DUALITY AND ASYMPTOTICS FOR A CLASS OF NONNEUTRAL DISCRETE MORAN MODELS

THIERRY HUILLET,* *Université de Cergy-Pontoise*

MARTIN MÖHLE,** *Eberhard Karls Universität Tübingen*

Abstract

A Markov chain X with finite state space $\{0, \ldots, N\}$ and tridiagonal transition matrix is considered, where transitions from i to $i-1$ occur with probability $(i/N)(1-p(i/N))$ and transitions from i to $i+1$ occur with probability $(1-i/N)p(i/N)$. Here $p \colon [0,1] \to [0,1]$ is a given function. It is shown that if p is continuous with $p(x) \le p(1)$ for all $x \in [0,1]$ then, for each N, a dual process Y to X (with respect to a specific duality function) exists if and only if $1-p$ is completely monotone with $p(0) = 0$. A probabilistic interpretation of Y in terms of an ancestral process of a mixed multitype Moran model with a random number of types is presented. It is shown that, under weak conditions on p, the process Y, properly time and space scaled, converges to an Ornstein–Uhlenbeck process as N tends to ∞. The asymptotics of the stationary distribution of Y is studied as N tends to ∞. Examples are presented involving selection mechanisms.

Keywords: Ancestral process; complete monotonicity; descendants; duality; multitype Moran model; Ornstein–Uhlenbeck process

2000 Mathematics Subject Classification: Primary 60K35
Secondary 92D10; 92D25

1. Introduction

Assume that $X = (X_t)_{t \in T}$ and $Y = (Y_t)_{t \in T}$ are two time-homogeneous Markov processes with state spaces $(E_1, \mathscr{F}_1)$ and $(E_2, \mathscr{F}_2)$, respectively. Typical time sets are $T = \{0, \ldots, n\}$ for some $n \in \mathbb{N} := \{1, 2, \ldots\}$, countable sets $T = \mathbb{N}_0 := \{0, 1, 2, \ldots\}$, or continuous-time sets such as the unit interval $T = [0, 1]$ or $T = [0, \infty)$. Let $B(E_1 \times E_2)$ denote the set of all real-valued, bounded, measurable functions on $E_1 \times E_2$. We recall the definition of duality of Markov processes in the sense of Liggett [21].

Definition 1.1. The process X is said to be dual to Y with respect to $H \in B(E_1 \times E_2)$ if

$$\mathrm{E}^x(H(X_t, y)) = \mathrm{E}^y(H(x, Y_t))$$

for all $x \in E_1$, $y \in E_2$, and $t \in T$, where E^x denotes the expectation given that the process X starts in $X_0 = x$ and E^y denotes the expectation given that the process Y starts in $Y_0 = y$.

Dual processes occur in many applications, usually when considering some phenomena forwards and backwards in time. For typical dual processes, we refer the reader to the mathematics and physics literature on interacting particle systems; see [8], [17], [21], [22], [28],

[29], and the references therein. Other important examples occur in the context of mathematical population genetics [1], [2], [5], [6], [10]–[14], [18], [25], [26] and essentially go back to similar duality results for stochastically monotone Markov processes [27].

Here we are interested in a particular class of Markov chains X of the following form. Let $p: [0, 1] \to [0, 1]$ be some function. For fixed $N \in \mathbb{N}$, consider a Markov chain $X = (X_n)_{n \in \mathbb{N}_0}$ with state space $\{0, \ldots, N\}$ and tridiagonal transition matrix

$$\Pi = (\pi_{ij})_{i,j \in \{0,\ldots,N\}} = \begin{pmatrix} r_0 & p_0 & 0 & \cdots & & 0 \\ q_1 & r_1 & p_1 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & & 0 \\ \vdots & \ddots & & q_{N-1} & r_{N-1} & p_{N-1} \\ 0 & \cdots & & 0 & q_N & r_N \end{pmatrix}, \quad (1.1)$$

where, for $i \in \{0, \ldots, N\}$,

$$p_i := \left(1 - \frac{i}{N}\right) p\left(\frac{i}{N}\right), \qquad q_i := \frac{i}{N}\left(1 - p\left(\frac{i}{N}\right)\right), \quad (1.2)$$

and

$$r_i := 1 - p_i - q_i = \left(1 - \frac{i}{N}\right)\left(1 - p\left(\frac{i}{N}\right)\right) + \frac{i}{N} p\left(\frac{i}{N}\right). \quad (1.3)$$

For particular choices of the function p, the Markov chain X is well known from the literature on mathematical population genetics. For $p = \mathrm{Id}$ (identity), the chain X counts the number of descendants in the classical haploid Moran model [23] with population size N (see also Equations (3.45)–(3.47) of [15, p. 105]). For $p(x) = 1 - (1 - x)^2$, $x \in [0, 1]$, the chain X is the forward process of a two-sex model introduced in [19]. For closely related Moran models in continuous time, we refer the reader to [7], [9], and [12]. Note that $E(X_{n+1} \mid X_n = i) = i + p(i/N) - i/N$ for all $i \in \{0, \ldots, N\}$. Thus, if $p(x) \le x$ for all $x \in [0, 1]$ then X is a supermartingale. If $p(x) \ge x$ for all $x \in [0, 1]$ then X is a submartingale.

The paper is organised as follows. In Section 2 we provide (for a suitable subclass of functions p) a multitype (K-sex) population model such that X counts the number of 'mating units' forwards in time in this model. We also introduce the ancestral chain Y which counts the number of 'mating units' backwards in time. In Section 3, the main duality result (Theorem 3.1) and its proof are given. Section 4 deals with the convergence of Y, properly time and space scaled, to an Ornstein–Uhlenbeck limiting process as N tends to ∞ (Theorem 4.1). In Section 5 we provide details of the extinction probabilities of the forward chain X and of the stationary distribution of the ancestral chain Y. The asymptotics of the mean and the variance of the stationary distribution of Y are given. We state a conjecture on the asymptotic normality (Conjecture 5.1) of the stationary distribution. The paper concludes with a collection of typical examples (Section 6) involving selection mechanisms.

Throughout the paper, we will use the notation $(x)_0 := 1$ and $(x)_k := x(x-1)\cdots(x-k+1)$, $x \in \mathbb{R}$, $k \in \mathbb{N}$, for the descending factorials. Furthermore, $S(i, j)$, $i, j \in \mathbb{N}_0$, denote the Stirling numbers of the second kind.

2. A multitype Moran model

Before we introduce the multitype Moran model we briefly recall some basics about completely monotone functions. In the literature (see, for example, [4, Section 2]), complete

monotonicity is mostly considered for functions q with domain $(0, \infty)$ or $[0, \infty)$. Here we are interested in functions q with domain $[0, 1]$, the unit interval.

Definition 2.1. A function $q \colon [0, 1] \to \mathbb{R}$ is called completely monotone if it is C^∞ on $(0, 1)$ with $(-1)^k q^{(k)}(x) \geq 0$ for all $k \in \mathbb{N}_0$ and all $x \in (0, 1)$.

The following lemma is essentially a version of the classical Bernstein theorem.

Lemma 2.1. *A continuous function $q \colon [0, 1] \to \mathbb{R}$ is completely monotone if and only if there exists a finite measure μ on $\mathbb{N}_0$ with $q(1 - x) = \sum_{n=0}^{\infty} x^n \mu(\{n\})$ for all $x \in [0, 1]$.*

Proof. If there exists a finite measure μ on $\mathbb{N}_0$ with $q(1 - x) = \sum_{n=0}^{\infty} x^n \mu(\{n\})$ for all $x \in [0, 1]$, then $(-1)^k q^{(k)}(x) = \sum_{n=k}^{\infty} n!/(n - k)! \, (1 - x)^{n-k} \mu(\{n\}) \geq 0$ for all $k \in \mathbb{N}_0$ and all $x \in (0, 1)$, i.e. q is completely monotone. Conversely, if q is completely monotone then the function $u \colon [0, 1] \to \mathbb{R}$, defined via $u(x) := q(1 - x)$ for $x \in [0, 1]$, satisfies $u^{(n)}(x) = (-1)^n q^{(n)}(1 - x) \geq 0$ for all $n \in \mathbb{N}_0$ and all $x \in (0, 1)$. Bernstein's theorem [16, Chapter VII.2, Theorem 2] for bounded functions u ensures the existence of a measure μ on $\mathbb{N}_0$ such that u has power series representation $u(x) = \sum_{n=0}^{\infty} x^n \mu(\{n\})$, $x \in [0, 1]$. The measure μ is finite as $\mu(\mathbb{N}_0) = u(1) < \infty$. □

Remark. Suppose that $q \colon [0, 1] \to \mathbb{R}$ is a continuous, completely monotone function. The measure μ in Lemma 2.1 is a probability measure if and only if $q(0) = 1$. In this case, there exists a random variable η with distribution $\mathrm{P}(\eta = n) = \mu(\{n\})$, $n \in \mathbb{N}_0$, and $u(x) := q(1 - x) = \mathrm{E}(x^\eta)$ for all $x \in [0, 1]$, i.e. u is the probability generating function (PGF) of η.

We now turn to the definition of the multitype Moran model. Fix a constant $K \in \mathbb{N}$, and consider a population where each individual has one of K possible types. Each generation consists of $N \in \mathbb{N}$ mating units, where a mating unit is (by definition) a set of K individuals of different types. Hence, the total population size is KN. In each generation, K children are born, one of each type, and each of these K children randomly chooses one of the ancestral mating units as its parental unit. These K newly born individuals form a new mating unit of the following generation. One of the N parental mating units is chosen at random and removed from the population. For $K = 1$, this model coincides with the standard haploid Moran model with population size N (see, for example, Equations (3.45)–(3.47) of [15, p. 105]). For $K = 2$, we arrive at Kämmerle's two-sex Moran model [19]. Some more details for this particular model are given in Section 6 (Example 6.1). We may also think of this model as a K-sex model by interpreting different types as different genders.

Fix $i \in \{1, \ldots, N\}$. A descendant-unit of the mating units $1, \ldots, i$ of generation 0 is any mating unit of any generation which has at least one member descending from one of these i mating units of generation 0. If X_n denotes the number of descendant-units in generation $n \in \mathbb{N}_0$ then it is easily seen that $X := (X_n)_{n \in \mathbb{N}_0}$ is a Markov chain with state space $\{0, \ldots, N\}$ and transition matrix Π of the form (1.1) with

$$p(x) := 1 - (1 - x)^K, \qquad x \in [0, 1].$$

If we instead look back into the past and let Y_n denote the number of ancestral mating units n generations backwards in time, we obtain the so-called ancestral chain $Y := (Y_n)_{n \in \mathbb{N}_0}$, sometimes also called the backward chain. The following lemma provides the transition probabilities of the Markov chain Y.

Lemma 2.2. *The transition probabilities $p_{ij} := P(Y_{n+1} = j \mid Y_n = i)$, $i, j \in \{0, \ldots, N\}$, of the ancestral chain $Y = (Y_n)_{n \in \mathbb{N}_0}$ are*

$$p_{i,i-1} = \frac{i}{N}\left(\frac{i-1}{N}\right)^K, \qquad i \in \{1, \ldots, N\},$$

$$p_{00} = 1,$$

$$p_{ii} = \frac{N-i}{N} + \frac{i}{N}(N-i+1)\left(\left(\frac{i}{N}\right)^K - \left(\frac{i-1}{N}\right)^K\right), \qquad i \in \{1, \ldots, N\},$$

$$p_{ij} = \frac{i}{N}\binom{N-i+1}{j-i+1}\sum_{k=0}^{j-i+1}(-1)^{j-i+1-k}\binom{j-i+1}{k}\left(\frac{i-1+k}{N}\right)^K$$

for $j \in \{i+1, \ldots, i+K-1\}$, and $p_{ij} = 0$ otherwise.

Proof. Assume that there are i ancestral mating units present in generation n. One generation backwards in time there will be $i-1$ ancestral mating units present if and only if one of the i mating units is removed and all K newly born individuals choose their parental unit among the $i-1$ remaining ancestral units. One of the i mating units is removed with probability i/N and each newly born individual chooses its parental unit among the $i-1$ remaining units with probability $(i-1)/N$. Thus, the formula for $p_{i,i-1}$ follows immediately.

One generation backwards in time there will be i ancestral mating units present if and only if either one of the $N-i$ nonancestral units is removed or if one of the i ancestral units is removed and, for some $l \in \{1, \ldots, K\}$, exactly $K-l$ of the K newly born individuals choose their ancestral unit among the remaining $i-1$ ancestral units and the other l newly born individuals all choose the same unit among the other $N-(i-1)$ mating units. Thus,

$$p_{ii} = \frac{N-i}{N} + \frac{i}{N}\sum_{l=1}^K \binom{K}{l}\frac{N-(i-1)}{N^l}\left(\frac{i-1}{N}\right)^{K-l}$$

$$= \frac{N-i}{N} + \frac{i}{N}(N-i+1)\sum_{l=1}^K \binom{K}{l}\left(\frac{1}{N}\right)^l\left(\frac{i-1}{N}\right)^{K-l}$$

$$= \frac{N-i}{N} + \frac{i}{N}(N-i+1)\left(\left(\frac{i}{N}\right)^K - \left(\frac{i-1}{N}\right)^K\right).$$

Similarly, there will be j ($> i$) ancestral mating units present one generation backwards in time if and only if one of the i ancestral mating units is removed and, for some $l \in \{0, \ldots, K\}$, exactly $K-l$ of the newly born individuals choose their ancestral unit among the remaining $i-1$ ancestral units and the other l newly born individuals altogether choose $j-i+1$ among the other $N-(i-1)$ mating units. Thus,

$$p_{ij} = \frac{i}{N}\sum_{l=0}^K \binom{K}{l}\frac{(N-(i-1))_{j-i+1}S(l, j-i+1)}{N^l}\left(\frac{i-1}{N}\right)^{K-l}.$$

Note that $p_{ij} = 0$ for $j > i+K-1$. Substituting

$$S(l, j-i+1) = \frac{1}{(j-i+1)!}\sum_{k=0}^{j-i+1}(-1)^{j-i+1-k}\binom{j-i+1}{k}k^l,$$

the explicit formula for the Stirling numbers of the second kind, into the above equation and interchanging the sums yields

$$p_{ij} = \frac{i}{N}\binom{N-i+1}{j-i+1}\sum_{k=0}^{j-i+1}(-1)^{j-i+1-k}\binom{j-i+1}{k}\sum_{l=0}^{K}\binom{K}{l}\left(\frac{k}{N}\right)^{l}\left(\frac{i-1}{N}\right)^{K-l}$$

$$= \frac{i}{N}\binom{N-i+1}{j-i+1}\sum_{k=0}^{j-i+1}(-1)^{j-i+1-k}\binom{j-i+1}{k}\left(\frac{i-1+k}{N}\right)^{K},$$

which completes the proof.

We can also include the case in which $K = 0$ ($p(x) \equiv 0$) in our model with the interpretation of a population consisting of N individuals of the same type in each generation. In each step (forwards in time), one randomly chosen individual is (by definition) replaced by a nondescendant individual. In this case we have $p_{i,i-1} = i/N$ and $p_{ii} = 1 - i/N$, in agreement with Lemma 2.2 for $K = 0$.

The model can be easily extended to a larger class of functions p as follows. Suppose that $p: [0, 1] \to [0, 1]$ is any continuous function satisfying $p(0) = 0$ and that $q := 1 - p$ is completely monotone. Then (see the remark after Lemma 2.1) there exists a random variable η taking values in $\mathbb{N}_0$ such that $p(x) = 1 - E((1-x)^{\eta})$ for all $x \in [0, 1]$. In this situation, the Markov chains X and Y can be interpreted probabilistically in the following way. Consider a population where each individual has one of $\max(\eta, 1)$ possible types. The population consists in each generation of N mating units, where a mating unit is (by definition) a set of $\max(\eta, 1)$ individuals of different types. Assume that, conditioned on $\eta = K$ ($K \in \mathbb{N}_0$), the population evolves according to the K-type Moran model introduced before. Then, X counts the number of mating units forwards in time and Y counts the number of mating units backwards in time. We therefore have found—for continuous functions $p: [0, 1] \to [0, 1]$ satisfying $p(0) = 0$ and such that $q := 1 - p$ is completely monotone—a meaningful probabilistic interpretation for the chains X and Y. In particular, the random variable Y_n can be interpreted as the number of ancestral mating units n generations backwards in time under the multitype Moran model with a random number, η, of types. The following corollary is a direct consequence of Lemma 2.2.

Corollary 2.1. *For the mixed multitype Moran model with a random number, η, of types, the transition probabilities $p_{ij} := P(Y_{n+1} = j \mid Y_n = i)$, $i, j \in \{0, \ldots, N\}$, of the ancestral chain $Y = (Y_n)_{n \in \mathbb{N}_0}$ are*

$$p_{i,i-1} = \frac{i}{N}\left(1 - p\left(\frac{N-i+1}{N}\right)\right), \qquad i \in \{1, \ldots, N\},$$

$$p_{00} = 1,$$

$$p_{ii} = \frac{N-i}{N} + \frac{i}{N}(N-i+1)\left(p\left(\frac{N-i+1}{N}\right) - p\left(\frac{N-i}{N}\right)\right), \qquad i \in \{1, \ldots, N\},$$

$$p_{ij} = \frac{i}{N}\binom{N-i+1}{j-i+1}\sum_{k=0}^{j-i+1}(-1)^{j-i-k}\binom{j-i+1}{k}p\left(\frac{N-i+1-k}{N}\right)$$

for $j \in \{i+1, i+2, \ldots, N\}$, and $p_{ii} = 0$ otherwise.

Proof. From Lemma 2.2 and the definition of the mixed multitype Moran model, it follows that, for $i \in \{1, \ldots, N\}$,

$$p_{i,i-1} = \sum_{K=0}^{\infty} \frac{i}{N} \left(\frac{i-1}{N} \right)^K P(\eta = K) = \frac{i}{N} E\left(\left(\frac{i-1}{N} \right)^{\eta} \right) = \frac{i}{N} \left(1 - p\left(\frac{N-i+1}{N} \right) \right).$$

The other transition probabilities p_{ij} are obtained similarly.

Note that Y is skip-free to the left, i.e. $p_{ij} = 0$ for $j < i - 1$. In the following section the main duality result (Theorem 3.1) is presented. An important consequence of this result is that, for the mixed multitype Moran model with a random number, η, of types, X is dual to Y with respect to a specific duality function H (see (3.1), below) which is universal in the sense that it does not depend on the function p.

3. A duality result

If the state spaces E_1 and E_2 of the processes X and Y in Definition 1.1 are finite, then the function $H \in B(E_1 \times E_2)$ in Definition 1.1 can be considered as a matrix $H = (h_{ij})_{i \in E_1, j \in E_2}$. In the literature on duality for Markov chains with identical finite state spaces $E_1 = E_2 = \{0, \ldots, N\}$, the particular nonsingular, symmetric, left-upper matrix $H = (h_{ij})_{i,j \in \{0,\ldots,N\}}$ with entries

$$h_{ij} := \binom{N-i}{j} \Big/ \binom{N}{j} = \binom{N-j}{i} \Big/ \binom{N}{i} = \prod_{k=0}^{i-1} \frac{N-j-k}{N-k} = \prod_{k=0}^{j-1} \frac{N-i-k}{N-k} \quad (3.1)$$

plays an important role. For example, this matrix turns out to be a suitable choice to obtain duality results for a large class of exchangeable population models [25]. It was also used (see also the remark at the end of this section) to derive duality results for a class of nonneutral Wright–Fisher models [18]. The following theorem shows that this matrix is also an appropriate choice for Moran models.

Theorem 3.1. *Let $p \colon [0, 1] \to [0, 1]$ be some continuous function satisfying $p(x) \leq p(1)$ for all $x \in [0, 1]$. Then, the following conditions are equivalent.*

(i) *There exists a random variable η taking values in $\mathbb{N}_0$ such that $1 - p(1 - x) = E(x^{\eta})$ for all $x \in [0, 1]$.*

(ii) *The function $q := 1 - p$ is completely monotone and p satisfies $p(0) = 0$.*

(iii) *For each $N \in \mathbb{N}$, there exists a Markov chain $Y = (Y_n)_{n \in \mathbb{N}_0}$ such that the Markov chain X (as defined in Section 1) is dual to Y with respect to $H = (h_{ij})_{i,j \in \{0,\ldots,N\}}$ with entries (3.1), i.e. $\Pi H = H P^{\top}$, where $P^{\top}$ denotes the transpose of the transition matrix $P = (p_{ij})_{i,j \in \{0,\ldots,N\}}$ of Y.*

Remark. It turns out (see (3.3), (3.4), (3.5), and (3.7), below) that the distribution of the chain Y coincides with that of the ancestral chain of the mixed multitype Moran model introduced at the end of Section 2 (see Corollary 2.1). However, note that we do not make use of the multitype Moran model, neither in the presentation of Theorem 3.1 nor in the following proof.

Proof of Theorem 3.1. The equivalence of (i) and (ii) follows immediately from the remark after Lemma 2.1.

We now verify that condition (ii) implies condition (iii). For $i, j \in \{0, \ldots, N\}$,

$$(\Pi H)_{ij} = q_i h_{i-1,j} + r_i h_{ij} + p_i h_{i+1,j}$$
$$= q_i \binom{N-i+1}{j} \bigg/ \binom{N}{j} + r_i \binom{N-i}{j} \bigg/ \binom{N}{j} + p_i \binom{N-i-1}{j} \bigg/ \binom{N}{j},$$

where p_i, q_i, and r_i are as defined in (1.2) and (1.3). Define the matrix $P := (p_{ij})_{i,j \in \{0,\ldots,N\}}$ recursively over its columns via

$$p_{i0} := \frac{(\Pi H)_{Ni}}{h_{N0}} = \begin{cases} 1 & \text{for } i = 0, \\ \dfrac{1 - p(1)}{N} & \text{for } i = 1, \\ 0 & \text{for } i \in \{2, \ldots, N\}, \end{cases}$$

and

$$p_{ij} := \frac{(\Pi H)_{N-j,i} - \sum_{k=0}^{j-1} h_{N-j,k} p_{ik}}{h_{N-j,j}}, \qquad i \in \{0, \ldots, N\}, \tag{3.2}$$

for $j \in \{1, \ldots, N\}$. From this definition, it follows immediately that

$$(\Pi H)_{N-j,i} = \sum_{k=0}^{j} h_{N-j,k} p_{ik} = \sum_{k=0}^{N} h_{N-j,k} p_{ik} = (HP^\top)_{N-j,i}$$

for all $i, j \in \{0, \ldots, N\}$. Thus, the duality equation $\Pi H = HP^\top$ is satisfied. It remains to verify that P is a stochastic matrix. Note that $p_{00} = 1$ and that $p_{ij} = 0$ for $j < i - 1$. Thus, the matrix P has the form

$$P = (p_{ij})_{i,j \in \{0,\ldots,N\}} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ \star & \star & \cdots & \star & \star \\ 0 & \star & \cdots & \star & \star \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \star & \star \end{pmatrix}, \tag{3.3}$$

where the $\star$ entries are those which are now studied in more detail. From (3.2), it follows that, for $i \in \{1, \ldots, N\}$,

$$p_{i,i-1} = \frac{(\Pi H)_{N-i+1,i}}{h_{N-i+1,i-1}}$$
$$= \frac{q_{N-i+1} h_{N-i,i}}{h_{N-i+1,i-1}}$$
$$= \frac{N-i+1}{N} \left(1 - p\left(\frac{N-i+1}{N}\right)\right) \binom{N}{i}^{-1} \binom{N}{i-1}$$
$$= \frac{i}{N} \left(1 - p\left(\frac{N-i+1}{N}\right)\right). \tag{3.4}$$

In particular, $p_{i,i-1} \geq 0$. For the diagonal entries of P, we obtain from (3.2) that $p_{00} = 1$ and that, for $i \in \{1, \ldots, N\}$,

$$p_{ii} = \frac{(\Pi H)_{N-i,i} - h_{N-i,i-1} p_{i,i-1}}{h_{N-i,i}}$$
$$= \frac{q_{N-i} h_{N-i-1,i} + r_{N-i} h_{N-i,i} + p_{N-i} h_{N-i+1,i} - h_{N-i,i-1} p_{i,i-1}}{h_{N-i,i}}$$

$$= (i+1)q_{N-i} + r_{N-i} - (N-i+1)p_{i,i-1}$$

$$= (i+1)\frac{N-i}{N}\left(1 - p\left(\frac{N-i}{N}\right)\right) + \frac{i}{N} + \frac{N-2i}{N}p\left(\frac{N-i}{N}\right)$$

$$- (N-i+1)\frac{i}{N}\left(1 - p\left(\frac{N-i+1}{N}\right)\right)$$

$$= \frac{N-i}{N} + \frac{i}{N}(N-i+1)\left(p\left(\frac{N-i+1}{N}\right) - p\left(\frac{N-i}{N}\right)\right). \tag{3.5}$$

In particular, $p_{ii} \geq 0$ as the function p is increasing by condition (ii). Proceeding in the same way we obtain

$$p_{i,i+1} = \frac{(\Pi H)_{N-i-1,i} - \sum_{k=i-1}^{i} h_{N-i-1,k}p_{ik}}{h_{N-i-1,i+1}}$$

$$= \frac{(\Pi H)_{N-i-1,i} - h_{N-i-1,i-1}p_{i,i-1} - h_{N-i-1,i}p_{ii}}{h_{N-i-1,i+1}}$$

$$= (N-i)\left(\frac{i+2}{2}q_{N-i-1} + r_{N-i-1} + \frac{1}{i+1}p_{N-i-1}\right)$$

$$- \binom{N-i+1}{2}p_{i,i-1} - \binom{N-i}{1}p_{ii}.$$

Substituting the expressions for p_{N-i-1}, q_{N-i-1}, and r_{N-i-1} into this as well as (3.4) and (3.5) for $p_{i,i-1}$ and p_{ii}, and sorting afterwards with respect to the values $p((N-i-1)/N)$, $p((N-i)/N)$, and $p((N-i+1)/N)$, we obtain, after some straightforward manipulation, for $i \in \{1, \dots, N-1\}$,

$$p_{i,i+1} = \frac{i}{N}\binom{N-i+1}{2}\left(2p\left(\frac{N-i}{N}\right) - p\left(\frac{N-i-1}{N}\right) - p\left(\frac{N-i+1}{N}\right)\right). \tag{3.6}$$

Since q is completely monotone, second-order differences of p are nonnegative from which it follows that $p_{i,i+1} \geq 0$. We now verify by induction on $j \in \{i+1, i+2, \dots, N\}$ that

$$p_{ij} = \frac{i}{N}\binom{N-i+1}{j-i+1}\sum_{k=0}^{j-i+1}\binom{j-i+1}{k}(-1)^{j-i-k}p\left(\frac{N-i+1-k}{N}\right). \tag{3.7}$$

For $j = i+1$, (3.7) coincides with (3.6). Thus, (3.7) is already proven for $j = i+1$. The induction step from $\{1, \dots, j-1\}$ to $j (> i)$ works as follows. For $j > i$, we have

$$p_{ij} = \frac{(\Pi H)_{N-j,i} - \sum_{k=i-1}^{j-1} h_{N-j,k}p_{ik}}{h_{N-j,j}}$$

$$= \frac{q_{N-j}h_{N-j-1,i} + r_{N-j}h_{N-j,i} + p_{N-j}h_{N-j+1,i}}{h_{N-j,j}} - \sum_{k=i-1}^{j-1}\frac{h_{N-j,k}}{h_{N-j,j}}p_{ik}$$

$$= \binom{N-i}{N-j}\left(q_{N-j}\frac{j+1}{j-i+1} + r_{N-j} + p_{N-j}\frac{j-i}{j}\right) - \sum_{k=i-1}^{j-1}\binom{N-k}{j-k}p_{ik}.$$

As before, substituting the expressions for p_{N-j}, q_{N-j}, and r_{N-j} into this as well as the (by induction already known) formulae for p_{ik}, $k \in \{i-1, i, \dots, j-1\}$, we find—after some

tedious but straightforward manipulations—that the latter expression coincides with the right-hand side of (3.7), which completes the induction.

Since q is completely monotone, the sum on the right-hand side of (3.7) is nonnegative, which shows that p_{ij} is nonnegative for $j > i$. For $i \in \{1, \ldots, N\}$, it follows furthermore from (3.4), (3.5), and (3.7) that, for $x \in [0, 1]$,

$$
\begin{aligned}
\sum_{j=0}^{N} p_{ij} x^j &= p_{i,i-1} x^{i-1} + p_{ii} x^i + \sum_{j=i+1}^{N} p_{ij} x^j \\
&= \frac{i x^{i-1}}{N}\left(1 - p\left(\frac{N-i+1}{N}\right)\right) \\
&\quad + \frac{N-i}{N} x^i + \frac{i x^i}{N}(N-i+1)\left(p\left(\frac{N-i+1}{N}\right) - p\left(\frac{N-i}{N}\right)\right) \\
&\quad + \sum_{j=i+1}^{N} \frac{i x^j}{N}\binom{N-i+1}{j-i+1}\sum_{k=0}^{j-i+1}\binom{j-i+1}{k}(-1)^{j-i-k}p\left(\frac{N-i+1-k}{N}\right) \\
&= \frac{i}{N} x^{i-1} + \frac{N-i}{N} x^i \\
&\quad + \frac{i}{N}\sum_{j=i-1}^{N} x^j\binom{N-i+1}{j-i+1}\sum_{k=0}^{j-i+1}\binom{j-i+1}{k}(-1)^{j-i-k}p\left(\frac{N-i+1-k}{N}\right) \\
&= \frac{i}{N} x^{i-1} + \frac{N-i}{N} x^i - \frac{i}{N}\sum_{k=0}^{N-i+1}\binom{N-i+1}{k}p\left(\frac{N-i+1-k}{N}\right)x^{i+k-1} \\
&\qquad\qquad\qquad\qquad\qquad \times \sum_{j=k+i-1}^{N}\binom{N-i+1-k}{j-i+1-k}(-x)^{j-i+1-k} \\
&= \frac{i}{N} x^{i-1} + \frac{N-i}{N} x^i \\
&\quad - \frac{i}{N}\sum_{k=0}^{N-i+1}\binom{N-i+1}{k}p\left(\frac{N-i+1-k}{N}\right)x^{i+k-1}(1-x)^{N-i+1-k} \\
&= \frac{i}{N} x^{i-1} + \frac{N-i}{N} x^i - \frac{i}{N}\sum_{l=0}^{N-i+1}\binom{N-i+1}{l}p\left(\frac{l}{N}\right)x^{N-l}(1-x)^l \\
&= \frac{N-i}{N} x^i + \frac{i}{N}\sum_{l=0}^{N-i+1}\binom{N-i+1}{l}q\left(\frac{l}{N}\right)x^{N-l}(1-x)^l. \qquad (3.8)
\end{aligned}
$$

For $x = 1$, only the summand for $l = 0$ differs from 0, and we obtain $\sum_{j=0}^{N} p_{ij} = (N-i)/N + i/Nq(0) = 1 - i/Np(0) = 1$, as $p(0) = 0$ by condition (ii). Thus, P is a stochastic matrix.

Conversely, let us now verify that condition (iii) implies condition (ii). If X is dual to Y with respect to H then recursion (3.2) holds. As in the previous part of the proof, it follows that the matrix P has the structure (3.3) with the $\star$ entries given by (3.4), (3.5), and (3.7). Consequently,

(3.8) is valid, and we obtain, for $x \in [0, 1]$,

$$
\begin{aligned}
\mathrm{E}(x^{Y_{n+1}} \mid Y_n = i) &= \sum_{j=0}^{N} p_{ij} x^j \\
&= \frac{N-i}{N} x^i + \frac{i}{N} x^{i-1} \sum_{l=0}^{N-i+1} \binom{N-i+1}{l} q\left(\frac{l}{N}\right) x^{N-i+1-l} (1-x)^l \\
&= \frac{N-i}{N} x^i + \frac{i}{N} x^{i-1} B_{N-i+1,q}(1-x),
\end{aligned}
$$

where $B_{N-i+1,q}(1-x)$ is the $(N-i+1)$th Bernstein polynomial of the function q evaluated at the point $1-x$. Now rewrite this equation for $i = 1$ as

$$
B_{N,q}(1-x) = x + N(\mathrm{E}(x^{Y_{n+1}} \mid Y_n = 1) - x). \tag{3.9}
$$

As q is continuous, the left-hand side converges uniformly for $x \in [0, 1]$ to $q(1-x)$ as $N \to \infty$. Therefore, the right-hand side of (3.9) also converges uniformly for $x \in [0, 1]$, and from the structure of this right-hand side, it follows that

$$
q(1-x) = \sum_{k=0}^{\infty} a_k x^k \tag{3.10}
$$

for some nonnegative real coefficients a_k, $k \in \mathbb{N}_0$. Note that, in particular, the coefficient a_1 is nonnegative, as it is the limit as $N \to \infty$ of

$$
1 + N(p_{11} - 1) = 1 + N\left(\frac{N-1}{N} + p(1) - p\left(\frac{N-1}{N}\right) - 1\right) = N\left(p(1) - p\left(\frac{N-1}{N}\right)\right),
$$

which is nonnegative as p satisfies $p(x) \le p(1)$ for all $x \in [0, 1]$. Thus, q is completely monotone. Since P is a stochastic matrix we have $\sum_{j=0}^{N} p_{ij} = 1$ for all $i \in \{0, \dots, N\}$. On the other hand, as in the previous part of the proof, it follows that $\sum_{j=0}^{N} p_{ij} = 1 - (i/N)p(0)$ for all $i \in \{0, \dots, N\}$. Thus, the equality $p(0) = 0$ must hold, which completes the proof.

Remark. In Theorem 3.1 it is assumed that the function $p \colon [0, 1] \to [0, 1]$ satisfies

$$
p(x) \le p(1) \quad \text{for all } x \in [0, 1]. \tag{3.11}
$$

This condition is automatically satisfied if $p(1) = 1$. If (3.11) is not satisfied then the situation becomes more subtle. Consider, for example, the function $p(x) := x(1-x)$, $x \in [0, 1]$. In this case, 0 is the only absorbing state of the chain X and all other states $1, \dots, N$ are communicating. The matrix P in (3.3), with entries defined via (3.4), (3.5), and (3.7), is a stochastic matrix. More precisely, P is tridiagonal with entries

$$
p_{ij} = \begin{cases}
\dfrac{i}{N} - \dfrac{i(i-1)(N-i+1)}{N^3} & \text{for } j = i-1, \\[2mm]
\dfrac{N-i}{N} + \dfrac{i(N-i+1)(N-2i+1)}{N^3} & \text{for } j = i, \\[2mm]
\dfrac{i(N-i)(N-i+1)}{N^3} & \text{for } j = i+1, \\[2mm]
0 & \text{for } |i-j| > 1,
\end{cases} \qquad i, j \in \{0, \dots, N\}.
$$

Therefore, if Y is a Markov chain having this transition matrix P then Y is dual to X with respect to H. However, $1 - p$ is not completely monotone (even not monotone). The stationary distribution of Y is concentrated in 0.

In general, it can be deduced from (3.5) that certain negative values of the coefficient a_1 in (3.10) are allowed without destroying the property that P is a stochastic matrix. For simplicity, in Theorem 3.1 we restrict our considerations to the situation when p satisfies (3.11).

Remark. In [18] a class of Markov chains $X = (X_n)_{n \in \mathbb{N}_0}$ with state space $\{0, \ldots, N\}$ and modified Wright–Fisher transition probabilities $\pi_{ij} = \binom{N}{j}(p(i/N))^j (1 - p(i/N))^{N-j}$, $i, j \in \{1, \ldots, N\}$, was studied. It was shown that, if $p \colon [0, 1] \to [0, 1]$ is continuous then, for each $N \in \mathbb{N}$, there exists a Markov chain $Y = (Y_n)_{n \in \mathbb{N}_0}$ such that X is dual to Y with respect to $H = (h_{ij})_{i, j \in \{0, \ldots, N\}}$ with entries (3.1) if and only if the function $q := 1 - p$ is completely monotone and p satisfies $p(0) = 0$. This duality result essentially coincides with that of Theorem 3.1 and was the starting point to study similar properties for other models, which finally led us to Theorem 3.1. The transition matrix $P = (p_{ij})_{i, j \in \{0, \ldots, N\}}$ of the dual chain Y in [18] has entries

$$p_{ij} = \binom{N}{j} \sum_{k=0}^{j} (-1)^{j-k} \binom{j}{k} q^i \left(1 - \frac{k}{N}\right), \qquad i, j \in \{0, \ldots, N\}, \qquad (3.12)$$

where q^i denotes the ith power of q. There is the following alternative formula for these transition probabilities. The remark after Lemma 2.1 ensures that there exists a random variable η taking values in $\mathbb{N}_0$ such that $q(1-x) = \mathrm{E}(x^\eta)$ for all $x \in [0, 1]$. Let $\eta_1, \eta_2, \ldots$ be independent copies of η. Then,

$$p_{ij} = (N)_j \, \mathrm{E}(N^{-L_i} S(L_i, j)), \qquad i, j \in \{0, \ldots, N\}, \qquad (3.13)$$

where $L_0 := 0$ and $L_i := \eta_1 + \cdots + \eta_i$ for $i \in \mathbb{N}$, using the notation for the descending factorials and for the Stirling numbers of the second kind mentioned at the end of the introduction. Formula (3.13) follows easily from (3.12) using the convolution property $\mathrm{E}(x^{L_i}) = q^i(1-x)$ (applied to $x := k/N$) and the explicit formula $j! \, S(i, j) = \sum_{k=0}^{j} (-1)^{j-k} \binom{j}{k} k^i$ for the Stirling numbers of the second kind. From (3.13), it is obvious that P is a stochastic matrix. In this sense the Markov chain Y in [18] has a simpler structure than the chain Y in Theorem 3.1.

4. A weak convergence result for the ancestral process

In this section it is always assumed that the continuous function $p \colon [0, 1] \to [0, 1]$ satisfies $p(0) = 0$ and that $q := 1 - p$ is completely monotone such that the existence of the ancestral process $Y = (Y_n)_{n \in \mathbb{N}_0}$ is guaranteed by Theorem 3.1. We are interested in the asymptotic behaviour of Y as N tends to ∞. Before the main convergence result (Theorem 4.1) is presented, the moments of Y_{n+1}, conditioned on $Y_n = i$, are studied in more detail.

Lemma 4.1. *For all $i \in \{0, \ldots, N\}$,*

$$\mathrm{E}(Y_{n+1} - i \mid Y_n = i) = \frac{i(N - i + 1)}{N} p\left(\frac{1}{N}\right) - \frac{i}{N} \qquad (4.1)$$

and

$$\mathrm{E}((Y_{n+1} - i)^2 \mid Y_n = i) = \frac{i}{N} + \frac{i(N - i + 1)(2N - 2i - 1)}{N} p\left(\frac{1}{N}\right)$$

$$- \frac{i(N - i + 1)(N - i)}{N} p\left(\frac{2}{N}\right). \tag{4.2}$$

Proof. Taking the rth derivative ($r \in \mathbb{N}_0$) with respect to x in (3.8) yields (Leibniz's rule)

$$\left(\frac{\partial}{\partial x}\right)^r \mathrm{E}(x^{Y_{n+1}} \mid Y_n = i)$$

$$= \frac{N - i}{N} (i)_r x^{i-r} + \frac{i}{N} \sum_{l=0}^{N-i+1} \binom{N - i + 1}{l} q\left(\frac{l}{N}\right)$$

$$\times \sum_{k=0}^{r} \binom{r}{k} (N - l)_k x^{N-l-k} (l)_{r-k} (1 - x)^{l-r+k} (-1)^{r-k}.$$

For $x = 1$, only the $k = r - l$ index contributes to the last sum. Thus, Y_{n+1}, conditioned on $Y_n = i$, has descending factorial moments

$$\mathrm{E}((Y_{n+1})_r \mid Y_n = i) = \frac{N - i}{N} (i)_r + \frac{i}{N} \sum_{l=0}^{N-i+1} \binom{N - i + 1}{l} q\left(\frac{l}{N}\right) \binom{r}{l} (N - l)_{r-l} l! \, (-1)^l.$$

For $r = 1$,

$$\mathrm{E}(Y_{n+1} \mid Y_n = i) = \frac{i(N - i)}{N} + \frac{i}{N}\left(Nq(0) - (N - i + 1)q\left(\frac{1}{N}\right)\right)$$

$$= \frac{i}{N}\left(N - 1 + (N - i + 1)p\left(\frac{1}{N}\right)\right),$$

as $q(0) = 1$, and (4.1) follows immediately. For $r = 2$,

$$\mathrm{E}((Y_{n+1})_2 \mid Y_n = i)$$

$$= \frac{(N - i)(i)_2}{N}$$

$$+ \frac{i}{N}\left((N)_2 q(0) - 2(N - i + 1)(N - 1)q\left(\frac{1}{N}\right) + (N - i + 1)_2 q\left(\frac{2}{N}\right)\right)$$

$$= \frac{i}{N}\left((i - 1)(N - 2) + 2(N - i + 1)(N - 1)p\left(\frac{1}{N}\right) - (N - i + 1)_2 p\left(\frac{2}{N}\right)\right).$$

Therefore,

$$\mathrm{E}((Y_{n+1} - i)^2 \mid Y_n = i)$$

$$= \mathrm{E}(Y_{n+1}^2 \mid Y_n = i) - 2i \, \mathrm{E}(Y_{n+1} \mid Y_n = i) + i^2$$

$$= \mathrm{E}((Y_{n+1})_2 \mid Y_n = i) + (1 - 2i) \, \mathrm{E}(Y_{n+1} \mid Y_n = i) + i^2$$

$$= \frac{i}{N}\left((i-1)(N-2) + 2(N-i+1)(N-1)p\left(\frac{1}{N}\right) - (N-i+1)_2 p\left(\frac{2}{N}\right)\right)$$
$$+ (1-2i)\frac{i}{N}\left(N-1 + (N-i+1)p\left(\frac{1}{N}\right)\right) + i^2$$
$$= \frac{i}{N}\left(1 + (N-i+1)(2N-2i-1)p\left(\frac{1}{N}\right) - (N-i+1)_2 p\left(\frac{2}{N}\right)\right),$$

which completes the proof of the lemma.

We are able to verify a weak convergence result for the ancestral process Y under the additional condition that the random variable η in Theorem 3.1 satisfies $E(\eta) > 1$ and $E(\eta^3) < \infty$. In terms of the function p, it is therefore assumed that $p'(0+) > 1$ and that $p'''(0+) < \infty$.

For $N \in \mathbb{N}$ and $n \in \mathbb{N}_0$, define $\hat{Y}_n := (Y_n - \alpha N)/\sqrt{N}$, where $\alpha := 1 - 1/p'(0+) \in (0, 1)$. Consider the space- and time-scaled process $(V_t^{(N)})_{t \geq 0}$, defined via

$$V_t^{(N)} := \hat{Y}_{\lfloor Nt \rfloor} = \frac{Y_{\lfloor Nt \rfloor} - \alpha N}{\sqrt{N}}, \qquad N \in \mathbb{N}, \ t \geq 0, \tag{4.3}$$

where, for $x \in \mathbb{R}$, $\lfloor x \rfloor := \max\{z \in \mathbb{Z} \mid z \leq x\}$ denotes the largest integer being less than or equal to x.

Theorem 4.1. *Suppose (see Theorem 3.1) that the continuous function $p \colon [0, 1] \to [0, 1]$ satisfies $p(0) = 0$ and that $q := 1 - p$ is completely monotone. Furthermore, suppose that $p'(0+) > 1$ and that $p'''(0+) < \infty$. If the sequence $(V_0^{(N)})_{N \in \mathbb{N}}$ converges in distribution to a random variable V_0, then the time- and space-scaled backward process $(V_t^{(N)})_{t \geq 0}$, defined via (4.3), converges weakly as $N \to \infty$ to an Ornstein–Uhlenbeck process $(V_t)_{t \geq 0}$ with drift parameter $\mu(x) = -\mu x$ and diffusion parameter $\sigma^2(x) = \sigma^2$, where*

$$\mu := \frac{\alpha}{1-\alpha} = p'(0+) - 1 \in (0, \infty)$$

and

$$\sigma^2 := -\alpha(1-\alpha)^2 p''(0+) = \frac{(1 - p'(0+))p''(0+)}{(p'(0+))^3} \in (0, \infty).$$

Remark. In terms of η the parameters α, μ, and σ^2 are given by

$$\alpha = 1 - \frac{1}{E(\eta)}, \qquad \mu = E(\eta) - 1, \quad \text{and} \quad \sigma^2 = \frac{(E(\eta) - 1)\,E(\eta(\eta - 1))}{(E(\eta))^3}.$$

Proof of Theorem 4.1. We essentially generalize Kämmerle's proof on p. 883 of [19]. The process $(\hat{Y}_n)_{n \in \mathbb{N}_0}$ is a Markov chain with state space $E_N := \{(i - \alpha N)/\sqrt{N} \mid i \in \{0, \ldots, N\}\}$. For $x \in E_N$, let $i := \alpha N + x\sqrt{N} \in \{0, \ldots, N\}$ and note that

$$E(\hat{Y}_{n+1} - x \mid \hat{Y}_n = x) = E\left(\frac{Y_{n+1} - \alpha N - x\sqrt{N}}{\sqrt{N}} \;\middle|\; \frac{Y_n - \alpha N}{\sqrt{N}} = x\right)$$
$$= \frac{1}{\sqrt{N}}\,E(Y_{n+1} - i \mid Y_n = i).$$

From (4.1) and the expansion $p(1/N) = p'(0+)/N + O(1/N^2)$, it follows that

$$
\begin{aligned}
E(Y_{n+1} - i \mid Y_n = i) &= \frac{i(N-i+1)}{N} p\left(\frac{1}{N}\right) - \frac{i}{N} \\
&= \frac{i}{N} \frac{N-i+1}{N} p'(0+) - \frac{i}{N} + O\left(\frac{1}{N}\right) \\
&= \frac{i}{N} \frac{N-i}{N} p'(0+) - \frac{i}{N} + O\left(\frac{1}{N}\right) \\
&= \left(\alpha + \frac{x}{\sqrt{N}}\right)\left(1 - \alpha - \frac{x}{\sqrt{N}}\right) \frac{1}{1-\alpha} - \left(\alpha + \frac{x}{\sqrt{N}}\right) + O\left(\frac{1}{N}\right) \\
&= -\frac{\alpha}{1-\alpha} \frac{x}{\sqrt{N}} - \frac{x^2}{1-\alpha} \frac{1}{N} + O\left(\frac{1}{N}\right),
\end{aligned}
$$

where the $O(1/N)$ term holds uniformly for all $x \in E_N$. Now suppose, in addition, that $x \in K$ for some arbitrary but fixed compact set $K \subset \mathbb{R}$. Then,

$$
E(Y_{n+1} - i \mid Y_n = i) = -\frac{\alpha}{1-\alpha} \frac{x}{\sqrt{N}} + O\left(\frac{1}{N}\right) = -\mu \frac{x}{\sqrt{N}} + O\left(\frac{1}{N}\right)
$$

uniformly for all $x \in E_N \cap K$. Thus, it is shown that

$$
E(\hat{Y}_{n+1} \mid \hat{Y}_n = x) = -\mu \frac{x}{N} + O\left(\frac{1}{N^{3/2}}\right) \tag{4.4}
$$

uniformly for all $x \in E_N \cap K$. Analogously,

$$
\begin{aligned}
E((\hat{Y}_{n+1} - x)^2 \mid \hat{Y}_n = x) &= E\left(\left(\frac{Y_{n+1} - \alpha N - x\sqrt{N}}{\sqrt{N}}\right)^2 \;\middle|\; \frac{Y_n - \alpha N}{\sqrt{N}} = x\right) \\
&= \frac{1}{N} E((Y_{n+1} - i)^2 \mid Y_n = i).
\end{aligned}
$$

From (4.2) and the expansions $p(1/N) = p'(0+)/N + p''(0+)/(2N^2) + O(1/N^3)$ and $p(2/N) = 2p'(0+)/N + 2p''(0+)/N^2 + O(1/N^3)$, it follows that

$$
\begin{aligned}
&E((Y_{n+1} - i)^2 \mid Y_n = i) \\
&= \frac{i}{N} + \frac{i(N-i+1)(2N-2i-1)}{N} p\left(\frac{1}{N}\right) - \frac{i(N-i+1)(N-i)}{N} p\left(\frac{2}{N}\right) \\
&= \frac{i}{N} + \frac{i(N-i+1)(2N-2i-1)}{N}\left(\frac{p'(0+)}{N} + \frac{p''(0+)}{2N^2}\right) \\
&\quad - \frac{i(N-i+1)(N-i)}{N}\left(\frac{2p'(0+)}{N} + \frac{2p''(0+)}{N^2}\right) + O\left(\frac{1}{N}\right) \\
&= \frac{i}{N} - \frac{i(N-i+1)}{N^2} p'(0+) - \frac{i(N-i+1)(N-i+1/2)}{N^3} p''(0+) + O\left(\frac{1}{N}\right) \\
&= \underbrace{\alpha - \alpha(1-\alpha)p'(0+)}_{=0} - \alpha(1-\alpha)^2 p''(0+) + O\left(\frac{1}{\sqrt{N}}\right) \\
&= \sigma^2 + O\left(\frac{1}{\sqrt{N}}\right).
\end{aligned}
$$

uniformly for all $x \in E_N \cap K$. Thus,

$$\mathrm{E}((\hat{Y}_{n+1} - i)^2 \mid \hat{Y}_n = x) = \frac{\sigma^2}{N} + O\left(\frac{1}{N^{3/2}}\right) \tag{4.5}$$

uniformly for all $x \in E_N \cap K$ for any arbitrary but fixed compact set $K \subset \mathbb{R}$.

Let $C_c^\infty(\mathbb{R})$ denote the set of all functions $f \in C^\infty(\mathbb{R})$ with compact support. For $x \in E_N$ and $f \in C_c^\infty(\mathbb{R})$, let $T_N f(x) := \mathrm{E}(f(\hat{Y}_{n+1}) \mid \hat{Y}_n = x)$. Taylor expansion yields

$$T_N f(x) - f(x) = f'(x)\,\mathrm{E}(\hat{Y}_{n+1} - x \mid \hat{Y}_n = x) + \frac{f''(x)}{2}\,\mathrm{E}((\hat{Y}_{n+1} - x)^2 \mid \hat{Y}_n = x) + R_N(x)$$

with Lagrange remainder

$$R_N(x) := \frac{f'''(\xi)}{3!}\,\mathrm{E}((\hat{Y}_{n+1} - x)^3 \mid \hat{Y}_n = x),$$

where ξ is a (random) point between x and $\hat{Y}_{n+1}$. In the following it is verified that $R_N(x) = O(N^{-3/2})$ uniformly for all $x \in E_N$. Since $f \in C_c^\infty(\mathbb{R})$, there exists a $C > 0$ with $|f'''(x)| \leq C$ for all $x \in \mathbb{R}$. Thus,

$$|R_N(x)| \leq \frac{C}{3!}\,\mathrm{E}(|\hat{Y}_{n+1} - x|^3 \mid \hat{Y}_n = x) = \frac{C}{3!}\frac{1}{N^{3/2}}\,\mathrm{E}(|Y_{n+1} - i|^3 \mid Y_n = i).$$

From the results of Section 2 (see Lemma 2.2 and the remarks thereafter), it follows that, conditioned on $Y_n = i$ and $\eta = K \in \mathbb{N}$, the random variable Y_{n+1} can only take values in $\{i - 1, i, i + 1, \ldots, i + K - 1\}$ with positive probability. Conditioned on $Y_n = i$ and $\eta = 0$, Y_{n+1} can only take the two values $i - 1$ and i with positive probability. Thus, conditioned on $Y_n = i$, we have $|Y_{n+1} - i| \leq \max(\eta, 1) \leq \eta + 1$. Therefore,

$$|R_N(x)| \leq \frac{C}{3!}\frac{1}{N^{3/2}}\,\mathrm{E}((\eta + 1)^3).$$

In particular, it is shown that $R_N(x) = O(N^{-3/2})$ uniformly for all $x \in E_N$, since $\mathrm{E}(\eta^3) < \infty$ by assumption ($p'''(0+) < \infty$).

The generator A of the Ornstein–Uhlenbeck process with drift parameter $\mu(x) = -\mu x$ and diffusion parameter $\sigma^2(x) = \sigma^2$ satisfies $Af(x) = \frac{1}{2}\sigma^2 f''(x) - \mu x f'(x)$. Thus, for all $N \in \mathbb{N}$ and $x \in E_N$,

$$|N(T_N f(x) - f(x)) - Af(x)| \leq |f'(x)||N\,\mathrm{E}(\hat{Y}_{n+1} - x \mid \hat{Y}_n - x) + \mu x|$$
$$+ \frac{|f''(x)|}{2}|N\,\mathrm{E}((\hat{Y}_{n+1} - x)^2 \mid \hat{Y}_n = x) - \sigma^2|$$
$$+ N|R_N(x)|.$$

Since f has compact support, say K, the derivatives $f'(x)$ and $f''(x)$ are both equal to 0 for $x \notin K$, and these derivatives are both bounded for $x \in E_N \cap K$. Together with (4.4) and (4.5) it follows that

$$\lim_{N \to \infty} \sup_{x \in E_N} |N(T_N f(x) - f(x)) - Af(x)| = 0 \quad \text{for all } f \in C_c^\infty(\mathbb{R}).$$

As $C_c^\infty(\mathbb{R})$ is a core for A (see [14, p. 371, Theorem 2.1]), the statement follows from [14, p. 31, Theorem 6.5] and [14, p. 233, Corollary 8.9].

Remarks. 1. The stationary distribution of the limiting Ornstein–Uhlenbeck process $(V_t)_{t\geq 0}$ in Theorem 4.1 is the normal distribution $N(0, \tau^2)$ with

$$\tau^2 := \frac{\sigma^2}{2\mu} = \frac{-p''(0+)}{2(p'(0+))^3} = \frac{\mathrm{E}(\eta(\eta-1))}{2(\mathrm{E}(\eta))^3}. \tag{4.6}$$

The constant τ^2 will appear again later (see Conjecture 5.1, below).

2. The presented proof of Theorem 4.1 uses the technical condition $p'''(0+) < \infty$ in order to control the rest term $R_N(x)$. We leave open the question of whether or not Theorem 4.1 remains valid if this technical condition is replaced by the weaker condition $p''(0+) > -\infty$.

3. The condition $p'(0+) > 1$ in Theorem 4.1 excludes important models such as the standard Moran model ($p(x) = x$). For models satisfying $p'(0+) = 1$, the limiting behaviour of the chain $Y = (Y_n)_{n\in\mathbb{N}_0}$ differs significantly from that in Theorem 4.1. For example, for the standard Moran model, the transition matrix $P = (p_{ij})_{i,j\in\{0,\ldots,N\}}$ of Y has entries $p_{i,i-1} = i(i-1)/N^2$, $p_{ii} = 1 - i(i-1)/N^2$, and $p_{ij} = 0$ otherwise. Thus, for any function $f : \{0,\ldots,N\} \to \mathbb{R}$ and any $i \in \{0,\ldots,N\}$,

$$\begin{aligned}
T_N f(i) &:= \mathrm{E}(f(Y_{n+1}) \mid Y_n = i) \\
&= f(i)p_{ii} + f(i-1)p_{i,i-1} \\
&= f(i) + \frac{i(i-1)}{N^2}(f(i-1) - f(i)),
\end{aligned}$$

or, equivalently, $(N^2/2)(T_N f(i) - f(i)) = \binom{i}{2}(f(i-1) - f(i))$. It follows that if Y_0 converges in distribution to some random variable D_0 as $N \to \infty$ then the time-scaled ancestral process $(Y_{\lfloor N^2 t/2 \rfloor})_{t\geq 0}$ weakly converges to a pure-death process $D = (D_t)_{t\geq 0}$, jumping with rate $i(i-1)/2$ from state i to state $i-1$. This convergence result is well known from the literature. Note that D is the block counting process of the Kingman coalescent [20].

Example. For the K-type Moran model with parameter $K \in \{2, 3, \ldots\}$ introduced in Section 2, $p(x) = 1 - (1-x)^K$. Thus, $p'(0+) = K > 1$, $p''(0+) = -K(K-1)$, and $p'''(0+) = K(K-1)(K-2) < \infty$. If $(V_0^{(N)})_{N\in\mathbb{N}}$ converges in distribution to some V_0 then, by Theorem 4.1, the time- and space-scaled ancestral process $(V_t^{(N)})_{t\geq 0}$, defined via $V_t^{(N)} := (Y_{\lfloor Nt \rfloor} - (1-1/K)N)/\sqrt{N}$ for $N \in \mathbb{N}$ and $t \geq 0$, converges weakly to an Ornstein–Uhlenbeck process $(V_t)_{t\geq 0}$ with drift parameter $\mu(x) = -(K-1)x$ and diffusion parameter $\sigma^2(x) = \sigma^2 = (1 - 1/K)^2$.

5. Extinction probability and stationary distribution

In the following it is assumed that $p(0) = 0$ or, equivalently ($\pi_{00} = r_0 = 1 - p(0)$), that 0 is an absorbing state of the Markov chain X. For $i \in \{0, \ldots, N\}$, let

$$\beta_i := \mathrm{P}(X_n = 0 \text{ eventually} \mid X_0 = i) = \lim_{n\to\infty} \mathrm{P}(X_n = 0 \mid X_0 = i)$$

denote the extinction probability given that the chain X starts in $X_0 = i$. Note that $\beta_0 = 1$. It is straightforward to check that the column vector $\beta := (\beta_0, \ldots, \beta_N)^\top$ is a solution of the fixed point equation $\Pi\beta = \beta$, i.e. β is an eigenvector to the eigenvalue 1 of Π. Since Π

is tridiagonal, β can be computed. For example, if the state N is also absorbing (which is equivalent to $p(1) = 1$) and if $p(x) > 0$ for all $x > 0$, then (see Equation (2.158) of [15])

$$\beta_i = 1 - \frac{\varphi(i)}{\varphi(N)}, \qquad i \in \{0, \ldots, N\}, \tag{5.1}$$

where

$$\varphi(i) := \sum_{j=0}^{i-1} \prod_{k=1}^{j} \frac{q_k}{p_k} = \sum_{j=0}^{i-1} \binom{N-1}{j}^{-1} \prod_{k=1}^{j} \frac{1 - p(k/N)}{p(k/N)}, \qquad i \in \{0, \ldots, N\}, \tag{5.2}$$

with the convention that empty sums are equal to 0 and empty products are equal to 1. Note that $\varphi(0) = 0$, $\varphi(1) = 1$, and $\varphi(2) = 1 + q_1/p_1$, and that $\varphi(i)$ is increasing in i. For example, for the haploid Moran model, $\varphi(i) = i$ and, hence, $\beta_i = 1 - i/N$, $i \in \{0, \ldots, N\}$. The column vector $(\varphi(0), \ldots, \varphi(N))^\top$ is also an eigenvector to the eigenvalue 1 of Π. Thus,

$$q_i \varphi(i-1) + p_i \varphi(i+1) = (p_i + q_i)\varphi(i), \qquad i \in \{1, \ldots, N-1\}. \tag{5.3}$$

Lemma 5.1. *If $p(x) \geq x$ for all $x \in [0, 1]$ then the sequence $(\varphi(i))_{i \in \{0, \ldots, N\}}$ is log-concave, i.e. $\varphi(i-1)\varphi(i+1) \leq (\varphi(i))^2$ for all $i \in \{1, \ldots, N-1\}$.*

Proof. For $i \in \{1, \ldots, N\}$, define the differences $d_i := \varphi(i) - \varphi(i-1)$ and note that (5.3) implies that $d_{i+1} = h_i d_i$ for $i \in \{1, \ldots, N-1\}$, where $h_i := q_i/p_i$ for $i \in \{1, \ldots, N-1\}$. From $p(x) \geq x$, $x \in [0, 1]$, it follows that $h_i \leq 1$ for $i \in \{1, \ldots, N-1\}$. In the following let $i \in \{1, \ldots, N-1\}$ be fixed. It is easily checked that the inequality $\varphi(i-1)\varphi(i+1) \leq (\varphi(i))^2$ is equivalent to $d_{i+1}d_i \geq \varphi(i)(d_{i+1} - d_i)$. Using $d_{i+1} = h_i d_i$, this is in turn equivalent to $h_i d_i^2 \geq \varphi(i)(h_i d_i - d_i)$, and, hence, equivalent to $h_i d_i \geq \varphi(i)(h_i - 1)$, because $d_i > 0$. Substituting in $d_i = \varphi(i) - \varphi(i-1)$, it follows that this in turn is equivalent to $h_i \varphi(i-1) \leq \varphi(i)$. But this inequality is obviously satisfied, because $\varphi(i-1) \leq \varphi(i)$ and $h_i \leq 1$. $\qquad \square$

Assume now that $1 - p$ is completely monotone. Then, by Theorem 3.1, for each population size N, there exists a Markov chain Y which is dual to X with respect to H. We are now able to repeat the arguments already used in Section 6 of [24] and Section 5 of [25]. The choice $k = N$ in the duality relation

$$\sum_{j=0}^{N} \pi_{ij}^{(n)} h_{jk} = \sum_{j=0}^{N} h_{ij} p_{kj}^{(n)}$$

leads to

$$P(X_n = 0 \mid X_0 = i) = E\left(\binom{N - Y_n}{i}\right) \Big/ \binom{N}{i}. \tag{5.4}$$

The iNth component of the matrix $(P^n)^\top = H^{-1} \Pi^n H$ is

$$P(Y_n = i \mid Y_0 = N) = p_{Ni}^{(n)}$$

$$= \sum_{k,l=0}^{N} (H^{-1})_{ik} \pi_{kl}^{(n)} h_{lN}$$

$$= \sum_{k=0}^{N} (H^{-1})_{ik} \pi_{k0}^{(n)}$$

$$= \binom{N}{i} \sum_{k=N-i}^{N} (-1)^{i+k-N} \binom{i}{N-k} P(X_n = 0 \mid X_0 = k)$$

$$= \binom{N}{i} \sum_{j=0}^{i} (-1)^{j-i} \binom{i}{j} P(X_n = 0 \mid X_0 = N - j). \qquad (5.5)$$

By definition, the limit $\beta_i := \lim_{n\to\infty} P(X_n = 0 \mid X_0 = i)$ exists for each $i \in \{0, \ldots, N\}$. Hence, by (5.5), the limit $\pi_i := \lim_{n\to\infty} P(Y_n = i \mid Y_0 = N)$ also exists for each $i \in \{0, \ldots, N\}$. Moreover,

$$\sum_{i=0}^{N} \pi_i = \sum_{i=0}^{N} \lim_{n\to\infty} P(Y_n = i \mid Y_0 = N) = \lim_{n\to\infty} \sum_{i=0}^{N} P(Y_n = i \mid Y_0 = N) = 1.$$

Taking the limit $n \to \infty$ in (5.4) and (5.5) leads to the one-to-one correspondence

$$\pi_i = \binom{N}{i} \sum_{j=0}^{i} (-1)^{i-j} \binom{i}{j} \beta_{N-j}, \qquad i \in \{0, \ldots, N\}, \qquad (5.6)$$

and

$$\beta_i = \binom{N}{i}^{-1} \sum_{j=i}^{N} \binom{j}{i} \pi_{N-j}, \qquad i \in \{0, \ldots, N\}. \qquad (5.7)$$

Obviously, $\pi := (\pi_0, \ldots, \pi_N)$ is a stationary distribution of Y. Note that $\pi_0 = \beta_N = 0$.

If $p(1) = 1$ (N is an absorbing state of X) then the dual chain Y is not irreducible as the state 0 is disconnected from the rest of the state space. In this case, (5.6) and (5.7) relate the extinction probabilities $\beta_0, \ldots, \beta_N$ of X with the invariant measure of Y restricted to the connected states $1, \ldots, N$.

If $p(1) < 1$ (N is a partially reflecting state of X) then the dual chain Y becomes itself absorbing at 0. Let $T_0 := \inf\{n \in \mathbb{N}_0 \mid Y_n = 0\}$ denote the extinction time of Y. If we write down—in analogy to the calculations around (5.5)—the $0j$th entries on both sides of the duality equation $(P_n)^\top = H^{-1} \Pi^n H$, we obtain

$$P(T_0 \leq n \mid Y_0 = j) = P(Y_n = 0 \mid Y_0 = j)$$

$$= p_{j0}^{(n)}$$

$$= \sum_{k,l=0}^{N} (H^{-1})_{0k} \pi_{kl}^{(n)} h_{lj}$$

$$= \sum_{l=0}^{N} \pi_{Nl}^{(n)} h_{lj}$$

$$= \sum_{l=0}^{N} P(X_n = l \mid X_0 = N) \binom{N-l}{j} \bigg/ \binom{N}{j}$$

$$= E\left(\binom{N-X_n}{j} \,\bigg|\, X_0 = N \right) \bigg/ \binom{N}{j},$$

giving a relationship between the distribution function of the extinction time T_0 of the chain Y started in $Y_0 = j$ and the factorial moments of $N - X_n$ conditioned that the chain X is started in state $X_0 = N$.

From (5.1) and (5.6), it follows that

$$\pi_i = -\frac{\binom{N}{i}}{\varphi(N)} \sum_{j=0}^{i} (-1)^{i-j} \binom{i}{j} \varphi(N-j), \qquad i \in \{1, \ldots, N\}. \tag{5.8}$$

Using (5.2), the probability π_i can be expressed in terms of the function p. However, this does not seem to result in a simple formula for π_i. We verify the following property.

Lemma 5.2. *The sequence $(\pi_i)_{i \in \{1, \ldots, N\}}$ is log-concave (LC) and positive, and, hence, unimodal.*

Proof. For $j \in \{0, \ldots, N\}$, define $\alpha_j := \varphi(j)/\varphi(N)$ and $\tilde{\alpha}_j := \alpha_{N-j}/j!$, and rewrite (5.8) in the form

$$\pi_i = (N)_i \sum_{j=0}^{i} \frac{(-1)^{i-j+1}}{(i-j)!} \tilde{\alpha}_j, \qquad i \in \{1, \ldots, N\}.$$

The sequence $(\alpha_j)_{j \in \{0, \ldots, N\}}$ is LC because $(\varphi(j))_{j \in \{0, \ldots, N\}}$ is LC by Lemma 5.1. The reversed sequence $(\alpha_{N-j})_{j \in \{0, \ldots, N\}}$ is clearly LC. Thus, the sequence $\tilde{\alpha} := (\tilde{\alpha}_j)_{j \in \{0, \ldots, N\}}$ is LC as a product of two positive LC sequences. Now $(\sum_{j=0}^{i} \tilde{\alpha}_j (-1)^{i-j+1}/(i-j)!)_{i \in \{0, \ldots, N\}} = ((\tilde{\alpha} * u)_i)_{i \in \{0, \ldots, N\}}$ is the convolution of the two sequences $\tilde{\alpha}$ and $u := (u_i)_{i \in \{0, \ldots, N\}}$, where $u_i := (-1)^{i+1}/i!$ for $i \in \{0, \ldots, N\}$. Note that u is alternating but LC. The convolution of two LC sequences is LC (see [30] for instance). So $((\tilde{\alpha} * u)_i)_{i \in \{0, \ldots, N\}}$ is LC and positive (because π_i is positive). The sequence $((N)_i)_{i \in \{0, \ldots, N\}}$ is positive and LC, so $(\pi_i)_{i \in \{1, \ldots, N\}}$ is positive and LC as a product of two positive LC sequences. Positive LC sequences are unimodal. $\quad\square$

Let Z_N be a random variable with distribution π. As in Section 6 of [24], it follows that Z_N has PGF

$$\mathrm{E}(s^{Z_N}) = \sum_{j=0}^{N} \beta_j \binom{N}{j} s^{N-j} (1-s)^j, \qquad s \in \mathbb{C},$$

and factorial moments

$$\mathrm{E}((Z_N)_k) = (N)_k \sum_{j=0}^{k} \beta_j \binom{k}{j} (-1)^j, \qquad k \in \mathbb{N}_0. \tag{5.9}$$

In particular, $\mu_N := \mathrm{E}(Z_N) = N(1 - \beta_1) = N/\varphi(N)$ and $\sigma_N^2 := \mathrm{var}(Z_N) = N(\beta_1 - \beta_2 + N(\beta_2 - \beta_1^2))$, $N \in \mathbb{N}$. In the following we are interested in the distributional behaviour of Z_N as $N \to \infty$.

Proposition 5.1. (Weak law of large numbers for Z_N.) *Suppose that the continuous function $p : [0, 1] \to [0, 1]$ satisfies $p(0) = 0$ and $p(1) = 1$, and that $1 - p$ is completely monotone. If $p \neq \mathrm{Id}$ then $Z_N/\mu_N \to 1$ in probability and in L^r, $r \geq 1$, as $N \to \infty$, and $\mathrm{E}(Z_N^k) \sim (\alpha N)^k$ for all $k \in \mathbb{N}_0$, where $\alpha := 1 - 1/p'(0+) \in (0, 1]$.*

Proof. Theorem 3.1(i) together with $p(1) = 1$ ensures that there exists a random variable η taking values in $\mathbb{N}$ such that $p(x) = 1 - \mathrm{E}((1-x)^\eta)$ for all $x \in [0, 1]$. From $p \neq \mathrm{Id}$, it follows that $\mathrm{E}(\eta) \in (1, \infty]$ and $\mathrm{P}(\eta = 1) < 1$. In particular,

$$p'(0+) := \lim_{x \searrow 0} p'(x) = \lim_{x \searrow 0} \mathrm{E}(\eta(1-x)^{\eta-1}) = \mathrm{E}(\eta) \in (1, \infty]$$

and

$$p'(1-) := \lim_{x \nearrow 1} p'(x) = \lim_{x \nearrow 1} \mathrm{E}(\eta(1-x)^{\eta-1}) = \mathrm{P}(\eta = 1) \in [0, 1).$$

Furthermore, $p(x) = 1 - \mathrm{E}((1-x)^\eta) > 1 - \mathrm{E}(1-x) = x$ and, hence, $1 - p(x) < 1 - x$ for $x \in (0, 1)$. Therefore,

$$h(x) := \frac{x(1 - p(x))}{(1-x)p(x)} \in (0, 1) \quad \text{for all } x \in (0, 1). \tag{5.10}$$

Together with $\lim_{x \searrow 0} h(x) = 1/p'(0+) < 1$ and $\lim_{x \nearrow 1} h(x) = p'(1-) < 1$, it follows that $\gamma := \sup_{x \in (0,1)} h(x) < 1$. For each fixed $k \in \mathbb{N}$ and all $N \geq k$, $q_k/p_k = h(k/N) \to 1/p'(0+) =: \beta$ as $N \to \infty$. It follows that, for each fixed $i \in \mathbb{N}_0$,

$$\varphi(i) = \sum_{j=0}^{i-1} \prod_{k=1}^{j} \frac{q_k}{p_k} \to \sum_{j=0}^{i-1} \beta^j = \frac{1 - \beta^i}{\alpha}, \qquad N \to \infty,$$

where $\alpha := 1 - \beta \in (0, 1]$. We now verify that $\lim_{N \to \infty} \varphi(N) = 1/\alpha$. Clearly, $\varphi(N) \geq \varphi(i)$ for $N \geq i$ and, hence, $\liminf_{N \to \infty} \varphi(N) \geq \lim_{N \to \infty} \varphi(i) = (1 - \beta^i)/\alpha$ for all $i \in \mathbb{N}_0$, which implies that $\liminf_{N \to \infty} \varphi(N) \geq 1/\alpha$, since $0 \leq \beta < 1$. In order to see that $\limsup_{N \to \infty} \varphi(N) \leq 1/\alpha$, fix $\varepsilon > 0$ and choose $i = i(\varepsilon) \in \mathbb{N}$ such that $\gamma^i/(1 - \gamma) \leq \varepsilon$, which is possible as $\gamma < 1$. From $q_k/p_k = h(k/N) \leq \gamma$ we conclude that, for $N > i$,

$$\varphi(N) = \varphi(i) + \sum_{j=i}^{N-1} \prod_{k=1}^{j} \frac{q_k}{p_k} \leq \varphi(i) + \sum_{j=i}^{N-1} \gamma^j \leq \varphi(i) + \frac{\gamma^i}{1 - \gamma} \leq \varphi(i) + \varepsilon.$$

Therefore, $\limsup_{N \to \infty} \varphi(N) \leq (1 - \beta^i)/\alpha + \varepsilon \leq 1/\alpha + \varepsilon$. Since $\varepsilon > 0$ can be chosen arbitrarily, it follows that $\limsup_{N \to \infty} \varphi(N) \leq 1/\alpha$. Thus, $\lim_{N \to \infty} \varphi(N) = 1/\alpha$. Therefore, for each fixed $i \in \mathbb{N}$, the extinction probability β_i satisfies $\beta_i = 1 - \varphi(i)/\varphi(N) \to \beta^i$ as $N \to \infty$. From (5.9), it follows that $\mathrm{E}((Z_N/N)^k) \to \alpha^k$ as $N \to \infty$ for all $k \in \mathbb{N}_0$, which is equivalent to the assertion of the proposition, since $0 \leq Z_N/N \leq 1$ for all $N \in \mathbb{N}$. $\qquad \square$

The next lemma provides precise information about the asymptotics of the mean $\mu_N = \mathrm{E}(Z_N)$ and the variance $\sigma_N^2 := \mathrm{var}(Z_N)$ of the stationary distribution of the ancestral process.

Lemma 5.3. *Suppose that the continuous function $p \colon [0, 1] \to [0, 1]$ satisfies $p(0) = 0$ and $p(1) = 1$, and that $1 - p$ is completely monotone. If $-\infty < p''(0+) < 0$ then*

$$\lim_{N \to \infty} (\mu_N - \alpha N) = 1 - \frac{p''(0+)}{2p'(0+)(1 - p'(0+))} \quad \text{and} \quad \lim_{N \to \infty} \frac{\sigma_N^2}{N} = \tau^2,$$

where $\alpha := 1 - 1/p'(0+) \in (0, 1)$ and τ^2 is defined in (4.6).

Proof. As in the proof of Proposition 5.1, consider the auxiliary function $h\colon (0, 1) \to (0, 1)$ defined via (5.10). Recall that $h(0+) := \lim_{x \searrow 0} h(x) = 1/p'(0+) =: \beta \in (0, 1)$, as $p'(0+) = \mathrm{E}(\eta) \in (1, \infty)$. It is straightforward to check that

$$h'(x) = \frac{p(x)(1 - p(x)) - x(1 - x)p'(x)}{(1 - x)^2(p(x))^2}, \qquad x \in (0, 1).$$

Using $p(x)/x \sim p'(0+) = 1/\beta$ and applying l'Hôpitals' rule yields

$$
\begin{aligned}
h'(0+) &:= \lim_{x \searrow 0} h'(x) \\
&= \lim_{x \searrow 0} \frac{p(x)(1 - p(x)) - x(1 - x)p'(x)}{(p(x))^2} \\
&= -1 + \lim_{x \searrow 0} \frac{p(x) - x(1 - x)p'(x)}{(p(x))^2} \\
&= -1 + \beta^2 \lim_{x \searrow 0} \frac{p(x) - (x - x^2)p'(x)}{x^2} \\
&= -1 + \beta^2 \lim_{x \searrow 0} \frac{p'(x) - (1 - 2x)p'(x) - x(1 - x)p''(x)}{2x} \\
&= -1 + \beta^2 \lim_{x \searrow 0} \left(p'(x) - (1 - x)\frac{p''(x)}{2} \right) \\
&= -1 + \beta^2 \left(p'(0+) - \frac{p''(0+)}{2} \right) \\
&= -\alpha - \frac{\beta^2}{2} p''(0+),
\end{aligned}
$$

where $\alpha := 1 - \beta \in (0, 1)$. For each fixed $k \in \mathbb{N}$ and all $N \geq k$,

$$\frac{q_k}{p_k} = h\left(\frac{k}{N} \right) = h(0+) + h'(0+)\frac{k}{N} + O\left(\frac{1}{N^2} \right) = \beta + h'(0+)\frac{k}{N} + O\left(\frac{1}{N^2} \right),$$

and, therefore, for fixed $j \in \mathbb{N}_0$ and $N > j$,

$$\prod_{k=1}^{j} \frac{q_k}{p_k} = \beta^j + \frac{h'(0+)}{N} \frac{j(j + 1)}{2} \beta^{j-1} + O\left(\frac{1}{N^2} \right).$$

It follows that, for each fixed $i \in \mathbb{N}_0$,

$$
\begin{aligned}
\varphi(i) &= \sum_{j=0}^{i-1} \prod_{k=1}^{j} \frac{q_k}{p_k} \\
&= \sum_{j=0}^{i-1} \left(\beta^j + \frac{h'(0+)}{2N} j(j + 1)\beta^{j-1} + O\left(\frac{1}{N^2} \right) \right) \\
&= \frac{1 - \beta^i}{\alpha} + \frac{h'(0+)M_{i,\beta}}{2N} + O\left(\frac{1}{N^2} \right),
\end{aligned}
$$

where

$$
\begin{aligned}
M_{i,\beta} &:= \sum_{j=0}^{i-1} j(j+1)\beta^{j-1} \\
&= \frac{\partial^2}{\partial^2 \beta} \sum_{j=0}^{i-1} \beta^{j+1} \\
&= \frac{\partial^2}{\partial^2 \beta} \left(\frac{1 - \beta^{i+1}}{1 - \beta} - 1 \right) \\
&= \frac{\partial}{\partial \beta} \left(\frac{1 - \beta^{i+1}}{(1 - \beta)^2} - (i+1)\frac{\beta^i}{1 - \beta} \right) \\
&= \frac{2(1 - \beta^{i+1})}{(1 - \beta)^3} + (i+1)\left(\frac{\beta^{i-1}}{1 - \beta} - \frac{2\beta^i}{(1 - \beta)^2} \right) - (i+1)^2 \frac{\beta^{i-1}}{1 - \beta}.
\end{aligned}
$$

Note that $M_{1,\beta} = 0$ and $M_{2,\beta} = 2$, and that $\lim_{i\to\infty} M_{i,\beta} = 2/(1 - \beta)^3 = 2/\alpha^3$, since $\beta < 1$.

In the following it is shown that

$$
\varphi(N) = \frac{1}{\alpha} + \frac{h'(0+)}{\alpha^3}\frac{1}{N} + O\left(\frac{1}{N^2} \right), \qquad N \to \infty. \tag{5.11}
$$

As in the previous proof, let $\gamma := \sup_{x\in(0,1)} h(x) < 1$. For fixed $\varepsilon \in (0, 1)$ and all $N \in \mathbb{N}$,

$$
0 \le N^2(\varphi(N) - \varphi(\lfloor N^\varepsilon \rfloor)) = N^2 \sum_{j=\lfloor N^\varepsilon \rfloor}^{N-1} \prod_{k=1}^{j} h\left(\frac{k}{N} \right) \le N^2 \sum_{j=\lfloor N^\varepsilon \rfloor}^{N-1} \gamma^j \le N^2 \frac{\gamma^{\lfloor N^\varepsilon \rfloor}}{1 - \gamma} \to 0
$$

as $N \to \infty$. Thus, it suffices to show that, for some $\varepsilon \in (0, 1)$,

$$
\varphi(\lfloor N^\varepsilon \rfloor) = \frac{1}{\alpha} + \frac{h'(0+)}{\alpha^3}\frac{1}{N} + O\left(\frac{1}{N^2} \right), \qquad N \to \infty.
$$

In order to see this, fix some $\varepsilon \in (0, \frac{1}{2})$ (for example, $\varepsilon = \frac{1}{4}$) and adapt the previous arguments to the case when $i := i_N := \lfloor N^\varepsilon \rfloor$ depends on N. For all $k \in \{1, \ldots, i_N\}$,

$$
\frac{q_k}{p_k} = h\left(\frac{k}{N} \right) = \beta + h'(0+)\frac{k}{N} + O\left(\frac{k^2}{N^2} \right), \qquad N \to \infty.
$$

Therefore, for $j \in \{0, \ldots, i_N - 1\}$,

$$
\prod_{k=1}^{j} \frac{q_k}{p_k} = \beta^j + \frac{h'(0+)}{N}\frac{j(j+1)}{2}\beta^{j-1} + O\left(\frac{j^4 \beta^{j-2}}{N^2} \right), \qquad N \to \infty.
$$

It follows that

$$\varphi(i_N) = \sum_{j=0}^{i_N-1} \prod_{k=1}^{j} \frac{q_k}{p_k}$$

$$= \sum_{j=0}^{i_N-1} \left(\beta^j + \frac{h'(0+)}{2N} j(j+1)\beta^{j-1} + O\left(\frac{j^4 \beta^{j-2}}{N^2}\right) \right)$$

$$= \frac{1 - \beta^{i_N}}{\alpha} + \frac{h'(0+)M_{i_N,\beta}}{2N} + O\left(\frac{1}{N^2}\right)$$

$$= \frac{1}{\alpha} + \frac{h'(0+)}{\alpha^3} \frac{1}{N} + O\left(\frac{1}{N^2}\right),$$

since $\beta^{i_N} = O(1/N^2)$ and $M_{i_N,\beta} = 2/\alpha^3 + O(i_N^2 \beta^{i_N}) = 2/\alpha^3 + O(1/N)$. The proof of (5.11) is complete. Taylor expansion of $f(x) := 1/x$ at the point $1/\alpha$ yields

$$\frac{1}{\varphi(N)} = f(\varphi(N))$$

$$= f\left(\frac{1}{\alpha}\right) + f'\left(\frac{1}{\alpha}\right)\left(\varphi(N) - \frac{1}{\alpha}\right) + O\left(\left(\varphi(N) - \frac{1}{\alpha}\right)^2\right)$$

$$= \alpha - \alpha^2\left(\varphi(N) - \frac{1}{\alpha}\right) + O\left(\left(\varphi(N) - \frac{1}{\alpha}\right)^2\right)$$

$$= \alpha - \frac{h'(0+)}{\alpha} \frac{1}{N} + O\left(\frac{1}{N^2}\right), \qquad N \to \infty,$$

where the last equality follows from (5.11). Hence, for each fixed $i \in \mathbb{N}$, the extinction probability β_i satisfies

$$\beta_i = 1 - \frac{\varphi(i)}{\varphi(N)}$$

$$= 1 - \left(\frac{1 - \beta^i}{\alpha} + \frac{h'(0+)M_{i,\beta}}{2} \frac{1}{N} + O\left(\frac{1}{N^2}\right)\right)\left(\alpha - \frac{h'(0+)}{\alpha} \frac{1}{N} + O\left(\frac{1}{N^2}\right)\right)$$

$$= \beta^i + \left(\frac{1 - \beta^i}{\alpha^2} - \frac{\alpha}{2} M_{i,\beta}\right) \frac{h'(0+)}{N} + O\left(\frac{1}{N^2}\right), \qquad N \to \infty.$$

In particular, $\beta_1 = \beta + h'(0+)/(\alpha N) + O(1/N^2)$ and $\beta_2 = \beta^2 + \beta(3 - \beta)h'(0+)/(\alpha N) + O(1/N^2)$. The mean $\mu_N := E(Z_N)$ and the variance $\sigma_N^2 := \mathrm{var}(Z_N)$ of the stationary distribution therefore satisfy

$$\lim_{N \to \infty} (\mu_N - N\alpha) = \lim_{N \to \infty} N(1 - \beta_1 - \alpha)$$

$$= -\frac{h'(0+)}{\alpha}$$

$$= 1 + \frac{\beta^2}{2\alpha} p''(0+)$$

$$= 1 - \frac{p''(0+)}{2p'(0+)(1 - p'(0+))}$$

and

$$\lim_{N \to \infty} \frac{\sigma_N^2}{N} = \lim_{N \to \infty} (\beta_1 - \beta_2 + N(\beta_2 - \beta_1^2))$$
$$= \beta(\alpha + h'(0+))$$
$$= -\frac{\beta^3}{2} p''(0+)$$
$$= \frac{-p''(0+)}{2(p'(0+))^3}$$
$$= \tau^2,$$

which completes the proof.

As already mentioned at the end of Section 4, the limiting Ornstein–Uhlenbeck process in Theorem 4.1 has stationary distribution $N(0, \tau^2)$. Based on the intuition that it is allowed to interchange the time limit, $t \to \infty$, and the space limit, $N \to \infty$, it is therefore natural to state the following conjecture.

Conjecture 5.1. (*Central limit theorem for Z_N.*) *Suppose that the continuous function $p \colon [0, 1] \to [0, 1]$ satisfies $p(0) = 0$ and $p(1) = 1$, and that $1 - p$ is completely monotone. If $-\infty < p''(0+) < 0$ then $(Z_N - \alpha N)/\sqrt{N}$ weakly converges to the normal law $N(0, \tau^2)$ as $N \to \infty$, where $\alpha := 1 - 1/p'(0+) \in (0, 1)$ and $\tau^2 \in (0, \infty)$ is defined in (4.6).*

Remark. If Conjecture 5.1 holds then Lemma 2 of [3, p. 99] shows that Z_N is locally asymptotic normal, since $(\pi_i)_i$ is LC by Lemma 5.2.

6. Examples

In Section 2 we introduced a fundamental example, the multitype Moran model. In this section we provide some more details for this model. We then present further concrete examples, some of which involve selective forces acting on particular genes or genotypes. The functions p in these examples have also been considered in [18] in the context of a modified Wright–Fisher model.

Example 6.1. (*The K-type Moran model.*) For a constant $K \in \mathbb{N}$, consider the K-type Moran model already introduced in Section 2. Kämmerle [19] showed that, for $K = 2$, the Markov chain X has extinction probabilities $\beta_i = \binom{N-1}{i} / \binom{2N}{i}$, $i \in \{0, \ldots, N\}$. In particular,

$$\varphi(i) = \frac{2N}{N+1} \left(1 - \binom{N-1}{i} \bigg/ \binom{2N}{i} \right), \qquad i \in \{0, \ldots, N\},$$

and, hence, $\varphi(N) = 2N/(N + 1) \to 2$. Furthermore, Kämmerle verified that, for $K = 2$, the stationary distribution of the chain Y is the hypergeometric distribution $\pi_i = \binom{N}{i}\binom{N}{i-1} / \binom{2N}{N+1}$, $i \in \{1, \ldots, N\}$. In particular, $\mathrm{E}(Z_N) = (N+1)/2 \sim N/2$, $\mathrm{var}(Z_N) = (N^2 - 1)/(4(2N - 1)) \sim N/8$, $Z_N/N \to \frac{1}{2}$ in probability, and $(Z_N - N/2)/N^{1/2}$ weakly converges to the normal law with mean 0 and variance $\frac{1}{8}$ as $N \to \infty$, in agreement with Proposition 5.1 and Conjecture 5.1, as $p'(0+) = 2$ and $p''(0+) = -2$.

For an arbitrary parameter $K \in \{2, 3, \ldots\}$, we have $p'(0+) = K$. Therefore, $Z_N/N \to 1 - 1/K$ in probability and in L^r, $r \geq 1$, as $N \to \infty$ by Proposition 5.1. Moreover, $p''(0+) = -K(K-1)$. By Conjecture 5.1, we conjecture that $(Z_N - N(1 - 1/K))/N^{1/2}$ weakly converges to the normal law with mean 0 and variance $\tau^2 = -p''(0+)/(2(p'(0+))^3) = (K - 1)/(2K^2)$.

Example 6.2. (*A selection model.*) For some parameter $s > 0$, let $p(x) := (1 + s)x/(1 + sx)$. This choice of p corresponds to the well-known selection model in which two genes A and a are considered, where A is selectively advantageous with fitness $1 + s$ and a has standard fitness 1. The variable X_n of the Markov chain $X = (X_n)_{n \in \mathbb{N}_0}$ counts the number of A genes at time (generation) n of a population of size N. The case in which $s = 0$ is excluded because it corresponds to the standard Moran model. Negative values of s are also excluded because, in this case, the function $1 - p$ is not completely monotone. Note that $p(x) = 1 - \mathrm{E}((1 - x)^\eta)$, where η is a random variable with distribution $\mathrm{P}(\eta = k) = s^{k-1}/(1 + s)^k = \beta\alpha^{k-1}$, $k \in \mathbb{N}$, where $\alpha := s/(s + 1) \in (0, 1)$ and $\beta := 1 - \alpha$. We have

$$\varphi(i) = \frac{s + 1}{s}\left(1 - \left(\frac{1}{s + 1}\right)^i\right) = \frac{1 - \beta^i}{\alpha}, \qquad i \in \{0, \ldots, N\}.$$

For this particular example, $\varphi(i)$ does not depend on N. The extinction probability is $\beta_i = 1 - (1 - \beta^i)/(1 - \beta^N)$, $i \in \{0, \ldots, N\}$. The chain Y has stationary distribution

$$\pi_i = \frac{1}{1 - \beta^N}\binom{N}{i}\alpha^i \beta^{N-i}, \qquad i \in \{1, \ldots, N\},$$

which is the binomial distribution with parameters N and α, conditioned that it does not take the value 0. Consequently, $(Z_N - \alpha N)/(N\alpha\beta)^{1/2}$ weakly converges to the standard normal law as $N \to \infty$, in agreement with Conjecture 5.1, as $p'(0+) = 1 + s$ and $p''(0+) = -2s(s+1)$, and, therefore, $\tau^2 := -p''(0+)/(2(p'(0+))^3) = s/(s + 1)^2 = \alpha\beta$. Theorem 4.1 is applicable. The limiting Ornstein–Uhlenbeck process has drift parameter $\mu(x) = -sx$ and diffusion parameter $\sigma^2(x) = \sigma^2 = 2s^2/(s + 1)^2$. The PGF $z \mapsto \mathrm{E}(z^{Z_N}) = \sum_{i=1}^N \pi_i z^i = ((\alpha z + \beta)^N - \beta^N)/(1 - \beta^N)$, $z \in \mathbb{C}$, has at most two real roots, namely, always $z = 0$ and $z = -2\beta/\alpha$, if N is even. Both these roots have multiplicity 1. This example shows that, in general, not all the N roots of the PGF of Z_N are real. Thus, in general it is not possible to verify Conjecture 5.1 by showing that all the roots of the PGF of Z_N are real.

Example 6.3. (*Selection with dominance.*) Let $s, h \in \mathbb{R}$ be two parameters with $s > -1$ and $sh > -1$. Suppose that

$$p(x) = \frac{(1 + s)x^2 + (1 + sh)x(1 - x)}{1 + sx^2 + 2shx(1 - x)}. \tag{6.1}$$

In this model, genotypes AA, Aa, and aa with frequencies x^2, $2x(1 - x)$, and $(1 - x)^2$, respectively, have fitnesses $1 + s$, $1 + sh$, and 1, respectively. The parameter h measures the degree of dominance of the heterozygote Aa. Note that (6.1) can be put into the canonical form of deviation from neutrality as

$$p(x) = x + sx(1 - x)\frac{h - x(2h - 1)}{1 + sx^2 + 2shx(1 - x)},$$

where the fraction appearing on the right-hand side is the ratio of the difference in marginal fitnesses of A and a to their mean fitnesses. The case in which $s > 0$ and $h \in (0, 1)$ corresponds to directional selection, where genotype AA has highest fitness and Aa has intermediate fitness. In this situation, the marginal fitness of A exceeds the marginal fitness of a and selective sweeps are expected.

For $s > 0$ and $h \in (0, \frac{1}{2})$, allele A is dominant to a, whereas, for $s > 0$ and $h \in (\frac{1}{2}, 1)$, allele A is recessive to a (a stabilizing effect slowing down the sweep). The critical value $h = \frac{1}{2}$ corresponds to balancing selection with

$$p(x) = x + \frac{s}{2} \frac{x(1-x)}{1+sx}.$$

Note that if $p(x)$ is the selection mechanism with dominance (6.1) with parameters s and h, then $\bar{p}(x) := 1 - p(1-x)$ is a selection with dominance mechanism with parameters $\bar{s} := -s/(s+1)$ and $\bar{h} := 1 - h$.

It is now verified that $q := 1 - p$ is completely monotone if and only if $s = 0$ or $h = \frac{1}{2}$ or $h \in (0, \frac{1}{2})$ and $s \geq (1 - 2h)/h^2 \ (> 0)$. For $s = 0$, this is clear, because, in this case, the model reduces to the standard haploid Moran model ($p(x) = x$). For $h = \frac{1}{2}$, it is easily seen that $p(x) = 1 - \mathrm{E}((1-x)^\eta)$, where η is a random variable with distribution $\mathrm{P}(\eta = 1) = (2+s)/(2(1+s))$ and $\mathrm{P}(\eta = k) = s^{k-1}/(2(1+s)^k)$ for $k \in \{2, 3, \ldots\}$. In particular, q is completely monotone for $h = \frac{1}{2}$. Assume now that $s \neq 0$ and that $h \neq \frac{1}{2}$. Then we shall soon see that q is completely monotone if and only if the roots of the denominator D in (6.1) are real and negative. The roots are real if and only if $s^2h^2 \geq s(1-2h)$. We have $D := 1 + 2shx + s(1-2h)x^2 = s(1-2h)(x^2 + Sx + P)$, where $S := 2h/(1-2h)$ and $P := 1/(s(1-2h))$. Thus, the roots are both negative if and only if $S > 0$ and $P > 0$, or, equivalently, if and only if $s > 0$ or $h \in (0, \frac{1}{2})$. Thus, q is completely monotone if and only if $s = 0$ or $h = \frac{1}{2}$ or $h \in (0, \frac{1}{2})$ and $s \geq (1 - 2h)/h^2$.

In the following the case in which $h \in (0, \frac{1}{2})$ and $s \geq (1 - 2h)/h^2 \ (> 0)$ is studied in more detail. Let x_+ and x_- denote the two real negative roots, i.e.

$$x_\pm := \frac{-sh \pm \sqrt{s^2h^2 - s(1-2h)}}{s(1-2h)} = \frac{-h \pm \sqrt{h^2 - (1-2h)/s}}{1-2h}.$$

Note that $x_- \leq x_+ < 0$. Both roots coincide ($x_+ = x_- = -h/(1-2h)$) for $s = (1-2h)/h^2$. Two cases are now distinguished.

Case 1. Assume that $x_+ \neq x_-$. Then, expansion into partial fractions yields

$$q(x) = 1 - p(x) = \frac{(1-x)(1+shx)}{s(1-2h)(x-x_+)(x-x_-)} = A + \frac{B}{x-x_+} + \frac{C}{x-x_-},$$

where

$$A := \frac{-h}{1-2h}, \qquad B := \frac{(1-x_+)(1+shx_+)}{s(1-2h)(x_+-x_-)}, \quad \text{and} \quad C := \frac{(1-x_-)(1+shx_-)}{s(1-2h)(x_--x_+)}.$$

From $s \geq (1-2h)/h^2$, it follows that $1 + shx_+ \geq 0$ and $1 + shx_- \leq 0$. Thus, $B, C \geq 0$, from which it follows that q is completely monotone. The PGF of η is

$$\mathrm{E}(x^\eta) = q(1-x) = A + \frac{B}{1-x-x_+} + \frac{C}{1-x-x_-} = A + \frac{B\alpha_+}{1-x\alpha_+} + \frac{C\alpha_-}{1-x\alpha_-},$$

where $\alpha_\pm := 1/(1-x_\pm) \in (0, 1)$. Thus, $\mathrm{P}(\eta = 0) = A + B\alpha_+ + C\alpha_- = 0$ and $\mathrm{P}(\eta = k) = B\alpha_+^{k+1} + C\alpha_-^{k+1}$ for $k \in \mathbb{N}$.

Case 2. Assume that $x_+ = x_- = -h/(1-2h)$. Then, $s = (1-2h)/h^2$ and q reduces to

$$q(x) = \frac{(1-x)(1+shx)}{s(1-2h)(x-x_+)^2} = \frac{1-x}{sh(x-x_+)} = \frac{h(1-x)}{x(1-2h)+h} = A + \frac{B}{x-x_+},$$

where $A := -1/(sh) = -h/(1-2h)$, as before, but $B := (1-x_+)/(sh) = h(1-h)/(1-2h)^2$. Again, q is completely monotone, as $B \geq 0$. The PGF of η is

$$\mathrm{E}(x^\eta) = q(1-x) = A + \frac{B}{1-x-x_+} = A + \frac{B\alpha_+}{1-x\alpha_+},$$

where $\alpha_+ := 1/(1-x_+) \in (0, 1)$. Thus, $\mathrm{P}(\eta = 0) = A + B\alpha_+ = -h/(1-2h) + 1/(sh) = 0$ and $\mathrm{P}(\eta = k) = B\alpha_+^{k+1}$ for $k \in \mathbb{N}$.

Suppose now that $s \neq 0$ and that $h = \frac{1}{2}$ or that $h \in (0, \frac{1}{2})$ and $s \geq (1-2h)/h^2$. Then, the conditions of Proposition 5.1 are satisfied. Therefore, $Z_N/N \to \alpha$ in probability and in L^r, $r \geq 1$, where $\alpha := 1 - 1/p'(0+) = 1 - 1/(sh+1) = sh/(sh+1) \in (0, 1)$. It is straightforward to check that $p''(0+) = -2s(2sh^2 + 3h - 1)$. Thus, by Conjecture 5.1, we conjecture that $(Z_N - \alpha N)/N^{1/2}$ weakly converges to the normal distribution with mean 0 and variance $\tau^2 = -p''(0+)/(2(p'(0+))^3) = s(2sh^2 + 3h - 1)/(sh+1)^3 \in (0, \infty)$. The limiting Ornstein–Uhlenbeck process in Theorem 4.1 has drift parameter $\mu(x) = -\alpha x/(1-\alpha) = -shx$ and diffusion parameter $\sigma^2(x) = \sigma^2 = -\alpha(1-\alpha)^2 p''(0+) = 2s^2 h(2sh^2 + 3h - 1)/(sh+1)^3$.

Finally, an example is mentioned where all the moments of η are infinite.

Example 6.4. Fix a parameter $0 < \gamma < 1$ and consider $p(x) := x^\gamma$, or, equivalently, $p(x) = 1 - \mathrm{E}((1-x)^\eta)$, where η is a random variable with distribution $\mathrm{P}(\eta = k) = (-1)^{k-1}\binom{\gamma}{k}$, $k \in \mathbb{N}$. It seems that expression (5.1) for the extinction probability β_i cannot be further simplified in this case. Note that $p'(0+) = \infty$, and, hence, $Z_N/N \to 1$ in probability and in L^r, $r \geq 1$, as $N \to \infty$ by Proposition 5.1. However, Theorem 4.1 and Conjecture 5.1 are not applicable, as $p''(0+) = -\infty$. We therefore do not expect a normal limiting behaviour of the stationary distribution of Y. It would be interesting to find a nondegenerate weak limiting law of $f_N(Z_N)$ for an appropriate scaling function f_N.

References

[1] ALKEMPER, R. AND HUTZENTHALER, M. (2007). Graphical representation of some duality relations in stochastic population models. *Electron. Commun. Prob.* **12,** 206–220.

[2] ATHREYA, S. R. AND SWART, J. M. (2005). Branching-coalescing particle systems. *Prob. Theory Relat. Fields* **131,** 376–414.

[3] BENDER, E. A. (1973). Central and local limit theorems applied to asymptotic enumeration. *J. Combinatorial Theory A* **15,** 91–111.

[4] BERG, C., MATEU, J. AND PORCU, E. (2008). The Dagum family of isotropic correlation functions. *Bernoulli* **14,** 1134–1149.

[5] BERTOIN, J. AND LE GALL, J.-F. (2003). Stochastic flows associated to coalescent processes. *Prob. Theory Relat. Fields* **126,** 261–288.

[6] BLYTHE, R. A. (2007). The propagation of a cultural or biological trait by neutral genetic drift in a subdivided population. *Theoret. Pop. Biol.* **71,** 454–472.

[7] COOP, C. AND GRIFFITHS, R. C. (2004). Ancestral inference on gene trees under selection. *Theoret. Pop. Biol.* **66,** 219–232.

[8] DIACONIS, P. AND FILL, J. A. (1990). Strong stationary times via a new form of duality. *Ann. Prob.* **18,** 1483–1522.

[9] DONNELLY, P. (1984). The transient behavior of the Moran model in population genetics. *Math. Proc. Camb. Phil. Soc.* **95,** 349–358.

[10] DONNELLY, P. (1985). Dual processes and an invariance result for exchangeable models in population genetics. *J. Math. Biol.* **23,** 103–118.

[11] DONNELLY, P. (1986). Dual processes in population genetics. In *Stochastic Spatial Processes* (Lecture Notes Math. **1212**), Springer, Berlin, pp. 94–105.

[12] DONNELLY, P. AND RODRIGUES, E. R. (2000). Convergence to stationarity in the Moran model. *J. Appl. Prob.* **37,** 705–717.

[13] ETHIER, S. N. AND KRONE, S. M. (1995). Comparing Fleming–Viot and Dawson–Watanabe processes. *Stoch. Process. Appl.* **60,** 171–190.

[14] ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes*. John Wiley, New York.

[15] EWENS, W. J. (2004). *Mathematical Population Genetics. I*, 2nd edn. Springer, New York.

[16] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd edn. John Wiley, New York.

[17] GRIFFEATH, D. (1993). Frank Spitzer's pioneering work on interacting particle systems. *Ann. Prob.* **21,** 608–621.

[18] HUILLET, T. (2009). A duality approach to the genealogies of discrete non-neutral Wright–Fisher models. *J. Prob. Statist.* **2009,** 714701, 22 pp.

[19] KÄMMERLE, K. (1989). Looking forwards and backwards in a bisexual Moran model. *J. Appl. Prob.* **26,** 880–885.

[20] KINGMAN, J. F. C. (1982). The coalescent. *Stoch. Process. Appl.* **13,** 235–248.

[21] LIGGETT, T. M. (1985). *Interacting Particles Systems*. Springer, New York.

[22] LIGGETT, T. M. (1999). *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Springer, Berlin.

[23] MORAN, P. A. P. (1958). Random processes in genetics. *Proc. Camb. Phil. Soc.* **54,** 60–71.

[24] MÖHLE, M. (1994). Forward and backward processes in bisexual models with fixed population sizes. *J. Appl. Prob.* **31,** 309–332.

[25] MÖHLE, M. (1999). The concept of duality and applications to Markov processes arising in neutral population genetics models. *Bernoulli* **5,** 761–777.

[26] MÖHLE, M. (2001). Forward and backward diffusion approximations for haploid exchangeable population models. *Stoch. Process. Appl.* **95,** 133–149.

[27] SIEGMUND, D. (1976). The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes. *Ann. Prob.* **4,** 914–924.

[28] SPITZER, F. (1970). Interaction of Markov processes. *Adv. Math.* **5,** 246–290.

[29] SUDBURY, A. AND LLOYD, P. (1995). Quantum operators in classical probability theory. II. The concept of duality in interacting particle systems. *Ann. Prob.* **23,** 1816–1830.

[30] WANG, Y. AND YEH, Y.-N. (2007). Log-concavity and LC-positivity. *J. Combinatorial Theory A* **114,** 195–210.