}_t = i\}$, and $K$ is a positive constant that does not depend on $n$.

Then the process $Z$ is positive recurrent, and we apply Theorem 21.14 of [23] to conclude that there exists a unique stationary distribution for $Z$, here denoted by $\nu$. To finish the proof of the first statement, we use Markov's inequality and (2.16) to deduce that

$$\mathbf{P}_n(\tau_{n_0} > a) < \frac{K}{a} \qquad \text{for every} \quad n \ge n_0.$$

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

This allows us to compute the total variation distance between $\mathbf{P}_n(Z_t = \cdot)$ and $\nu(\cdot)$ by conditioning on the event $\{\tau_{n_0} > t_0\}$, where $t_0 \in (0, t)$. In other words,

$$\begin{aligned}||\mathbf{P}_n(Z_t = \cdot) - \nu(\cdot)|| &\leq ||\mathbf{P}_n(Z_t = \cdot, \tau_{n_0} \leq t_0) - \nu(\cdot)|| \\ &\quad + \mathbf{P}_n(\tau_{n_0} > t_0)||\mathbf{P}_n(Z_t = \cdot|\tau_{n_0} > t_0) - \nu(\cdot)|| \\ &\leq \int_0^{t_0} ||\mathbf{P}_{n_0}(Z_{t-s} = \cdot) - \nu(\cdot)||\mathbf{P}_n(\tau_{n_0} \in \mathrm{d}s) + \frac{K}{t_0}.\end{aligned}$$

Therefore, again using Theorem 21.14 of [23], we observe that

$$||\mathbf{P}_{n_0}(Z_{t-s} = \cdot) - \nu(\cdot)|| \to 0 \qquad \text{as} \quad t \to \infty$$

for all $s \in [0, t_0]$, and thus we conclude that

$$\lim_{t\to\infty} \sup_{n\in\mathbb{N}} ||\mathbf{P}_n(Z_t = \cdot) - \nu(\cdot)|| < \frac{K}{t_0}.$$

Since $t_0$ was taken to be arbitrary, we deduce (1.4).

For the case with catastrophes, we use an order-preserving coupling. In other words, we introduce two branching processes $Z$ and $\overline{Z}$ with the same parameters $d = \overline{d} = 0$, $c = \overline{c}$, $\pi_i = \overline{\pi}_i$, $b_i = \overline{b}_i$, for $i \in \mathbb{N}$, but the former takes catastrophes into account and the latter has no catastrophes. Again, we refer the reader to Section 2 of López *et al.* [26] for the conditions on the intensities $(q_{i,j})_{i,j\in\mathbb{N}_0}$ that ensure the existence of an order-preserving coupling, which are clearly satisfied by $Z$ and $\overline{Z}$. We denote by $\overline{\mathbf{P}}_n$ the law of $\overline{Z}$ starting at $n$. Hence, the order-preserving property and the upper bound in (2.16) imply

$$\mathbf{E}_n[\tau_{n_0}] \leq \overline{\mathbf{E}}_n[\overline{\tau}_{n_0}] \leq K \qquad \text{for} \quad n \geq n_0,$$

where $\overline{\tau}_i := \inf\{t > 0 : \overline{Z}_t = i\}$, and $K$ is a positive constant that does not depend on $n$. Then we proceed similarly as in the case with no catastrophes to prove uniform convergence to stationarity in the total variation norm. We leave the details to the reader.

In order to prove last statement of this theorem, we first assume that $\mathbb{P}_x(X_\infty \in \{0, 1\}) < 1$, where $X_\infty := \lim_{t\to\infty} X_t$. This implies that $\lim_{t\to\infty} \mathbb{E}_x[X_t^n] < \lim_{t\to\infty} \mathbb{E}_x[X_t]$, since for every $x \in (0, 1)$, $x^n$ is strictly smaller than $x$. From the moment duality between $Z$ and $X$ and the fact that $\nu$ is the invariant distribution of $Z$, we obtain

$$\sum_{i\geq 0} x^i \nu(i) = \lim_{t\to\infty} \mathbf{E}_n\left[x^{Z_t}\right] = \lim_{t\to\infty} \mathbb{E}_x[X_t^n] < \lim_{t\to\infty} \mathbb{E}_x[X_t] = \lim_{t\to\infty} \mathbf{E}_1[x^{Z_t}] = \sum_{i\geq 0} x^i \mu(i),$$

which contradicts our hypothesis. In other words, we conclude that $\mathbb{P}_x(X_\infty \in \{0, 1\}) = 1$ for any starting point $x \in (0, 1)$. Finally, we use the left-hand side of the identity (2.13) to deduce that $\nu$ satisfies

$$\sum_{i\geq 0} x^i \nu(i) = \mathbb{P}_x(X_\infty = 1)$$

for any $x \in (0, 1)$. This completes the proof. □

Next we prove Corollary 2.

*Proof of Corollary 2.* From the dominated convergence theorem and the moment duality between $Z$ and $X$, we observe that for any $x \in (0, 1)$ and $n \geq 1$, we have

$$\mathbb{E}_x\left[\lim_{t\to\infty} X_t^n\right] = \lim_{t\to\infty} \mathbb{E}_x[X_t^n] = \lim_{t\to\infty} \mathbf{E}_n\left[x^{Z_t}\right] = \mathbf{E}_n\left[\lim_{t\to\infty} x^{Z_t}\right].$$

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

From Theorem 3, we observe that the limit of the generating function of $Z$ does not depend on the starting state $n$. This implies that all the moments of $X_\infty = \lim_{t\to\infty} X_t$ must be equal. Since the limit of $X$ must be in $[0, 1]$, from the latter observation we deduce that $\mathbb{P}_x(X_\infty \in \{0, 1\}) = 1$, and thus

$$\mathbb{P}_x\,(X_\infty = 1) = \mathbb{E}_x\left[\lim_{t\to\infty} X_t^n\right] = \mathbf{E}_n\left[\lim_{t\to\infty} x^{Z_t}\right].$$

The proof is completed once we compute $\mathbf{E}_n[\lim_{t\to\infty} x^{Z_t}]$ in each of the three cases of Theorem 3. □

### 2.4. Proof of Theorem 4

For the proof of Theorem 1, we need Lemma 4 and the following result.

**Lemma 5.** *Assume that* $\varsigma < 0$, $a = 0$, $d > 0$, *and* $\rho > 0$. *Then*

$$\sup_{n\in\mathbb{N}} \mathbf{E}_n\big[\tau_0\big] < \infty,$$

*where* $\tau_0 = \inf\{t > 0 : Z_t = 0\}$.

*Proof.* To prove this lemma, we again use a coupling method. From our assumptions $Z$ has parameters $d > 0$, $c$, $\pi_i$, $b_i$, for $i \in \mathbb{N}$, and $\lambda_{i,k}$, for $k \in \{2, \dots, i\}$. For our purposes we also consider a branching process with interactions $\overline{Z} = (\overline{Z}_t, t \in \mathbb{R}_{+,0})$ with parameters $\overline{d} = 0$, $\overline{c} = c$, $\overline{\pi_i} = \pi_i$, $\overline{b_i} = b_i$, for all $i \in \mathbb{N}$, and $\overline{\lambda}_{i,k} = \lambda_{i,k}$, for $k \in \{2, \dots, i\}$. We denote by $\overline{\mathbf{P}}_n$ its law starting from $\overline{Z}_0 = n$. Observe that the process $\overline{Z}$ fulfils the conditions of Theorem 3, and thus there exists a unique stationary distribution $\nu$. In particular,

$$\lim_{t\to\infty} \overline{\mathbf{P}}_n(\overline{Z}_t = 1) = \nu(\{1\}) > 0,$$

and this is a uniform limit on the variable $n$. Thus, for every $\epsilon \in (0, \nu(\{1\}))$ there exists $t_0 > 0$ such that

$$\inf_{n\in\mathbb{N}} \overline{\mathbf{P}}_n(\overline{Z}_{t_0} = 1) > \epsilon.$$

Since the processes differ only in the death parameter, the order-preserving coupling is straightforward. Let us denote this coupling by $\{(U_t, \overline{U}_t), t \in \mathbb{R}_{+,0}\}$. In particular, we have that $U_t$ and $\overline{U}_t$ are equal in distribution to $Z_t$ and $\overline{Z}_t$, respectively. As we said before, such a coupling preserves order, meaning that

$$U_t \le \overline{U_t} \qquad \text{for } t \in \mathbb{R}_{+,0}, \qquad \text{a.s.}$$

On the other hand, we observe that $\mathbb{P}(U_h = 0|U_0 = 1) > 0$ for $h \in \mathbb{R}_+$. Hence we conclude that for $n \in \mathbb{N}$,

$$\begin{aligned}\mathbf{P}_n(Z_{t_0+h} = 0) &= \mathbb{P}(U_{t_0+h} = 0|U_0 = n)\\ &> \mathbb{P}(U_{t_0+h} = 0|U_{t_0} = 1)\mathbb{P}(U_{t_0} = 1|U_0 = n)\\ &> \mathbb{P}(U_h = 0|U_0 = 1)\mathbb{P}(\overline{U}_{t_0} = 1|\overline{U}_0 = n)\\ &> \mathbf{P}_1(Z_h = 0)\epsilon.\end{aligned}$$

Again, it is important to note that the right-hand side of the previous equation does not depend on $n$. From the Markov property we deduce that

$$\begin{aligned}\mathbf{P}_n(\tau_0 > t) &\le \mathbf{P}_n(\tau_0 > \lfloor t/(t_0 + h)\rfloor t_0)\\ &< (1 - \mathbf{P}_1(Z_h = 0)\epsilon)^{\lfloor t/(t_0+h)\rfloor}.\end{aligned}$$

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

This allow us to conclude that stochastically $\tau_0 \leq (t_0 + h)G$, where $G$ is a geometric random variable with parameter $\mathbf{P}_1(Z_h = 0)\epsilon$. Note that $G$ is independent of the starting condition $n$. This leads to the conclusion that, for every $n \in \mathbb{N}$,

$$\mathbf{E}_n[\tau_0] < \frac{t_0 + h}{\mathbf{P}_1(Z_h = 0)\epsilon} < \infty,$$

which completes the proof. □

*Proof of Theorem 4.* In order to prove our result, we first deduce that under our assumptions the sequence $(\mathbf{P}_n)_{n\in\mathbb{N}}$ converges weakly in the space of probability measures on $D(\overline{\mathbb{N}}, [0, T])$, the Skorokhod space of càdlàg functions on $[0, T]$ with values in $\overline{\mathbb{N}}$. We follow the tightness argument provided in the proof of Theorem 1 in Donelly [10]. We observe that the process does not possess instantaneous states and also that it is stochastically monotone with respect to the starting point. In other words, the condition (A1) of Theorem 1 in [10] is satisfied. Moreover, under our assumptions the process does not explode, and the condition (A2) is also satisfied, if we denote by $B_n^N$ the branching process with interactions starting from $n$ and stopped at state $N$. Then tightness holds and we identify the finite marginal distributions by noticing that for $k \geq 1$, for $t_1, \ldots t_k \in \mathbb{R}_{+,0}$, and for $n_1, \ldots, n_k \in \mathbb{N}$, the probabilities $\mathbf{P}_n(Z_{t_1} \leq n_1, \cdots, Z_{t_k} \leq n_k)$ are nonincreasing with respect to $n$.

Next, since $Z$ and $X$ are moment duals, for any $t \in \mathbb{R}_+$ we have

$$\mathbf{P}_\infty(Z_t < \infty) = \lim_{x\to 1} \lim_{n\to\infty} \mathbf{E}_n\left[x^{Z_t}\right] = \lim_{x\to 1} \lim_{n\to\infty} \mathbb{E}_x[X_t^n] = \lim_{x\to 1} \mathbb{P}_x(X_t = 1), \tag{2.17}$$

where in the first equality we have used that $(\mathbf{P}_n)_{n\in\mathbb{N}}$ converges weakly in the space of probability measures. In order to deduce that the previous probability equals 1, we first consider the case $\Lambda \equiv 0$, $d = 0$, and $\rho = 0$. From Lemma 3, we deduce that

$$\lim_{x\to 1} \mathbb{P}_x(X_t = 1) = \lim_{x\to 1} \mathbb{P}_x(X_{T_{0,1}} = 1, T_{0,1} < t) = 1.$$

By a comparison result for diffusions (see for instance Theorem 2.2 in Dawson and Li [9]), we deduce that the latter also holds for the case $\rho > 0$, $\Lambda \equiv 0$, and $d = 0$. In other words, when $\Lambda \equiv 0$ and $d = 0$, we deduce from (2.17) that

$$\mathbf{P}_\infty(Z_t < \infty) = 1 \qquad \text{for any} \quad t \in \mathbb{R}_+,$$

and conclude that the process $Z$ comes down from infinity. Finally, the remaining cases (i.e. $d > 0$ and/or catastrophes) follow from an order-preserving coupling.

Next, we study the limit of $\mathbf{E}_n[\tau_1]$ as $n$ goes to $\infty$. If the parameters $d$, $c$, $\pi_i$, $b_i$, for $i \in \mathbb{N}$, and $\lambda_{i,k}$ for $k \in \{2, \ldots, i\}$ are as in Lemmas 4 and 5, then it is clear that $\lim_{n\to\infty} \mathbf{E}_n[\tau_1] < \infty$. If the parameters are as in Theorem 3, then the identity (1.4) implies that there exists a time $t_0$ such that $\sup_{n\in\mathbb{N}} \mathbf{P}_n(Z_{t_0} = 1) > \nu(1)/2$. By the Markov property, for every $n \geq 1$, we observe that

$$\mathbf{P}_n(\tau_1 > kt_0) < \mathbf{P}_n(Z_{it_0} \neq 1, \text{ for all } i \in \{1, 2, ..., k\}) < (1 - \nu(1)/2)^k.$$

This implies that $\mathbf{E}_n[\tau_1] < 2/\nu(1) < \infty$ for all $n \in \mathbb{N}$; in other words we have $\lim_{n\to\infty} \mathbf{E}_n[\tau_1] < 2/\nu(1)$. This completes the proof. □

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

### 2.5. Proof of Proposition 2

*Proof of Proposition 2.* We first deduce that for any choice of parameter it holds that

$$\sup_{n\in\mathbb{N}} \mathbf{E}_n[\tau_{0,1}] < \infty.$$

Let us first assume that $a = 0$. If $d > 0$, or $d = 0$ and $\pi_i = 0$ for all $i \in \mathbb{N}$, then the result follows immediately from Lemmas 4 and 5. If $d = 0$ and $\pi_i > 0$ for some $i \in \mathbb{N}$, then the identity (1.4) guarantees that for every $\epsilon < \nu(1)$ there exists $t_\epsilon > 0$ such that $\inf_{n\in\mathbb{N}} \mathbf{P}_n(Z_{t_\epsilon} = 1) > \epsilon$. Thus, by the strong Markov property we conclude that

$$\sup_{n\in\mathbb{N}} \mathbf{E}_n[\tau_{0,1}] \le \sup_{n\in\mathbb{N}} \mathbf{E}_n[\tau_1] \le \frac{t_\epsilon}{\epsilon} < \infty.$$

Next, we assume $a > 0$. In this case, we will again use an order-preserving coupling. Let $Z$ be a branching process with interactions with parameters $a > 0$, $d$, $c$, $\pi_i$, $b_i$, and $\lambda_{i,k}$ for $k \in \{2, \dots, i\}$, which represents the catastrophes. Recall that $\mathbf{P}_n$ denotes its law starting from $Z_0 = n$. We also consider another branching process with interactions $\overline{Z} = (\overline{Z}_t, t \in \mathbb{R}_{+,0})$, with parameters $\overline{a} = 0$, $\overline{d} = d$, $\overline{\pi_i} = \pi_i$, $\overline{c} = c + a$, $\overline{b_i} = b_i$, for all $i \in \mathbb{N}$, and $\overline{\lambda}_{i,k}$ for $k \in \{2, \dots, i\}$. We denote by $\overline{\mathbf{P}}_n$ its law starting from $\overline{Z}_0 = n$. Since the only parameters in which the processes $Z$ and $\overline{Z}$ differ are the annihilation and competition parameters, the coupling is straightforward (every time there is an annihilation event in $Z$ there is a competition event in $\overline{Z}$).

Let us denote by $\{(U_t, \overline{U}_t), t \in \mathbb{R}_{+,0}\}$ such a coupling, and observe that

$$U_t \le \overline{U_t} \qquad \text{for } t \in \mathbb{R}_{+,0}, \qquad \text{a.s.}$$

Since $\overline{U}$ has no annihilation, we conclude by the previous argument that

$$\sup_{n\in\mathbb{N}} \mathbf{E}_n[\tau_{0,1}] < \sup_{n\in\mathbb{N}} \mathbf{E}_n[\overline{\tau}_{0,1}] < \infty,$$

where $\overline{\tau}_{0,1}$ denotes the first hitting time of the states $\{0, 1\}$ of the process $\overline{Z}$. Whenever $d > 0$, we have

$$\sup_{n\in\mathbb{N}} \mathbf{E}_n[\tau_0] < \sup_{n\in\mathbb{N}} \mathbf{E}_n[\overline{\tau}_0] < \infty.$$

This proves Part (ii).

From the first part of the proof, we know that $Z$ visits $\{0, 1\}$. Observe that if $Z$ visits the state $\{0\}$ it will get absorbed, but if it visits the state $\{1\}$ it will get absorbed if and only if $d = 0$ and $\pi_i = 0$ for all $i \in \mathbb{N}$. The proofs of Parts (a) and (b) rely on verifying whether the states $\{0\}$ and $\{1\}$ are absorbing states or not.

For Part (a), the assumptions guarantee that for every $n \in \mathbb{N} \setminus \{1\}$ there exists some number $i \in \mathbb{Z}$ such that $Z$ goes from $n$ to $n + 2i - 1$ at a positive rate. Since $a > 0$, we have that $Z$ goes from $n$ to $n - 2$ at a positive rate. In other words, the states $\{0\}$ and $\{1\}$ are accessible.

On the other hand, we deduce that $\{1\}$ is an absorbing state if and only if $\sum_{i\in\mathbb{N}} \pi_i = 0$. If $\{1\}$ is an absorbing state, then from the first part of the proof, we deduce that $\sup_{n\in\mathbb{N}} \mathbf{E}_n[\tau_{0,1}] < \infty$. If $\{1\}$ is not absorbing, then $\{0\}$ is accessible from $\{1\}$, which implies that there exist $\epsilon > 0$ and $t_\epsilon \in \mathbb{R}_+$ such that

$$\mathbf{P}_1(Z_{t_\epsilon} = 0) > \epsilon.$$

Using the strong Markov property we conclude that

$$\sup_{n\ge 1} \mathbf{E}_n[\tau_0] < \frac{1}{\epsilon}\Big(\sup_{n\ge 1} \mathbf{E}_n[\tau_{0,1}] + t_\epsilon\Big) < \infty.$$

This completes the proof of Part (a).

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

For Part (b), we observe that $Z$ only makes jumps of even size, so the support of $\mathbf{P}_n(Z_t = \cdot)$ consists of the even or the odd numbers, depending the value of $n$. If $n$ is even, then $\{1\}$ is not accessible and $\{0\}$ is accessible. In other words, $\tau_{0,1} = \tau_0$, and by the first part of our proof, we conclude that $\sup_{n\in\mathbb{N}} \mathbf{E}_n[\tau_0] < \infty$. Similarly, if $n$ is odd then $\{1\}$ is accessible but $\{0\}$ is not. In this case, $\tau_{0,1} = \tau_1$, and again from the first part of the proof we conclude that $\sup_{n\in\mathbb{N}} \mathbf{E}_n[\tau_1] < \infty$.

On the other hand, under our assumptions, we deduce that $\{1\}$ is an absorbing state if and only if $\sum_{i\in\mathbb{N}} \pi_i = 0$. If $\sum_{i\in\mathbb{N}} \pi_i > 0$, we have that $\{1\}$ is positive recurrent. The uniform convergence to the stationary distribution in this case follows from arguments similar to those used for Theorem 3, Part (ii). We leave the details to the reader. □

## Acknowledgements

All authors would like to thank Gabriel Berzunza for many useful discussions, as well as two anonymous referees, whose careful reading led to significant improvements. A. G. C. acknowledges support from UNAM through the grant PAPIIT IA100419 and from the German Research Foundation through the Priority Programme 1590, *Probabilistic Structures in Evolution*. J. C. P. and J. L. P. also acknowledge support from the Royal Society and CONACyT-MEXICO.

## References

[1] Alkemper, R. and Hutzenthaler, M. (2007). Graphical representation of some duality relations in stochastic populations models. *Electron. Commun. Prob.* **12**, 206–220.
[2] Athreya, S. R. and Swart, J. M. (2005). Branching–coalescing particle systems. *Prob. Theory Relat. Fields* **131**, 376–414.
[3] Athreya, S. R. and Swart, J. M. (2012). Systems of branching, annihilating and coalescing particles. *Electron. J. Prob.* **17**, 1–32.
[4] Blath, J., Etheridge, A. M. and Meredith, M. (2007). Coexistence in competing populations and branching annihilating random walk. *Ann. Appl. Prob.* **17**, 1474–1507.
[5] Blath, J. and Kurt, N. (2011). Survival and extinction of caring double-branching annihilating random walk. *Electron. Comm. Prob.* **16**, 271–282.
[6] Bramson, M. and Gray, L. (1985). The survival of branching annihilating random walk. *Z. Wahrscheinlichkeitsth.* **68**, 447–460.
[7] Carinci, G., Giardina, C., Giberti, C. and Redig, F. (2015). Dualities in population genetics: a fresh look with new dualities. *Stoch. Process. Appl.* **125**, 941–969.
[8] Chen, M. F. (1991). On three classical problems for Markov chains with continuous time parameters. *J. Appl. Prob.* **28**, 305–320.
[9] Dawson, D. A. and Li, Z. (2012). Stochastic equations, flows and measure-valued processes. *Ann. Prob.* **40**, 813–857.
[10] Donnelly, P. (1991). Weak convergence to a Markov chain with an entrance boundary: ancestral processes in population genetics. *Ann. Prob.* **19**, 1102–1117.
[11] Ethier, S. N. and Kurtz, T. G. (1986). *Markov processes. Characterization and convergence.* John Wiley, New York.
[12] Foucart, C. (2013). The impact of selection in the Λ-Wright–Fisher model. *Electron. Commun. Prob.* **18**, 1–10.
[13] Foucart, C. (2019). Continuous-state branching processes with competition: duality and reflection at infinity. *Electron. J. Prob.* **24**, 1–38.
[14] González Casanova, A., Miró Pina, V. and Pardo, J. C. (2020). The Wright–Fisher model with efficiency. *Theoret. Pop. Biol.* **132**, 33–46.
[15] González Casanova, A. and Spanó, D. (2018). Duality and fixation in Ξ-Wright–Fisher processes with frequency-dependent selection. *Ann. Appl. Prob.* **28**, 250–284.
[16] Griffiths, R. C. (2014). The Λ-Fleming–Viot process and a connection with Wright–Fisher diffusion. *Adv. Appl. Prob.* **46**, 1009–1035.
[17] Hairer, M. (2016). *Lectures Notes on the Convergence of Markov Processes.* Available at http://www.hairer.org/notes/Convergence.pdf.

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press

[18] JAGERS, P. (1997). Towards dependence in general branching processes. In *Classical and Modern Branching Processes,* eds. K. B. Athreya and P. Jagers, Springer, New York, pp. 127–139.
[19] JANSEN, S. AND KURT, N. (2014). On the notion(s) of duality for Markov processes. *Prob. Surveys* **11**, 59–120.
[20] KRONE, S. AND NEUHAUSER, C. (1997). The genealogy of samples in models with selection. *Genetics* **145**, 519–534.
[21] KRONE, S. AND NEUHAUSER, C. (1997). Ancestral processes with selection. *Theoret. Pop. Biol.* **51**, 210–237.
[22] LAMBERT, A. (2005). The branching process with logistic growth. *Ann. Appl. Prob.* **15**, 1506–1535.
[23] LEVIN, D. A., PERES, Y. AND WILMER, E. L. (2009). *Markov Processes and Mixing Times.* American Mathematical Society, Providence, RI.
[24] LI, Z. AND PU, F. (2012). Strong solutions of jump-type stochastic equations. *Electron. Commun. Prob.* **17**, 1–13.
[25] LINDVALL, T. (1992). *Lectures on the Coupling Method.* John Wiley, New York.
[26] LÓPEZ, F. J., MARTÍNEZ, S. AND SANZ, G. (2000). Stochastic domination and Markovian couplings. *Adv. Appl. Prob.* **32**, 1064–1076.
[27] MATIS, J. H. AND KIFFE, T. R. (2004). On stochastic logistic population growth models with immigration and multiple births. *Theoret. Pop. Biol.* **65**, 89–104.
[28] MEYN, S. P. AND TWEEDIE, R. L. (1993). Stability of Markovian processes III: Foster–Lyapunov criteria for continuous-time processes. *Adv. Appl. Prob.* **25**, 518–548.
[29] MIJATOVIĆ, A. AND URUSOV, M. (2012). Convergence of integral functionals of one-dimensional diffusions. *Electron. Commun. Prob.* **17**, 1–13.
[30] MÖHLE, M. (1999). The concept of duality and applications to Markov processes arising in neutral population genetics models. *Bernoulli* **5**, 761–777.
[31] NÅSELL, I. (2001). Extinction and quasi-stationarity in the Verhulst logistic model. *J. Theoret. Biol.* **211**, 11–27.
[32] NORRIS, J. R. (1997). *Markov Chains.* Cambridge University Press.
[33] PITMAN, J. (1999). Coalescents with multiple collisions. *Ann. Prob.* **27**, 1870–1902.
[34] REVUZ, D. AND YOR, M. (1991). *Continuous Martingales and Brownian Motion.* Springer, Berlin, Heidelberg.
[35] SAGITOV, S. (1999). The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Prob.* **36**, 1116–1125.
[36] STURM, A. AND SWART, J. M. (2015). A particle system with cooperative branching and coalescence. *Ann. Appl. Prob.* **25**, 1616–1649.
[37] TWEEDIE, R. L. (1975). Sufficient conditions for regularity, recurrence and ergodicity of Markov processes. *Math. Proc. Camb. Phil. Soc.* **78**, 125–136.

https://doi.org/10.1017/apr.2020.59 Published online by Cambridge University Press