INRIA

The convergence to equilibrium of neutral genetic models

Pierre Del Moral — Laurent Miclo — Frédéric Patras — Sylvain Rubenthaler

N° 6442

Rapport de recherche

The convergence to equilibrium of neutral genetic models

Pierre Del Moral[*], Laurent Miclo [†] , Frédéric Patras[‡] , Sylvain Rubenthaler [§]

Abstract: This article is concerned with the long time behavior of neutral genetic population models, with fixed population size. We design an explicit, finite, exact, genealogical tree based representation of stationary populations that holds both for finite and infinite types (or alleles) models. We then analyze the decays to the equilibrium of finite populations in terms of the convergence to stationarity of their first common ancestor. We estimate the Lyapunov exponent of the distribution flows with respect to the total variation norm. We give bounds on these exponents only depending on the stability with respect to mutation of a single individual; they are inversely proportional to the population size parameter.

Key-words: Wright-Fisher model, neutral genetic models, coalescent trees, Lyapunov exponent, stationary distribution.

[*] Centre INRIA Bordeaux et Sud-Ouest & Institut de Mathématiques de Bordeaux , Université de Bordeaux I, 351 cours de la Libération 33405 Talence cedex, France, Pierre.Del-Moral@inria.fr.

[†] Laboratoire d'Analyse, Topologie, Probabilités UMR 6632, Université de Provence, Technopôle Château-Gombert, 39, rue F. Joliot Curie, 13453 Marseille Cedex 13

[‡] CNRS UMR 6621, Université de Nice, Laboratoire de Mathématiques J. Dieudonné, Parc Valrose, 06108 Nice Cedex 2, France.

[§] CNRS UMR 6621, Université de Nice, Laboratoire de Mathématiques J. Dieudonné, Parc Valrose, 06108 Nice Cedex 2, France.

Convergence à l'équilibre d'algorithmes génétiques neutres

Résumé : Cet article analyse le comportement en temps long de dynamiques de populations génétiques neutres et de taille fixée. Nous confectionnons une réprésentation exacte et explicite de la mesure stationnaire des populations en termes d'arbres généalogiques. Nous analysons la vitesse de convergence à l'équilibre des populations finies en fonction de la vitesse à l'équilibre de l'unique ancêtre commun. Nous proposons une estimation des exposants de Lyapunov des flots de distributions des populations, par rapport à la distance en variation totale. Les bornes obtenues sont reliées aux propriétés de stabilité des mutations d'un seul individu, et elles sont inversement proportionnelles à la taille des populations.

Mots-clés : Modèle de Wright-Fisher, modèle génétique neutre, arbres de coalescence, exposants de Lyapunov, distributions stationnaires.

1 Introduction

Stochastic models for population dynamics provide a mathematical framework for the analysis of genetic variations in biological populations that evolve under the influence of evolutionary type forces such as selections and mutations. The common models, such as the (generalized) Wright-Fisher models, allow for many variations in the fundamental assumptions. For example, one may consider a finite population, or a sample out of an infinite population; there might be a finite or infinite number of types (or alleles) of individuals; the genetics involved may refer to monoecious (i.e. where there is only one sex) or dioecious phenomena, and so on.

Dealing with finite number of individuals without sampling in an infinite population is usually difficult: although the problems are easily formulated, computations become soon very involved. Here, we are interested in finite population models where the evolution is driven by a selection/mutation process and derive explicit formulas for the stationary state of the population or the decay to the equilibrium. We do not put any restriction on the number of types or alleles, that may be finite or infinite. The main restriction to a full generality of the model is neutrality of the selection process: that is, we assume that all individuals have the same reproduction rate (we refer e.g. to the monograph of M. Kimura [7] for a detailed account on the neutral theory of molecular evolution).

At the molecular level, the evolutionary models we consider correspond therefore to situations where the genetic drifts that govern the dynamics are the mutations combined with a uniform reproduction rate. Genealogical tree evolution models arise then in a natural way, when considering the complete past history of the individuals. These path space models can be described forwards with respect to the time parameter. Conversely, we can also trace back in time the complete ancestral line of all the individuals. This backward view of the ancestral structures is then interpreted as a stochastic coalescent process.

The important questions that arise then are the precise description of the asymptotic behavior of evolution processes, and the corresponding time to equilibrium analysis. They are at the core of the modern development of mathematical biology. Let us also point out that, apart from their importance in biology, they are also related to the convergence of a class of genealogical tree algorithms used in advanced stochastic engineering, and in Bayesian statistics. For example, a full understanding of the long time behavior of genetic type branching models is essential in the designing of genealogical particle filters and smoothers, as well as for the tuning of genetic algorithms for solving global optimization problems. These features of evolutionary models are another strong motivation for the present work and, although we emphasize mainly the applications to genetics, the interested reader should keep in mind these other application areas (for further informations and references on the subject, see for instance the pair of recent books [2, 4]).

During the last three decades, many efforts have been made to tackle these questions. Although several natural Markov chain models of genetic processes have been developed, their combinatorial complexity makes both the intuitive, and the rigorous understanding of the long time behavior of evolution mechanisms difficult. Several reduced models have been developed to obtain rigorous, but partial theoretical results. These simplified models are usually based on large population approximations, and appropriate rescaling of the time parameter in units of the total population size.

Let us review very briefly three main directions of research in the field. The first one studies the diffusion limit model arising, through appropriate time rescaling techniques, when considering gene type fractions decompositions on large population sizes models (see for instance the seminal book of S. N. Ethier and T. Kurtz [9], and the references therein). A second idea is to consider the backward ancestral lineage in infinitely many allele models. This genealogical process is expressed in terms of rescaled Poisson coalescence epochs, and superimposed Poisson mutation events. Running back in time this population model, the coalescent of Kingman describes the ancestral genealogy of the individuals in terms of a simplified binary ancestral lineage tree. The Ewens' sampling formula describes the limiting distribution of the type spectrum in finite samples of the infinitely many allele models (see for instance the Saint Flour's lecture notes of S. Tavaré [12], and references therein). A third, more recent, idea is essentially based on the mean field interpretation of genetic models. In this context, the occupation measures of simple genetic populations converge, as the size of the systems tends to infinity, to a non linear Feynman-Kac semigroup in the distribution space. For a rather detailed account of these mean field limits, we refer the reader to the first author monograph [2]. In this interpretation, and in the case of reversible mutations, the long time behavior of an infinite population model corresponds to the ground state of Schrödinger type operators.

In the present article we depart from these techniques and tackle directly the combinatorial complexity of finite population dynamics. This leads to a fine asymptotic analysis of a general class

of neutral genetic models on arbitrary state spaces, with fixed population size. In contrast to the existing literature on the asymptotic behavior of neutral genetic models, our approach is not based on some had-oc rescaling of the time parameter, and it applies to evolutionary models with fixed population size.

Firstly, we provide an explicit functional representation of their invariant measures in terms of planar genealogical trees. We then use this representation to analyze the decays to stationary populations in terms of the convergence to the equilibrium of the first common ancestor and show, for example, that the Lyapunov exponent of the genetic distribution flow is inversely proportional to the population size of the model. This estimate is of course sharper than the one we would obtain using crude minorization techniques of the transitions probabilities of the genetic model, such as those presented in [5]. We refer to Section 2.3, Theorems 2.1 and 2.2 for a precise statement of the main results of the article.

Unfortunately, these rather natural genealogical techniques are restricted to neutral genetic population models, but it leads to conjecture that the same type of decays holds true for more general models. Besides, some of the ideas and techniques developed in the article can certainly be adapted to cover inhomogeneous selection rates -this is a point we leave deliberately out of the scope of the present work.

We finally mention that the genealogical invariant measure derived in the present article belongs to the same class of tree based measures as the ones studied in [1] to derive coalescent tree based functional representations of genetic type particle models. The symmetry properties of the invariant measure can be combined with the combinatorial analysis presented in [1] to simplify notably the formula for the distributions of stationary populations. We briefly present these constructions in the last section.

To the best of our knowledge, these genealogical tree based representations of stationary populations, as well as of the corresponding convergence decays to the equilibrium, are the first of this kind, for this class of evolutionary models.

2   Neutral Genetic Models

2.1   Conventions

Let us introduce some notation. We denote respectively by $\mathcal{M}(E)$, $\mathcal{P}(E)$, and $\mathcal{B}(E)$, the set of all finite signed measures on some measurable space $(E, \mathcal{E})$, the convex subset of all probability measures, and the Banach space of all bounded and measurable functions f on E, equipped with the uniform norm $\|f\| = \sup_{x \in E} |f(x)|$. The space E will stand for the set of types or alleles of the genetic model.

We let $\mu(f) = \int \mu(dx) \, f(x)$, be the Lebesgue integral of a function $f \in \mathcal{B}(E)$, with respect to a measure $\mu \in \mathcal{M}(E)$, and we equip $\mathcal{M}(E)$ with the total variation norm $\|\mu\|_{\mathrm{tv}} = \sup_{f \in \mathcal{B}(E) : \|f\| \leq 1} |\mu(f)|$.

We recall that a Markov transition M from E into itself, is an integral probability operator such that the functions

$$M(f) \, : \, x \in E \mapsto M(f)(x) = \int_E M(x, dy) \, f(y) \, \in \mathbb{R}$$

are $\mathcal{E}$-measurable and bounded, for any $f \in \mathcal{B}(E)$. It generates a dual operator $\mu \mapsto \mu M$ from $\mathcal{M}(E)$ into itself defined by $(\mu M)(f) := \mu(M(f))$. For a pair of Markov transitions M_1, and M_2, we denote by $M_1 M_2$ the composition integral operator from E into itself, defined for any $f \in \mathcal{B}_b(E)$ by $(M_1 M_2)(f) := M_1(M_2(f))$. The tensor power $M^{\otimes N}$ represents the bounded integral operator on E^N, defined for any $F \in \mathcal{B}(E^N)$ by

$$M^{\otimes N}(F)(x^1, \ldots, x^N) = \int_{E^N} \left[M(x^1, dy^1) \, \ldots \, M(x^N, dy^N) \right] \, F(y^1, \ldots, y^N)$$

We let N be a fixed integer parameter, and we set $[N] = \{1, \ldots, N\}$. The integer N will stand for the number of individuals in the population. We denote by $m(x) = \frac{1}{N} \sum_{i=1}^{N} \delta_{x^i}$ the empirical measure associated with an N-uple $x = (x^i)_{1 \leq i \leq N} \in E^N$. The notation δ_x stands in general for the Dirac measure at the point x.

We let $\mathcal{A} = [N]^{[N]}$, be the set of mappings from $[N]$ into itself. We associate with a mapping $a \in \mathcal{A}$, the Markov transition D_a on E^N defined by $D_a(x, dy) = \delta_{x^a}(dy)$, with $x^a = (x^{a(i)})_{i \in [N]}$. The transition $\mathcal{D}_a$ can be interpreted as a coalescent, or a selection type transition on the set E^N.

In this interpretation, the population of individuals $x^a = (x^{a(i)})_{i \in [N]}$ results from a selection in $x = (x^i)_{i \in [N]}$ of the individuals with labels $(a(i))_{i \in [N]}$. Also notice that the N-tensor product of an empirical measure can be expressed in terms of these selection type transitions :

$$m(x)^{\otimes N}(dy) = \frac{1}{|\mathcal{A}|} \sum_{a \in \mathcal{A}} D_a(x, dy) =: \mathcal{D}(x, dy)$$

We consider now a Markov transition M, and a probability measure η on E. The Markov transition M will represent the mutation transition process, whereas η will stand for the initial distribution of the population. An N-neutral genetic model, with these parameters, can be represented by a pair of E^N-valued Markov chains $(\xi_n)_{n \in \mathbb{N}} = ((\xi_n^i)_{i \in [N]})_{n \in \mathbb{N}}$ and $(\widehat{\xi}_n)_{n \in \mathbb{N}} = ((\widehat{\xi}_n^i)_{i \in [N]})_{n \in \mathbb{N}}$, together with transitions

$$\xi_n \xrightarrow{selection/reproduction} \widehat{\xi}_n \xrightarrow{mutation} \xi_{n+1} \tag{2.1}$$

The initial configuration ξ_0 consists of N independent and identically distributed random variables with common distribution η.

The selection/reproduction transition $\xi_n \rightsquigarrow \widehat{\xi}_n$ corresponds to a simple Wright-Fisher selection model: it consists in sampling N conditionally independent random variables $(\widehat{\xi}_n^i)_{i \in [N]}$, with common distribution $m(\xi_n)$. Equivalently, the selected population $\widehat{\xi}_n$ is a random variable with distribution $\mathcal{D}(\xi_n, dx)$; in other words, we chose randomly a mapping A_n in $\mathcal{A}$, and we set $\widehat{\xi}_n = \xi_n^{A_n} \left(= (\xi_n^{A_n(i)})_{i \in [N]} \right)$. As the number of offsprings of each individuals have all the same law, the model is called "neutral genetic" and from now on, we will use the word "reproduction" instead of "selection" or "selection/reproduction". In the literature on branching processes, the above random reproduction mapping A_n is often expressed in terms of a multinomial branching rules. Assume, for instance that all the ξ_n^i are different, if we let $L_n = (L_n^i)_{i \in [N]}$ be the number of offsprings of the individuals ξ_n

$$\forall i \in [N] \qquad L_n^i := \left| \left\{ j \in [N] \ : \ \widehat{\xi}_n^j = \xi_n^i \right\} \right|$$

then, L_n has a symmetric multinomial distribution

$$\mathbb{P} \left(L_n^1 = l^1, \dots, L_n^N = l^N \mid \xi_n \right) = \frac{N!}{l^1! \dots l^N!} \ \frac{1}{N^N}$$

for any $l = (l^i)_{i \in [N]} \in \mathbb{N}^N$, with $\sum_{i \in [N]} l^i = N$.

During the mutation, each of the selected particles $\widehat{\xi}_n^i$ evolves to a new location ξ_{n+1}^i, randomly chosen with distribution $\delta_{\widehat{\xi}_n^i} M$, with $i \in [N]$. Equivalently, the population ξ_{n+1} after mutation is a random variable with distribution $\mathcal{M}(\widehat{\xi}_n, dx)$, with $\mathcal{M} := M^{\otimes N}$.

The distribution laws of the populations before and after the reproduction step are defined by the values, for any function $F \in \mathcal{B}_b(E^N)$, of

$$\Gamma_{\eta,n}(F) := \mathbb{E}\left(F(\xi_n)\right) \quad \text{and} \quad \widehat{\Gamma}_{\eta,n}(F) := \mathbb{E}\left(F(\widehat{\xi}_n)\right)$$

Notice that $\Gamma_{\eta,0} = \eta^{\otimes N}$, and $\widehat{\Gamma}_{\eta,0} = \Gamma_{\eta,0}\mathcal{D}$. By construction, we also have that $\widehat{\Gamma}_{\eta,n} = \Gamma_{\eta,n}\mathcal{D}$, and $\Gamma_{\eta,n+1} = \widehat{\Gamma}_{\eta,n}\mathcal{M}$. This clearly gives the dynamical structure of the pair of distributions

$$\Gamma_{\eta,n+1} = \Gamma_{\eta,n}\left(\mathcal{D}\mathcal{M}\right) \quad \text{and} \quad \widehat{\Gamma}_{\eta,n+1} = \widehat{\Gamma}_{\eta,n}\left(\mathcal{M}\mathcal{D}\right)$$

In particular, we have:

$$\Gamma_{\eta,n}(F) = \Gamma_{\eta,0}\left(\mathcal{D}\mathcal{M}\right)^n(F) = \mathbb{E}\left(\Gamma_{\eta,0}D_{A_0}\mathcal{M}D_{A_1}\dots D_{A_{n-1}}\mathcal{M}(F)\right)$$

and

$$\widehat{\Gamma}_{\eta,n}(F) = \widehat{\Gamma}_{\eta,0}\left(\mathcal{M}\mathcal{D}\right)^n(F) = \mathbb{E}\left(\Gamma_{\eta,0}D_{A_0}\mathcal{M}D_{A_1}\dots\mathcal{M}D_{A_n}(F)\right)$$

with a sequence of independent random variables $(A_p)_{0 \leq p \leq n}$, uniformly chosen in the set $\mathcal{A}$.

2.2   Genealogic trees representations

In neutral reproduction models, the pair of mutation-reproduction processes can be separated in order to describe the mutation scenarios and the reproductions on two different levels. Traditionally, these neutral genetic models are sampled in the following way. The genealogical structure of the individuals is first modelized. Then, the genetic mutations are superimposed on the sampled genealogy.

In the present section, we design a representation of this pair of genetic processes in terms of random trees. The reproductions are encoded by random mappings; their composition gives rise to random coalescent trees. The representation of the process is complete when mutation scenarios are taken into account. In contrast to traditional mutation-reproduction decoupling models, non necessarily neutral selection processes could also be described forward in terms of these random trees. In this situation, the random reproduction type mappings would depend on the configuration of the genes.

We let $X_n^{\mathbf{i}_n}$, with $\mathbf{i}_n = (i_0, \ldots, i_n) \in [N]^{n+1}$, and $n \geq 0$, be the collection of E-valued random variables defined inductively as follows: at rank $n = 0$, $X_0^{\mathbf{i}_0}$, with $\mathbf{i}_0 = i_0 \in [N]$, represents a collection of N independent, and identically distributed random variables with common distribution η. Given the random variable $X_{n-1}^{\mathbf{i}_{n-1}}$, for some multi index $\mathbf{i}_{n-1} \in [N]^n$, the sequence of random variables $X_n^{\mathbf{i}_n}$, with $\mathbf{i}_n = (\mathbf{i}_{n-1}, i_n)$, and $i_n \in [N]$ consists in N conditionally independent random variables, with common distribution $\delta_{X_{n-1}^{\mathbf{i}_{n-1}}} M$.

Following the convention of [6], the collection of random variables $X_p^{\mathbf{i}_p}$, with $\mathbf{i}_p \in [N]^{p+1}$, and $0 \leq p \leq n$, can be associated to the vertices of a planar forest of height n. The sequence of integers $\mathbf{i}_n$ represents the complete genealogy of an individual at level n. For instance, an individual $X_2^{\mathbf{i}_2}$ in the second generation is associated with the triplet $\mathbf{i}_2 = (i_0, i_1, i_2)$ of integers that indicates that he his the i_2-th child of the i_1-th child of the i_0-th root ancestor individual. Running back in time, we can trace back the complete ancestral line of a given current individual $X_n^{\mathbf{i}_n}$ at the n-th generation

$$X_0^{\mathbf{i}_0} \longleftarrow X_1^{\mathbf{i}_1} \longleftarrow \ldots \longleftarrow X_{n-1}^{\mathbf{i}_{n-1}} \longleftarrow X_n^{\mathbf{i}_n}$$

The reproduction step, and the mutation transition introduced in (2.1) have a natural interpretation in terms of these random forests. Roughly speaking, the random forests introduced above represent all the transitions of a reproduction-mutation genetic algorithm.

To be more precise, it is convenient to introduce some additional algebraic structures. We equip $\mathcal{A} = [N]^{[N]}$ with the unital semigroup structure associated with the composition operation $ab := a \circ b$, and the identity element Id. We equip the set $\mathcal{A}$ with the partial order relation defined for any pair of mappings $a, c \in \mathcal{A}$ by the following formula

$$a \leq c \Longleftrightarrow \exists b \in \mathcal{A} \quad a = bc$$

For any collection of mappings $(a_p)_{0 \leq p \leq n}$, we notice that

$$a_{0,n} \leq a_{1,n} \leq \ldots \leq a_{n-1,n} \leq a_n \leq Id$$

with the composition semigroup $a_{p,n} = a_p a_{p+1,n}$, $0 \leq p < n$; and the convention $a_{n,n} = a_n$. For any weakly decreasing sequence of mappings $(b_0, \ldots, b_n) \in \mathcal{A}^{n+1}$ (that is, any sequence s.t. $b_i \geq b_{i+1}$, we set

$$\mathcal{X}_n^{(b_n, \ldots, b_0)} := \left(X_n^{b_n(i), \ldots, b_1(i), b_0(i)} \right)_{1 \leq i \leq N}$$

In this notation, it is now easy to prove that the neutral genetic model can be seen as a particular way to explore the random forest introduced above

$$\mathcal{X}_0^{Id} \xrightarrow{reproduction} \mathcal{X}_0^{A_0} \xrightarrow{mutation} \mathcal{X}_1^{A_0, Id} \xrightarrow{reproduction} \mathcal{X}_1^{A_0 A_1, A_1} \xrightarrow{mutation} \mathcal{X}_2^{A_0 A_1, A_1, Id} \longrightarrow \ldots$$

with a sequence of independent random variables $(A_p)_{0 \leq n}$, randomly chosen in the set $\mathcal{A}$. More generally, the distribution laws of the neutral genetic model are given by the formulae

$$\widehat{\Gamma}_{\eta,n}(F) = \mathbb{E}_\eta \left(F \left(\mathcal{X}_n^{A_{0,n}, A_{1,n}, \ldots, A_{n-1,n}, A_n} \right) \right)$$

with the semigroup $A_{p,n} := A_p A_{p+1,n}$, with $0 \leq p < n$, and the convention $A_{n,n} = Id$. In much the same way, we also have that

$$\Gamma_{\eta,n+1}(F) = \widehat{\Gamma}_{\eta,n} \mathcal{M}(F) = \mathbb{E}_\eta \left(F \left(\mathcal{X}_{n+1}^{A_{0,n}, A_{1,n}, \ldots, A_{n-1,n}, A_n, Id} \right) \right)$$

2.3 Asymptotic behavior

The permutation mappings $\sigma \in \mathcal{G}_N$ are the largest elements of $\mathcal{A}$, while the smallest elements of $\mathcal{A}$ are given by the constants elementary mappings $e_i(j) = i$, for any $j \in [N]$, with $1 \leq i \leq N$. The set of these lower bounds is denoted by

$$\partial \mathcal{A} = \{a \in \mathcal{A} \ : \ |a| = 1\} = \{e_i \ : \ i \in [N]\},$$

where $|a|$ stands for the number of elements in the image of a. We also let $(B_p)_{0 \leq p \leq n}$ be the weakly decreasing Markov chain on $\mathcal{A}$ defined by

$$\forall n \geq 1 \qquad B_n := A_n B_{n-1} \tag{2.2}$$

with the initial condition $B_0 = A_0$, where $(A_p)_{0 \leq p \leq n}$ stands for a sequence of independent and uniformly distributed random variables on the set of mappings $\mathcal{A}$. We let T be the first time the chain B_n enters in the set $\partial \mathcal{A}$.

$$T := \inf \{n \geq 0 \ : \ B_n \in \partial \mathcal{A}\} \tag{2.3}$$

In terms of genealogy, the Markov chain (2.2) represent the ancestral branching process from the present generation at time 0 back into the past. In this interpretation, the mapping A_n in formula (2.2) represents the way the individuals choose their parents in the previous ancestral generation. The range of A_n represents successful parents with direct descendants, whereas the range of B_n represents successful ancestors with descendants in all the generations till time 0. The random variable T represents the time to most recent common ancestor of an initial population with $|B_0|$ individuals.

We are now in position to state the two main results of this article.

Theorem 2.1 *The Markov chain B_n is absorbed by the boundary $\partial \mathcal{A}$ in finite time, and B_T is uniformly distributed in $\partial \mathcal{A}$. In addition, for any time horizon $n \geq 0$, we have*

$$\mathbb{P}(T > n) \leq K \left(\frac{n}{N} \vee 1 \right) \exp \left(- \left(\frac{n}{N} - 1 \right)_+ \right) \quad \text{with} \quad K := e \prod_{l \in \mathbb{N}^*} \left(1 - \frac{2}{(l+1)(l+2)} \right)^{-1}$$

The Theorem follows from (4.1).

We further assume that the Markov transition M has an invariant probability measure μ on E, and we denote by $\widehat{\Gamma}_\mu$ the probability measure on E^N defined by

$$\forall F \in \mathcal{B}_b(E^N) \qquad \widehat{\Gamma}_\mu(F) := \mathbb{E}_\mu \left(F \left(\mathcal{X}_T^{B_T, \dots, B_1, B_0} \right) \right) \tag{2.4}$$

Theorem 2.2 *The measure $\widehat{\Gamma}_\mu$ is an invariant measure of the N-neutral genetic model $\widehat{\xi}_n$. In addition, if the mutation transition M satisfies the following regularity condition :*

(H) *There exists some $\lambda > 0$, and some finite constant $\delta < \infty$ such that for any $n \geq 0$*

$$\beta(M^n) := \sup_{(x,y) \in E^2} \|M^n(x, \boldsymbol{\cdot}) - M^n(y, \boldsymbol{\cdot})\|_{\mathrm{tv}} \leq \delta \ e^{-\lambda n} \tag{2.5}$$

then, there exists $K' \geq 0$ (a universal constant) such that for $n \geq N + \frac{1}{\lambda}$, we have the estimate

$$\left\| \widehat{\Gamma}_{\eta,n} - \widehat{\Gamma}_\mu \right\|_{\mathrm{tv}} \quad \leq \quad \delta K' \frac{n}{N} \exp \left(- \frac{n}{N + \frac{1}{\lambda}} \right)$$

This Theorem follows from Theorem 2.1, Equation (3.5) and Corollary 4.3. Notice that the regularity condition is a weak condition, most often satisfied in reproduction-mutation genetic models -degenerate cases should be avoided such as, for example, in the finite state case, the splitting of the matrix representation of M into a direct sum, or the existence of several fixed points of the Markov transition. The regularity hypothesis is satisfied for example if $M(x, dy) = \alpha \delta_x(dy) + (1 - \alpha)\mu(dy)$ with $0 < \alpha < 1$ and μ a Lebesgue-type density on E.

We end this section with some consequence of these two theorems. Firstly, we notice that the first assertion of the theorem can alternatively be expressed in terms of the measure Γ_μ defined by

$$\forall F \in \mathcal{B}_b(E^N) \qquad \Gamma_\mu(F) := \mathbb{E}_\mu \left(F \left(\mathcal{X}_T^{B_T,\dots,B_1,B_0,Id} \right) \right)$$

More precisely we notice that, by construction, $\Gamma_\mu(F) = \widehat{\Gamma}_\mu(\mathcal{M}F)$. This readily implies that Γ_μ is an invariant measure of the neutral genetic model ξ_n. Indeed, we have

$$\Gamma_\mu(\mathcal{D}\mathcal{M}(F)) = \widehat{\Gamma}_\mu(\mathcal{M}\mathcal{D}(\mathcal{M}F)) = \widehat{\Gamma}_\mu\mathcal{M}(F) = \Gamma_\mu(F)$$

The reverse assertion follows the same line of arguments.

Notice also that the second assertion of the theorem can be used to estimate the Lyapunov exponent of the distribution semigroup of the neutral genetic particle model; that is we have that

$$\liminf_{n \to \infty} -\frac{1}{n} \log \left\| \widehat{\Gamma}_{\eta,n} - \widehat{\Gamma}_\mu \right\|_{\mathrm{tv}} \geq \frac{\lambda}{\lambda N + 1} \wedge \frac{1}{N}$$

3   Random mappings and coalescent tree based measures

Recall that we denote by $|a|$ we denote the cardinality of the set $a([N])$. Notice that $a \leq c \Rightarrow |a| \leq |c|$, and

$$\mathcal{A}_c := \{a \in \mathcal{A} \ : \ a \leq c\}$$

has size $N^{|c|}$, the number of mappings of $c([N])$ into $[N]$. Also observe that $|\partial\mathcal{A}| = N$, and $\mathcal{A}_{e_i} = \partial\mathcal{A}$, for any $i \in [N]$, where $|\delta A|$ stands for the cardinality of δA. In this notation, the transitions probabilities of the chain B_n introduced in (2.2) are given by

$$\mathbb{P}(B_n = b \mid B_{n-1} = a) = \frac{1}{N^{|a|}} \ 1_{\mathcal{A}_a}(b) \tag{3.1}$$

The uniform distribution $\nu_0(a) := \frac{1}{|\mathcal{A}|}$ on $\mathcal{A}$ is such that $\nu_0(a) = K(Id, a)$, and we also have that $K(e_i, b') = \frac{1}{N} \ 1_{\partial\mathcal{A}}(b')$, where K stands for the Markov transition on $\mathcal{A}$ defined by

$$K(\phi) : a \in \mathcal{A} \longmapsto K(\phi)(a) := \frac{1}{|\mathcal{A}|} \sum_{b \in \mathcal{A}} \phi(ba)$$

for any function ϕ on $\mathcal{A}$. This implies that the uniform measure $\nu_1(a) := \frac{1}{N} 1_{\partial A}(a)$ on the set ∂A is an invariant measure of K. Therefore $\nu_1 = \nu_1 K$.

For any weakly decreasing sequence of mappings $\mathbf{b} = (b_0, \dots, b_n) \in \mathcal{A}^{n+1}$, with a finite length $l(\mathbf{b}) = n$, we write

$$|\mathbf{b}| := \sum_{0 \leq p \leq n} |b_p|$$

For any $\mathbf{c} = (c_0, \dots, c_n) \in \mathcal{A}^{n+1}$, any mapping $a \in \mathcal{A}$, and any $b \geq c_0$, we also write

$$(b, \mathbf{c}) = (b, c_0, \dots, c_n) \qquad \mathbf{c}a := (c_0 a, \dots, c_n a) \quad \text{and} \quad \mathbf{c} \star a := (c_0 a, \dots, c_{t_n} a)$$

where $t_n = \inf\{0 \leq p \leq n \ : \ c_p a \in \partial\mathcal{A}\}$; with the convention $t_n = n$, if $c_p a \notin \partial\mathcal{A}$ for any $0 \leq p \leq n$.

The set of all weakly decreasing excursions from $\mathcal{A}$ into $\partial\mathcal{A}$ is given by

$$\mathcal{C} := \cup_{n \geq 0} \mathcal{C}_n$$

with the sets $\mathcal{C}_n$ of all weakly decreasing excursions $\mathbf{c}$, with length $l(\mathbf{c}) = n \geq 0$, and defined by

$$\mathcal{C}_n = \{\mathbf{c} = (c_0, \dots, c_n) \in (\mathcal{A} - \partial\mathcal{A})^n \times \partial\mathcal{A} \ : \ \forall 0 \leq p < n \ \ c_p \geq c_{p+1}\}$$

We use the convention $\mathcal{C}_0 = \partial\mathcal{A}$. We associate with the random excursion

$$\mathbf{B} = (B_0, B_1, \dots, B_T) \in \mathcal{C}$$

the pair of random excursions

$$\mathbf{B}' = (Id, \mathbf{B}) \in \mathcal{C} \quad \text{and} \quad \mathbf{B}' \star A \in \mathcal{C},$$

where A stands for a uniformly distributed $\mathcal{A}$-valued random variable.

Lemma 3.1 *The Markov chain B_n is absorbed by the boundary $\partial \mathcal{A}$ in finite time, and B_T is uniformly distributed in $\partial \mathcal{A}$. Furthermore, the excursions $\mathbf{B}$ and $\mathbf{B}' \star A$ are distributed on $\mathcal{C}$ according to the same distribution $\mathbf{p} : \mathbf{c} \in \mathcal{C} \mapsto \mathbf{p}(\mathbf{c}) := 1/N^{N+(|\mathbf{c}|-1)}$.*

From

$$\mathbb{P}(T > n) \leq \mathbb{P}(\forall 0 \leq p \leq n \ \ A_p \notin \partial \mathcal{A}) = \left(1 - \frac{1}{N^{N-1}}\right)^{n+1} \tag{3.2}$$

it follows that $\mathbb{P}(T < \infty) = 1$. Furthermore, by symmetry arguments, the entrance point B_T is uniformly distributed in $\partial \mathcal{A}$, that is we have that $\mathbb{P}(B_T = a) = \nu_1(a)$.

From the computation of the conditional expectation in (3.1), we have that $\mathbf{p}$ is the distribution of the excursion $\mathbf{B}$. Notice that it is a well defined probability measure on the set of excursions $\mathcal{C}$ since

$$\mathbf{p}(\mathcal{C}_n) = \mathbb{P}(T = n) \Rightarrow \mathbf{p}(\mathcal{C}) = \mathbb{P}(T < \infty) = 1$$

By construction, we have

$$\mathbf{B}' = (B_0', \ldots, B_T', B_{T+1}') = (Id, \mathbf{B}) = (Id, B_0, \ldots, B_T) .$$

This implies that

$$\mathbf{B}' \star A := (B_0'A, B_1'A, \ldots, B_S'A) = (A, B_0A, B_1A, \ldots, B_{S-1}A)$$

with the stopping time

$$S = \inf \{n \geq 0 \ : \ B_n'A \in \partial \mathcal{A}\} = \inf \{n \geq 0 \ : \ B_{n-1}A \in \partial \mathcal{A}\}$$

and the convention $B_{-1} = Id$. The Lemma follows since, by the very definition of B_n, the Markov chain $B_{n-1}A$, $n \geq 0$ has the same distribution as the Markov chain B_n. We notice that, for any test function f we have:

$$
\begin{aligned}
\mathbb{E}(f(\mathbf{B}' \star A)) &= \sum_{n \geq 0} \mathbb{E}(f(B_{-1}A, B_0A, B_1A, \ldots, B_{n-1}A) \ 1_{S=n}) \\
&= \sum_{n \geq 0} \sum_{(c_0, \ldots, c_n) \in \mathbf{C}_n} f(c_0, \ldots, c_n) \nu(c_0) K(c_0, c_1) \ldots K(c_{n-1}, c_n) \\
&= \sum_{\mathbf{c} \in \mathbf{C}} f(\mathbf{c}) \ \mathbf{p}(\mathbf{c}) = \mathbb{E}(f(\mathbf{B}))
\end{aligned}
\tag{3.3}
$$

(where ν stand for the uniform probability measure on $\mathcal{A}$). $\blacksquare$

Definition 3.2 *We associate with a probability measure $\eta \in \mathcal{P}(E)$, and a weakly decreasing sequence of mappings $\mathbf{b} = (b_0, \ldots, b_n) \in \mathcal{A}^{n+1}$, a probability measure $\eta_{(b_0, \ldots, b_n)} \in \mathcal{P}(E^N)$ defined for any $F \in \mathcal{B}_b(E^N)$ by*

$$\eta_{(b_0, \ldots, b_n)}(F) := \mathbb{E}_\eta \left(F\left(\mathcal{X}_n^{b_n, \ldots, b_0}\right)\right)$$

By the definition of the neutral genetic model, for any weakly decreasing sequence of mappings $\mathbf{b}$ in $\mathcal{A}$, and any mapping $a \in \mathcal{A}$ we have

$$\eta_{\mathbf{b}}\mathcal{M} = \eta_{(Id, \mathbf{b})} \quad \text{and} \quad \eta_{\mathbf{b}}D_a = \eta_{\mathbf{b}a} \tag{3.4}$$

In addition, for any excursion $\mathbf{c} \in \mathcal{C}$, any path $\mathbf{a} = (a_1, \ldots, a_m) \in \partial \mathcal{A}^m$, and any $a \in \mathcal{A}$ we have

$$\eta_{(\mathbf{c}, \mathbf{a})} = (\eta M^m)_{\mathbf{c}} \qquad \mu_{(\mathbf{c}, \mathbf{a})} = \mu_{\mathbf{c}} \quad \text{and therefore} \quad \mu_{\mathbf{c}a} = \mu_{\mathbf{c} \star a}$$

Since $(A_0, \ldots, A_n)$ has the same distribution as the reversed sequence $(A_n, \ldots, A_0)$, we also find that

$$
\begin{aligned}
\widehat{\Gamma}_{\eta, n}(F) &= \mathbb{E}_\eta \left(F\left(\mathcal{X}_n^{A_{0,n}, A_{1,n}, \ldots, A_{n-1,n}, A_n}\right)\right) \\
&= \mathbb{E}_\eta \left(F\left(\mathcal{X}_n^{A_n \ldots A_0, A_{n-1} \ldots A_0, \ldots, A_1 A_0, A_0}\right)\right) \\
&= \mathbb{E}_\eta \left(F\left(\mathcal{X}_n^{B_n, B_{n-1}, \ldots, B_1, B_0}\right)\right) = \mathbb{E}\left(\eta_{(B_0, \ldots, B_n)}(F)\right)
\end{aligned}
$$

Proposition 3.3 *If the Markov transition M has an invariant probability measure μ on E, then the probability measure*

$$\widehat{\Gamma}_\mu(F) := \sum_{\mathbf{c} \in \mathbf{C}} \mathbf{p}(\mathbf{c}) \ \mu_{\mathbf{c}}(F) = \mathbb{E}\left(\mu_{(B_0,\ldots,B_T)}(F)\right)$$

is an invariant measure of the neutral genetic model $(\widehat{\xi}_n)_{n \geq 0}$. Under the regularity condition **(H)**, *the neutral genetic model has a unique invariant measure $\widehat{\Gamma}_\mu$, and we have the estimate*

$$\forall n \geq 0 \qquad \left\|\widehat{\Gamma}_{\eta,n} - \widehat{\Gamma}_\mu\right\|_{\mathrm{tv}} \leq \delta \ \mathbb{E}(e^{-\lambda(n-T)} \ 1_{T \leq n}) + 2 \ \mathbb{P}(T > n) \tag{3.5}$$

with the pair of parameters (δ, λ) introduced in (2.5).

Indeed, if we take the excursion $\mathbf{B} = (B_0, B_1, \ldots, B_T)$, then by (3.3)

$$
\begin{aligned}
\mathbb{E}\left(\mu_{\mathbf{B}} M^{\otimes N} D_A(F)\right) &= \mathbb{E}\left(\mu_{(Id,\mathbf{B})} D_A(F)\right) \\
&= \mathbb{E}\left(\mu_{(Id,\mathbf{B})A}(F)\right) = \mathbb{E}\left(\mu_{(Id,\mathbf{B})\star A}(F)\right) = \mathbb{E}\left(\mu_{\mathbf{B}}(F)\right)
\end{aligned}
$$

Since $\mathbb{E}\left(\mu_B(F)\right) = \widehat{\Gamma}_\mu(F)$, the end of the proof of the first assertion of the theorem is completed. To prepare the proof of the second assertion, firstly we notice that

$$\eta_{(\mathbf{c},a)}(F) = \mathbb{E}_\eta\left(F\left(\mathcal{X}_{n+1}^{(a,c_n,\ldots,c_0)}\right)\right) = \mathbb{E}_{\eta M}\left(F\left(\mathcal{X}_n^{(c_n,\ldots,c_0)}\right)\right) = (\eta M)_{\mathbf{c}}(F)$$

for any excursion $\mathbf{c} = (c_0, \ldots, c_n) \in \mathcal{C}$, with lenght $l(\mathbf{c}) = n$, and any $a \in \partial \mathcal{A}$. More generally, for any path $\mathbf{a} = (a_1, \ldots, a_m) \in \partial \mathcal{A}^m$, we have

$$\eta_{(\mathbf{c},\mathbf{a})} = (\eta M^m)_{\mathbf{c}}$$

This clearly implies that

$$
\begin{aligned}
\widehat{\Gamma}_{\eta,n}(F) &= \mathbb{E}\left(\eta_{(B_0,\ldots,B_n)}(F) \ 1_{T \leq n}\right) + \mathbb{E}\left(\eta_{(B_0,\ldots,B_n)}(F) \ 1_{T > n}\right) \\
&= \mathbb{E}\left((\eta M^{n-T})_{(B_0,\ldots,B_T)}(F) \ 1_{T \leq n}\right) + \mathbb{E}\left(\eta_{(B_0,\ldots,B_n)}(F) \ 1_{T > n}\right)
\end{aligned}
$$

To take the final step, we consider the decomposition

$$
\begin{aligned}
\widehat{\Gamma}_{\eta,n}(F) - \widehat{\Gamma}_\mu(F) &= \mathbb{E}\left(\left[(\eta M^{n-T})_{(B_0,\ldots,B_T)} - \mu_{(B_0,\ldots,B_T)}\right](F) \ 1_{T \leq n}\right) \\
&\qquad + \mathbb{E}\left(\eta_{(B_0,\ldots,B_n)}(F) - \mu_{(B_0,\ldots,B_n)}(F)\right) \ 1_{T > n}
\end{aligned}
$$

For any excursion $\mathbf{c} = (c_0, \ldots, c_n) \in \mathcal{C}$, with lenght $l(\mathbf{c}) = n$, we notice that

$$[\eta_{\mathbf{c}} - \mu_{\mathbf{c}}](F) = \int_E [\eta - \mu](dx) \times \mathbb{E}_{\delta_x}\left(F\left(\mathcal{X}_n^{(c_n,\ldots,c_0)}\right)\right)$$

Since $\mu = \mu M^{n-T}$, applying (2.5), we get

$$\left|\widehat{\Gamma}_{\eta,n}(F) - \widehat{\Gamma}_\mu(F)\right| \leq \delta \ \mathbb{E}(e^{-\lambda(n-T)} \ 1_{T \leq n}) + 2\mathbb{P}(T > n)$$

for any $\|F\| \leq 1$. This ends the proof of the proposition. ∎

4 Absorption times behavior

We study here the renormalized time T/N for large integers N, where T is defined in (2.3), when the mapping-valued Markov chain $(B_n)_{n \in \mathbb{N}}$ has the identity as initial state. Let us recall that T can be interpreted as the time a neutral genetic model with N particles has to look backward to encounter its first common ancestor. The convergence in law of T/N shows that T is of order N.

We notice that T only depends on $(|B_n|)_{n \in \mathbb{N}}$ and that this $[N]$-valued stochastic chain is Markovian, with transitions described by

$$\forall n \in \mathbb{N}, \forall 1 \leq p \leq q \leq N, \qquad \mathbb{P}(|B_{n+1}| = p \mid |B_n| = q) = S(q,p)\frac{(N)_p}{N^q}$$

(where $S(q,p)$ is the Stirling number of the second kind giving the number of ways of partitioning the set $[q]$ into p non empty blocks and where $(N)_p := N!/(N-p)!$) and starting from N if $B_0 = \mathrm{Id}$.

This observation leads us to define for $N \in \mathbb{N}^*$, a triangular transition matrix $M^{(N)}$ by

$$\forall\ 1 \leq p, q \leq N, \qquad M^{(N)}_{q,p} \quad := \quad S(q,p)\frac{(N)_p}{N^q}$$

and to consider for any $1 \leq i \leq N$, a Markov chain $R^{(N,i)} := (R^{(N,i)}_n)_{n\in\mathbb{N}}$ starting from i and whose transitions are governed by $M^{(N)}$. Such a Markov chain is (a.s.) non-increasing and 1 is an absorption state. We denote

$$S^{(N,i)} \quad := \quad \inf\{n \in \mathbb{N}\ :\ R^{(N,i)}_n = 1\}$$

so that T has the same law as $S^{(N,N)}$.

The following convergence in law is a consequence of [3], Th. 3.

Theorem 4.1 *Let $(i_N)_{N\in\mathbb{N}^*}$ be a sequence of integers satisfying $1 \leq i_N \leq N$ for any $N \in \mathbb{N}^*$ and diverging to infinity. Then*

$$\frac{S^{(N,i_N)}}{N} \quad \xrightarrow[N\to+\infty]{\mathcal{L}} \quad \sum_{l\in\mathbb{N}} \frac{2}{(l+1)(l+2)}\mathcal{E}_l$$

where $(\mathcal{E}_l)_{l\in\mathbb{N}}$ is an independent family of exponential variables of parameter 1.

The goal of this section is to prove the

Theorem 4.2 *Let $(i_N)_{N\in\mathbb{N}^*}$ be a sequence of integers satisfying $1 \leq i_N \leq N$ for any $N \in \mathbb{N}^*$ and diverging to infinity. We have for any fixed $0 \leq \alpha < 1$,*

$$\sup_{N\in\mathbb{N}^*} \mathbb{E}[\exp(\alpha S^{(N,i_N)}/N)] \quad \leq \quad \exp(\alpha)\Pi(\alpha)$$

where

$$\forall\ 0 \leq \alpha < 1, \qquad \Pi(\alpha) \quad := \quad \prod_{l\in\mathbb{N}} \frac{1}{1 - \frac{2\alpha}{(l+1)(l+2)}}$$

is the Laplace transform of the above limit law. Thus for any continuous function $f : \mathbb{R}_+ \to \mathbb{R}_+$ verifying $\lim_{\alpha\to 1_-} \limsup_{s\to+\infty} \exp(-\alpha s)f(s) = 0$, we are insured of

$$\lim_{N\to\infty} \mathbb{E}[f(S^{(N,i_N)}/N)] \quad = \quad \mathbb{E}\left[f\left(\sum_{l\in\mathbb{N}} \frac{2}{(l+1)(l+2)}\mathcal{E}_l\right)\right]$$

In particular, for any given $0 \leq \alpha < 1$, via a Markov inequality, we deduce the exponential upper bound

$$\forall\ N \in \mathbb{N}^*, \forall\ 1 \leq i \leq N, \forall\ n \in \mathbb{N}, \qquad \mathbb{P}[S^{(N,i)} \geq n] \quad \leq \quad \exp(\alpha)\Pi(\alpha)\exp(-\alpha n/N)$$
$$\leq \quad \frac{K}{1-\alpha}\exp(-\alpha n/N)$$

with $K = e\left(\prod_{l\in\mathbb{N}^*}\left(1 - \frac{2}{(l+1)(l+2)}\right)\right)^{-1}$. Optimizing this inequality with respect to $0 \leq \alpha < 1$, we obtain that for any positive integers n and N,

$$\forall\ 1 \leq i \leq N, \qquad \mathbb{P}[S^{(N,i)} \geq n] \quad \leq \quad K\left(\frac{n}{N} \vee 1\right)\exp(-(n/N-1)_+) \tag{4.1}$$

(of course this bound does not give any relevant information for $n \leq N$, when the l.h.s. probability is not small). As it was explained in the second section, such an inequality is useful to estimate convergence to equilibrium for neutral genetic models and the results presented in the introduction follow from it.

Indeed, in view of proposition 3.3, we still need another estimate, but it is an immediate consequence of the above bound:

Corollary 4.3 *There exists a constant $K' \geq 0$ such that for any $\lambda > 0$, any $N \in \mathbb{N}^*$, any $1 \leq i \leq N$ and any $n \geq N + 1/\lambda$, we have*

$$\mathbb{E}\left[\exp(-\lambda(n - S^{(N,i)}))\mathbf{1}_{\{S^{(N,i)} \leq n\}}\right] \quad \leq \quad K'\frac{n}{N}\exp\left(-\left(1 + \frac{1}{\lambda N}\right)^{-1}\frac{n}{N}\right)$$

Indeed, let $m \in [n]$ be given, we have

$$
\begin{aligned}
&\mathbb{E}\left[\exp(-\lambda(n - S^{(N,i)}))\mathbf{1}_{\{S^{(N,i)} \leq n\}}\right] \\
&= \quad \mathbb{E}\left[\exp(-\lambda(n - S^{(N,i)}))\mathbf{1}_{\{S^{(N,i)} \leq m\}}\right] + \mathbb{E}\left[\exp(-\lambda(n - S^{(N,i)}))\mathbf{1}_{\{m \leq S^{(N,i)} \leq n\}}\right] \\
&\leq \quad \exp(-\lambda(n - m)) + \mathbb{P}(S^{(N,i)} > m) \\
&\leq \quad \exp(-\lambda(n - m)) + K\left(\frac{m}{N} \vee 1\right)\exp(-(m/N - 1)_+) \\
&\leq \quad (1 + K)\frac{n}{N}\max(\exp(-\lambda(n - m)), \exp(-(m/N - 1)_+))
\end{aligned}
$$

Where we have used that $m \leq n$ and that $n \geq N$. Optimizing the last term, we are led to consider

$$m \quad = \quad \left\lfloor \frac{n}{1 + \frac{1}{\lambda N}} \right\rfloor$$

Note that this integer number is larger than N for $n \geq N + 1/\lambda$ and the corollary's inequality follows easily from the obvious bounds

$$\frac{n}{1 + \frac{1}{\lambda N}} - 1 \quad \leq \quad m \quad \leq \quad \frac{n}{1 + \frac{1}{\lambda N}}$$

∎

Remark 4.4 *The limit distribution appearing in Theorem 4.2 is the same as the law of the coalescence time for the Kingman process (see for instance [8]). This could have been expected, since it is known that, suitably "rearranged", the mapping-valued Markov process $(B_{\lfloor Nt \rfloor})_{t \in \mathbb{R}_+}$ converges to the Kingman coalescent when $N \to +\infty$. Nevertheless, at our best knowledge, this convergence takes place in a weak sense which does not permit to deduce the results presented here. In fact, the latter could serve to strengthen the previous convergence.*

Remark 4.5 *An exact formula for the distribution of $S^{(N,i)}$ is known (see Lemma 3.1 in [11]) but it is not sufficient to deduce the above bounds (some different bounds are deduced in [11], see Equations (4.1), (4.2), (4.4)).*

Before proving Theorem 4.2, we will investigate the simpler problem where only negative jumps of unit length are permitted. More precisely, let $\widetilde{M}^{(N)}$ be the transition matrix defined by

$$\forall\, 1 \leq q, p \leq N, \qquad \widetilde{M}_{q,p}^{(N)} \quad = \quad \begin{cases} \dfrac{(N)_q}{N^q} & \text{, if } p = q \\[2mm] 1 - \dfrac{(N)_q}{N^q} & \text{, if } p = q - 1 \\[2mm] 0 & \text{, otherwise} \end{cases}$$

Every corresponding notion will be overlined by a tilde. Then we have a result which is similar to Theorem 4.2, with nevertheless some slight differences:

Proposition 4.6 *Let $(i_N)_{N \in \mathbb{N}^*}$ be a sequence of integers satisfying $1 \leq i_N \leq N$ for any $N \in \mathbb{N}^*$, diverging to infinity and such that $a := \lim_{N \to \infty} i_N/N$ exists in $[0, 1]$. Then*

$$\frac{\widetilde{S}^{(N,i_N)}}{N} \quad \underset{N \to +\infty}{\overset{\mathcal{L}}{\longrightarrow}} \quad a + \sum_{l \in \mathbb{N}} \frac{2}{(l+1)(l+2)}\mathcal{E}_l$$

where $(\mathcal{E}_l)_{l \in \mathbb{N}}$ is an independent family of exponential variables of parameter 1. Furthermore, we have for any fixed $0 \leq \alpha < 1$,

$$\sup_{N \in \mathbb{N}^*} \mathbb{E}\left[\exp\left(\alpha\frac{\widetilde{S}^{(N,i_N)} - i_N + 1}{N}\right)\right] \quad = \quad \Pi(\alpha)$$

Remark 4.7 *The above convergence in law is a consequence of Th. 3 of [3] but we will write a simpler proof below.*

Let $2 \leq i \leq N$ be given. By a backward iteration and with the convention that $\widetilde{T}_i^{(N)} = 0$, we define for $1 \leq j \leq i - 1$,

$$\widetilde{T}_j^{(N)} \quad := \quad \inf\left\{ n \in \mathbb{N} \,:\, \widetilde{R}_{\widetilde{T}_{j+1}^{(N)} + n}^{(N,i)} = j \right\}$$

So $\widetilde{T}_j^{(N)}$ is the time necessary for the Markov chain $(\widetilde{R}_n^{(N,i)})_{n \in \mathbb{N}}$ to jump from $j + 1$ to j. For $1 \leq j \leq i - 1$, we set $\widehat{T}_j^{(N)} := \widetilde{T}_j^{(N)} - 1$ and $p_{N,j} := \widetilde{M}_{j+1,j}^{(N)} = (N)_{j+1}/N^{j+1}$. It is clear that $\widehat{T}_j^{(N)}$ is distributed as a geometric law of parameter $p_{N,j}$, namely,

$$\forall\, m \in \mathbb{N}, \qquad \mathbb{P}[\widehat{T}_j^{(N)} = m] \quad = \quad (1 - p_{N,j}) p_{N,j}^m$$

Furthermore, the variables $\widetilde{T}_j^{(N)}$, for $1 \leq j \leq i - 1$, are independent and we can write

$$\widetilde{S}^{(N,i_N)} - i_N + 1 \quad = \quad \sum_{1 \leq j \leq i_N - 1} \widehat{T}_j^{(N)}$$

Thus we get for any $\alpha < 1$,

$$\begin{aligned}
\mathbb{E}\left[\exp\left(\alpha \frac{\widetilde{S}^{(N,i_N)} - i_N + 1}{N} \right) \right] &= \prod_{1 \leq j \leq i_N - 1} \frac{1 - p_{N,j}}{1 - \exp(\alpha/N) p_{N,j}} \\
&\leq \prod_{1 \leq j \leq i_N - 1} \frac{1}{1 - \frac{2\alpha}{j(j+1)}} \\
&\leq \Pi(\alpha)
\end{aligned}$$

Let $i \in \mathbb{N}^*$ be fixed (at first). By assumption, for N large enough, we will have $i_N \geq i$ and quite obviously, $\widetilde{S}^{(N,i_N)} - i_N + 1$ will stochastically dominate $\widetilde{S}^{(N,i)} - i + 1$, which implies, for $0 \leq \alpha < 1$,

$$\mathbb{E}\left[\exp\left(\alpha \frac{\widetilde{S}^{(N,i_N)} - i_N + 1}{N} \right) \right] \quad \geq \quad \mathbb{E}\left[\exp\left(\alpha \frac{\widetilde{S}^{(N,i)} - i + 1}{N} \right) \right]$$

and thus

$$\begin{aligned}
\liminf_{N \to \infty} \mathbb{E}\left[\exp\left(\alpha \frac{\widetilde{S}^{(N,i_N)} - i_N + 1}{N} \right) \right] &\geq \lim_{N \to \infty} \mathbb{E}\left[\exp\left(\alpha \frac{\widetilde{S}^{(N,i)} - i + 1}{N} \right) \right] \\
&= \prod_{1 \leq j \leq i - 1} \frac{1}{1 - \frac{2\alpha}{j(j+1)}}
\end{aligned}$$

Letting i go to infinity, it appears that

$$\lim_{N \to \infty} \mathbb{E}\left[\exp\left(\alpha \frac{\widetilde{S}^{(N,i_N)} - i_N + 1}{N} \right) \right] \quad = \quad \Pi(\alpha)$$

and the last part of Proposition 4.6 follows. We also obtain that for $0 \leq \alpha < 1$,

$$\lim_{N \to \infty} \mathbb{E}\left[\exp\left(\alpha \frac{\widetilde{S}^{(N,i_N)}}{N} \right) \right] \quad = \quad \exp(\alpha a) \Pi(\alpha)$$

and to conclude to the announced convergence in law, it remains to check that above convergence is uniform in some compact right neighborhood of 0. But this is a consequence of Dini's theorem, since the r.h.s. is continuous in $\alpha \in [0, 1)$ and for all fixed $N \in \mathbb{N}^*$, the mapping $[0, 1) \ni \alpha \mapsto \mathbb{E}[\exp(\alpha \widetilde{S}^{(N,i_N)}/N)]$ is increasing. ∎

We now proceed to the proof of Theorem 4.2. Since $S^{(N,i_N)}$ is stochastically dominated by $\widetilde{S}^{(N,i_N)}$:

$$\forall\, n \in \mathbb{N}, \qquad \mathbb{P}[S^{(N,i_N)} \geq n] \quad \leq \quad \mathbb{P}[\widetilde{S}^{(N,i_N)} \geq n]$$

This fact is based on two observations: on one hand, assuming that the Markov chain $R^{(N,i_N)}$ is in state $j \in [i_N]$ at some time n, it will wait the same time to jump out of it as $\widetilde{R}^{(N,i_N)}$, but $R^{(N,i_N)}$ will visit less states in $[i_N]$ than $\widetilde{R}^{(N,i_N)}$. Taking into account these two facts, the above inequalitiy follows immediately. Details are left to the reader: one can e.g. construct a coupling between $R^{(N,i_N)}$ and $\widetilde{R}^{(N,i_N)}$ such that $\mathbb{P}[\widetilde{S}^{(N,i_N)} \geq S^{(N,i_N)}] = 1$ (for a general reference on the subject, see for instance [10]). The same trick is used implicitly in [11], p. 83.

In particular we get that for any $\alpha \geq 0$,

$$\sup_{N \in \mathbb{N}^*} \mathbb{E}[\exp(\alpha S^{(N,i_N)}/N)] \quad \leq \quad \sup_{N \in \mathbb{N}^*} \mathbb{E}[\exp(\alpha \widetilde{S}^{(N,i_N)}/N)]$$

so the wanted upper bound follows from Proposition 4.6.

∎

5 Planar genealogical tree based representations

As we promised in the end of the introduction, this last section is concerned with designing a description of the invariant measure $\widehat{\Gamma}_\mu$ in terms of planar genealogical trees. Most of the arguments are only sketched, since they follow the same lines as the ones developed in [1].

We start by recalling that for any pair of mappings $a \leq b$, we have $a = cb$ for some non unique $c \in \mathcal{A}$. The mapping b induces the following equivalence relation $\sim_b$ on $\mathcal{A}$

$$c \sim_b c' \iff cb = c'b$$

Equivalently, $c \sim_b c'$ if and only if the restriction of c and c' to the image of a are equal, so that the equivalence class $\overline{c} \in \mathcal{A}_{/\sim_b}$ can be seen as the corresponding map $\overline{c}$ from $b([N])$ into $[N]$; and we therefore have

$$\left| \mathcal{A}_{/\sim_b} \right| = N^{|b|}$$

In terms of backward genealogies, for a given fixed pair of mappings $a \leq b$, the mapping $\overline{c}$ represents the way the individuals $b([N])$ choose their parents, and the composition mapping $a = \overline{c}b$ represents the way the individuals $[N]$ choose their grand parents in $a([N])$. Notice that in this interpretation, not all the individuals in the previous generations are ancestors, the individuals that do not leave descendants are not counted.

We further suppose that we are given a function

$$\theta \; : \; (a,b) \in \{(a',b') \in \mathcal{A}^2 \; : \; a' \leq b'\} \mapsto \theta(a,b) \in \mathbb{R}$$

such that $\theta(a,b) = \theta(a',b')$, as soon as the pair of mappings (a,a'), and resp. (b,b'), only differs in the way the sets $a([N])$ and $a'([N])$, and resp. $b([N])$ and $b'([N])$, are labeled. That is, $\theta(a,b) = \theta(a',b')$ whenever there exist increasing bijections α resp. β from $a([N])$ to $a'([N])$ resp. $b([N])$ to $b'([N])$ s.t. $a' = \alpha a$ and $b' = \beta b$. We write $(a',b') \equiv (a,b)$ when this property is satisfied. The relation $\equiv$ is an equivalence relation. We also mention that, for a given pair (a,b), there are

$$\binom{N}{|a|} \times \binom{N}{|b|} = \frac{(N)_{|a|}}{|a|!} \times \frac{(N)_{|b|}}{|b|!}$$

pairs of mappings s.t. $(a',b') \equiv (a,b)$. Moreover, there is a unique pair (a',b') with the property $a'([N]) = [|a|]$ and $b'([N]) = [|b|]$. We use the familiar combinatorial terminology and refer to this process (the replacement of the set $a([N])$ by $[|a|]$, of a by a', and so on), as the *standardization process* (of subsets of $\mathbb{N}$, of maps between these subsets...) Therefore, to evaluate $\theta(a,b)$, up to a change of indexes we can replace the pair of sets $(a([N]),b([N]))$, by the pair $([|a|],[|b|])$, and the pair of mappings $(a,\overline{c})$, by a unique pair of surjections $(a',c') \in \left([|a|]^{[N]} \times [|b|]^{[|a|]}\right)$. In other terms, the pair of mappings $b \leq a$ is canonically associated to a pair of surjections

$$[|b|] \xleftarrow{\;c'\;} [|a|] \xleftarrow{\;a'\;} [N]$$

To describe precisely the planar tree representations of the stationary population of the neutral genetic model (2.1), it is convenient to introduce a series of multi index notation.

Definition 5.1 *We let $\mathcal{G}_q$ be the symmetric group of all permutations of the set $[q]$, and for any weakly decreasing sequence of integers*

$$\mathbf{q} \in \mathcal{Q}_n := \left\{ (q_0, \ldots, q_n) \in [N]^{n+1} \; : \; q_0 = N \geq q_1 \geq \ldots \geq q_n > 1 \right\}$$

we use the multi index notation

$$\mathbf{q}! := \prod_{0 \leq k \leq n} q_k! \qquad (N)_{\mathbf{q}} = \prod_{0 \leq k \leq n} (N)_{q_k} \quad \text{and} \quad |\mathbf{q}| := \sum_{0 \leq k \leq n} q_k$$

We also introduce the sets

$$\mathcal{C}_n(\mathbf{q}) \quad := \quad \left\{ \mathbf{c} = (c_0, \ldots, c_n) \in \mathcal{C}_n \; : \; \forall 0 \leq p \leq n \;\; |c_p| = q_{p+1} \right\} \quad \text{and} \quad \mathcal{G}_n(\mathbf{q}) := \prod_{0 \leq p \leq n} \mathcal{G}_{q_p}$$

with the convention $q_{n+1} = 1$. Finally, we denote by $\mathcal{S}_n(\mathbf{q})$ the set of all sequences of sujections $\mathbf{a} = (a_0, \ldots, a_n) \in \left([q_1]^{[N]} \times \ldots \times [q_n]^{[q_{n-1}]} \times [1]^{[q_n]} \right)$.

Running back in time the arguments given above, any coalescent sequence of mappings $\mathbf{c} = (c_0, \ldots, c_n) \in \mathcal{C}_n(\mathbf{q})$ is associated in a canonical way to a unique sequence of surjective mappings

$$\mathbf{a} = (a_0, \ldots, a_n) \in \mathcal{S}_n(\mathbf{q})$$

so that, up to standardization, we have that

$$\mathbf{c} = \pi(\mathbf{a}) := (a_0, a_1 a_0, a_2 a_1 a_0, \ldots, a_n a_{n-1} \ldots a_0)$$

To put this again in another way, the coalescence sequence $\mathbf{c}$ has, up to standardization, the following backward representation

$$\{1\} \xleftarrow{a_n} [q_n] \xleftarrow{a_{n-1}} [q_{n-1}] \longleftarrow \ldots \longleftarrow [q_2] \xleftarrow{a_1} [q_1] \xleftarrow{a_0} [N] \tag{5.1}$$

In terms of genealogical trees, the mapping a_0 describes the way the N individuals in the present generation choose their parents among q_1 ancestors; the mapping a_1 describes the way these q_1 individuals again choose their parents among q_2 ancestors, and so on.

From these considerations, by symmetry arguments, the invariant measure $\widehat{\Gamma}_\mu$ of Prop. 3.3 can be expressed in the following way

$$\widehat{\Gamma}_\mu(F) = \sum_{n \geq 0} \sum_{\mathbf{q} \in \mathcal{Q}_n} \frac{1}{N^{|\mathbf{q}|}} \frac{(N)_{\mathbf{q}}}{\mathbf{q}!} \sum_{\mathbf{a} \in \mathcal{S}_n(\mathbf{q})} \mu_{\pi(\mathbf{a})}(F)$$

We also have the more synthetic representation formula

$$\widehat{\Gamma}_\mu(F) = \sum_{\mathbf{a} \in \mathcal{S}} \frac{1}{\rho(\mathbf{a})!} \frac{(N)_{\rho(\mathbf{a})}}{N^{|\rho(\mathbf{a})|}} \mu_{\pi(\mathbf{a})}(F) \quad \text{with} \quad \mathcal{S} := \cup_{n \geq 0} \cup_{\mathbf{q} \in \mathcal{Q}_n} \mathcal{S}_n(\mathbf{q})$$

and the multi index mapping $\rho(\mathbf{a}) := \mathbf{q}$, for any $\mathbf{a} = (a_0, \ldots, a_n) \in \mathcal{S}_n(\mathbf{q})$.

Although this parametrization by sequences of surjective maps of the expansion of the invariant measure is already much better than the one in Prop. 3.3, it does not take into account the full symmetry properties of $\widehat{\Gamma}_\mu$. To take fully advantage of it, notice first of all that the labelling of individuals in a genetic population is arbitrary. From the algebraic point of view, the invariant probability measure constructed in Prop. 3.3 inherits this property in the sense that

$$\widehat{\Gamma}_\mu(F) = \widehat{\Gamma}_\mu(F^\sigma)$$

for any $\sigma \in \mathcal{G}_n$, where $F^\sigma(x_1, ..., x_n) := F(x_{\sigma(1)}, ..., x_{\sigma(n)})$. In particular, when computing $\widehat{\Gamma}_\mu(F)$ we will always assume from now on that F is symmetry invariant, that is, that $F^\sigma = F$ for all $\sigma \in \mathcal{G}_n$ -a property that we write $F \in B_b^{sym}(E^N)$. If F does not have this property, its symmetrization $\overline{F}$ has it, where

$$\overline{F} := \frac{1}{n!} \sum_{\sigma \in \mathcal{G}_n} F^\sigma,$$

and the invariance properties of $\widehat{\Gamma}_\mu$ insure that $\widehat{\Gamma}_\mu(F) = \widehat{\Gamma}_\mu(\overline{F})$

We then consider the following natural action of the permutation group $\mathcal{G}_{\mathbf{q}}$ on the set $\mathcal{S}_n(\mathbf{q})$, defined for any $\mathbf{s} = (s_p)_{0 \le p \le n} \in \mathcal{G}_n(\mathbf{q})$, and $\mathbf{a} \in \mathcal{S}_n(\mathbf{q})$ by

$$\mathbf{s}(\mathbf{a}) := (s_1 a_0 s_0^{-1}, \dots, a_n s_n^{-1})$$

We let $\mathcal{Z}(\mathbf{a}) := \{\mathbf{s} : \mathbf{s}(\mathbf{a}) = \mathbf{a}\}$ be the stabilizer of $\mathbf{a}$.

This group action induces a partition of the set $\mathcal{S}_n(\mathbf{q})$ into orbit sets or equivalent classes, where $\mathbf{a}$ and $\mathbf{a}'$ are equivalent if, and only if, there exists $\mathbf{s} \in \mathcal{G}_n(\mathbf{q})$ such that $\mathbf{a}' = \mathbf{s}(\mathbf{a})$. The set of equivalence classes is written $I_n(\mathbf{q})$. In terms of ancestral lines, the genealogical trees associated with the sequences of mappings $\mathbf{s}(\mathbf{a})$ and $\mathbf{a}$ only differ by a change of labels of the ancestors, at each level set. By induction on n, it can be shown that each orbit, or equivalence class, contains an element in the subset $\mathcal{S}'_n(\mathbf{q}) \subset \mathcal{S}_n(\mathbf{q})$ of all sequences of weakly increasing surjections. This observation can be used, together with standard techniques such as lexicographical ordering of sequences of sequences of integers to construct a canonical representative in $\mathcal{S}'_n(\mathbf{q})$ for each equivalence class. Also notice the mappings $\mathbf{a} \mapsto \rho(\mathbf{a})$, and $\mathbf{a} \mapsto \mu_{\pi(\mathbf{a})}$ are invariant; that is we have that $\rho(\mathbf{a}) = \rho(\mathbf{a}')$, and $\mu_{\pi(\mathbf{a}')}(F) = \mu_{\pi(\mathbf{a})}(F)$ for any $F \in B_b^{sym}(E^N)$, as soon as $\mathbf{a}' = \mathbf{s}(\mathbf{a})$, for some relabeling permutation sequence $\mathbf{s}$. Thus, applying the class formula, we find the following formula.

Proposition 5.2 *For any function $F \in \mathcal{B}_b^{sym}(E^N)$, we have*

$$\widehat{\Gamma}_\mu(F) = \sum_{\mathbf{a} \in \mathcal{S}'} \frac{1}{|\mathcal{Z}(\mathbf{a})|} \frac{(N)_{\rho(\mathbf{a})}}{N^{|\rho(\mathbf{a})|}} \mu_{\pi(\mathbf{a})}(F)$$

with the set $\mathcal{S}' := \cup_{n \ge 0} \cup_{\mathbf{q} \in \mathcal{Q}_n} I_n(\mathbf{q})$.

This functional representation formula can be interpreted in terms of genealogical trees. The precise description of this alternative interpretation is notationally consuming, so that it will be only sketched. The reader is refered to [1] for details on the subject.

Recall that a rooted tree $\mathbf{t}$ is an acyclic, connected, directed graph in which any vertex has at most an outgoing edge. The paths are oriented from the vertices to the root, and a leaf in a tree is a vertex without any incoming edge. Notice first that any sequence of mappings $\mathbf{a} = (a_0, \dots, a_n)$ in $\mathcal{S}_n(\mathbf{q})$ gives rise to a directed graph -defined as usual: that is, to any element k in $[q_i]$ we associate a vertex of the graph and a directed arrow to the vertex associated to the element $a_i(k)$ in $[q_{i+1}]$. Two sequences in $S_n(\mathbf{q})$ are equivalent under the action of $\mathcal{G}_n(\mathbf{q})$ if and only if they have the same underlying abstract graph -see [1]. It follows that $I_n(\mathbf{q})$ is isomorphic to the set $\mathcal{T}_n(\mathbf{q})$ of all rooted trees $\mathbf{t}$, with height $ht(\mathbf{t}) = (n+1)$, with $q_{n-(p-1)}$ vertices at each level $p = 1, \dots, (n+1)$, and with leaves only at level $(n+1)$. The set of all planar trees is denoted by $\mathcal{T} := \cup_{n \ge 0} \cup_{\mathbf{q} \in \mathcal{Q}_n} \mathcal{T}_n(\mathbf{q})$. Next, with a slight abuse of notation, for any choice $\mathbf{a} \in \mathcal{S}'_n(\mathbf{q})$ of a representative of a tree $\mathbf{t} \in \mathcal{T} := \cup_{n \ge 0} \cup_{\mathbf{q} \in \mathcal{Q}_n} \mathcal{T}_n(\mathbf{q})$ we set

$$\mu_{\mathbf{t}} = \mu_{\pi(\mathbf{a})} \qquad \mathcal{Z}(\mathbf{t}) = \mathcal{Z}(\mathbf{a}) \quad \text{and} \quad \rho(\mathbf{t}) = \rho(\mathbf{a})$$

In terms of genealogical trees, we finally have that

$$\widehat{\Gamma}_\mu(F) = \sum_{\mathbf{t} \in \mathcal{T}} \frac{1}{|\mathcal{Z}(\mathbf{t})|} \frac{(N)_{\rho(\mathbf{t})}}{N^{|\rho(\mathbf{t})|}} \mu_{\mathbf{t}}(F)$$

A closed formula of the cardinal of the stabilizer $\mathcal{Z}(\mathbf{t})$ can be easily derived using the combinatorial techniques developed in the article [1]. Firstly, we recall some notions.

Definition 5.3 *A forest $\mathbf{f}$ is a multiset of trees, that is an element of the commutative monoid on the set of trees. We denote by $B(\mathbf{t})$ the forest obtained by cutting the root of tree $\mathbf{t}$; that is, removing its root vertex, and all its incoming edges. Conversely, we denote by $B^{-1}(\mathbf{f})$ the tree deduced from the forest $\mathbf{f}$ by adding a common root to its rooted tree. We write*

$$\mathbf{f} := \mathbf{t}_1^{m_1} \dots \mathbf{t}_k^{m_k} \tag{5.2}$$

for the forest with the trees $\mathbf{t}_i$ appearing with multiplicity m_i, with $1 \le i \le k$. When the trees $(\mathbf{t}_i)_{i=1 \dots k}$ are pairwise distinct, we say that the forest is written in normal form.

Definition 5.4 *The symmetry multiset of a tree is defined as follows*

$$\mathbf{t} = B^{-1}(\mathbf{t}_1^{m_1} \dots \mathbf{t}_k^{m_k}) \Longrightarrow \mathbf{S}(\mathbf{t}) := (m_1, \dots, m_k)$$

The symmetry multiset of a forest is the disjoint union of the symmetry multisets of its trees

$$\mathbf{S}(\mathbf{t}_1^{m_1} \dots \mathbf{t}_k^{m_k}) := \left(\underbrace{\mathbf{S}(\mathbf{t}_1), \dots, \mathbf{S}(\mathbf{t}_1)}_{m_1 - terms}, \dots, \underbrace{\mathbf{S}(\mathbf{t}_k), \dots, \mathbf{S}(\mathbf{t}_k)}_{m_k - terms} \right)$$

Following the proof of theorem 3.8 in [1], we find the following closed formula

$$|\mathcal{Z}(\mathbf{t})| = \prod_{0 \le i < ht(\mathbf{t})} \mathbf{S}(B^i(\mathbf{t}))!,$$

where we use the multiset notation $\mathbf{S}(\mathbf{t})! := \prod_{i=1}^{k} m_i!$ if $\mathbf{S}(\mathbf{t}) = (m_1, ..., m_k)$. The above discussion is summarized in the following proposition.

Proposition 5.5

$$\widehat{\Gamma}_\mu(F) = \sum_{\mathbf{t} \in \mathcal{T}} \frac{(N)_{\rho(\mathbf{t})}}{N^{|\rho(\mathbf{t})|} \prod_{0 \le i < ht(\mathbf{t})} \mathbf{S}(B^i(\mathbf{t}))!} \ \mu_{\mathbf{t}}(F)$$

References

[1] Del Moral, P., Patras, F. and Rubenthaler, S., Coalescent Tree Based Functional Representations for some Feynman-Kac Particle Models, arXiv:math.PR/0607453

[2] Del Moral, P., *Feynman-Kac formulae. Genealogical and interacting particle systems with applications*, Probability and its Applications, Springer Verlag, New York, 2004.

[3] Donnelly, P., *Weak convergence to a Markov chain with an entrance boundary: ancestral processes in population genetics*, Ann. Probab. 19 (1991), no. 3, 1102–1117.

[4] Doucet, A., de Freitas, N. and Gordon, N., editors, *Sequential Monte Carlo Methods in Pratice*, Statistics for Engineering and Information Science, Springer, New York, 2001.

[5] Gao, Y., An Upper Bound on the Convergence Rates of Canonical Genetic Algorithms, *Complexity International*, 5, 1998.

[6] Harris, T.E., Branching processes, *Ann. Math. Statist.*, 19, pp. 474–494, 1948.

[7] Kimura, M., *The neutral theory of molecular evolution*, Cambridge University Press, Cambridge, 1983.

[8] Kingman, J. F. C., Exchangeability and the evolution of large populations, *Exchangeability in probability and statistics (Rome, 1981)*, pp. 97–112, North-Holland, Amsterdam-New York, 1982.

[9] Ethier, S.N. and Kurtz, T., *Markov Processes*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1986.

[10] Lindvall, Torgny, *Lectures on the coupling method*, Corrected reprint of the 1992 original, Dover Publications Inc., Mineola, NY, 2002.

[11] Möhle, M., *The time back to the most recent common ancestor in exchangeable population models*, Adv. in Appl. Probab. 36 (2004), no. 1, 78–97.

[12] Tavaré S., Ancestral inference in population genetics, *Lectures on probability theory and statistics, Saint-Flour 2001, Lecture Notes in Math.,* vol. 1837, pp. 1-188, Springer, Berlin, 2004.