Kingman's model with random mutation probabilities: convergence and condensation I

Linglong Yuan

Department of Mathematical Sciences
University of Liverpool

Department of Mathematical Sciences
Xi'an Jiaotong-Liverpool University

January 22, 2020

Contents

Abstract

For a one-locus haploid infinite population with discrete generations, the celebrated Kingman's model describes the evolution of fitness distributions under the competition of selection and mutation, with a constant mutation probability. Letting mutation probabilities vary on generations reflects the influence of a random environment. This paper generalises Kingman's model by using a sequence of i.i.d. random mutation probabilities. For any distribution of the sequence, the weak convergence of fitness distributions to the globally stable equilibrium for any initial fitness distribution is proved. We define the condensation of the random model as that almost surely a positive proportion of the population travels to and condensates on the largest fitness value. The condensation may occur when selection is more favoured than mutation. A criterion is given to tell whether the condensation occurs or not.

Keywords and phrases. Population dynamics, Mutation-Selection balance, House of Cards, Fitness distribution, Size-biased distribution, Distributional equation.
MSC Subject classification: 60F05 (primary), 60G10, 92D15, 92D25 (secondary)

1 Motivation and background

Various biological forces interact with each other and drive the evolution of population all together. One important competing pair consists of selection and mutation. It was as early as 1937 that Haldane [14] put forward the concept of mutation-selection balance. The foundations of this subject was given by Crow and Kimura [7], Ewens [11], and Kingman [19]. For more details on this topic, we refer to Bürger [5, 6].

A simple setting is to consider a one-locus haploid infinite population with discrete generations under selection and mutation. The locus is assumed to have infinitely many possible alleles which have continuous effects on a quantitative type. The continuum-of-alleles models were introduced by Crow and Kimura [7] and Kimura [16] and are used frequently in quantitative genetics.

Kingman [17] suggested to explain the tendency that most mutations are deleterious by the assumption of the independence of the gene before and after mutation. This feature was named "House of Cards", as the mutation destroys the biochemical house of cards built up by evolution, by Kingman in [18] where the most famous one-locus model was proposed. In this model, a population is characterised by its type distribution, which is a probability measure on $[0, 1]$ and any $x \in [0, 1]$ is a type value. In Kingman's setting, an individual with a larger type value is fitter, which means more productive. So the type value can also be named fitness value. Kingman's model can be seen as the limit of a finite population model, see [13].

Bürger [4] generalised the selection mechanism which allows the gene after mutation to depend on that before and proved the convergence in total variation. The genetic variation of the equilibrium distribution was computed and discussed. I proposed [21] a more general selection mechanism which can model general macroscopic epistasis, with the other settings the same as in Kingman's model. This model was applied to the modelling of the Lenski experiment (see [12] for a description of the experiment).

There are also many models on the balance of mutation and selection in the setting of continuous generations. Bürger [3] provided an exact mathematical analysis of Kimura's continuum-of-alleles model, focusing on the equilibrium genetic variation. Steinsaltz *et al* [20] proposed a multi-loci model using a differential equation to study the ageing effect. Later on the recombination was incorporated to the model [10]. Betz *et al*'s model [2] generalised a continuous-time version of Kingman's model and other models arising from physics.

However to the best of my knowledge, Kingman's model has never been generalised to a random version. In this paper we will assume that the mutation probabilities of all generations form an i.i.d. sequence with the other settings unchanged. So the feature of "House of Cards" is retained. Biologically we think of a stable random environment such that the mutation probabilities are different but independently sampled from the same distribution.

In Kingman's model, the condensation occurs if a certain proportion of the population travels to and condensates on the largest fitness value. This is due to the dominance of selection over mutation. Kingman [18] called the regime with condensation *Meritocracy* or *Aristocracy* depending on contexts, and the regime without condensation *Democracy*. In the random model, we consider also the convergence of (random) fitness distributions to the equilibrium and the condensation phenomenon. Moreover, King-

Address: Department of Mathematical Sciences, The University of Liverpool, Liverpool L69 7ZL, UK.s
Email: yuanlinglongcn@gmail.com
This paper is the first half split from an article submitted to Arxiv: https://arxiv.org/abs/1903.10993

man's model has been revisited recently in terms of the travelling wave of mass to the largest fitness value [8]. The random model provides another example for consideration in this direction.

2 Models

2.1 Kingman's model with time-varying mutation probabilities

Consider a haploid population of infinite size and discrete generations under the competition of selection and mutation. We use a sequence of probability measures $(P_n) = (P_n)_{n\geq 0}$ on $[0,1]$ to describe the distribution of fitness values in the nth generation. Individuals in the nth generation are children of the $n-1$th generation. First of all, the fitness distribution of children is initially P_{n-1} (an exact copy from parents). Then selection takes effect, such that the fitness distribution is updated from P_{n-1} to the size-biased distribution

$$\frac{xP_{n-1}(dx)}{\int yP_{n-1}(dy)}.$$

Here we use $\int$ to denote $\int_0^1$. Basically the new population is resampled from the existing population by using their fitness as a selective criterion. Next, each individual mutates independent with the same mutation probability which we denote by b_n and which takes values in $[0,1)$. Each mutant has the fitness value sampled independently from a common mutant distribution, that we denote by Q, a probability measure on $[0,1]$. Then the resulting distribution is the distribution in the nth generation

$$P_n(dx) = (1-b_n)\frac{xP_{n-1}(dx)}{\int yP_{n-1}(dy)} + b_nQ(dx). \tag{1}$$

The fact that we exclude the case that a mutation probability can be 1 is because in this situation $P_n = Q$ which loses accumulated evolutionary changes. This is not interesting neither biologically nor mathematically.

Expanding (1), we can also obtain

$$P_n(dx) = \left(\prod_{l=0}^{n-1}\frac{1-b_{l+1}}{\int yP_l(dy)}\right)x^nP_0(dx) + \sum_{j=1}^{n}\left(\prod_{l=j}^{n-1}\frac{1-b_{l+1}}{\int yP_l(dy)}\right)b_jm_{n-j}Q^{n-j}(dx) \tag{2}$$

where

$$Q^k(dx) := \frac{x^kQ(dx)}{\int y^kQ(dy)}, \quad m_k := \int x^kQ(dx), \quad \forall k\geq 0.$$

In particular if $Q=\delta_0$, the dirac measure on $\{0\}$, $Q^k = \delta_0$.

When all b_n's are equal to the same number, that we denote by $b\in[0,1)$, this is the model introduced by Kingman [18]. In the general setting we allow mutation probabilities to be different. We call it *Kingman's model with time-varying mutation probabilities* or *the general model* for short.

We introduce a few more notations. Let M be the space of (nonnegative) measures on $[0,1]$ and M_1 the subspace of M consisting of probability measures. Let M, M_1 be endowed with the topology of weak convergence T. We use $\xrightarrow{d}$ to denote weak convergence. We say a sequence of measures (u_n) converges in total variation to a measure u, denoted by

$$u_n \xrightarrow{TV} u,$$

if the total variation $\|u_n - u\|$ converges to 0.

For any $u\in M_1$, define

$$S_u := \sup\{x : u[x,1] > 0\}. \tag{3}$$

So S_u is interpreted as the *largest fitness value* in a population of distribution u. Since $S_{P_1} \geq S_Q$, we can always assume $S_{P_0} \geq S_Q$, otherwise we take P_1 as P_0. We summarise it as follows

$$[S_Q, 1] \ni h := S_{P_0} \geq S_Q.$$

It is straightforward to see that $S_{P_n} = h$ for any $n \geq 0$.

The general model has parameters $(b_n)_{n\geq 1}, Q, P_0, h$. We use $\{(b_n), Q, P_0, h\}$ to denote the general model and call (P_n) *the forward sequence* or just *the sequence* of $\{(b_n), Q, P_0, h\}$. We use $\{(b), Q, P_0, h\}$ for Kingman's model.

2.2 Convergence and condensation in Kingman's model

Kingman [18] proved the convergence of (P_n) when all mutation probabilities are equal.

Theorem 1 (Kingman's theorem, [18])**.** *1. If $\int \frac{Q(dx)}{1-x/h} \geq b^{-1}$, (P_n) converges in total variation to*

$$\mathcal{K}(dx) = \frac{b\theta_b Q(dx)}{\theta_b - (1-b)x},$$

with θ_b, as a function of b, being the unique solution of

$$\int \frac{b\theta_b Q(dx)}{\theta_b - (1-b)x} = 1. \tag{4}$$

2. If $\int \frac{Q(dx)}{1-x/h} < b^{-1}$, (P_n) converges weakly to

$$\mathcal{K}(dx) = \frac{bQ(dx)}{1-x/h} + \Big(1 - \int \frac{bQ(dy)}{1-y/h}\Big)\delta_h(dx).$$

We say *there is a condensation on h in Kingman's model* if $Q(h) = 0$ but $\mathcal{K}(h) > 0$. We call $\mathcal{K}(h)$ *the condensate size* when $Q(h) = 0$. In the case 1 above, there is no condensation. The condition is satisfied only if b is big and/or Q is fit (i.e., having more mass on larger fitness values). It means mutation is stronger against selection, so that the limit does not depend on P_0.

In the case 2, the condition $\int \frac{Q(dx)}{1-x/h} < b^{-1}$ implies $Q(h) = 0$, but we see that $\mathcal{K}(h) > 0$. So there is a condensation. Contrarily to the first case, selection is more favoured so that the limit depends on P_0 through h. If $P_0(h) = 0$ (implying $S_{P_n} = h$ and $P_n(h) = 0$ for any n), a certain amount of mass $\Big(1 - \int \frac{bQ(dy)}{1-y/h}\Big)$ travels to the largest fitness value h, by the force of selection.

Next we introduce the random model, which is the main object of study in this paper.

2.3 Kingman's model with random mutations probabilities

Let β be a random variable taking values in $[0, 1)$ with $\mathbb{P}(\beta = 0) \neq 1$. Let $(\beta_n)_{n\geq 0}$ be an i.i.d. sequence sampled from the distribution of β. The *Kingman's model with random mutation probabilities* or simply *the random model* is defined by the following iteration:

$$P_n(dx) = (1 - \beta_n)\frac{xP_{n-1}(dx)}{\int yP_{n-1}(dy)} + \beta_n Q(dx), \quad n \geq 1. \tag{5}$$

We use $\{(\beta_n), Q, P_0, h\}$ to denote the random model. This model is a generalisation of Kingman's model, as β can be defined to be equal to b with probability 1.

We are interested in the convergence of (P_n) to the equilibrium and the phenomenon of condensation, resulting from the competition of selection and mutation. Since we are dealing with random probability measures, let us recall the definition of weak convergence in this context. Random probability measures (μ_n) supported on $[0,1]$ converge weakly to a limit μ if and only if for any continuous function f on $[0,1]$ we have

$$\lim_{n\to\infty} \mathbb{E}\left[f(x)\mu_n(dx)\right] = \mathbb{E}\left[f(x)\mu(dx)\right].$$

As the sequence (P_n) is completely determined by $(\beta_n), Q$ and h, the only randomness arises from (β_n). In comparison to the terminology in statistical physics, the weak limit of (P_n) is an *annealed* limit, which is obtained given the law of (β_n). A *quenched* limit, which is obtained by conditioning on the values of (β_n), is impossible, as the independent β_n's destroy the accumulated evolution, unless $P_0 = Q = \delta_0$. However we will see that it is possible to obtain a quenched limit if the evolution is seen backwards.

Consider the particular case that $Q = \delta_0$. It can be easily deduced that

$$P_n(dx) = (1-\beta_n)\frac{x^n P_0(dx)}{\int y^n P_0(dy)} + \beta_n \delta_0(dx),$$

which converges weakly to

$$(1-\beta)\delta_h(dx) + \beta\delta_0(dx).$$

So we assume from now on $Q \neq \delta_0$.

3 Main results

3.1 Weak convergence

In this part we show the convergence of (P_n) in the random model. But to explain what the limit is, we need some notations and some small results.

For the general model, we introduce the finite backward sequence $(P_j^n) = (P_j^n)_{0\le j\le n}$ which has parameters $n, (b_j)_{1\le j\le n}, Q, P_n^n, h$ with $h = S_{P_n^n}$:

$$P_j^n(dx) = (1-b_{j+1})\frac{xP_{j+1}^n(dx)}{\int yP_{j+1}^n(dy)} + b_{j+1}Q(dx), \quad \forall\, 0 \le j \le n-1. \tag{6}$$

Lemma 1. *In the general model, for the finite backward sequence with $P_n^n = \delta_h$, P_j^n converges in total variation to a limit, denoted by $\mathcal{G}_j = \mathcal{G}_{j,h}$ (and $\mathcal{G} = \mathcal{G}_0, \mathcal{G}_Q = \mathcal{G}_{0,S_Q}$), as n goes to infinity with j fixed, such that*

$$\mathcal{G}_{j-1}(dx) = (1-b_j)\frac{x\mathcal{G}_j(dx)}{\int y\mathcal{G}_j(dy)} + b_jQ(dx), \quad j \ge 1. \tag{7}$$

As a consequence, $\mathcal{G} : [0,1)^\infty \to M_1$ is a measurable function, with $\mathcal{G}_j = \mathcal{G}(b_{j+1}, b_{j+2,\ldots})$ supported on $[0, S_Q] \cup \{h\}$ for any $j \ge 0$.

Note that either $\mathcal{G}_j(h)$'s are all zero or all strictly positive. Note also that $(\mathcal{G}_j)$ has parameters $(b_{j+1}, b_{j+2}, \cdots)$ and Q, h. Define

$$\mathcal{I}_j = \mathcal{I}_{j,h} = \mathcal{G}(\beta_{j+1}, \beta_{j+2}, \cdots), \quad \mathcal{I} = \mathcal{I}_0 \text{ and } \mathcal{I}_Q = \mathcal{I}_{0,S_Q}.$$

Therefore $\mathcal{I}_j$ is the quenched limit of the finite backward sequence with $P_n^n = \delta_h$ in the random model. More generally a dynamical system is easier to handle if we take a backward point of view, see Diaconis and Freedman [9].

Corollary 1. *The sequence $(\mathcal{I}_j) = (\mathcal{I}_j)_{j\ge 0}$ is stationary ergodic and satisfies*

$$\mathcal{I}_{j-1}(dx) = (1-\beta_j)\frac{x\mathcal{I}_j(dx)}{\int y\mathcal{I}_j(dy)} + \beta_jQ(dx), \quad j \ge 1. \tag{8}$$

The convergence theorem is as follows.

Theorem 2. *For the random model $\{(\beta_n), Q, P_0, h\}$, the sequence (P_n) converges weakly to $\mathcal{I}$ whose distribution depends on β, Q, h but not on P_0.*

3.2 Condensation criterion

It is clear that if $h = S_Q$ and $Q(S_Q) > 0$, we have $\mathcal{I}(h) > 0, a.s.$. In general, we have $\mathbb{P}(\mathcal{I}(h) > 0) \in \{0, 1\}$. A justification will be provided in Corollary 3. Then we can introduce the definition of condensation in the random model.

Definition 1. *For the random model, we say there is a condensation on $\{h\}$ if $Q(h) = 0$ but $\mathcal{I}(h) > 0$ a.s..*

Theorem 3 (Condensation criterion)**.** *1. If $h = S_Q$, then there is no condensation on S_Q if*

$$\mathbb{E}\left[\ln \frac{S_Q(1-\beta)}{\int y\mathcal{I}_Q(dy)}\right] < 0. \tag{9}$$

2. If $h > S_Q$, then there is no condensation on h if and only if

$$\mathbb{E}\left[\ln \frac{h(1-\beta)}{\int y\mathcal{I}_Q(dy)}\right] \leq 0. \tag{10}$$

By Corollary 4 to be given later, if $h = S_Q$, we can only have $\mathbb{E}\left[\ln \frac{S_Q(1-\beta)}{\int y\mathcal{I}_Q(dy)}\right] \leq 0$. The fact that we cannot say anything about the occurrence of condensation by $\mathbb{E}\left[\ln \frac{S_Q(1-\beta)}{\int y\mathcal{I}_Q(dy)}\right] = 0$ can be better understood in Kingman's model, which is a special random model. In Kingman's model, $\mathbb{E}\left[\ln \frac{S_Q(1-\beta)}{\int y\mathcal{I}_Q(dy)}\right] = 0$ becomes

$$\ln \frac{S_Q(1-b)}{\int y\mathcal{K}_Q(dy)} = 0.$$

By some simple computations using Theorem 1, the above display is equivalent to

$$\int \frac{Q(dx)}{1 - x/S_Q} \leq b^{-1}.$$

But it covers cases with and without condensation.

Theorem 3 provides a condensation criterion in the random model relying on the limit which however has no explicit expression. Therefore we cannot know how the limit looks like and whether the condensation really happens in concrete cases. This problem will be solved in a follow-up paper based on a matrix approach.

3.3 Invariant measure

We introduce the notion of *invariant measure*, which includes the limit $\mathcal{I}$. We will heavily use the invariant measures in the proofs which possess interesting properties by themselves.

Definition 2 (Invariant measure)**.** *A random probability measure ν is invariant if it is supported on $[0, 1]$ and satisfies*

$$\nu(dx) \overset{d}{=} (1-\beta)\frac{x\nu(dx)}{\int y\nu(dy)} + \beta Q(dx) \tag{11}$$

where β is independent of ν.

Theorem 4 (Compoundness of invariant measures)**.** *For any invariant measure ν,*

$$\nu \overset{d}{=} \mathcal{I}_{0,S_\nu}.$$

Using the notion of invariant measures, we can solve a distributional equation in the following example. For a survey on distributional equations, we refer to Aldous and Bandyopadhyay [1].

Example 1. Consider a particular case: Q is supported only on $\{c\}$ for some $c \in (0,1)$, with $h \in (c,1)$. Let ν be an invariant measure supported on $\{c\} \cup \{h\}$. Then ν can be written as $\nu = X\delta_c + (1-X)\delta_h$ where X is a random variable taking values in $[0,1]$, and satisfies

$$X\delta_c + (1-X)\delta_h \stackrel{d}{=} (1-\beta)\frac{cX\delta_c + h(1-X)\delta_h}{cX + h(1-X)} + \beta\delta_c,$$

where β is independent of X. The above display is equivalent to

$$X \stackrel{d}{=} \frac{c + (h\beta - c)(1-X)}{c + (h-c)(1-X)}.$$

We are interested in a necessary and sufficient condition for the above equation to have a solution X with $0 \le X < 1$ a.s. (i.e., $\nu(h) > 0$ a.s.). In this case, by Theorem 4, $\nu \stackrel{d}{=} \mathcal{I}_{0,h}$. So it is equivalent to saying that $\mathcal{I}_{0,h}(h) > 0$ a.s., which means a condensation occurs on h. By Theorem 3, the condition is simply $\mathbb{E}[\ln(h(1-\beta)/c)] > 0$. Moreover as such ν is unique, the solution X is also unique.

4 Proofs

4.1 Relations between measures

We introduce firstly some notations to describe relations between measures.
1). For measures $u, v \in M$, we say u is a component of v on $[0,a]$ (resp. $[0,a)$), denoted by $u \preceq_a v$ (resp. $\preceq_{a-}$), if

$$u(A) \le v(A), \qquad \text{for any measurable set } A \subset [0,a] \text{ (resp. } [0,a)).$$

For random measures $\mu, \nu \in M$, we write $\mu \preceq_a \nu$ if there exists a couple (μ', ν') with $\mu', \nu' \in M$ such that

$$\mu' \preceq_a \nu' \text{ a.s. and } \mu' \stackrel{d}{=} \mu, \ \nu' \stackrel{d}{=} \nu. \tag{12}$$

The relation $\mu \preceq_{a-} \nu$ is defined in a similar way.
2). For measures (u_n) and u, we introduce a notation

$$u_n \preceq_a \xrightarrow{TV} u$$

which means that $u_n \preceq_a u_{n+1}$ for any n, and u_n converges in total variation to u. We define similarly $\preceq_{a-}\xrightarrow{TV}$. We also define $\preceq_a\xrightarrow{TV}$ and $\preceq_{a-}\xrightarrow{TV}$ for random measures in a similar way as in (12).
3). For any $u \in M_1$, let the distribution function of u be

$$D_u(x) := u([0,x]), \quad \forall x \in [0,1].$$

For any $u, v \in M_1$, we write $u \le v$ if $D_u(x) \ge D_v(x)$ for any $x \in [0,1]$. For random probability measures μ, ν, we also define $\mu \le \nu$ similarly as in (12).
4). For real-valued random variables ξ, η, we write $\xi \le_d \eta$ if

$$\mathbb{P}(\xi \le x) \ge \mathbb{P}(\eta \le x), \quad \forall x \in \mathbb{R}.$$

4.2 Proofs of Lemma 1 and Corollary 1

Proof of Lemma 1. We prove a stronger version below

$$\text{For any } j, \quad P_j^n \preceq_{h-} \xrightarrow{TV} \mathcal{G}_j, \text{ as } n \to \infty. \tag{13}$$

It suffices to show that

$$P_j^n \preceq_{h-} P_j^{n+1}$$

as P_j^n's are all supported on $[0, S_Q] \cup \{h\}$.

First of all, $P_n^n = \delta_h \preceq_{h-} P_n^{n+1}$. Assume for some $1 \le j \le n$, we have $P_j^n \preceq_{h-} P_j^{n+1}$. By definition

$$P_{j-1}^n(dx) = (1-b_j)\frac{xP_j^n(dx)}{\int yP_j^n(dy)} + b_jQ(dx), \;\; P_{j-1}^{n+1}(dx) = (1-b_j)\frac{xP_j^{n+1}(dx)}{\int yP_j^{n+1}(dy)} + b_jQ(dx). \tag{14}$$

Since $P_j^n \preceq_{h-} P_j^{n+1}$, we have

$$\int yP_j^n(dy) \ge \int yP_j^{n+1}(dy)$$

and thus

$$\frac{x}{\int yP_j^n(dy)} \le \frac{x}{\int yP_j^{n+1}(dy)}, \quad \forall x \in [0,1].$$

Together with $P_j^n \preceq_{h-} P_j^{n+1}$ and (14), we get $P_{j-1}^n \preceq_{h-} P_{j-1}^{n+1}$. By induction,

$$P_j^n \preceq_{h-} P_j^{n+1}, \text{ for any } 0 \le j \le n, n \ge 0. \tag{15}$$

$\square$

The monotonicity analysis in the above proof will be used many times in this paper. An immediate application is the following: we can compare $(\mathcal{G}_j)$ and $(\mathcal{G}'_j = \mathcal{G}_{j,h'})$ for $h < h'$ with the same $(b_j), Q$ as follows.

Corollary 2. *Let $(\mathcal{G}_j)$ and $(\mathcal{G}'_j)$ be the above sequences. Then we have*

$$\mathcal{G}'_j \preceq_{h-} \mathcal{G}_j, \quad \mathcal{G}'_j(h') \ge \mathcal{G}_j(h), \quad \forall j \ge 0. \tag{16}$$

Moreover we have the exact equalities in the above display if and only if $\mathcal{G}'_0(h') = 0$.

Before proving Corollary 1, we need the following lemma which is proved by Lemma 9.5 in [15].

Lemma 2. *Let $(S, \mathscr{S})$ and $(S', \mathscr{S}')$ be measurable spaces. Let $(\alpha_j) \in S^\infty$ be a stationary ergodic sequence of random variables. Let $f : S^\infty \to S'$ be a measurable function. Then $(f(\alpha_j, \alpha_{j+1}, \cdots))$ is also stationary ergodic.*

Proof of Corollary 1. Since (β_j) is i.i.d., it is stationary ergodic. As $\mathcal{G}$ is a measurable function from $[0,1)^\infty$ to M_1, we apply Lemma 2 to obtain that $(\mathcal{I}_n)$ is also stationary ergodic. The iteration equation (8) is inherited from (7). $\square$

4.3 Limits of the finite backward sequences

The reason to consider finite backward sequences is due to a simple observation: Let (P_n) be a forward sequence and (P_j^n) be the finite backward sequence with $P_n^n = P_0$, both in the random model with the same Q and (β_j). Since (β_j) is i.i.d., we have

$$(P_0, P_1, \cdots, P_n) \overset{d}{=} (P_n^n, P_{n-1}^n, \cdots, P_0^n). \tag{17}$$

So showing the convergence of (P_n) is equivalent to showing that of (P_0^n). But investigating the finite backward sequences will appear to be more convenient.

We start with the general model. Specifically, we consider (P_j^n) with $P_n^n = \delta_h$, the one studied in Lemma 1. Developing (6) we obtain

$$P_0^n(dx) = \left(\prod_{l=1}^{n} \frac{1-b_l}{\int y P_l^n(dy)}\right) x^n P_n^n(dx) + \sum_{j=0}^{n-1} \left(\prod_{l=1}^{j} \frac{1-b_l}{\int y P_l^n(dy)}\right) b_{j+1} m_j Q^j(dx) \tag{18}$$

$$= \left(\prod_{l=1}^{n} \frac{h(1-b_l)}{\int y P_l^n(dy)}\right) \delta_h(dx) + \sum_{j=0}^{n-1} \left(\prod_{l=1}^{j} \frac{1-b_l}{\int y P_l^n(dy)}\right) b_{j+1} m_j Q^j(dx). \tag{19}$$

Proposition 1. *Let (P_j^n) be the finite backward sequence in the general model with $P_n^n = \delta_h$. Then for the sequence $(\mathcal{G}_j)$, we have*

$$\mathcal{G}_0(dx) = G_0 \delta_h(dx) + \sum_{j=0}^{\infty} \prod_{l=1}^{j} \frac{(1-b_l)}{\int y \mathcal{G}_l(dy)} b_{j+1} m_j Q^j(dx), \tag{20}$$

where the second term on the right side of (19) converges to that of (20):

$$\sum_{j=0}^{n-1} \left(\prod_{l=1}^{j} \frac{1-b_l}{\int y P_l^n(dy)}\right) b_{j+1} m_j Q^j(dx) \preceq_{S_Q} \xrightarrow{d} \sum_{j=0}^{\infty} \prod_{l=1}^{j} \frac{(1-b_l)}{\int y \mathcal{G}_l(dy)} b_{j+1} m_j Q^j(dx) \tag{21}$$

and $G_0 = G_{0,h}$ is

$$\text{the limit to which } \prod_{l=1}^{n} \frac{h(1-b_l)}{\int y P_l^n(dy)} \text{ decreases in } n \tag{22}$$

$$\text{equal to } 1 - \sum_{j=0}^{\infty} \prod_{l=1}^{j} \frac{(1-b_l)}{\int y \mathcal{G}_l(dy)} b_{j+1} m_j \in [0,1] \tag{23}$$

$$\text{equal to } \mathcal{G}_0(h) \text{ if } Q(S_Q) = 0 \tag{24}$$

$$\text{the limit to which } \int \left(\frac{y}{h}\right)^n \mathcal{G}_n(dy) \prod_{l=1}^{n} \frac{h(1-b_l)}{\int y \mathcal{G}_l(dy)} \text{ decreases in } n, \text{ if } G_0 > 0. \tag{25}$$

Moreover if we define G_j for $\mathcal{G}_j$ similarly as G_0 for $\mathcal{G}_0$ (we denote $G = G_0$), we have

$$G_{j-1} = G_j \frac{h(1-b_j)}{\int y \mathcal{G}_j(dy)}, \quad \forall j \geq 1. \tag{26}$$

As a consequence G_j's are either all 0 or all strictly positive.

Proof. By (13) and the expression (19), we obtain (21), (22) and hence(20). From (20) we observe (23) and (24). To show (25), we develop (7) as follows

$$\begin{aligned}
&\mathcal{G}_0(dx) \\
&= \left(\prod_{l=1}^{n} \frac{1-b_l}{\int y \mathcal{G}_l(dy)}\right) x^n \mathcal{G}_n(dx) + \sum_{j=0}^{n-1} \left(\prod_{l=1}^{j} \frac{1-b_l}{\int y \mathcal{G}_l(dy)}\right) b_{j+1} m_j Q^j(dx) \\
&= \left(\int (\frac{y}{h})^n \mathcal{G}_n(dy) \prod_{l=1}^{n} \frac{h(1-b_l)}{\int y \mathcal{G}_l(dy)}\right) \frac{x^n \mathcal{G}_n(dx)}{\int y^n \mathcal{G}_n(dy)} + \sum_{j=0}^{n-1} \left(\prod_{l=1}^{j} \frac{1-b_l}{\int y \mathcal{G}_l(dy)}\right) b_{j+1} m_j Q^j(dx).
\end{aligned}$$

Comparing the above display and (20) we obtain (25) and also that $\frac{x^n \mathcal{G}_n(dx)}{\int y^n \mathcal{G}_n(dy)}$ converges weakly to δ_h. Finally, combining (7) and (20) we obtain (26). ∎

Remark 1. *The results and notations in the proposition can be carried over to the random model. Therefore (P_0^n) in the random model with $P_n^n = \delta_h$ converges in total variation to $\mathcal{I}$ conditional on (β_j). By (17), (P_n) in the random model with $P_0 = \delta_h$ converges weakly to $\mathcal{I}$.*

Moreover we have the following corollary.

Corollary 3. *The process $(I_j)_{j\geq 0}$ is stationary ergodic. Moreover $\mathbb{P}(\{I_j = 0, \forall j\}) = \mathbb{P}(I_0 = 0) \in \{0, 1\}$.*

Proof. By Proposition 1, $G = G(b_1, b_2, \cdots)$ is a measurable function from $[0,1)^\infty$ to $[0,1]$. For any j, we have

$$I_j = G(\beta_{j+1}, \beta_{j+2}, \cdots).$$

As (β_j) is i.i.d., we obtain that (I_j) is stationary ergodic, thanks to Lemma 2.

By (26), for any k, $\{I_k = 0\} = \{I_j = 0, \forall j\}$. Note that $\{I_j = 0, \forall j\}$ is an invariant set in the sigma algebra generated by (I_j). By ergodicity of (I_j), $\mathbb{P}(\{I_j = 0, \forall j\}) = \mathbb{P}(I_0 = 0) \in \{0, 1\}$. □

The following result provides us a tool to have some more information about $\mathcal{I}$ and Q.

Corollary 4. *The following statements about $\mathbb{E}\left[\frac{(1-\beta)}{\int y\mathcal{I}(dy)}\right]$ hold:*

1). $\mathbb{E}\left[\frac{1-\beta}{\int y\mathcal{I}(dy)}\right]$ exists, taking values in $[-\infty, -\ln\int yQ(dy)]$, and does not depend on the joint law of $(\beta, \mathcal{I})$.

2). If $Q(h) = 0$, then

$$\mathbb{E}\left[\ln \frac{h(1-\beta)}{\int y\mathcal{I}(dy)}\right] \leq 0.$$

3). If $\mathcal{I}(h) > 0$ a.s. and $Q(h) = 0$, then

$$\mathbb{E}\left[\frac{h(1-\beta)}{\int y\mathcal{I}(dy)}\right] = 0.$$

4). If $h = S_Q$ and $Q(S_Q) > 0$, then

$$\mathbb{E}\left[\ln \frac{S_Q(1-\beta)}{\int y\mathcal{I}(dy)}\right] < 0 \text{ and } I = 0, a.s..$$

Proof. 1). By (20), $\mathcal{G} = \mathcal{G}_0$ is a convex combination of probability measures $\{\delta_h, Q, Q^1, Q^2, \cdots\}$. As $Q^j \leq Q^{j+1} \leq \delta_h$ for any $j \geq 0$, we have

$$Q \leq \mathcal{I} \leq \delta_h. \tag{27}$$

Then

$$\ln \int yQ(dy) \leq \mathbb{E}\left[\ln \int y\mathcal{I}(dy)\right] \leq \ln h.$$

In consequence

$$\begin{aligned}\mathbb{E}\left[\ln \frac{1-\beta}{\int y\mathcal{I}(dy)}\right] &= \mathbb{E}\left[\ln(1-\beta) - \ln \int y\mathcal{I}(dy)\right] \\ &= \mathbb{E}\left[\ln(1-\beta)\right] - \mathbb{E}\left[\ln \int y\mathcal{I}(dy)\right] \in \left[-\infty, -\ln \int yQ(dy)\right].\end{aligned}$$

We observe that the above display does not depend on the joint law of $(\beta, \mathcal{I})$.

2). Let (P_j^n) be the finite backward sequence in the random model with $P_n^n = \delta_h$. By assumption, $Q(h) = 0$. Adapting (19) into the random model and taking the expectation of the mass on $\{h\}$ we obtain

$$\mathbb{E}[P_0^n(h)] = \mathbb{E}\left[\left(\prod_{l=1}^{n} \frac{h(1-\beta_l)}{\int y P_l^n(dy)}\right)\right] \geq \exp\left(\sum_{l=1}^{n} \mathbb{E}\left[\ln \frac{h(1-\beta_l)}{\int y P_l^n(dy)}\right]\right)$$

where the inequality is due to Janson's inequality. By (13)

$$\mathbb{E}\Big[\ln \frac{h(1-\beta_l)}{\int y P_l^n(dy)}\Big] \text{ increases in } n \text{ to } \mathbb{E}\Big[\ln \frac{h(1-\beta_l)}{\int y \mathcal{I}_l(dy)}\Big] = \mathbb{E}\Big[\ln \frac{h(1-\beta)}{\int y \mathcal{I}(dy)}\Big].$$

Combining the above two displays, it must be that $\mathbb{E}[\ln \frac{h(1-\beta)}{\int y\mathcal{I}(dy)}] \leq 0$.

3). Lemma 1 implies that, for any j, $\frac{h(1-b_j)}{\int y \mathcal{G}_j(dy)}$ is the value which is mapped to by the same measurable function from $(b_j, b_{j+1}, \cdots)$. By Lemma 2,

$$\left(\frac{h(1-\beta_j)}{\int y \mathcal{I}_j(dy)}\right) \text{ is stationary ergodic.}$$

By (25)

$$\lim_{n\to\infty} (I_0)^{1/n} = \lim_{n\to\infty} \exp\left(\frac{1}{n} \ln \int \left(\frac{y}{h}\right)^n \mathcal{I}_n(dy) + \frac{1}{n}\sum_{l=1}^{n} \ln \frac{h(1-\beta_l)}{\int y \mathcal{I}_l(dy)}\right). \tag{28}$$

Since $\mathcal{I}(h) = \mathcal{I}_0(h) = I_0 > 0$ a.s., the above two limiting terms equal 1 a.s.. As $(\mathcal{I}_j)$ is stationary ergodic, $\int \left(\frac{y}{h}\right)^n \mathcal{I}_n(dy) \in [I_n, 1]$ converges weakly to I_0, which is strictly positive almost surely. So $\frac{1}{n} \ln \int (\frac{y}{h})^n \mathcal{I}_n(dy) \in [\frac{1}{n} \ln I_n, 0]$ converges weakly to 0. Thus the above display, together with the ergodicity of $\left(\frac{h(1-\beta_j)}{\int y\mathcal{I}_j(dy)}\right)$, yield

$$1 = \exp\left(\mathbb{E}\left[\ln \frac{h(1-\beta)}{\int y \mathcal{I}(dy)}\right]\right) \quad \text{or equivalently} \quad \mathbb{E}\left[\ln \frac{h(1-\beta)}{\int y \mathcal{I}(dy)}\right] = 0.$$

4). We show $I_0 (= I) = 0$ a.s. by contradiction. Adapting (19) in the random model

$$P_0^n(dx) = \left(\prod_{l=1}^{n} \frac{S_Q(1-\beta_l)}{\int y P_l^n(dy)}\right) \delta_{S_Q}(dx) + \sum_{j=0}^{n-1} \left(\prod_{l=1}^{j} \frac{1-\beta_l}{\int y P_l^n(dy)}\right) \beta_{j+1} m_j Q^j(dx).$$

If $I_0 > 0$ a.s., we consider the mass on $\{S_Q\}$ in the above display. Note that $m_j Q^j(S_Q) = S_Q^j Q(S_Q)$. Together with (22) we obtain

$$P_0^n(S_Q) \geq Q(S_Q) \sum_{j=0}^{n-1} \left(\prod_{l=1}^{j} \frac{S_Q(1-\beta_l)}{\int y P_l^n(dy)}\right) \beta_{j+1} \geq Q(S_Q) \sum_{j=0}^{n-1} I_0 \beta_{j+1} \xrightarrow{d} \infty, \quad a.s..$$

This is a contradiction. So $I_0 = 0, a.s.$. Note that by (27), $\mathcal{I}_0(S_Q) \geq Q(S_Q) > 0$. Then we get $\mathbb{E}\left[\ln \frac{h(1-\beta)}{\int y \mathcal{I}(dy)}\right] < 0$ using (28) and the arguments thereafter.

□

4.4 Proof of Theorem 3

Proof of Theorem 3. The first assertion holds due to assertion 3) of Corollary 4. We consider the second one.

If there is a condensation on $\{h\}$, then $\mathcal{I}_{0,S_Q} \neq \mathcal{I}_{0,h}$. By assertion 3) of Proposition 1 and Corollary 2,

$$\mathbb{E}\left[\ln \frac{h(1-\beta)}{\int y\mathcal{I}_{0,S_Q}(dy)}\right] > \mathbb{E}\left[\ln \frac{h(1-\beta)}{\int y\mathcal{I}_{0,h}(dy)}\right] = 0.$$

If there is no condensation on $\{h\}$, again by Corollary 2, $\mathcal{I}_{0,h} = \mathcal{I}_{0,S_Q}$. By assertion 2) of Corollary 4,

$$\mathbb{E}\left[\ln \frac{h(1-\beta)}{\int y\mathcal{I}_{0,S_Q}(dy)}\right] = \mathbb{E}\left[\ln \frac{h(1-\beta)}{\int y\mathcal{I}_{0,h}(dy)}\right] \leq 0.$$

□

4.5 Some properties of invariant measures

In this section, we prove some results concerning invariant measures. But we leave the proof of Theorem 4 to the end. Those measures will play important roles in the proof of Theorem 2.

Lemma 3. *For any invariant measure ν, $\mathbb{E}\left[\frac{1-\beta}{\int y\nu(dy)}\right]$ exists, taking values in $[-\infty, -\ln\int yQ(dy)]$, and does not depend on the joint law of (β,ν).*

Proof. By the definition of invariant measure

$$\begin{aligned}\mathbb{E}\left[\int y\nu(dy)\right] &= (1-\mathbb{E}[\beta])\mathbb{E}\left[\frac{\int y^2\nu(dy)}{\int y\nu(dy)}\right] + \mathbb{E}[\beta]\mathbb{E}\left[\int yQ(dy)\right] \\ &\geq (1-\mathbb{E}[\beta])\mathbb{E}\left[\int y\nu(dy)\right] + \mathbb{E}[\beta]\mathbb{E}\left[\int yQ(dy)\right]\end{aligned}$$

where the inequality is due to the fact that $\int y^2\nu(dy) \geq (\int y\nu(dy))^2$. Then we obtain

$$\int yQ(dy) \leq \mathbb{E}\left[\int y\nu(dy)\right] \leq 1.$$

Proceeding similarly as in the proof of assertion 1) in Corollary 4, we conclude that this lemma holds. □

Corollary 5. *The invariant measure supported on $[0, S_Q]$ is unique which is $\mathcal{I}_Q$.*

Proof. Let (P_n) and (P'_n) be two forward sequences as in the Appendix with

$$Q = Q'; \quad h = h' = S_Q; \quad P_0 \stackrel{d}{=} \nu, \quad P'_0 = \delta_h$$

and P_0 is independent of (β_n). Using the notations in the Appendix, and by (17) and the monotonicity analysis as in the proof of Lemma 1

$$\int \mathcal{M}_n(dx) \leq \int \mathcal{M}'_n(dx), \quad \mathcal{F}'_n \preceq_{S_Q} \mathcal{F}_n. \tag{29}$$

If $I_Q = 0$ a.s., by Proposition 1 and Remark 1,

$$\mathcal{F}'_n \xrightarrow{d} \mathcal{I}_Q, \quad \int \mathcal{M}'_n(dx) \xrightarrow{d} 0.$$

Due to (29) and (37), (38),

$$\nu \stackrel{d}{=} P_n \xrightarrow{d} \mathcal{I}_Q$$

implying $\nu \stackrel{d}{=} \mathcal{I}_Q$.

If $I_Q > 0, a.s.$, then by assertion 4) of Corollary 4, $Q(S_Q) = 0$. Again by monotonicity analysis, contionally on (β_j)

$$P'_n \preceq_{S_Q-} P_n, \quad P_n(S_Q) \le P'_n(S_Q). \tag{30}$$

The above entails, using Remark 1

$$\mathcal{I}_0 \preceq_{S_Q-} \nu, \quad \nu(S_Q) \le_d I_0 = \mathcal{I}_0(S_Q).$$

By assertion 3) of Corollary 4, we have $\mathbb{E}\left[\frac{S_Q(1-\beta)}{\int y\mathcal{I}_Q(dy)}\right] = 0$. Assume that ν is not equal to $\mathcal{I}_Q$ in distribution, then by the above display, we obtain

$$\mathbb{E}\left[\frac{S_Q(1-\beta)}{\int y\nu(dy)}\right] > 0.$$

The inequality implies that if $\varepsilon > 0$ is small enough, we also have

$$\mathbb{E}\left[\frac{(S_Q-\varepsilon)(1-\beta)}{\int y\nu(dy)}\right] > 0.$$

By definition of invariant measures, it is straightforward to see that $S_\nu = S_Q$ almost surely. Consequently $\int (\frac{x}{S_Q-\varepsilon})^n P_0(dx) > 0, a.s.$, for any n. Using again (37), we get

$$\begin{aligned}
1 &= \mathbb{E}\left[\int P_0(dx)\right] \ge \mathbb{E}\left[\int \mathcal{M}_n(dx)\right] \\
&= \mathbb{E}\left[\exp\left(\sum_{l=0}^{n-1} \mathbb{E}\Big[\ln \frac{(S_Q-\varepsilon)(1-\beta_{l+1})}{\int yP_l^n(dy)}\Big] + \int \left(\frac{x}{S_Q-\varepsilon}\right)^n P_0(dx)\right)\right] \\
&\ge \mathbb{E}\left[\exp\left(\sum_{l=0}^{n-1} \ln \frac{(S_Q-\varepsilon)(1-\beta_{l+1})}{\int yP_l^n(dy)}\right)\right] \ge \exp\left(n\mathbb{E}\left[\frac{(S_Q-\varepsilon)(1-\beta)}{\int y\nu(dy)}\right]\right) \overset{n\to\infty}{\longrightarrow} \infty
\end{aligned}$$

where the third inequality is due to Janson's inequality and Lemma 4.5. So this is a contradiction, which means that ν is equal in distribution to $\mathcal{I}_Q$. □

4.6 Proof of Theorem 2

The proof is given in 4 subcases.

Case 1. $P_0 = \delta_h$.

Proof of Theorem 2. This is shown in Remark 1.

□

Case 2. $I_{0,h} = 0$ **a.s..**

Proof of Theorem 2. Let $(P_n)_{n\ge0}$, $(P'_n)_{n\ge0}$ be two forward sequences in the Appendix with

$$Q = Q', \quad h = h', \quad P'_0 = \delta_h.$$

Next it suffices to follow the same procedure as in the proof of Corollary 5 in the case $I_Q = 0$ a.s.. The proof is omitted. □

Case 3. $I_{0,h} > 0$ **a.s. and** $P_0(h) > 0$.

First of all, we recall a result from ([21], p.10), where only $h = 1$ is considered. But it is easily generalised to any h.

Lemma 4. *Let two measures $u_1, u_2 \in M_1$ such that $S_{u_1} = S_{u_2} = h$ and $u_1 \preceq_{h-} u_2$. If for some $\varepsilon > 0$ there exists $a \in (0, h)$, such that $D_{u_1}(a) + \varepsilon \le D_{u_2}(a)$, then*

$$\int y u_1(dy) \ge c(a, \varepsilon) \int y u_2(dy)$$

where $c(a, \varepsilon) = \frac{1}{1-\varepsilon(h-a)} > 1$.

Proof of Theorem 2. Let $(P_n), (P'_n)$ be the two forward sequences in the proof of Case 2 of this theorem. Similarly as (30), conditionally on (β_j) we have

$$P'_n \preceq_{h-} P_n, \quad P_n(h) \le P'_n(h) \tag{31}$$

implying

$$\int y P'_j(dy) \ge \int y P_j(dy), \quad \forall j \ge 0.$$

For any $\varepsilon > 0, a \in (0, h)$, let

$$\kappa_n := \#\{n : D_{P'_j}(a) + \varepsilon \le D_{P_j}(a), 0 \le j \le n\}.$$

Note that by assertion 4) of Proposition 1, $Q(h) = 0$. Using (37) and (38)

$$P'_n(h) = \prod_{l=0}^{n-1} \frac{h(1-\beta_{l+1})}{\int y P'_l(dy)}, \quad P_n(h) = \left(\prod_{l=0}^{n-1} \frac{h(1-\beta_{l+1})}{\int y P_l(dy)} \right) P_0(h).$$

Then by Lemma 4,

$$P'_n(h) \le \frac{1}{c(a, \varepsilon)^{\kappa_n} P_0(h)}.$$

But (22) of Proposition 1 and (17) entail that $P'_n(h)$ converges weakly to $I_{0,h}$ which is by assumption non-zero almost surely. Then $\lim_{n\to\infty} \kappa_n < \infty$ a.s.. As ε can be any small positive value and by Case 1 of this theorem $P'_n \xrightarrow{d} \mathcal{I}_{0,h}$, we use (31) to conclude that P_n also converges weakly to $\mathcal{I}_{0,h}$. □

Case 4. $I_{0,h} > 0$ a.s. and $P_0(h) = 0$.

Proof of Theorem 2. The idea is to use a tripling argument similarly as in the proof of Theorem 5 in [21]. For any $u \in M_1$ and any $a \in [0, 1]$, define

$$u^a = u_{[0,a)} + u([a, 1])\delta_a, \quad a < h$$

where $u_{[0,a)}$ is the restriction of u on $[0, a)$. Let $(P_n), (P'_n), (P''_n)$ be three forward sequences as defined in the Appendix with

$$h'' < h = h'; \quad Q' = Q, \quad Q'' = Q^{h''}; \quad P'_0 = \delta_h, \quad P''_0 = P_0^{h''}.$$

By monotonicity analysis, conditionally on (β_j)

$$P'_n \preceq_{h-} P_n \preceq_{h''-} P''_n. \tag{32}$$

Consider first the case $S_Q < h'' < h = h'$. Note that $P''_0(h'') > 0$. Now using Case 1 and Case 3 of this theorem,

$$P'_n \xrightarrow{d} \mathcal{I}_{0,h}, \quad P''_n \xrightarrow{d} \mathcal{I}_{0,h''}. \tag{33}$$

Next we show that $\mathcal{I}_{0,h''}$ converges weakly to $\mathcal{I}_{0,h}$ as h'' approaches h. Since $I_{0,h} > 0$, *a.s.* and $h > S_Q$, by Theorem 3,

$$\mathbb{E}\left[\frac{h(1-\beta)}{\int y \mathcal{I}_{0,S_Q}(dy)} \right] > 0.$$

Then there exists a small number ε with $S_Q < h - \varepsilon$ such that

$$\mathbb{E}\left[\frac{(h-\varepsilon)(1-\beta)}{\int y\mathcal{I}_{0,S_Q}(dy)}\right] > 0.$$

We take $h'' \in [h-\varepsilon, h)$. Using the above display, we have

$$\mathbb{E}\left[\frac{h''(1-\beta)}{\int y\mathcal{I}_{0,S_Q}(dy)}\right] > 0.$$

Then by Theorem 3

$$I_{0,h''} > 0, \quad a.s.$$

which entails, due to assertion 3) of Corollary 4

$$\mathbb{E}\left[\frac{h''(1-\beta)}{\int y\mathcal{I}_{0,h''}(dy)}\right] = 0.$$

Letting $h'' \to h$ and using Corollary 2, $\mathcal{I}_{0,h''}$ converges weakly to a limit, denoted by ν. Since $\mathcal{I}_{0,h''}$ is an invariant measure, ν is still an invariant measure. Using the above display and Corollary 2,

$$\mathbb{E}\left[\frac{h(1-\beta)}{\int y\nu(dy)}\right] = 0, \quad \nu(h) > 0 \ \ a.s.. \tag{34}$$

Using Corollary 2 again, we know

$$\mathcal{I}_{0,h} \preceq_{h''-} \mathcal{I}_{0,h''}, \quad \mathcal{I}_{0,h''}(h'') \leq_d \mathcal{I}_{0,h}(h)$$

implying

$$\mathcal{I}_{0,h} \preceq_{h-} \nu, \quad \nu(h) \leq_d I_{0,h}. \tag{35}$$

The above display yields $I_{0,h} > 0, a.s.$. By assertion 3) of Corollary 4,

$$\mathbb{E}\left[\frac{h(1-\beta)}{\int y\mathcal{I}_{0,h}(dy)}\right] = 0.$$

Together with (34) and (35) we obtain

$$\nu \overset{d}{=} \mathcal{I}_{0,h} \tag{36}$$

Let f be any continuous function on $[0,1]$. By (32), we have

$$\mathbb{E}[f(x)P''_n(dx)] \leq \mathbb{E}[f(x)P_n(dx)] \leq \mathbb{E}[f(x)P'_n(dx)].$$

Moreover by (36) and (33),

$$\mathbb{E}[f(x)P''_n(dx)] \overset{n\to\infty}{\longrightarrow} \mathbb{E}[f(x)\mathcal{I}_{0,h''}(dx)] \overset{h''\to h}{\longrightarrow} \mathbb{E}[f(x)\mathcal{I}_{0,h}(dx)] = \lim_{n\to\infty} \mathbb{E}[f(x)P'_n(dx)].$$

We conclude that (P_n) converges to $\mathcal{I}_{0,h}$.

If $h = S_Q$, we follow the same procedure, except that to prove (36), we require Corollary 5. $\square$

4.7 Proof of Theorem 4

Firstly we prove two lemmas. Recall the definition of S_u for $u \in M_1$.

Lemma 5. *$S_{(\cdot)}$ is a continuous (hence measurable) function on (M_1, T).*

Proof. Assume that a sequence (u_n) converges weakly to u. If S_{u_n} does not converge to S_u, then there exists a subsequence (u_{n_k}) such that $S_{u_{n_k}}$ converges to a limit a with $a < S_u$ or $a > S_u$. Without loss of generality, assume $a < S_u$. We take any positive and continuous function f supported on $(\frac{a+S_u}{2}, S_u]$ and then $\int f(x)u(dx) > 0$. But $\int f(x)u_{n_k}(dx)$ converges to 0. This is against the weak convergence. We conclude that $S_{(\cdot)}$ is a continuous function on (M_1, T). $\square$

The next lemma generalises Corollary 5.

Lemma 6. *For any invariant measure ν with $S_\nu = h$ a.s., we have $\nu \stackrel{d}{=} \mathcal{I}$.*

Proof. Let (P_n) be the forward sequence in the random model with $P_0 \stackrel{d}{=} \nu$ and P_0 independent of (β_n). By Theorem 2, conditionally on P_0, P_n converges in distribution to the same random measure $\mathcal{I}$. Then unconditionally $P_n \stackrel{d}{=} \nu$ converges in distribution to $\mathcal{I}$, implying $\nu \stackrel{d}{=} \mathcal{I}$. $\square$

Proof of Theorem 4. Let ν be an invariant measure. By definition (11), $S_\nu \in [S_Q, 1]$. By Lemma 5, S_ν is a random variable. Applying Theorem 5.3 in [15], there exists a probability kernel $U(\cdot,\cdot)$ from $[0,1]$ to M_1 such that

$$(\nu(dx)|S_\nu) = U(S_\nu, dx), \quad a.s..$$

Applying it to the right side of (11) we get

$$\left((1-\beta)\frac{x\nu(dx)}{\int y\nu(dy)} + \beta Q(dx)\Big|S_\nu\right) = (1-\beta)\frac{xU(S_\nu, dx)}{\int yU(S_\nu, dy)} + \beta Q(dx), \quad a.s..$$

The above two displays show that, in order that ν be an invariant measure, conditionally on S_ν, $U(S_\nu,\cdot)$ must be an invariant measure almost surely. Conditionally on S_ν, By Lemma 6, $U(S_\nu,\cdot) \stackrel{d}{=} \mathcal{I}_{0,S_\nu}$. Then unconditionally $\nu \stackrel{d}{=} U(S_\nu,\cdot) \stackrel{d}{=} \mathcal{I}_{0,S_\nu}$.

$\square$

5 Appendix

Let $(P_n), (P'_n), (P''_n)$ be three forward sequences corresponding respectively to

$$((\beta_n), Q, P_0, h), \quad ((\beta_n), Q', P'_0, h'), \quad ((\beta_n), Q'', P''_0, h'').$$

Using (2), we write

$$P_n(dx) = \mathcal{M}_n(dx) + \mathcal{F}_n(dx) \tag{37}$$

with

$$\mathcal{M}_n(dx) = \Big(\prod_{l=0}^{n-1} \frac{1-\beta_{l+1}}{\int yP_l(dy)}\Big)x^n P_0(dx)$$

and

$$\mathcal{F}_n(dx) = \sum_{j=1}^{n} \Big(\prod_{l=j}^{n-1} \frac{1-\beta_{l+1}}{\int yP_l(dy)}\Big)b_j m_{n-j} Q^{n-j}(dx).$$

Similarly we introduce

$$P'_n(dx) = \mathcal{M}'_n(dx) + \mathcal{F}'_n(dx) \tag{38}$$

$$P''_n(dx) = \mathcal{M}''_n(dx) + \mathcal{F}''_n(dx) \tag{39}$$

with $\mathcal{M}'_n, \mathcal{F}'_n, \mathcal{M}''_n, \mathcal{F}''_n$ defined correspondingly.

6 Acknowledgement

The author thanks Takis Konstantopoulos, Götz Kersting and Pascal Grange for discussions. The author acknowledges the support of the National Natural Science Foundation of China (Youth Program, Grant: 11801458), and the XJTLU RDF-17-01-39.

References

[1] Aldous, D.J., Bandyopadhyay, A. A survey of max-type recursive distributional equations. *Ann. Probab.* 15(2), 1047-1110, 2005.

[2] Betz, V., Dereich, S., Mörters, P. The shape of the emerging condensate in effective models of condensation. *Ann. Inst. Henri Poincaré.* 19(6), 1869-1889, 2018.

[3] Bürger, R. On the maintenance of genetic variation: global analysis of Kimuras continuum-of-alleles model. *J. Math. Biol.* 24, 341-351, 1986.

[4] Bürger, R. Mutation-selection balance and continuum-of-alleles models. *Math. Biosci.* 12(9), 67-83, 1989.

[5] Bürger, R. Mathematical properties of mutation-selection models. *Genetica.* 102, 279-298, 1998.

[6] Bürger, R. *The Mathematical Theory of Selection, Recombination, and Mutation.* Wiley, Chichester, NewYork. 2000.

[7] Crow, J.F., Kimura, M. *An Introduction to Population Genetics Theory.* Harper & Row, New York. 1970.

[8] Dereich, S., Mörters, P. Emergence of condensation in Kingmans model of selection and mutation. *Acta Appl. Math.* 127, 17-26, 2013.

[9] Diaconis, P., Freedman, D. Iterated random functions. *SIAM Rev.* 41(1), 45-76, 1999.

[10] Evans, S.N., Steinsaltz, D., Wachter, K.W. *A Mutation-Selection Model with Recombination for General Genotypes.* American Mathematical Soc. 2013.

[11] Ewens, W.J. *Mathematical Population Genetics.* Springer-Verlag, New York, Heidelberg, Berlin. 1979.

[12] Gonzalez-Casanova, A., Kurt, N., Wakolbinger, A., Yuan, L. An individual-based model for the Lenski experiment, and the deceleration of the relative fitness. *Stochastic Process. Appl.* 126(8), 2211-2252, 2016.

[13] Grange, P. Steady states in a non-conserving zero-range process with extensive rates as a model for the balance of selection and mutation. *J. Phys. A.* 52(36), 365601, 2019.

[14] Haldane, J.B.S. The effect of variation on fitness. *Amer. Naturalist.* 71, 337-349, 1937.

[15] Kallenberg, O. *Foundations of Modern Probability*, Springer Science & Business Media., 1997.

[16] Kimura, M. A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proc. Natl. Acad. Sci.* 54(3), 731-736, 1965.

[17] Kingman, J.F.C. On the properties of bilinear models for the balance between mutation and selection. *Proc. Camb. Phil. Soc.* 80, 443-453, 1977.

[18] Kingman, J.F.C. A simple model for the balance between selection and mutation. *J. Appl. Prob.* 15, 1-12, 1978.

[19] Kingman, J.F.C. *Mathematics of Genetic Diversity.* SIAM. 1980.

[20] Steinsaltz, D., Evans, S.N., Wachter, K.W. A generalized model of mutation-selection balance with applications to aging. *Adv. Appl. Math.* 35(1), 16-33, 2005.

[21] Yuan, L. A generalization of Kingmans model of selection and mutation and the Lenski experiment. *Math. Biosci.* 285, 61-67, 2017.