

K-Branching Random Walk with Noisy Selection: Large Population Limits and Phase Transitions

Colin Desmarais*[†] Emmanuel Schertzer* Zsófia Talyigás*

Abstract

We analyze a variant of the Noisy *K*-Branching Random Walk, a population model that evolves according to the following procedure. At each time step, each individual produces a large number of offspring that inherit the fitness of their parents up to independent and identically distributed fluctuations. The next generation consists of a random sample of all the offspring so that the population size remains fixed, where the sampling is made according to a parameterized Gibbs measure of the fitness of the offspring. Our model interpolates between classical models of fitness waves and exhibits a novel phase transition in the propagation of the wave. By employing a stochastic Hopf-Cole transformation, we show that as we increase the population size, the random dynamics of the model can be described by deterministic operations acting on the limiting population densities. We then show that for fitness fluctuations with exponential tails, these operations admit a unique traveling wave solution with local stability. The traveling wave solution undergoes a phase transition when changing selection pressure, revealing a complex interaction between evolution and natural selection.

Contents

1	Introduction	2
1.1	The model	2
1.2	Large population limit	3
1.3	Traveling wave solutions and the phase transition	7
1.4	Conjectures and related work	11
1.5	Outline of the paper	12
1.6	Some notation	12
2	Dynamics of the model as operators on limiting log-profiles	13
2.1	Technical Lemmas	14
2.1.1	A Laplace's principle type Lemma	14
2.1.2	Concentration inequalities	17
2.2	Limiting log-profiles after reproduction	19
2.3	Limiting log-profiles after sampling	22
2.4	Proof of Theorem 1.3	25
2.4.1	Example with no limiting log-profile after selection	27

*Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria

[†]Faculty of Mathematics and Geoinformation, Vienna University of Technology, Wiedner Hauptstraße 8–10, 1040 Vienna, Austria

3	Piecewise linear log-profiles	28
3.1	Operators on piecewise linear log-profiles	29
3.2	The space of piecewise linear log-profiles is attractive	33
4	Traveling Wave Solutions	37
4.1	Piecewise affine transformations of the front	38
4.2	Proof of Theorem 1.4	41
4.3	Proof of Theorem 1.5	44
4.3.1	Proofs of Continuity Lemmas	46

1 Introduction

Branching particle systems with selection have attracted significant attention in recent years as powerful models for understanding the interplay between reproduction, competition, and evolutionary dynamics. In these systems, the spatial location of a particle encodes its fitness, with particles further to the right corresponding to fitter individuals. Branching corresponds to reproduction: each particle gives rise to a fixed or random number of children, whose positions (fitness values) are random modifications of their parents' positions. The selection mechanism enforces that only the rightmost (fittest) particles survive and reproduce. We call this procedure truncation selection. A seminal contribution came from Brunet and Derrida [5, 6], who introduced a class of such systems, formulated key conjectures about their behavior, and analyzed an exactly solvable version that has since served as a benchmark for both rigorous and heuristic studies.

A natural extension of this framework is to introduce noisy selection, where survival is no longer restricted to the very fittest individuals but instead determined probabilistically. In this setting, one samples from the population in a way that favors fitter particles: those further to the right have a higher chance of being chosen, while still allowing less fit individuals to occasionally persist. This stochastic relaxation of truncation selection better reflects biological and evolutionary contexts where randomness (luck) influences reproductive success. Such models have been studied by Cortines and Mallein [9], by Schertzer and Wences [26], and in a companion paper which investigates the biological implications of the present rigorous mathematical results on a slightly different model [21]. These works, as well as the present article, highlight how noise affects the long-term behavior of different branching-selection particle systems.

1.1 The model

We consider an asexual population of fixed size K in discrete time. To each individual in the population, we associate a real number, which represents the individual's evolutionary fitness. The population evolves at each time step according to two substeps: reproduction and selection. Let X be a real-valued random variable with cumulative distribution function F_X , let $\beta > 0$ and $m \in \mathbb{N} \setminus \{0\}$. We will use the notation $[n] := \{1, 2, \dots, n\}$ for $n \in \mathbb{N}$. We iterate the following steps:

Reproduction. Each of the K individuals produces a fixed number of offspring m . These children inherit the fitness of their parent up to an independent random fluctuation. The fitness values of the children have independent i.i.d. displacements from their

parents' fitness with distribution X . That is, the child of an individual with fitness x has fitness distributed as $x + X$.

Random selection. Following reproduction, the population consists of $N := mK$ individuals. To regulate the size of the population, we sample K individuals without replacement according to the Gibbs sampling weights $(e^{\beta y_i})_{i=1}^N$, where $(y_i)_{i=1}^N$ denotes the fitness values of the offspring.

One may also think of this process as a particle system with K particles on the real line at all times, where each particle location represents the fitness value of an individual of the population. In order to define the process more formally, we let $(X_{i,j}^t)_{i \in [K], j \in [m], t \in \mathbb{N}}$ be i.i.d. random variables distributed as X . Let the initial particle configuration be given by $\mathcal{Z}_K^0 := (Z_i^0)_{i=1}^K$. Now the process $(\mathcal{Z}_K^t)_{t \in \mathbb{N}}$ is defined inductively as follows. Assume that at some time $t \in \mathbb{N}$, the configuration is given by $\mathcal{Z}_K^t = (Z_i^t)_{i=1}^K$. Then, after the **reproduction** step, the locations of the offspring particles are given by the collection $(Z_i^t + X_{i,j}^t)_{i \in [K], j \in [m]}$. Now in the **selection** step, out of the $N = Km$ offspring, we sample K without replacement in such a way that the probability of picking an offspring particle at location $y \in \mathbb{R}$ is proportional to $e^{\beta y}$. The configuration $\mathcal{Z}_K^{t+1} = (Z_i^{t+1})_{i=1}^K$ at time $t+1$ is then given by the locations of the sampled particles.

The case $\beta = \infty$ is a version of the K -Branching Random Walk (K -BRW), see e.g. [5, 6, 7], in which only the K fittest offspring survive after selection. We will refer to the present model as the noisy K -BRW with parameter (β, m) and reproduction law F_X . The aim of the present article is to study the case of high reproduction bias ($m \gg 1$) and show that an interesting phase transition occurs in the large population limit. See the next section for an interpretation of this parameter regime.

Our work relates to another group of fitness wave models, considered, for example, in [11, 8, 28, 29, 23], where reproduction and selection happen in one step: a new individual picks a parent with probability proportional to the parent's fitness (or alternatively, fitter individuals reproduce at higher rates). Although the details of this scheme differ from those of the K -BRW type models (where reproduction and selection are performed in two substeps), both approaches are believed to capture the same universal large-population behavior. By varying the parameters of the noisy K -BRW we will see that it exhibits behaviors similar to the above classical models in one parameter regime, while we find new phenomena in the other.

1.2 Large population limit

Our first theorem says that, in the noisy K -BRW model, the fitness distribution (fitness wave) evolves according to a deterministic dynamical system in the large population limit. Before stating this theorem, we specify our assumptions on the parameters of the model (F_X, β, m) and on the class of functions which will describe the fitness distribution.

Tail assumption. Recall that $N = mK \in \mathbb{N}$ denotes the total number of offspring at each time step. Let F_X be the reproduction law of the noisy K -BRW and let

$$c_N := (1 - F_X)^{-1}(1/N) := \inf\{x : 1 - F_X(x) \geq 1/N\}. \quad (1.1)$$

Assume that $c_N \rightarrow \infty$, as $N \rightarrow \infty$, and for $\rho > 0$ and $c_- \in \mathbb{R}_+ \cup \{+\infty\}$ we assume the existence of a function

$$h(x) := \begin{cases} -x^\rho & x \geq 0, \\ -c_-|x|^\rho & x < 0, \end{cases} \quad (1.2)$$

such that

$$\begin{aligned} \forall x \geq 0, \quad \lim_{N \rightarrow \infty} \frac{\log(1 - F_X(xc_N))}{\log N} &= h(x), \quad \text{and} \\ \forall x < 0, \quad \lim_{N \rightarrow \infty} \frac{\log(F_X(xc_N))}{\log N} &= h(x). \end{aligned} \tag{1.3}$$

This amounts to considering a sub ($\rho < 1$) or super ($\rho \geq 1$) exponential tail on the right. When $c_- = +\infty$, the tail is lighter on the left, while the decay is of the same order on both sides when $c_- < \infty$.

Log-profiles. Let $\beta > 0$, $\gamma \in (0, 1)$ and c_N as in (1.1). We consider a noisy $\lfloor N^\gamma \rfloor$ -BRW with Gibbs parameter β_N and fertility m_N defined as

$$\beta_N := \frac{\beta}{c_N} \log N, \quad m_N := N / \lfloor N^\gamma \rfloor. \tag{1.4}$$

Let us briefly comment on our choice of parameters. First, we note that $m_N \rightarrow \infty$. Biologically, this means that every individual reproduces a large number of offspring (or gametes) but only a few of those are selected to the next generation (r-selected population [19]). Children of highly fit individuals will have a better chance to survive selection than their peers, leading to high reproductive bias.

The biological significance of the parameter γ is as follows. For a fixed value of N , a lower γ entails a higher selection pressure since only a substantially reduced number ($K := \lfloor N^\gamma \rfloor$) of individuals can reproduce; whereas a high γ (close to 1) corresponds to a mild selection scheme where a larger set of children survive to the next generation. As a consequence, γ can be interpreted as capturing the selection pressure in the population. When γ is small, selection pressure is high; when γ is close to 1, selection pressure is low.

The parameter β measures how noisy the system is: large β corresponds to low noise, small β corresponds to high noise. Furthermore, by (1.4), the Gibbs sampling weight of a particle at a location xc_N is a constant power of N : $N^{\beta x}$. Our scaling parameter c_N is defined in (1.1) in such a way that, for any $y > 0$, the tail probability $\mathbb{P}(|X| > yc_N)$ is roughly of order N^{-y^ρ} ; a constant power of N .

We now describe a stochastic Hopf-Cole transformation of the system, which allows us to capture the time evolution of the fitness distribution in terms of powers of N . In order to justify the upcoming definition, let us use the informal notation $a_N \approx b_N$ to say that a_N and b_N are close to each other in some loose sense when N is large. Let $\mathcal{Z}_N^t$ refer to the configuration of particles at time t . Roughly speaking, we will say that $(\mathcal{Z}_N^t)_{N \in \mathbb{N}}$ admits a *limiting log-profile* if and only if there exists a deterministic function $g^t : \mathbb{R} \rightarrow \mathbb{R}_+ \cup \{-\infty\}$ such that the number of particles in an interval of size $c_N dx$ around xc_N is close to $N^{g^t(x)} dx$, and so the number of particles in an interval $[ac_N, bc_N]$ is close to $\int_a^b N^{g^t(x)} dx \approx N^{\sup_{[a,b]} g^t}$. A more formal definition is given below. Note that when $g^t(x) = -\infty$ there are no particles around xc_N .

Definition 1.1. A function $g : \mathbb{R} \rightarrow \mathbb{R}_+ \cup \{-\infty\}$ belongs to the class $\mathcal{D}$ if and only if

- (a) g admits a left limit and a right limit for all $x \in \mathbb{R}$, and
- (b) g has bounded support $\text{Supp}(g) = \{x \in \mathbb{R} : g(x) \neq -\infty\}$, in which case we define $L(g) \leq U(g)$ such that $[L(g), U(g)]$ is the smallest interval containing $\text{Supp}(g)$. We call $L(g)$ and $U(g)$ the *lower edge* and *upper edge* of g respectively.

A function $g : \mathbb{R} \rightarrow \mathbb{R}_+ \cup \{-\infty\}$ belongs to the class $\mathcal{C}$ if and only if $g \in \mathcal{D}$ and

(c) g is concave on its support.

We now formally define what it means for a sequence of point measures to have a limiting log-profile. We denote the discontinuity points of g by D_g .

Definition 1.2 (limiting log-profile). A sequence of (random) point measures $(\mu_N)_{N \in \mathbb{N}}$ is said to have a *limiting log-profile* if there exists a deterministic function $g \in \mathcal{D}$ such that for all $a, b \in \mathbb{R} \setminus D_g$ with $a < b$,

$$\frac{\log \mu_N((a, b])}{\log N} \xrightarrow{\mathbb{P}} \sup_{x \in (a, b]} g(x), \text{ as } N \rightarrow \infty. \quad (1.5)$$

Next, we define two operators r and s_σ on the space $\mathcal{D}$, which describe how the limiting log-profile of the configuration changes after a reproduction and a selection step, respectively. We give a heuristic explanation after the statement of Theorem 1.3. Let us define the function π by

$$\pi(x) := \begin{cases} x & \text{if } x \geq 0 \\ -\infty & \text{otherwise.} \end{cases} \quad (1.6)$$

For all $g \in \mathcal{D}$ and $x \in \mathbb{R}$, we let

$$r(g)(x) := \pi \left(\sup_{z \in \mathbb{R}} (1 - \gamma + g(z) + h(x - z)) \right), \quad (1.7)$$

$$\text{and } \forall \sigma > 0, \quad s_\sigma(g)(x) := \pi(g(x) + \beta(x - \sigma)_-). \quad (1.8)$$

We can now state the first main result of this paper. Analogous to the notation in Section 1.1, let $(Z_{i,N}^t)_{i=1}^{N^\gamma}$ denote the particle locations of the noisy $\lfloor N^\gamma \rfloor$ -BRW at time t . Recall the definitions of c_N and h from (1.1) and (1.2) and recall β_N and m_N from (1.4).

Theorem 1.3. For all $t \geq 0$, we define the sequence of point measures $(M_N^t)_{N \in \mathbb{N}}$ by

$$M_N^t := \sum_{i=1}^{N^\gamma} \delta_{Z_{N,i}^t/c_N}. \quad (1.9)$$

Assume that

(i): the sequence of initial configurations $(M_N^0)_{N \in \mathbb{N}}$ admits a limiting log-profile $g^0 \in \mathcal{C}$,

(ii): the function h is concave on its support, that is $\rho \geq 1$.

Then, for all $t \in \mathbb{N}$, $(M_N^t)_{N \in \mathbb{N}}$ admits a limiting log-profile $g^t \in \mathcal{C}$. Further, (g^t) satisfies a recursive equation in terms of the following discrete “free boundary” problem,

$$\forall x \in \mathbb{R}, \quad g^{t+1}(x) = s_{\sigma_t} \circ r(g^t)(x), \quad (1.10)$$

where σ^t is defined implicitly through the relation

$$\sigma^t := \inf \left\{ \sigma : \sup_{z \in \mathbb{R}} s_\sigma(r(g^t))(z) \leq \gamma \right\}. \quad (1.11)$$

Heuristics. Suppose that the sequence of configurations $(M_N^t)_{N \in \mathbb{N}}$ admits a limiting log-profile g^t ; that is, the number of particles around a location zc_N at time t is roughly $N^{g^t(z)}dz$. Then a first moment argument yields that the averaged number of offspring after reproduction (but before selection) at a location xc_N is

$$\approx \int N^{g^t(z)} N^{1-\gamma} N^{h(x-z)} dz \approx N^{\sup_z (1-\gamma+g^t(z)+h(x-z))}, \quad (1.12)$$

where the integral counts the average contribution of every possible parent location, using the fact that we have $\approx N^{g^t(z)}dz$ particles at zc_N , and each of them produces about $N^{1-\gamma}$ offspring with $\approx N^{1-\gamma+h(x-z)}$ of those offspring around xc_N . A Laplace's principle type of argument together with the application of concentration results yields that the stochastic exponent at x converges in probability to the deterministic value $r(g^t)(x)$ defined in (1.7). The projector π encodes the fact that, when the average number of particles at a location is a negative power of N , then with high probability, there are no particles at that location. Note that the total number of offspring particles after reproduction is N , which is reflected by the fact that $\sup_{x \in \mathbb{R}} r(g^t)(x) = 1$ (when $\sup_{x \in \mathbb{R}} g^t(x) = \gamma$).

Next, finding the averaged number of particles sampled in the selection step at every location will lead us to determining the limiting log-profile at time $t+1$. We need to sample $\lfloor N^\gamma \rfloor$ particles without replacement according to the Gibbs weights. For this purpose, we will use an equivalent sampling procedure involving exponential competing clocks, which is convenient for our analysis. We endow each of the N offspring particles with an exponential random variable (clock). If $(x_i^t c_N)_{i=1}^N$ is the set of particle locations after reproduction, then the respective rates of the exponential clocks are $e^{\beta_N x_i^t c_N} = N^{\beta x_i^t}$. We then select the $\lfloor N^\gamma \rfloor$ particles with the earliest ring times. We will now argue that the time when $\lfloor N^\gamma \rfloor$ clocks have rung, is given by $N^{-\beta \sigma_N}$ for some $\sigma_N \in \mathbb{R}$.

To determine the limiting log-profile after selection, we first find the limiting log-profile of the configuration of particles, whose clocks rang before time $N^{-\beta \sigma}$ for any $\sigma \in \mathbb{R}$. We show that most particles at a location $x_i^t c_N > \sigma c_N$ will have their clocks ring before time $N^{-\beta \sigma}$, while a particle $x_i^t c_N < \sigma c_N$ has probability $\approx N^{\beta(x_i^t - \sigma)}$ that their clock will ring before time $N^{-\beta \sigma}$. The effect is that, for any fixed σ , the average number of offspring around xc_N , whose clock rang before time $N^{-\beta \sigma}$, is

$$\approx N^{r(g^t)(x) + \beta(x - \sigma) -} dx. \quad (1.13)$$

Applying Laplace's principle and concentration results once more, we conclude that the stochastic exponent of the number of particles at xc_N whose clock rang before time $N^{-\beta \sigma}$ converges in probability to $s_\sigma \circ r(g^t)(x)$ (see (1.8)). Thus, we find the total number of such particles by integrating over all locations; and since we need this to be $\sim N^\gamma$, we solve

$$N^\gamma \approx \int N^{(s_{\sigma_t} \circ r)(g^t)(x)} dx \approx N^{\sup_x (s_{\sigma_t} \circ r)(g^t)(x)} \quad (1.14)$$

for σ_t , and we will find that $N^{-\beta \sigma_t}$ is indeed the deterministic limit of the times $(N^{-\beta \sigma_N})_N$, when N^γ clocks have rung. Therefore, we will be able to conclude that the limiting log-profile after selection, at time $t+1$ is

$$g^{t+1} = s_{\sigma_t} \circ r(g^t),$$

where σ_t is given by the implicit relation (1.11).

The heuristics above only apply where the first moment argument yields a non-zero exponent; otherwise we lack the information needed to apply our second moment computation and we cannot determine whether or not there are any particles around that location. This is illustrated in Section 2.4.1. As a benefit of assuming (i) and (ii) of Theorem 1.3, we show that the functions in the exponents of (1.12) and (1.13) after our first moment argument are also concave, and within any interval whose endpoints are continuity points of the functions $r(g^t)$ and $s_\sigma \circ r(g^t)$ respectively, the maximum value attained is non-zero. As we have defined limiting log-profiles in Definition 1.2 outside of discontinuity points, we may indeed apply our second moment computation and prove that $r(g^t) \in \mathcal{C}$ and $s_\sigma \circ r(g^t) \in \mathcal{C}$ are the limiting log-profiles of our processes after reproduction and sampling according to exponential clocks respectively.

We show in Section 2.2 that the argument laid out above for reproduction generalizes to $g^t \in \mathcal{D}$ and $\rho > 0$, and a limiting log-profile exists. However, a limiting log-profile after selection cannot be guaranteed in general, as illustrated in Section 2.4.1. Under some additional conditions, our argument for selection extends to limiting log-profiles $g^t \in \mathcal{D}$, and $\rho > 0$ which is captured in Theorem 2.10.

1.3 Traveling wave solutions and the phase transition

Theorem 1.3 establishes the dynamical system that describes the behavior of the fitness wave in the noisy $\lfloor N^\gamma \rfloor$ -BRW model. Therefore, in order to characterize this stochastic model, we may analyze the dynamical system (1.10). We are particularly interested in how the noise parameter β and the selection pressure γ affect the shape and speed of the fitness wave. Note that the latter can be interpreted as the rate of adaptation of the population to its environment.

We have explicit results for the particular case when the displacement distribution F_X has exponential tails. Throughout this section, we assume that $\rho = 1$ for h in (1.2). We say that the dynamics (1.10) admits a traveling wave solution if and only if there exists (G, ν) such that if we initialize the profile g^0 at G , i.e. $g^0 = G$, then

$$\forall x \in \mathbb{R}, \forall t \geq 0, \quad g^{t+1}(x + t\nu) = G(x).$$

Traveling wave solutions are always defined up to translation and provided that $\text{Supp}(G)$ is bounded, we take the convention that $U(G) = 0$, where we recall from Definition 1.1 that $U(G)$ is the upper edge of the profile. For the space of functions in $\mathcal{D}$, we define the metric

$$\phi(g_1, g_2) := \sup_{x \in \mathbb{R}} |g_1(x)_+ - g_2(x)_+| \vee |U(g_1) - U(g_2)| \vee |L(g_1) - L(g_2)|.$$

Selection of the fittest and luckiest phase transition. Theorem 1.4 says that there is a unique traveling wave solution to the dynamics (1.10). Moreover, later in this manuscript, we provide an explicit expression for this traveling wave solution. We will prove that it can be described as concave piecewise linear functions as illustrated in Fig. 1. The other message of Theorem 1.4 is that the noisy $\lfloor N^\gamma \rfloor$ -BRW exhibits a phase transition. This phase transition is characterized in two ways: by the rate of adaptation and by the geometry of the wave.

Theorem 1.4. *Let $\beta > 0$, $\gamma \in (0, 1)$ and assume that $\rho = 1$ in (1.2). Then, the dynamical system (1.10) admits a unique traveling wave solution G (explicitly given by Proposition 4.6) with speed ν . Furthermore, let*

$$k := k(\beta) = \begin{cases} \lfloor \frac{1}{\beta} \rfloor & \text{if } \frac{1}{\beta} \notin \mathbb{N} \\ \frac{1}{\beta} - 1 & \text{otherwise} \end{cases} \quad \text{and} \quad \gamma_c(\beta) := \frac{k(2 - (k+1)\beta)}{(k+1)(2 - k\beta)}. \quad (1.15)$$

Then the traveling wave solution exhibits the following phase transition. For a fixed value of β ,

(a) **Selection of the luckiest:** *If $\gamma < \gamma_c(\beta)$,*

(a.1) *the speed of the wave is given by*

$$\nu(\gamma, \beta) = \frac{2\gamma}{k(2 - (k+1)\beta)}.$$

In particular, it is increasing in γ .

(a.2) *If $g^0 = G$, then $\forall x \in \text{Supp}(g^1)$, $r(g^0)(x) > g^1(x)$.*

(b) **Selection of the fittest:** *If $\gamma \geq \gamma_c(\beta)$,*

(b.1) *the speed of the wave is given by*

$$\nu(\gamma, \beta) = 1 - \gamma.$$

In particular, it is decreasing in γ .

(b.2) *If $g^0 = G$, then $\forall x \in [-\chi + 1 - \gamma, 1 - \gamma]$, $r(g^0)(x) = g^1(x)$ where*

$$\chi(\gamma, \beta) := \gamma - \sum_{j=1}^k (1 - j\beta)(1 - \gamma), \quad (1.16)$$

which is strictly positive on $(\gamma_c, 1)$.

Local convergence. Our next theorem shows local convergence to the traveling wave solution G of (1.10). We discuss some of the implications after the statement of Theorem 1.5. Recall the critical value $\gamma_c(\beta)$ from (1.15).

Theorem 1.5. *Let $\beta > 0$, $\gamma \in (0, 1)$ and assume that $\rho = 1$ in (1.2). If we also assume the technical conditions $\gamma \neq \gamma_c(\beta)$ and $1/\beta \notin \mathbb{N}$ whenever $\gamma > \gamma_c(\beta)$, then the traveling wave solution G of (1.10) is locally stable in the sense that there exists $\delta > 0$ such that if $g^0 \in \mathcal{D}$ and*

$$\phi(g^0, G) \leq \delta$$

then there exists $c \in \mathbb{R}, C > 0$ and $r \in (0, 1)$ such that

$$\forall t \in \mathbb{N}, \quad \phi(g^t(\cdot + \nu t + c), G) \leq Cr^t.$$

Interpretation. To understand the significance of the theorems above, we recall that γ captures the selection pressure (high γ means low selection pressure). Theorem 1.4 says that for any $\beta > 0$ there is a critical value of γ that maximizes the speed of the

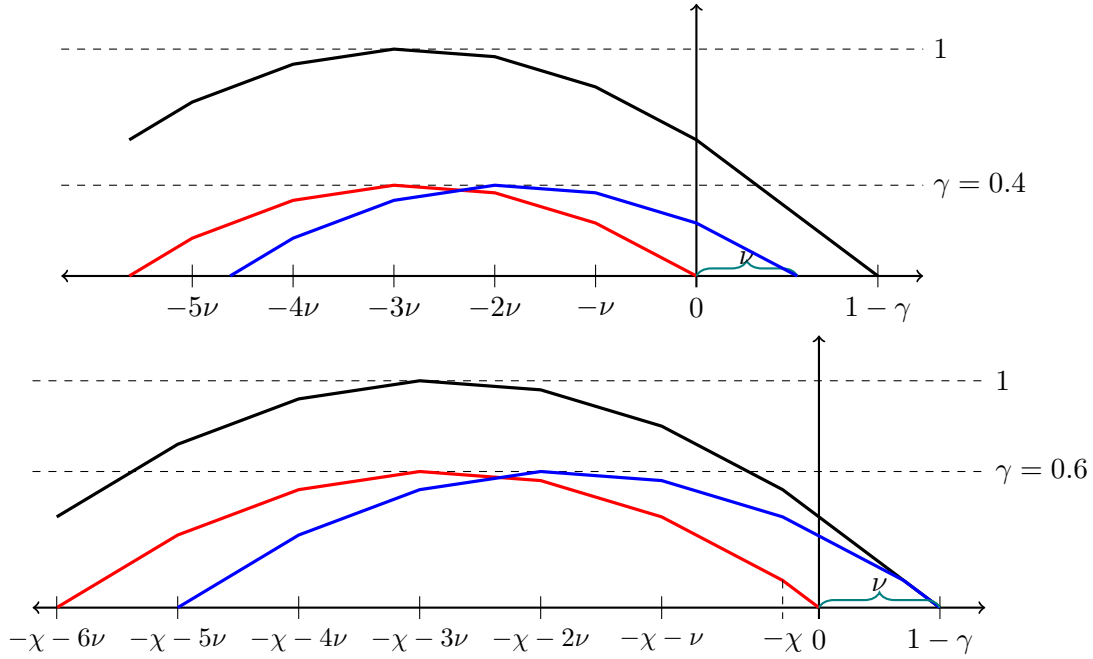


Figure 1: Two examples of G (in red) where $c_- = \infty$ and $\beta = 0.3$, along with the limiting log-profiles after reproduction $r(G)$ (in black) and the limiting log-profile after selection $s \circ r(G) = G(\cdot - \nu)$ in blue. The first plot corresponds to the selection of the luckiest regime with $\gamma = 0.4$, and the second plot corresponds to the selection of the fittest regime with $\gamma = 0.6$. In both plots, G consists of line segments with slopes $-(1 - 6\beta), \dots, -(1 - \beta)$, and a further line segment with slope -1 in the second plot.

traveling wave. That is, the function $\gamma \mapsto \nu(\gamma, \beta)$ is maximized at $\gamma = \gamma_c(\beta)$ from (1.15) (see also Fig. 2). In the regime to the left of γ_c , as we decrease γ from γ_c to 0, selection becomes stricter: fewer of the N children are allowed to reproduce. We say that selection is ‘too strict’ in this case, in the sense that the individual with the largest fitness (i.e. the rightmost particle) after reproduction is not selected. This is the message of the statement (a.2) and it is illustrated in the first plot of Fig 1. The rightmost selected particle (right edge of the blue curve) is to the left of the rightmost particle after reproduction (right edge of the black curve): the few fittest individuals after reproduction do not have enough weight to survive under high selection pressure. As a result, as Theorem 1.4 and Fig. 2 show, higher selection pressure is detrimental to the rate of adaptation, that is, the speed of the wave is increasing in γ . We call this regime selection of the luckiest, because it is the noise that determines which of the children survive to the next generation rather than the actual fitness of the children.

On the other hand, when $\gamma \geq \gamma_c$, more of the N children are allowed to reproduce and enough for the fittest to be selected, hence, we call this regime selection of the fittest. The second plot of Fig 1 shows that there is an interval of length χ , in which the log-profile after reproduction (black) agrees with the log-profile after selection (blue). That is, there is an interval at the front of the wave, in which every offspring particle is selected to survive. This is the message of statement (b.2) in Theorem 1.4. Once γ is large enough that the fittest individual survives, further increasing γ will mean that selection is ‘not strict enough’: the average fitness of the selected individuals decreases by increasing γ and

therefore the speed decreases, as shown in Fig. 2.

Comparing the geometry of the traveling wave solutions in the two regimes, we find that the slope of G at the front (near zero) differ in the two cases: in the selection of the luckiest case it is $-(1 - \beta)$, whereas in the selection of the fittest regime it is -1 . This property also reflects the fact that in the latter case there is a segment at the front of the population where every offspring survives selection.

Theorem 1.4 and Fig 2 also show that the speed of the traveling wave is monotone increasing in β , which is not surprising: as noise decreases speed increases. We also observe in this figure that $\gamma_c(\beta)$ decreases in β , which means that, if we optimize on the speed, we have to be more selective when selection is less noisy.

The paper [21] is closely related to our work and provides a more detailed description of the two regimes, focusing on the biological implications for a different version of our model. In that paper the noisy $\lfloor N^\gamma \rfloor$ -BRW is studied with the same reproduction step as in the present article, but with a selection step different from Gibbs sampling. Each individual is assumed to have a genotype and a phenotype, both represented by a real number. The genotype and phenotype of a child are distributed as $x + X$ and $x + X + Y$ respectively, where x denotes the genotype of the parent, $X \sim \text{Laplace}(1)$ represents random mutations, and $Y \sim \text{Laplace}(1/\mu)$ represents the noise between genotype and phenotype, and where the Laplace distribution with parameter $b > 0$ has density $\frac{1}{2b}e^{-|x|/b}$. Then, in the selection step, out of the N offspring, the offspring with the $\lfloor N^\gamma \rfloor$ largest phenotypes are selected. The next generation inherits the genotypes of the selected individuals up to random mutation.

This model also exhibits the phase transition between selection of the fittest and selection of the luckiest regimes, however, as opposed to the present work, the paper [21] does not provide a rigorous proof for the convergence to the deterministic dynamics or for the local convergence to the traveling wave. The selection of the luckiest regime in this model has an interesting interpretation; namely, it can be thought of as a Goodhart's law [16] in evolution. In general, Goodhart's law says that 'When a measure becomes a target, it ceases to be a good measure.' That is, when selection is highly noisy, selecting individuals only with the few largest observed phenotypes is not a good way to find individuals with the largest genotypes. One can certainly think of the same interpretation of the present article.

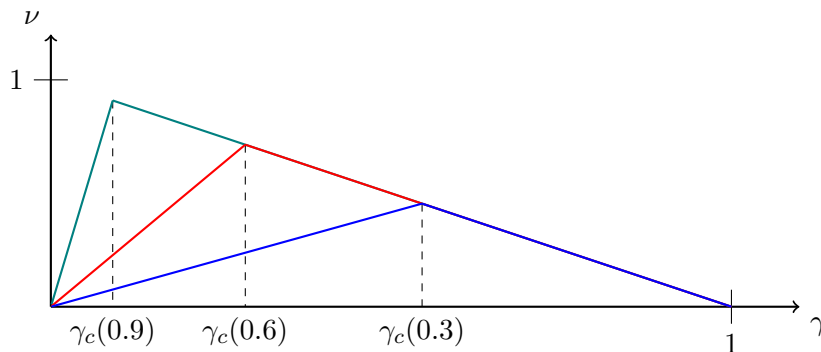


Figure 2: Plots of the speed of the traveling wave $\nu(\beta, \gamma)$ for 3 different values of β : $\beta = 0.3$ (in blue), $\beta = 0.6$ (in red), and $\beta = 0.9$ (in teal).

1.4 Conjectures and related work

Brunet and Derrida [5, 6] considered a version of a K -BRW and (heuristically) argued that the distribution of particles should be given by a deterministic FKPP-type equation, which admits traveling wave solutions; and that the speed of the stochastic system should be given by the minimal velocity solution of the deterministic equation. These conjectures led to several rigorous results, for example on the convergence of branching-selection particle systems to free boundary problems [13] and [3, 10]; on the speed of the traveling wave solution of the FKPP equation with cut-off [12] and with random noise [22]; and on the limiting velocity of branching-selection particle systems [1, 2, 9, 26]. Our results fit into this literature and leave some open questions, which are subject to future work.

Rate of adaptation. Our Theorems 1.3, 1.4, 1.5 show that under some general conditions on the tails of the reproduction law, the noisy $\lfloor N^\gamma \rfloor$ -BRW admits a limiting log-profile, which changes according to a deterministic dynamical system. For $\rho = 1$ (see (1.2)), we show local convergence to a unique traveling wave solution of this system, where the shape and speed of the wave are explicitly given. It is then natural to ask what we can say about the speed (rate of adaptation) of the original stochastic particle system. Does it converge to the same speed as the traveling wave solution? A first step towards such a result would be to show that for fixed values of N , an asymptotic speed $\nu_N \in \mathbb{R}$ exists, satisfying

$$\lim_{t \rightarrow \infty} \frac{\max_{i=1, \dots, \lfloor N^\gamma \rfloor} Z_{N,i}^t}{t} = \lim_{t \rightarrow \infty} \frac{\min_{i=1, \dots, \lfloor N^\gamma \rfloor} Z_{N,i}^t}{t} = \nu_N \quad \text{a.s.}$$

Assuming such an asymptotic speed exists one may investigate its behavior, as $N \rightarrow \infty$. Our results (in the case $\rho = 1$) suggest that we should see

$$\frac{\nu_N}{\log N} \rightarrow \nu$$

as $N \rightarrow \infty$, where ν is defined in Theorem 1.4.

Genealogical structure. Brunet, Derrida, Mueller and Munier [7] predicted that the Bolthausen-Sznitman coalescent [4, 24] is a universal limiting genealogy for branching-selection particle systems. We conjecture that the noisy $\lfloor N^\gamma \rfloor$ -BRW only belongs to this universality class in the selection of the fittest regime; and we see a different coalescent in the limit in the selection of the luckiest regime, when noise has the more significant role. This conjecture is in line with the results of [26].

In [26], a variation of the noisy $\lfloor N^\gamma \rfloor$ -BRW model was considered, which was introduced by Cortines and Mallein [9] (who proved convergence to the Bolthausen-Sznitman coalescent for $\beta > 1$) as a generalization of the exponential model of Brunet and Derrida [5, 7]. It is assumed in [26] that at each reproduction step, every individual produces an infinite number of offspring according to an exponential Poisson point process centered around the parental value. Then, $\lfloor N^\gamma \rfloor$ individuals are selected using a sampling scheme interpolating between truncation selection and Gibbs sampling.

The critical assumption of this model is the choice of the exponential reproduction law. In particular, one crucial observation (already made by Brunet and Derrida) is that, up to translation, the system reaches stationarity after a *single* step. (Compared to Theorem 1.5, the system reaches its travelling wave regime at infinite rate instead of geometric rate.)

This makes the model fully integrable and allows to make precise predictions on the large-population limit and in particular, on the genealogical structure of the model. In [26], the existence of a critical $\gamma_c(\beta) \in [0, 1]$ was proved, segregating between two different limiting regimes.

- If $\gamma < \gamma_c(\beta)$, then the genealogy of the population is given by the Poisson-Dirichlet coalescent with parameter $(0, \beta)$ [25, 27].
- If $\gamma > \gamma_c(\beta)$, then, after rescaling time by $c \log(N)$ (for some explicit choice $c \equiv c(\beta, \gamma)$), the genealogy of the population is given by the Bolthausen-Sznitman coalescent.

We conjecture that the same transition occurs in the noisy $\lfloor N^\gamma \rfloor$ -BRW. However, we emphasize that the proof in [26] heavily relies on the integrability of the exponential model and more specifically, on the property that the system reaches stationarity after only a single step. Thus, the geometric rate of convergence in Theorem 1.5 appears to be a first important step in the direction of proving the universality of the phase transition observed in the model.

1.5 Outline of the paper

This work proceeds as follows: Section 2 is dedicated to proving Theorem 1.3 and the more general Theorem 2.10. The proof relies on investigating the limiting log-profile after reproduction in Section 2.2 and after selection in Section 2.3. We examine an example of where a limiting log-profile after selection fails to materialize in Section 2.4.1.

In Sections 3 and 4 we prove that the deterministic system (1.10) admits a unique traveling wave solution, which is given by a piecewise linear function, which exhibits a phase transition as we change the selection pressure γ . In Section 3 we introduce the space $\mathcal{T}$ of concave piecewise linear functions, where the traveling wave solution will belong. In Section 3.1 we describe how the operators defined in (1.7) and (1.8) act on such piecewise linear functions. In Section 3.2 we prove that the space $\mathcal{T}$ is strongly attractive with respect to (1.10); starting from any $g^0 : \mathbb{R} \rightarrow \mathbb{R}_+ \cup \{-\infty\}$ with bounded support such that $\sup_x g^0(x) = \gamma$, after finitely many iterations of (1.10), every g^t is a concave piecewise linear function. Using the results of Section 3 we manage to translate (1.10) to a finite-dimensional system describing the evolution of the “break-points” where the slopes of the linear segments change. Analyzing this finite-dimensional problem, we prove Theorems 1.4 and 1.5 in Sections 4.2 and 4.3.

1.6 Some notation

Throughout the rest of this work, for ease of reading, we will assume that N^γ and $N^{1-\gamma}$ are integers. Specifically, we will write N^γ and $N^{1-\gamma}$ for $\lfloor N^\gamma \rfloor$ and $N/\lfloor N^\gamma \rfloor$ respectively. For an integer K , we use the notation $[K] = \{1, \dots, K\}$, and so we write $[N^\gamma] = \{1, \dots, \lfloor N^\gamma \rfloor\}$ and $[N^{1-\gamma}] = \{1, \dots, N/\lfloor N^\gamma \rfloor\}$.

We will write $\bar{\mathbb{R}} = \{-\infty\} \cup \{+\infty\} \cup \mathbb{R}$.

For a function $f : \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty\}$, we will say that f has a left limit at x if and only if $\lim_{z \rightarrow x^-} f(z)$ exists and is real, or if there exists $\delta > 0$ such that $f(z) = -\infty$ for all $z \in (x - \delta, x)$, in which case we say that $\lim_{z \rightarrow x^-} f(x) = -\infty$ exists. Then f is left continuous at x if and only if $\lim_{z \rightarrow x^-} f(z) = f(x)$. Similarly we define right limits and

right continuity, and say that f is continuous at x if and only if f is both left and right continuous at x , and discontinuous otherwise. We will denote the set of discontinuity points of f by D_f . Recall that any function f which admits a left and right limit at every $x \in \mathbb{R}$ has at most countably many discontinuities.

In a similar vein as càdlàg functions, we say that a function f is *càllàl* (continue à l'un, limite à l'autre) if and only if for all $x \in \mathbb{R}$, f is continuous on one side of x and has a limit on the other.

We extend the definition of concavity to functions $f \in \mathcal{D}$. Notice that such a function that is concave on a compact support $\text{Supp}(f) := \{x : f(x) \neq -\infty\}$ is also concave on the real line; if x or y is taken outside of $\text{Supp}(f)$, then for all $\alpha \in [0, 1]$,

$$f(\alpha x + (1 - \alpha)y) \geq \alpha f(x) + (1 - \alpha)f(y) = -\infty.$$

For any point measure M_N , we write

$$\forall a < b, \quad M_N(a, b) = M_N((a, b]).$$

For sequences of random variables $(X_N)_{N \in \mathbb{N}}$ and $(Y_N)_{N \in \mathbb{N}}$ we write $X_N \lesssim_{\mathbb{P}} Y_N \in \mathbb{R}$ as $N \rightarrow \infty$, if and only if, for all $\varepsilon > 0$,

$$\mathbb{P}(X_N - Y_N > \varepsilon) \rightarrow 0,$$

as $N \rightarrow \infty$. We often use the notation with X_N or Y_N being simply a deterministic constant. In particular, if $X_N \lesssim_{\mathbb{P}} c \in \mathbb{R}$ and $X_N \gtrsim_{\mathbb{P}} c$ as $N \rightarrow \infty$, then $X_N \xrightarrow{\mathbb{P}} c$ as $N \rightarrow \infty$.

We will also use the following property. Suppose that $X_N \lesssim_{\mathbb{P}} Y_N$ as $N \rightarrow \infty$, and that

$$Y_N \xrightarrow{\mathbb{P}} Y^*$$

as $N \rightarrow \infty$ for some random (or deterministic) variable Y^* . Then we have

$$X_N \lesssim_{\mathbb{P}} Y^*,$$

as $N \rightarrow \infty$. If we assume $X_N \gtrsim_{\mathbb{P}} Y_N$ instead, then it follows that $X_N \gtrsim_{\mathbb{P}} Y^*$.

2 Dynamics of the model as operators on limiting log-profiles

This section is divided as follows. We relegate our technical lemmas to Section 2.1. In Section 2.2, we prove that a deterministic operator r describes the change of the limiting log-profile for point measures after reproduction, while a similar result for the selection step under technical conditions is proven in Section 2.3. The conditions are needed since though we get a quite detailed perspective of the process when looking at a logarithmic scale, there are some exceptional cases where the scale does not offer enough precision to describe a limiting log-profile after sampling. These minor conditions, however, are always satisfied for limiting log-profiles in $\mathcal{C}$ and when h is concave. We then provide a proof of Theorem 1.3 in Section 2.4.

2.1 Technical Lemmas

2.1.1 A Laplace's principle type Lemma

We prove a version of Laplace's principle for the type of Riemann-Stieltjes integrals that we encounter throughout this work. Recall the definition of the space of functions $\mathcal{D}$ (Definition 1.1), and recall that a function f is càllàl if and only if for every $x \in \mathbb{R}$, the function f is continuous on one side and has a limit on the other. Before stating the main result of this section, we will need to prove an easy lemma.

Lemma 2.1. *Assume $(M_N)_{N \in \mathbb{N}}$ has limiting log-profile $g \in \mathcal{D}$ with lower and upper edges L, U . For any $a, b \in \bar{\mathbb{R}} \setminus D_g$ with $a < b$, if $\sup_{x \in (a, b]} g(x) = -\infty$, then as $N \rightarrow \infty$,*

$$M_N((a, b]) \xrightarrow{\mathbb{P}} 0.$$

Furthermore, for the leftmost point $\rho_{N,L}$ and rightmost point $\rho_{N,U}$ of M_N , as $N \rightarrow \infty$,

$$\rho_{N,L} \xrightarrow{\mathbb{P}} L, \quad \text{and} \quad \rho_{N,U} \xrightarrow{\mathbb{P}} U.$$

Proof. Since $M_N((a, b])$ takes integer values,

$$\mathbb{P}(M_N((a, b]) > 0) = \mathbb{P}\left(\frac{\log M_N((a, b])}{\log N} \geq 0\right).$$

Thus if $\log M_N((a, b]) / \log N \xrightarrow{\mathbb{P}} -\infty$, then $M_N((a, b]) \xrightarrow{\mathbb{P}} 0$.

Now we examine $\rho_{N,U}$. Using the above result and definition of $\mathcal{D}$, for any $\delta > 0$,

$$\mathbb{P}(\rho_{N,U} > U + \delta) = \mathbb{P}(M_N((U + \delta, \infty)) > 0) \xrightarrow{N \rightarrow \infty} 0. \quad (2.1)$$

Since g has countably many discontinuity points, we may always find $0 < \delta' \leq \delta$ such that g is continuous at $U - \delta'$, and so

$$\frac{\log M_N((U - \delta', \infty))}{\log N} \xrightarrow{\mathbb{P}} \sup_{x \in (U - \delta', \infty)} g(x) > 0$$

and as such,

$$\mathbb{P}(\rho_{N,U} < U - \delta) \leq \mathbb{P}(M_N((U - \delta', \infty)) = 0) \xrightarrow{N \rightarrow \infty} 0. \quad (2.2)$$

Then $\rho_{N,U} \xrightarrow{\mathbb{P}} U$ follows from (2.1) and (2.2). A symmetric argument holds for $\rho_{N,L}$. $\square$

Lemma 2.2. *Let $f_N : \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty\}$ be a sequence of functions and let $f : \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty\}$ be a càllàl function with discontinuity set D_f such that for all $x \in \mathbb{R}$,*

$$\lim_{\delta \rightarrow 0} \lim_{N \rightarrow \infty} \sup_{z \in B(x, \delta)} f_N(z) = \lim_{\delta \rightarrow 0} \sup_{z \in B(x, \delta)} f(z), \quad (2.3)$$

$$\lim_{\delta \rightarrow 0} \lim_{N \rightarrow \infty} \inf_{z \in B(x, \delta)} f_N(z) = \lim_{\delta \rightarrow 0} \inf_{z \in B(x, \delta)} f(z). \quad (2.4)$$

Let $(M_N)_{N \in \mathbb{N}}$ be a sequence of (random) point measures on $\mathbb{R}$ which has a limiting log-profile $g \in \mathcal{D}$ with discontinuity set D_g . If $D_f \cap D_g = \emptyset$, then for all $a, b \in \bar{\mathbb{R}} \setminus (D_f \cup D_g)$, as $N \rightarrow \infty$,

$$\frac{\log \left(\int_{(a, b]} N^{f_N(x)} dM_N(x) \right)}{\log N} \xrightarrow{\mathbb{P}} \sup_{x \in (a, b]} (f(x) + g(x)) \in \mathbb{R} \cup \{-\infty\}.$$

Remark 2.3. The conditions (2.3) and (2.4) hold trivially if f_N converges locally uniformly to f , while these conditions imply pointwise convergence at all continuity points of f .

Proof. Note throughout that (1.5) can be rewritten as

$$\frac{\log \left(\int_{(a,b]} dM_N(x) \right)}{\log N} \xrightarrow{\mathbb{P}} \sup_{x \in (a,b]} g(x) \quad (2.5)$$

for $a, b \notin D_g$.

We will first assume that all f_N and f are real-valued functions. Let L and U be the lower and upper edges of g . First suppose that $a = -\infty$ and $b < L$. From Lemma 2.1, $M_N((-\infty, b]) \xrightarrow{\mathbb{P}} 0$, and so in this case,

$$\mathbb{P} \left(\int_{(-\infty, b]} N^{f_N(x)} dM_N(x) > 0 \right) \leq \mathbb{P}(M_N(-\infty, b) > 0) \rightarrow 0$$

as $N \rightarrow \infty$, and so

$$\frac{\log \left(\int_{(-\infty, b]} N^{f_N(x)} dM_N(x) \right)}{\log N} \xrightarrow{\mathbb{P}} \sup_{x \in (-\infty, b]} g(x) = -\infty.$$

A similar argument holds if $a > U$ and $b = \infty$, completing the proof in these cases.

Now suppose $a, b \in \mathbb{R} \setminus (D_f \cup D_g)$ and fix $\varepsilon > 0$. From (2.3) and (2.4), for all $x \in \mathbb{R}$, there exists a $\delta'_x > 0$ such that for all $\delta_x \leq \delta'_x$,

$$\sup_{z \in B(x, \delta_x)} f_N(z) \leq \sup_{z \in B(x, \delta_x)} f(z) + \varepsilon \quad \text{and} \quad \inf_{z \in B(x, \delta_x)} f_N(z) \geq \inf_{z \in B(x, \delta_x)} f(z) - \varepsilon \quad (2.6)$$

holds for all N large enough. Since f is càllàl, g has left and right limits for every $x \in \mathbb{R}$, and $D_f \cap D_g = \emptyset$, then for every $x \in [a, b]$, there exists $0 < \delta_x \leq \delta'_x$ such that for the open neighbourhood $B(x, \delta_x)$, one of the two holds:

(a) for all $y, z \in B(x, \delta_x)$, $|f(y) - f(z)| < \varepsilon$ and if $(y-x)(z-x) > 0$, then $|g(y) - g(z)| < \varepsilon$,
or

(b): for all $y, z \in B(x, \delta_x)$, $|g(y) - g(z)| < \varepsilon$ and if $(y-x)(z-x) > 0$, then $|f(y) - f(z)| < \varepsilon$.

The condition $(y-x)(z-x) > 0$ ensures y and z are on the same side of x . Since g and f have left and right limits for all $x \in \mathbb{R}$, $D_g \cup D_f$ is countable, and so we may also specify that δ_x is chosen in such a way that in addition to the above, both f and g are continuous at $x - \delta_x$ and $x + \delta_x$. Since $[a, b]$ is compact and $\{B(x, \delta_x)\}_{x \in [a, b]}$ is an open cover of $[a, b]$, there is a finite subcover $\{B(x_i, \delta_{x_i})\}_{i=1}^K$ of $[a, b]$. Let

$$B_i := (x_i - \delta_{x_i}, x_i + \delta_{x_i}] \cap (a, b],$$

so that $\bigcup B_i = (a, b]$. We show that for all $i = 1, \dots, K$,

$$\sup_{x \in B_i} (f(x) + g(x)) - 5\varepsilon \lesssim_{\mathbb{P}} \frac{\log \left(\int_{B_i} N^{f_N(x)} dM_N(x) \right)}{\log N} \lesssim_{\mathbb{P}} \sup_{x \in B_i} (f(x) + g(x)) + 3\varepsilon. \quad (2.7)$$

Let $s_i^f := \sup_{x \in B_i} f(x)$ and $s_i^g := \sup_{x \in B_i} g(x)$. By the definition of δ_{x_i} and the continuity of f and g at $b \wedge (x_i + \delta_{x_i, \varepsilon})$, we see that

$$\left| \sup_{x \in B_i} (f(x) + g(x)) - (s_i^f + s_i^g) \right| \leq 2\varepsilon. \quad (2.8)$$

For the upper bound of (2.7), using (2.6), since $a \vee (x_i - \delta_{x_i})$ and $b \wedge (x_i + \delta_{x_i})$ are continuity points of g , (2.5) guarantees that for all N large enough

$$\frac{\log \left(\int_{B_i} N^{f_N(x)} dM_N(x) \right)}{\log N} \leq \frac{\log \left(N^{s_i^f + \varepsilon} \int_{B_i} dM_N(x) \right)}{\log N} \xrightarrow{\mathbb{P}} s_i^f + s_i^g + \varepsilon,$$

which along with (2.8) proves the upper bound of (2.7).

For the lower bound, we consider the two cases (a) and (b) for $B(x_i, \delta_{x_i})$. Note that if $a = x_i$ or $b = x_i$, since f is continuous at a and b , we may assume $B(x_i, \delta_{x_i})$ is of type (a) in this case.

If $B(x_i, \delta_{x_i})$ is of type (a), then $\inf_{x \in B(x_i, \delta_{x_i})} f(x) \geq s_i^f - \varepsilon$, which along with (2.6) guarantees that for all N large enough, $f_N(x) \geq s_i^f - 2\varepsilon$ for all $x \in B(x_i, \delta_{x_i})$. By using (2.5) again,

$$\frac{\log \left(\int_{B_i} N^{f_N(x)} dM_N(x) \right)}{\log N} \geq \frac{\log \left(N^{s_i^f - 2\varepsilon} \int_{B_i} dM_N(x) \right)}{\log N} \xrightarrow{\mathbb{P}} s_i^f - 2\varepsilon + s_i^g,$$

which along with (2.8) completes the lower bound of (2.7) in this case.

Now suppose $B(x_i, \delta_{x_i})$ is of type (b). We define a new value y_i in the following way; if $f(x) \geq s_i^f - \varepsilon$ for all $x < x_i$ in B_i , let $y_i = x_i - \delta_{x_i}/2$, while if $f(x) \geq s_i^f - \varepsilon$ for all $x > x_i$ in B_i , let $y_i = x_i + \delta_{x_i}/2$ (at least one of these two cases must hold from the definition of $B(x_i, \delta_{x_i})$ and since f is either right continuous or left continuous at x_i). Repeating similar arguments as above, from (2.4), we may find $\delta_{y_i} \leq \delta_{x_i}/2$ such that $y_i - \delta_{y_i}, y_i + \delta_{y_i} \notin D_g$ and for all N large enough, $f_N(x) \geq s_i^f - 2\varepsilon$ for all $x \in B(y_i, \delta_{y_i})$. Since $B(x_i, \delta_{x_i})$ is of type (b),

$$\sup_{x \in B(y_i, \delta_{y_i})} g(x) \geq s_i^g - \varepsilon, \quad (2.9)$$

and so from (2.5) once more,

$$\frac{\log \left(\int_{B_i} N^{f_N(x)} dM_N(x) \right)}{\log N} \geq \frac{\log \left(N^{s_i^f - 2\varepsilon} \int_{B(y_i, \delta_{y_i})} dM_N(x) \right)}{\log N} \xrightarrow{\mathbb{P}} s_i^f - 2\varepsilon + \sup_{x \in B(y_i, \delta_{y_i})} g(x).$$

Substituting (2.9) and (2.8) into the above equation completes the lower bound of (2.7) in this case as well.

Suppose j is such that

$$\sup_{x \in B_j} (f(x) + g(x)) = \max_{i=1}^K \left\{ \sup_{x \in B_i} (f(x) + g(x)) \right\}.$$

With (2.7) in hand and since there are only finitely many B_i , we see that

$$\begin{aligned} \frac{\log \left(\int_{(a,b]} N^{f_N(x)} dM_N(x) \right)}{\log N} &\leq \frac{\log \left(\sum_{i=1}^K \int_{B_i} N^{f_N(x)} dM_N(x) \right)}{\log N} \\ &\leq \frac{\log(K) + \log \left(\max_{i=1}^K \int_{B_i} N^{f_N(x)} dM_N(x) \right)}{\log N} \\ &\lesssim_{\mathbb{P}} \sup_{x \in B_j} (f(x) + g(x)) + 3\varepsilon, \end{aligned}$$

while

$$\frac{\log \left(\int_{(a,b]} N^{f_N(x)} dM_N(x) \right)}{\log N} \geq \frac{\log \left(\int_{B_j} N^{f_N(x)} dM_N(x) \right)}{\log N} \gtrsim_{\mathbb{P}} \sup_{x \in B_j} (f(x) + g(x)) - 5\varepsilon.$$

Since $\bigcup B_i = (a, b]$, we see that

$$\sup_{x \in (a,b]} (f(x) + g(x)) = \sup_{x \in B_j} (f(x) + g(x)),$$

and so letting $\varepsilon \rightarrow 0$ proves the statement of the lemma in this case.

Now consider any $a, b \in \overline{\mathbb{R}} \setminus (D_f \cup D_g)$ such that $a < b$. Let $x_L < L$ and $x_U > U$ be continuity points of f and g . Then splitting the integrals

$$\begin{aligned} \int_{(a,b]} N^{f(x)} dM_N(x) &= \int_{(a \wedge x_L, x_L]} N^{f(x)} dM_N(x) \\ &\quad + \int_{(a \vee x_L, b \wedge x_U]} N^{f(x)} dM_N(x) + \int_{(x_U, b \vee x_U]} N^{f(x)} dM_N(x), \end{aligned}$$

and applying the previous cases proves the lemma for real-valued functions.

Now we allow $f_N : \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty\}$ and $f : \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty\}$. The statement of the lemma holds by truncating the functions and applying the argument above. For all $M \in \mathbb{R}$, let $f_N^M(x) = -M \vee f_N(x)$ and $f^M(x) = -M \vee f(x)$. It is evident that the conditions (2.3) and (2.4) hold for the truncated functions as well, and so in particular

$$\frac{\log \left(\int_{(a,b]} N^{f_N^M(x)} dM_N(x) \right)}{\log N} \xrightarrow{\mathbb{P}} \sup_{x \in (a,b]} (f^M(x) + g(x)) = \sup_{x \in (a,b]} (-M \vee f(x) + g(x)). \quad (2.10)$$

Since $g(x)$ is bounded above, if $\sup_{x \in (a,b]} (f(x) + g(x)) = -\infty$, then (2.10) tends to $-\infty$ as M increases, while if $\sup_{x \in (a,b]} (f(x) + g(x)) \in \mathbb{R}$, then (2.10) is precisely equal to this value for all M large enough. In either case, letting $M \rightarrow \infty$ completes the proof. $\square$

2.1.2 Concentration inequalities

We adapt classic concentration inequalities to be applied for our results. We will use the following version of McDiarmid's inequality.

Lemma 2.4 (McDiarmid, [20]). *Let $X_1, \dots, X_n$ be independent random variables with $X_k \in [0, 1]$ for all $k \in [n]$. Let $S_n := \sum_{k=1}^n X_k$ and let $\mu = \mathbb{E}[S_n]$. Then for any $\epsilon \in (0, 1)$,*

$$\mathbb{P}(|S_n - \mu| > \epsilon\mu) < e^{-\frac{1}{4}\epsilon^2\mu}.$$

Let $(\ell_N)_{N \in \mathbb{N}}$ and $(a_j^N)_{N,j \in \mathbb{N}}$ be sequences of random variables. For all $N \in \mathbb{N}$, let $\mathcal{F}_N$ be a σ -algebra such that $\ell_N, a_1^N, \dots, a_{\ell_N}^N \in \mathcal{F}_N$. Let $(\xi_j^N)_{N,j \in \mathbb{N}}$ be random variables with $\xi_j^N \in \{0, 1\}$ for all $j, N \in \mathbb{N}$. Suppose that, conditioned on $\mathcal{F}_N$, the random variables $(\xi_j^N)_{N,j \in \mathbb{N}}$ are independent, and they are distributed as $\text{Ber}(a_j^N \wedge 1)$. We also let $S_{\ell_N} := \sum_{j=1}^{\ell_N} \xi_j^N$ and $\mathbb{E}[S_{\ell_N} | \mathcal{F}_N] =: \mu_N$ for all $N \in \mathbb{N}$.

Lemma 2.5. (a) *If there exists a constant $C_1 > 0$ such that $\frac{\log \mu_N}{\log N} \xrightarrow{\mathbb{P}} C_1$, then*

$$\frac{\log S_{\ell_N}}{\log N} \xrightarrow{\mathbb{P}} C_1$$

as $N \rightarrow \infty$.

(b) *If there exists a constant $C_2 \in \mathbb{R}$ such that $\frac{\log \mu_N}{\log N} \lesssim_{\mathbb{P}} C_2$, then, as $N \rightarrow \infty$,*

$$\frac{\log S_{\ell_N}}{\log N} \lesssim_{\mathbb{P}} C_2, \quad \text{if } C_2 \geq 0,$$

and

$$S_{\ell_N} \xrightarrow{\mathbb{P}} 0, \quad \text{if } C_2 < 0.$$

Proof. We first prove part (a). Take $0 < \varepsilon < 1 \wedge C_1/2$. It is enough to show that the probability of the event $\{S_{\ell_N} \notin [N^{C_1-2\varepsilon}, N^{C_1+2\varepsilon}]\}$ converges to zero as $N \rightarrow \infty$. Splitting this event based on the value of μ_N we get

$$\begin{aligned} \mathbb{P}(S_{\ell_N} \notin [N^{C_1-2\varepsilon}, N^{C_1+2\varepsilon}]) &\leq \mathbb{P}(S_{\ell_N} \notin [N^{C_1-2\varepsilon}, N^{C_1+2\varepsilon}], \mu_N \in [N^{C_1-\varepsilon}, N^{C_1+\varepsilon}]) \\ &\quad + \mathbb{P}(\mu_N \notin [N^{C_1-\varepsilon}, N^{C_1+\varepsilon}]). \end{aligned} \quad (2.11)$$

Notice that for N sufficiently large we have

$$\{S_{\ell_N} + 1 \notin [N^{C_1-2\varepsilon}, N^{C_1+2\varepsilon}], \mu_N \in [N^{C_1-\varepsilon}, N^{C_1+\varepsilon}]\} \subseteq \left\{|S_{\ell_N} - \mu_N| > \frac{\mu_N}{2}\right\}.$$

Note furthermore, that the second term on the right-hand side of (2.11) converges to zero as $N \rightarrow \infty$ by the assumption of the lemma on μ_N . Hence, continuing (2.11), and first conditioning on $\mathcal{F}_N$, and then applying Lemma 2.4, for N sufficiently large we obtain

$$\begin{aligned} \mathbb{P}(S_{\ell_N} \notin [N^{C_1-2\varepsilon}, N^{C_1+2\varepsilon}]) &\leq \mathbb{E} \left[\mathbb{P} \left(|S_{\ell_N} - \mu_N| > \frac{\mu_N}{2} \mid \mathcal{F}_N \right) \mathbf{1}_{\{\mu_N \in [N^{C_1-\varepsilon}, N^{C_1+\varepsilon}]\}} \right] + o(1) \\ &\leq \mathbb{E} \left[e^{-\frac{\mu_N}{16}} \mathbf{1}_{\{\mu_N \in [N^{C_1-\varepsilon}, N^{C_1+\varepsilon}]\}} \right] + o(1) \\ &\leq \mathbb{E} \left[e^{-N^{C_1-\varepsilon}/16} \right] + o(1), \end{aligned}$$

which converges to zero as $N \rightarrow \infty$, showing part (a).

We now move on to part (b). Take $0 < \varepsilon < 1 \wedge |C_2|/2$. Similarly to part (a), in this case we have

$$\mathbb{P}(S_{\ell_N} > N^{C_2+2\varepsilon}) \leq \mathbb{E}[\mathbb{P}(S_{\ell_N} > N^{C_2+2\varepsilon} \mid \mathcal{F}_N) \mathbf{1}_{\mu_N < N^{C_2+\varepsilon}}] + \mathbb{P}(\mu_N > N^{C_2+\varepsilon}).$$

By Markov's inequality and our assumption on μ_N , and recalling that $\mu_N = \mathbb{E}[S_{\ell_N} \mid \mathcal{F}_N]$, we arrive at

$$\mathbb{P}(S_{\ell_N} > N^{C_2+2\varepsilon}) \leq \mathbb{E}\left[\frac{\mu_N \mathbf{1}_{\{\mu_N < N^{C_2+\varepsilon}\}}}{N^{C_2+2\varepsilon}}\right] + o(1) \leq N^{-\varepsilon} + o(1) \rightarrow 0, \quad (2.12)$$

as $N \rightarrow \infty$. Note that (2.12) holds for any $C_2 \in \mathbb{R}$. If $C_2 \geq 0$, then this shows the first statement of part (b). If $C_2 < 0$, then the second statement follows from (2.12) and from the fact that S_{ℓ_N} takes integer values. $\square$

2.2 Limiting log-profiles after reproduction

For a function $g \in \mathcal{D}$, define $\bar{r}(g)$ and $r(g)$ by

$$\bar{r}(g)(x) := 1 - \gamma + \sup_{z \in \mathbb{R}} (g(z) + h(x - z)), \quad \text{and} \quad r(g)(x) := \pi(\bar{r}(g)(x)). \quad (2.13)$$

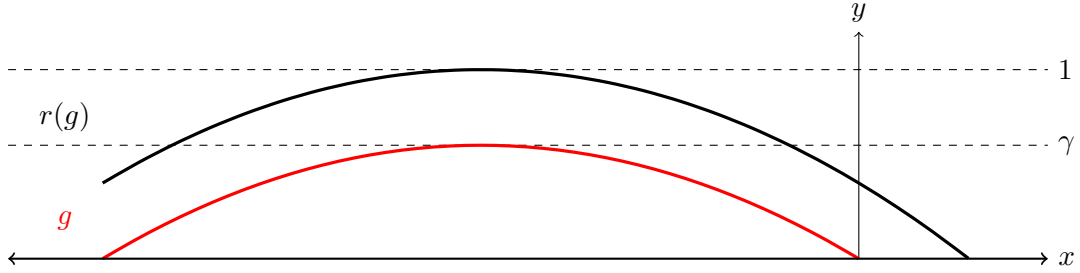


Figure 3: A potential outcome of applying the operator r to the log-profile g (in red), producing the functions $r(g)$ (in black).

Let $(X_{i,j} : i \in [N^\gamma], j \in [N^{1-\gamma}])$ be an array of i.i.d. random variables distributed as X . Consider a sequence of point measures $(\mathcal{Z}_N = (Z_{N,i})_{i=1}^{N^\gamma})_{N \in \mathbb{N}}$. We assume that the sequence of measure

$$M_N := \sum_{i=1}^{N^\gamma} \delta_{Z_{N,i}/c_N}$$

has a limiting log-profile $g \in \mathcal{D}$. Finally, we define the profile after reproduction as

$$M_N^R = \sum_{i=1}^{N^\gamma} \sum_{j=1}^{N^{1-\gamma}} \delta_{(Z_{N,i} + X_{i,j})/c_N}.$$

The goal of this section is to prove the following Lemma:

Lemma 2.6. *The sequence of measures $(M_N^R)_{N \in \mathbb{N}}$ has limiting log-profile $r(g) \in \mathcal{D}$.*

To prove Lemma 2.6 we make use of Lemma 2.2 and Lemma 2.5 to prove a convergence in probability after scaling for $M_N^R(a, b)$. However, to apply Lemma 2.2, certain continuity conditions at a and b must be satisfied, while we require that the supremum of $\bar{r}$ in the interval $(a, b]$ be nonzero to apply Lemma 2.5. Therefore, we first apply our probabilistic argument where these conditions are met to achieve Lemma 2.7 below, and finally use some technical continuity arguments to finish the proof of Lemma 2.6.

Lemma 2.7. *For all $a, b \in \mathbb{R} \setminus D_g$, if $\sup_{x \in (a, b]} \bar{r}(g)(x) \neq 0$, then*

$$\frac{\log M_N^R(a, b)}{\log N} \xrightarrow{\mathbb{P}} \sup_{x \in (a, b]} r(g)(x).$$

Proof. Set $\bar{r} := \bar{r}(g)$ and $r := r(g)$. We have

$$M_N^R(a, b) = \int_{(ac_N, bc_N]} \sum_{i=1}^{N^\gamma} \sum_{j=1}^{N^{1-\gamma}} \delta_{X_{ij} + Z_{N,i}}(dx) = \sum_{i=1}^{N^\gamma} \sum_{j=1}^{N^{1-\gamma}} \mathbb{1}_{\{X_{i,j} \in (ac_N - Z_{N,i}, bc_N - Z_{N,i}]\}}. \quad (2.14)$$

We first show that for all $a, b \in \mathbb{R} \setminus D_g$ with $a < b$,

$$A_N(a, b) := \frac{\log(\mathbb{E}[M_N^R(a, b) \mid \sigma(\mathcal{Z}_N)])}{\log N} \xrightarrow{\mathbb{P}} \sup_{x \in (a, b]} \bar{r}(x). \quad (2.15)$$

For all $a < b$, define

$$p_N(a, b) := \frac{\log(F_X(bc_N) - F_X(ac_N))}{\log N}.$$

From our assumptions of the tail of X , one can verify that for all $z \in \mathbb{R}$,

$$p_N(a - z, b - z) \xrightarrow{N \rightarrow \infty} H_{a,b}(z) := \sup_{x \in (a - z, b - z]} h(x),$$

and that for $a - z, b - z \neq 0$,

$$\begin{aligned} \lim_{\delta \rightarrow 0} \lim_{N \rightarrow \infty} \sup_{x \in B(z, \delta)} p_N(a - x, b - x) &= \lim_{\delta \rightarrow 0} \sup_{x \in B(z, \delta)} H_{a,b}(x), \\ \lim_{\delta \rightarrow 0} \lim_{N \rightarrow \infty} \inf_{x \in B(z, \delta)} p_N(a - x, b - x) &= \lim_{\delta \rightarrow 0} \inf_{x \in B(z, \delta)} H_{a,b}(x). \end{aligned} \quad (2.16)$$

The function $H_{a,b}(z)$ has at most two points of discontinuity (when $z = a$ or $z = b$) and is càllàl. Since all $X_{i,j}$ are i.i.d., (2.14) gives us

$$\begin{aligned} A_N(a, b) &= \frac{1}{\log N} \log \left(\sum_{i=1}^{N^\gamma} N^{1-\gamma} \mathbb{P}(X \in (ac_N - Z_{N,i}, bc_N - Z_{N,i}] \mid \sigma(\mathcal{Z}_N)) \right) \\ &= \frac{1}{\log N} \log \left(\int_{\mathbb{R}} N^{1-\gamma} N^{p_N(a-z, b-z)} dM_N(z) \right). \end{aligned}$$

From (2.16), so long as $a, b \notin D_g$ (and so $D_{H_{a,b}} \cap D_g = \emptyset$), we may apply Lemma 2.2 and see that

$$A_N(a, b) \xrightarrow{\mathbb{P}} \sup_{z \in \mathbb{R}} (1 - \gamma + g(z) + H_{a,b}(z)).$$

Observe that by the definition of the function $H_{a,b}$ and by swapping the two suprema, we have

$$\begin{aligned} \sup_{z \in \mathbb{R}} (1 - \gamma + g(z) + H_{a,b}(z)) &= \sup_{z \in \mathbb{R}} \left(1 - \gamma + g(z) + \sup_{x \in (a-z, b-z]} h(x) \right) \\ &= \sup_{x \in (a,b]} \left(1 - \gamma + \sup_{z \in \mathbb{R}} \{g(z) + h(x-z)\} \right) \\ &= \sup_{x \in (a,b]} \bar{r}(x), \end{aligned}$$

and so (2.15) holds.

Now apply Lemma 2.5 with $\ell_N = N$, $a_{i,j}^N = \mathbb{P}(X \in (ac_N - Z_{N,i}, bc_N - Z_{N,i}])$ for all $i \in [N^\gamma]$ and $j \in [N^{1-\gamma}]$, $\xi_{i,j} = \mathbf{1}_{\{X_{i,j} \in (ac_N - Z_{N,i}, bc_N - Z_{N,i}]\}}$, with $\mathcal{F}_N = \sigma(\mathcal{Z}_N)$. By Lemma 2.5(a) we conclude that

$$\frac{\log M_N^R(a, b)}{\log N} \xrightarrow{\mathbb{P}} \sup_{x \in (a,b]} \bar{r}(x) = \sup_{x \in (a,b]} r(x),$$

if $\sup_{x \in (a,b]} \bar{r}(x) > 0$; and by Lemma 2.5(b) we see that

$$M_N^R(a, b) \xrightarrow{\mathbb{P}} 0,$$

if $\sup_{x \in (a,b]} \bar{r}(x) < 0$, which concludes the proof. $\square$

Proof of Lemma 2.6. Once more let $\bar{r} := \bar{r}(g)$ and $r := r(g)$. It is an elementary argument to show that $r(g) \in \mathcal{D}$, that is left to the reader. We now use the previous lemma to prove that r is the limiting log-profile of M_N^R .

Assume that $a, b \in \mathbb{R}$ such that $\sup_{x \in (a,b]} \bar{r}(x) = 0$. We show that either a or b must be a discontinuity point of r . Recalling h from (1.2), $h^{-1}(\{x \geq -(1-\gamma)\})$ is a closed interval containing 0, call it $[-p, q]$. Note that $q > 0$ and that $p > 0$ if $c_- \neq \infty$ and $p = 0$ otherwise, while $h(x) > -(1-\gamma)$ for all $x \in (-p, q)$. Therefore, for any $x \in \mathbb{R}$,

if $g(z) \geq 0$ for some $z \in (x-q, x+p)$, then $\bar{r}(x) \geq 1 - \gamma + \sup_{z \in (x-q, x+p)} (g(z) + h(x-z)) > 0$.

Fix $x \in (a, b]$. Since $\bar{r}(x) \leq 0$, the above implies $g(z) = -\infty$ for all $x \in (x-q, x+p)$ and

$$\bar{r}(x) = 1 - \gamma + \sup_{z \leq x-q} (g(z) + h(x-z)) \quad \text{or} \quad \bar{r}(x) = 1 - \gamma + \sup_{z \geq x+p} (g(z) + h(x-z)). \quad (2.17)$$

Suppose $\bar{r}(x) \neq -\infty$ and is given by the first expression of (2.17). Since the function h is strictly decreasing on $\mathbb{R}_+$ and g has bounded support, we see that for all $q > \delta > 0$,

$$\sup_{z \in \mathbb{R}} (g(z) + h(x - \delta - z)) \geq \sup_{z \leq x-q} (g(z) + h(x - \delta - z)) > \sup_{z \leq x-q} (g(z) + h(x - z)),$$

and so $\bar{r}(x - \delta) > \bar{r}(x)$. If $\bar{r}(x) \neq -\infty$ and is given by the second expression of (2.17), then $c_- \neq \infty$, and since the function h is strictly increasing on $\mathbb{R}_-$, a similar argument shows that $\bar{r}(x + \delta) > \bar{r}(x)$ for all $p > \delta > 0$. In all cases, we see that since $\sup_{x \in (a,b]} \bar{r}(x) = 0$,

$$\forall \delta > 0 \text{ small enough,} \quad \sup_{x \in (a+\delta, b-\delta)} \bar{r}(x) < 0 \quad \text{and} \quad \sup_{x \in (a-\delta, b+\delta)} \bar{r}(x) > 0. \quad (2.18)$$

The function π sends negative values to $-\infty$, and so (2.18) implies that at least one of a or b is a discontinuity point of $r = \pi(\bar{r})$. Therefore,

$$a, b \in \mathbb{R} \setminus D_r \implies \sup_{x \in (a, b]} \bar{r}(x) \neq 0. \quad (2.19)$$

Now take any continuity points $a, b \in \mathbb{R} \setminus D_r$. Since $D_g \cup D_r$ is countable, we may find $a^+ > a$ and $b^- < b$ such that $a^+, b^- \notin D_g \cup D_r$ and

$$\sup_{x \in (a^+, b^-]} r(x) = \sup_{x \in (a, b]} r(x) = -\infty, \quad \text{or} \quad \sup_{x \in (a, b]} r(x) - \sup_{x \in (a^+, b^-]} r(x) < \varepsilon.$$

Since $\sup_{x \in (a^+, b^-]} r(x) \neq 0$, we use Lemma 2.7 and get

$$\frac{\log M_N^R(a, b)}{\log N} \geq \frac{\log M_N^R(a^+, b^-)}{\log N} \xrightarrow{\mathbb{P}} \sup_{x \in (a^+, b^-]} r(x).$$

A similar argument for an upper bound, and letting $\varepsilon \rightarrow 0$, we see that indeed

$$\frac{\log M_N^R(a, b)}{\log N} \xrightarrow{\mathbb{P}} \sup_{x \in (a, b]} r(x).$$

This holds for all $a, b \in \mathbb{R} \setminus D_r$, and so along with (2.19), r is the limiting log-profile of $(M_N^R)_{N \in \mathbb{N}}$, as desired. $\square$

2.3 Limiting log-profiles after sampling

Let $r \in \mathcal{D}$. For any $\sigma \in \mathbb{R}$, define $\bar{s}_\sigma(r)$ and $s_\sigma(r)$ by

$$\bar{s}_\sigma(r)(x) := r(x) + \beta(x - \sigma)_-, \quad \text{and} \quad s_\sigma(r)(x) := \pi(\bar{s}_\sigma(r)(x)). \quad (2.20)$$

We then define

$$\sigma^*(r) := \inf \left\{ \sigma : \sup_{z \in \mathbb{R}} s_\sigma(r)(z) \leq \gamma \right\} \quad (2.21)$$

and finally, we define $\bar{s}(r)$ and $s(r)$ by

$$\bar{s}(r) := \bar{s}_{\sigma^*(r)}(r), \quad \text{and} \quad s(r) := s_{\sigma^*(r)}(r) = \pi(\bar{s}(r)). \quad (2.22)$$

Let $\mathcal{R}_N = (R_{N,j})_{j \in [N]}$ denote a random set of points. We think of $\mathcal{R}_N$ as the point set after the reproduction step and we wish to sample N^γ particles $\mathcal{S}_N = (S_{N,i})_{i \in [N^\gamma]}$ from this set without replacement, where each particle $R_{N,j}$ has relative weight

$$w_{N,j} := e^{\beta_N R_{N,j}}, \quad (2.23)$$

and β_N is given by (1.4). We define the sequences of measures

$$M_N^R := \sum_{j=1}^N \delta_{R_{N,j}/c_N}, \quad \text{and} \quad M_N^S := \sum_{i=1}^{N^\gamma} \delta_{S_{N,i}/c_N},$$

and assume that $(M_N^R)_{N \in \mathbb{N}}$ has limiting log-profile $r \in \mathcal{D}$. Our goal in this section is to prove the following:

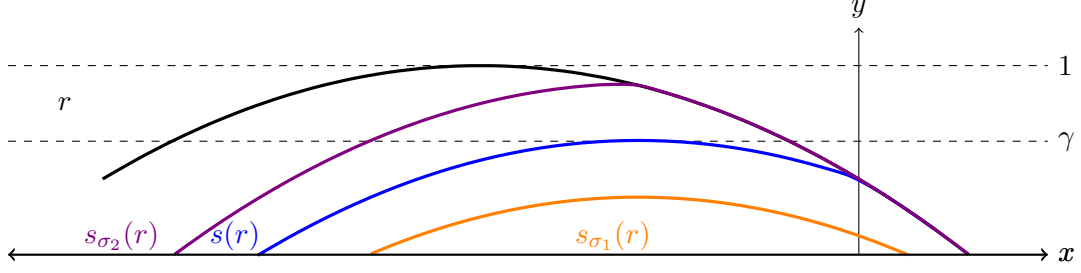


Figure 4: Applications of the operator s_σ to r (in black). The functions $s_\sigma(r)$ are portrayed for three values $\sigma_1 > \sigma^*(r) > \sigma_2$, where $s_{\sigma_1}(r)$ is in violet, $s(r) = s_{\sigma^*(r)}(r)$ is in blue, and $s_{\sigma_2}(r)$ is in orange.

Lemma 2.8. Assume that $\sigma^* := \sigma^*(r)$ is the unique value of σ such that $\sup s_\sigma(r) = \gamma$. For all $a, b \in \mathbb{R} \setminus D_r$ such that $\sup_{x \in (a,b]} \bar{s}(r)(x) \neq 0$,

$$\frac{\log M_N^S(a, b)}{\log N} \xrightarrow{\mathbb{P}} \sup_{x \in (a,b]} s(r)(x). \quad (2.24)$$

The conditions imposed on σ^* and $\bar{s}$ in Lemma 2.8 are required to guarantee the existence of a limiting log-profile for M_N^S . We explore what can fail if these conditions are not satisfied in Section 2.4.1.

To prove Lemma 2.8, we perform our sampling without replacement using exponential clocks: a clock associated to a particle at location $R_{N,j}$ is an exponential random variable $\tau_{N,j} \sim \text{Exp}(w_{N,j})$ for $w_{N,j}$ defined in (2.23). Then the particles corresponding to the N^γ smallest values of the $\tau_{N,j}$'s (i.e. the N^γ particles whose clocks rang the fastest) are selected. In analyzing the measures M_N^S we will first investigate the number of clocks that have 'gone off' by a certain time. For any $\sigma \in \mathbb{R}$, define

$$M_{N,\sigma}^S := \sum_{j=1}^N \delta_{R_{N,j}/c_N} \mathbf{1}_{\{\tau_{N,j} \leq N^{-\beta\sigma}\}}.$$

Lemma 2.9. Suppose the sequence of point measures $(M_N^R)_{N \in \mathbb{N}}$ has limiting log-profile $r \in \mathcal{D}$. For all $a, b \in \mathbb{R} \setminus D_r$, if $\sup_{x \in (a,b]} \bar{s}_\sigma(r)(x) \neq 0$, then

$$\frac{\log M_{N,\sigma}^S(a, b)}{\log N} \xrightarrow{\mathbb{P}} \sup_{x \in (a,b]} s_\sigma(r)(x).$$

Proof. Throughout we let $\bar{s}_\sigma := \bar{s}_\sigma(r)$ and $s_\sigma := s_\sigma(r)$. For a particle at location $R_{N,j} = xc_N$, we have $w_{N,j} = N^{\beta x}$, and

$$\mathbb{P}(\tau_{N,j} \leq N^{-\beta\sigma} | \sigma(\mathcal{R}_N)) = 1 - \exp(-w_{N,j} N^{-\beta\sigma}) = 1 - \exp(-N^{\beta(x-\sigma)}). \quad (2.25)$$

For all $a, b \in \mathbb{R}$, if we let $B_N(a, b) := \mathbb{E}[M_{N,\sigma}^S(a, b) | \sigma(\mathcal{R}_N)]$, then

$$B_N(a, b) = \sum_{R_{N,j} \in (ac_N, bc_N]} \mathbb{P}(\tau_{N,j} \leq N^{-\beta\sigma} | \sigma(\mathcal{R}_N)) = \int_{(a,b]} \left(1 - \exp(-N^{\beta(x-\sigma)})\right) dM_N^R(x). \quad (2.26)$$

It is easy to see that

$$\frac{\log(1 - \exp(-N^{\beta(x-\sigma)}))}{\log N} \xrightarrow{N \rightarrow \infty} \beta(x - \sigma)_-$$

uniformly in x . From Remark 2.3, we may apply Lemma 2.2, and so for all $a, b \in \mathbb{R} \setminus D_r$,

$$B_N(a, b) \xrightarrow{\mathbb{P}} \sup_{x \in (a, b]} \bar{s}_\sigma(x), \text{ as } N \rightarrow \infty. \quad (2.27)$$

Now we apply Lemma 2.5 with $\ell_N = N$, $a_j^N = (1 - \exp(-w_{N,j}N^{-\beta\sigma}))\mathbb{1}_{\{R_{N,j} \in (ac_N, bc_n]\}}$, $\xi_j^N = \mathbb{1}_{\{\tau_{N,j} \leq N^{-\beta\sigma}\}}$, and $\mathcal{F}_N = \sigma(\mathcal{R}_N)$. By (2.27) and Lemma 2.5(a), we conclude that

$$\frac{\log M_{N,\sigma}^S(a, b)}{\log N} \xrightarrow{\mathbb{P}} \sup_{x \in (a, b]} \bar{s}_\sigma(x) = \sup_{x \in (a, b]} s_\sigma(x), \text{ as } N \rightarrow \infty,$$

if $\sup_{x \in (a, b]} \bar{s}_\sigma(x) > 0$, and by Lemma 2.5(b), we see that

$$M_{N,\sigma}^S \xrightarrow{\mathbb{P}} 0, \text{ as } N \rightarrow \infty,$$

if $\sup_{x \in (a, b]} \bar{s}_\sigma(x) < 0$, which concludes the proof. $\square$

For all $N \in \mathbb{N}$, let us write

$$\sigma_N^* := \sigma_N^*(\mathcal{R}_N) = \inf \left\{ \sigma : M_{N,\sigma}^S(\mathbb{R}) \leq N^\gamma \right\} \quad (2.28)$$

as an analog to $\sigma^*(r)$ from (2.21). Then as discussed above, $M_{N,\sigma_N^*}^S(a, b) = M_N^S(a, b)$.

Proof of Lemma 2.8. Let $s_\sigma := s_\sigma(r)$ and $s := s(r)$. Since we assume σ^* is the unique value for which $\sup s_{\sigma^*} = \gamma$, then for all $\varepsilon > 0$ and by Lemma 2.9,

$$\frac{\log M_{N,\sigma^*-\varepsilon}^S(-\infty, \infty)}{\log N} \xrightarrow{\mathbb{P}} \sup s_{\sigma^*-\varepsilon} > \gamma$$

and so from the definition of σ_N^* , as $N \rightarrow \infty$,

$$\begin{aligned} \mathbb{P}(\sigma^* - \varepsilon > \sigma_N^*) &\leq \mathbb{P}(M_{N,\sigma^*-\varepsilon}^S(-\infty, \infty) \leq M_N^S(-\infty, \infty) = N^\gamma) \\ &\leq \mathbb{P}\left(\frac{\log M_{N,\sigma^*-\varepsilon}^S(-\infty, \infty)}{\log N} \leq \gamma\right) \rightarrow 0. \end{aligned}$$

By a symmetric argument, $\mathbb{P}(\sigma_N^* > \sigma^* + \varepsilon) \xrightarrow{N \rightarrow \infty} 0$, and so $\sigma_N^* \xrightarrow{\mathbb{P}} \sigma^*$ as $N \rightarrow \infty$. As a consequence, for all $a, b \in \mathbb{R}$,

$$\mathbb{P}(M_{N,\sigma^*+\varepsilon}^S(a, b) \leq M_N^S(a, b) \leq M_{N,\sigma^*-\varepsilon}^S(a, b)) \geq \mathbb{P}(\sigma^* - \varepsilon \leq \sigma_N^* \leq \sigma^* + \varepsilon) \xrightarrow{N \rightarrow \infty} 1. \quad (2.29)$$

Note that for any $x, \sigma \in \mathbb{R}$ and any $\varepsilon > 0$, $|\bar{s}_\sigma(x) - \bar{s}_{\sigma \pm \varepsilon}(x)| \leq \varepsilon$, and so it is evident that for all $a, b \in \mathbb{R} \setminus D_r$ such that $\sup_{x \in (a, b]} \bar{s}(x) \neq 0$, the function $\sigma \mapsto \sup_{x \in (a, b]} s_\sigma(x)$ is continuous around σ^* . Then applying Lemma 2.9 and letting $\varepsilon \rightarrow 0$ in (2.29) completes the proof. $\square$

2.4 Proof of Theorem 1.3

Recall that M_N^t is the point measure associated with $\mathcal{Z}_N^t = (Z_{N,i}^t)_{i \in [N\gamma]}$, see (1.9). For $t \in \mathbb{N}$, we introduce $M_N^{R,t}$ and $M_N^{S,t} = M_N^{t+1}$ as the analog point measures after reproduction and selection (between times t and $t+1$) as defined in Sections 2.2 and 2.3. We also recall the definitions of the operators, $\bar{r}, r$ defined in (2.13) and $\bar{s}, s$ defined in (2.22). First, we prove a more general theorem with technical conditions, which we then apply to prove Theorem 1.3.

Theorem 2.10. *Assume that $(M_N^0)_{N \in \mathbb{N}}$ has limiting log-profile $g^0 \in \mathcal{D}$ and let $(g^t)_{t \geq 0}$ be the dynamics as defined in Theorem 1.3. Suppose further that*

(i) $g^t \in \mathcal{D}$,

(ii) for all $t \geq 0$,

$$\forall a, b \in \mathbb{R} \setminus D_{g^{t+1}}, \quad \sup_{x \in (a,b]} \bar{s}(r^t)(x) \neq 0, \quad (2.30)$$

(iii) σ^t is the unique σ for which $\sup s_\sigma(r^t) = \gamma$.

Then for all $t \in \mathbb{N}$, the sequences of point measures $(M_N^t)_{N \in \mathbb{N}}$ have limiting log-profile g^t .

Proof. Let $t \in \mathbb{N}$. Suppose the result is true for all $0 \leq n \leq t$. By Lemma 2.6, $(M_N^{R,t})_{N \in \mathbb{N}}$ has limiting log-profile r^t . Let $a, b \in \mathbb{R} \setminus D_{g^{t+1}}$, then (2.30) guarantees that $\sup_{x \in (a,b]} \bar{s}(r^t)(x) \neq 0$. For any $\varepsilon > 0$, we may find $a^-, b^+ \in \mathbb{R} \setminus D_{r^t}$ such that $a^- \leq a$, $b \leq b^+$,

$$\sup_{x \in (a^-, b^+]} \bar{s}(r^t)(x) \neq 0, \quad \text{and} \quad \sup_{x \in (a^-, b^+]} g^{t+1}(x) - \sup_{x \in (a,b]} g^{t+1}(x) \leq \varepsilon. \quad (2.31)$$

We assumed that σ^t is the unique value for which $\sup s_\sigma(r^t) = \gamma$, and so from (2.31) and Lemma 2.8,

$$\frac{\log M_N^{t+1}(a, b)}{\log N} \leq \frac{\log M_N^{t+1}(a^-, b^+)}{\log N} \xrightarrow{\mathbb{P}} \sup_{x \in (a^-, b^+]} g^{t+1}(x).$$

A symmetric argument holds for a lower bound, and letting $\varepsilon \rightarrow 0$, we conclude that

$$\frac{\log M_N^{t+1}(a, b)}{\log N} \xrightarrow{\mathbb{P}} \sup_{x \in (a,b]} g^{t+1}(x).$$

The above holds for all $a, b \in D_{g^{t+1}}$, and so the sequence of point measures $(M_N^{t+1})_{N \in \mathbb{N}}$ has limiting log-profile g^{t+1} . Recursively applying the above argument completes the proof. $\square$

In order to prove Theorem 1.3 we will need one more technical lemma. Recall the definition of the function class $\mathcal{C}$ from Definition 1.1.

Lemma 2.11. *Suppose h from (1.2) is concave ($\rho \geq 1$). If $g \in \mathcal{C}$ and $\sup g = \gamma$, then $\bar{r}(g), \bar{s}(r(g))$ are concave and*

$$r(g), s(r(g)) \in \mathcal{C}.$$

Proof. Set $\bar{r} := \bar{r}(g)$ and $r := r(g)$. For any $x, x' \in \mathbb{R}$, by using the concavity of g and h ,

$$\begin{aligned}\bar{r}(g)(\alpha x + (1 - \alpha)x') &= 1 - \gamma + \sup_{z \in \mathbb{R}} \{g(z) + h(\alpha x + (1 - \alpha)x' - z)\} \\ &= 1 - \gamma + \sup_{z, z' \in \mathbb{R}} \{g(\alpha z + (1 - \alpha)z') + h(\alpha(x - z) + (1 - \alpha)(x' - z'))\} \\ &\geq 1 - \gamma + \alpha \sup_{z \in \mathbb{R}} \{g(z) + h(x - z)\} + (1 - \alpha) \sup_{z' \in \mathbb{R}} \{g(z') + h(x' - z')\} \\ &= \alpha \bar{r}(x) + (1 - \alpha) \bar{r}(x'),\end{aligned}$$

and so $\bar{r}$ is concave. Since π is non-decreasing and concave, then $r = \pi(\bar{r})$ is also concave. Finally, the function $x \mapsto \beta(x - \sigma^*(r))_-$ is concave, and so $\bar{s}(r)(x) = r(g)(x) + \beta(x - \sigma^*(r))_-$ is also concave, while once more, the composition $\pi(\bar{s}(r)) = s(r)$ is also concave. $\square$

Proof of Theorem 1.3. By Lemma 2.11, for all $t \geq 0$, we have that $r^t, g^t \in \mathcal{C} \subset \mathcal{D}$. To apply Theorem 2.10, we then only need to show that (2.30) holds and that there is a unique σ^t for which $\sup s_{\sigma^t}(r^t) = \gamma$. For $t \in \mathbb{N}$ let $L^t := L(g^t)$ and $U^t := U(g^t)$ from Definition 1.1.

From Lemma 2.11, $\bar{s}(r^t)$ and g^{t+1} are concave, and are therefore continuous on the interior of their support, so $D_{g^{t+1}} = \{L^{t+1}, U^{t+1}\}$. By the concavity of $\bar{s}(r^t)$, we also see that $\bar{s}(r^t)(x) < 0$ for all $x \in \mathbb{R} \setminus [L^{t+1}, U^{t+1}]$, and since $\bar{s}(r^t) > 0$ on (L^{t+1}, U^{t+1}) we see that

$$a, b \in \mathbb{R} \setminus \{L^{t+1}, U^{t+1}\} \implies \sup_{x \in (a, b]} \bar{s}(r^t)(x) \neq 0,$$

and so (2.30) holds.

We now prove the uniqueness of σ^t . Suppose otherwise that there exists two values $\sigma_1 < \sigma_2$ such that $\sup s_{\sigma_1}(r^t) = \sup s_{\sigma_2}(r^t) = \gamma$. By the definition of the operator s_{σ_1} ,

$$\gamma \geq \sup_{x \in (\sigma_1, \infty)} s_{\sigma_1}(r^t)(x) = \sup_{x \in (\sigma_1, \infty)} r^t(x).$$

Since r^t is concave and $\sup r^t = 1 > \gamma$, the above implies that r^t is strictly decreasing for $x \in (\sigma_1, \infty) \cap \text{Supp}(r^t)$. Therefore, if we let $\hat{\sigma}$ be such that $\sigma_1 < \hat{\sigma} < \sigma_2$, then

$$\sup_{x \in (\hat{\sigma}, \infty)} s_{\sigma_2}(r^t)(x) \leq \sup_{x \in (\hat{\sigma}, \infty)} r^t(x) < \gamma \quad (2.32)$$

and

$$\sup_{x \in (\sigma_1, \hat{\sigma}]} s_{\sigma_2}(r^t)(x) = \sup_{x \in (\sigma_1, \hat{\sigma}]} \pi(r^t(x) + \beta(x - \sigma_2)) \leq \sup_{x \in (\sigma_1, \hat{\sigma}]} r^t(x) + \beta(\hat{\sigma} - \sigma_2) < \gamma. \quad (2.33)$$

Now let $x \in (-\infty, \sigma_1]$. Since $\beta(x - \sigma_2) < \beta(x - \sigma_1)$, then $s_{\sigma_2}(r^t)(x) < s_{\sigma_1}(r^t)(x) \leq \gamma$ on $\text{Supp}(s_{\sigma_2}(r^t))$, while $s_{\sigma_2}(r^t)(x) = -\infty$ otherwise. Along with (2.32) and (2.33), we see that $\sup s_{\sigma_2}(r^t) < \gamma$, and a contradiction is reached. This completes the proof of Theorem 2.10. $\square$

2.4.1 Example with no limiting log-profile after selection

Example 2.12. We now construct a counterexample where $\mathcal{S}_N$ – the point measure after selection step – does not admit a limiting log-profile coinciding with the prediction of Theorem 1.3. To do so, we consider a population whose log-profile after selection has an isolated 0, so that it becomes impossible to deduce the presence or absence of particles from the log-profile alone.

Let $\rho \in (0, 1)$, and X be a random variable with probability density function $f_X = c(\rho) \exp(-|x|^\rho)$ for some normalizing constant $c(\rho)$, so that $h(x) = -|x|^\rho$. We see that $c_N \approx (\ln N)^{1/\rho}$. Starting with N^γ particles at 0, the population after reproduction $\mathcal{R}_N$ is given by N i.i.d. copies of X . In turn, the limiting log-profile is given by $r(x) := \pi(1 + h(x))$.

Suppose β, γ satisfy $\beta = 1 - \gamma \in (\rho, 1)$. An easy computation yields $\sigma^* := \sigma^*(r) = 1$, and there exist $a < b < 1$ such that

$$\bar{s}_{\sigma^*}(x) = r(g)(x) + \beta(x - \sigma)_- = 1 - x^\rho + \beta(x - 1)$$

is greater or equal to 0 on $[a, b] \cup \{1\}$ and negative otherwise, see Figure 5. Since $\pi(0) = 0$,

$$\text{Supp}(s(r)) = [a, b] \cup \{1\}. \quad (2.34)$$

Let us now show that the limiting log-profile (if there exists any) after selection can not coincide with $s(r)$. We only provide an outline of the proof and leave the details to the reader. Let $\mathcal{S}_N$ be the point measure after the selection step. The key point is that for $\varepsilon > 0$ small enough, the random variable

$$\#\{x \in \mathcal{S}_N \in c_N(1 - \varepsilon, 1 + \varepsilon)\}$$

is tight and any subsequential limit has a positive probability to be 0. If the limiting r.v. is 0, the limiting log-profile must be $-\infty$ at 1, in contradiction with (2.34).

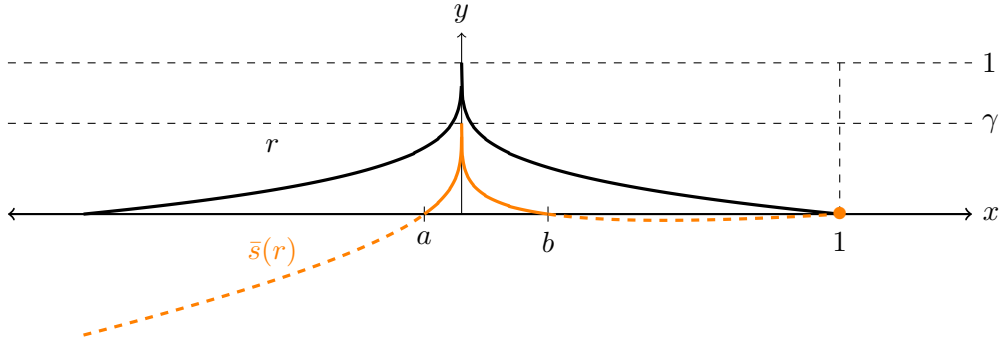


Figure 5: After applying the operator $\bar{s}$ to $r(x) = \pi(1 - |x|^{1/4})$ (in black) with $\sigma^*(r) = 1$, the function $\bar{s}(r)$ (in orange) has a local maximum $\bar{s}(r)(1) = 0$, and $s(r)$ (in solid orange) is not the limiting log-profile of the process after selection.

3 Piecewise linear log-profiles

From now on, we consider the special class of limiting log-profiles when the function h has the form

$$h(x) = \begin{cases} -x & x \geq 0 \\ c_- x & x < 0, \end{cases} \quad (3.1)$$

where $c_- \in \mathbb{R}_+ \cup \{+\infty\}$. Define the parameters

$$k := k(\beta) = \begin{cases} \lfloor \frac{1}{\beta} \rfloor & \text{if } \frac{1}{\beta} \notin \mathbb{N} \\ \frac{1}{\beta} - 1 & \text{otherwise} \end{cases}, \quad K := K(\beta) = \left\lfloor \frac{1 + c_-}{\beta} \right\rfloor + 2, \quad (3.2)$$

where we take the convention that $K = +\infty$ if $c_- = \infty$. Since $c_- > 0$, we have that $K \geq k + 2$, while we also note that when $c_- \neq +\infty$ then

$$(K - 1)\beta - 1 \leq c_- + \beta \leq K\beta - 1. \quad (3.3)$$

Throughout the rest of the article, we will use the convention that $\mathbb{R}^{K+1} := \mathbb{R}^{\mathbb{N}}$ when $K = +\infty$. We are interested in the following class of functions:

Definition 3.1. Suppose $\mathbf{x} \in \mathbb{R}^{K+1}$ is a non-increasing vector, that is,

$$x_0 \geq x_1 \geq x_2 \geq \cdots \geq x_K. \quad (3.4)$$

Define the function $f_{\mathbf{x}} \in \mathcal{C}$ by

$$f_{\mathbf{x}}(x) := \pi \left(\int_x^{x_0} (-c_- - \beta) \vee \left(1 - \beta \sum_{j=1}^K \mathbf{1}_{\{z < x_j\}}(z) \right) dz \right). \quad (3.5)$$

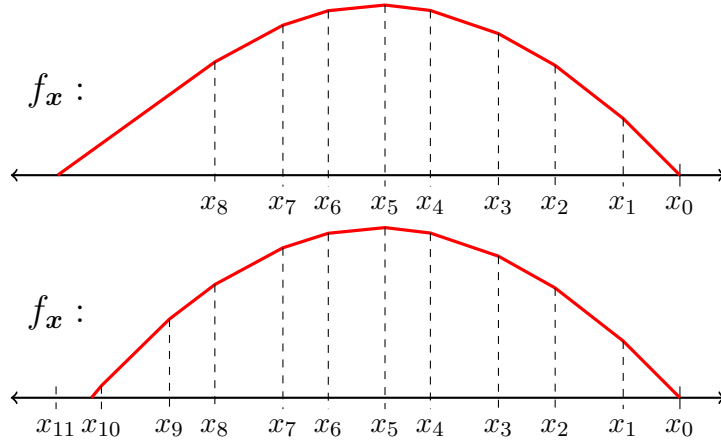


Figure 6: Two examples of functions $f_{\mathbf{x}}$. In the first plot, $c_- = 0.5$ and so $k = 4, K = 8$, and the function $f_{\mathbf{x}}$ has slopes $-1, -(1 - \beta), \dots, -(1 - 7\beta)$ and $c_- + \beta$. In the second plot, $c_- = \infty$ and so $k = 4, K = \infty$, and the function $f_{\mathbf{x}}$ may have slope $-1, -(1 - \beta), -(1 - 2\beta), \dots$

The function $f_{\mathbf{x}}$ consists of the non-negative parts of the linear functions with possible slopes

$$c_- + \beta, \text{ on } (-\infty, x_K], \text{ and } -(1 - j\beta) \text{ on } [x_{j+1}, x_j] \text{ for } j \in \{0\} \cup [K - 1].$$

Because of (3.3), the function $f_{\mathbf{x}}$ is concave on its support for all non-increasing vectors $\mathbf{x} \in \mathbb{R}^{K+1}$, and so $f_{\mathbf{x}} \in \mathcal{C}$ and $x_0 = U(f_{\mathbf{x}})$ is the upper edge of $f_{\mathbf{x}}$. We note from (3.5) that $f_{\mathbf{x}}$ is not necessarily defined by a unique vector $\mathbf{x}$, though as we state in the following lemma, we may determine when two vectors $\mathbf{x}$ and $\mathbf{x}'$ define the same function $f_{\mathbf{x}} = f_{\mathbf{x}'}$.

Lemma 3.2. *Let $\mathbf{x} \in \mathbb{R}^{K+1}$ be non-increasing. For any non-increasing $\mathbf{x}' \in \mathbb{R}^{K+1}$, $f_{\mathbf{x}'} = f_{\mathbf{x}}$ if and only if the following are satisfied:*

$$\begin{aligned} x'_j &= x_j & \text{for all } x_j > L(f_{\mathbf{x}}), \\ x'_j &\leq L(f_{\mathbf{x}}) & \text{for all } x_j \leq L(f_{\mathbf{x}}). \end{aligned}$$

Proof. The result follows immediately from the definition of $f_{\mathbf{x}}$ and $f_{\mathbf{x}'}$ in (3.5). $\square$

Remark 3.3. The definition of k entails that the function $f_{\mathbf{x}}$ is decreasing for $x > x_{k+1}$ and non-decreasing on $x \leq x_{k+1}$, and so

$$\sup f_{\mathbf{x}}(x) = f_{\mathbf{x}}(x_{k+1}) = \sum_{j=0}^k (1 - j\beta)(x_j - x_{j+1}). \quad (3.6)$$

Therefore $L(f_{\mathbf{x}}) \leq x_{k+1}$ and so by Lemma 3.2,

$$f_{\mathbf{x}} = f_{\mathbf{x}'} \implies x_0 = x'_0, x_1 = x'_1, \dots, x_{k+1} = x'_{k+1}. \quad (3.7)$$

In Section 3.1, we introduce a vector operator ψ_{σ} and prove that under some conditions on σ , we may translate the operator $s_{\sigma} \circ r$ acting on $f_{\mathbf{x}}$ as a transformation $\mathbf{x} \mapsto \psi_{\sigma}(\mathbf{x})$ so that $s_{\sigma} \circ r(f_{\mathbf{x}}) = f_{\psi_{\sigma}(\mathbf{x})}$. The main result of this section is Theorem 3.9, which is contained in Section 3.2. We show that for any $g \in \mathcal{D}$, after finitely many applications of $s \circ r$, the resulting function is a piecewise linear function belonging to the following class of functions:

Definition 3.4. A function g belongs to $\mathcal{T}$ if and only if $g = f_{\mathbf{x}}$ for some non-increasing vector $\mathbf{x} \in \mathbb{R}^{K+1}$ and $\sup g = \gamma$.

We will assume throughout this section that $g^0 \in \mathcal{D}$ with $\sup g^0 = \gamma$. As in the dynamics of Theorem 1.3 we recursively define for all $t \geq 0$,

$$\sigma^t := \sigma^*(r(g^t)), \quad g^{t+1} := s \circ r(g^t) = s_{\sigma^t} \circ r(g^t), \quad (3.8)$$

and let $L^t := L(g^t)$, $U^t := U(g^t)$ be the lower and upper edges of g^t .

Remark 3.5. A careful reading of the proofs in this section will reveal that our arguments are valid for a much broader class of functions than $\mathcal{D}$. Namely, it is enough to assume that $g^0 : \mathbb{R} \rightarrow \mathbb{R}_+ \cup \{-\infty\}$ have bounded support with $\sup g^0 = \gamma$.

3.1 Operators on piecewise linear log-profiles

Starting from $f_{\mathbf{x}}$ for non-increasing $\mathbf{x} \in \mathbb{R}^{K+1}$, we show that $r(f_{\mathbf{x}})$ is a piecewise linear function, and that for a range of values of σ , $s_{\sigma} \circ r(f_{\mathbf{x}}) = f_{\mathbf{y}}$ for some non-increasing vector $\mathbf{y} \in \mathbb{R}^{K+1}$. We evaluate how the operator $s_{\sigma} \circ r$ acts on $f_{\mathbf{x}}$ by evaluating the affects on the vector $\mathbf{x}$. More precisely, we introduce in Lemma 3.7 below a vector operator ψ_{σ} and prove that under some conditions on σ , $s_{\sigma} \circ r(f_{\mathbf{x}}) = f_{\psi_{\sigma}(\mathbf{x})}$.

We start with a description of how the operator r behaves when h satisfies (3.1). For any $g \in \mathcal{D}$, we bound $r(g)$ from above and show that $r(g)$ is a line with slope -1 on the interval $[U(g), U(r(g))]$. We also provide a direct computation for $r(f_{\mathbf{x}})$.

Lemma 3.6. For $g \in \mathcal{D}$, define $\omega(g) := \sup_{z \in \mathbb{R}} (g(z) + z)$. The following hold:

(i) For all $x \in \mathbb{R}$,

$$r(g)(x) \leq \pi(1 - \gamma + \omega(g) - x),$$

with equality for all $x \geq U(g)$, that is, $r(g)$ is a line with slope -1 on $[U(g), U(r(g))]$.

(ii) For $\mathbf{x} \in \mathbb{R}^{K+1}$ non-increasing, let $L' := L(\mathbf{f}_{\mathbf{x}}) \vee x_{K-1}$ when $K \in \mathbb{N}$ and $L' = L(\mathbf{f}_{\mathbf{x}})$ when $K = \infty$. Then $\omega(\mathbf{f}_{\mathbf{x}}) = U(\mathbf{f}_{\mathbf{x}}) = x_0$ and

$$r(\mathbf{f}_{\mathbf{x}})(x) = \begin{cases} \pi(1 - \gamma - (x - x_0)) & x > x_0, \\ 1 - \gamma + \mathbf{f}_{\mathbf{x}}(x) & x \in [L', x_0], \\ \pi(1 - \gamma + \mathbf{f}_{\mathbf{x}}(L') - c_-(L' - x)) & x < L'. \end{cases} \quad (3.9)$$

Proof. Let $U = U(g)$ and $L = L(g)$. Note that $h(x - z) \leq z - x$ for all $z, x \in \mathbb{R}$, and so immediately

$$1 - \gamma + \sup_{z \in \mathbb{R}} (g(z) + h(x - z)) \leq 1 - \gamma + \sup_{z \in \mathbb{R}} (g(z) + z - x) = 1 - \gamma + \omega(g) - x. \quad (3.10)$$

Applying π to both sides proves the inequality of (i). For the equality, since $g(z) = -\infty$ for all $z > U$, then $\sup_{z \in \mathbb{R}} (g(z) + h(x - z)) = \sup_{z \in \mathbb{R}} (g(z) + z - x)$ for all $x \geq U$. Thus for all $x \geq U$, we replace the inequality of (3.10) with equality, and applying π to both sides once more completes the proof of (i).

For any non-increasing sequence $\mathbf{x} \in \mathbb{R}^{K+1}$, the slopes of $\mathbf{f}_{\mathbf{x}}$ are all at least -1 . As a consequence, we see that $\omega(\mathbf{f}_{\mathbf{x}}) = U(\mathbf{f}_{\mathbf{x}}) = x_0$, and the first case of (3.9) follows from (i). Let $f := \mathbf{f}_{\mathbf{x}}$ and suppose $x \in [L', x_0]$. If $x \in [x_{j+1}, x_j]$ for $j = 0, 1, \dots, K - 2$, then by the definition of $\mathbf{f}_{\mathbf{x}}$,

$$f(y) \leq f(x) - (1 - j\beta)(y - x) \text{ for all } y \in \mathbb{R},$$

and since $-1 \leq -(1 - j\beta) \leq c_-$ for these choices of j , $f(z) + h(x - z)$ reaches its maximum at x and

$$r(f)(x) = 1 - \gamma + \sup_{z \in \mathbb{R}} (f(z) + h(x - z)) = 1 - \gamma + f(x).$$

Next for $x < L'$ and recalling (3.3), we see that $f(x) \leq f(L') - c_-(L' - x)$, and so

$$r(f)(x) = \pi \left(1 - \gamma + \sup_{z \in \mathbb{R}} (f(z) + h(x - z)) \right) = \pi(1 - \gamma + f(L') - c_-(L' - x)),$$

completing the proof. $\square$

We now focus on introducing the vector operator ψ_σ that satisfies $s_\sigma \circ r(\mathbf{f}_{\mathbf{x}}) = \mathbf{f}_{\psi_\sigma(\mathbf{x})}$. We have already seen above in Lemma 3.6 that $r(\mathbf{f}_{\mathbf{x}})$ is once more a piecewise linear function, while the addition of $\beta(x - \sigma)_-$ changes the slopes of $r(\mathbf{f}_{\mathbf{x}})$ by β for all $x < \sigma$ and leaves the slopes unchanged for $x \geq \sigma$. Therefore, defining ψ_σ is a matter of finding in which linear interval of $r(\mathbf{f}_{\mathbf{x}})$ lies σ , or if σ lies outside the support of $r(\mathbf{f}_{\mathbf{x}})$.

Defining ψ_σ for all possible σ would be quite cumbersome. We therefore restrict ourselves to values of σ that we will encounter in later proofs. For $g \in \mathcal{D}$, define

$$a_-(g) := L(g) + \frac{1 - \gamma}{1 \wedge \beta} \quad \text{and} \quad a_+(g) := \omega(g) - \gamma + \frac{1 - \gamma}{1 \wedge \beta}. \quad (3.11)$$

The values above are not arbitrary, indeed we will see later (in Lemma 3.11) that $a_-(g^t) \leq \sigma^t \leq a_+(g^t)$ for all $t \geq 0$.

For a non-increasing vector $\mathbf{x} \in \mathbb{R}^{K+1}$, define

$$\rho(\mathbf{x}) \in \mathbb{R}^{K+1} := \begin{cases} (x_0 + 1 - \gamma, x_1, \dots, x_{K-1}, x_{K-1}) & K \in \mathbb{N} \\ (x_0 + 1 - \gamma, x_1, x_2, \dots) & K = +\infty. \end{cases} \quad (3.12)$$

The vector $\rho(\mathbf{x})$ can be seen as a description of $r(f_{\mathbf{x}})$. From (3.9), we see that $r(f_{\mathbf{x}})$ may have a linear section with slope c_- , and so $r(f_{\mathbf{x}})$ may not necessarily be expressed as $f_{\mathbf{y}}$ for some non-increasing vector $\mathbf{y}$ (as defined in Definition 3.1). However, it is true that the linear intervals of $r(\mathbf{x})$ are bounded by the values $L \vee \rho(\mathbf{x})_j$. Thus, as discussed above, the different cases of ψ_{σ} depend on σ and $\rho(\mathbf{x})$.

We refer the reader to Figure 7 for a visual aid in understanding the definition of ψ_{σ} below. Lemma 3.7 may be proved by a case by case analysis for the different values of σ , though by examining the integral definition of $f_{\mathbf{x}}$ in (3.5), the proof becomes somewhat more straightforward.

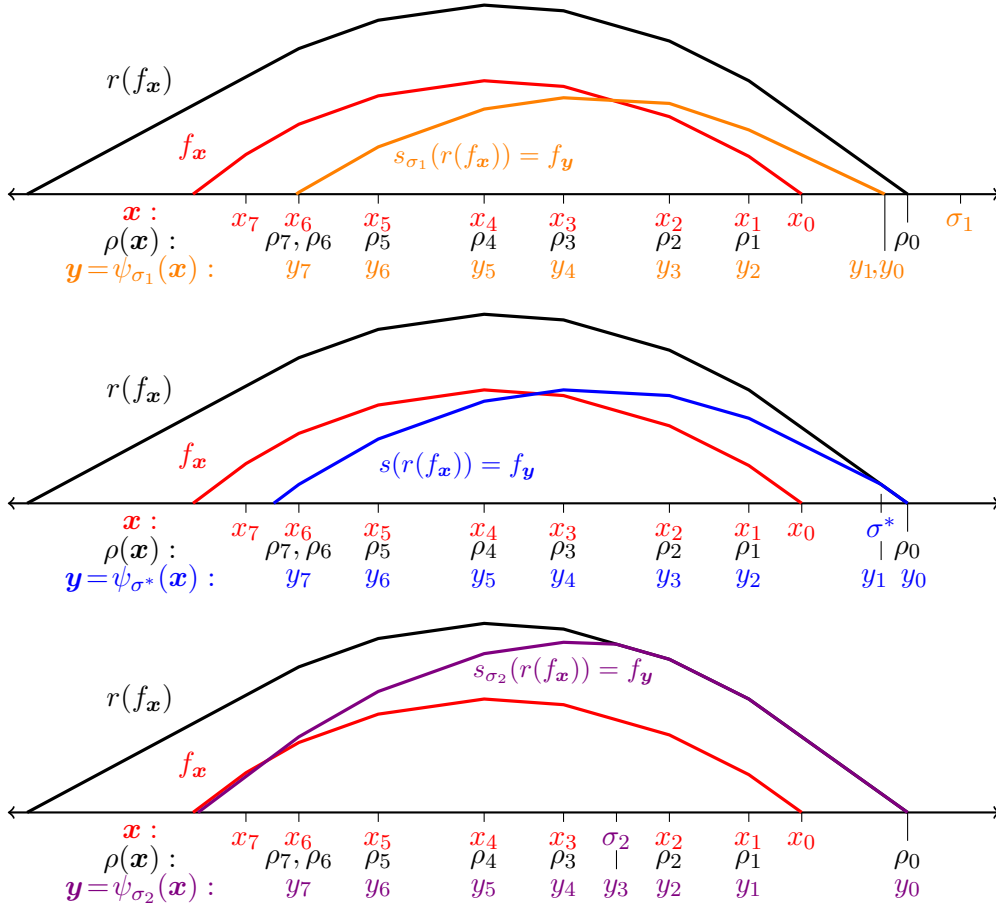


Figure 7: Let $\beta = 0.3$, $\gamma = 0.6$, and $c_- = 0.75$, and so $k = 4$ and $K = 7$. The functions $f_{\mathbf{x}}$, $r(f_{\mathbf{x}})$ and $s_{\sigma}(r(f_{\mathbf{x}})) = f_{\mathbf{y}}$ are illustrated along with the positions of the entries of the vectors $\mathbf{x}$, $\rho(\mathbf{x})$, and $\mathbf{y} = \psi_{\sigma}(\mathbf{x})$ for $\sigma \in \{\sigma_1, \sigma^*, \sigma_2\}$, where $\sigma^* := \sigma^*(r(f_{\mathbf{x}}))$.

Lemma 3.7. Recall the definition of $a_-(g)$ and $a_+(g)$ from (3.11). If $\mathbf{x} \in \mathbb{R}^{K+1}$ is non-increasing and $\sigma \in \mathbb{R}$ satisfies

$$a_-(f_{\mathbf{x}}) \vee x_{k+1} \leq \sigma \leq a_+(f_{\mathbf{x}}), \quad (3.13)$$

then $s_\sigma(r(f_{\mathbf{x}})) = f_{\psi_\sigma(\mathbf{x})}$, where

$$\psi_\sigma(\mathbf{x}) := \begin{cases} \left(\frac{\rho(\mathbf{x})_0 - \beta\sigma}{1-\beta}, \frac{\rho(\mathbf{x})_0 - \beta\sigma}{1-\beta}, x_1, x_2, x_3, \dots, x_{K-1} \right) & \rho(\mathbf{x})_0 < \sigma, \\ (\rho(\mathbf{x})_0, x_1, \dots, x_i, \sigma, x_{i+1}, x_{i+2}, \dots, x_{K-1}) & \rho(\mathbf{x})_{i+1} < \sigma \leq \rho(\mathbf{x})_i \text{ for } 0 \leq i \leq k. \end{cases} \quad (3.14)$$

Remark 3.8. Note from (3.11), if $\beta \geq 1$, then $a_+(f_{\mathbf{x}}) = x_0 + 1 - 2\gamma < \rho(\mathbf{x})_0$ and so the first case of (3.14) does not apply. Therefore, (3.14) is well defined for all σ satisfying (3.13).

Proof. Denote $f := f_{\mathbf{x}}$ and $L := L(f)$, $U := U(f)$. We start by assuming that $x_{K-1} \geq L$, and so $L' = x_{K-1}$ as defined in Lemma 3.6 (ii). From (3.9) and the formula $f_{\mathbf{x}}$ from (3.5), we may write

$$r(f)(x) = \pi \left(\int_x^{\rho(\mathbf{x})_0} -c_- \vee \left(1 - \beta \sum_{j=1}^{K-1} \mathbb{1}_{\{z < x_j\}}(z) \right) dz \right). \quad (3.15)$$

Since $\beta(x - \sigma)_- = - \int_x^\infty \beta \mathbb{1}_{\{z < \sigma\}}(z) dz$, we see that

$$s_\sigma \circ r(f)(x) = \pi \left(\int_x^{\rho(\mathbf{x})_0} -c_- \vee \left(1 - \beta \sum_{j=1}^{K-1} \mathbb{1}_{\{z < x_j\}}(z) \right) dz - \int_x^\sigma \beta \mathbb{1}_{\{z < \sigma\}}(z) dz \right). \quad (3.16)$$

We need only show that the expression in (3.16) is equal to the expression of $f_{\psi_\sigma(\mathbf{x})}$ from (3.5). Along with the restriction $\sigma \geq x_{k+1}$ from (3.13), this follows when $\sigma \leq \rho(\mathbf{x})_0$ by replacing the upper bound of the second integral in (3.16) with $\rho(\mathbf{x})_0$. Now let $\rho(\mathbf{x})_0 < \sigma \leq a_+(f)$, which by Remark 3.8 can only occur if $\beta < 1$. Let $y_0 = U(s_\sigma \circ r(f))$ be the upper edge of $s_\sigma \circ r(f)$. Then $y_0 \leq \rho(\mathbf{x})_0 < \sigma$ and so $\beta(x - \sigma)_-$ can be rewritten as

$$\int_x^\sigma \beta \mathbb{1}_{\{z < \sigma\}}(z) dz = \int_x^{y_0} \beta \mathbb{1}_{\{z < y_0\}}(z) dz + \int_{y_0}^\sigma \beta \mathbb{1}_{\{z < \sigma\}}(z) dz.$$

From Lemma 3.6(ii) and the restriction $\sigma \leq a_+(f) < x_0 + \frac{1-\gamma}{\beta}$, we see that

$$s_\sigma \circ r(f)(x_0) = r(f)(x_0) + \beta(x_0 - \sigma) = 1 - \gamma + \beta x_0 - \beta \sigma > 0,$$

and so $y_0 \geq x_0$. This implies the integrand of the first integral in (3.16) is simply 1 for all $x \in [y_0, \rho(\mathbf{x})_0]$. Thus, so long as

$$\int_{y_0}^{\rho(\mathbf{x})_0} dz - \int_{y_0}^\sigma \beta \mathbb{1}_{\{z < \sigma\}}(z) dz = \rho(\mathbf{x})_0 - \beta \sigma - (1 - \beta)y_0 = 0, \quad (3.17)$$

then (3.16) can be rewritten as

$$s_\sigma \circ r(f)(x) = \pi \left(\int_x^{y_0} -c_- \vee \left(1 - \beta \sum_{j=1}^{K-1} \mathbb{1}_{\{z < x_j\}}(z) \right) dz - \int_x^{y_0} \beta \mathbb{1}_{\{z < y_0\}}(z) dz \right). \quad (3.18)$$

Solving (3.17) yields $y_0 = \frac{\rho(\mathbf{x})_0 - \beta\sigma}{1-\beta}$, and we see that (3.18) is indeed equal to the integral expression for $f_{\psi_\sigma(\mathbf{x})}$ from (3.5).

Now suppose $L \geq x_{K-1}$. If we let $\mathbf{x}' = (x'_0, x'_1, \dots, x'_K)$ where $x'_j = L \vee x_j$, then $f_{\mathbf{x}'} = f$ by Lemma 3.2. Then we may revert to the case above for $f_{\mathbf{x}'}$ and see that

$$s_\sigma \circ r(f) = s_\sigma \circ r(f_{\mathbf{x}'}) = f_{\psi_\sigma(\mathbf{x}')}.$$

Since we assume from (3.13) that $\sigma > L$, we see directly from (3.14) that $\psi_\sigma(\mathbf{x}')_j = L \vee \psi_\sigma(\mathbf{x})_j$. Therefore, from Lemma 3.2, if $L(f_{\psi_\sigma(\mathbf{x}')}) \geq L$, then $f_{\psi_\sigma(\mathbf{x})} = f_{\psi_\sigma(\mathbf{x}')}.$

Since $L \geq x_{K-1}$, then $f(L) = 0$ and by (3.9), $r(f)(x) < r(f)(L) = 1 - \gamma$ for all $x < L$. Since we restrict $\sigma \geq a_-(f) = L + \frac{1-\gamma}{1 \wedge \beta}$, then

$$x < L \implies s_\sigma \circ r(f)(x) = \pi \left(\underbrace{r(f)(x) + \beta(x - \sigma)}_{< 0} \right) = -\infty,$$

and so $L(f_{\psi(\mathbf{x}')}) \geq L$. □

3.2 The space of piecewise linear log-profiles is attractive

The goal of this section is to prove that $\mathcal{T}$ is attractive in the following sense:

Theorem 3.9. *There exists a number $t_\ell := t_\ell(\gamma, \beta)$ (depending only on γ and β) such that for all $g^0 \in \mathcal{D}$, then $g^t \in \mathcal{T}$ for all $t \geq t_\ell$.*

In proving Theorem 3.9, we will first prove the following comparison principle:

Lemma 3.10 (comparison principle). *Define $\mathbf{x}^0 = (\omega(g^0), L^0, L^0, \dots, L^0) \in \mathbb{R}^{K+1}$. The recursion*

$$\mathbf{x}^t := \psi_{\sigma^{t-1}}(\mathbf{x}^{t-1}), \quad t \geq 1$$

is well defined, in that σ^{t-1} satisfies the conditions (3.13) to apply $\psi_{\sigma^{t-1}}$ to $\mathbf{x}^{t-1}$. Furthermore, for all $t \geq 1$, $g^t \leq f_{\mathbf{x}^t}$ and $g^t(x) = f_{\mathbf{x}^t}(x)$ for all $x \geq U^0$.

To complete the argument of Theorem 3.9, we show in the proof of Theorem 3.9 that the supports of the profiles $(g^t)_{t \geq 0}$ are increasing. Specifically we show that the upper edges $(U^t)_{t \geq 0}$ are always increasing by at least some constant C . By applying the comparison principle described above, this implies that more and more of the fronts of the profiles $(g^t)_{t \geq 0}$ are piecewise linear until eventually, $g^t = f_{\mathbf{x}^t} \in \mathcal{T}$ for all t large enough.

Before proving Lemma 3.10, we begin with a technical lemma that provides bounds for σ^t . These bounds will imply that σ^t satisfies the conditions of Lemma 3.7 so that we may apply ψ_{σ^t} to $\mathbf{x}^t$ and guarantee that $s_{\sigma^t} \circ r(f_{\mathbf{x}^t}) = f_{\mathbf{x}^{t+1}}$.

Lemma 3.11. *Recall the sequence $(\sigma^t)_{t \geq 0}$ from (3.8). The following hold:*

(i) *Recall a_- and a_+ from (3.11). For all $t \geq 0$, $\omega(g^t) \in [U^t, U^t + \gamma]$*

$$a_-(g^t) \leq \sigma^t \leq a_+(g^t).$$

(ii) *The sequence $(\sigma^t)_{t=0}^\infty$ is non-decreasing.*

Proof. By definition of U^t being the upper edge of g^t and since $\sup g^t = \gamma$, we see immediately that $U^t \leq \omega(g^t) \leq U^t + \gamma$. Recall that since $\sigma^t := \sigma^*(r(g^t))$, then $s \circ r(g) = s_{\sigma^t}(r(g^t))$.

Starting with the lower bound of σ^t , we note that for all $x \geq \sigma^t$, since $h(0) = 0$,

$$\gamma \geq s \circ r(g^t)(x) = r(g^t)(x) \geq 1 - \gamma + g^t(x),$$

and so $g^t(x) \leq 2\gamma - 1 < \gamma$. Since $\sup g^t = \gamma$, this implies that $\sup_{z \in [L^t, \sigma^t]} g^t(z) = \gamma$, and

$$\begin{aligned} \gamma \geq 0 \vee s \circ r(g^t)(\sigma^t) &= 0 \vee r(g^t)(\sigma^t) \geq 1 - \gamma + \sup_{z \in [L^t, \sigma^t]} (g^t(z) + h(\sigma^t - z)) \\ &\geq 1 - \gamma + \sup_{z \in [L^t, \sigma^t]} g^t(z) + L^t - \sigma^t = 1 + L^t - \sigma^t, \end{aligned}$$

as well as

$$\begin{aligned} \gamma \geq \sup_{z \in [L^t, \sigma^t]} (s \circ r(g^t)(z)) &= \sup_{z \in [L^t, \sigma^t]} (r(g^t)(z) + \beta(z - \sigma^t)) \\ &\geq 1 - \gamma + \sup_{z \in [L^t, \sigma^t]} g^t(z) + \beta(L^t - \sigma^t) = 1 + \beta(L^t - \sigma^t). \end{aligned}$$

The lower bound of (ii) then follows by solving for σ^t in the inequalities above and taking the maximum of the two values.

Now onto the upper bound of σ^t . When $\sup g^t = \gamma$, then $\sup r(g^t) = 1$ which along with Lemma 3.6(i) implies that for all $x \in \mathbb{R}$,

$$r(g^t)(x) \leq r_{\max}^t(x) := 1 \wedge \pi(1 - \gamma + \omega(g^t) - x).$$

We can verify that

$$\sup s_{a_+(g^t)}(r_{\max}^t) = \sup_{x \in \mathbb{R}} (r_{\max}^t(x) + \beta(x - a_+(g^t))_-) = \gamma.$$

Therefore $\sigma^t \leq \sigma^*(r_{\max}^t) \leq a_+(g^t)$ since $r(g^t) \leq r_{\max}^t$.

The proof of (ii) follows by contradiction, and so suppose $\sigma^{t+1} < \sigma^t$. For all $x \in \mathbb{R}$,

$$g^{t+1}(x) = \pi(r(g^t)(x) + \beta(x - \sigma^t)_-) \leq 1 + \beta(x - \sigma^t)_- \implies \beta(x - \sigma^t)_- \geq g^{t+1}(x) - 1.$$

Let $\varepsilon = (1 - \gamma) \wedge \beta(\sigma^t - \sigma^{t+1})$. Since $\sup g^{t+1} = \gamma$, the above implies that for all $x \in \mathbb{R}$,

$$\beta(x - \sigma^{t+1})_- \geq g^{t+1}(x) - 1 + \varepsilon.$$

By evaluating g^{t+2} and recalling that $h(0) = 0$, we then see that

$$\begin{aligned} \sup g^{t+2} &= \sup_{x \in \mathbb{R}} (r(g^{t+1})(x) + \beta(x - \sigma^{t+1})_-) \\ &\geq \sup_{x \in \mathbb{R}} (g^{t+1}(x) + 1 - \gamma + g^{t+1}(x) - 1 + \varepsilon) = 2 \sup g^{t+1} - \gamma + \varepsilon = \gamma + \varepsilon, \end{aligned}$$

a contradiction to $\sup g^{t+2} = \gamma$. □

Recalling the conditions imposed on σ to apply Lemma 3.7, the following is a corollary to Lemma 3.11(i):

Corollary 3.12. *If $g = f_{\mathbf{x}} \in \mathcal{T}$ for some non-increasing $\mathbf{x} \in \mathbb{R}^{K+1}$, then $\sigma^* := \sigma^*(r(g))$ satisfies the conditions (3.13) of Lemma 3.7 and*

$$s \circ r(g) = f_{\psi_{\sigma^*}(\mathbf{x})} \in \mathcal{T}.$$

Proof. Since $g \in \mathcal{T}$, then $g(x_{k+1}) = \gamma$. By using Lemma 3.6(ii) for $r(g)$, for all $\sigma < x_{k+1}$,

$$s_{\sigma} \circ r(g)(x_{k+1}) = r(g)(x_{k+1}) = 1 - \gamma + g(x_{k+1}) = 1 > \gamma.$$

Since $\sup s \circ r(g) = \gamma$, we conclude that $\sigma^*(r(g)) \geq x_{k+1}$. The result then follows from Lemma 3.11(i) by setting $g^0 = g$. $\square$

We now move on to the proof of our comparison principle Lemma 3.10.

Proof of Lemma 3.10. We prove the statement by induction. We see that $g^0 \leq f_{\mathbf{x}^0}$, and so $r(g^0) \leq r(f_{\mathbf{x}^0})$ and by Lemma 3.11(i),

$$a_-(f_{\mathbf{x}^0}) \leq a_-(g^0) \leq \sigma^0 \leq a_+(g^0) = a_+(f_{\mathbf{x}^0}).$$

Therefore, σ^0 satisfies the conditions (3.13) and $s_{\sigma^0} \circ r(f_{\mathbf{x}^0}) = f_{\mathbf{x}^1}$. By Lemma 3.6 we also see that $r(g^0)(x) = r(f_{\mathbf{x}^0})(x)$ for all $x \geq U^0$. This equality as well as the inequality $g^0 \leq f_{\mathbf{x}^0}$ are maintained when applying the operator s_{σ^0} to both $r(g^0)$ and $r(f_{\mathbf{x}^0})$, which completes the base case $t = 1$.

Now suppose the claim holds for some $t \geq 1$, and so $g^t \leq f_{\mathbf{x}^t}$, with equality for $x \geq U^0$. We then clearly have that

$$\sup_{z \in \mathbb{R}} (g^t(z) + h(x - z)) \leq \sup_{z \in \mathbb{R}} (f_{\mathbf{x}^t}(z) + h(x - z)), \quad (3.19)$$

and so $r(g^t) \leq r(f_{\mathbf{x}^t})$; applying the operator s_{σ^t} to both functions maintains the inequality, and so $g^{t+1} \leq s_{\sigma^t} \circ r(f_{\mathbf{x}^t})$.

By Lemma 3.6(ii), $\omega(f_{\mathbf{x}^t}) = x_0^t$, while $x_0^t = U^t \leq \omega(g^t) \leq \omega(f_{\mathbf{x}^t})$, and so by Lemma 3.6, $r(g^t)(x) = r(f_{\mathbf{x}^t})(x)$ for all $x \geq x_0^t$. For all $x \leq x_0^t$, from the slopes of $f_{\mathbf{x}^t}$, we know that $f_{\mathbf{x}^t}(x) \leq f_{\mathbf{x}^t}(y) + y - x$ for all $y > x$. In particular, this implies (along with the induction hypothesis) that for all $x \in [U^0, x_0^t]$,

$$\begin{aligned} r(g^t)(x) &\geq \pi \left(1 - \gamma + \sup_{z \geq x} (g^t(z) + h(x - z)) \right) \\ &= \pi \left(1 - \gamma + \sup_{z \geq x} (f_{\mathbf{x}^t}(z) + h(x - z)) \right) = r(f_{\mathbf{x}^t})(x), \end{aligned}$$

The above along with (3.19) implies that $r(g^t)(x) = r(f_{\mathbf{x}^t})(x)$ for all $x \geq U^0$; the equality is maintained by applying the operation s_{σ^t} to both functions, and so $g^{t+1}(x) = s_{\sigma^t} \circ r(f_{\mathbf{x}^t})(x)$ for $x \geq U^0$.

The inductive step is now complete if we show that σ^t satisfies (3.13), and so $s_{\sigma^t} \circ r(f_{\mathbf{x}^t}) = f_{\mathbf{x}^{t+1}}$. We have already argued that $\omega(g^t) = \omega(f_{\mathbf{x}^t}) = x_0^t$ while $L(f_{\mathbf{x}^t}) \leq L^t$ is evident since $g^t \leq f_{\mathbf{x}^t}$, and applying Lemma 3.11(i) to g^t guarantees

$$a_-(f_{\mathbf{x}^t}) \leq a_-(g^t) \leq \sigma^t \leq a_+(g^t) = a_+(f_{\mathbf{x}^t}).$$

All that is left is to show that $\sigma^t \geq x_{k+1}^t$. By examining (3.14) of Lemma 3.7, the induction hypothesis implies that $\sigma^{t-1} \geq x_{k+1}^t$, while $\sigma^t \geq \sigma^{t-1}$ is guaranteed by Lemma 3.11(ii). $\square$

With the comparison principle of Lemma 3.10 in hand, all that is left is to show that the front edges of g^t are increasing, and so more and more of g^t is piecewise linear until $g^t \in \mathcal{T}$ for all t large enough.

Proof of Theorem 3.9. Recall the vectors $(\mathbf{x}^t)_{t=0}^\infty$ from Lemma 3.10. We have that $x_0^t = \omega(g^0) \geq U^0$, and since $\mathbf{x}^{t+1} = \psi_{\sigma^t}(\mathbf{x}^t)$ and σ^t satisfies the upper bound of (3.13), a quick calculation via (3.14) of Lemma 3.7 reveals

$$x_0^{t+1} - x_0^t = \psi_{\sigma^t}(\mathbf{x}^t)_0 - x_0^t \geq \begin{cases} (1 - \gamma) \wedge \left(\frac{\beta\gamma}{1-\beta}\right) & \beta < 1 \\ 1 - \gamma & \beta \geq 1. \end{cases} \quad (3.20)$$

By repeatedly applying (3.20), there exists $\tau_1 := \tau_1(\gamma, \beta)$ such that for all $t \geq \tau_1$,

$$x_0^t > U^0 + \frac{\gamma}{1 - k\beta}. \quad (3.21)$$

Between x_{k+1}^t and x_0^t , the function $f_{\mathbf{x}^t}$ has slopes $-1 + j\beta$ for $j = 0, \dots, k$, and so

$$x \in [x_{k+1}^t, x_0^t] \implies f_{\mathbf{x}^t}(x) \geq (1 - k\beta)(x_0^t - x). \quad (3.22)$$

Since $f_{\mathbf{x}^t}$ and g^t coincide for all $x \geq U^0$, if $x_{k+1}^t \leq U^0$ for some $t \geq \tau_1$, then (3.21) and (3.22) would imply $g^t(U^0) = f_{\mathbf{x}^t}(U^0) > \gamma$, contradicting $\sup g^t = \gamma$. Thus for all $t \geq \tau_1$,

$$x_{k+1}^t > U^0.$$

Since $f_{\mathbf{x}^t} \geq g^t$, then $\sup f_{\mathbf{x}^t} \geq \sup g^t = \gamma$. Additionally, since the two functions coincide for $x \geq U^0$ and $f_{\mathbf{x}^t}$ reaches its maximum at x_{k+1}^t , we see that for all $t \geq \tau_1$,

$$\sup f_{\mathbf{x}^t} = f_{\mathbf{x}^t}(x_{k+1}^t) = g(x_{k+1}^t) \leq \gamma, \quad (3.23)$$

thus $\sup f_{\mathbf{x}^t} = \gamma$ and $f_{\mathbf{x}^t} \in \mathcal{T}$ for all $t \geq \tau_1$.

Let $t \geq \tau_1$ for the rest of the proof. From Lemma 3.10, σ^t satisfies the lower bound of (3.13). Therefore $\sigma^t \geq x_{k+1}^t$, and since $U^0 < x_{k+1}^t$,

$$g^{t+1}(U^0) \leq \pi(r(g^t)(U^0) + \beta(U^0 - \sigma^t)) \leq \pi(1 + \beta(U^0 - \sigma^t)) \leq \pi(1 + \beta(U^0 - x_{k+1}^t)). \quad (3.24)$$

From (3.22) and (3.23), we conclude that

$$x_{k+1}^t \geq x_0^t - \frac{\gamma}{1 - k\beta}.$$

By once more repeatedly applying (3.20), the above inequality yields that for all $t \geq t_\ell$ for some $t_\ell := t_\ell(\gamma, \beta) \geq \tau_1$,

$$x_{k+1}^t > U^0 + \frac{1}{\beta}. \quad (3.25)$$

Substituting (3.25) into (3.24), $f_{\mathbf{x}^t}(U^0) = g^t(U^0) = -\infty$ for all $t \geq t_\ell$. Since $f_{\mathbf{x}^t} \in \mathcal{C}$ is concave, this implies that $f_{\mathbf{x}^t}(x) = -\infty$ for all $x < U^0$, and since $g^t \leq f_{\mathbf{x}^t}$, we see that the two functions coincide on the entire real line for all $t \geq t_\ell$. Since $f_{\mathbf{x}^t} \in \mathcal{T}$, so is $g^t \in \mathcal{T}$. $\square$

4 Traveling Wave Solutions

For a non-increasing vector $\mathbf{x} \in \mathbb{R}^{K+1}$, recall the parameters k and K from (3.2), the definition of $f_{\mathbf{x}}$ from Definition 3.1, as well as the definition of $\mathcal{T}$ from Definition 3.4. For $g = f_{\mathbf{x}} \in \mathcal{T}$, we define

$$\mathbf{v}(g) := (x_1 - x_2, \dots, x_k - x_{k+1}) \in \mathbb{R}^k. \quad (4.1)$$

For $j = 1, \dots, k$, $\mathbf{v}(g)_j$ consists of the linear segment of g with slope $-(1 - j\beta)$, and by (3.6), the length of the linear segment with slope -1 is given by

$$x_0 - x_1 = \gamma - \langle \mathbf{u}, \mathbf{v}(g) \rangle, \quad \text{where} \quad \mathbf{u} := (1 - j\beta)_{j=1}^k. \quad (4.2)$$

Since all other slopes of g are increasing, $\mathbf{v}(g)$ and (4.2) describe g from the point where it reaches its maximum γ to its upper edge. We also note by (3.7), that $\mathbf{v}(g)$ is uniquely defined even if there are potentially several $\mathbf{x}$'s such that $g = f_{\mathbf{x}}$. When $\beta \geq 1$ (and so $k = 0$), then $\mathbf{v}(g)$ is simply the empty vector; we will assume in this case that $\langle \mathbf{u}, \mathbf{v}(g) \rangle = 0$, and so (4.2) simplifies to $x_0 - x_1 = \gamma$.

As is stated in the following lemma, any traveling wave solution must be a piecewise linear function $g \in \mathcal{T}$ for which $\mathbf{v}(g)$ is invariant under the operation $s \circ r$.

Lemma 4.1. *If (g, ν) is a traveling wave solution, then $g \in \mathcal{T}$ and $\mathbf{v}(s \circ r(g)) = \mathbf{v}(g)$.*

Proof. The first part of the lemma follows from Theorem 3.9. Let $\mathbf{x}$ be a non-increasing vector such that $g = f_{\mathbf{x}}$ and let $\sigma^* := \sigma^*(r(g))$. By Corollary 3.12 and the definition of traveling waves,

$$f_{\psi_{\sigma^*}(\mathbf{x})} = s \circ r(g) = g(\cdot - \nu) = f_{\mathbf{x}}(\cdot - \nu).$$

The result now follows by Lemma 3.2 and the definitions of $\mathbf{v}(g)$ and $\mathbf{v}(s \circ r(g))$. $\square$

In seeking traveling wave solutions, the following lemma implies that we can in fact limit our search to a subclass of functions $g \in \mathcal{T}$.

Lemma 4.2. *For $g \in \mathcal{T}$, if $\mathbf{v}(s \circ r(g)) = \mathbf{v}(g)$, then*

$$\sigma^* := \sigma^*(r(g)) > U(g) + \langle \mathbf{u}, \mathbf{v}(g) \rangle - \gamma. \quad (4.3)$$

Proof. Let $g = f_{\mathbf{x}}$ for a non-increasing vector $\mathbf{x} \in \mathbb{R}^{K+1}$ and denote $\mathbf{v} := \mathbf{v}(g)$. Recalling (4.2), if (4.3) does not hold, then

$$\sigma^* \leq x_0 - \gamma + \langle \mathbf{u}, \mathbf{v} \rangle = x_1 = \rho(\mathbf{x})_1.$$

Therefore, from Lemma 3.7 and (4.2) again,

$$\gamma - \langle \mathbf{u}, \mathbf{v}(s \circ r(g)) \rangle = \psi_{\sigma^*}(\mathbf{x})_0 - \psi_{\sigma^*}(\mathbf{x})_1 = x_0 + 1 - \gamma - x_1 > x_0 - x_1 = \gamma - \langle \mathbf{u}, \mathbf{v}(g) \rangle,$$

and so $\mathbf{v}(s \circ r(g)) \neq \mathbf{v}(g)$. $\square$

We will show in Section 4.1 that the mapping $\mathbf{v} \mapsto \mathbf{v}(s \circ r(g))$ is a piecewise affine transformation that can be derived from Lemma 3.7. In light of Lemma 4.2, we will only

prove the transformations for $g \in \mathcal{T}$ satisfying (4.3), though we provide without proof the mapping $\mathbf{v} \mapsto \mathbf{v}(s \circ r(g))$ in terms of affine transformations for all $g \in \mathcal{T}$ (see Remark 4.5).

The proof of Theorem 1.4 is contained in Section 4.2. We prove convergence results for the affine transformations of Section 4.1, the limits of which corresponding to potential vectors $\mathbf{v}(g)$ of functions $g \in \mathcal{T}$ satisfying $\mathbf{v}(s \circ r(g)) = \mathbf{v}(g)$. These limiting vectors inform the parameters $\nu(\gamma, \beta)$ and $\chi(\gamma, \beta)$ of Theorem 1.4

After stating the continuity of certain maps in Section 4.3, Theorem 1.5 follows by translating the vector convergence results of Section 4.2 to the convergence of translations of g^t to the traveling wave solution. The continuity lemmas needed have technical proofs that are left to Section 4.3.1.

4.1 Piecewise affine transformations of the front

From (3.6), since $\sup g = \gamma$ for $g \in \mathcal{T}$, we see that $\mathbf{v}(g)$ must belong to the following subset of $\mathbb{R}^k$:

$$\Delta := \left\{ \mathbf{v} \in \mathbb{R}^k : v_i \geq 0 \text{ and } \langle \mathbf{u}, \mathbf{v} \rangle \leq \gamma \right\}, \quad \text{where } \mathbf{u} = (1 - j\beta)_{j=1}^k. \quad (4.4)$$

We wish to express $\mathbf{v}(g) \mapsto \mathbf{v}(s \circ r(g))$ in terms of vector transformations (independent of g), and to that end, we start with expressing $\sigma^*(r(g)) - U(g)$ in terms of $\mathbf{v}(g)$ alone.

Lemma 4.3. *Recall $\mathbf{u}$ from (4.4) and define the piecewise linear operator $\Sigma : \mathbb{R}^k \rightarrow \mathbb{R}$ by*

$$\Sigma(\mathbf{v}) := \begin{cases} \langle \mathbf{u}, \mathbf{v} \rangle - \gamma - \sum_{i=1}^{k-1} v_i - \frac{1-k\beta}{1\wedge\beta} v_k + \frac{1-\gamma}{1\wedge\beta} & \text{if } v_k < \frac{1-\gamma}{1-k\beta}, \\ \langle \mathbf{u}, \mathbf{v} \rangle - \gamma - \sum_{i=1}^k v_i + \frac{1-\gamma}{1-k\beta} & \text{otherwise.} \end{cases} \quad (4.5)$$

For all $g \in \mathcal{T}$,

$$\sigma^*(r(g)) = U(g) + \Sigma(\mathbf{v}(g)). \quad (4.6)$$

Proof. Let $\mathbf{x} \in \mathbb{R}^{K+1}$ such that $f_{\mathbf{x}} = g$, and let $\sigma^* := \sigma^*(r(g))$. From Corollary 3.12, we have $s \circ r(g) = f_{\psi_{\sigma^*}(\mathbf{x})} \in \mathcal{T}$.

We start with $k = 0$ (and so $\beta \geq 1$), then $\mathbf{v}$ is empty and both cases of (4.5) simplify to $1 - 2\gamma$. By (4.2), $\gamma = x_0 - x_1$ and

$$\gamma = \psi_{\sigma^*}(\mathbf{x})_0 - \psi_{\sigma^*}(\mathbf{x})_1 = x_0 + 1 - \gamma - \sigma^*.$$

Solving for σ^* yields $\sigma^* = 1 - 2\gamma$, completing the argument in this case.

For the remainder of the proof, assume $\beta \in (0, 1)$. Recalling (3.6), g achieves its maximum of γ at x_{k+1} and $s \circ r(g)$ achieves its maximum at $\psi_{\sigma^*}(\mathbf{x})_{k+1}$. Since g has slope $-(1 - k\beta)$ on (x_{k+1}, x_k) ,

$$x \in (x_{k+1}, x_k] \implies g(x) = \gamma - (1 - k\beta)(x - x_{k+1}). \quad (4.7)$$

By Lemma 3.7, if $\sigma^* > x_k$, then $\psi_{\sigma^*}(\mathbf{x})_{k+1} = x_k$ and by using Lemma 3.6(ii) to calculate $r(g)(x_k)$,

$$\gamma = s \circ r(g)(x_k) = r(g)(x_k) + \beta(x_k - \sigma^*) = 1 - \gamma + g(x_k) + \beta(x_k - \sigma^*), \quad (4.8)$$

while if $\sigma^* \in (x_{k+1}, x_k]$, then $\psi_{\sigma^*}(\mathbf{x})_{k+1} = \sigma^*$ and

$$\gamma = s \circ r(g)(\sigma^*) = r(g)(\sigma^*) = 1 - \gamma + g(\sigma^*). \quad (4.9)$$

Substituting (4.7) into (4.8) and (4.9), and solving for σ^* yields

$$\sigma^* = \begin{cases} x_k - \frac{1-k\beta}{\beta}(x_k - x_{k+1}) + \frac{1-\gamma}{\beta} & \text{if } \sigma^* > x_k \\ x_{k+1} + \frac{1-\gamma}{1-k\beta} & \text{if } \sigma^* \leq x_k. \end{cases} \quad (4.10)$$

By evaluating both cases of (4.10), we see that

$$\begin{aligned} \sigma^* > x_k &\implies x_k - x_{k+1} = \frac{\beta}{1-k\beta} \left(\sigma^* - x_k + \frac{1-\gamma}{\beta} \right) < \frac{1-\gamma}{1-k\beta}, \\ \sigma^* \leq x_k &\implies x_k - x_{k+1} \geq \sigma^* - x_{k+1} = \frac{1-\gamma}{1-k\beta}. \end{aligned}$$

Thus $v_k = x_k - x_{k+1} < \frac{1-\gamma}{1-k\beta}$ if and only if $\sigma^* > x_k$. From (4.1) and (4.2),

$$\text{for } i \geq 1, \quad x_i = x_0 - (\gamma - \langle \mathbf{u}, \mathbf{v} \rangle) - \sum_{j=1}^{i-1} v_j, \quad (4.11)$$

and replacing x_j with (4.11) in (4.10) completes the proof. $\square$

From Lemma 4.2 and applying Lemma 4.3 to express $\sigma^*(r(g))$, we see that if $g \in \mathcal{T}$ satisfies $\mathbf{v}(s \circ r(g)) = \mathbf{v}(g)$, then

$$\begin{aligned} \mathbf{v}(g) \in \Delta_{-1} &:= \left\{ \mathbf{v} \in \Delta : 1 - \gamma < \Sigma(\mathbf{v}) \right\}, \quad \text{or} \\ \mathbf{v}(g) \in \Delta_0 &:= \left\{ \mathbf{v} \in \Delta : \langle \mathbf{u}, \mathbf{v} \rangle - \gamma < \Sigma(\mathbf{v}) \leq 1 - \gamma \right\}. \end{aligned} \quad (4.12)$$

Thus we focus on expressing the transformation $\mathbf{v}(g) \rightarrow \mathbf{v}(s \circ r(g))$ as piecewise linear transformations when $\mathbf{v}(g) \in \Delta_{-1} \cup \Delta_0$. Define the matrices

$$A_{-1} := \begin{pmatrix} \frac{\beta}{1-\beta} & \frac{\beta}{1-\beta} & \cdots & \frac{\beta}{1-\beta} & \frac{1-k\beta}{1-\beta} \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \quad A_0 := \begin{pmatrix} -1 & -1 & \cdots & -1 & \frac{-(1-k\beta)}{\beta} \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

and let $\{\mathbf{e}_i\}_{i=1}^k$ be the standard basis of $\mathbb{R}^k$.

Lemma 4.4. *For $g \in \mathcal{T}$, let $\mathbf{v} := \mathbf{v}(g)$.*

(i) *If $\mathbf{v} \in \Delta_{-1}$, then*

$$\mathbf{v}(s \circ r(g)) = A_{-1}\mathbf{v} + \frac{\gamma - \langle \mathbf{u}, \mathbf{v} \rangle}{1-\beta} \mathbf{e}_1 \quad \text{and} \quad U(s \circ r(g)) = U(g) + \frac{1-\gamma - \beta \Sigma(\mathbf{v})}{1-\beta}. \quad (4.13)$$

(ii) *If $\mathbf{v} \in \Delta_0$, then*

$$\mathbf{v}(s \circ r(g)) = A_0\mathbf{v} + \frac{1-\gamma}{\beta} \mathbf{e}_1, \quad \text{and} \quad U(s \circ r(g)) = U(g) + 1 - \gamma. \quad (4.14)$$

Proof. Let $g = f_{\mathbf{x}}$ and $\sigma^* := \sigma^*(r(g))$. From Lemma 4.3, the case $\mathbf{v}(g) \in \Delta_{-1}$ is equivalent to $\rho(\mathbf{x})_0 < \sigma^*$, and the case $\mathbf{v}(g) \in \Delta_0$ is equivalent to $\rho(\mathbf{x})_1 < \sigma^* \leq \rho(\mathbf{x})_0$. We apply Corollary 3.12 and Lemma 3.7. In both cases, we see that $\psi_{\sigma^*}(\mathbf{x})_{j+1} = x_j$ for $j = 1, \dots, k$, and so

$$\text{for } j = 2, \dots, k, \quad \mathbf{v}(s \circ r(g))_j = \psi_{\sigma^*}(\mathbf{x})_j - \psi_{\sigma^*}(\mathbf{x})_{j+1} = x_{j-1} - x_j = \mathbf{v}(g)_{j-1}.$$

For $\mathbf{v}(s \circ r(g))_1 = \psi_{\sigma^*}(\mathbf{x})_1 - \psi_{\sigma^*}(\mathbf{x})_2$, if $\rho(\mathbf{x})_0 < \sigma^*$, then

$$\mathbf{v}(s \circ r(g))_1 = \frac{x_0 + 1 - \gamma - \beta(x_0 + \Sigma(\mathbf{v}))}{1 - \beta} - x_1 = \frac{\gamma - \langle \mathbf{u}, \mathbf{v}(g) \rangle}{1 - \beta} + \frac{\beta}{1 - \beta} \sum_{i=1}^{k-1} v_i + \frac{1 - k\beta}{1 - \beta} v_k,$$

and if $\rho(\mathbf{x})_1 < \sigma^* \leq \rho(\mathbf{x})_0$, then

$$\mathbf{v}(s \circ r(g))_1 = \sigma^* - x_1 = x_0 - x_1 + \Sigma(\mathbf{v}) = - \sum_{i=1}^{k-1} v_i - \frac{1 - k\beta}{\beta} v_k + \frac{1 - \gamma}{\beta}.$$

Finally the results for $U(s \circ r(g)) - U(g)$ follow directly from $\psi_{\sigma^*}(\mathbf{x})_0 - x_0$. $\square$

Remark 4.5. The affine transformations of Lemma 4.4 are the only ones needed for future proofs, but it may be of interest to the reader that affine transformations analogous to those in Lemma 4.4 exist if $g \in \mathcal{T}$ and $\mathbf{v}(g)$ belongs to neither Δ_{-1} nor Δ_0 . We partition Δ into $\Delta = \Delta_{-1} \cup \Delta_0 \cup \Delta_1 \cup \dots \cup \Delta_k$ by further defining

$$\text{for } i = 1, \dots, k, \quad \Delta_i := \left\{ \mathbf{v} \in \Delta : - \sum_{j=1}^i v_j < \Sigma(\mathbf{v}) - \langle \mathbf{u}, \mathbf{v} \rangle + \gamma \leq - \sum_{j=1}^{i-1} v_j \right\}.$$

For $i = 1, \dots, k-1$, let $I_i \in \mathbb{R}^{i \times i}$ be the identity matrix and define the $k \times k$ matrix A_i by

$$A_i := \begin{pmatrix} I_{i-1} & 0 \\ 0 & M_i \end{pmatrix} \quad \text{where} \quad M_i := \begin{pmatrix} 1 & 1 & 1 & \dots & 1 & \frac{1-k\beta}{\beta} \\ 0 & -1 & -1 & \dots & -1 & \frac{-(1-k\beta)}{\beta} \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

From case by case calculations using Lemmas 3.7 and 4.3, one may show that if $\mathbf{v}(g) \in \Delta_i$ for $i = 1, \dots, k-1$, then

$$\mathbf{v}(s \circ r(g)) = A_i \mathbf{v} + \frac{1 - \gamma}{\beta} (\mathbf{e}_{i+1} - \mathbf{e}_i) \quad \text{and} \quad U(s \circ r(g)) = U(g) + 1 - \gamma,$$

while if $\mathbf{v}(g) \in \Delta_k$, then

$$\mathbf{v}(s \circ r(g)) = I_k \mathbf{v} - \frac{1 - \gamma}{1 - k\beta} \mathbf{e}_k \quad \text{and} \quad U(s \circ r(g)) = U(g) + 1 - \gamma.$$

4.2 Proof of Theorem 1.4

The aim of this section is to prove the following proposition which establishes the existence and uniqueness of a traveling wave solution; the proof of Theorem 1.4 will then follow as a consequence.

Proposition 4.6. *There is a unique traveling wave solution (G, ν) , which is given by*

$$G = f_{\mathbf{y}}, \quad \text{where } \mathbf{y} := \left(0, (-\chi \vee 0) - (j-1)\nu\right)_{j=1}^K, \quad (4.15)$$

and $\nu = \nu(\gamma, \beta)$, $\chi = \chi(\gamma, \beta)$ are defined in the statement of Theorem 1.4.

We postpone the proof of Proposition 4.6 until the end of the section. Recalling Lemmas 4.1 and 4.2, we first seek vectors that are fixed points to the affine transformations of Lemma 4.4. In the following two lemmas, we prove that starting from any non-negative vector $\mathbf{v} \in \mathbb{R}^k$, after successive applications of (4.13) or (4.14), the resulting sequence of vectors converge at geometric rate to a unique fixed point.

Lemma 4.7. *Let $\beta \in (0, 1)$, and define*

$$\boldsymbol{\mu}_L := (\nu_L, \nu_L, \dots, \nu_L) \in \mathbb{R}^k, \quad \text{where } \nu_L = \frac{\gamma}{\sum_{j=1}^k (1-j\beta)} = \frac{2\gamma}{k(2-(k+1)\beta)}.$$

For $\mathbf{v} \in \mathbb{R}^k$, define $\mathbf{v}^t$ recursively by

$$\mathbf{v}^0 = \mathbf{v}, \quad \text{and} \quad \forall t \geq 1, \quad \mathbf{v}^t = A_{-1} \mathbf{v}^{t-1} + \frac{\gamma - \langle \mathbf{u}, \mathbf{v}^{t-1} \rangle}{1-\beta} \mathbf{e}_1. \quad (4.16)$$

There exists constants $C > 0$ and $0 < r < 1$ such that for all non-negative $\mathbf{v} \in \mathbb{R}^k$,

$$\forall t \geq 1, \quad \|\mathbf{v}^t - \boldsymbol{\mu}_L\|_\infty \leq Cr^t \|\mathbf{v} - \boldsymbol{\mu}_L\|_\infty.$$

Proof. From (4.16), we see that for all $t \geq 0$,

$$\gamma - \langle \mathbf{u}, \mathbf{v}^{t+1} \rangle = \gamma - \beta \sum_{i=1}^{k-1} v_i^t - (1-k\beta)v_k^t - \gamma + \langle \mathbf{u}, \mathbf{v}^t \rangle - \sum_{i=1}^{k-1} (1-(i+1)\beta)v_i^t = 0$$

We may therefore assume that $\langle \mathbf{u}, \mathbf{v} \rangle = \gamma$ in (4.16) (and so $\mathbf{v}^t = A_{-1}^t \mathbf{v}$). Note that A_{-1} is an irreducible and aperiodic Leslie matrix with leading eigenvalue $\lambda_1 = 1$; A_{-1} is a row stochastic matrix and since its first row is positive, A_{-1} is also primitive (and so aperiodic). The vectors $\mathbf{u}$ and $\boldsymbol{\mu}_L$ are left and right eigenvectors associated with λ_1 respectively. It follows from standard Perron-Frobenius Theory (see for example [17]) that if $\mathbf{v}$ is a non-negative vector satisfying $\langle \mathbf{u}, \mathbf{v} \rangle = \gamma$, then $A_{-1}^t \mathbf{v}$ converges to $\boldsymbol{\mu}_L$ at a geometric rate, completing the proof. $\square$

Lemma 4.8. *Let $\beta \in (0, 1)$ and define*

$$\boldsymbol{\mu}_F := (1-\gamma, \dots, 1-\gamma) \in \mathbb{R}^k.$$

For $\mathbf{v} \in \mathbb{R}^k$, define $\mathbf{v}^t$ recursively by

$$\mathbf{v}^0 = \mathbf{v}, \quad \text{and} \quad \forall t \geq 1, \quad \mathbf{v}^t = A_0 \mathbf{v}^{t-1} + \frac{1-\gamma}{\beta} \mathbf{e}_1. \quad (4.17)$$

If $\beta^{-1} \notin \mathbb{N}$, then there exists constants $C > 0$ and $0 < r < 1$ such that for all $\mathbf{v} \in \mathbb{R}^k$,

$$\forall t \geq 1, \quad \|\mathbf{v}^t - \boldsymbol{\mu}_F\|_\infty \leq Cr^t \|\mathbf{v} - \boldsymbol{\mu}_F\|_\infty.$$

Proof. We can directly verify that

$$\boldsymbol{\mu}_F = A_0 \boldsymbol{\mu}_F + \frac{1-\gamma}{\beta} \mathbf{e}_1. \quad (4.18)$$

Define $\boldsymbol{\varepsilon}^t := \mathbf{v}^t - \boldsymbol{\mu}_F$. We can then rework (4.17) with the aid of (4.18) to get

$$\boldsymbol{\varepsilon}^{t+1} = A_0 \boldsymbol{\varepsilon}^t. \quad (4.19)$$

The matrix A_0 is a permuted polynomial companion matrix, and so by reading off the first row of A_0 we see immediately that its characteristic polynomial is

$$\lambda^k + \lambda^{k-1} + \cdots + \lambda + \frac{1-k\beta}{\beta}. \quad (4.20)$$

From (4.20), we see that 1 is not an eigenvalue of A_0 , and therefore the unique solution to $\boldsymbol{\varepsilon} = A_0 \boldsymbol{\varepsilon}$ is the zero vector; and so from (4.19), $\boldsymbol{\mu}_F$ is the unique solution to (4.18).

Suppose $\beta \notin \mathbb{N}$. From the Eneström-Keakeya Theorem [14, 18], if $a_0 + a_1x + \cdots + a_nx^n$ is a polynomial with real coefficients such that $0 \leq a_0 \leq \cdots \leq a_n$, then all of its roots lie in the unit disk $|z| \leq 1$. Therefore, the eigenvalues λ of A_0 satisfy $|\lambda| \leq 1$. Let $c = \frac{1-k\beta}{\beta}$, suppose $0 < c < 1$ (which holds when $\beta^{-1} \notin \mathbb{N}$), and assume there is an eigenvalue with $|\lambda| = 1$, i.e., $\lambda = e^{i\theta}$ for some θ . Then

$$\sum_{n=0}^k e^{in\theta} = \lambda^k + \lambda^{k-1} + \cdots + \lambda + c + (1-c) = 1 - c,$$

and it can be worked out that

$$\sum_{n=0}^k e^{in\theta} = \frac{\sin(\frac{1}{2}(k+1)\theta)}{\sin(\frac{1}{2}\theta)} e^{\frac{ik\theta}{2}}.$$

Since $1 - c$ is real, this means that in the exponent of e , $k\theta = 2M\pi$ for some integer M . Then $\sin(\frac{1}{2}(k+1)\theta) = \sin(M\pi + \frac{1}{2}\theta) = \sin(\frac{1}{2}\theta)$, which then means $\sum_{n=0}^k e^{in\theta} = \pm 1 \neq 1 - c$, and so $e^{i\theta}$ cannot be an eigenvalue of A_0 . Therefore, $|\lambda| < 1$ for all eigenvalues of A_0 . Since A'_0 has spectral radius less than 1, we see that recursive applications of (4.19) yields that $\|\boldsymbol{\varepsilon}^t\|_\infty$ vanishes at a geometric rate. Since $\mathbf{v}^t = \boldsymbol{\mu}_F + \boldsymbol{\varepsilon}^t$, the statement of the lemma follows. $\square$

Remark 4.9. If $\beta^{-1} \in \mathbb{N}$, then $\frac{1-k\beta}{\beta} = 1$, and $\lambda^k + \cdots + \lambda + 1$ has as its roots all the $(k+1)^{th}$ roots of unity except 1, and so $|\lambda| = 1$ for all eigenvalues of A_0 . In fact in this case, we may verify that A_0 has period k , and so as $t \rightarrow \infty$, we see that $\boldsymbol{\varepsilon}^t$ will simply cycle through $\boldsymbol{\varepsilon}^0, \dots, \boldsymbol{\varepsilon}^k$. Thus $\mathbf{v}^t$ does not converge in this case.

From Lemmas 4.7 and 4.8, we have shown that there is at most one fixed point for each of the two affine transformations of Lemma 4.4. To prove Proposition 4.6, we need to show that these solutions exist by showing that either

$$\boldsymbol{\mu}_L \in \Delta_{-1} \quad \text{or} \quad \boldsymbol{\mu}_F \in \Delta_0. \quad (4.21)$$

We show in the next lemma that $\gamma_c(\beta)$ as defined in (1.15) is the critical value of γ for which we transition from the first inclusion of (4.21) holding (when $\gamma < \gamma_c(\beta)$) to the second inclusion of (4.21) holding (when $\gamma \geq \gamma_c(\beta)$).

Lemma 4.10. For $\gamma, \beta \in (0, 1)$,

(i) $\mu_L \in \Delta_{-1}$ iff $\gamma < \gamma_c(\beta)$, in which case $\Sigma(\mu_L) = \beta^{-1}((\beta - 1)\nu_L + 1 - \gamma)$.

(ii) $\mu_F \in \Delta_0$ iff $\gamma \geq \gamma_c(\beta)$, in which case $\Sigma(\mu_F) = \langle \mathbf{u}, \mu_F \rangle + 1 - 2\gamma$.

Proof. The key observation underlying the proof is that our choice of $\gamma_c := \gamma_c(\beta)$ entails that

$$\text{for } \gamma \in (0, 1), \quad \begin{array}{ll} \nu_L < 1 - \gamma & \text{if and only if } \gamma < \gamma_c(\beta), \\ \gamma \geq \langle \mathbf{u}, \mu_F \rangle & \text{if and only if } \gamma \geq \gamma_c(\beta), \end{array} \quad (4.22)$$

which follows by rewriting (1.15) as

$$\gamma_c = \frac{\sum_{j=1}^k (1 - j\beta)}{1 + \sum_{j=1}^k (1 - j\beta)} \implies (1 - \gamma_c) \sum_{j=1}^k (1 - j\beta) = \gamma_c.$$

Let $\gamma < \gamma_c(\beta)$. Then by (4.22), $\langle \mathbf{u}, \mu_F \rangle > \gamma$, and so does not belong to $\Delta_0 \subset \Delta$. In order to prove that $\mu_L \in \Delta_{-1}$, we need to show that $\Sigma(\mu_L) > 1 - \gamma$. From (4.22), the first condition in (4.5) is satisfied and a direct computation shows that

$$\Sigma(\mu_L) = \frac{(\beta - 1)\nu_L + 1 - \gamma}{\beta}. \quad (4.23)$$

Since $\nu_L < 1 - \gamma$ and $\beta - 1 < 0$, this yields $\Sigma(\mu_L) > 1 - \gamma$ and $\mu_L \in \Delta_{-1}$.

Let us now take $\gamma \geq \gamma_c(\beta)$. For μ_F , The first condition of (4.5) is satisfied when $\beta^{-1} \notin \mathbb{N}$, while the second condition is satisfied when $\beta^{-1} \in \mathbb{N}$. Both cases yield

$$\Sigma(\mu_F) = -\gamma + \langle \mathbf{u}, \mu_F \rangle + 1 - \gamma,$$

and so $-\gamma + \langle \mathbf{u}, \mu_F \rangle < \Sigma(\mu_F) \leq 1 - \gamma$ and $\mu_F \in \Delta_0$. It remains to prove that $\mu_L \notin \Delta_{-1}$. If the first condition of (4.5) is satisfied, we have the same expression as (4.23), and since $\nu_L \geq 1 - \gamma$, it follows that $\Sigma(\mu_L) \leq 1 - \gamma$. If the second condition of (4.5) is satisfied, then $\nu_L \geq \frac{1-\gamma}{1-k\beta}$, and $\Sigma(\mu_L) \leq -\nu_L + \frac{1-\gamma}{1-k\beta} < 0$. Both conditions yield $\Sigma(\mu_L) \leq 1 - \gamma$, and so $\mu_L \notin \Delta_{-1}$. $\square$

The proofs of Proposition 4.6 and Theorem 1.4 now follow from the previous lemmas.

Proof of Proposition 4.6. For $\beta \in (0, 1)$, Lemmas 4.1, 4.2, 4.4, along with Lemmas 4.7–4.10 imply that

$$\mathbf{v}(G) = \begin{cases} \mu_L & \text{if } \gamma < \gamma_c(\beta), \\ \mu_F & \text{if } \gamma \geq \gamma_c(\beta), \end{cases} \quad (4.24)$$

while $\mathbf{v}(G) \in \Delta_0$ is the empty vector when $\beta \geq 1$. By Lemma 4.4 and Lemma 4.10, a quick calculation yields

$$\nu := U(s \circ r(G)) - U(G) = \begin{cases} \nu_L & \text{if } \gamma < \gamma_c(\beta), \\ 1 - \gamma & \text{if } \gamma \geq \gamma_c(\beta), \end{cases}$$

and so $\nu = \nu(\gamma, \beta)$ as defined in Theorem 1.4. Recalling μ_L from Lemma 4.7 and μ_F from Lemma 4.8, we thus have $\mathbf{v}(G) = (\nu, \dots, \nu) \in \mathbb{R}^k$ and

$$\gamma - \langle \mathbf{u}, \mathbf{v}(G) \rangle = \begin{cases} 0 & \text{if } \gamma < \gamma_c(\beta), \\ -\chi & \text{if } \gamma \geq \gamma_c(\beta), \end{cases}$$

where $\chi := \chi(\gamma, \beta)$ is defined in (1.16). Thus if $G = f_{\mathbf{y}}$, we see from (4.1) and (4.2) that $y_0 = 0$ and $y_j = -(\chi \vee 0) - (j-1)\nu$ for $j = 1, \dots, k+1$. We need only find y_j for $j > k+1$. By Corollary 3.12 and Lemma 3.7, we see that $\psi_{\Sigma(\mathbf{v}(G))}(\mathbf{y})_j = y_{j-1}$ for all $j = k+1, \dots$. Since (G, ν) is a traveling wave, this implies $\psi_{\Sigma(\mathbf{v}(G))}(\mathbf{y})_j = y_j + \nu$, and so

$$\mathbf{y} = \left(0, (-(\chi \vee 0) - (j-1)\nu)_{j=1}^K \right),$$

completing the proof. $\square$

Proof of Theorem 1.4. The statements on the speed of the wave (a.1) and (b.1) follow directly from Proposition 4.6. Let $g^0 = G$ and $g^1 = s \circ r(g^0)$; the statements (a.2) and (b.2) follow by evaluating $\sigma^* = \Sigma(\mathbf{v}(G))$ with the help of Lemma 4.10 and (4.24). If $\gamma < \gamma_c$ then $\Sigma(\mathbf{v}(G)) > 1 - \gamma$ and so $g^1 = r(g^0)(x) + \beta(x - \sigma^*) < r(g^0)(x)$ for all $x \in \text{Supp}(g^1)$, and if $\gamma \geq \gamma_c$ then $\Sigma(\mathbf{v}(G)) = -\chi + 1 - \gamma$ and so $g^1(x) = r(g^0)(x)$ for all $x \in \text{Supp}(g^1) \cap [-\chi + 1 - \gamma, \infty) = [-\chi + 1 - \gamma, 1 - \gamma]$. $\square$

4.3 Proof of Theorem 1.5

Recall the definition of ϕ :

$$\phi(g_1, g_2) := \sup_{x \in \mathbb{R}} |g_1(x)_+ - g_2(x)_+| \vee |U(g_1) - U(g_2)| \vee |L(g_1) - L(g_2)|.$$

Throughout this section, we will denote $\Sigma := \Sigma(\mathbf{v}(G))$, where $G = f_{\mathbf{y}}$ is the traveling wave solution defined in (4.15) of Proposition 4.6. To prove Theorem 1.5, we will make use of the following continuity lemmas, the proofs of which follow by technical applications of previous lemmas and are relegated to Section 4.3.1.

Lemma 4.11. *The transformation $s \circ r$ is continuous (with respect to ϕ) at G .*

Lemma 4.12. *There exists a constant $C > 0$ such that for any $g, g' \in \mathcal{T}$,*

$$\|\mathbf{v}(g) - \mathbf{v}(g')\|_{\infty} \leq C\phi(g, g').$$

Lemma 4.13. *Let $\mathbf{x} \in \mathbb{R}^{K+1}$ be non-increasing. There exists a constant C such that*

(i): *For $K \in \mathbb{N}$ and $\|\mathbf{y} - \mathbf{x}\|_{\infty}$ small enough,*

$$\phi(G, f_{\mathbf{x}}) \leq C\|\mathbf{y} - \mathbf{x}\|_{\infty}.$$

(ii): *For $K = \infty$ with $M = \min\{j \in \mathbb{N} : y_{j+1} < L(G)\}$ and $\max_{j=0, \dots, M} |y_j - x_j|$ small enough,*

$$\phi(G, f_{\mathbf{x}}) \leq C \max_{j=0, \dots, M} |y_j - x_j|.$$

To outline the proof of Theorem 1.5, Lemmas 4.11 and 4.12 allow us to guarantee that once $g^{t_\ell} \in \mathcal{T}$ (recall Theorem 3.9), then $\mathbf{v}(g^{t_\ell})$ is close to $\mathbf{v}(G)$ whenever $\phi(G, g^0)$ is small enough. We then apply the convergence Lemmas 4.7 and 4.8 of Section 4.2 to show that $\mathbf{v}(g^t)$ converges at a geometric rate to $\mathbf{v}(G)$. To apply these convergence lemmas we need (from Lemma 4.4) that $\mathbf{v}(g^t)$ and $\mathbf{v}(G)$ belong to the same set Δ_i for all $t \geq t_\ell$; this condition prevents us from being able to prove a global convergence result in this article, but does hold if $\mathbf{v}(G)$ is an interior point of either Δ_{-1} or Δ_0 and $\mathbf{v}(g^{t_\ell})$ is close enough to $\mathbf{v}(G)$. Since $g^t \in \mathcal{T}$ then $g^t = f_{\mathbf{x}^t}$ for some $\mathbf{x}^t$, and we show that since $\mathbf{v}(g^t)$ converges to $\mathbf{v}(G)$ at a geometric rate, so does $\mathbf{x}^t$ converge to a translation of $\mathbf{y}$ at a geometric rate. The result then follows by applying Lemma 4.13.

Proof of Theorem 1.5. Throughout, let $\mathbf{y}^t := (y_0 + t\nu, y_1 + t\nu, \dots, y_K + t\nu)$ so that $G^t := f_{\mathbf{y}^t} = G(\cdot - t\nu)$, and recall t_ℓ from Theorem 3.9.

By applying Lemma 4.11 (and translations thereof), we see that

$$g^0 \mapsto g^{t_\ell}$$

is a composition of finitely many continuous mappings (w.r.t. ϕ) around $G = G^0$, and since $G^{t_\ell}, g^{t_\ell} \in \mathcal{T}$, we may apply Lemma 4.12 and guarantee that $\|\mathbf{v}(G^{t_\ell}) - \mathbf{v}(g^{t_\ell})\|_\infty$ is arbitrarily small for all $\phi(G, g^0)$ small enough.

For the rest of the proof, let $t \geq t_\ell$ so that $g^t \in \mathcal{T}$. The next step is to prove the existence of constants $C_1 > 0$ and $0 < r < 1$ such that

$$\|\mathbf{v}(g^t) - \mathbf{v}(G)\|_\infty < C_1 r^t. \quad (4.25)$$

When $\beta \geq 1$, (4.25) is vacuously true since $\mathbf{v}(G)$ and $\mathbf{v}(g^t)$ are both empty, so now let $\beta \in (0, 1)$. From Lemma 4.3 (and since $1 - j\beta < 1$ for $j = 0, \dots, k$) we can see from (4.5) that there exists a universal constant C' such that

$$|\Sigma(\mathbf{v}(g^t)) - \Sigma| \leq C' \|\mathbf{v}(g^t) - \mathbf{v}(G)\|_\infty. \quad (4.26)$$

Thus, recalling the definition of Δ_i from (4.12), if $\|\mathbf{v}(G) - \mathbf{v}(g^t)\|_\infty$ is small enough, then (4.26) guarantees that $\mathbf{v}(g^t)$ and $\mathbf{v}(G)$ belong to the same set Δ_i (recall that we specify that $\gamma \neq \gamma_c$, and so $\mathbf{v}(G)$ is an interior point of the set Δ_i it belongs to). Lemma 4.4 and Lemma 4.10, along with Lemma 4.7 when $\gamma < \gamma_c$ and Lemma 4.8 when $\gamma > \gamma_c$, $\beta^{-1} \notin \mathbb{N}$ guarantee that if $\|\mathbf{v}(g^{t_\ell}) - \mathbf{v}(G)\|_\infty$ is small enough, then $\|\mathbf{v}(g^t) - \mathbf{v}(G)\|_\infty$ remains small enough such that for all $t \geq t_\ell$, $\mathbf{v}(g^t)$ and $\mathbf{v}(G)$ belong to the same set Δ_i and that (4.25) holds.

For all values of $\beta > 0$ and $t \geq t_\ell$, since $\mathbf{v}(G)$ and $\mathbf{v}(g^t)$ belong to the same set Δ_i , we see from Lemma 4.4 and (4.26) that $|(U^{t+1} - U^t) - \nu| = 0$ when $\gamma \geq \gamma_c(\beta)$ and is bounded by a constant times $\|\mathbf{v}(g^t) - \mathbf{v}(G)\|_\infty$ when $\gamma < \gamma_c(\beta)$. From (4.27), if we let

$$c := \sum_{t=0}^{\infty} ((U^{t+1} - U^t) - \nu) < \infty,$$

then there exists a constant $C_2 > 0$ such that

$$|(U^t - c) - t\nu| = \left| \sum_{s=t+1}^{\infty} ((U^{s+1} - U^s) - \nu) \right| < C_2 r^t. \quad (4.27)$$

Since $g^t \in \mathcal{T}$ for all $t \geq t_\ell$, let $\mathbf{x}^t \in \mathbb{R}^{K+1}$ be a non-increasing vector for which $g^t = f_{\mathbf{x}^t}$. For all $i = 0, 1, \dots, k+1$, from (4.25), (4.27) and (4.2), along with the definition of $\mathbf{v}(g^t)$ in (4.1),

$$|(x_i^t - c) - y_i^t| \leq |(U^t - c) - t\nu| + |\langle \mathbf{u}, \mathbf{v}(g^t) - \mathbf{v}(G) \rangle| + \sum_{j=0}^i |\mathbf{v}(g^t)_j - \mathbf{v}(G)_j| \leq C_3 r^t \quad (4.28)$$

for a constant C_3 . Since $\mathbf{v}(g^t)$ belongs to Δ_{-1} or Δ_0 (and so $\sigma^*(r(g^t)) > x_1$), by Lemma 3.7,

$$x_{j+1}^{t+1} = x_j^t \quad \text{for } j = 1, 2, \dots, K-1. \quad (4.29)$$

Starting with the case $K \in \mathbb{N}$, we see from (4.28) and (4.29) that for all t large enough, $\|(\mathbf{x}^t - c(1, 1, \dots, 1)) - \mathbf{y}^t\|_\infty \leq C' r^t$, and applying Lemma 4.13 and a translation yields

$$\phi(g^t(\cdot + t\nu + c), G) \leq C r^t$$

for some constant C . For $K = \infty$, the argument is similar with K replaced by M as defined in Lemma 4.13. $\square$

4.3.1 Proofs of Continuity Lemmas

In this section we provide proofs to the continuity lemmas 4.11 – 4.13.

Proof of Lemma 4.13. We start with $K \in \mathbb{N}$. For $\|\mathbf{x} - \mathbf{y}\|_\infty$ small enough, we may guarantee that on $\text{Supp}(G) \cap \text{Supp}(f_{\mathbf{x}})$, the slopes of the linear sections of G and $f_{\mathbf{x}}$ differ by at most β , in which case the slopes only differ on the intervals $(y_j \wedge x_j, y_j \vee x_j)$. Thus for $x \in \text{Supp}(G) \cap \text{Supp}(f_{\mathbf{x}})$,

$$|G(x) - f_{\mathbf{x}}(x)| \leq \int_x^{x_0 \wedge y_0} |G'(z) - f'_{\mathbf{x}}(z)| dz + \int_{x_0 \wedge y_0}^{x_0 \vee y_0} dz \leq \beta \sum_{j=1}^K |y_j - x_j| + |y_0 - x_0|.$$

For $x \notin \text{Supp}(G) \cap \text{Supp}(f_{\mathbf{x}})$, since both $G, f_{\mathbf{x}}$ are concave and $0 \vee G, 0 \vee f_{\mathbf{x}}$ are continuous, we see immediately that $|0 \vee G(x) - 0 \vee f_{\mathbf{x}}(x)|$ is bounded by the maximum value of this difference in $\text{Supp}(G) \cap \text{Supp}(f_{\mathbf{x}})$, completing the argument.

For $K = \infty$, the argument is similar to the one above but with K replaced by M . The condition that $\max_{j=0, \dots, M} |y_j - x_j|$ be small enough is used to guarantee that $L(f_{\mathbf{x}}) > x_M$, and so $0 \vee f_{\mathbf{x}}$ is continuous. $\square$

Proof of Lemma 4.11. Before dwelling into the proof, the careful reader may notice that the transformation $s \circ r$ is not continuous everywhere in $\mathcal{D}$. However, around G , we will sandwich the function with two piecewise linear functions to control the effect of the transformation.

For conciseness, we assume that $K < \infty$. The continuity of the following maps follow from Lemmas 3.7 and 4.13.

- (1): $(\sigma, \mathbf{x}) \mapsto f_{\psi_\sigma(\mathbf{x})}$ is continuous at $(\Sigma, \mathbf{y})$.
- (2): For non-increasing $\mathbf{x} \in \mathbb{R}^{K+1}$ let $\sigma_{\mathbf{x}} := \sigma^*(r(f_{\mathbf{x}}))$, then $\mathbf{x} \mapsto \sigma_{\mathbf{x}}$ is continuous at $\mathbf{y}$.

Let $\sigma^* := \sigma^*(r(g))$. To proceed, it is evident that for all δ we may find non-increasing vectors $\mathbf{y}^{\pm\delta} \in \mathbb{R}^{K+1}$ near $\mathbf{y}$ such that $f_{\mathbf{y}^{-\delta}} \leq g \leq f_{\mathbf{y}^{+\delta}}$ for all $\phi(G, g) < \delta$. For all $\sigma \in \mathbb{R}$, the operator $s_\sigma \circ r$ maintains the inequality

$$s_\sigma \circ r(f_{\mathbf{y}^{-\delta}}) \leq s_\sigma \circ r(g) \leq s_\sigma \circ r(f_{\mathbf{y}^{+\delta}}), \quad (4.30)$$

from which we may also conclude that $\sigma_{\mathbf{y}^{-\delta}} \leq \sigma^* \leq \sigma_{\mathbf{y}^{+\delta}}$. By applying (2) to $\mathbf{y}^{\pm\delta}$, we see that we may choose δ small enough to guarantee that $|\Sigma - \sigma^*|$ is arbitrarily small. From this arbitrarily small bound on $|\Sigma - \sigma^*|$ as well as applying (1) to $(\sigma^*, \mathbf{y}^{\pm\delta})$, we conclude from (4.30) the statement of the lemma. $\square$

Proof of Lemma 4.12. Let $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^{K+1}$ such that $g = f_{\mathbf{x}}$ and $g' = f_{\mathbf{x}'}$. The upper edges of the functions differ by $|x_0 - x'_0|$, and so

$$\phi(f_{\mathbf{x}}, f_{\mathbf{x}'}) \geq |x_0 - x'_0|.$$

Fix $j = 1, \dots, k+1$, and assume without loss of generality that $x_j < x'_j$. We have by definition of the functions that for all $x_j < z \leq x_0 \wedge x'_j$,

$$f_{\mathbf{x}'}(x_j) - f_{\mathbf{x}'}(z) \leq (1 - j\beta)(z - x_j), \quad f_{\mathbf{x}}(x_j) - f_{\mathbf{x}}(z) \geq (1 - (j-1)\beta)(z - x_j),$$

Thus,

$$\begin{aligned} 2\phi(f_{\mathbf{x}}, f_{\mathbf{x}'}) &\geq |f_{\mathbf{x}}(z) - f_{\mathbf{x}'}(z)| + |f_{\mathbf{x}}(x_j) - f_{\mathbf{x}'}(x_j)| \\ &\geq (f_{\mathbf{x}}(x_j) - f_{\mathbf{x}}(z)) - (f_{\mathbf{x}'}(x_j) - f_{\mathbf{x}'}(z)) \geq \beta(z - x_j). \end{aligned} \quad (4.31)$$

If $x'_j \leq x_0$, then (4.31) implies $\phi(f_{\mathbf{x}}, f_{\mathbf{x}'}) \geq \beta(x'_j - x_j)/2$. If $x_0 < x'_j$, then with (4.31),

$$\begin{aligned} 3\phi(f_{\mathbf{x}}, f_{\mathbf{x}'}) &\geq |f_{\mathbf{x}}(x_0) - f_{\mathbf{x}'}(x_0)| + |f_{\mathbf{x}}(x_j) - f_{\mathbf{x}'}(x_j)| + |x'_0 - x_0| \\ &\geq \beta(x_0 - x_j) + (x'_j - x_0) \geq (1 \wedge \beta)(x'_j - x_j), \end{aligned}$$

and so $\phi(f_{\mathbf{x}}, f_{\mathbf{x}'}) \geq (1 \wedge \beta)(x'_j - x_j)/3$. We therefore have a constant c such that

$$\phi(f_{\mathbf{x}}, f_{\mathbf{x}'}) \geq c \max_{j=0, \dots, k+1} |x_j - x'_j|.$$

The statement of the lemma then follows from the above, recalling the definition of $\mathbf{v}(g)$ from (4.1), and the triangle inequality. $\square$

Acknowledgement

E.S. gratefully acknowledges support from the FWF project PAT3816823 “Waves in Population Genetics”.

References

- [1] Jean Bérard and Jean-Baptiste Gouéré, Brunet-Derrida behavior of branching-selection particle systems on the line, *Comm. Math. Phys.*, 298(2):323–342, 2010.
- [2] Jean Bérard and Pascal Maillard, The limiting process of N-particle branching random walk with polynomial tails, *Electron. J. Probab.*, 19:1–17, 2014.

- [3] Julien Berestycki, Éric Brunet, and Sarah Penington, Global existence for a free boundary problem of Fisher–KPP type, *Nonlinearity*, 32(10):3912, 2019.
- [4] Eric Bolthausen and Alain-Sol Sznitman, On ruelle’s probability cascades and an abstract cavity model, *Comm. Math. Phys.*, 197(2):247–276, 1998.
- [5] Éric Brunet and Bernard Derrida, Shift in the velocity of a front due to a cut-off, *Phys. Rev. E*, 56:2597–2604, 1997.
- [6] Éric Brunet and Bernard Derrida, Microscopic models of traveling wave equations, *Comput. Phys. Commun.*, 121:376–381, 1999.
- [7] Éric Brunet, Bernard Derrida, Alfred H. Mueller, and Stéphane Munier, Effect of selection on ancestry: An exactly soluble case and its phenomenological generalization, *Phys. Rev. E.*, 76(4):041104, 2007.
- [8] Éric Brunet, Igor M. Rouzine, and Claus O. Wilke, The stochastic edge in adaptive evolution, *Genetics*, 179(1):603–620, 2008
- [9] Aser Cortines and Bastien Mallein, A N -branching random walk with random selection, *ALEA Lat. Am. J. Probab. Math. Stat.*, 92:117–137, 2017.
- [10] Anna De Masi, Pablo A. Ferrari, Errico Presutti, and Nahuel Soprano-Loto, Hydrodynamics of the N-BBM process, *In International workshop on Stochastic Dynamics out of Equilibrium*, 523–549, 2017
- [11] Michael M. Desai, Daniel S. Fisher, Beneficial mutation selection balance and the effect of linkage on positive selection, *Genetics*, 176(3):1759–1798, 2007
- [12] Freddy Dumortier, Nikola Popović, and Tasso J. Kaper. The critical wave speed for the Fisher–Kolmogorov–Petrovskii–Piscounov equation with cut-of, *Nonlinearity*, 20(4):855, 2007.
- [13] Rick Durrett and Daniel Remenik, Brunet–Derrida particle systems, free boundary problems and Wiener–Hopf equations. *Ann. Probab.*, 39(6):2043–2078, 2011.
- [14] Gustaf Eneström, Remarque sur un théorème relatif aux racines de l’équation $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$ où tous les coefficients a sont réels et positifs, *Tôhoku Math. J.*, 18:34–36, 1920.
- [15] William Feller, *An Introduction to Probability Theory and its Applications, Vol II*, Second edition, John Wiley & Sons, Inc., 1971.
- [16] Charles AE Goodhart, Problems of monetary management: the UK experience, *Monetary theory and practice: The UK experience*, 91–121, 1984.
- [17] Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Second edition, Cambridge University Press, 2013.
- [18] Sôichi Takeya, On the limits of roots of an algebraic equation with positive coefficients, *Tôhoku Math. J.*, 2:140–142, 1912.

- [19] Robert H. MacArthur and Edward O. Wilson, The theory of island biogeography, Princeton university press, 2001
- [20] Colin McDiarmid, Concentration, In *Probabilistic methods for algorithmic discrete mathematics*, 196–248, Springer, 1998.
- [21] Bastien Mallein, Francesco Paparella, Emmanuel Schertzer, and Zsófia Talyigás, Selection of the fittest or selection of the luckiest: the emergence of Goodhart’s law in evolution, arXiv preprint arXiv:2503.21849, 2025.
- [22] Carl Mueller, Leonid Mytnik, and Jeremy Quastel, Small noise asymptotics of traveling waves, *Markov Process. Related Fields*, 14(3):333–342, 2008.
- [23] Richard A. Neher and Oskar Hallatschek, Genealogies of rapidly adapting populations, *Proc. Natl. Acad. Sci. USA*, 110 (2):437–442, 2013.
- [24] Jim Pitman, Coalescents with multiple collisions, *Ann. Probab.*, 27(4):1870–1902, 1999.
- [25] Jim Pitman and Marc Yor, The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator, *Ann. Probab.*, 25(2):855–900, 1997.
- [26] Emmanuel Schertzer and Alejandro H. Wences, Relative vs absolute fitness in a population genetics model. How stronger selection may promote genetic diversity, Preprint arXiv:2301.07762, 2023.
- [27] Jason Schweinsberg, Coalescent processes obtained from supercritical Galton-Watson processes. *Stochastic Process. Appl.*, 106(1):107–139, 2003.
- [28] Jason Schweinsberg, Rigorous results for a population model with selection I: evolution of the fitness distribution, *Electron. J. Probab.*, 22:1–94, 2017.
- [29] Jason Schweinsberg, Rigorous results for a population model with selection II: genealogy of the population, *Electron. J. Probab.*, 22:1–54, 2017.