Chapter 4

Branching Processes II: Convergence of critical branching to Feller's CSB

Figure 4.1: Feller

4.1 Birth and Death Processes

4.1.1 Linear birth and death processes

Branching processes can be studied in discrete or continuous time. We now consider a classical continuous time version. This is a continuous time Markov chain, $\{X_t\}_{t\geq 0}$ with state space $\mathbb{N}_0$ and with linear birth and death rates, b and d and let $V = b + d \geq 0$. This corresponds to a branching system in which (independently) each particle can die or be replaced by two offspring in the interval $[t, t + \Delta t)$ with probability $V\Delta t + o(\Delta t)$. This means that the time until the first branch (birth-death event) is an exponential random variable with mean $\frac{1}{V}$. V is called the branching rate. When the particle "branches" it dies with probability $\frac{d}{b+d}$ and is replaced by two descendants with probability $\frac{b}{b+d}$. Note that this process can be built directly on a probability space containing a sequence of iid exponential (1) rv's and a sequence of iid Bernoulli ($p = \frac{b}{b+d}$) rv's (or a sequence of iid Uniform $[0, 1]$ rv's) and this description can be used to generate a simulation of the model. The special case in which $d = 0$ is called the *Yule* process.

The birth and death process can also be realized on a probability space $(\Omega, \mathcal{F}, P)$ on which independent Poisson random measures N_1, N_2 on $\mathbb{R}_+^2$ are defined. Then the birth and death process is defined via a stochastic differential equation driven by the Poisson noises, namely,

(4.1)
$$X_t = x_0 + \int_0^t \int_0^{bX(s_-)} N_1(du, ds) - \int_0^t \int_0^{dX(s-)} N_2(du, ds).$$

This equation has a pathwise unique càdlàg solution which is a continuous time Markov chain with the required transition rates. See Li-Ma [430].

Let P_{x_0} denote the resulting probability law on $D_{\mathbb{N}_0}([0, \infty))$, the space of càdlàg functions from $[0, \infty)$ to $\mathbb{N}_0$.

4.1.2 Semigroups and generating functions

Given the Markov process X_t we can associate a Markov semigroup $\{T_t : t \geq 0\}$ of operators on the Banach space $C_0(\mathbb{N}_0)$ (the space of bounded functions on $\mathbb{N}_0$, with limits at infinity) as follows:

$$T_t f(x_0) := E_{x_0}(f(X_t)) = \int f(x) P_{x_0}(X_t \in dx).$$

This semigroup determines the finite dimensional distributions of the Markov chain. This semigroup satisfies the conditions of the Hille-Yosida theorem with generator given by

$$\begin{aligned} Gf(n) &= \frac{dT_t f(n)}{dt}|_{t=0} \\ &= bn(f(n+1) - f(n)) + dn(f(n-1) - f(n)) \end{aligned}$$

Now consider the *Laplace function* of X_t starting with one particle at time 0:

$$L(t, \theta) := E_{x_0}(e^{-\theta X_t}), \quad \text{with } x_0 = 1,\ \theta \geq 0$$

Noting the outcome at the first branching time and using the independence of the particle and its offspring when a birth occurs, we obtain the nonlinear renewal-type equation

$$L(t, \theta) = e^{-Vt} e^{-\theta} + \frac{d}{b+d}(1 - e^{-Vt}) + V \frac{b}{b+d} \int_0^t e^{-Vu} L^2(t-u, \theta) du$$

Alternately, note that we can represent the jump in X_t at a branching time by the addition of an independent random variable ζ with Laplace transform $E(e^{-\theta\zeta}) = \frac{b}{b+d}e^{-\theta} + \frac{d}{b+d}e^{\theta}$. Since the branching occurs at linear rate VX_t at time t, we get

$$\begin{aligned}
\frac{\partial L(t,\theta)}{\partial t} &= \lim_{\Delta\to 0}\frac{L(t+\Delta,\theta)-L(t,\theta)}{\Delta}\\
&= \lim_{\Delta\to 0}\left(\frac{E(E(e^{-\theta X_{t+\Delta}}|X_{t-}))-E(e^{-\theta X_{t-}})}{\Delta}\right)\\
&= \lim_{\Delta\to 0}\frac{V\Delta[E(X_{t-}E(e^{-\theta(X_{t-}+\zeta)}|X_{t-}))-E(X_{t-}e^{-\theta X_{t-}})]+o(\Delta)}{\Delta}\\
&= -\{V\frac{b}{b+d}(e^{-\theta}-1)+V\frac{d}{b+d}(e^{\theta}-1)\}\frac{\partial L(t,\theta)}{\partial\theta}.
\end{aligned}$$

Here we have used $E(Xe^{-\theta X}) = -\frac{\partial L(\theta)}{\partial\theta}$. So we then have the first order PDE

$$(4.2)\quad \frac{\partial L(t,\theta)}{\partial t} + V[\frac{b}{b+d}(e^{-\theta}-1)+\frac{d}{b+d}(e^{\theta}-1)]\frac{\partial L(t,\theta)}{\partial\theta} = 0,\quad L(0,\theta) = e^{-\theta}.$$

We can solve this by finding the *characteristic curves* $(t(s),\theta(s))$ in the (t,θ) plane along which $L(t(s),\theta(s))$ is constant (refer to Garabedian (1964) [258], John (1982) [350] or Delgado (1997) [157]). We write this as

$$(4.3)\quad \frac{\partial}{\partial s}L(t(s),\theta(s)) = L_1\frac{\partial t(s)}{\partial s} + L_2\frac{\partial\theta(s)}{\partial s} = 0$$

where L_1, L_2 denote the first partial derivatives with respect to t, θ respectively. Comparing (4.3) with (4.2) leads to the *characteristic equations*

$$(4.4)\quad \frac{\partial L(s)}{\partial s} = 0,\quad \frac{\partial t(s)}{\partial s} = 1,\quad \frac{\partial\theta(s)}{\partial s} = h(\theta) = b(e^{-\theta}-1)+d(e^{\theta}-1)$$

For $b \neq d$ we obtain the characteristic curve

$$(4.5)\quad \frac{(e^{-\theta}-1)e^{(b-d)t}}{be^{-\theta}-d} = \text{ constant.}$$

and general solution

$$(4.6)\quad L(t,\theta) = \Psi(\frac{(e^{-\theta}-1)e^{(b-d)t}}{be^{-\theta}-d})$$

where Ψ is a differentiable function. From the initial condition we have

$$(4.7)\quad \Psi(\frac{e^{-\theta}-1}{be^{-\theta}-d}) = e^{-\theta X_0}.$$

Solving for Ψ we obtain for $b \neq d$ the solution

$$(4.8)\quad L(t,\theta) = \left(\frac{d(e^{-\theta}-1)e^{(b-d)t} - (be^{-\theta}-d)}{b(e^{-\theta}-1)e^{(b-d)t} - (be^{-\theta}-d)} \right)^{X_0}$$

and for $b = d$

$$(4.9)\quad L(t,\theta) = \left(\frac{1-(bt-1)(e^{-\theta}-1)}{1-bt(e^{-\theta}-1)} \right)^{X_0}.$$

Remark 4.1 *Note that the form of the Laplace transforms (4.8), (4.9)implies the* branching property*, namely, if* $X_0 = X_{0,1} + X_{0,2}$*, then the probability law of* X_t *is identical to the distribution of the sum of independent random variables* $X_{t,1} + X_{t,2}$ *where* $X_{t,i}$ *are versions of the linear birth and death process with initial conditions* $X_{0,1}, X_{0,2}$*.*

Distribution function, moments, extinction probability

Setting $b,\ d$ as the birth and death rates. Then replacing θ by $-\ln z$ in $L_t(\theta)$ we obtain the probability generating function

$$(4.10)\quad G_t(z) = L(t, -\ln z) = \sum_{k=0}^{\infty} z^k p_k(t).$$

Then expanding in a power series in z we can obtain the standard formula

$$(4.11)\quad p_0(t) = f(t),$$

$$(4.12)\quad p_n(t) = (1-f(t))(1-g(t))g(t)^{n-1},\ n \geq 1$$

where

$$(4.13)\quad f(t) = \frac{d(e^{(b-d)t}-1)}{be^{(b-d)t}-d}, \quad g(t) = \frac{b(e^{(b-d)t}-1)}{be^{(b-d)t}-d}.$$

Similarly if $b = d = \frac{V}{2}$, then

$$(4.14)\quad p_n(t) = \frac{(bt)^{n-1}}{(1+bt)^{n+1}},\ n \geq 1,$$

$$(4.15)\quad p_0(t) = \frac{bt}{1+bt}.$$

Then the extinction probability is

$$(4.16)\quad \lim_{t\to\infty} p_0(t) = \lim_{t\to\infty} \frac{d(e^{(b-d)t}-1)}{be^{(b-d)t}-d} = \begin{cases} 1 & \text{if } b \leq d, \\ \frac{d}{b} & \text{if } b > d. \end{cases}$$

Recalling that

$$(4.17)\quad E(X_t) = -\left.\frac{\partial L_t(\theta)}{d\theta}\right|_{\theta=0},$$

$$(4.18)\quad E((X_t)^2) = \left.\frac{\partial^2 L_t(\theta)}{d\theta^2}\right|_{\theta=0}.$$

we can obtain

$$(4.19)\quad E(X_t) = X_0 e^{(b-d)t},$$

$$(4.20)\quad E((X_t)^2) = (X_0)^2 e^{2(b-d)t} + \frac{X_0(b+d)}{b-d} e^{(b-d)t}(e^{(b-d)t} - 1),\ b \neq d$$

$$(4.21)\quad E((X_t)^2) = (X_0)^2 + 2bt, \quad \text{if } b = d.$$

4.2 Critical branching

Exponential growth of a population is unrealistic and therefore supercritical branching models describe only the growth of a population as long as the resources are unlimited. Otherwise logistic competition comes into play. We will return to this circle of questions throughout this course.

Only critical branching processes have the property that the mean population size is stable but as shown above the critical branching process actually suffers extinction with probability one. Nevertheless critical branching processes have played a key role in the development stochastic population models. We will later see that a key feature of critical branching is the limiting behavior of the process conditioned on non-extinction up to time t and letting $t \to \infty$. We now give two formulations of the resulting behavior.

Theorem 4.2 *Consider the BGW process Z_n with mean offspring size $m = 1$. Suppose that $\sigma^2 := \mathrm{Var}(\xi) = \mathrm{E}[\xi^2] - 1 \leq \infty$. Then*

(i) Kolmogorov

$$(4.22)\quad \lim_{n\to\infty} nP[Z_n > 0] = \frac{2}{\sigma^2}$$

(ii) Yaglom: If $\sigma < \infty$, then the conditional distribution of $\frac{Z_n}{n}$ given $Z_n > 0$ converges as $n \to \infty$ to an exponential law with mean $\frac{\sigma^2}{2}$.

We refer the proof of this to the literature [9], [438].

Theorem 4.3 *Consider the critical linear birth and death process,* $\{X_t\}$ *with* $\alpha = \frac{1}{2}$, $b = d = \frac{V}{2}$. *Then*

(i) Extinction probability: $\lim_{t\to\infty} p_0(t) = 1$.

(ii) Expected extinction time: Let $\tau := \inf\{t : X_t = 0\}$ *Then* $E[\tau] = \infty$.

(iii) Exponential limit law: conditioned on $X_t \neq 0$,

$$\frac{X_t}{t} \Rightarrow Y$$

where Y *is exponential with mean* b.

Proof. The proof is based on the explicit form of the generating function (4.9).

(i) From (4.15), $p_0(t) = \frac{bt}{bt+1} \to 1$ as $t \to \infty$.

(ii) The expected extinction time is infinite

$$E(\tau) = \int_0^\infty (1 - p_0(t))dt = \int_0^\infty \frac{1}{bt+1}dt = \infty.$$

(iii) From (4.9),

(4.23)

$$E(e^{-\frac{X_t\theta}{t}}|X_t \neq 0) = \frac{L(t, \frac{\theta}{t}) - P(X_t = 0)}{1 - P(X_t = 0)}$$

$$= \frac{\frac{1-(bt-1)(e^{-\theta}-1)}{1-bt(e^{-\theta}-1)} - \frac{bt}{bt+1}}{\frac{1}{bt+1}}$$

$$\lim_{t\to\infty} E(e^{-\frac{X_t\theta}{t}}|X_t \neq 0) = \frac{1}{1+b\theta}$$

and which is the Laplace transform of the exponential distribution with mean b.
■

4.3 Feller's continuous state branching process (CSBP)

Consider the Itô stochastic differential equation (SDE)

$$dX_t = mX_tdt + \sqrt{\gamma X_t}dW_t, \quad X_0 = x \geq 0$$

where $\{W_t\}$ is a standard Brownian motion. This equation has a non-Lipschitz coefficient but its pathwise uniqueness follows from the Yamada-Watanabe theorem [617].

Using Itô's lemma one can then check that the generator of the resulting diffusion process acting on $D(G) = \{f \in C_0^2(\mathbb{R}_+),\ xf_x,\ xf_{xx} \in C_0(\mathbb{R}_+^d)\}$ satisfies

$$(4.24)\quad Gf(x) = mx\frac{\partial f}{\partial x} + \frac{1}{2}\gamma x\frac{\partial^2 f}{\partial x^2}$$

and therefore X_t is a realization of the Feller CSBP process.

Proposition 4.4 *(Laplace transform and extinction probability)*
(a)The Laplace transform is given by

$$(4.25)\quad L(\theta, t) = \mathbb{E}_x \exp(-\theta X_t) = \exp(-u(t)x)$$

where $u(s)$ satisfies the equation:

$$(4.26)\quad \frac{\partial u}{\partial s} = mu - \frac{\gamma}{2}u^2 \quad u(0) = \theta.$$

(b) In the critical case $m = 0$

$$(4.27)\quad P_x(x_t = 0) = \exp\left(-\frac{x}{\gamma t}\right).$$

Proof. Assume that $\theta(s) \geq 0$ is differentiable. Then applying Itô's lemma ([517], Theorem 3.3, Remark 1) to $F(\theta, x) = e^{-\theta x}$ we have

$$\begin{aligned}(4.28)\quad & F(\theta(t), X_t) - F(\theta(0), X_0) \\ & = m\int_0^t X_s F_2(\theta(s), X_s)ds + \int_0^t F_2(\theta(s), X_s)\sqrt{\gamma X_s}dW_s \\ & + \int_0^t F_1(\theta(s), X_s)d\theta(s) + \frac{\gamma}{2}\int_0^t X_s F_{22}(\theta(s), X_s)ds\end{aligned}$$

Noting that $E(X_s e^{-\theta X_s}) = -L_1(\theta, s)$ we obtain

$$\begin{aligned}(4.29)\quad & \frac{\partial L(\theta(s), s)}{\partial s} = L_1(\theta(s), s)\frac{d\theta(s)}{ds} - m\theta(s)L_1(\theta(s), s) + \frac{\gamma}{2}\theta(s)^2 L_1(\theta(s), s) \\ & \quad \text{with} \quad L(\theta, 0) = e^{-\theta x}\end{aligned}$$

If u is a solution of

$$(4.30)\quad \frac{\partial u(\theta, s)}{\partial s} = mu(\theta, s) - \frac{\gamma}{2}u^2(\theta, s), \qquad u(\theta, 0) = \theta$$

then the derivative with respect to s

$$\frac{\partial}{\partial s}L(u(\theta, t - s), s) = 0, \quad 0 \leq s \leq t$$

and therefore

$$(4.31)\quad \mathbb{E}_x(e^{-\theta X_t}) = L(\theta, t) = L(u(\theta, t), 0) = e^{-u(\theta,t)x}.$$

(b) Solving (4.30) we get

$$(4.32)\quad u(\theta,t)=\frac{\theta}{(1+t\gamma\theta)},\ \text{ if } m=0$$

$$(4.33)\quad u(\theta,t)=\frac{\theta m e^{mt}}{m+\gamma\theta(e^{mt}-1)},\ \text{ if } m\neq 0.$$

If $m=0$

$$(4.34)\quad P_{X_0}(x_t=0)=\lim_{\theta\to\infty}e^{-x_0u(\theta,t)}=e^{-\frac{x_0}{\gamma t}}.$$

■

Remark 4.5 *An immediate consequence of (4.31) is that for each t, X_t is an infinitely divisible random variable. In fact the law of X_t corresponds to the law of the sum of a Poisson distributed number of independent exponential random variables. These facts will provide an important tool for the study of these processes and their infinite dimensional generalizations.*

Feller CSBP with immigration

Adding an immigration term ct to X_t, one obtains the continuous state branching with immigration process (CBI), and can verify (see e.g. Li (2006) [426]) the following:

Proposition 4.6 *Consider the continuous subcritical branching process with immigration (CBI), given by the SDE:*

$$(4.35)\quad dY_t=cdt-bY_tdt+\sqrt{\gamma Y_t}dW_t,\quad Y_0=y_0,\quad b,c>0.$$

(a) The Laplace transform of the distribution of Y_t is given by:

$$(4.36)\quad \mathbb{E}_{y_0}\exp(-\theta Y(t))=e^{-y_0u(t)-\int_0^t cu(s)\,ds};\quad \frac{\partial u}{\partial t}=-bu-\frac{\gamma}{2}u^2\quad u(0)=\theta>0.$$

(b) In the subcritical case Y_t converges to equilibrium, $Y_t\Rightarrow Y_\infty$ as $t\to\infty$, where Y_∞ has the gamma distribution *with Laplace transform*

$$(4.37)\quad L(\theta)=\frac{c}{[(b+\gamma\theta)-\gamma\theta e^{-bt}]^{1/\gamma}}.$$

Proof. (a) This can be proved using the method of Theorem 4.4. Alternately, we can prove this by consider the process with immigrants coming according to $\frac{1}{K}\sum\delta_{y_i}$ where $\{y_i\}$ are the points of a Poisson process with rate K and letting $K\to\infty$.

(b) follows from (a) by simple integration of (4.36). ■

Remark 4.7 *The critical Feller CSBP with immigration*

$$(4.38)\quad \begin{aligned} dY_t &= \beta dt + 2\sqrt{Y_t}dW_t \\ Y_0 &= y_0 \end{aligned}$$

is the square of a β-dimensional Bessel process. (See Revuz Yor [517] where this is called a $BESQ^\beta$ process). For $\beta \geq 2$, $\{0\}$ is polar. For $0 < \beta < 2$, $\{0\}$ is instantaneously reflecting. For $0 < \beta < 1$ the set $\{t : X_t = 0\}$ is a perfect set. (See Revuz Yor [517] Chap. XI.)

4.4 Diffusion limits of critical and nearly critical branching processes

4.4.1 Convergence to Feller's continuous state branching process

In a celebrated paper Feller (1951) [242] developed the diffusion approximation to branching processes using semigroup methods.

Theorem 4.8 *(Convergence of B+D and BGW processes to Feller CSBP)*

(a) Consider the sequence of birth and death process, $\{X_t^K\}$, $K \in \mathbb{N}$, with linear birth and death rates $b_K = 1 + \frac{m}{2K}$, $d_K = 1 - \frac{m}{2K}$ with $X_0^K = \lfloor Kz \rfloor$. Assume that $\frac{\lfloor Kz \rfloor}{K} \to x$ and let

$$(4.39)\quad Z_t^K := \frac{1}{K} X_{Kt}^K.$$

Then as $K \to \infty$

$$(4.40)\quad \{Z_t^K\}_{t \geq 0} \Longrightarrow \{Z_t\}_{t \geq 0},$$

where $\{Z_t\}_{t\geq 0}$ is a CSBP with generator G given by (4.24) with $\gamma = 1$ and $Z_0 = x$. The convergence is in the sense of weak convergence of probability measures on $D_{[0,\infty)}([0,\infty))$ and the limiting process is a.s. continuous.

(b) Consider a sequence of BGW processes $\{X_k^N\}$ with mean offspring sizes

$$(4.41)\quad E(\xi^N) = m_N = 1 + \frac{m}{N}$$

and offspring variances

$$(4.42)\quad \mathrm{Var}(\xi^{\mathrm{N}}) = \gamma > 0.$$

Let

$$(4.43)\quad Z_t^N := \frac{1}{N} X_{\lfloor Nt \rfloor}^N.$$

Assume that $Z_0^N \to Z_0$ *as* $N \to \infty$*. Then*

(4.44) $\{Z_t^N\}_{t\geq 0} \Longrightarrow \{Z_t\}_{t\geq 0},$

that is, Z_t^N *converges in distribution on* $D_{[0,\infty)}([0,\infty))$ *to a Markov diffusion process,* $\{Z_t\}_{t\geq 0}$*, called the Feller continuous state branching process (CSBP). The generator of the CSBP* $\{Z_t\}$ *acting on functions* $f \in C_0^2([0,\infty))$ *is given by*

(4.45) $$Gf(x) = mx\frac{\partial f}{\partial x} + \frac{1}{2}\gamma x\frac{\partial^2 f}{\partial x^2}.$$

Proof. (a) The proof follows a standard program for weak convergence of processes, namely,

- the convergence of the finite dimensional distributions,
- proof that the laws of the processes $P^K \in \mathcal{P}(D_{[0,\infty)}([0,\infty))$ are tight.

To show that the finite dimensional distributions converge, first substitute birth and death rates $b = 1 + \frac{m}{2K}$, $d = 1 - \frac{m}{2K}$, in (**??**) to obtain the Laplace transform of Z_t^K with $Z_0^K = \lfloor Kz \rfloor$ as follows:

(4.46)
$$\begin{aligned}
&E(e^{-\theta Z_t^K}) = L^K(t,\theta)\\
&= \left(-\frac{K(e^{-\frac{\theta}{K}}-1)(e^{mt}-1) - \frac{m}{2}(e^{-\frac{\theta}{K}}-1)e^{mt} - \frac{m}{2}(e^{-\frac{\theta}{K}}+1)}{K(e^{-\frac{\theta}{K}}-1)(e^{mt}-1) + \frac{m}{2}(e^{-\frac{\theta}{K}}-1)e^{mt} - \frac{m}{2}(e^{-\frac{\theta}{K}}+1)}\right)^{\lfloor Kz \rfloor}\\
&= \left(-\frac{(-\theta)(e^{mt}-1) + \frac{\theta^2}{2K} - \frac{m}{2}(e^{-\frac{\theta}{K}}-1)e^{mt} - \frac{m}{2}(e^{-\frac{\theta}{K}}+1) + O(K^{-2})}{-\theta(e^{mt}-1) + \frac{\theta^2}{2K} + \frac{m}{2}(e^{-\frac{\theta}{K}}-1)e^{mt} - \frac{m}{2}(e^{-\frac{\theta}{K}}+1) + O(K^{-2})}\right)^{\lfloor Kz \rfloor}\\
&\longrightarrow \exp\left(-\frac{m\theta z e^{mt}}{m + \theta(e^{mt}-1)}\right).
\end{aligned}$$

This coincides (see Proposition 4.4) with the Laplace transform at time t of the diffusion process with $Z_0 = x$ and with generator

(4.47) $$Gf(x) = mx\frac{\partial f}{\partial x} + \frac{1}{2}x\frac{\partial^2 f}{\partial x^2}.$$

Using the Markov property and the continuity of the transition probability in x we can then obtain convergence of the finite dimensional distributions.

To complete the proof we must verify that the probability laws of $\{Z_t^K\}_{t\geq 0}$ denoted by $P^K \in \mathcal{P}(D_{[0,\infty)}([0,\infty))$ are tight. We will use the Aldous condition. We first verify that given $\delta > 0$ there exists $0 < L < \infty$ such

(4.48) $$\sup_K P^K(\sup_{0\leq t\leq T} X^K(t) > L) \leq \delta.$$

Note that the generator of $Z_t^K = \frac{X_{Kt}^K}{K}$ is

(4.49)
$$G^K f(\frac{n}{K}) = \frac{n}{K} \cdot K^2 [f(\frac{n+1}{K}) + f(\frac{n-1}{K}) - 2f(\frac{n}{K})] + \frac{mn}{2K} \cdot K [f(\frac{n+1}{K}) - f(\frac{n-1}{K})].$$

Then

(4.50) $$M_t^K := Z_t^K - m \int_0^t Z_s^K ds \quad \text{is a martingale.}$$

By Gronwall's inequality

(4.51) $$\sup_{0 \leq t \leq T} Z_t^K \leq \sup_{0 \leq t \leq T} |M_t^K| e^{mt}.$$

Applying Doob's maximal inequality to M_t^K

(4.52) $$P(\sup_{0 \leq t \leq T} |M_t^K| \geq R) \leq \frac{E((M_T^K)^2)}{R^2}.$$

It remains to compute $E((M_T^K)^2)$. We have

(4.53) $$E((M_T^K)^2) \leq E(Z_T^K)^2) + 2|m| \int_0^T E(Z_s^K Z_T^K) ds + \int_0^T \int_0^T E(Z_s^K Z_t^K) ds dt.$$

Using (4.20) we can check that

(4.54) $$E[(Z_t^K)^2] \leq Z_0^K e^{mt} \frac{(e^{mt} - 1)}{m} + (Z_0^K)^2 e^{2mt}.$$

A simple calculation then yields

(4.55) $$E((M_T^K)^2) \leq C(T, z)$$

where $C(T, z)$ does not depend on K which proves (4.48).

We can then apply the Aldous sufficient condition for tightness, namely, given stopping times $\tau_K \leq T$ and $\delta_K \downarrow 0$ as $K \to \infty$

(4.56) $$\lim_{K \to \infty} P^K(|Z_{\tau_K + \delta_K}^K - Z_{\tau_K}^K| > \varepsilon) = 0.$$

First note that $X_{\tau_K}^K$ is tight so we can take a convergent subsequence. Then by Skorohod's representation we can put these on a common probability space so that there is a.s. convergence. In this setting assume that $X_{\tau_{K_n}}^{K_n} \to x$. In now suffices to prove that $X_{\tau_{K_n} + \delta_{K_n}}^{K_n}$ converges in distribtion to x. Then by the strong

Markov property we have

(4.57)

$$E(e^{-\theta(Z^K_{\tau_K+\delta_K})}-e^{-\theta Z^K_{\tau_K}}|Z^K_{\tau_K})$$
$$=\left(-\frac{K(e^{-\frac{\theta}{K}}-1)(e^{m\delta_K}-1)-\frac{m}{2}(e^{-\frac{\theta}{K}}-1)e^{m\delta_K}-\frac{m}{2}(e^{-\frac{\theta}{K}}+1)}{K(e^{-\frac{\theta}{K}}-1)(e^{m\delta_K}-1)+\frac{m}{2}(e^{-\frac{\theta}{K}}-1)e^{m\delta_K}-\frac{m}{2}(e^{-\frac{\theta}{K}}+1)}\right)^{\lfloor KZ^K_{\tau_K}\rfloor}-e^{-\theta Z^K_{\tau_K}}$$
$$\longrightarrow 0 \text{ on } \{\sup_{0\le t\le T} Z^K(t)\le L\} \text{ as } K\to\infty.$$

Therefore $Z^K_{\tau_K+\delta_K}-Z^K_{\tau_K}\to 0$ in distribution and for $\varepsilon,\eta>0$ we can find K_0 such that

$$(4.58)\quad P^K(|Z^K_{\tau_K+\delta_K}-Z^K_{\tau_K}|>\varepsilon)<2\eta,\quad \forall\, K\ge K_0.$$

This completes the proof of tightness.

(b) See Ethier and Kurtz ([222] Chapter 9, Theorem 1.3) for a proof based on a semigroup convergence theorem (e.g. [222], Chap. 1, Theorem 6.5). This involves showing that

$$(4.59)\quad \lim_{N\to\infty}\sup_{x=\frac{\ell}{N},\,\ell\in\mathbb{N}}|N(T_Nf(x)-f(x))-Gf(x)|=0\quad \forall\, f\in C_c^\infty([0,\infty)),$$

where

$$(4.60)\quad T_Nf(x)=E[f(\frac{1}{N}\sum_{k=1}^{Nx}\xi_k^N)],\quad x\in\{\frac{\ell}{N},\ \ell\in\mathbb{N}\}$$

and where $\{\xi_k^N\}$ are i.i.d. satisfy (4.41), (4.42).

■

Remark 4.9 *These results can also be proved using the martingale problem formulation in the same way as is carried out below for the Wright-Fisher model.*

4.5 The critical BGW tree

4.5.1 The rooted BGW tree as a metric space

We begin by recalling that given a BGW tree $\mathcal{T}\in\mathbf{T}$ with root $\varnothing$ we can embed it in the plane the edges appear according to the lexicographic order to produce a plane tree. This is given by a subset of the set of potential individuals $\mathcal{I}$ satisfying (3.1.2) and in which each parent is connected to its offspring by an edge. If we assign length 1 to each edge then a metric $d(x,y)$ can be defined on $\mathcal{T}$ by

$$(4.61)\quad d(x,y):=\text{the length of the shortest path in } \mathcal{T} \text{ from } x \text{ to } y.$$