A Branching Diffusion Model of Selection: from the Neutral Wright-Fisher Case to the One Including Mutations

Thierry E. Huillet

Laboratoire de Physique Théorique et Modélisation
CNRS-UMR 8089 et Université de Cergy-Pontoise
2 Avenue Adolphe Chauvin, F-95302, Cergy-Pontoise, France
Thierry.Huillet@u-cergy.fr

Abstract. We consider diffusion processes x_t on the unit interval. Doob-transformation techniques consist of a selection of x_t-paths procedure. The law of the transformed process is the one of a branching diffusion system of particles, each diffusing like a new process $\widetilde{x}_t$, superposing an additional drift to the one of x_t. Killing and/or branching of $\widetilde{x}_t-$particles occur at some space-dependent rate λ. For this transformed process, so in the class of branching diffusions, the question arises as to whether the particle system is sub-critical, critical or super-critical. In the first two cases, extinction occurs with probability one.

We apply this circle of ideas to diffusion processes arising in population genetics. In this setup, the process x_t is a Wright-Fisher (WF) diffusion, either neutral or with mutations.

We study a particular Doob transform which is based on the exponential function in the usual fitness parameter σ. We have in mind that this is an alternative way to introduce selection or fitness in both WF-like diffusions, leading to branching diffusion models ideas. For this Doob-transform model of fitness, the usual selection drift $\sigma x (1 - x)$ should be superposed to the one of x_t to form $\widetilde{x}_t$ which is the process that can branch, binarily.

In the first neutral case, there is a trade-off between branching events giving birth to new particles and absorption at the boundaries, killing the particles. Under our assumptions, the branching diffusion process gets eventually globally extinct in finite time with exponential tails.

In the second case with mutations, there is a trade-off between killing events removing some particles from the system and reflection at the boundaries where

the particles survive. This branching diffusion process also gets eventually globally extinct but in very long finite time with power-law tails.

Our approach relies on the spectral expansion of the transition probability kernels of both x_t and $\widetilde{x}_t$.

PACS: 87.23.Cc, 02.50.Ey, 87.23

Keywords: Diffusions, Doob transform, killing, branching, quasi-stationarity, Wright-Fisher model, neutral or with mutation, selection

1. Introduction

We consider diffusion processes on the unit interval with in mind a series of elementary stochastic models arising chiefly in population dynamics. Special emphasis is put on Doob-transformation techniques of the diffusion processes under concern.

Most of the manuscript's content focuses on the specific Wright-Fisher (WF) diffusion model and some of its variations, describing the evolution of one two-locus colony undergoing random mating, possibly under the additional actions of mutation and selection. These models found their way over the last sixty years, chiefly in mathematical population genetics. We refer to the general monographs [3], [21], [8], [5] and [9].

We now describe the content of this work in some more details.

Section 2 is devoted to generalities on one-dimensional diffusions on the unit interval $[0,1]$, say $(x_t; t \geq 0)$. Special emphasis is put on the Kolmogorov backward and forward equations, while stressing the crucial role played by the boundaries in such one-dimensional diffusion problems. Some questions such as the meaning of speed and scale functions, existence of an invariant measure, random time change... are addressed in the light of the Feller classification of boundaries. When the boundaries are absorbing, the important problem of evaluating additive functionals α along sample paths is then briefly discussed, emphasizing the prominent role played by the Green function of the model.

So far, we have dealt with a given diffusion process x_t, and recalled the various ingredients for computing the expectations of various quantities of interest, summing up over the history of its paths. In this setup, there is no distinction among paths with different destinations, nor did we allow for annihilation or creation of paths inside the domain. The Doob transform of paths is an invitation to do so. This important class of transformations is a particular instance of a more general construction based on multiplicative functionals. We fix the background.

Roughly speaking, the Doob transformation of paths procedure allows to select sample paths $x \rightarrow y$ within any laps of time t, favoring large values of the ratio $\alpha\left(y\right)/\alpha\left(x\right)$ for some specific functional $\alpha > 0$ that fixes the selection problem under study. The process solving this selection of paths procedure belongs to a class of branching diffusion processes, where independent particles diffusing like a new process $\widetilde{x}_t$ inside the interval are allowed to duplicate would the visited region of the state-space fulfill the selection of paths criterion or to die, if not. In the process, advantageous regions of the state-space are reinforced while unfavorable ones are left unexplored which is a reasonable physical way to look at selection of paths. The new process $\widetilde{x}_t$ alluded to is obtained from x_t just after superposing an additional suitable drift to the latter process. An important parameter is the state-dependent rate λ at which killing and/or branching occur. Depending on λ and on the type of boundaries which $\{0,1\}$ are to $\widetilde{x}_t$, the full transformed process can have two stopping times: the time to absorption at the boundaries and the killing time inside the domain. Besides, there is or not an opportunity that the $\widetilde{x}_t-$particles duplicate, leading or not to a daughters particle system evolving independently starting from where their mother particle died. The killing/branching issues depend on the sign of λ.

It turns out that the same diffusion methods used in the previous discussion on simple diffusions apply to the transformed processes obtained after the induced change of measure. We develop this circle of ideas.

We next apply these general ideas to diffusion processes arising in population genetics.

In Section 3 we start recalling that WF diffusion models with various drifts are continuous space-time models which can be obtained as scaling limits of a biased discrete Galton-Watson model with a conservative number of offsprings over the generations. Sections 4 and 5 are devoted to a detailed study of both the neutral WF diffusion process (WFN), the WF diffusion with selection (WFS), the WF diffusion with mutations (WFM) and the WF diffusion with mutations and selection (WFMS) respectively.

In this context, our suggestion is the following one: we can view the action of selection (or fitness) on the evolution of the allele frequency distribution, either neutral or with mutations, as a functional deformation of the sample paths of the original process, say x_t, favoring initial values $x_0 = x$ with small $\alpha\left(x\right)$ and terminal values $x_t = y$ with large $\alpha\left(y\right)$, for each t. In our construction, $\alpha\left(x\right) = e^{\sigma x}$, $\sigma > 0$, is the chosen exponential fitness functional. Stated differently and more precisely, if $p\left(x;t,y\right)$ is the transition probability density of x_t (either

WFN or WFM), our model of the action of fitness is

$$p\left(x;t,y\right) \overset{\text{fitness}}{\to} \overline{p}\left(x;t,y\right) = \frac{e^{\sigma y}}{e^{\sigma x}} p\left(x;t,y\right).$$

With this choice of α, the modification consists of selecting those paths $x \to y$ of x_t for which $e^{\sigma(y-x)}$ is large. As a result of this transformation of paths, the usual positive selection drift $\sigma x\left(1-x\right)$ has to be superposed to the one of x_t to form the new process $\widetilde{x}_t$, but our functional definition of fitness also generates an additional branching multiplicative term, translating that a particle system pops in: The resulting transformed process is a (binary) branching diffusion of WF diffusions $\widetilde{x}_t$. We may call the obtained processes the branching neutral Wright-Fisher process and the branching Wright-Fisher process with mutations. This point of view seems to be new, to the best of the author's knowledge.

Because the spectral representation of both transition probability densities of WFN or WFM are known explicitly from the works of Crow and Kimura (see [14], [3] and [4]), some easy consequences on the spectral structures of the branching transformed processes are available. For instance, it is possible to decide whether the BD process is sub-critical, critical or super-critical in the sense of ([1] and [2]).

In Section 6, we therefore give a detailed study of the binary branching diffusion process obtained while using the Doob transform $\alpha\left(x\right) = e^{\sigma x}$ when the starting point process is a WFN diffusion process. We end up with a branching particle system, each diffusing according to the WF model with a selection drift, but branching at a bounded rate $b > 0$. In this setup, the particles cannot get killed, rather they are allowed either to survive or to split: the transformed process is a pure binary branching diffusion. For this super-critical binary branching diffusion process, there is a trade-off between branching events giving birth to new particles and absorption at the boundaries, killing the particles. Thanks to the spectral representation of the WFN process, this problem is amenable to the results obtained in ([1] and [2]). Under our assumptions, the branching diffusion process turns out to be globally sub-critical: the branching diffusion process gets eventually globally extinct in finite exponential time. This requires the computation of the ground states associated with the smallest nonnegative eigenvalue of the infinitesimal generator of the transformed process which are here shown to be explicit. In particular, the expression of the quasi-stationary distribution of the particle system can be obtained in closed-form.

In Section 7, we study the binary branching diffusion process obtained while using the same Doob transform, when the starting point process is now a WF diffusion process with mutations, assuming reflecting boundaries. We end up in a branching particle system, each diffusing according to the WF model with

a mutation and selection drift, but branching at quadratic rate λ, which is bounded from below and above. Although the particles are still allowed to split, they can now also get killed at the branching times: the transformed process is again a binary branching diffusion but with killing now allowed. In this setup, there is a competition between branching/ killing events and reflection at the boundaries where the particles survive. This problem is also amenable to the results obtained in ([1] and [2]) and we end up now in a globally critical branching particle system, each diffusing according to the WF model with a mutation and selection drift. This branching diffusion process turns out to be globally critical: it also gets eventually globally extinct but now in long finite time, with power-law tails.

2. Diffusion processes on the unit interval and Doob transforms

We start with generalities on one-dimensional diffusions with the WF model and its relatives in mind. For more technical details, we refer to [6], [7], [13] and [19]. We also introduce Doob transforms as particular instances of the modification of the original diffusion process through a multiplicative functional.

2.1. **One-dimensional diffusions on the interval** $[0, 1]$. Let $(w_t; t \geq 0)$ be a standard one-dimensional Brownian (Wiener) motion. We consider a $1-$dimensional Itô diffusion driven by $(w_t; t \geq 0)$ on the interval say $[0, 1]$, see [11]. We assume it has locally Lipschitz continuous drift $f(x)$ and local standard deviation (volatility) $g(x)$, namely we consider the stochastic differential equation (SDE):

$$(1) \qquad dx_t = f(x_t)\, dt + g(x_t)\, dw_t, \ x_0 = x \in I := (0, 1)\,.$$

The condition on $f(x)$ and $g(x)$ guarantees in particular that there is no point x_* in I for which $|f(x)|$ or $|g(x)|$ would blow up and diverge as $|x - x_*| \to 0$.

The Kolmogorov backward infinitesimal generator of (1) is $G = f(x)\,\partial_x + \frac{1}{2}g^2(x)\,\partial_x^2$. As a result, for all suitable ψ in the domain of the operator $S_t := e^{tG}$, $u := u(x, t) = \mathbf{E}\psi(x_{t \wedge \tau_x})$ satisfies the Kolmogorov backward equation (KBE)

$$\partial_t u = G(u)\,;\ u(x, 0) = \psi(x)\,.$$

In the definition of the mathematical expectation u, we have $t \wedge \tau_x := \inf(t, \tau_x)$ where τ_x indicates a random time at which the process should possibly be stopped (absorbed), given the process was started in x. The description of this (adapted) absorption time is governed by the type of boundaries which $\partial I := \{0, 1\}$ are to $(x_t; t \geq 0)\,.$ A classification of the boundaries exists, due to Feller (see [13] pp. 226): they can be either accessible (namely exit or regular), or inaccessible (namely entrance or natural).

2.2. **Natural coordinate, scale, speed measure, time change.** For such Markovian diffusions, it is interesting to consider the $G-$harmonic coordinate $\varphi \in C^2$ belonging to the kernel of G, i.e. satisfying $G(\varphi) = 0$. For φ and its derivative $\varphi' := d\varphi/dy$, with $(x_0, y_0) \in (0,1)$, one finds

$$
\begin{aligned}
\varphi'(y) &= \varphi'(y_0) e^{-2\int_{y_0}^{y} \frac{f(z)}{g^2(z)} dz} \\
\varphi(x) &= \varphi(x_0) + \varphi'(y_0) \int_{x_0}^{x} e^{-2\int_{y_0}^{y} \frac{f(z)}{g^2(z)} dz} dy.
\end{aligned}
$$

One should choose a version of φ satisfying $\varphi'(y) > 0$, $y \in I$. The function φ kills the drift f of $(x_t; t \geq 0)$ in the sense that, considering the change of variable $y_t = \varphi(x_t)$,

$$
dy_t = (\varphi'g)\left(\varphi^{-1}(y_t)\right) dw_t, \ y_0 = \varphi(x).
$$

The drift-less diffusion $(y_t; t \geq 0)$ is often termed the diffusion in natural coordinates with state-space $[\varphi(0), \varphi(1)] =: \varphi(I)$. Its volatility is $\widetilde{g}(y) := (\varphi'g)(\varphi^{-1}(y))$. The function φ is often called the scale function.

Whenever $\varphi(0) > -\infty$ and $\varphi(1) < +\infty$, one can choose the integration constants defining $\varphi(x)$ so that

$$
\varphi(x) = \frac{\int_0^x e^{-2\int_0^y \frac{f(z)}{g^2(z)} dz} dy}{\int_0^1 e^{-2\int_0^y \frac{f(z)}{g^2(z)} dz} dy},
$$

with $\varphi(0) = 0$ and $\varphi(1) = 1$. In this case, the state-space of $(y_t; t \geq 0)$ is again $[0,1]$, the same as for $(x_t; t \geq 0)$.

Finally, considering the random time change $t \to \theta_t$ with inverse: $\theta \to t_\theta$ defined by $\theta_{t_\theta} = \theta$ and

$$
\theta = \int_0^{t_\theta} \widetilde{g}^2(y_s) ds,
$$

the novel diffusion $(w_\theta := y_{t_\theta}; \theta \geq 0)$ is easily checked to be identical in law to a standard Brownian motion on $\varphi(I)$. The random time t_θ can be expressed as

$$
t_\theta = \int_0^\theta m\left(\varphi^{-1}(w_\tau)\right) \left(\varphi^{-1}\right)'(w_\tau) d\tau
$$

where $m(x) := 1/(g^2\varphi')(x)$ is the (positive) speed density at $x = \varphi^{-1}(y)$. Both the scale function φ and the speed measure $d\mu = m(x) \cdot dx$ are therefore essential ingredients to reduce the original stochastic process $(x_t; t \geq 0)$ to the standard Brownian motion $(w_\theta; \theta \geq 0)$. The Kolmogorov backward infinitesimal generator G may then be written in Feller form

$$
G(\cdot) = \frac{1}{2} \frac{d}{d\mu} \left(\frac{d}{d\varphi} \cdot\right).
$$

Examples (from population genetics):
• Assume $f(x) = 0$ and $g^2(x) = x(1-x)$. This is the neutral WF model discussed at length later. This diffusion is already in natural scale and $\varphi(x) = x$, $m(x) = [x(1-x)]^{-1}$. The speed measure is not integrable.
• With $\pi_1, \pi_2 > 0$, assume $f(x) = \pi_1 - (\pi_1 + \pi_2)x$ and $g^2(x) = x(1-x)$. This is the WF model with mutation. The parameters π_1, π_2 can be interpreted as mutation rates. The drift vanishes when $x = \pi_1/\pi$ (where $\pi := \pi_1 + \pi_2$ is the total mutation pressure) which is an attracting point for the dynamics. Here:
$$\varphi'(y) = \varphi'(y_0) y^{-2\pi_1} (1-y)^{-2\pi_2}, \quad \varphi(x) = \varphi(x_0) + \varphi'(y_0) \int_{x_0}^x y^{-2\pi_1} (1-y)^{-2\pi_2} dy,$$
with $\varphi(0) = -\infty$ and $\varphi(1) = +\infty$ if $\pi_1, \pi_2 > 1/2$. The speed measure density is $m(x) \propto x^{2\pi_1 - 1}(1-x)^{2\pi_2 - 1}$ and so is always integrable. After normalization to 1, $m(x)$ is the beta$(2\pi_1, 2\pi_2)$ density.
• With $\sigma \in \mathbf{R}$, assume a model with quadratic logistic drift $f(x) = \sigma x(1-x)$ and local variance $g^2(x) = x(1-x)$. This is the WF model with selection or selection. For this diffusion (see [16]), $\varphi(x) = \frac{1-e^{-2\sigma x}}{1-e^{-2\sigma}}$ and $m(x) \propto [x(1-x)]^{-1} e^{2\sigma x}$ is not integrable. Here, σ is a selection or fitness parameter.

Time change and subordination. We start from the diffusion (1) with infinitesimal generator $G = f\partial_x + \frac{1}{2}g^2\partial_x^2$ and consider the time change problem without passing first in natural coordinate. Let the random time change

$$t \to \theta_t = \int_0^t g^2(x_s)\,ds.$$

Its inverse: $\theta \to t_\theta$ defined by $\theta_{t_\theta} = \theta$ is given by $\theta = \int_0^{t_\theta} g^2(x_s)\,ds$.
In this new stochastic time clock, the subordinated diffusion $(y_\theta := x_{t_\theta}; \theta \geq 0)$ obeys the Langevin SDE with potential $U(y) := -2\int_0^y \frac{f(z)}{g^2(z)}dz$

$$dy_\theta = \frac{f}{g^2}(y_\theta)\,d\theta + dw_\theta,$$

with backward infinitesimal generator $\widetilde{G} = g^{-2}G = \frac{f}{g^2}\partial_x + \frac{1}{2}\partial_x^2$ (See [12], pp. 164-169).

We have $\dot{\theta}_t = g^2(x_t)$ meaning that at each point x_t of the former motion, the motion of the path is accelerated or decelerated, depending on the rate $g^2(x_t) \lessgtr 1$. Note that conversely $\dot{t}_\theta = 1/g^2(y_\theta)$. Under the time substitutions, the road maps of the paths of both $(x_t; t \geq 0)$ and $(y_\theta; \theta \geq 0)$ remain exactly the same. If a path of the former process is accelerated or decelerated by its squared volatility g^2 (its local variance) at each locality, then this process boils down to the latter one. Stated differently, if we measure time by the amount of

squared volatility accumulated within each of its path, the process $(x_t; t \geq 0)$ becomes $(y_\theta; \theta \geq 0)$, both with state-space I.

2.3. **The transition probability density.** Assume that $f(x)$ and $g(x)$ are now differentiable in I. Let then $p(x; t, y)$ stand for the transition probability density function of x_t at y given $x_0 = x$. Then $p := p(x; t, y)$ is the smallest solution to the Kolmogorov forward (Fokker-Planck) equation (KFE):

$$(2) \qquad\qquad \partial_t p = G^*(p), \; p(x; 0, y) = \delta_y(x)$$

where $G^*(\cdot) = -\partial_y(f(y)\cdot) + \frac{1}{2}\partial_y^2(g^2(y)\cdot)$ is the adjoint of G (G^* acts on the terminal value y whereas G acts on the initial value x). The way one can view this partial differential equation (PDE) depends on the type of boundaries that $\{0, 1\}$ are.

Suppose for example that the boundaries $\circ := 0$ or 1 are both exit (or absorbing) boundaries. From the Feller classification of boundaries, this will be the case if $\forall y_0 \in (0, 1)$:

$$(3) \qquad (i) \; m(y) \notin L_1(y_0, \circ) \; \text{and} \; (ii) \; \varphi'(y) \int_{y_0}^y m(z)\,dz \in L_1(y_0, \circ),$$

where a function $f(y) \in L_1(y_0, \circ)$ if $\int_{y_0}^\circ |f(y)|\,dy < +\infty$.

In this case, a sample path of $(x_t; t \geq 0)$ can reach $\circ$ from the inside of I in finite time but cannot reenter. The sample paths are absorbed at $\circ$. There is an absorption at $\circ$ at time $\tau_{x, \circ} = \inf(t > 0 : x_t = \circ \mid x_0 = x)$ and $\mathbf{P}(\tau_{x, \circ} < \infty) = 1$. Whenever both boundaries $\{0, 1\}$ are absorbing, the diffusion x_t should be stopped at $\tau_x := \tau_{x, 0} \wedge \tau_{x, 1}$. Would none of the boundaries $\{0, 1\}$ be absorbing, then $\tau_x = +\infty$. This occurs when the boundaries are inaccessible.

Examples of diffusion with exit boundaries are the WF model and the WF model with selection. In the WF model including mutations, the boundaries are entrance boundaries and so are not absorbing.

When the boundaries are absorbing, then $p(x; t, y)$ is a sub-probability. Letting $\rho_t(x) := \int_0^1 p(x; t, y)\,dy$, we clearly have $\rho_t(x) = \mathbf{P}(\tau_x > t)$. Such models are non-conservative.

For one-dimensional diffusions, the transition density $p(x; t, y)$ is reversible with respect to the speed density ([13], Chapter 15, Section 13) and so detailed balance holds:

$$(4) \qquad\qquad m(x)\,p(x; t, y) = m(y)\,p(y; t, x), \; 0 < x, y < 1.$$

The speed density $m(y)$ satisfies $G^*(m) = 0$. It may be written as a Gibbs measure with density: $m(y) \propto \frac{1}{g^2(y)}e^{-U(y)}$ where the potential function $U(y)$ reads:

$$(5) \qquad\qquad U(y) = -2 \int_0^y \frac{f(z)}{g^2(z)}\,dz, \; 0 < y < 1$$

and with the measure $\frac{dy}{g^2(y)}$ standing for the reference measure.

Furthermore, if $p(s, x; t, y)$ is the transition probability density from (s, x) to (t, y), $s < t$, then $-\partial_s p = G(p)$, with terminal condition $p(t, x; t, y) = \delta_y(x)$ and so $p(s, x; t, y)$ also satisfies the KBE when looking at it backward in time. The Feller evolution semigroup being time-homogeneous, one may as well observe that with $p := p(x; t, y)$, operating the time substitution $t-s \to t$, p itself solves the KBE

(6)
$$\partial_t p = G(p), \ \ p(x; 0, y) = \delta_y(x).$$

In particular, integrating over y, $\partial_t \rho_t(x) = G(\rho_t(x))$, with $\rho_0(x) = \mathbf{1}(x \in (0, 1))$.

$p(x; t, y)$ being a sub-probability, we may define the normalized conditional probability density $q(x; t, y) := p(x; t, y) / \rho_t(x)$, now with total mass 1. We get

$$\partial_t q = -\partial_t \rho_t(x) / \rho_t(x) \cdot q + G^*(q), \ \ q(x; 0, y) = \delta_y(x).$$

The term $b_t(x) := -\partial_t \rho_t(x) / \rho_t(x) > 0$ is the time-dependent birth rate at which mass should be created to compensate the loss of mass of the original process due to absorption of $(x_t; t \geq 0)$ at the boundaries. In this creation of mass process, a diffusing particle started in x dies at rate $b_t(x)$ at point (t, y) where it is duplicated in two new independent particles both started at y (resulting in a global birth) evolving in the same diffusive way [1]. The birth rate function $b_t(x)$ depends here on x and t, not on y.

When the boundaries of x_t are absorbing, the spectra of both $-G$ and $-G^*$ are discrete (see [13] pp. 330): There exist positive eigenvalues $(\lambda_k)_{k \geq 1}$ ordered in ascending sizes and eigenvectors $(v_k, u_k)_{k \geq 1}$ of both $-G^*$ and $-G$ satisfying $-G^*(v_k) = \lambda_k v_k$ and $-G(y_k) = \lambda_k u_k$ such that, with $\langle u_k, v_k \rangle := \int_0^1 u_k(x) v_k(x) dx$ and $b_k := \langle u_k, v_k \rangle^{-1}$, the spectral representation

(7)
$$p(x; t, y) = \sum_{k \geq 1} b_k e^{-\lambda_k t} u_k(x) v_k(y)$$

holds.

Let $\lambda_1 > \lambda_0 = 0$ be the smallest non-null eigenvalue of the infinitesimal generator $-G^*$ (and of $-G$). Clearly, $-\frac{1}{t} \log \rho_t(x) \underset{t \to \infty}{\to} \lambda_1$ and by L' Hospital rule therefore $b_t(x) \underset{t \to \infty}{\to} \lambda_1$. Putting $\partial_t q = 0$ in the latter evolution equation,

---

[1] Consider a diffusion process with forward infinitesimal generator G^* governing the evolution of $p(x; t, y)$. Suppose that a sample path of this process has some probability that it will be killed or create a new copy of itself, and that the killing and birth rates d and b depend on the current location y of the path. Then the process with the birth and death opportunities of a path has the infinitesimal generator $\lambda(y) \cdot + G^*(\cdot)$, where $\lambda(y) = b(y) - d(y)$. The rate can also depend on t and x.

independently of the initial condition x

(8) $$q(x; t, y) \underset{t \to \infty}{\to} q_\infty(y) = v_1(y),$$

where v_1 is the eigenvector of $-G^*$ associated to λ_1, satisfying $-G^* v_1 = \lambda_1 v_1$. The limiting probability v_1/norm (after a proper normalization) is called the quasi-stationary Yaglom limit law of $(x_t; t \geq 0)$ conditioned on being currently alive at all time t (see [23]).

2.4. **Additive functionals along sample paths.** Let $(x_t; t \geq 0)$ be the diffusion model defined by (1) on the interval I where both endpoints are assumed absorbing (exit). This process is thus transient and non-conservative. We wish to evaluate the nonnegative additive quantities

$$\alpha(x) = \mathbf{E}\left(\int_0^{\tau_x} c(x_s)\, ds + d(x_{\tau_x})\right),$$

where the functions c and d are both assumed nonnegative on I and $\partial I = \{0, 1\}$. The functional $\alpha(x) \geq 0$ solves the Dirichlet problem:

$$-G(\alpha) = c \text{ if } x \in I$$
$$\alpha = d \text{ if } x \in \partial I,$$

and α is a super-harmonic function for G, satisfying $-G(\alpha) \geq 0$.

Some examples:

1. Assume $c = 1$ and $d = 0$: here, $\alpha = \mathbf{E}(\tau_x)$ is the mean time of absorption (average time spent in $(0, 1)$ before absorption).

2. Whenever both $\{0, 1\}$ are exit boundaries, it is of interest to evaluate the probability that x_t first hits $[0, 1]$ (say) at 1, given $x_0 = x$. This can be obtained by choosing $c = 0$ and $d(\circ) = \mathbf{1}(\circ = 1)$.

3. Let $y \in I$ and put $c = \frac{1}{2\varepsilon}\mathbf{1}(x \in (y - \varepsilon, y + \varepsilon))$ and $d = 0$. As $\varepsilon \to 0$, c converges weakly to $\delta_y(x)$ and, $\alpha =: \mathfrak{g}(x, y) = \mathbf{E}\left(\lim \frac{1}{2\varepsilon}\int_0^{\tau_x}\mathbf{1}_{(y-\varepsilon, y+\varepsilon)}(x_s)\, ds\right) = \int_0^\infty p(x; s, y)\, ds$ is the Green function, solution to:

$$-G(\mathfrak{g}) = \delta_y(x) \text{ if } x \in I$$
$$\mathfrak{g} = 0 \text{ if } x \in \partial I.$$

$\mathfrak{g}$ is therefore the mathematical expectation of the local time at y, starting from x (the sojourn time density at y). The solution is known to be (see [13], pp. 198 or [5], pp. 280)

$$\mathfrak{g}(x, y) = 2m(y)(\varphi(1) - \varphi(y))\frac{\varphi(x) - \varphi(0)}{\varphi(1) - \varphi(0)} \text{ if } x \leq y$$

(9) $$\mathfrak{g}(x, y) = 2m(y)(\varphi(y) - \varphi(0))\frac{\varphi(1) - \varphi(x)}{\varphi(1) - \varphi(0)} \text{ if } x \geq y.$$

The Green function is of particular interest to solve the general problem of evaluating additive functionals $\alpha(x)$. Indeed, as is well-known, see [13] for example, the integral operator with respect to the Green kernel inverts the second order operator $-G$ leading to

$$\alpha(x) = \int_I \mathfrak{g}(x,y)c(y)\,dy \text{ if } x \in I$$
$$\alpha = d \text{ if } x \in \partial I.$$

Under this form, $\alpha(x)$ appears as a potential function and any potential function is super-harmonic. Note that for all harmonic function $h \geq 0$ satisfying $-G(h) = 0$,

$$\alpha_h(x) := \int_I \mathfrak{g}(x,y)c(y)\,dy + h(x)$$

is again super-harmonic because $-G(\alpha_h) = c \geq 0$.

2.5. **Transformation of sample paths (Doob-transform) producing killing and/or branching.** In the preceding Subsections, we have dealt with a given process and recalled the various ingredients for the expectations of various quantities of interest, summing over the history of paths. In this setup, there is no distinction among paths with different destinations nor did we allow for annihilation or creation of paths inside the domain before the process reached one of the boundaries. The Doob transform of paths allows to do so.

Consider a one-dimensional diffusion $(x_t; t \geq 0)$ as in (1). Let $p(x;t,y)$ be its transition probability. Let $\alpha(x) \geq 0$ as $x \in [0,1]$.

Define a new transformed stochastic process $(\overline{x}_t; t \geq 0)$ by its transition probability

$$(10) \qquad\qquad \overline{p}(x;t,y) = \frac{\alpha(y)}{\alpha(x)}p(x;t,y).$$

In this construction of $(\overline{x}_t; t \geq 0)$ through a change of measure, sample paths $x \to y$ of $(x_t; t \geq 0)$ with a large value of the ratio $\alpha(y)/\alpha(x)$ are favored. This is a selection of paths procedure due to Doob (see [6]).

The KFE for $\overline{p}$ clearly is $\partial_t \overline{p} = \overline{G}^*(\overline{p})$, with $p(x;0,y) = \delta_y(x)$ and $\overline{G}^*(\overline{p}) = \alpha(y)G^*(\overline{p}/\alpha(y))$. The adjoint Kolmogorov backward operator of the transformed process is therefore by duality

$$(11) \qquad\qquad \overline{G}(\cdot) = \frac{1}{\alpha(x)}G(\alpha(x)\cdot).$$

Developing, with $\alpha'(x) := d\alpha(x)/dx$ and $\widetilde{G}(\cdot) := \frac{\alpha'}{\alpha}g^2\partial_x(\cdot) + G(\cdot)$, we get

$$(12) \qquad\qquad \overline{G}(\cdot) = \frac{1}{\alpha}G(\alpha)\cdot + \widetilde{G}(\cdot) =: \lambda(x)\cdot + \widetilde{G}(\cdot)$$

and the new KB operator can be obtained from the latter by adding a drift term $\frac{\alpha'}{\alpha}g^2\partial_x$ to the one in G of the original process to form a new process $(\widetilde{x}_t; t \geq 0)$ with the KB operator $\widetilde{G}$ and by killing or branching its sample paths at rate $\lambda(x) := G(\alpha)/\alpha$. In others words, with $\widetilde{f}(x) := f(x) + \frac{\alpha'}{\alpha}g^2(x)$, the novel time-homogeneous SDE to consider is

$$(13) \qquad d\widetilde{x}_t = \widetilde{f}(\widetilde{x}_t)\,dt + g(\widetilde{x}_t)\,dw_t, \ \widetilde{x}_0 = x \in (0,1),$$

possibly killed or branching at rate $\lambda(x)$ as soon as $\lambda \neq 0$. Whenever $(\widetilde{x}_t; t \geq 0)$ is killed, it enters conventionally into some coffin state $\{\partial\}$ added to the state-space.

Let us look at special cases:

(i) Suppose $\alpha \geq 0$ is such that $-G(\alpha) \geq 0$ (By $\alpha \geq 0$, we mean $\alpha > 0$ in I, possibly with $\alpha(0)$ or $\alpha(1)$ equal 0). Then α is called a super-harmonic (or excessive) function for the process with infinitesimal generator G.

In this case, the rate $\lambda(x) =: -d(x)$ satisfies $\lambda(x) \leq 0$ and only killing occurs at rate $d(x)$. Let $\widetilde{\tau}_x$ be the new absorption time at the boundaries of $(\widetilde{x}_t; t \geq 0)$ started at x (with $\widetilde{\tau}_x = \infty$ would the boundaries be inaccessible to the new process $\widetilde{x}_t$). Let $\widetilde{\tau}_{x,\partial}$ be the killing time of $(\widetilde{x}_t; t \geq 0)$ started at x (the hitting time of ∂), with $\widetilde{\tau}_{x,\partial} = \infty$ if $G(\alpha) \equiv 0$. Then $\overline{\tau}_x := \widetilde{\tau}_x \wedge \widetilde{\tau}_{x,\partial}$ is the novel stopping time for $(\widetilde{x}_t; t \geq 0)$. The SDE for $(\widetilde{x}_t; t \geq 0)$, together with its global stopping time $\overline{\tau}_x$ characterize the new process $(\overline{x}_t; t \geq 0)$ with full generator $\overline{G}$ to consider.

(ii) Suppose $\alpha \geq 0$ is such that $-G(\alpha) \leq 0$. Then α is called a sub-harmonic function for the process with generator G.

In this case, the rate $\lambda(x) =: b(x)$ satisfies $\lambda(x) \geq 0$ and only branching occurs at rate $b(x)$. The transformed process (with infinitesimal backward generator $\overline{G}$) accounts for a branching diffusion where a diffusing mother particle (with generator $\widetilde{G}$ and started at x) lives a random exponential time with constant rate 1. When the mother particle dies, it gives birth to a spatially dependent random number $M(x)$ of particles, with mean $\mu(x) = 1 + \lambda(x)$ (where $M(x) \overset{d}{=} 1 + \Delta(\lambda(x))$ and $\Delta(\lambda(x))$ is a geometrically distributed random variable on $\{0, 1, 2, ...\}$ with mean $\lambda(x)$). Then $M(x)$ independent daughter particles are started afresh where their mother particle died, with the event $M(x) = 0$ impossible; they move along a diffusion governed by $\widetilde{G}$ and reproduce, independently and so on for the subsequent particles.

(iii) If $\lambda(x) =: b(x)$ is bounded above, λ may be put under the alternative form

$$\lambda(x) = \lambda_*(\mu(x) - 1),$$

where $\lambda_* = \sup_{x \in [0,1]} \lambda(x)$ and $1 \leq \mu(x) \leq 2$. In this case, we can assume that $M(x)$ can only take the values 1 or 2 with probability $p_1(x)$ and $p_2(x)$ respectively, with $p_1(x) + p_2(x) = 1$. Then, $\mu(x) = \mathbf{E}(M(x)) = p_1(x) + 2p_2(x) = 1 + p_2(x)$ and

$$\lambda(x) = \lambda_* p_2(x).$$

Note that $\mu(x) - 1 = p_2(x) > p_1(x) = 2 - \mu(x)$. We get a binary branching process at rate λ_* with the event to produce two particles being more likely than the one to produce a single one, whatever is x.

(iv) Whenever α is such that $-G(\alpha)$ has no constant sign, then both killing and branching can simultaneously occur at the death of the mother particle. $\lambda(x)$ may be put under the form $\lambda(x) = b(x) - d(x)$ where $b(x)$ and $d(x)$ are the birth (branching) and death (killing) components of $\lambda(x)$.

(v) Suppose $\lambda(x)$ is bounded below and let $\lambda_* = -\inf_{x \in [0,1]} \lambda(x) > 0$. Then one may view λ as

$$\lambda(x) = \lambda_* (\mu(x) - 1),$$

where $\mu(x) \geq 0$. In this case, branching occurs at rate λ_*. When the mother particle dies, it gives birth to a spatially dependent random number $M(x)$ of particles (where $M(x) \stackrel{d}{=} \Delta(\mu(x))$ and $\Delta(\mu(x))$ is a geometrically distributed random variable on $\{0, 1, 2, ...\}$ with mean $\mu(x) = 1 + \lambda(x)/\lambda_*$). With $p_m(x) = \mathbf{P}(M(x) = m) = p_0(x) q(x)^m$, $m \geq 0$, $p_0(x) = \frac{1}{1+\mu(x)}$, $q_0(x) = 1 - p_0(x)$

$$\lambda(x) = \lambda_* \left(\sum_{m \geq 1} m p_m(x) - 1 \right) = \lambda_* \left(\sum_{m \geq 2} (m-1) p_m(x) - p_0(x) \right).$$

Thus, the decomposition $\lambda(x) = b(x) - d(x)$ holds, where b and d can be read from

$$\lambda(x) = \lambda_* \left(\frac{\mu(x)^2}{1 + \mu(x)} - \frac{1}{1 + \mu(x)} \right).$$

(vi) In some other examples, the killing/branching rate $\lambda = G(\alpha)/\alpha$ is bounded above and below. Then λ may be put under the form

$$\lambda(x) = \lambda_* (\mu(x) - 1),$$

where $\lambda_* = \sup_{x \in [0,1]} |\lambda(x)|$ and $0 \leq \mu(x) \leq 2$. In this case, we can assume that $M(x)$ can only take the values 0 or 2 with probability $p_0(x)$ and $p_2(x)$

respectively, with $p_0(x) + p_2(x) = 1$. Then, $\mu(x) = \mathbf{E}(M(x)) = 2p_2(x)$ and

$$\lambda(x) = \lambda_*(2p_2(x) - 1) = \lambda_*(p_2(x) - p_0(x)),$$

giving a simple decomposition of λ in the form $\lambda(x) = b(x) - d(x)$ with the mother particle living a random exponential time now with constant rate λ_* before giving birth to none or two descending particles (a binary branching process). Note that $p_0(x) \geq p_2(x)$ (respectively $p_0(x) \geq p_2(x)$) when $\mu(x) \leq 1$ ($\mu(x) \geq 1$).

Examples of α. When $(x_t; t \geq 0)$ is non-conservative, consider

$$\alpha(x) = \mathbf{E}\left(\int_0^{\tau_x} c(x_s)\, ds + d(x_{\tau_x})\right),$$

where the functions c and d are both assumed nonnegative on I and $\partial I = \{0, 1\}$. Then $\alpha \geq 0$ solves the Dirichlet equation $-G\alpha(x) = c(x) \geq 0$ on I ($= d(x)$ on ∂I) and so α is super-harmonic or excessive. We refer to [10] for examples of Doob transforms based on such super-harmonic functions allowing to understand various conditionings of interest when the starting point process $(x_t; t \geq 0)$ is a neutral WF diffusion or a WF diffusion with selection.

Whenever α is super-harmonic for G, then $\beta = 1/\alpha \geq 0$ is sub-harmonic for $\widetilde{G} = G + \frac{\alpha'}{\alpha}g^2\partial_x$. This results from the obvious identity

$$\beta^{-1}\widetilde{G}(\beta) = -\alpha^{-1}G(\alpha),$$

showing that $-G\alpha \geq 0$ entails $-\widetilde{G}(\beta) \leq 0$.

Whenever $(x_t; t \geq 0)$ is conservative and ergodic

$$\frac{1}{t}\mathbf{E}_x\int_0^t c(x_s)\, ds \underset{t\to\infty}{\to} \mu(c) := \int_0^1 c(y)\, d\mu(y)$$

where $d\mu = m(y)\, dy$ is the invariant probability measure of $(x_t; t \geq 0)$. Define

$$\lim_{t\to\infty}\mathbf{E}_x\int_0^t c(x_s)\, ds - t\mu(c) = \alpha(x).$$

Thus

$$\alpha(x) := \int_0^\infty (\mathbf{E}_x(c(x_s)) - \mu(c))\, ds$$

solves the Poisson equation

$$-G\alpha(x) = \widetilde{c}(x) := c(x) - \mu(c).$$

We conclude that α is $G-$super-harmonic if ever $c(x) \geq \mu(c), \forall x.$ ⋄

Background (multiplicative functional and path integral). The Doob transforms used here are particular instances of more general transformations based on multiplicative functionals. Let x_t be the diffusion process (1) governed by $G = f\partial_x + \frac{1}{2}g^2\partial_x^2$ with $x_0 = x$.

Define the multiplicative functional M_t as the solution of the differential equation

$$dM_t = M_t \cdot (a(x_t)\,dt + b(x_t)\,dw_t), \ \ M_0 = 1,$$

where a and b are arbitrary twice differentiable functions. Integrating, we get

$$M_t = e^{\int_0^t \left(a - \frac{1}{2}b^2\right)(x_s)ds + \int_0^t b(x_s)dw_s}.$$

Let B be a Borel subset of I. Define a new process whose density $\overline{p}$ is obtained after a modification of the original one while using the multiplicative modulation factor M_t as

$$\int_B \overline{p}(x;t,y)\,dy := \mathbf{E}_x\left[M_t \mathbf{1}(x_t \in B)\right] = \int_B \mathbf{E}_x\left[M_t \mid x_t = y\right] p(x;t,y)\,dy.$$

Integrating M_t over paths with fixed two endpoints x and y, $\mathbf{E}_x\left[M_t \mid x_t = y\right]$ can be interpreted as the Radon-Nykodym derivative of $\overline{p}$ with respect to p, the density of x_t. By duality, let

$$v(x,t) = \mathbf{E}_x\left[M_t \psi(x_t)\right], \ v(x,0) = \psi(x).$$

Applying Itô calculus, we get

$$\partial_t v = \overline{G}(v) = (G + gb\partial_x + a)(v) =: \left(\widetilde{G} + a\right)(v),$$

where the modified backward infinitesimal generator $\overline{G}$ is obtained by adding a drift term $gb\partial_x$ to G to produce $\widetilde{G}$ and a multiplicative part $a\cdot$. The adjoint KFE giving the evolution of $\overline{p}$ is thus

$$\partial_t \overline{p} = \overline{G}^*(\overline{p}) = \left(\widetilde{G}^* + a\right)(\overline{p}), \ \overline{p}(x;0,y) = \delta_y(x).$$

- (Cameron-Martin-Girsanov) For instance, when $a = 0$ and $b = -f/g$, the generator of the transformed diffusion is $\overline{G} = \frac{1}{2}g^2\partial_x^2$ killing the drift term of the original process governed by G. In this case,

$$M_t = e^{-\frac{1}{2}\int_0^t \left[\left(\frac{f}{g}\right)^2(x_s)ds + 2\frac{f}{g}(x_s)dw_s\right]}.$$

Clearly in this case M_t is a martingale with $\mathbf{E}_x(M_t) = 1$, assuming b to be bounded. This construction kills the drift of the original process while using a change of measure.

- (Feynman-Kac) When $b = 0$, the generator of the transformed diffusion is $\overline{G} = G + a$ adding a multiplicative component a to the one G governing the original process. In this case

$$M_t = e^{\int_0^t a(x_s)ds}$$

is the exponential of the integrated rate. If $v(x, t) = \mathbf{E}_x[M_t \psi(x_t)]$, $v(x, 0) = \psi(x)$, then v solves

$$\partial_t v = \overline{G}(v) = (G + a)(v), \quad v(x, 0) = \psi(x).$$

In particular, if $v(x, t) = \mathbf{E}_x[M_t]$, $v(x, 0) = \mathbf{1}(x \in (0, 1))$, then v solves

$$\partial_t v = \overline{G}(v) = (G + a)(v), \quad v(x, 0) = \mathbf{1}(x \in (0, 1)).$$

- (Doob) Suppose now

$$dM_t = M_t\left(\alpha^{-1}(x_t)\,d\alpha(x_t)\right), \quad M_0 = 1.$$

This M_t is a particular instance of the general M_t introduced above. Indeed, applying Itô calculus,

$$\alpha^{-1}(x_t)\,d\alpha(x_t) = \alpha^{-1}\alpha'\left[fdt + gdw\right] + \frac{1}{2}\alpha^{-1}\alpha''g^2dt,$$

leading to

$$
\begin{aligned}
a &= \alpha^{-1}\left(f\alpha' + \frac{g^2}{2}\alpha''\right) = G(\alpha)/\alpha =: \lambda(x) \\
b &= \alpha^{-1}\alpha'g.
\end{aligned}
$$

Thus $\overline{G} = G + gb\partial_x + a = G + \alpha^{-1}\alpha'g^2\partial_x + \lambda(x)$ as already observed earlier.

Now, from the differential generation of M_t,

$$M_t = \frac{\alpha(x_t)}{\alpha(x)}, \quad M_0 = 1$$

only depends on the terminal and initial values of $(x_s; 0 \leq s \leq t)$ and not on its intermediate values (such a particular Doob transformation is thus a gauge). Thus here $\mathbf{E}_x[M_t \mid x_t = y] = \frac{\alpha(y)}{\alpha(x)}$ consistently with the definition $\overline{p}(x; t, y) = \frac{\alpha(y)}{\alpha(x)}p(x; t, y)$.

A super-harmonic example.

Although this work chiefly focuses on Doob-transforms where branching is present in λ, let us give a significant example where the Doob transform just produces killing like in (i). Suppose $(x_t; t \geq 0)$ is a non-conservative diffusion. Let λ_1 be the smallest non-null eigenvalue of the infinitesimal generator G of $(x_t; t \geq 0)$. Let $\alpha = u_1$ be the corresponding eigenvector, that is satisfying $-Gu_1 = \lambda_1 u_1 \geq 0$ with boundary conditions $u_1(0) = u_1(1) = 0$. Then $c =$

$\lambda_1 u_1$. The new KB operator associated to the transformed process $(\overline{x}_t; t \geq 0)$ is

(14) $$\overline{G}(\cdot) = \frac{1}{\alpha} G(\alpha) \cdot + \widetilde{G}(\cdot) = -\lambda_1 \cdot + \widetilde{G}(\cdot),$$

obtained while killing the sample paths of the process $(\widetilde{x}_t; t \geq 0)$ governed by $\widetilde{G}$ at constant death rate $d = \lambda_1$. The transition probability of the transformed stochastic process $(\overline{x}_t; t \geq 0)$ is

$$\overline{p}(x; t, y) = \frac{u_1(y)}{u_1(x)} p(x; t, y).$$

Define $\widetilde{p}(x; t, y) = e^{\lambda_1 t} \overline{p}(x; t, y)$. It is the transition probability of the process $(\widetilde{x}_t; t \geq 0)$ governed by $\widetilde{G}$; it corresponds to the original process $(x_t; t \geq 0)$ conditioned on never hitting the boundaries $\{0, 1\}$ (the so-called $Q-$process of $(x_t; t \geq 0)$, see [18]). It is simply obtained from $(x_t; t \geq 0)$ by adding the additional drift term $\frac{u_1'}{u_1} g^2$ to f, where u_1 is the eigenvector of G associated to its smallest non-null eigenvalue. The determination of $\alpha = u_1$ is a Sturm-Liouville problem. When t is large, to the dominant order

$$p(x; t, y) \sim e^{-\lambda_1 t} \frac{u_1(x) v_1(y)}{\langle u_1, v_1 \rangle},$$

where v_1 is the Yaglom limit law of $(x_t; t \geq 0)$. Therefore

(15) $$\widetilde{p}(x; t, y) \sim e^{\lambda_1 t} \frac{u_1(y)}{u_1(x)} e^{-\lambda_1 t} \frac{u_1(x) v_1(y)}{\langle u_1, v_1 \rangle} = \frac{u_1(y) v_1(y)}{\langle u_1, v_1 \rangle}.$$

Thus the limit law of the $Q-$process $(\widetilde{x}_t; t \geq 0)$ is the normalized Hadamard product of the eigenvectors u_1 and v_1 associated respectively to G and G^*. On the other hand, the limit law of $(\widetilde{x}_t; t \geq 0)$ is directly given by

(16) $$\widetilde{p}(x; t, y) \underset{t \to \infty}{\to} \widetilde{p}(y) = \frac{1}{Z g^2(y)} e^{2 \int_0^y \frac{f(z) + \left(\frac{u_1'}{u_1} g^2\right)(z)}{g^2(z)} dz} = \frac{u_1^2(y)}{Z g^2(y)} e^{2 \int_0^y \frac{f(z)}{g^2(z)} dz},$$

where Z is the appropriate normalizing constant. Comparing (15) and (16)

$$v_1(y) = \frac{u_1(y)}{g^2(y)} e^{2 \int_0^y \frac{f(z)}{g^2(z)} dz} = u_1(y) m(y).$$

The eigenvector v_1 associated to G^* is therefore equal to the eigenvector u_1 associated to G times the speed density of $(x_t; t \geq 0)$.

When dealing for example with the neutral WF diffusion (see Section 4 for additional details), it is known that $\lambda_1 = 1$ with $u_1 = x(1 - x)$ and $v_1 \equiv 1$. The $Q-$process $(\widetilde{x}_t; t \geq 0)$ in this case obeys

(17) $$d\widetilde{x}_t = (1 - 2\widetilde{x}_t) dt + \sqrt{\widetilde{x}_t(1 - \widetilde{x}_t)} dw_t,$$

with an additional stabilizing drift toward $1/2$: $\widetilde{f}(x) = \frac{u_1'}{u_1} g^2(x) = 1 - 2x$.

The limit law of the $Q-$process $(\widetilde{x}_t; t \geq 0)$ in this case is $6y(1-y)$. The latter conditioning is more stringent than the Yaglom conditioning and so the limiting law has more mass away from the boundaries (compare with the uniform quasi-stationary Yaglom limit (8) with $v_1 \equiv 1$).

3. The Wright-Fisher example

In this Section, we briefly and informally recall that the celebrated WF diffusion process with or without a drift may be viewed as a scaling limit of a simple two alleles discrete space-time branching process preserving the total number N of individuals in the subsequent generations (see [13], [7], for example).

3.1. **The neutral Wright-Fisher model.** Consider a discrete-time Galton Watson branching process preserving the total number of individuals in each generation. We start with N individuals. The initial reproduction law is defined as follows: Let $|\mathbf{k}_N| := \sum_{m=1}^{N} k_m = N$ and $\mathbf{k}_N := (k_1, ..., k_N)$ be integers. Assume the first-generation random offspring numbers $\boldsymbol{\nu}_N := (\nu_N(1), ..., \nu_N(N))$ admit the following joint exchangeable polynomial distribution on the discrete simplex $|\mathbf{k}_N| = N$:

$$\mathbf{P}(\boldsymbol{\nu}_N = \mathbf{k}_N) = \frac{N! \cdot N^{-N}}{\prod_{n=1}^{N} k_n!}.$$

This distribution can be obtained by conditioning N independent Poisson distributed random variables on summing to N. Assume subsequent iterations of this reproduction law are independent so that the population is with constant size for all generations.

Let $N_r(n)$ be the offspring number of the n first individuals at the discrete generation $r \in \mathbf{N}_0$ corresponding to (say) allele A_1 (the remaining number $N - N_r(n)$ counts the number of alleles A_2 at generation r). This sibship process is a discrete-time Markov chain with binomial transition probability given by:

$$\mathbf{P}(N_{r+1}(n) = k' \mid N_r(n) = k) = \binom{N}{k'}\left(\frac{k}{N}\right)^{k'}\left(1 - \frac{k}{N}\right)^{N-k'}.$$

Assume next that $n = [Nx]$ where $x \in (0, 1)$. Then, as well-known, the dynamics of the continuous space-time re-scaled process $x_t := N_{[Nt]}(n)/N$, $t \in \mathbf{R}_+$ can be approximated for large N, to the leading term in N^{-1}, by a Wright-Fisher-Itô diffusion on $[0, 1]$ (the purely random genetic drift case):

(18) $$dx_t = \sqrt{x_t(1 - x_t)}dw_t, \ x_0 = x.$$

Here $(w_t; t \geq 0)$ is a standard Wiener process. For this scaling limit process, a unit laps of time $t = 1$ corresponds to a laps of time N for the original discrete-time process; thus time is measured in units of N. If the initial condition is

$x = N^{-1}$, x_t is the diffusion approximation of the offspring frequency of a singleton at generation $[Nt]$.

Equation (18) is a $1-$dimensional diffusion as in (1) on $[0, 1]$, with zero drift $f(x) = 0$ and volatility $g(x) = \sqrt{x(1-x)}$. This diffusion is already in natural coordinate and so $\varphi(x) = x$. The scale function is x and the speed measure $[x(1-x)]^{-1} dx$. One can check that both boundaries are exit in this case: The stopping time is $\tau_x = \tau_{x,0} \wedge \tau_{x,1}$ where $\tau_{x,0}$ is the extinction time and $\tau_{x,1}$ the fixation time. The corresponding infinitesimal generators are $G(\cdot) = \frac{1}{2}x(1-x)\partial_x^2(\cdot)$ and $G^*(\cdot) = \frac{1}{2}\partial_y^2(y(1-y)\cdot)$.

3.2. **Non-neutral cases.** Two alleles WF models (with non-null drifts) are classically obtained by considering the binomial transition probabilities $\text{bin}(N, p_N)$:

$$\mathbf{P}\left(N_{r+1}(n) = k' \mid N_r(n) = k\right) = \binom{N}{k'}\left(p_N\left(\frac{k}{N}\right)\right)^{k'}\left(1 - p_N\left(\frac{k}{N}\right)\right)^{N-k'}$$

where

$$p_N(x) : x \in (0, 1) \to (0, 1)$$

is now some state-dependent probability (which is different from the identity x) reflecting some deterministic evolutionary drift from the allele A_1 to the allele A_2. For each r, we have

$$\mathbf{E}\left(N_{r+1}(n) \mid N_r(n) = k\right) = Np_N\left(\frac{k}{N}\right)$$

$$\sigma^2\left(N_{r+1}(n) \mid N_r(n) = k\right) = Np_N\left(\frac{k}{N}\right)\left(1 - p_N\left(\frac{k}{N}\right)\right)$$

which is amenable to a diffusion approximation in terms of $x_t := N_{[Nt]}(n)/N$, $t \in \mathbf{R}_+$ under suitable conditions.

For instance, taking

$$p_N(x) = (1 - \pi_{2,N})x + \pi_{1,N}(1 - x)$$

where $(\pi_{1,N}, \pi_{2,N})$ are small (N-dependent) mutation probabilities from A_2 to A_1 (respectively A_1 to A_2). Assuming $(N \cdot \pi_{1,N}, N \cdot \pi_{2,N}) \underset{N\to\infty}{\to} (\pi_1, \pi_2)$, leads after scaling to the drift of WF model with positive mutations rates (π_1, π_2).

Taking

$$p_N(x) = \frac{(1 + s_{1,N})x}{1 + s_{1,N}x + s_{2,N}(1 - x)}$$

where $s_{i,N} > 0$ are small $N-$dependent selection parameter satisfying $N \cdot s_{i,N} \underset{N\to\infty}{\to} \sigma_i > 0$, $i = 1, 2$, leads, after scaling, to the WF model with selective

drift $f(x) = \sigma x (1 - x)$, where $\sigma := \sigma_1 - \sigma_2$. Typically, the drift $f(x)$ is a large N approximation of the bias: $N(p_N(x) - x)$. The WF diffusion with selection is thus:

$$(19) \qquad dx_t = \sigma x_t (1 - x_t) dt + \sqrt{x_t (1 - x_t)} dw_t$$

where time is measured in units of N. Letting $\theta_t = Nt$ define a new time-scale with inverse $t_\theta = \theta/N$, the time-changed process $y_\theta = x_{\theta/N}$ now obeys the SDE

$$dy_\theta = s y_\theta (1 - y_\theta) d\theta + \sqrt{\frac{1}{N} y_\theta (1 - y_\theta)} dw_\theta,$$

with a small diffusion term. Here $s = s_1 - s_2$ and time θ is the usual time-clock.

The WF diffusion with selection (19) tends to drift to $\circ = 1$ (respectively $\circ = 0$) if allele A_1 is selectively advantageous over $A_2 : \sigma_1 > \sigma_2$ (respectively $\sigma_1 < \sigma_2$) in the following sense: if $\sigma > 0$ (respectively < 0), the fixation probability at $\circ = 1$, which is [16]

$$\mathbf{P}(\tau_{x,1} < \tau_{x,0}) = \frac{1 - e^{-2\sigma x}}{1 - e^{-2\sigma}},$$

increases (decreases) with σ taking larger (smaller) values.

The usual way to look at the WF diffusion with mutation and selection is to compose the two above mechanisms $p_N(x)$ corresponding to mutation and selection respectively. In the scaling limit, one obtains the standard WF diffusion model including mutations and selection as:

$$(20) \qquad dx_t = [(\pi_1 - (\pi_1 + \pi_2) x_t) + \sigma x_t (1 - x_t)] dt + \sqrt{x_t (1 - x_t)} dw_t.$$

4. The neutral WF model

In this Section, we particularize the general ideas developed in the introductory Section 2 to the neutral WF diffusion (18) and draw some straightforward conclusions most of which are known which illustrate the use of Doob transforms.

4.1. **Explicit solutions of the neutral KFE.** As shown by Kimura in ([15]), the Kolmogorov forward (and backward) equation is exactly solvable in this case, using spectral theory. The solutions involve a series expansion in terms of eigen-functions of the KB infinitesimal generator with discrete eigenvalues spectrum.

Let $\lambda_k = k(k+1)/2$, $k \geq 0$. There exist $u_k = u_k(x)$ and $v_k = v_k(y)$ solving the eigenvalue problem: $-G(u_k) = \lambda_k u_k$ and $-G^*(v_k) = \lambda_k v_k$. With $\langle v_k, u_k \rangle = \int_0^1 u_k(x) v_k(x) dx$, the transition probability density $p(x; t, y)$ of the

neutral WF models admits the spectral expansion

$$p(x; t, y) = \sum_{k \geq 1} b_k e^{-\lambda_k t} u_k(x) v_k(y) \text{ where } b_k = \frac{1}{\langle v_k, u_k \rangle}$$

where $u_k(x)$ are the Gegenbauer polynomials rescaled on $[0, 1]$ and normalized to have value 1 at $x = 0$. In particular, $u_0(x) = x$, $u_1(x) = x - x^2$, $u_2(x) = x - 3x^2 + 2x^3$, $u_3(x) = x - 6x^2 + 10x^3 - 5x^4$, $u_4(x) = x - 10x^2 + 30x^3 - 35x^4 + 14x^5$, ...

Next, $v_k(y) = m(y) u_k(y)$ where $m(y) = 1/(y(1-y))$ is the speed density of the neutral WF diffusion. For instance, $v_0(y) = \frac{1}{1-y}$, $v_1(y) = 1$, $v_2(y) = 1 - 2y$, $v_3(y) = 1 - 5y + 5y^2$, $v_4(y) = 1 - 9y + 21y^2 - 14y^3$,...

Although $\lambda_0 = 0$ really constitutes an eigenvalue, only $v_0(y)$ is not a polynomial and the spectral expansion of p should start at $k = 1$, expressing that p is a sub-probability. When $k \geq 1$, from their definition, the $u_k(x)$ polynomials satisfy $u_k(0) = u_k(1) = 0$ in such a way that $v_k(y) = m(y) \cdot u_k(y)$, $k \geq 1$ is a polynomial with degree $k - 1$.

The series expansion for $p(x; t, y)$ solves the KFE of the WF model.
We have $\mathbf{P}(\tau_x > t) = \int_0^1 \mathbf{P}(x_t \in dy)$ and so

$$\rho_t(x) := \mathbf{P}(\tau_x > t) = \sum_{k \geq 1} \frac{\int_0^1 v_k(y) \, dy}{\langle v_k, u_k \rangle} e^{-\lambda_k t} u_k(x)$$

is the exact tail distribution of the absorption time.
Since $v_1(y) = 1$, to the leading order in t, for large time

$$\mathbf{P}(x_t \in dy) = 6e^{-t} \cdot x(1-x) \, dy + \mathcal{O}(e^{-3t})$$

which is independent of y. Integrating over y, $\rho_t(x) := \mathbf{P}(\tau_x > t) \sim 6e^{-t} \cdot x(1-x)$ so that the conditional probability

(21) $$\mathbf{P}(x_t \in dy \mid \tau_x > t) \underset{t \to \infty}{\sim} dy$$

is asymptotically uniform in the Yaglom limit. As time passes by, given absorption did not occur in the past, $x_t \xrightarrow{d} x_\infty$ (as $t \to \infty$) which is a uniformly distributed random variable on $[0, 1]$.

4.2. **Additive functionals for the neutral WF and Doob transforms.**
Let $(x_t; t \geq 0)$ be the neutral WF diffusion model defined by (18) on the interval $I = [0, 1]$ where both endpoints are absorbing (exit). Consider the additive quantities

$$\alpha(x) = \mathbf{E}\left(\int_0^{\tau_x} c(x_s) \, ds + d(x_{\tau_x}) \right),$$

where functions c and d are both nonnegative. With $G = \frac{1}{2}x(1-x)\partial_x^2$, $\alpha(x)$ solves:

$$-G(\alpha) = c \text{ if } x \in I$$
$$\alpha = d \text{ if } x \in \partial I.$$

Therefore α is a super-harmonic function for G.

Take $c = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon}\mathbf{1}_{(y-\varepsilon,y+\varepsilon)}(x) =: \delta_y(x)$ and $d = 0$, when $y \in I$: in this case, $\alpha := \mathfrak{g}(x,y)$ is the Green function (the mean local time at y given the process started at x). The solution takes the simple form

$$\mathfrak{g}(x,y) = 2\frac{x}{y} \text{ if } x < y$$
$$\mathfrak{g}(x,y) = 2\frac{1-x}{1-y} \text{ if } x > y.$$

The Green function solves the above general problem of evaluating additive functionals $\alpha(x)$:

$$\alpha(x) = \int_I \mathfrak{g}(x,y)c(y)\,dy \text{ if } x \in I$$
$$\alpha = d \text{ if } x \in \partial I.$$

There are many interesting choices of c therefore leading to α, allowing to compute for example the mean time till absorption for the neutral WF diffusion, the probability to hit state 1 before 0... For each choice of α, it is interesting to study the transformed process $(\overline{x}_t; t \geq 0)$ whose transition probability is given by

$$\overline{p}(x;t,y) = \frac{\alpha(y)}{\alpha(x)}p(x;t,y),$$

in terms of the original process transition probability $p(x;t,y)$. This allows for example to understand the neutral WF process conditioned on exit at some boundary and to evaluate for this new process interesting average additive functionals such as the mean time needed to hit the exit boundary...For detailed similar examples arising in the context of WF diffusions and related ones, see [10].

5. THE WF MODEL WITH SELECTION

Now we briefly focus on the diffusion process (19). Let $(v_k(y))_{k \geq 1}$ be the Gegenbauer eigen-polynomials of the KF operator corresponding to the neutral WF diffusion (18), so with eigenvalues $\lambda_k = k(k+1)/2$, $k \geq 1$. Define the oblate spheroidal wave functions on $[0,1]$ as

(22) $$w_k^\sigma(y) = \sum_{l \geq 1}{}' f_k^l v_l(y),$$

where f_k^l obey the three-term recurrence defined in [20]. In the latter equality, the l summation is over odd (even) values if k is even (odd).

Define $v_k^\sigma (y) = e^{\sigma y} w_k^\sigma (y)$ and $u_k^\sigma (x) = \frac{1}{m(x)} v_k^\sigma (x)$ where $m(x) = e^{2\sigma x} / (x(1-x))$ is the speed measure density of the WF model with selection (19).

The system $(u_k^\sigma (x), v_k^\sigma (x))_{k \geq 1}$ constitute a system of eigen-functions for the WF with selection generators $-G$ and $-G^*$ with eigenvalues λ_k^σ implicitly defined in [20], thus with $-G(u_k^\sigma) = \lambda_k^\sigma u_k^\sigma$ and $-G^*(v_k^\sigma) = \lambda_k^\sigma v_k^\sigma$. The eigen-function expansion of the transition probability density of the WF model with selection is thus, [14]:

$$(23) \qquad p(x; t, y) = \sum_{k \geq 1} b_k^\sigma e^{-\lambda_k^\sigma t} u_k^\sigma (x) v_k^\sigma (y)$$

where $b_k^\sigma = \langle v_k^\sigma, u_k^\sigma \rangle^{-1}$. The WF model with selection can be viewed as a perturbation problem of the neutral WF model (see [21]). There exist perturbation developments of λ_k^σ around λ_k with respect to σ^2, [14]. They are valid and useful for small σ.

The WF diffusion process x_t with selection (19) is non-conservative, with finite hitting time τ_x of one of the boundaries. Following the general arguments developed in Section 2, the Yaglom limit of x_t conditioned on $\tau_x > t$ is the normalized version of

$$(24) \qquad v_1^\sigma (y) = e^{\sigma y} w_1^\sigma (y).$$

The limit law of x_t conditioned on never hitting the boundaries in the remote future is the normalized version of

$$(25) \qquad u_1^\sigma (y) v_1^\sigma (y) = \frac{1}{m(y)} v_1^\sigma (x)^2 = y(1-y) w_1^\sigma (y)^2.$$

Because the latter conditioning is more stringent than the former, the probability mass of (25) is more concentrated inside the interval than (24). Compare with the statements at the end of Section 2 concerning the neutral WF diffusion.

6. Doob transform of the neutral WF model: sub-critical BD

In this Section, we define the branching WF diffusion model with selection while applying a Doob transform to the neutral WF model, based on the sub-harmonic additive functional $\alpha(x) = e^{\sigma x}$, say with $\sigma > 0$. We then study in detail the obtained branching process.

The starting point is thus the neutral WF diffusion: $dx_t = \sqrt{x_t (1 - x_t)} dw_t$, $x_0 = x \in (0, 1)$.

For this model, $G = \frac{1}{2} x(1-x) \partial_x^2$ and both boundaries are exit. With $\lambda_k = k(k+1)/2$, $k \geq 0$, its transition density $p(x; t, y)$ admits the spectral

representation

(26) $$p(x; t, y) = \sum_{k \geq 1} b_k e^{-\lambda_k t} u_k(x) v_k(y),$$

in terms of the Gegenbauer eigen-polynomials (see Subsection 4.1). We shall consider the following transformation of paths on the neutral WF model: Let $\alpha(x) = e^{\sigma x}$, $\sigma > 0$ and consider $\overline{G}(\cdot) = \alpha^{-1} G(\alpha \cdot) = \widetilde{G}(\cdot) + b(x) \cdot$. We now have $G(\alpha) = \frac{1}{2}\sigma^2 x(1-x) e^{\sigma x}$ and so $b(x) = G(\alpha)/\alpha = \frac{\sigma^2}{2} x(1-x) \geq 0$.

Note that $-G(\alpha) \leq 0$ indicating that α is sub-harmonic for G.

In this case study, one selects sample paths of $(x_t; t \geq 0)$ with large $\alpha(y)$ and we claim that this is an alternative interesting way to introduce selection in the neutral WF diffusion process.

The dynamics of $(\widetilde{x}_t; t \geq 0)$ governed by $\widetilde{G}$ is easily seen to be the standard WF with selection dynamics (19)

$$d\widetilde{x}_t = \sigma \widetilde{x}_t(1 - \widetilde{x}_t)\,dt + \sqrt{\widetilde{x}_t(1 - \widetilde{x}_t)}dw_t,$$

subject to additional quadratic branching at rate $b(x) = \frac{1}{2}\sigma^2 x(1-x)$ inside I. We indeed have

$$\overline{G}(\cdot) = e^{-\sigma x} G(e^{\sigma x} \cdot) = b(x) \cdot + \widetilde{G}(\cdot),$$

where

$$\widetilde{G} =: \widetilde{f}\partial_x + \frac{1}{2}\widetilde{g}^2\partial_x^2 = \sigma x(1-x)\partial_x + \frac{1}{2}x(1-x)\partial_x^2$$

is the KBE operator of the dynamics $(\widetilde{x}_t; t \geq 0)$. Recall that $\widetilde{x}_t$ is transient and so hits one of the boundaries $\{0, 1\}$ in finite time $\widetilde{\tau}_x$.

To summarize, in our branching diffusion way to look at selection, we move from the neutral WF diffusion $(x_t; t \geq 0)$ to the standard WF diffusion with selection $(\widetilde{x}_t; t \geq 0)$ but subject to additional branching at rate $b(x)$.

Remark. With $\beta(x) := \alpha(x)^{-1} = e^{-\sigma x}$, we clearly have

$$\overline{G}(\beta(x)) = 0$$

and β is an harmonic function for $\overline{G}$ and as a result, Doob-transforming $\overline{G}$ by β, we get

$$\beta^{-1}\overline{G}(\beta \cdot) = (\alpha\beta)^{-1}\overline{G}(\alpha\beta \cdot) = G(\cdot)$$

which is the infinitesimal generator of the original neutral WF martingale. $\diamond$

The birth (creating) rate $b \geq 0$ in $\overline{G}$ is bounded from above on $(0, 1)$. It may be put into the canonical form $b(x) = b_*(\mu(x) - 1)$ where $b_* = \max_{x \in [0,1]} (b(x)) =$

$\frac{\sigma^2}{8} > 0$ and

$$(27) \qquad\qquad \mu(x) = 1 + 4x(1 - x),$$

whose range is the interval $[1, 2]$ as $x \in [0, 1]$.

The density of the transformed process is $\overline{p}(x; t, y) = \frac{\alpha(y)}{\alpha(x)} p(x; t, y)$. It is exactly known because so is p is from (26).

The transformed process (with infinitesimal backward generator $\overline{G}$) accounts for a branching diffusion (BD) where a diffusing mother particle (with generator $\widetilde{G}$ and started at x) lives a random exponential time with constant rate b_*. When the mother particle dies, it gives birth to a spatially dependent random number $M(x)$ of particles (with mean $\mu(x)$). $M(x)$ independent daughter particles are started where their mother particle died; they move along a WF diffusion with selection and reproduce, independently and so on.

Because $\mu(x)$ is bounded above by 2 and larger than 1 (indicating a supercritical branching process), we actually get a BD with binary scission whose random offspring number satisfies ('w.p.' meaning 'with probability')

$$M(x) = 0 \text{ w.p. } p_0(x) = 0$$

$$M(x) = 1 \text{ w.p. } p_1(x) = 2 - \mu(x)$$

$$M(x) = 2 \text{ w.p. } p_2(x) = \mu(x) - 1,$$

with $p_2(x) \geq p_1(x)$ (the event that 2 particles are generated in a splitting event is more probable than a single one).

For such a transformed process, the trade-off is as follows: there is a competition between the boundaries $\{0, 1\}$ which are absorbing for the particle system and the number of particles $N_t(x)$ in the system at each time t, which may grow due to binary branching events (or remain steady when $M(x) = 1$).

The density $\overline{p}$ of the transformed process has the following interpretation

$$(28) \qquad\qquad \overline{p}(x; t, y) = \mathbf{E}\left[\sum_{n=1}^{N_t(x)} p^{(n)}(x; t, y)\right],$$

where $p^{(n)}(x; t, y)$ is the density at (t, y) of the nth alive particle descending from the ancestral one (Eve), started at x. In the latter formula, the sum vanishes if $N_t(x) = 0$. A particle is alive at time t if it came to birth before t and has not been yet absorbed by the boundaries.

Let $\overline{\rho}_t(x) = \int_{(0,1)} \overline{p}(x; t, y)\, dy$. Then $\overline{\rho}_t(x)$ is the expected number of particle alive at time t. We have

$$\partial_t \overline{\rho}_t(x) = \overline{G}(\overline{\rho}_t(x)), \ \overline{\rho}_0(x) = \mathbf{1}\,(x \in (0, 1)).$$

Remark. From the Feynman-Kac formula, $\overline{p}$ in (28) is also

$$\overline{p}\left(x; t, y\right) = \mathbf{E}_x\left(e^{\int_0^{t \wedge \widetilde{\tau}_x} b(\widetilde{x}_s)ds} \mid \widetilde{x}_t = y\right) p\left(x; t, y\right)$$

and

$$\overline{\rho}_t\left(x\right) = \mathbf{E}_x\left(e^{\int_0^{t \wedge \widetilde{\tau}_x} b(\widetilde{x}_s)ds}\right). \quad \diamond$$

But then $\overline{q}\left(x; t, y\right) := \overline{p}\left(x; t, y\right)/\overline{\rho}_t\left(x\right)$ obeys the forward PDE

$$\partial_t \overline{q}\left(x; t, y\right) = \left(-\frac{\partial_t \overline{\rho}_t\left(x\right)}{\overline{\rho}_t\left(x\right)} + b\left(y\right)\right) \overline{q}\left(x; t, y\right) + \widetilde{G}^*\left(\overline{q}\left(x; t, y\right)\right)$$

as a result of $\partial_t \overline{p}\left(x; t, y\right) = \overline{G}^*\left(\overline{p}\left(x; t, y\right)\right)$. We have

(29)
$$\overline{q}\left(x; t, y\right) = \frac{\mathbf{E}\left[\sum_{n=1}^{N_t(x)} p^{(n)}\left(x; t, y\right)\right]}{\mathbf{E}\left[N_t\left(x\right)\right]}$$

showing that $\overline{q}\left(x; t, y\right)$ is the average presence density at (t, y) of the system of particles all descending from Eve started at x.

Clearly $-\frac{\log \overline{\rho}_t(x)}{t} \underset{t \to \infty}{\to} \lambda_1 = 1$ (and therefore also $-\frac{\partial_t \overline{\rho}_t(x)}{\overline{\rho}_t(x)}$ by L' Hospital rule), because

$$\overline{\rho}_t\left(x\right) = \frac{1}{\alpha\left(x\right)} \sum_{k \geq 1} b_k e^{-\lambda_k t} u_k\left(x\right) \int_0^1 \alpha\left(y\right) v_k\left(y\right) dy.$$

The expected number of particles in the system decays globally and exponentially at rate λ_1.

The BD transformed process therefore admits an integrable Yaglom limit $\overline{q}_\infty$, solution to $-\widetilde{G}^*\left(\overline{q}_\infty\right) = \left(\lambda_1 + b\left(y\right)\right) \overline{q}_\infty$ or $-\overline{G}^*\left(\overline{q}_\infty\right) = \lambda_1 \overline{q}_\infty$. With $v_1\left(y\right) = 1$, the first eigenvector of $-G^*$ associated to the smallest positive eigenvalue $\lambda_1 = 1$, $\overline{q}_\infty$ is of the product form

(30)
$$\overline{q}_\infty\left(y\right) = C_* e^{\sigma y} v_1\left(y\right) = \frac{\sigma e^{\sigma y}}{e^\sigma - 1}.$$

The arbitrary multiplicative constant C_* was chosen in such a way that $\overline{q}_\infty\left(y\right)$ is a probability.

By analogy with the Yaglom construction, this limiting probability $\overline{q}_\infty$ can be called the quasi-stationary Yaglom average density at (t, y) for the BD particle system (it is also the ground state for $\overline{G}^*$).

There is also a natural eigenvector $\overline{\phi}_\infty$ of the backward operator $-\overline{G}$, satisfying $-\overline{G}\left(\overline{\phi}_\infty\right) = \lambda_1 \overline{\phi}_\infty$ (the ground state for $\overline{G}$). It is explicitly here

(31)
$$\overline{\phi}_\infty\left(x\right) = \frac{C}{\alpha\left(x\right)} u_1\left(x\right) = \frac{6\left(e^\sigma - 1\right)}{\sigma} e^{-\sigma x} x\left(1 - x\right).$$

The arbitrary multiplicative constant $C = 6/C_*$ was chosen in such a way that $\int_0^1 \overline{q}_\infty\left(y\right) \overline{\phi}_\infty\left(y\right) dy = 1$. Note that the spectral structures of both $\overline{G}^*$ and $\overline{G}$

are easily obtainable from the ones of G^* and G thanks to the Doob transform structure.

In the terminology of [22], both operators $\overline{G}\left(\cdot\right)+\lambda_1\cdot$ and its adjoint are critical [2]. In this context, the constant λ_1 is called the generalized principal eigenvalue. The eigen-functions $\left(\overline{\phi}_\infty, \overline{q}_\infty\right)$ are their associated ground states. We note that we have the L^1-product property (See [22], Subsection 4.9).

$$\int_0^1 \overline{\phi}_\infty\left(x\right)\overline{q}_\infty\left(x\right)dx = 6\int_0^1 u_1\left(x\right)v_1\left(x\right)dx = 1 < \infty.$$

Remark. Using the Feynman-Kac representation of $\overline{\rho}_t\left(x\right)$, we get

$$-\frac{1}{t}\log\mathbf{E}_x\left(e^{\int_0^{t\wedge\widetilde{\tau}_x}b(\widetilde{x}_s)ds}\right)\underset{t\to\infty}{\to}\lambda_1 = 1 \text{ and}$$

$$e^{\lambda_1 t}\mathbf{E}_x\left(e^{\int_0^{t\wedge\widetilde{\tau}_x}b(\widetilde{x}_s)ds}\right)\underset{t\to\infty}{\to}\overline{\phi}_\infty\left(x\right). \quad \diamond$$

With $p_m\left(x\right) = \mathbf{P}\left(M\left(x\right) = m\right)$, let

$$l\left(x\right) = \sum_{m\geq 1}p_m\left(x\right)m\log m = 2\log 2p_2\left(x\right).$$

We have the $x\log x$ condition:

$$(32) \quad \int_0^1 l\left(x\right)\overline{\phi}_\infty\left(x\right)\overline{q}_\infty\left(x\right)dx = 48\log 2\int_0^1 x\left(1-x\right)u_1\left(x\right)v_1\left(x\right)dx < \infty.$$

We conclude (following [1] and [2]) that, as a result of the condition (32) being trivially satisfied, global extinction holds in the following sense:

(i) $\mathbf{P}\left(N_t\left(x\right) = 0\right)\underset{t\to\infty}{\to}1$, uniformly in x.

(ii) there exists a constant $\gamma > 0$: $e^{\lambda_1 t}\left[1 - \mathbf{P}\left(N_t\left(x\right) = 0\right)\right]\underset{t\to\infty}{\to}\gamma\overline{\phi}_\infty\left(x\right)$, uniformly in x.

(iii) For all bounded measurable function ψ on I :

$$\mathbf{E}\left[\sum_{n=1}^{N_t(x)}\psi\left(\widetilde{x}_t^{(n)}\right) \mid N_t\left(x\right) > 0\right]\underset{t\to\infty}{\to}\gamma^{-1}\int_{(0,1)}\psi\left(y\right)\overline{q}_\infty\left(y\right)dy.$$

---

[2]$\overline{G}\left(\cdot\right)+\lambda_1\cdot$ $\left(\overline{G}^*\left(\cdot\right)+\lambda_1\cdot\right)$ is said to be critical if there exists some function $\overline{\phi}_\infty \in C^2$ (respectively $\overline{q}_\infty \in C^2$), strictly positive in $\left(0,1\right)$, such that: $\overline{G}\left(\overline{\phi}_\infty\right)+\lambda_1\overline{\phi}_\infty = 0$ (respectively $\overline{G}^*\left(\overline{q}_\infty\right)+\lambda_1\overline{q}_\infty = 0$) and the operators do not possess a minimal positive Green function.

From (i), it is clear that the process gets ultimately extinct with probability 1. In the trade-off between pure branching and absorption at the boundaries, all particles get eventually absorbed and the global BD process turns out be sub-critical (even though $\mu(x) = \mathbf{E}M(x) > 1$ for all $x \in (0,1)$): Probability mass escapes out of I although the BD survives with positive probability.

In the statement (ii), the quantity $1-\mathbf{P}(N_t(x) = 0) = \mathbf{P}(N_t(x) > 0)$ is also $\mathbf{P}(T(x) > t)$ where $T(x)$ is the global extinction time of the particle system descending from an Eve particle started at x. The number $-\lambda_1$ is the usual Malthus exponential decay rate parameter. From (ii), $\overline{\phi}_\infty(x)$ has a natural interpretation in terms of the propensity of the particle system to survive to its extinction fate: the so-called reproductive value in demography.

(iii) with $\psi = 1$ reads $\mathbf{E}[N_t(x) \mid N_t(x) > 0] \underset{t\to\infty}{\to} \gamma^{-1}$ giving an interpretation of the constant γ (which may be hard to evaluate in practise).

The ground states of $\overline{G} + \lambda_1$ and its adjoint are thus $(\overline{\phi}_\infty, \overline{q}_\infty)$ and explicit here. It is useful to consider the process whose infinitesimal generator is given by the Doob-transform

$$\overline{\phi}_\infty^{-1}(\overline{G} + \lambda_1)(\overline{\phi}_\infty \cdot) = \overline{\phi}_\infty^{-1}\left(\widetilde{G} + b + \lambda_1\right)(\overline{\phi}_\infty \cdot),$$

because product-criticality is preserved under this transformation. The ground states associated to this new operator and its dual are $(1, \overline{\phi}_\infty \overline{q}_\infty)$. Developing, we obtain a process whose infinitesimal generator is

$$\widetilde{G} + \frac{\overline{\phi}_\infty'}{\overline{\phi}_\infty}g^2\partial_x = G + \frac{u_1'}{u_1}g^2\partial_x,$$

with no multiplicative part. In our case study, we get $\frac{1}{2}x(1-x)\partial_x^2 + (1-2x)\partial_x$ adding a stabilizing drift towards $1/2$ to the original neutral WF model. The associated diffusion process is positive recurrent and so its invariant measure $\overline{\phi}_\infty \overline{q}_\infty = 6u_1v_1 = 6y(1-y)$ is integrable with mass 1. It is the beta$(2,2)$ limit law of the $Q-$process (see (17) and the comments at the end of Section 2 relative to the neutral WF diffusion).

Remarks.

(i) At time t, let $\left(\widetilde{x}_t^{(n)}\right)_{n=1}^{N_t(x)}$ denote the positions of the BD particle system. Let $u(x,t;z) = \mathbf{E}\left[\prod_{n=1}^{N_t(x)} z^{\psi\left(\widetilde{x}_t^{(n)}\right)}\right]$ stand for the functional generating function ($|z| \leq 1$) of the measure-valued branching particle system. $u(x,t;z)$ obeys the nonlinear (quadratic) Kolmogorov-Petrovsky-Piscounoff PDE, [17]:

$$\partial_t u(x,t;z) = b_*\theta(x, u(x,t;z)) + \widetilde{G}(u(x,t;z)); \; u(x,0;z) = z^{\psi(x)},$$

where $\theta\left(x,z\right)=\mathbf{E}\left[z^{M\left(x\right)}\right]-z=\left(p_2\left(x\right)z^2+p_1\left(x\right)z\right)-z$ or

$$\theta\left(x,z\right)=4x\left(1-x\right)z\left(z-1\right)$$

is the shifted probability generating function of the branching law of $M\left(x\right)$. Thus, the nonlinear part reads $b_*\theta\left(x,u\left(x,t;z\right)\right)=b\left(x\right)u\left(x,t;z\right)\left(u\left(x,t;z\right)-1\right)$ which is quadratic in u.

In particular, if $u\left(x,t\right):=\partial_z u\left(x,t;z\right)_{z=1}=\mathbf{E}\left[\sum_{n=1}^{N_t\left(x\right)}\psi\left(\widetilde{x}_t^{\left(n\right)}\right)\right]$, $u\left(x,t\right)$ obeys the linear backward PDE

$$\partial_t u\left(x,t\right)=b\left(x\right)u\left(x,t\right)+\widetilde{G}\left(u\left(x,t\right)\right);\ \ u\left(x,0\right)=\psi\left(x\right)$$

involving $\overline{G}\left(\cdot\right)=\widetilde{G}\left(\cdot\right)+b\left(x\right)\cdot$. We have the Feynman-Kac interpretation

$$u\left(x,t\right)=\mathbf{E}_x\left(e^{\int_0^{t\wedge\widetilde{\tau}_x}b\left(\widetilde{x}_s\right)ds}\psi\left(\widetilde{x}_t\right)\right).$$

The latter evolution equation is the backward version of the forward PDE giving the evolution of $\overline{p}\left(x;t,y\right)$ as $\partial_t\overline{p}\left(x;t,y\right)=\overline{G}^*\left(\overline{p}\left(x;t,y\right)\right)$, $\overline{p}\left(x;0,y\right)=\delta_x\left(y\right)$.

$\left(ii\right)$ Let us look at the branching diffusion process governed by $\overline{G}$ would time be measured using the time substitution

$$\theta_t=\int_0^t g^2\left(\widetilde{x}_s\right)ds=\int_0^t\widetilde{x}_s\left(1-\widetilde{x}_s\right)ds$$

for each of the particles that came to birth before t.

Then $\overline{G}\to\overline{\mathcal{G}}:=\frac{1}{x\left(1-x\right)}\overline{G}=\sigma\partial_x+\frac{1}{2}\partial_x^2+\frac{1}{2}\sigma^2\cdot$. In particular, each motion $y_\theta=\widetilde{x}_{t_\theta}$ is a Brownian motion with constant drift (a Gaussian process). This new $\overline{\mathcal{G}}$ is the one of absorbing Brownian motion with drift σ on $[0,1]$, including branching at constant rate $\frac{1}{2}\sigma^2$. The Sturm-Liouville problem for $\overline{\mathcal{G}}$ admits the eigenvalues $\lambda_k=\frac{k^2+\sigma^2}{2}$, $k\geq1$ with eigen-states $u_k\left(x\right)\propto e^{-\sigma x}\sin\left(k\pi x\right)$ and $v_k\left(y\right)\propto e^{\sigma y}\sin\left(k\pi y\right)$. The spectral gap is $\lambda_1=\frac{1+\sigma^2}{2}>0$ and the time-changed branching diffusion also becomes eventually extinct, sub-critically: The time substitution changes the spectral structure of the model but not its qualitative features. $\diamond$

7. Doob transform of the WF model with mutations: critical BD

In this Section, we start from the WF model with mutations. Using the same Doob transform based on the additive functional $\alpha\left(x\right)=e^{\sigma x}$ to introduce selection, we end up with a WF diffusion process with killing and branching describing the effect of selection on the WF model in the presence of mutations. We show that in this setup, the resulting branching diffusion process is no longer sub-critical; rather, it turns out to be critical.

Suppose the starting point model is now the WF diffusion with mutations:

$$dx_t = (\pi_1 - \pi x_t)\, dt + \sqrt{x_t(1 - x_t)} dw_t, x_0 = x \in (0,1),$$

with $\pi := \pi_1 + \pi_2$. For this model, $G = (\pi_1 - \pi x)\,\partial_x + \frac{1}{2}x(1-x)\,\partial_x^2$ and both boundaries are chosen as being entrance (reflecting)[3]. The WF diffusion process with mutations is now ergodic. With

$$\lambda_k = \frac{k(k-1+\pi)}{2},\ k \geq 0,$$

its transition density $p(x; t, y)$ now admits the discrete spectral representation

(33)
$$p(x; t, y) = \sum_{k \geq 0} b_k e^{-\lambda_k t} u_k(x)\, v_k(y).$$

Here, $u_k(x)$ are the Jacobi polynomials rescaled on $[0,1]$ and normalized to have value 1 at $x = 0$. In particular, $u_0(x) = 1$, $u_1(x) = 1 - \frac{\pi}{\pi_2}x$, $u_2(x) = 1 - \frac{2(1+\pi)}{\pi_2}x + \frac{(1+\pi)(2+\pi)}{\pi_2(1+\pi_2)}x^2$,... Next, $v_k(y) = m(y)\, u_k(y)$ where

$$m(y) = \frac{\Gamma(2\pi)}{\Gamma(2\pi_1)\,\Gamma(2\pi_2)} y^{2\pi_1 - 1}(1 - y)^{2\pi_2 - 1}$$

is the speed density of the ergodic WF diffusion with mutations (its normalized invariant measure). Note that the $k = 0$ term in (33) is precisely $m(y)$ as required. Because the transition probability density of the WF diffusion with mutations has also a discrete spectral representation, this model is amenable to a similar analysis than the neutral WF diffusion.

Proceeding as for the neutral case, we shall consider the following transformation of paths for the WF model with mutations: Let $\alpha(x) = e^{\sigma x}$ and consider a transformed process with infinitesimal generator $\overline{G}(\cdot) = \alpha^{-1}G(\alpha \cdot)$. The multiplicative part of $\overline{G}$ is now

$$\lambda(x) = G(\alpha)/\alpha = \sigma(\pi_1 - \pi x) + \frac{\sigma^2}{2}x(1 - x).$$

Note that now α is neither sub-harmonic nor super-harmonic for the infinitesimal generator G including mutations because the sign of $\lambda(x)$ varies as x varies.

In this case study, one selects sample paths of the WF diffusion model with mutations $(x_t; t \geq 0)$ with large terminal values of $\alpha(y)$. The dynamics of

---

[3]When both the mutation rates u_1 and u_2 are greater than $1/2$, the boundaries are entrance. When either u_1 or u_2 is smaller than $1/2$ the corresponding boundary is regular and one needs to specify whether it is reflecting or absorbing or a mixture of the two. We force here the regular boundaries to be entrance.

$(\widetilde{x}_t; t \geq 0)$ is easily seen to be the WF with mutation and selection dynamics of the type (20)

$$d\widetilde{x}_t = [(\pi_1 - \pi\widetilde{x}_t) + \sigma\widetilde{x}_t (1 - \widetilde{x}_t)] \, dt + \sqrt{\widetilde{x}_t (1 - \widetilde{x}_t)} dw_t,$$

subject to additional quadratic killing and branching at rate $\lambda(x)$ inside I. We indeed have

$$\overline{G}(\cdot) = e^{-\sigma x} G(e^{\sigma x} \cdot) = \lambda(x) \cdot + \widetilde{G}(\cdot),$$

where $\widetilde{G} = [(\pi_1 - \pi x) + \sigma x (1 - x)] \partial_x + \frac{1}{2} x (1 - x) \partial_x^2$ is the KBE operator of the dynamics $(\widetilde{x}_t; t \geq 0)$.

To summarize, in our branching diffusion way to look at the action of selection, we move from the WF diffusion with mutations $(x_t; t \geq 0)$ to the standard WF diffusion with mutation and selection $(\widetilde{x}_t; t \geq 0)$, but subject to additional killing/branching at rate $\lambda(x)$.

Remark. With $\beta = \alpha^{-1} = e^{-\sigma x}$, again $\overline{G}(\beta) = 0$ and $\beta^{-1}\overline{G}(\beta \cdot) = G(\cdot)$ is the infinitesimal generator of the original WF model, now with mutations. $\diamond$

The birth (creating) and death (annihilating) rate λ in $\overline{G}$ is bounded from above and below on $(0, 1)$. It may now be put into the canonical form $\lambda(x) = \lambda_* (\mu(x) - 1)$ where $\lambda_* = \max_{x \in [0,1]} (|\lambda(x)|)$ and

(34) $$\mu(x) = 1 + \frac{\lambda(x)}{\lambda_*}$$

whose range belongs to the interval $[0, 2]$ as $x \in [0, 1]$.

Note that when $\pi > \sigma/2$, $\lambda_* = \sigma(\pi_1 \vee \pi_2)$ whereas when $\pi < \sigma/2$, $\lambda_* = \lambda(x_*) \vee \sigma\pi_2$ where $x_* = 1/2 - \pi/\sigma > 0$.

The density of the transformed process is $\overline{p}(x; t, y) = \frac{\alpha(y)}{\alpha(x)} p(x; t, y)$. It is exactly known because p is known from (33).

The transformed process (with infinitesimal backward generator $\overline{G}$) accounts for a branching diffusion (BD) where a diffusing mother particle (with generator $\widetilde{G}$ and started at x) lives a random exponential time with constant rate λ_*. When the mother particle dies, it gives birth to a spatially dependent random number $M(x)$ of particles (with mean $\mu(x)$). If $M(x) \neq 0$, $M(x)$ independent daughter particles are started where their mother particle died; they move along a WF diffusion with mutation and selection (with generator $\widetilde{G}$) and reproduce independently, and so on.

Because $\mu(x)$ is bounded above by 2 and larger than 0, we actually get a BD with binary scission whose random offspring number satisfies

$$M(x) = 0 \text{ w.p. } p_0(x) = 1 - \mu(x)/2$$

$$M(x) = 1 \text{ w.p. } p_1(x) = 0$$

$$M(x) = 2 \text{ w.p. } p_2(x) = \mu(x)/2.$$

Note that

$$\lambda(x) = \lambda_* (p_2(x) - p_0(x)) =: b(x) - d(x)$$

identifying the birth and death components of the full multiplicative rate $\lambda(x)$.

For such a transformed process, the trade-off is of a different nature: there is a competition between the boundaries $\{0, 1\}$ which are now reflecting for the system of particles and the number of particles $N_t(x)$ in the system at each time t, which may grow or diminish due either to branching or killing events. In the presence of mutations, the particles are no longer killed once they hit the boundaries, suggesting that there should be a greater amount of them alive in the system. However, in this new model, there is an opportunity to kill the particles inside the definition domain, when they branch. The question now being: does the new trade-off result in global extinction or global explosion of the particle system? We will now show that critical global extinction occurs.

The density $\overline{p}$ of the transformed process again has the interpretation (28), where $p^{(n)}(x; t, y)$ is the density at (t, y) of the nth alive particle descending from the ancestral one (Eve), started at x. In the latter formula, the sum vanishes if $N_t(x) = 0$. A particle is alive at time t if it came to birth before t and has not yet been killed by a killing event.

Let $\overline{\rho}_t(x) = \int_{(0,1)} \overline{p}(x; t, y)\, dy$. Then $\overline{\rho}_t(x)$ is the expected number of particle alive at time t. We have

$$\partial_t \overline{\rho}_t(x) = \overline{G}(\overline{\rho}_t(x)), \ \overline{\rho}_0(x) = \mathbf{1}\, (x \in (0, 1)).$$

But then $\overline{q}(x; t, y) := \overline{p}(x; t, y)/\overline{\rho}_t(x)$ obeys the forward PDE

$$\partial_t \overline{q}(x; t, y) = \left(-\frac{\partial_t \overline{\rho}_t(x)}{\overline{\rho}_t(x)} + b(y) \right) \overline{q}(x; t, y) + \widetilde{G}^* (\overline{q}(x; t, y))$$

as a result of $\partial_t \overline{p}(x; t, y) = \overline{G}^* (\overline{p}(x; t, y))$. We again have (29), with $\overline{q}(x; t, y)$ the average presence density at (t, y) of the system of particles all descending from Eve started at x.

Clearly $-\frac{\log \overline{\rho}_t(x)}{t} \underset{t \to \infty}{\to} \lambda_0 = 0$ (and therefore also $-\frac{\partial_t \overline{\rho}_t(x)}{\overline{\rho}_t(x)}$), because

$$\overline{\rho}_t(x) = \frac{1}{\alpha(x)} \sum_{k \geq 0} b_k e^{-\lambda_k t} u_k(x) \int_0^1 \alpha(y) v_k(y)\, dy.$$

The expected number of particles in the system decays globally at rate λ_1 towards the non-zero limiting value

$$\overline{\rho}_\infty (x) := \alpha (x)^{-1} b_0 u_0 (x) \int_0^1 \alpha (y) v_0 (y) \, dy = e^{-\sigma x} \int_0^1 e^{\sigma y} m (y) \, dy.$$

The BD transformed process therefore admits an integrable Yaglom limit $\overline{q}_\infty$, solution to $-\widetilde{G}^* (\overline{q}_\infty) = \lambda (y) \overline{q}_\infty$ or $-\overline{G}^* (\overline{q}_\infty) = 0$. With $v_0 (y) = m (y)$, the first eigenvector of $-G^*$ associated to the smallest positive eigenvalue $\lambda_0 = 0$ (the equilibrium density of the WF diffusion with mutations), $\overline{q}_\infty$ is of the product form

$$(35) \qquad \overline{q}_\infty (y) = \frac{e^{\sigma y} m (y)}{\int_0^1 e^{\sigma y} m (y) \, dy}.$$

This explicit limiting probability $\overline{q}_\infty$ is the Yaglom limiting average presence density at (t, y) for the BD system of particles (it is also the ground state for $\overline{G}^*$).

There is also a natural eigenvector $\overline{\phi}_\infty$ of the backward operator $-\overline{G}$, satisfying $-\overline{G} (\overline{\phi}_\infty) = 0$ (the ground state for $\overline{G}$). It is explicitly here

$$(36) \qquad \overline{\phi}_\infty (x) = \frac{1}{\alpha (x)} u_0 (x) \int_0^1 e^{\sigma y} m (y) \, dy = e^{-\sigma x} \int_0^1 e^{\sigma y} m (y) \, dy.$$

Both operators $\overline{G} (\cdot)$ and its adjoint are again critical. The constant $\lambda_0 = 0$ is the new generalized principal eigenvalue; The eigen-functions $(\overline{\phi}_\infty, \overline{q}_\infty)$ are the new associated ground states.

We note that we have the L^1-product property

$$\int_0^1 u_0 (x) v_0 (x) \, dx = \int_0^1 \overline{\phi}_\infty (x) \overline{q}_\infty (x) \, dx = 1 < \infty.$$

Clearly the ground states of $-\overline{G}^*$ and $-\overline{G}$ are defined up to arbitrary multiplicative constants. Note that we chose these constants in such a way that $\int_0^1 \overline{q}_\infty (y) \, dy = 1$ and $\int_0^1 \overline{\phi}_\infty (x) \overline{q}_\infty (x) \, dx = 1$.

With $p_m (x) = \mathbf{P} (M (x) = m)$, let

$$\kappa (x) = \sum_{m \geq 2} m (m - 1) p_m (x) = 2 p_2 (x).$$

Because $p_2 (x)$ is a degree two polynomial in x, we have the condition:

$$(37) \qquad \int_0^1 \kappa (x) \overline{\phi}_\infty (x) \overline{q}_\infty (x) \, dx = 2 \int_0^1 p_2 (x) u_0 (x) v_0 (x) \, dx < \infty.$$

We conclude (following [1] and [2]) that, as a result of the condition (37) being trivially satisfied, global extinction holds critically, in the following sense:

(i) $\mathbf{P} (N_t (x) = 0) \underset{t \to \infty}{\to} 1$, uniformly in x.

(ii) Let $\mu_t = \sum_{n=1}^{N_t(\cdot)} \delta_{x_t^{(n)}}$, with $\mu_t(\psi) = \sum_{n=1}^{N_t(\cdot)} \psi\left(x_t^{(n)}\right)$.
There exists a finite positive constant :

$$\mu = \frac{1}{2t} \int_0^1 \mathbf{E}_x \left[\mu_t(\phi)^2 - \mu_t(\phi^2)\right] \overline{q}_\infty(x)\, dx = \frac{1}{2t}\mathbf{E}_{\overline{q}_\infty}\left[\mu_t(\phi)^2 - \mu_t(\phi^2)\right]$$

such that $t\left[1 - \mathbf{P}\left(N_t(x) = 0\right)\right] \underset{t\to\infty}{\to} \mu^{-1}\overline{\phi}_\infty(x)$, uniformly in x.

(iii) For all bounded measurable function ψ on I :

$$\frac{1}{t}\mathbf{E}\left[\sum_{n=1}^{N_t(x)} \psi\left(\widetilde{x}_t^{(n)}\right) \mid N_t(x) > 0\right] \underset{t\to\infty}{\to} \mu \int_{(0,1)} \psi(y)\overline{q}_\infty(y)\, dy.$$

From (i), it is clear that the process gets ultimately extinct with probability 1. In the trade-off between killing-branching and reflection at the boundaries, all particles get eventually absorbed but the global BD process turns out be critical. Thus, the killing part of $\lambda(x)$ is strong enough to avoid the explosion of the number of particles inside the unit interval, resulting in an overall critical process where global extinction still holds.

In the statement (ii), $1 - \mathbf{P}\left(N_t(x) = 0\right) = \mathbf{P}\left(T(x) > t\right)$ where $T(x)$ is the global extinction time of the particle system. The Pareto tails of $T(x)$ decay like t^{-1}, thus algebraically slowly: the time till extinction in this critical model is much longer than in the previous neutral sub-critical case (with exponential tails). From (ii), $\overline{\phi}_\infty(x)$ has again a natural interpretation in terms of the propensity of the particle system to survive to its extinction fate.

(iii) with $\psi = 1$ reads $\frac{1}{t}\mathbf{E}\left[N_t(x) \mid N_t(x) > 0\right] \underset{t\to\infty}{\to} \mu$ giving an interpretation of the constant μ. The constant μ is also ([2], page 287)

$$\mu = \frac{1}{2}\lambda_* \int_0^1 \kappa(x)\overline{\phi}_\infty(x)^2 \overline{q}_\infty(x)\, dx = \lambda_* \int_0^1 p_2(x)\overline{\phi}_\infty(x)^2 \overline{q}_\infty(x)\, dx < \infty$$

and so is explicitly available in our case.

The ground states of $\overline{G} + \lambda_0$ and its adjoint are thus $\left(\overline{\phi}_\infty, \overline{q}_\infty\right)$ and explicit here. It is also useful to consider the process whose infinitesimal generator is given by the Doob-transform

$$\overline{\phi}_\infty^{-1}\overline{G}\left(\overline{\phi}_\infty\cdot\right) = \overline{\phi}_\infty^{-1}\left(\widetilde{G} + \lambda\right)\left(\overline{\phi}_\infty\cdot\right),$$

because product-criticality is preserved under this transformation. The ground states associated to this new operator and its dual are $\left(1, \overline{\phi}_\infty\overline{q}_\infty\right)$. Developing,

we obtain a process whose infinitesimal generator is

$$\widetilde{G} + \frac{\overline{\phi}'_\infty}{\overline{\phi}_\infty} g^2 \partial_x = G + \frac{u'_0}{u_0} g^2 \partial_x = G,$$

with no multiplicative part. The associated diffusion process is the starting point WF diffusion with mutations, which is positive recurrent and so its invariant measure $\overline{\phi}_\infty \overline{q}_\infty = u_0 v_0 = m(y)$ is integrable.

Remark. The functional generating function $u(x,t;z) = \mathbf{E}\left[\prod_{n=1}^{N_t(x)} z^{\psi\left(\widetilde{x}_t^{(n)}\right)}\right]$ of the measure-valued branching particle system obeys now the nonlinear (quadratic) PDE:

$$\partial_t u(x,t;z) = \lambda_* \theta(x, u(x,t;z)) + \widetilde{G}(u(x,t;z)); \ u(x,0;z) = z^{\psi(x)},$$

where $\theta(x,z) = \mathbf{E}\left[z^{M(x)}\right] - z = (p_2(x) z^2 + p_0(x)) - z$ or

$$\theta(x,z) = (z-1)(p_2(x) z - p_0(x))$$

is the shifted probability generating function of the branching law of $M(x)$.

If $u(x,t) := \partial_z u(x,t;z)_{z=1} = \mathbf{E}\left[\sum_{n=1}^{N_t(x)} \psi\left(\widetilde{x}_t^{(n)}\right)\right]$, recalling $\lambda(x) = \lambda_*(p_2(x) - p_0(x))$, $u(x,t)$ obeys the linear backward PDE

$$\partial_t u(x,t) = \lambda(x) u(x,t) + \widetilde{G}(u(x,t)); \ u(x,0) = \psi(x)$$

involving $\overline{G}(\cdot) = \widetilde{G}(\cdot) + \lambda(x) \cdot$. It holds that

$$u(x,t) = \mathbf{E}_x\left(e^{\int_0^t \lambda(\widetilde{x}_s)ds} \psi(\widetilde{x}_t)\right). \ \diamond$$

References

[1] Asmussen, S.; Hering, H. Strong limit theorems for general supercritical branching processes with applications to branching diffusions. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 36, no. 3, 195-212, 1976.

[2] Asmussen, S.; Hering, H. Some modified branching diffusion models. Math. Biosci. 35, no. 3-4, 281-299, 1977.

[3] Crow, J.; Kimura, M. *An introduction to population genetics theory.* Harper & Row, Publishers, New York-London 1970.

[4] Crow, J.; Kimura, M. Some genetic problems in natural populations. Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. IV, pp. 1–22. University of California Press, Berkeley and Los Angeles, 1956.

[5] Durrett, R. Probability models for DNA sequence evolution. Second edition. Probability and its Applications (New York). Springer, New York, 2008.

[6] Dynkin, E. B. *Markov processes.* Vols. I, II. Translated with the authorization and assistance of the author by J. Fabius, V. Greenberg, A. Maitra, G. Majone. Die Grundlehren der Mathematischen Wi ssenschaften, Bände 121, 122 Academic Press Inc., Publishers, New York; Springer-Verlag, Berlin-Göttingen-Heidelberg 1965 Vol. I: xii+365 pp.; Vol. II: viii+274 pp.

[7] Ethier, S. N.; Kurtz, T. G. *Markov processes. Characterization and convergence.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986.

[8] Ewens, W. J. *Mathematical population genetics. I. Theoretical introduction.* Second edition. Interdisciplinary Applied Mathematics, 27. Springer-Verlag, New York, 2004.

[9] Gillespie, J. H. *The Causes of Molecular Evolution.* New York and Oxford: Oxford University Press, 1991.

[10] Huillet, T. On Wright-Fisher diffusion and its relatives. J. Stat. Mech., Th. and Exp. P11006, vol.11, 2007.

[11] Itô, K. On stochastic differential equations. Mem. Am. Math. Soc. No. 4 (1951).

[12] Itô, K.; McKean, H. P. *Diffusion processes and their sample paths.* Die Grundlehren der Mathematischen Wissenschaften, Band 125 Academic Press, Inc., Publishers, New York; Springer-Verlag, Berlin-New York 1965.

[13] Karlin, S.; Taylor, H. M. *A second course in stochastic processes.* Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1981.

[14] Kimura, M. Diffusion Models in Population Genetics. J. Appl. Prob. 1, 177-232, 1964.

[15] Kimura, M. Stochastic processes and distribution of gene frequencies under natural selection. Cold Spring Harbor Symposia on Quantitative Biology 20, 33-53, 1955. Also, contribution 114 of the National Institute of Genetics, Mishima-shi, Japan.

[16] Kimura, M. On the probability of fixation of mutant genes in a population. Genetics, 47, 713-19, 1962.

[17] Kolmogorov, A., Petrovsky, I.; Piscounov, N. Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique. Moscou Univ. Bull. Math. 1, pp. 125, 1937.

[18] Lambert, A. Population dynamics and random genealogies. Stochastic Models 24, suppl. 1, 45–163, (2008).

[19] Mandl, P. *Analytical treatment of one-dimensional Markov processes.* Die Grundlehren der mathematischen Wissenschaften, Band 151 Academia Publishing House of the Czechoslovak Academy of Sciences, Prague; Springer-Verlag New York Inc., New York 1968.

[20] Mano, S. Duality, ancestral and diffusion processes in models with selection. Theoretical Population Biology 75, 164-175, 2009.

[21] Maruyama, T. *Stochastic problems in population genetics.* Lecture Notes in Biomathematics, 17. Springer-Verlag, Berlin-New York, 1977.

[22] Pinsky, R. G. *Positive harmonic functions and diffusion.* Cambridge Studies in Advanced Mathematics, 45. Cambridge University Press, Cambridge, 1995.

[23] Yaglom, A. M. Certain limit theorems of the theory of branching random processes. Doklady Akad. Nauk SSSR (N.S.) 56, 795–798, (1947).