# A look–down model with selection

B. Bah, E. Pardoux and A. B. Sow

**Abstract** The goal of this paper is to study a new version of the look-down construction with selection. We show (see Theorem 2) convergence in probability, locally uniformly in $t$, as the population size $N$ tends to infinity, towards the Wright-Fisher diffusion with selection.

## 1 Introduction and Preliminaries

In this paper we consider the simplest look–down (also called by some authors the "modified look–down") model with selection. We consider the case of two alleles $b$ and $B$, where $B$ has a selective advantage over $b$. This selective advantage is modelled by a death rate $\alpha$ for the type $b$ individuals, while the type $B$ individuals are not subject to that specific death mechanism. The look–down construction is due to Donnelly and Kurtz, see [3] and [4] in the neutral case. Those authors extended their construction to the selective case in [5].

Our selective look–down construction is slightly different from theirs. We will consider the proportion of $b$ individuals. Hence type $b$ individuals are coded by 1, and $B$ by 0. We assume that individuals are placed at time 0 on levels 1,2,..., each one being, independently from the others, 1 with probability $x$, 0 with probability $1-x$, for some $0 < x < 1$. For any $t \geq 0$, $i \geq 1$, let $\eta_t(i)$ denote the type of the individual

B. Bah
LERSTAD, UFR S.A.T, Université Gaston Berger, BP 234, Saint-Louis, SENEGAL. e-mail: :bbah12@yahoo.fr

E. Pardoux
CMI, LATP-UMR 6632, Université de Provence, 39 rue F. Joliot Curie, Marseille cedex 13, FRANCE. e-mail: pardoux@cmi.univ-mrs.fr

A. B. Sow
LERSTAD, UFR S.A.T, Université Gaston Berger, BP 234, Saint-Louis, SENEGAL. e-mail: ahmadou-bamba.sow@ugb.edu.sn

sitting on site $i$ at time $t$. Clearly $\eta_t(i) \in \{0,1\}$. The evolution of the population is governed by the two following mechanisms.

1. *Births* For any $1 \le i < j$, arrows are placed from $i$ to $j$ according to a rate one Poisson process, independently of the other pairs $i' < j'$. Suppose there is an arrow from $i$ to $j$ at time $t$. Then a descendent (of the same type) of the individual sitting on level $i$ at time $t^-$ occupies the level $j$ at time $t$, while for any $k \ge j$, the individual occupying the level $k$ at time $t^-$ is shifted to level $k+1$ at time $t$. In other words, $\eta_t(k) = \eta_{t^-}(k)$ for $k < j$, $\eta_t(j) = \eta_{t^-}(i)$, $\eta_t(k) = \eta_{t^-}(k-1)$ for $k > j$.
2. *Deaths* Any type 1 individual dies at rate $\alpha$, his vacant level being occupied by his right neighbor, who himself is replaced by his right neighbor, etc. In other words, independently of the above arrows, crosses are placed on each level according to a rate $\alpha$ Poisson process, independently of the other levels. Suppose there is a cross at level $i$ at time $t$. If $\eta_{t^-}(i) = 0$, nothing happens. If $\eta_{t^-}(i) = 1$, then $\eta_t(k) = \eta_{t^-}(k)$ for $k < i$, and $\eta_t(k) = \eta_{t^-}(k+1)$ for $k \ge i$.

This model has been formulated by Anton Wakolbinger in an oral presentation [8]. In contradiction with the models studied in [3], [4] and [5], the evolution of the $N$ first individuals $\eta_t(1), \ldots, \eta_t(N)$ depends upon the next ones, and $X_t^N = N^{-1}(\eta_t(1) + \cdots + \eta_t(N))$ is not a Markov process. We will show however that for each $t > 0$ the $\{\eta_t(k),\ k \ge 1\}$ constitute an exchangeable sequence of $\{0,1\}$–valued random variables, to which we can apply de Finetti's theorem, and that $X_t^N \to X_t$ in probability, locally uniformly in $t \ge 0$, where $X_t$ is a $[0,1]$–valued Markov process, solution of the Wright–Fisher SDE with selection (1).

In fact $\{X_t^N,\ t \ge 0\}$ is approximately Markovian, in a sense which will be clear below. It is possible, also no certain, that the techniques of proof from [3], [4] and [5] might be adaptable in the present situation. We rather prefer to use a quite different approach. In particular, there is no mention of a generator in this paper. Rather, we use extensively de Finetti's theorem, tightness, and a duality argument between the Wright–Fisher diffusion with selection, and what could be thought of as an ancestral recombination graph. We do not claim any superiority of our method of proof over that of [3], [4] and [5]. We just think that new approaches may be interesting in that they bring new insights into the problem.

Our paper is organized as follows. Section 2 presents the duality relation between a birth-death process (which could be viewed as related to an ARG) and Wright-Fisher's diffusion with selection. We both construct our process, and establish a crucial exchangeability property satisfied by our look-down model with selection in section 3. We prove the convergence result in section 4.

## 2 Wright-Fisher diffusion with selection and duality

Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbf{P})$, be a stochastic basis on which a $d-$dimensional Brownian motion $(B_t)_{t\geq 0}$ is defined. We assume that $\mathscr{F}_t = \sigma\{B_s, 0 \leq s \leq t\} \vee \mathscr{N}$, where $\mathscr{N}$ is the class of $\mathbf{P}-$ null sets of $\mathscr{F}$.

**Definition 1.** A Wright-Fisher diffusion with selection is a $[0,1]$–valued Markov process $Y = \{Y_t, t \geq 0\}$ with continuous paths, solution of the following stochastic differential equation

$$\begin{cases} dY_t = -\alpha Y_t(1-Y_t)dt + \sqrt{Y_t(1-Y_t)}dB_t, \quad t \geq 0, \\ Y_0 = y, \quad 0 < y < 1, \end{cases} \tag{1}$$

where $B$ is a realization of the standard Brownian motion, $\alpha \in \mathbb{R}$.

In all what follows, $\alpha > 0$. $Y_t$ will denote the proportion of non-advantageous alleles.

In this section we study the duality between a jump Markov process and Wright-Fisher diffusion with selection.

Let $\{R_t, t \geq 0\}$ be a $\mathbf{N}^*$-valued jump Markov process which, when in state $k$, jumps to

1. $k-1$ at rate $\binom{k}{2}$;
2. $k+1$ at rate $\alpha k, \quad \alpha > 0$.

In other words, the infinitesimal generator of $\{R_t, t \geq 0\}$ is given by:

$$Qf(n) = \frac{n(n-1)}{2}[f(n-1) - f(n)] + \alpha n[f(n+1) - f(n)]$$

for any $f : \mathbf{N} \to \mathbb{R}$.

**Proposition 1.** *Let $(Y_t)_{t\geq 0}$ given by* (1). *Then for any $n \geq 1$ and $t \geq 0$ we have*

$$\mathbf{E}[Y_t^n | Y_0 = y] = \mathbf{E}[y^{R_t} | R_0 = n], \quad 0 \leq y \leq 1.$$

PROOF : We fix $n \geq 1$ and we consider the function $u : \mathbf{R}_+ \times [0,1] \to \mathbf{R}$ given by

$$u(t,y) = \mathbf{E}(y^{R_t} | R_0 = n), \quad t \geq 0, \quad 0 \leq y \leq 1.$$

Let $f : \mathbf{N} \to \mathbf{R}$. The process $(M_t^f)_{t\geq 0}$ given by

$$M_t^f = f(R_t) - f(R_0) - \int_0^t \left[ \binom{R_s}{2} [f(R_s - 1) - f(R_s)] + \alpha R_s [f(R_s + 1) - f(R_s)] \right] ds \tag{2}$$

is a local martingale. Applying (2) with the particular choice $f(n) = y^n$ for each $y \in [0,1]$, there exists a local martingale $(M_t^{(1)})_{t\geq 0}$ such that $M_0^{(1)} = 0$ and

$$y^{R_t} = y^{R_0} + y(1-y)\int_0^t \left(\frac{1}{2}R_s(R_s-1)y^{R_s-2} - \alpha R_s y^{R_s-1}\right) ds + M_t, \quad t \geq 0. \quad (3)$$

Applying (2) with $f(n) = y^{2n}$ and comparing with Itô formula for the square of $y^{R_t}$, we deduce that

$$< M >_t = (y-1)^2 \int_0^t y^{2R_s-2}\left[\binom{R_s}{2} + \alpha R_s y^2\right] ds.$$

Moreover, $M_t$ is in fact a square integrable martingale. Indeed, by a natural coupling, one may stochastically upper bound the birth and death process $\{R_t; t \geq 0\}$ by the Yule process $\{Z_t; t \geq 0\}$ issued from $R_0$, which jumps from $k$ to $k+1$, at the birth times of $\{R_t; t \geq 0\}$. It is easy to show that each term in (2) is integrable.
Taking the conditional expectation $\mathbf{E}(\cdot|R_0 = n)$, we deduce from (3)

$$\begin{aligned} u(t,y) &= u(0,y) + y(1-y)\int_0^t \mathbf{E}\left(\frac{1}{2}R_s(R_s-1)y^{R_s-2} - \alpha R_s y^{R_s-1}|R_0 = n\right) ds \\ &= u(0,y) + y(1-y)\int_0^t \left(\frac{1}{2}\frac{\partial^2 u}{\partial y^2}(s,y) - \alpha\frac{\partial u}{\partial y}(s,y)\right) ds. \end{aligned}$$

Hence $u$ solves the following linear parabolic PDE

$$\begin{cases} \partial_t u(t,y) = \dfrac{1}{2}y(1-y)\partial^2_{yy}u(t,y) - \alpha y(1-y)\partial_y u(t,y) \quad t \geq 0, \quad 0 < y < 1, \\ u(0,y) = y^n, \quad u(t,0) = 0, \quad u(t,1) = 1. \end{cases} \quad (4)$$

It is easy to check that $u$ is of class $C^{1,2}(\mathbf{R} \times (0,1))$. Itô's formula applied to the function $(s,y) \longmapsto u(t-s,y)$ yields

$$\begin{aligned} &u(0,Y_t) = u(t,Y_0) + \int_0^t \frac{\partial u}{\partial x}(t-r,Y_r)\sqrt{Y_r(1-Y_r)}\, dB_r \\ &+ \int_0^t \left[-\frac{\partial u}{\partial s}(t-r,Y_r) - \alpha X_r(1-X_r)\frac{\partial u}{\partial x}(t-r,Y_r) + \frac{1}{2}Y_r(1-Y_r)\frac{\partial^2 u}{\partial x^2}(t-r,Y_r)\right] dr \end{aligned}$$

Using (4), we deduce that

$$u(0,Y_t) = u(t,Y_0) + N_t$$

where $(N_t)_{t \geq 0}$ is a zero-mean martingale. It remains to take the expectation in the last identity to get the desired result. □

*Remark 1.* Strong uniqueness of (1) is well known. Weak uniqueness follows from that result as well as from the duality argument in Proposition 1.

# 3 Look-down with selection, exchangeability

## 3.1 Construction of our process

We consider the look-down model with selection defined in the introduction. We first need to give a construction of our $\{\eta_t(i),\ i \geq 1,\ t \geq 0\}$. For each $N$, consider the process $\{\eta_t^N(i),\ i \geq 1,\ t \geq 0\}$, obtained by applying only the arrows between $1 \leq i < j \leq N$, and the crosses on levels 1 to $N$. In other words, we disregard all the arrows pointing to levels above $N$, as well as all the crosses on levels above $N$. We then have a finite number of arrows and crosses on any finite time interval, and $\{\eta_t^N(i),\ i \geq 1,\ t \geq 0\}$ is constructed in an obvious way, by implementing the effect of the arrows and crosses, in the order in which they are met.

In the rest of this sub–section, we refer to $\eta^N$ as the just defined process. It follows from the Borel–Cantelli Lemma and the next Proposition that for $N$ large enough (depending upon $\omega$)$\{(\eta_t^{2N+k}(1),\dots,\eta_t^{2N+k}(N)),\ t \geq 0\}$ does depend upon $k \geq 1$, hence $\eta^N$ converges to a limit $\eta$ as $N \to \infty$.

**Proposition 2.** *There exists a constant C such that*

$$\mathbf{P}\left(\exists 1 \leq i \leq N, k \geq 1, t > 0 \textit{ such that } \eta_t^{2N}(i) \neq \eta_t^{2N+k}(i)\right) \leq Ce^{-N^3/4}.$$

PROOF : For each $i \geq 1, t > 0$, let $\xi_t^{i,2N}$ denote the level on which the individual who was sitting on level $i$ at time $t = 0$ sits at time $t$, where the evolution corresponds to the "$2N$–model", i. e. all arrows pointing to levels above $2N$, and all crosses on levels above $2N$ have been erased. Each time there is a birth on a level smaller that or equal to $\xi_{t^-}^{i,2N}$, $\xi_t^{i,2N}$ has a jump of size +1. Each time there is a death on a level smaller than or equal to $\xi_{t^-}^{i,2N}$, $\xi_t^{i,2N}$ has a jump of size -1. In other words, $\xi_t^{i,2N}$ follows the position of the individual who was sitting on level $i$ at time $t = 0$ until his possible death, then follows the position of his left neighbor, etc.. We have

$$\begin{aligned}
&\left\{\exists 1 \leq i \leq N, k \geq 1, t > 0 \text{ such that } \eta_t^{2N}(i) \neq \eta_t^{2N+k}(i)\right\} \\
&\subset \left\{\exists i \geq 1, 0 \leq t < t' \text{ such that } \xi_t^{i,2N} > 2N,\ \xi_{t'}^{i,2N} = N\right\} \\
&\subset \left\{\exists 1 \leq i \leq 2N+1, 0 \leq t < t' \text{ such that } \xi_t^{i,2N} > 2N,\ \xi_{t'}^{i,2N} = N\right\}.
\end{aligned}$$

In other words, for the crosses and arrows on levels higher that $2N$ to interfere with the behavior of the population at levels 1 to $N$, we need that at least one individual visit the level $N$, after having visited the level $2N+1$, and the second inclusion follows from the following monotonicity property : $i < j \Rightarrow \xi_t^{i,2N} \leq \xi_t^{j,2N}$ a. s. for all $t > 0$. Consequently

$$
\begin{aligned}
&\mathbf{P}\left(\exists 1 \le i \le N, k \ge 1, t > 0 \text{ such that } \eta_t^{2N}(i) \neq \eta_t^{2N+k}(i)\right) \\
&\qquad \le \sum_{i=1}^{2N+1} \mathbf{P}\left(\exists 0 \le t < t' \text{ such that } \xi_t^{i,2N} > 2N,\ \xi_{t'}^{i,2N} = N\right).
\end{aligned} \tag{5}
$$

We first show that there exists $C > 0$ such that

$$
\mathbf{P}\left(\exists t > 0 \text{ such that } \xi_t^{2N+1,2N} = N\right) \le \frac{C}{N} e^{-N^3/4}. \tag{6}
$$

We can couple the process $\xi_t^{2N+1,2N}$ with a birth and death process $\rho_t^N$, with birth rate $N(N+1)/2$ and death rate $\alpha(2N+1)$, with the properties

$$
\rho_0^N = 2N+1,\ \rho_t^N \le \xi_t^{2N+1,2N},\ 0 \le t \le \tau_N,
$$

where

$$
\tau_N = \inf\{t > 0,\ \rho_t^N = N\}.
$$

Clearly

$$
\mathbf{P}(\exists t > 0 \text{ such that } \xi_t^{2N+1,2N} = N) \le \mathbf{P}(\tau_N < \infty),
$$

hence (6) follows from

**Lemma 1.** *There exists a constant $C > 0$ such that*

$$
\mathbf{P}(\tau_N < \infty) \le \frac{C}{N} e^{-N^3/4}.
$$

PROOF : Let $\{X_n,\ n \ge 1\}$ and $\{Y_n,\ n \ge 1\}$ be two mutually independent sequences of i. i. d. r. v.'s, the $X_n$'s being exponential with parameter bounded from below by $N^2/2$, the $Y_n$'s being exponential with parameter bounded from above by $2\alpha N$. We have

$$
\mathbf{P}(\tau_N < \infty) \le \sum_{n=1}^{\infty} \mathbf{P}(X_1 + \cdots + X_n > Y_1 + \cdots + Y_n + N).
$$

Now

$$
\begin{aligned}
&\mathbf{P}(X_1 + \cdots + X_n > Y_1 + \cdots + Y_n + N) \\
&\qquad = \mathbf{P}\left(\exp[N^2(X_1 + \cdots + X_n - Y_1 - \cdots - Y_n)/4] > e^{N^3/4}\right) \\
&\qquad \le e^{-N^3/4}\left(\mathbf{E}\left[e^{N^2 X_1/4}\right]\mathbf{E}\left[e^{-N^2 Y_1/4}\right]\right)^n \\
&\qquad \le e^{-N^3/4}\left(\frac{4\alpha N}{2\alpha N + N^2/4}\right)^n \\
&\qquad \le e^{-N^3/4}\left(\frac{16\alpha}{N}\right)^n.
\end{aligned}
$$

Summing from $n = 1$ to $\infty$ yields the result of the Lemma. □

We can now conclude the proof of Proposition 2. We note that with the same constant as that appearing in (6), for any $1 \le i \le N$,

$$\mathbf{P}\left(\exists 0 < t < t' \text{ such that } \xi_t^{i,2N} = 2N+1, \xi_{t'}^{i,2N} = N\right) \le \frac{C}{N} e^{-N^3/4}.$$

Indeed, wait until $\theta_{i,2N} = \inf\{t > 0,\ \xi_t^{i,2N} = 2N+1\}$, which is a stopping time at which the Markov process $\{\eta_t^{2N}(j),\ j \ge 1\}_{t\ge 0}$ starts afresh, and then use the same argument as that of Lemma 1. Consequently each term in the right hand side of (5) can be estimated as the last one in (6), hence the Proposition. □

From now on, we equip the probability space $(\Omega, \mathscr{F}, \mathbf{P})$ with the filtration defined by $\mathscr{F}_t = \sigma\{\eta_s(i),\ i \ge 1,\ 0 \le s \le t\}$. Any stopping time will be defined with respect to that filtration.

### 3.2 Exchangeability

Our goal in this subsection is to show that for all $t > 0$, the sequence $\{\eta_t(i),\ i \ge 1\}$ is exchangeable. It in fact suffices to show that for all $t > 0$, any $n \ge 1$, $\eta_t^n := (\eta_t(1), \dots, \eta_t(n))$ is an exchangeable sequence of $\{0,1\}-$ valued r. v.'s.

For any $t \ge 0$, $n \ge 1$, $\eta_t^n$ is a $\{0,1\}^n-$valued random vector. Let $S_n$ denote the group of permutations of $\{1, 2, \dots, n\}$.

For $\pi \in S_n$ and $a^n = (a_i)_{1\le i\le n} \in \{0,1\}^n$, we define the vectors

$$\pi^{-1}(a^n) = (a_{\pi^{-1}(1)}, \dots, a_{\pi^{-1}(n)}) = (a_i^\pi)_{1\le i\le n}$$
$$\pi(\eta_t^n) = (\eta_t(\pi(1)), \dots, \eta_t(\pi(n)))$$

We should point out that $\pi(\eta_t^n)$ is a permutation of $(\eta_t(1), \dots, \eta_t(n))$ and it is clear from the definitions that

$$\{\pi(\eta_t^n) = a^n\} = \{\eta_t^n = \pi^{-1}(a^n)\}, \quad \text{for any} \quad \pi \in S_n. \tag{7}$$

We want to prove the

**Proposition 3.** *Suppose that $\{\eta_0(i), i \ge 1\}$ are i. i. d. random variables. Then for all $t > 0$, $\{\eta_t(i), i \ge 1\}$ is an exchangeable sequence of $\{0,1\}$–valued random variables.*

We first establish two Lemmas.

**Lemma 2.** *For any stopping time $\mathscr{S}$, any $\mathbf{N}$–valued $\mathscr{F}_{\mathscr{S}}$–measurable random variable $\mathbf{n}$, if the random vector $\eta_{\mathscr{S}}^{\mathbf{n}} = (\eta_{\mathscr{S}}(1), \dots, \eta_{\mathscr{S}}(\mathbf{n}))$ is exchangeable, and $T$ is the first time after $\mathscr{S}$ of an arrow pointing to a level $\le \mathbf{n}$, then the random vector $\eta_T^{\mathbf{n}+1} = (\eta_T(1), \dots, \eta_T(\mathbf{n}), \eta_T(\mathbf{n}+1))$ is exchangeable.*

PROOF : For the sake of simplifying the notations, we condition upon $\mathbf{n}=n$ and $T=t$. We start with some notation.

$$A_t^{i,j} := \{\text{The arrow at time } t \text{ is drawn from level } i \text{ to level } j\},\ 1\le i<j\le n.$$

We define

$$\widehat{\mathbf{P}}_{t,n}[.]=\mathbf{P}(.|T=t,\mathbf{n}=n)$$

Thanks to (7), we deduce that, for $\pi\in S_{n+1}$

$$\begin{aligned}\widehat{\mathbf{P}}_{t,n}(\pi(\eta_t^{n+1})=a^{n+1}) &= \sum_{1\le i<j\le n}\widehat{\mathbf{P}}_{t,n}\left(\eta_t^{n+1}=\pi^{-1}(a^{n+1}),A_t^{i,j}\right)\\ &= \sum_{1\le i<j\le n}\widehat{\mathbf{P}}_{t,n}\left(\eta_t(1)=a_1^\pi,\dots,\eta_t(n+1)=a_{n+1}^\pi,A_t^{i,j}\right),\end{aligned}\tag{8}$$

On the event $A_t^{i,j}$, we have :

$$\eta_t(k)=\begin{cases}\eta_{t^-}(k), & \text{if } 1\le k<j\\ \eta_{t^-}(i), & \text{if } k=j\\ \eta_{t^-}(k-1), & \text{if } j<k\le n+1\end{cases}$$

This implies that

$$A_t^{i,j}\cap\{\eta_t^{n+1}=(a_1^\pi,\dots,a_{n+1}^\pi)\}\subset\{a_i^\pi=a_j^\pi\}.$$

For $1\le j\le n$, define the mapping $\rho_j:\{0,1\}^{n+1}\longrightarrow\{0,1\}^n$ by :

$$\rho_j(b_1,\dots,b_{n+1})=(b_1,\dots,b_{j-1},b_{j+1},\dots,b_{n+1}).$$

The second term of the hand right side of (8) is equal to

$$\sum_{1\le i<j\le n}\mathbf{1}_{\{a_i^\pi=a_j^\pi\}}\widehat{\mathbf{P}}_{t,n}\left(\eta_{t^-}^n=\rho_j(\pi^{-1}(a^{n+1})),A_t^{i,j}\right),$$

It is easy to see that the events $(\eta_{t^-}^n=\rho_j(\pi^{-1}(a^{n+1})))$ and $A_t^{i,j}$ are independent. Thus

$$\begin{aligned}\widehat{\mathbf{P}}_{t,n}(\pi(\eta_t^{n+1})=a^{n+1}) &= \sum_{1\le i<j\le n}\mathbf{1}_{\{a_i^\pi=a_j^\pi\}}\widehat{\mathbf{P}}_{t,n}\left(\eta_{t^-}^n=\rho_j(\pi^{-1}(a^{n+1}))\right)\widehat{\mathbf{P}}_{t,n}(A_t^{i,j})\\ &= \frac{2}{n(n-1)}\sum_{1\le i<j\le n}\mathbf{1}_{\{a_i^\pi=a_j^\pi\}}\mathbf{P}\left(\eta_{t^-}^n=\rho_j(\pi^{-1}(a^{n+1}))\right).\end{aligned}$$

If $a_j=a_{\pi(j)}$, $\rho_j(a^{n+1})$ and $\rho_j(\pi^{-1}(a^{n+1}))$ contain the same number of $0's$ and $1's$. This implies that

$$\mathbf{1}_{\{a_i^\pi=a_j^\pi=\gamma\}}\mathbf{P}\left(\eta_{t^-}^n=\rho_j(\pi^{-1}(a^{n+1}))\right)=\mathbf{1}_{\{a_i=a_j=\gamma\}}\mathbf{P}\left(\eta_{t^-}^n=\rho_j(a^{n+1})\right),\quad \gamma\in\{0,1\}$$

On the other hand, we have $\#\{1 \le i < j \le n : a_i^\pi = a_j^\pi\} = \#\{1 \le i < j \le n : a_i = a_j\}$. Let $k = \pi(i) \wedge \pi(j)$ and $\ell = \pi(i) \vee \pi(j)$. If $a_i = a_j$, then we have $a_k^\pi = a_\ell^\pi = a_i = a_j$. Finally, we obtain

$$\begin{aligned}
\widehat{\mathbf{P}}_{t,n}(\pi(\eta_t^{n+1}) = a^{n+1}) &= \frac{2}{n(n-1)} \sum_{1 \le k < \ell \le n} \mathbf{1}_{\{a_k^\pi = a_\ell^\pi\}} \mathbf{P}\left(\eta_{t^-}^n = \rho_\ell(\pi^{-1}(a^{n+1})\right) \\
&= \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} \mathbf{1}_{\{a_i = a_j\}} \mathbf{P}\left(\eta_{t^-}^n = \rho_j(a^{n+1})\right) \\
&= \widehat{\mathbf{P}}_{t,n}(\eta_t^{n+1} = a^{n+1})
\end{aligned}$$

We have proved that for any $\pi \in S_{n+1}$ and $t \ge 0$, $\widehat{\mathbf{P}}_{t,n}(\pi(\eta_t^{n+1}) = a^{n+1}) = \widehat{\mathbf{P}}_{t,n}(\eta_t^{n+1} = a^{n+1})$. The result follows.

**Lemma 3.** *For any stopping time $\mathscr{S}$, any $\mathbf{N}$–valued $\mathscr{F}_{\mathscr{S}}$–measurable random variable $\mathbf{n}$, if the random vector $\eta_{\mathscr{S}}^{\mathbf{n}} = (\eta_{\mathscr{S}}(1), \ldots, \eta_{\mathscr{S}}(\mathbf{n}))$ is exchangeable, and $T$ is the first time after $\mathscr{S}$ of a death at a level $\le \mathbf{n}$, then the random vector $\eta_T^{\mathbf{n}-1} = (\eta_T(1), \ldots, \eta_T(\mathbf{n}-1))$ is exchangeable.*

PROOF : For the sake of simplifying the notations, we condition upon $\mathbf{n} = n$ and $T = t$. Let $\pi \in S_{n-1}$ be arbitrary. We consider the events :

$$B_t^i := \{\text{the level of the dying individual at time } t \text{ is } i\}.$$

Let $\widehat{\mathbf{P}}_{t,n}[.] = \mathbf{P}(.|T = t, \mathbf{n} = n)$. Using eq. (7) we deduce that

$$\begin{aligned}
\widehat{\mathbf{P}}_{t,n}(\pi(\eta_t^{n-1}) = a^{n-1}) &= \sum_{1 \le i \le n} \widehat{\mathbf{P}}_{t,n}\left(\eta_t^{n-1} = \pi^{-1}(a^{n-1}), B_t^i\right) \\
&= \sum_{1 \le i \le n} \widehat{\mathbf{P}}\left(\eta_t(1) = a_1^\pi, \ldots, \eta_t(n-1) = a_{n-1}^\pi, B_t^i\right).
\end{aligned}$$

Define

$$c_i^{\pi,n} = (a_1^\pi, \ldots, a_{i-1}^\pi, 1, a_i^\pi, \ldots, a_{n-1}^\pi), \qquad c_i^n = (a_1, \ldots, a_{i-1}, 1, a_i, \ldots, a_{n-1}).$$

Using the property of the look-down with selection, the last term in the previous relation is equal to

$$\begin{aligned}
\sum_{1 \le i \le n} \widehat{\mathbf{P}}\left(\eta_{t^-}^n = c_i^{\pi,n}, B_t^i\right) &= \sum_{1 \le i \le n} \widehat{\mathbf{E}}_{t,n}\left(\mathbf{1}_{\{\eta_{t^-}^n = c_i^{\pi,n}\}} \mathbf{1}_{B_t^i}\right) \\
&= \sum_{1 \le i \le n} \mathbf{P}\left(\eta_{t^-}^n = c_i^{\pi,n}\right) \widehat{\mathbf{P}}_{t,n}\left(B_t^i \mid \eta_{t^-}^n = c_i^{\pi,n}\right) \\
&= \frac{1}{1 + \sum_{j=1}^n a_j^\pi} \sum_{1 \le i \le n} \mathbf{P}\left(\eta_{t^-}^n = c_i^{\pi,n}\right).
\end{aligned}$$

Thanks to the exchangeability of $(\eta_{t^-}(1), \ldots, \eta_{t^-}(n))$, we have

$$\widehat{\mathbf{P}}_{t,n}(\pi(\eta_t^{n-1}) = a^{n-1}) = \frac{1}{1+\sum_{j=1}^n a_j} \sum_{1\le i\le n} \mathbf{P}\left(\eta_{t^-}^n = c_i^n\right)$$

since $\sum_{j=1}^{n-1} a_j^\pi = \sum_{j=1}^{n-1} a_j$ and $c_i^{\pi,n}$ is a permutation of $c_i^\pi$. The result follows. □

We can now proceed with the

PROOF OF PROPOSITION 3 For each $N \ge 1$, let $\{V_t^N,\ t \ge 0\}$ denote the $\mathbf{N}$–valued process which describes the position at time $t$ of the individual sitting on level $N$ at time 0, with the convention that, if that individual dies, we replace him by his left neighbor. When $V_t^N = k$, $V_t^N$ is shifted to $k+1$ at rate $k(k-1)/2$, and shifted to $k-1$ at rate $\alpha(\eta_t(1)+\cdots+\eta_t(k))$. Since the rate of increase is quadratic in $k$, and the rate of decrease in bounded from above by $\alpha k$, the same argument which insures that Kingman's coalescent comes down from infinity (see [6]) and that the ARG also comes down from infinity (see [7]), implies that $V_t^N \to \infty$ in finite time ($N$ fixed). Moreover $\inf_{t\ge 0} V_t^N \to \infty$, as $N \to \infty$.

It follows from Lemma 2 and 3 that for each $t > 0$, $N \ge 1$, $(\eta_t(1),\ldots,\eta_t(V_t^N))$ is an exchangeable random vector.

Consequently, for any $t > 0$, $n \ge 1$, $\pi \in S_n$, $a^n \in \{0,1\}$,

$$|\mathbf{P}(\eta_t^n = a^n) - \mathbf{P}(\eta_t^n = \pi^{-1}(a^n))| \le \mathbf{P}(V_t^N < n),$$

which goes to zero, as $N \to \infty$. The result follows □

*Remark 2.* The collection of random process

$$\{\eta_t(i), t \ge 0\}_{i\ge 1} \text{ is not exchangeable.}$$

Indeed, $\eta_t(1)$ can jump from 1 to 0, but never from 0 to 1, while the other $\eta_t(i)$ do not have that property.

For $N \ge 1$ and $t \ge 0$, denote by $X_t^N$ the proportion of type $b$ individuals at time $t$ among the first $N$ individuals, i.e.

$$X_t^N = \frac{1}{N}\sum_{i=1}^N \eta_t(i) \tag{9}$$

We are interested in the asymptotic properties of $(X_t^N)_{t\ge 0}$ as $N$ tends to infinity. For this, let us recall the following useful result due to de Finetti (see e. g. [1] ).

**Theorem 1.** *An exchangeable (countably infinite) sequence $\{X_n, n \ge 1\}$ of random variables is a mixture of i.i.d. sequences, in the sense that conditionnally upon $\mathscr{G}$ (the tail $\sigma$–field of the sequence $\{X_n, n \ge 1\}$) the $X_n$'s are i.i.d.*

As a consequence, we have the following asymptotic property for fixed $t$ of the sequence $(X_t^N)_{N\ge 1}$ defined by (9).

**Corollary 1.** *For each $t \ge 0$,*

$$X_t = \lim_{N\to\infty} X_t^N \quad \textit{exist a.s.} \tag{10}$$

PROOF : Let $t \geq 0$ and $n \geq 1$. Let us introduce the filtration $\mathscr{F}_n = \sigma(\eta_t(n+1), \eta_t(n+2), \dots)$. We have (here "converges" means "converges as $N \to \infty$")

$$\mathbf{P}\left(N^{-1}\sum_{i=1}^{N}\eta_t(i) \text{ converges}\right) = \mathbf{E}\left[\mathbf{P}\left(N^{-1}\sum_{i=1}^{N}\eta_t(i) \text{ converges}|\bigcap_{n=0}^{\infty}\mathscr{F}_n\right)\right]$$

From Proposition 3 and theorem 1, conditionally upon $\bigcap_{n=0}^{\infty}\mathscr{F}_n$, $\eta_t(i), i \geq 1$ are i.i.d. random variables. Thanks to the law of large numbers, $N^{-1}\sum_{i=1}^{N}\eta_t(i)$ converge a.s. as $N \to \infty$. This implies

$$\mathbf{P}\left(N^{-1}\sum_{i=1}^{N}\eta_t(i) \text{ converges}|\bigcap_{n=0}^{\infty}\mathscr{F}_n\right) = 1,$$

which establishes the desired result. □

# 4 Convergence to the Wright-Fisher diffusion with selection

## *4.1 Preliminary results*

Before stating the main theorem of this section, let us establish some auxiliary results which we shall need in its proof.

**Proposition 4.** *Let $\{\xi_1, \xi_2, \dots, \}$ be a countable exchangeable sequence of $\{0,1\}$-valued random variables and $\mathscr{G}$ denote its tail $\sigma$-field. Let $\mathscr{H}$ be some additional $\sigma$-algebra. If conditionaly upon $\mathscr{G} \vee \mathscr{H}$, the $\xi_i$'s are exchangeable, then their conditional law given $\mathscr{G} \vee \mathscr{H}$ is their conditional law given $\mathscr{G}$.*

PROOF : Let $n \geq 1$ and $f : \{0,1\}^n \to \mathbf{R}$ be an arbitrary mapping. It follows from the assumption that

$$\begin{aligned}\mathbf{E}(f(\xi_1, \dots, \xi_n)|\mathscr{G} \vee \mathscr{H}) &= \mathbf{E}\left(N^{-1}\sum_{k=1}^{N} f(\xi_{(k-1)n+1}, \dots, \xi_{kn})|\mathscr{G} \vee \mathscr{H}\right) \\ &= \mathbf{E}[f(\xi_1, \dots, \xi_n)|\mathscr{G}],\end{aligned}$$

where the second equality follows from the fact that the quantity inside the previous conditionally expectation converges a.s. to $\mathbf{E}[f(\xi_1, \dots, \xi_n)|\mathscr{G}]$ as $N \to \infty$, as a consequence of exchangeability and de Finetti's theorem. □

Let us look backwards from time $s$ to time 0. For each $0 \leq r \leq s$, we denote by $Z_r^{N,s}$ the highest level occupied by the ancestors at time $r$ of the $N$ first individuals at time $s$. We have

**Lemma 4.** *For any $0 < r-h \le r$, $0 < c < (1-e^{-1})/2$, $N$ large enough (depending only upon $c$ and $\alpha$),*

$$\mathbf{P}(Z^{N,r}_{r-h} > N) \le e^{-chN^2}$$

PROOF : Let $X^N_h$ (resp. $Y^N_h$) denote the number of birth (resp. death ) events between time $r-h$ and time $r$ on leveles $1,2,\dots,N$. $X^N_h$ is Poisson with parameter $N(N-1)h/2$. Moreover there is a random variable $Y'^N_h$, $Y'^N_h$ being Poisson with parameter $\alpha Nh$ and independent of $X^N_h$, such that $Y^N_h \le Y'^N_h a.s.$ Indeed, $Y'^N_h$ represents the number of deaths between time $r$ and time $r-h$ when all the individuals sitting on positions $\{1,\dots,N\}$ are $b$. We have

$$\begin{aligned}
\mathbf{P}(X_h < Y_h) &\le \mathbf{P}(X_h < Y'_h)\\
&= \sum_{k=0}^{\infty} \mathbf{P}(Y'_h > k)\mathbf{P}(X_h = k)\\
&\le \mathbf{E}[e^{Y'_h}] \sum_{k=0}^{\infty} \frac{(N(N-1)h/2e)^k}{k!} e^{-N(N-1)h/2}\\
&= \exp\left(-\frac{N(N-1)}{2}h(1-e^{-1}) + \alpha Nh(e-1)\right)\\
&\le e^{-cN^2h}
\end{aligned}$$

for $N$ large enough, provided $0 < c < (1-e^{-1})/2$. The result follows. □

If $\pi \in S_n, a \in \{0,1\}^n$, we shall write $\pi^*(a) = (a_{\pi(1)},\dots,a_{\pi(n)})$. Recall that a partition $P$ of $\{1,\dots,n\}$ induces an equivalence relation, whose equivalence classes are the blocks of the partition. Hence we shall write $i \simeq_P j$ whenever $i$ and $j$ are in the same block of $P$. Finally we write $\#P$ for the number of blocks of the partition $P$.

We have the

**Proposition 5.** *For all $N \ge 1$, $k \ge 1$, $0 \le r_k < r_{k-1} < \cdots < r_1 < r_0 = s$, $a \in \{0,1\}^N$, $p_1, p_2, \dots, p_k$ such that $0 \le Np_1, Np_2, \dots, Np_k \le N$ is an integer, $\pi \in S_N$,*

$$\begin{aligned}
&\mathbf{P}\left(\eta^N_s = a, \bigcap_{i=1}^{k}\{X^N_{r_i} = p_i\}, \bigcap_{i=1}^{k}\{Z^{N,r_{i-1}}_{r_i} \le N\}\right)\\
&\qquad = \mathbf{P}\left(\eta^N_s = \pi^*(a), \bigcap_{i=1}^{k}\{X^N_{r_i} = p_i\}, \bigcap_{i=1}^{k}\{Z^{N,r_{i-1}}_{r_i} \le N\}\right).
\end{aligned}$$

PROOF :

We prove this in the case $k=2$, the case $k>2$ is similar.

For all $a^1 \in \{0,1\}^N$, we denote by $\mathscr{P}_{a^1}$ the set of partitions $P$ of $\{1,2,\dots,N\}$ which are such that $i \simeq_P j \Rightarrow a^1_i = a^1_j$.

For any $i = 1,2$, let $\eta^N_{r_{i-1}} = a^1$, where $a^1 \in \{0,1\}^N$. let $P^i \in \mathscr{P}_{a^1}$ be a partition of $\{1,\dots,N\}$. This partition represents the genealogy at time $r_i$ of the individuals sitting on positions $\{1,\dots,N\}$ at time $r_{i-1}$. $|P^i|$, the number of blocks of the partition

$P^i$, is the number of ancestor at time $r_i$ of the individuals $\{1,\dots,N\}$ at time $r_{i-1}$. Each block of the partition $P^i$ is a subset of $\{1,\dots,N\}$ consisting of those individuals who have the same ancestor at time $r_i$. We assume that the blocks of $P^i$ are arranged in increasing order of their smallest element.

For $1 \le j \le |P^i|$, let us define

$$c_j^{P^i} = \begin{cases} 1, & \text{if the } j-\text{th block of } P^i \text{ consist of type } b \text{ individuals} \\ 0, & \text{otherwise .} \end{cases}$$

Let

$$I^i = \{\ell_1^{I^i},\dots,\ell_{|P^i|}^{I^i}\}, \qquad J(I^i) = \{1 \le j < \ell_{|P^i|}^{I^i};\ j \notin \{\ell_1^{I^i},\dots,\ell_{|P^i|}^{I^i}\}\},$$

where $i = 1,2$ and $1 \le \ell_1^{I^i} < \ell_2^{I^i} < \cdots < \ell_{|P^i|}^{I^i}$ denote the levels of the $|P^i|$ ancestors at time $r_i$. Note that on the set $\{Z_{r_{i-1}}^{N,r_i} \le N\}, |P^i| \le \ell_{|P^i|}^{I^i} \le N$.

For any $i = 1,2$, we define

$$\mathscr{H}_{P^i} = \Big\{(\ell_1,\dots,\ell_{|P^i|});\ 1 \le \ell_1 < \cdots < \ell_{|P^i|} \le N\Big\},$$

$$\mathscr{A}_{r_i,r_{i-1}}(a,P^i,I^i,p_i) = \{b \in \{0,1\}^N : \sum_{k=1}^N b_k = Np_i, \forall 1 \le j \le |P^i|, b_{\ell_j^{I^i}} = c_j^{P^i}, \forall j \in J(I^i), b_j = 1\}.$$

Note that the set $\mathscr{H}_{P^i}$ depends only upon $|P^i|$. Consider the event

$$A_{I^i}^{P^i} := \{\text{the succession of births and deaths between } r \text{ and } s \text{ produces } P^i \text{ and } I^i\}.$$

For any $i = 1,2$ and $a^1 \in \{0,1\}^N$, we then have

$$\{\eta_{r_{i-1}}^N = a^1, X_{r_i}^N = p_i, Z_{r_{i-1}}^{N,r_i} \le N\} = \bigcup_{P^i \in \mathscr{P}_a} \bigcup_{I^i \in \mathscr{H}_{P^i}} \bigcup_{b \in \mathscr{A}_{r_i,r_{i_1}}(a,P^i,I^i,p_i)} \{A_{I^i}^{P^i}, \eta_{r_i}^N = b\},$$

from which, we deduce that

$$\begin{aligned}
&\{\eta_s^N = a, X_{r_1}^N = p_1, X_{r_2}^N = p_2, Z_{r_1}^{N,s} \le N, Z_{r_2}^{N,r_1} \le N\} \\
&\quad = \bigcup_{P^1 \in \mathscr{P}_a} \bigcup_{I^1 \in \mathscr{H}_{P^1}} \bigcup_{a^1 \in \mathscr{A}_{r_1,s}(a,P^1,I^1,p_1)} \bigcup_{P^2 \in \mathscr{P}_{a1}} \bigcup_{I^2 \in \mathscr{H}_{P^2}} \bigcup_{a^2 \in \mathscr{A}_{r_2,r_1}(a_2,P^2,I^2,p_2)} \{A_{I^1}^{P^1}, A_{I^2}^{P^2}, \eta_{r_2}^N = a^2\}
\end{aligned}$$

and from the independence of $A_{I^1}^{P^1}, A_{I^2}^{P^2}$ and $\eta_{r_2}^N$

$$\begin{aligned}
&\mathbf{P}(\eta_s^N = a, X_{r_1}^N = p_1, X_{r_2}^N = p_2, Z_{r_1}^{N,s} \le N, Z_{r_2}^{N,r_1} \le N) \\
&\quad = \sum_{P^1 \in \mathscr{P}_a} \sum_{I^1 \in \mathscr{H}_{P^1}} \sum_{a^1 \in \mathscr{A}_{r_1,s}(a,P^1,I^1,p_1)} \sum_{P^2 \in \mathscr{P}_{a^1}} \sum_{I^2 \in \mathscr{H}_{P^2}} \sum_{a^2 \in \mathscr{A}_{r_2,r_1}(a^1,P^2,I^2,p_2)} \mathbf{P}(A_{I^1}^{P^1})\mathbf{P}(A_{I^2}^{P^2})\mathbf{P}(\eta_{r_2}^N = a^2)
\end{aligned}$$

Similarly, for any $\pi \in S_N$ and $1 \le i \le 2$, if $\pi^*(P^i)$ is defined by $k \simeq_{P^i} j \Leftrightarrow \pi(k) \simeq_{\pi^*(P^i)} \pi(j)$

$$\begin{aligned}
&\mathbf{P}(\eta_s^N = \pi^*(a), X_{r_1}^N = p_1, X_{r_2}^N = p_2, Z_{r_1}^{N,s} \le N, Z_{r_2}^{N,r_1} \le N) \\
&\quad = \sum_{P^1 \in \mathscr{P}_{\pi^*(a)}} \sum_{I^1 \in \mathscr{H}_{P^1}} \sum_{a^1 \in \mathscr{A}_{r_1,s}(\pi^*(a),P^1,I^1,p_1)} \sum_{P^2 \in \mathscr{P}_{a^1}} \sum_{I^2 \in \mathscr{H}_{P^2}} \sum_{a^2 \in \mathscr{A}_{r_2,r_1}(a^1,P^2,I^2,p_2)} \mathbf{P}(A_{I^1}^{P^1}) \mathbf{P}(A_{I^2}^{P^2}) \mathbf{P}(\eta_{r_2}^N = a^2) \\
&\quad = \sum_{P^1 \in \mathscr{P}_a} \sum_{I^1 \in \mathscr{H}_{P^1}} \sum_{a^1 \in \mathscr{A}_{r_1,s}(\pi^*(a),\pi^*(P^1),I^1,p_1)} \sum_{P^2 \in \mathscr{P}_{a^1}} \sum_{I^2 \in \mathscr{H}_{P^2}} \sum_{a^2 \in \mathscr{A}_{r_2,r_1}(a^1,P^2,I^2,p_2)} \mathbf{P}(A_{I^1}^{\pi^*(P^1)}) \mathbf{P}(A_{I^2}^{P^2}) \mathbf{P}(\eta_{r_2}^N = a^2)
\end{aligned}$$

For any $i = 1,2$ and $a^1 \in \{0,1\}^N$, we now describe a one-to-one correspondence $\rho_\pi$ between $\mathscr{A}_{r_i,r_{i-1}}(a^1,P^i,I^i,p_i)$ and $\mathscr{A}_{r_i,r_{i-1}}(\pi^*(a^1),\pi^*(P^i),I^i,p_i)$. Let $b \in \mathscr{A}_{r_i,r_{i-1}}(a^1,P^i,I^i,p_i)$. We define $b^{'} = \rho_\pi(b)$ as follows.

$$b_j^{'} = \begin{cases} 1, & \text{if } j \in J(I^i) \\ b_j, & \text{if } j > l_k^{I^i} \end{cases},$$

and

$$b^{'}_{\ell_j^{I^i}} = c_j^{\pi^*(P^i)}, \text{ for all } 1 \le j \le |P^i|.$$

It is plain that $\sum_{j=1}^N b_j = \sum_{j=1}^N b_j^{'}$. Clearly there exists $\pi' \in S_N$ such that $b^{'} = \pi'^*(b)$. Moreover, for any $\pi \in S_N$ and $i = 1,2$, $\mathbf{P}(A_{I^i}^{P^i}) = \mathbf{P}(A_{I^i}^{\pi^*(P^i)})$

Consequently

$$\begin{aligned}
&\sum_{a^1 \in \mathscr{A}_{r_1,s}(\pi^*(a),\pi^*(P^1),I^1,p_1)} \sum_{P^2 \in \mathscr{P}_{a^1}} \sum_{I^2 \in \mathscr{H}_{P^2}} \sum_{a^2 \in \mathscr{A}_{r_2,r_1}(a^1,P^2,I^2,p_2)} \mathbf{P}(A_{I^1}^{\pi^*(P^1)}) \mathbf{P}(A_{I^2}^{P^2}) \mathbf{P}(\eta_{r_2}^N = a^2) \\
&= \sum_{a^1 \in \mathscr{A}_{r_1,s}(a,P^1,I^1,p_1)} \sum_{P^2 \in \mathscr{P}_{a^1}} \sum_{I^2 \in \mathscr{H}_{P^2}} \sum_{a^2 \in \mathscr{A}_{r_2,r_1}(\pi'(a^1),\pi'^*(P^2),I^2,p_2)} \mathbf{P}(A_{I^1}^{\pi^*(P^1)}) \mathbf{P}(A_{I^2}^{\pi'^*(P^2)}) \mathbf{P}(\eta_{r_2}^N = a^2) \\
&= \sum_{a^1 \in \mathscr{A}_{r_1,s}(a,P^1,I^1,p_1)} \sum_{P^2 \in \mathscr{P}_{a^1}} \sum_{I^2 \in \mathscr{H}_{P^2}} \sum_{a^2 \in \mathscr{A}_{r_2,r_1}(a^1,P^2,I^2,p_2)} \mathbf{P}(A_{I^1}^{P^1}) \mathbf{P}(A_{I^2}^{P^2}) \mathbf{P}(\eta_{r_2}^N = \pi''(a^2)) \\
&= \sum_{a^1 \in \mathscr{A}_{r_1,s}(a,P^1,I^1,p_1)} \sum_{P^2 \in \mathscr{P}_{a^1}} \sum_{I^2 \in \mathscr{H}_{P^2}} \sum_{a^2 \in \mathscr{A}_{r_2,r_1}(a^1,P^2,I^2,p_2)} \mathbf{P}(A_{I^1}^{P^1}) \mathbf{P}(A_{I^2}^{P^2}) \mathbf{P}(\eta_{r_2}^N = a^2)
\end{aligned}$$

Where the last identity follows from the fact $\eta_r^N$ is exchangeable. The result follows. □

### 4.2 *Tightness of* $(X_t^N)_{t \ge 0}$

Before we establish tightness, we collect some results which will required for its proof.

**Lemma 5.** *For any* $0 < r \le h$, $N \ge 1$, $\varphi : [0,1]^3 \to \mathbf{R}$ *Borel measurable, any* $1 \le i < N$,

$$\Big| \mathbf{E}(\varphi(X^N_{t-h}, X^N_t, X^N_{t+r}); \eta_{t+r}(i) = 0, \eta_{t+r}(N) = 1) - \mathbf{E}(\varphi(X^N_{t-h}, X^N_t, X^N_{t+r}); \eta_{t+r}(i) = 1, \eta_{t+r}(N) = 0) \Big| \le \mathbf{E}(|\varphi(X^N_{t-h}, X^N_t, X^N_{t+r})|; \{Z^{N,t+r}_t > N\} \cup \{Z^{N,t}_{t-h} > N\}).$$

PROOF : Define

$$\begin{aligned} Z &= \varphi(X^N_{t-h}, X^N_t, X^N_{t+r}), \\ A &= \{\eta_{t+r}(i) = 0, \eta_{t+r}(N) = 1\}, \\ B &= \{\eta_{t+r}(i) = 1, \eta_{t+r}(N) = 0\}, \\ C &= \{Z^{N,t+r}_t > N\} \cup \{Z^{N,t}_{t-h} > N\}. \end{aligned}$$

We have shown in Proposition 5 that

$$\mathbf{E}[Z(\mathbf{1}_A - \mathbf{1}_B)] = \mathbf{E}[Z(\mathbf{1}_A - \mathbf{1}_B)\mathbf{1}_C].$$

But

$$|\mathbf{E}[Z(\mathbf{1}_A - \mathbf{1}_B)\mathbf{1}_C]| \le \mathbf{E}[|Z|; C].$$

The result follows. □

**Lemma 6.** *There exists* $a, b > 0$*, which depend only upon the parameter* $\alpha$*, such that for N large enough*

$$\mathbf{P}(\{Z^{N,t+r}_t > N\} \cup \{Z^{N,t}_{t-h} > N\}) \le a\left(Nre^{-bN^2 r} + Nhe^{-bN^2 h}\right).$$

PROOF : We have

$$\mathbf{P}(\{Z^{N,t+r}_t > N\} \cup \{Z^{N,t}_{t-h} > N\}) \le \mathbf{P}(Z^{N,t+r}_t > N) + \mathbf{P}(Z^{N,t}_{t-h} > N).$$

It suffices to estimate the second term. Using the same argument as in the proof of Lemma 4, we have

$$\begin{aligned} \mathbf{P}(X_h < Y_h) \le \mathbf{P}(X_h < Y'_h) &= \sum_{k=0}^{\infty} \mathbf{P}(Y'_h > k)\mathbf{P}(X_h = k) \\ &\le \mathbf{E}[e^{Y'_h}; Y'_h > 0] \sum_{k=0}^{\infty} \frac{(N(N-1)h/2e)^k}{k!} e^{-\mathbf{N}(N-1)h/2} \\ &= \alpha N h \exp\left(-\frac{N(N-1)}{2} h(1-e^{-1}) + \alpha N h(e-1)\right) \\ &\le \alpha N h e^{-cN^2 h} \end{aligned}$$

for $N$ large enough, provided $0 < c < (1-e^{-1})/2$. □

We first deduce from the above estimates with $h=0$

**Corollary 2.** *For any $t,r>0$, $N\ge 1$, $1\le i<N$,*

$$\Big|\mathbf{E}\left((X^N_{t+r}-X^N_t);\eta_{t+r}(i)=0,\eta_{t+r}(N)=1\right) \\ -\mathbf{E}\left((X^N_{t+r}-X^N_t);\eta_{t+r}(i)=1,\eta_{t+r}(N)=0\right)\Big| \\ \le aNr\, e^{-bN^2r}.$$

We now deduce

**Corollary 3.** *If $a'=\sqrt{a}$, $b'=b/2$, where $a$ and $b$ are the constants in the statement of Lemma 6,*

$$\Big|\mathbf{E}\left((X^N_t-X^N_{t-h})^2(X^N_{t+r}-X^N_t);\eta_{t+r}(i)=0,\eta_{t+r}(N)=1\right) \\ -\mathbf{E}\left((X^N_t-X^N_{t-h})^2(X^N_{t+r}-X^N_t);\eta_{t+r}(i)=1,\eta_{t+r}(N)=0\right)\Big| \\ \le a'\left(\sqrt{Nr}\, e^{-b'N^2r}+\sqrt{Nh}\, e^{-b'N^2h}\right)\sqrt{\mathbf{E}\left[(X^N_t-X^N_{t-h})^4\right]}.$$

PROOF : Combining the two above Lemmas with Schwarz's inequality in the form

$$\mathbf{E}(Z;C)\le\sqrt{\mathbf{E}(Z^2)\times\mathbf{P}(C)}$$

yields the result. □

We are going to invoke a tightness criterium which involves an estimate of $\mathbf{E}\left[(X^N_t-X^N_{t-h})^2(X^N_{t+h}-X^N_t)^2\right]$. More precisely, we have

**Proposition 6.** *For any $T>0$, there exists a constant $K>0$ which depends only upon $\alpha$ and $T$ such that for all $N\ge 1$, $0<h<t\le T$,*

$$\mathbf{E}\left[\left(X^N_t-X^N_{t-h}\right)^2\left(X^N_{t+h}-X^N_t\right)^2\right]\le Kh^{5/4}$$

PROOF : We have

$$dX^N_r=\frac{1}{N}\left[\sum_{1\le i<j\le N}\xi^i_{r^-}dP^{i,j}_r+\sum_{1\le i\le N}\theta^i_{r^-}dP^i_r\right],$$

where $P^{i,j}$, $1\le i<j$, $P$, $i\ge 1$ are mutually independent Poisson processes, the $P^{i,j}$'s being standard, and the $P^i$'s having intensity $\alpha$,

$$\xi^i_r=\mathbf{1}_{\{\eta_r(i)=1,\eta_r(N)=0\}}-\mathbf{1}_{\{\eta_r(i)=0,\eta_r(N)=1\}},\\ \theta^i_r=-\mathbf{1}_{\{\eta_r(i)=1,\eta_r(N+1)=0\}}.$$

Let

$$\gamma_t^N := \left(X_t^N - X_{t-h}^N\right)^2.$$

It now follows from Corollary 3

$$\begin{aligned}
\gamma_t^N \left(X_{t+h}^N - X_t^N\right)^2 &= \frac{2}{N}\left[\sum_{1\le i<j\le N}\int_t^{t+h}\gamma_t^N\left(X_{r^-}^N - X_t^N\right)\xi_{r^-}^i\,dP_r^{i,j} + \sum_{1\le i\le N}\int_t^{t+h}\gamma_t^N\left(X_{r^-}^N - X_t^N\right)\theta_{r^-}^i\,dP_r^i\right]\\
&\quad + \frac{1}{N^2}\left[\sum_{1\le i<j\le N}\int_t^{t+h}\gamma_t^N(\xi_{r^-}^i)^2 dP_r^{i,j} + \sum_{1\le i\le N}\int_t^{t+h}\gamma_t^N(\theta_{r^-}^i)^2 dP_r^i\right]\\
\mathbf{E}\left[\gamma_t^N\left(X_{t+h}^N - X_t^N\right)^2\right] &= \frac{2}{N}\mathbf{E}\left[\sum_{1\le i<j\le N}\int_t^{t+h}\gamma_t^N\left(X_r^N - X_t^N\right)\xi_r^i dr + \alpha\sum_{1\le i\le N}\int_t^{t+h}\gamma_t^N\left(X_r^N - X_t^N\right)\theta_r^i dr\right]\\
&\quad + \frac{1}{N^2}\mathbf{E}\left[\sum_{1\le i<j\le N}\int_t^{t+h}\gamma_t^N(\xi_r^i)^2 dr + \sum_{1\le i\le N}\int_t^{t+h}\gamma_t^N(\theta_r^i)^2 dr\right]\\
&\le a'\left[N^{3/2}\int_0^h\sqrt{r}e^{-b'N^2r}dr + N^{3/2}h^{3/2}e^{-b'N^2h}\right]\sqrt{\mathbf{E}\left[(\gamma_t^N)^2\right]} + C_\alpha h\mathbf{E}(\gamma_t^N),
\end{aligned}$$

with $C_\alpha = 3\alpha + 1/2$. But from Hölder's inequality with $p = 4$, $q = 4/3$,

$$\begin{aligned}
N^{3/2}\int_0^h\sqrt{r}e^{-b'N^2r}dr &\le N^{3/2}\left(\int_0^h r^2\right)^{1/4}\left(\int_0^h e^{-4b'N^2r/3}dr\right)^{3/4}\\
&\le C\frac{N^{3/2}}{N^{3/2}}h^{3/4} \le Ch^{3/4},
\end{aligned}$$

for some $C > 0$, while

$$N^{3/2}h^{3/2}e^{-b'N^2h} \le C'h^{3/4},$$

with $C' = \sup_{x>0} x^{3/4}e^{-b'x} < \infty$. We have shown that

$$\mathbf{E}\left[\gamma_t^N\left(X_{t+h}^N - X_t^N\right)^2\right] \le C''h^{3/4}\sqrt{\mathbf{E}\left[(\gamma_t^N)^2\right]} + C_\alpha h\mathbf{E}(\gamma_t^N). \tag{11}$$

It remains to estimate $\mathbf{E}(\gamma_t^N)$. The computations are quite similar to the previous ones, but simpler. We use Corollary 2, but with the interval $[t, t+h]$ replaced by the interval $[t-h, t]$.

$$\begin{aligned}
\left(X_t^N - X_{t-h}^N\right)^2 &= \frac{2}{N}\left[\sum_{1\le i<j\le N}\int_{t-h}^t \left(X_{r^-}^N - X_{t-h}^N\right)\xi_{r^-}^i\, dP_r^{i,j} + \sum_{1\le i\le N}\int_{t-h}^t \left(X_{r^-}^N - X_{t-h}^N\right)\theta_{r^-}^i\, dP_r^i\right]\\
&\quad + \frac{1}{N^2}\left[\sum_{1\le i<j\le N}\int_{t-h}^t (\xi_{r^-}^i)^2 dP_r^{i,j} + \sum_{1\le i\le N}\int_{t-h}^t (\theta_{r^-}^i)^2 dP_r^i\right]\\
\mathbf{E}\left[\left(X_t^N - X_{t-h}^N\right)^2\right] &= \frac{2}{N}\mathbf{E}\left[\sum_{1\le i<j\le N}\int_{t-h}^t \left(X_r^N - X_{t-h}^N\right)\xi_{r^-}^i\, dr + \sum_{1\le i\le N}\int_{t-h}^t \left(X_r^N - X_{t-h}^N\right)\theta_{r^-}^i\, dr\right]\\
&\quad + \frac{1}{N^2}\mathbf{E}\left[\sum_{1\le i<j\le N}\int_{t-h}^t (\xi_r^i)^2 dr + \sum_{1\le i\le N}\int_{t-h}^t (\theta_r^i)^2 dr\right]\\
&\le aN^2\int_0^h re^{-bN^2 r}dr + C_\alpha h\\
&\le \frac{a}{b}h + C_\alpha h\\
&\le C''' h.
\end{aligned}$$

Moreover

$$\mathbf{E}\left[(\gamma_t^N)^2\right] \le \mathbf{E}[\gamma_t^N] \le C''' h.$$

The result follows if we combine this last estimate with (11), keeping in mind that $h \le T$, and $K$ may depend upon $T$. $\square$

It now follows from Proposition 6 and Theorem 13.5 in [2] that the collection of random processes $\{X_t^N,\ t\ge 0\}_{N\ge 1}$ is tight in $D([0,\infty))$. Since we already know that for all $k\ge 1$, all $0\le t_1, t_2, \dots, t_k < \infty$,

$$(X_{t_1}^N, X_{t_2}^N, \dots, X_{t_k}^N) \to (X_{t_1}, X_{t_2}, \dots, X_{t_k}) \quad a.s., as\ N\to\infty,$$

we have that $X^N \Rightarrow X$ weakly in $D([0,\infty))$. Moreover, since $\sup_t |X_t^N - X_{t^-}^N| = 1/N$, it follows from Theorem 13.4 in [2] that $X$ possesses an a. s. continuous modification, and the weak convergence holds for the topology of locally uniform convergence in $[0,+\infty)$.

We have in fact a slightly stronger result.

**Corollary 4.** *The process $X$ possessses an a. s. continuous modification, and for all $T>0$,*

$$\sup_{0\le t\le T} |X_t^N - X_t| \to 0 \text{ in probability, as } N\to\infty.$$

PROOF : To each $\delta>0$, we associate $n\ge 1$ and $0=t_0<t_1<\cdots<t_n=T$, such that $\sup_{1\le i\le n}(t_i - t_{i-1}) \le \delta$. We have

$$
\begin{aligned}
\sup_{0\le t\le T}|X_t^N - X_t| &\le \sup_i \sup_{t_{i-1}\le t\le t_i} |X_t^N - X_{t_{i-1}}^N| \wedge |X_t^N - X_{t_i}^N| + \sup_i |X_{t_i}^N - X_{t_i}| \\
&\quad + \sup_i \sup_{t_{i-1}\le t\le t_i} |X_t - X_{t_i}| \\
&\le w_T''(X^N,\delta) + \sup_i |X_{t_i}^N - X_{t_i}| + w_T(X,\delta),
\end{aligned}
$$

where

$$
\begin{aligned}
w_T(x,\delta) &= \sup_{0\le s,t\le T, |s-t|\le T} |x(t)-x(s)|, \\
w_T''(x,\delta) &= \sup_{0\le t_1<t<t_2\le T, t_2-t_1\le\delta} |x(t)-x(t_1)| \wedge |x(t)-x(t_2)|.
\end{aligned}
$$

From the proof of Theorem 13.5 in [2], we know that Proposition 6 implies that

$$
\mathbf{P}(w_T''(X^N,\delta) > \varepsilon) \le \varepsilon^{-4} C_T (2\delta)^{1/4}.
$$

Since $X$ is continuous a. s., for each $\varepsilon > 0$,

$$
\mathbf{P}(w_T(X,\delta) > \varepsilon) \to 0, \quad \text{as } \delta \to 0.
$$

Moreover

$$
\sup_i |X_{t_i}^N - X_{t_i}| \to 0 \quad \text{a. s., as } N \to \infty.
$$

The result follows □

## 4.3 The main result

In this section, we prove our main result. Before let us etablish

**Lemma 7.** $\forall s > 0, k \ge 1, 0 \le r_k < r_{k-1} < \cdots < r_1 < r_0 = s$, $\forall N \ge 1$ , $\forall a \in \{0,1\}^N, \forall \pi \in S_N, \forall A_{r_j} \in \sigma(X_{r_j}), 0 \le j \le k$,

$$
\mathbf{P}\left(\{\eta_s^N = a\}, \bigcap_{j=1}^k A_{r_j}\right) = \mathbf{P}\left(\{\eta_s^N = \pi^*(a)\}, \bigcap_{j=1}^k A_{r_j}\right).
$$

PROOF : $\forall m \ge N$, $p_0, p_1, \ldots, p_k \in [0,1]$ such that $mp_j \in \mathbf{N}$ for $0 \le j \le k$,

$$\mathbf{P}\left(\{\eta_s^N=\pi^*(a)\},\bigcap_{j=0}^{k}\{X_{r_j}^m=p_j\},\bigcap_{i=1}^{k}\{Z_{r_i}^{m,r_{i-1}}\le m\}\right)$$
$$=\sum_{b'\in\mathscr{A}(a,p_0)}\mathbf{P}\left(\eta_s^m=(\pi^*(a),b'),\bigcap_{j=1}^{k}\{X_{r_j}^m=p_j\},\bigcap_{i=1}^{k}\{Z_{r_i}^{N,r_{i-1}}\le m\}\right)$$
$$=\sum_{b'\in\mathscr{A}(a,p_0)}\mathbf{P}\left(\eta_s^m=(\pi'^*(a,b'),\bigcap_{j=1}^{k}\{X_{r_j}^m=p_j\},\bigcap_{i=1}^{k}\{Z_{r_i}^{N,r_{i-1}}\le m\}\right)$$

where $\mathscr{A}(a,p_0)=\{b'\in\{0,1\}^{m-N}:\sum_{i=1}^N a_i+\sum_{j=1}^{m-N}b'_j=mp_0\},\ \ \pi'\in S_m$. Thanks to Proposition 5, we deduce that

$$\mathbf{P}\left(\{\eta_s^N=\pi^*(a)\},\bigcap_{j=0}^{k}\{X_{r_j}^m=p_j\},\bigcap_{i=1}^{k}\{Z_{r_i}^{m,r_{i-1}}\le m\}\right)$$
$$=\mathbf{P}\left(\{\eta_s^N=(a)\},\bigcap_{j=0}^{k}\{X_{r_j}^m=p_j\},\bigcap_{i=1}^{k}\{Z_{r_i}^{m,r_{i-1}}\le m\}\right)$$

wich implies that for all $f_j\in C_b([0,1]),1\le j\le k$,

$$\mathbf{E}\left[\prod_{j=1}^{k}f_j(X_{r_j}^m);\{\eta_s^N=a\};\bigcap_{i=1}^{k}\{Z_{r_i}^{m,r_{i-1}}\le m\}\right]=\mathbf{E}\left[\prod_{j=1}^{k}f_j(X_{r_j}^m);\{\eta_s^N=\pi^*(a)\};\bigcap_{i=1}^{k}\{Z_{r_i}^{m,r_{i-1}}\le m\}\right]$$

from which we deduce by

$$\left|\mathbf{E}\left(\prod_{j=1}^{k}f_j(X_{r_j}^m);\{\eta_s^N=a\}\right)-\mathbf{E}\left(\prod_{j=1}^{k}f_j(X_{r_j}^m);\{\eta_s^N=\pi^*(a)\}\right)\right|\le 2\left(\prod_{j=1}^{k}||f||_\infty\right)\mathbf{P}(\cup_{i=1}^{k}\{Z_{r_i}^{m,r_{i-1}}\ge m\})$$
$$\le 2k\prod_{j=1}^{k}||f||_\infty e^{-chm^2},$$

where the last line follows the lemma 4 and $h=\inf_{1\le j\le k}\{r_j-r_{j-1}\}$. Letting $m\to\infty$, we deduce that

$$\mathbf{E}\left[\prod_{j=1}^{k}f_j(X_{r_j});\{\eta_s^N=a\}\right]=\mathbf{E}\left[\prod_{j=1}^{k}f_j(X_{r_j});\{\eta_s^N=\pi^*(a)\}\right].$$

The lemma has been established. □

We are now in position to prove our main result.

**Theorem 2.** *The* $[0,1]-$ *valued process* $\{X_t,t\ge 0\}$ *defined by* (10) *admits a continuous version which is a weak solution of the Wright-Fisher equation* (1).

PROOF : We already know from Corollary 4 that $\{X_t,t\ge 0\}$ defined by (10) possesses a continuous modification. The proof of Theorem 2 is structured as follows.

In step 1 we show that $\{X_t, t \geq 0\}$ is a Markov process. In step 2 we show that $X_t$ is a weak solution of the Wright-Fisher equation (1) .

STEP 1: We want to show that $\{X_t, t \geq 0\}$ defined by (3.3) is a Markov process. For $0 \leq s < t$, let $H_{s,t}$ denote the history between $s$ and $t$ with affects the vector $\{\eta_s(i), i \geq 1\}$. For all $N \geq 1$, the history $H^N_{s,t}$ is described by the time ordered sequence of all birth and death events affecting the levels between 1 and $N$, from time $s$ to time $t$ . $H_{s,t}$ is the union over $N \in \mathbf{N}$ of the $H^N_{s,t}$'s. $X_t$ is a function of $\{\eta_s(i), i \geq 1\}$ and $H_{s,t}$. But $H_{s,t}$ is independent of $\sigma(X_r,\ 0 \leq r \leq s) \vee \sigma(\eta_s(i), i \geq 1)$. Consequently, for any $0 < x \leq 1$, there exists a measurable function $G_x : \{0,1\}^{\mathbf{N}} \to [0,1]$ such that

$$\mathbf{P}(X_t \leq x | \sigma(X_r,\ 0 \leq r \leq s)) = \mathbf{E}(G_x(\eta_s(i), i \geq 1) | \sigma(X_r,\ 0 \leq r \leq s)).$$

We know that conditionally upon $X_s = x$, the $\eta_s = \{\eta_s(i), i \geq 1\}$ are i. i. d. Bernoulli with parameter $x$. So all we need to show is that is that conditionally upon $\sigma(X_r,\ 0 \leq r \leq s\}$, the $\{\eta_s(i), i \geq 1\}$ are i. i. d Bernoulli with parameter $X_s$. In view of Proposition 4, it suffices to prove that conditionally upon $\sigma(X_r,\ 0 \leq r \leq s\}$, the $\eta_s(i)$ are exchangeable. This will follow from the fact that the same is true conditionally upon $\sigma(X_{r_1}, \ldots, X_{r_k}, X_s\}$, for all $k \geq 1, 0 \leq r_k < r_{k-1} < \ldots r_1 < r_0 = s$.

Hence it suffices to show is that for all $N \geq 1$, $\eta^N_s = (\eta_s(1), \ldots, \eta_s(N))$ is conditionally exchangeable, given $\sigma(X_{r_1}, \ldots, X_{r_k}, X_s\}$. This is established in lemma 7. The Markov
property of the process $\{X_t, t \geq 0\}$ follows.

STEP 2: We now finally show that $\{X_t, t \geq 0\}$ has the right transition probability. Since $X_t$ takes values in the compact set $[0,1]$, the conditional law of $X_t$, given that $X_0 = x$ is determined by its moments. Hence all we have to show is that for all $t > 0, x \in [0,1]$, $n \geq 1$,

$$\mathbf{E}[X^n_t | X_0 = x] = \mathbf{E}[Y^n_t | Y_0 = x], \tag{12}$$

where $\{Y_t,\ t \geq 0\}$ solves (1).

From de Finneti's theorem [1], we deduce that conditionally upon $X_t$, the $\{\eta_s(i), i \geq 1\}$ are i. i. d. Bernoulli with parameter $X_t$. Consequently, for all $n \geq 1$,

$$X^n_t = \mathbf{P}(\eta_t(1) = \cdots = \eta_t(n) = 1 \mid X_t)$$

This implies that

$$\begin{aligned}
\mathbf{E}_x[X^n_t] &= \mathbf{E}_x[\mathbf{P}(\eta_t(1) = \cdots = \eta_t(n) = 1 | X_t)] \\
&= \mathbf{P}_x(\eta_t(1) = \cdots = \eta_t(n) = 1) \\
&= \mathbf{P}_x(\text{the } 1 \ldots \widetilde{R}(t) \text{ individuals at time 0 are } b) \\
&= \mathbf{E}_n[x^{\widetilde{R}_t}],
\end{aligned}$$

where $\widetilde{R}_s = Z^{N,t}_{t-s}$, for all $0 \le s \le t$. And it is easy to see that $\widetilde{R}_t$ and $R_t$ defined in section 2 have the same law. (12) then follows from Proposition 1. The result is proved. $\square$

*Remark 3.* For any $N \ge 1$, the process $\{X^N_t, t \ge 0\}$ is not a Markov process. Indeed, the past values $\{X^N_s, 0 \le s < t\}$ give us some clue as to what the values of $\eta_t(N+1), \eta_t(N+2), \dots$ may be, and this influences the law of the future values $\{X^N_{t+r}, r > 0\}$.

## References


1. D. Aldous, Exchangeability and related topics, in *Ecole d'été St Flour 1983*, Lectures Notes in Math. **1117**, 1–198, 1985.
2. P. Billingsley, *Convergence of Probability Measures*, 2d ed., Wiley Inc., NewYork, 1999.
3. P. Donnelly and T.G. Kurtz. A countable representation of the Fleming Viot measure- valued diffusion. *Ann. Probab.* **24**, 698–742, 1996.
4. P. Donnelly and T.G. Kurtz. Particle representations for measure-valued population models. *Ann. Probab.* **27**, 166–205, 1999.
5. P. Donnelly and T.G. Kurtz. Genealogical processes for Fleming-Viot models with selection and recombination, *Ann. Appl. Probab.* **9**, 1091–1148, 1999.
6. J.F.C Kingman, The coalescent, *Stoch. Proc. Appl.*, **13**, 235–248, 1982.
7. É. Pardoux, M. Salamat, On the height and length of the ancestral recombination graph, *J. Appl. Prob.* **46**, 669–689, 2009.
8. A. Wakolbinger, Lectures at the "Evolutionary Biology and Probabilistic Models", Summer School ANR MAEV, La Londe Les Maures, unpublished 2008.