Siegmund duality with applications to the neutral Moran model conditioned on never being absorbed

Thierry Huillet

SIEGMUND DUALITY WITH APPLICATIONS TO THE NEUTRAL MORAN MODEL CONDITIONED ON NEVER BEING ABSORBED

THIERRY HUILLET

Abstract. We shall first consider the classical neutral Moran model with two alleles whose fate is either to become extinct or to reach fixation. We will study an ergodic version of the Moran model obtained by conditioning it to never hit the boundaries, making use of a Doob transform. We shall call it the recurrent Moran model. We will show that the Siegmund dual of the recurrent Moran process exists and is a substochastic birth and death chain. Conditioning this process to exit in its natural absorbing state, we construct a process with a unique absorbing state which is intertwined to the original recurrent Moran process. The time needed for the intertwined process to first hit its absorbing state is related to the time to reach stationarity for the recurrent Moran process. Using spectral information on the intertwined chain, we extract limiting information on this first hitting time that shows that there is no abrupt relaxation to equilibrium for the recurrent Moran chain. This makes use of the relation between duality and intertwining and strong stationary times. Other related transition times of the recurrent Moran chain are also briefly investigated, namely: the first return time to the ground state and the expected time needed to move from one end to the other end of the state-space.

Running title: Siegmund duality and the recurrent Moran Model

Keywords: Population dynamics, Moran model, Birth and death chains, Fluctuation theory, Relaxation to equilibrium, Siegmund dual, Strong stationarity times.

AMS 60J10, 60J70, 92B05; PACS numbers: 02.50.Ey, 87.23.Kg, 87.10.Mn

1. Introduction and outline of the results

Consider a population constituted of N individuals (alleles) of two types. Suppose at all time n there are Z_n individuals of type 1 and $N - Z_n$ individuals of type 2. Elementary population genetics aims at elucidating the fate of the type 1 allele frequency Z_n, when various driving 'forces' are at stake in the gene pool. One major driving source we shall stick to in this manuscript is the genetic drift whose nature is exclusively random. In this context, we shall first consider the classical *neutral* Moran model with two alleles (see [14] or [28]). This process is a well-known birth and death chain on the state-space $\{0, .., N\}$. The process Z_n either becomes extinct or it fixes; it has two natural boundaries $\{0, N\}$ which are both absorbing. We will next introduce an ergodic version of Z_n obtained by conditioning Z_n to never hit the boundaries, making use of a Doob transform. We shall call the resulting process, say X_n, the recurrent Moran process. Its state-space is now $\{1, .., N-1\}$.

The recurrent Moran process X_n will be the main concern of this paper. Its main features are first derived in Section 2.

Then we run into the rich duality idea between Markov chains. Dual processes occur in many applications; for typical examples and use of dual processes, we refer the reader to the mathematics and physics literature on interacting particle systems (see for example [7], [32], [27], [31], [18], [4] and [22]). Other examples occur in the population genetics context (see for example [10], [11], [29], [19] and [20]) and essentially go back to Siegmund similar early concerns for stochastically monotone processes (see [30] and [5]).

In this Note, we will focus on the Siegmund duality and on its use in our genetics problem. We will show that the Siegmund dual of the recurrent Moran process X_n exists and is a substochastic birth and death chain. It loses mass at state $\{1\}$ and is absorbed in state $\{N-1\}$. We shall call this process $\widehat{X}_n$. By adding an extra coffin state, say $\{0\}$, we will consider an enlarged proper Markov chain $\widehat{X}_n^+$, now with the two absorbing states $\{0\}$ and $\{N-1\}$. While conditioning the enlarged process $\widehat{X}_n^+$ to exit in its natural absorbing state $\{N-1\}$, we shall construct another proper stochastic process $\widetilde{X}_n$, but now with this unique absorbing state. This new process $\widetilde{X}_n$ turns out to be intertwined with the original recurrent Moran process X_n. It is again a birth and death chain.

The intertwining construction shows that the original positive recurrent birth and death chain X_n may also be viewed as the output (through a stochastic link) of a dual hidden Markov chain $\widetilde{X}_n$ which is absorbed in a single state. Its peculiarity is that X_n is a Markov output of $\widetilde{X}_n$ which is itself Markov. As a result of this intertwining construction, the time needed for $\widetilde{X}_n$ to first hit its absorbing state is related to the time to reach stationarity for the recurrent Moran process X_n under study. This is when both intertwined processes X_n and $\widetilde{X}_n$ start in the same state $\{1\}$. Because we deal here with a birth and death chain X_n with known spectral characteristics, the estimation of the hitting time for $\widetilde{X}_n$ is somehow made easier (see [24] and [16]). These constructions are described in Section 3; they are in the spirit of [1] and [7] and actually serve as an additional illustration of the theory.

Using spectral information on the intertwined chain $\widetilde{X}_n$, we will show in Section 4 that there is no abrupt relaxation to equilibrium for X_n (no cutoff phenomenon[1] in the sense of Diaconis, [6]). More precisely, we will show that the expected value of the time $\widetilde{\tau}_{1,N-1}$ for $\widetilde{X}_n$ to first hit its absorbing state $\{N-1\}$, starting from state $\{1\}$, grows like N^2, just like its standard deviation. We will identify the limit law of $\widetilde{\tau}_{1,N-1}$ (normalized by its mean value) as a quadratically weighted sum of exponentially distributed random variables. The random time $\widetilde{\tau}_{1,N-1}$ being a strong stationarity time, the ergodic chain X_n enters gradually into its invariant measure after about N^2 steps. By the end of this Section 4, we add some considerations on the scaling limit attached to the recurrent Moran model X_n. These follow from the observation that the shifted recurrent Moran model is a Moran model with mutations.

In Section 5, other transition times of the chain X_n are briefly investigated. It is shown that the first return time to state $\{1\}$ grows like N^2 whereas the expected time for X_n to move from one end $\{1\}$ to the other end $\{N-1\}$ of the state-space grows like N^3.

Remark[1]: Similar considerations for the *nonneutral* ergodic Moran model including mutations (described in [23]) can be found in [12] and [11]. It includes the Ehrenfest urn model and the Bernoulli-Laplace urn model (see [15], [9]), both being simple heat exchange models of statistical physics. In this context, the conclusions are radically different: the Moran model with mutations exhibits a cutoff phenomenon. This ergodic chain (of a different nature from the X_n under study here) enters abruptly into its invariant measure after about $\frac{N}{2\mu}\log N$ steps, where μ is the total mutation pressure.

2. The neutral Moran model conditioned on non-absorption

Consider a population constituted of N individuals (alleles) of two types. Suppose at time n there are $Z_n = x$ individuals of type 1 and $N-x$ individuals of type 2. We are first interested in the random Moran neutral evolution of type 1 individuals, namely of Z_n, $n \geq 0$. It can be described as follows: Pick at random 2 distinct individuals from the whole 2-types population at time n. Assume that one is bound to die, the other one is bound to survive in the next generation and produce an additional offspring so that, at step $n+1$, the total population size remains constant. From this model, given $Z_n = x$, the updated number of type 1 individuals is

$$Z_{n+1} = x+1 \text{ with probability } \alpha_x = \left(1-\frac{x}{N}\right)\frac{x}{N} \tag{1}$$

$$Z_{n+1} = x-1 \text{ with probability } \beta_x = \frac{x}{N}\left(1-\frac{x}{N}\right) \tag{2}$$

$$Z_{n+1} = x \text{ with probability } \rho_x = \left(\frac{x}{N}\right)^2 + \left(1-\frac{x}{N}\right)^2 \tag{3}$$

The neutral Moran process Z_n is thus a transient birth and death Markov chain. It is a cornerstone model of population genetics (the so-called random genetic drift model). See [14], p. 104, [28] or [17] for example. Its stochastic transition matrix reads (the empty entries are 0).

$$\Pi = \begin{bmatrix} 1 & 0 & & & & \\ \beta_1 & \rho_1 & \alpha_1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \beta_{N-1} & \rho_{N-1} & \alpha_{N-1} & \\ & & & 0 & 1 & \end{bmatrix}.$$

The eigenvalues of Π are easily seen to be $\theta_k = 1 - \frac{k(k-1)}{N^2}$, $k = 0, .., N$ (see [14], p. 106). Note that $\theta_0 = \theta_1 = 1$, translating the fact that the two states $\{0, N\}$ are absorbing for Z_n. We shall let $\tau_{x,0}$, $\tau_{x,N}$ be the first random times at which Z_n hits $\{0\}$ and $\{N\}$ respectively, starting from $Z_0 = x$. We shall also let $\tau_x = \tau_{x,0} \wedge \tau_{x,N}$,

with $\tau_x < \infty$ almost surely. Let

$$\overline{\Pi} = \begin{bmatrix} \rho_1 & \alpha_1 & & & \\ \beta_2 & \rho_2 & \alpha_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_{N-2} & \rho_{N-2} & \alpha_{N-2} \\ & & & \beta_{N-1} & \rho_{N-1} \end{bmatrix}$$

be obtained from Π after removing its first and last rows and columns. This matrix of size $N-1$ is substochastic in the sense that it loses mass at states $\{1, N-1\}$.

Let $\overline{\boldsymbol{\pi}}^{'}$ be the transpose row-vector of some column-vector $\overline{\boldsymbol{\pi}}$. Let $\overline{\boldsymbol{\pi}}^{'}$ and $\overline{\boldsymbol{\psi}}$ be defined as the left and right (row and column) eigenvectors associated to the spectral radius (largest eigenvalue) θ_2 of $\overline{\Pi}$, namely:

$$\overline{\boldsymbol{\pi}}'\overline{\Pi} = \theta_2 \overline{\boldsymbol{\pi}}' \text{ and } \overline{\Pi}\bar{\boldsymbol{\psi}} = \theta_2 \overline{\boldsymbol{\psi}}. \tag{4}$$

We have $\theta_2 = 1 - 2/N^2$ and $\overline{\boldsymbol{\psi}} = \left(\overline{\psi}_x,\ x = 1, .., N-1\right)$, where $\overline{\psi}_x = \frac{x}{N}\left(1 - \frac{x}{N}\right)$. Note also that $\overline{\boldsymbol{\pi}} = (\overline{\pi}_x,\ x = 1, .., N-1)$, where $\overline{\pi}_x = \frac{1}{N-1}$. Regardless of the starting point x, we therefore have

$$\mathbb{P}_x\left(Z_n = z \mid \tau_x > n\right) \to \overline{\pi}_z = \frac{1}{N-1} \text{ as } n \to \infty,$$

and $\overline{\boldsymbol{\pi}}$ has an interpretation in terms of the uniform Yaglom limiting measure of Z_n on $\{1, .., N-1\}$.

Let P be the birth and death stochastic matrix defined from $\overline{\Pi}$ by applying the Doob transform

$$P = \theta_2^{-1} D_{\overline{\boldsymbol{\psi}}}^{-1} \overline{\Pi} D_{\overline{\boldsymbol{\psi}}}, \tag{5}$$

where $D_{\overline{\boldsymbol{\psi}}} \equiv \mathrm{diag}\left(\overline{\psi}_1, .., \overline{\psi}_{N-1}\right)$. We have

$$P = \begin{bmatrix} r_1 & p_1 & & & \\ q_2 & r_2 & p_2 & & \\ & \ddots & \ddots & \ddots & \\ & & q_{N-2} & r_{N-2} & p_{N-2} \\ & & & q_{N-1} & r_{N-1} \end{bmatrix}$$

where

$$p_x = \theta_2^{-1}\alpha_x \frac{\overline{\psi}_{x+1}}{\overline{\psi}_x} = \theta_2^{-1}\frac{x+1}{N}\left(1 - \frac{x+1}{N}\right) \tag{6}$$

$$q_x = \theta_2^{-1}\beta_x \frac{\overline{\psi}_{x+1}}{\overline{\psi}_x} = \theta_2^{-1}\frac{x-1}{N}\left(1 - \frac{x-1}{N}\right) \tag{7}$$

$$r_x = \theta_2^{-1}\rho_x = \theta_2^{-1}\left[\left(\frac{x}{N}\right)^2 + \left(1 - \frac{x}{N}\right)^2\right] \tag{8}$$

The invariant measure $\boldsymbol{\pi} = (\pi_x,\ x = 1, .., N-1)$ of the corresponding random walk (say X_n) governed by P is given by $\pi_x = \pi_1 \prod_{y=1}^{x-1} \frac{p_y}{q_{y+1}}$. π_x can also be expressed

in terms of the normalized Hadamard product of $\overline{\boldsymbol{\pi}}$ and $\overline{\boldsymbol{\psi}}$:

$$\pi_x = \frac{\overline{\pi}_x \overline{\psi}_x}{\sum_{x=1}^{N-1} \overline{\pi}_x \overline{\psi}_x} = \frac{6x\,(N-x)}{N\,(N-1)\,(N+1)}. \tag{9}$$

This invariant probability measure is a special incarnation of the bivariate generalized hypergeometric distribution (a discrete version of the beta$(2,2)$ distribution supported by the unit interval). Note that the recurrent Moran model governed by P is reversible as a result of P obeying detailed balance which is true for all birth and death chains.

Clearly, the eigenvalues of P are

$$t_k = \theta_2^{-1}\left(1 - \frac{k\,(k+1)}{N^2}\right),\ k = 1, .., N-1$$

which may be recast as

$$t_k = 1 - \frac{(k-1)\,(k+2)}{N^2-2},\ k = 1, .., N-1. \tag{10}$$

Note that $t_1 = 1$, $t_2 = 1 - \frac{4}{N^2-2} < 1$ and the spectral gap of P is $1 - t_2 = \frac{4}{N^2-2}$. We have $t_{N-1} = \frac{N}{N^2-2} > 0$ so that all eigenvalues of P are > 0.

The random walk X_n, governed by P, corresponds to the random neutral Moran process Z_n conditioned to never become extinct nor fixed in the remote future[2]. See [25] for a continuous-time version of this construction. It is again a birth and death Markov chain but with the modified transition probabilities (6, 8). Because it is a positive recurrent ergodic Markov chain, we shall call the process X_n governed by P the recurrent neutral Moran process (or shortly the *recurrent* Moran model). Note that if $X_0 \overset{d}{\sim} \boldsymbol{\pi}$ (meaning that the distribution of X_0 is $\boldsymbol{\pi}$), then $X_n \overset{d}{\sim} \boldsymbol{\pi}$ for all $n \geq 1$: the distribution of X_n remains unchanged at all times.

We can check from (6, 7) that for all $N \geq 3$ and for all $x \in \{1, .., N-2\}$

$$p_x + q_{x+1} \leq 1 \tag{11}$$

which means that the random walk X_n is stochastically monotone in the sense that for all $n \geq 0$ and $y \geq 1$, the function $x \to \mathbb{P}_x\,(X_n \leq y)$ is a non-increasing function of x. This fact may also be viewed as a direct consequence of P having all its eigenvalues positive. This property will prove essential in the following statement that a Siegmund dual to X_n exists, [30]. See [5] for similar considerations for birth and death chains on the half-line.

Remark[2]: The original idea of conditioning a Markov chain to never become extinct in the remote future seems to be due to Spitzer (unpublished) in the sixties and then Lamperti-Ney, [26]; see Athreya-Ney [2], p $56-59$. It was first introduced in the context of the celebrated Bienyamé-Galton-Watson branching process.

3. The Siegmund dual to the recurrent Moran model

In this Section, we shall illustrate the power of the duality/intertwining relationship by considering the simplest Siegmund dual of a birth and death chain, with the recurrent neutral Moran birth and death example in mind. This will prove useful to address the question of computing the strong stationary time distribution that

helps quantifying the 'distance' to equilibrium of the original positive recurrent birth and death process.

Definition 1. [27]*: Two discrete-time Markov processes* $\left(X_n, \widehat{X}_n; n \geq 0\right)$, *with state-spaces* $(\mathcal{X}, \mathcal{Y})$, *possibly with substochastic transition kernels, are said to be dual with respect to some non-singular duality kernel* $H \geq 0$ *on the product space* $\mathcal{X} \times \mathcal{Y}$ *if* $\forall x \in \mathcal{X}$, $\forall y \in \mathcal{Y}$, $\forall n \in \mathbb{N}$:

$$\mathbb{E}_x H\left(X_n, y\right) = \mathbb{E}_y H\left(x, \widehat{X}_n\right). \tag{12}$$

When the state-spaces $(\mathcal{X}, \mathcal{Y}) = \{1, .., N-1\}^2$ are finite and identical, the duality kernel is a square-matrix and the transition matrix of the dual process $\widehat{X}_n$, say $\widehat{P}$, is obtained from the one P of the original process X_n by:

$$\widehat{P}' = H^{-1} P H, \tag{13}$$

where $\widehat{P}'$ stands for the transpose of $\widehat{P}$.

The Siegmund dual [30]**.** The Siegmund duality kernel is:

$$H\left(x, y\right) = \mathbf{1}\left(x \leq y\right). \tag{14}$$

If, for a given process X_n, a process $\widehat{X}_n$ exists satisfying the above conditions (12) or (13), then $\widehat{X}_n$ is called the Siegmund dual of X_n. Clearly, in the birth and death case for X_n, the condition is that X_n should be stochastically monotone in that, for all $y \geq 1$ and $n \geq 0$, the function $x \to \mathbb{P}_x\left(X_n \leq y\right)$ should be non-increasing with x. This is because

$$\mathbb{P}_x\left(X_n \leq y\right) = \mathbb{P}_y\left(\widehat{X}_n \geq x\right),$$

as a result of (12) and (14).

For the positive recurrent Moran process and for the Siegmund kernel, the transition matrix $\widehat{P}$ of the dual process $\widehat{X}_n$ reads:

$$\widehat{P} = \begin{bmatrix} r_1 - q_2 & q_2 & & & \\ p_2 & \widehat{r}_2 & q_3 & & \\ & \ddots & \ddots & \ddots & \\ & & p_{N-2} & \widehat{r}_{N-2} & q_{N-1} \\ & & & 0 & 1 \end{bmatrix},$$

where $\widehat{r}_x \equiv 1 - (p_x + q_{x+1})$, $x \in \{2, ..., N-2\}$ (and $\widehat{q}_x = p_x$, $x = 2, .., N-2$, $\widehat{p}_x = q_{x+1}$, $x = 1, ..., N-2$). It is again the one of a birth and death process. Indeed, in this case, H is an upper-right triangular matrix with non-zero entries 1, whereas the non-null entries of H^{-1} are the diagonal (with entries 1) and the upper-diagonal with entries -1. The displayed structure of $\widehat{P}$ follows from this and the duality relation (13).

For the dual process $\widehat{X}_n$ to exist, we thus need to make sure that $p_x + q_{x+1} \leq 1$ for $x \in \{1, .., N-2\}$. This is a necessary and sufficient condition to guarantee the stochastic monotonicity of X_n. But this holds true if X_n is the recurrent Moran model defined in the previous section.

From the structure of $\widehat{P}$, it is apparent that the dual process loses mass at $y=1$ and is absorbed in $y=N-1$. Let us therefore add a coffin state, say $+\equiv\{0\}$, and consider the enlarged stochastic matrix which we shall call $\widehat{P}_+$:

$$\widehat{P}_+=\begin{bmatrix} 1 & 0 & & & & & \\ 1-r_1 & r_1-q_2 & q_2 & & & & \\ & p_2 & \widehat{r}_2 & q_3 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & p_{N-2} & \widehat{r}_{N-2} & q_{N-1} & \\ & & & & 0 & 1 & \end{bmatrix},$$

The corresponding proper birth and death chain, call it $\widehat{X}_n^+$, has now two absorbing states, one at $\{0\}$, one at $\{N-1\}$. We shall now construct another random walk by conditioning $\widehat{X}_n^+$ to be absorbed in $\{N-1\}$. This will be done with the help of its scale function of $\widehat{X}_n^+$.

The scale function of $\widehat{X}_n^+$. Let now $\widehat{\varphi}(x)$, $x=0,1,...,N-1$ be the scale (or harmonic) function of $\widehat{X}_n^+$, solving $\widehat{P}_+\widehat{\varphi}=\widehat{\varphi}$ while imposing $\widehat{\varphi}(0)=0$. We can easily check that:

$$\widehat{\varphi}(0)=0,\ \widehat{\varphi}(x)=\frac{1}{\pi_1}\sum_{y=1}^{x}\pi_y,\ x=1,..,N-1. \tag{15}$$

The scale function of $\widehat{X}_n^+$ expresses in terms of the cumulative probability distribution $\pi_x^c\equiv\sum_{y=1}^{x}\pi_y$ of the invariant measure of the process X_n.

Let $\widehat{\tau}_x\equiv\widehat{\tau}_{x,1}\wedge\widehat{\tau}_{x,N-1}$ be the smallest of the first hitting time for $\widehat{X}_n^+$ of $\{0\}$ and $\{N-1\}$, starting from $x\in\{1,...,N-2\}$. We have:

$$\mathbb{P}_x\left(\widehat{X}^+_{\widehat{\tau}_x}=N-1\right)=\mathbb{P}\left(\widehat{\tau}_{x,N-1}<\widehat{\tau}_{x,1}\right)=\frac{\widehat{\varphi}(x)}{\widehat{\varphi}(N-1)}=\pi_x^c. \tag{16}$$

The probability that $\widehat{X}_n^+$ is absorbed in $\{N-1\}$ can therefore be expressed in terms of the scale function $\widehat{\varphi}$ of $\widehat{X}_n^+$. Let $\boldsymbol{\pi}^c=(\pi_x^c,\ x=0,..,N-1)$ and $D_{\boldsymbol{\pi}^c}=\operatorname{diag}\left(\pi_1^c,..,\pi_{N-1}^c\right)$.

Doob $h-$transform, [13]**.** Define a new transition matrix $\widetilde{P}_+$ by

$$\widetilde{P}_+=D_{\boldsymbol{\pi}^c}^{-1}\widehat{P}_+D_{\boldsymbol{\pi}^c}. \tag{17}$$

$\widetilde{P}_+$ is a stochastic matrix. The entries of $\widetilde{P}_+$ read

$$\widetilde{P}_+=\begin{bmatrix} 1 & 0 & & & & & \\ 0 & r_1-q_2 & p_1+q_2 & & & & \\ & \frac{\pi_1^c}{\pi_2^c}p_2 & \widehat{r}_2 & \frac{\pi_3^c}{\pi_2^c}q_3 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \frac{\pi_{N-3}^c}{\pi_{N-2}^c}p_{N-2} & \widehat{r}_{N-2} & \frac{\pi_{N-1}^c}{\pi_{N-2}^c}q_{N-1} & \\ & & & & 0 & 1 & \end{bmatrix},$$

where the state $\{0\}$ becomes isolated and disconnected. Deleting the first row and column (corresponding to the extra state$\{0\}$) of $\widetilde{P}_+$, we get a stochastic transition

matrix, call it $\widetilde{P}$, of some process $\widetilde{X}_n$ on the state-space $\{1,...,N-1\}$ which corresponds to $\widehat{X}_n^+$ conditioned to first hit state $\{N-1\}$ before state $\{0\}$. The state $\{1\}$ of this conditioned birth and death process is now partially reflecting whereas its only remaining absorbing state is $\{N-1\}$. See [21] for a similar conditioning in the context of the Ehrenfest urn model.

We now state that $\widetilde{P}$ and P (characterized by (6, 8)) are intertwined through a stochastic link Λ.

Proposition 1. (i) *The matrices $\widetilde{P}$ and P are similar (with the same eigenvalues), that is*

$$\widetilde{P} = \Lambda P \Lambda^{-1}. \tag{18}$$

With $\boldsymbol{\pi}$ as in (10), the link matrix Λ is given by: $\Lambda(\widetilde{x},x) = \frac{\pi_x}{\pi^c_{\widetilde{x}}}\mathbf{1}(x \leq \widetilde{x})$. It corresponds to the entries of a lower-triangular stochastic matrix. Its interpretation is in terms of the conditional probability

$$\Lambda(\widetilde{x},x) = \mathbb{P}\left(X_n = x \mid \widetilde{X}_n = \widetilde{x}\right), \text{ for all } n \geq 1. \tag{19}$$

(ii) *The link matrix Λ satisfies*

$$\Lambda(N-1,x) = \pi_x,\ x = 0,..,N. \tag{20}$$

(iii) *$\widetilde{\boldsymbol{\pi}}_0' = \boldsymbol{\pi}_0' = \mathbf{e}_0' \equiv (1,0,...,0)$ are admissible initial distributions for the chains $\widetilde{X}_n$ and X_n, satisfying*

$$\boldsymbol{\pi}_0' = \widetilde{\boldsymbol{\pi}}_0' \Lambda. \tag{21}$$

Proof: (i) Using (17) and (13)

$$\widetilde{P} = D_{\boldsymbol{\pi}^c}^{-1} \widehat{P} D_{\boldsymbol{\pi}^c} = D_{\boldsymbol{\pi}^c}^{-1} H' P' H'^{-1} D_{\boldsymbol{\pi}^c}.$$

The random walk X_n being reversible (detailed balance holds), $P' = D_{\boldsymbol{\pi}} P D_{\boldsymbol{\pi}}^{-1}$. Thus:

$$\widetilde{P} = \left(D_{\boldsymbol{\pi}^c}^{-1} H' D_{\boldsymbol{\pi}}\right) P \left(D_{\boldsymbol{\pi}}^{-1} H'^{-1} D_{\boldsymbol{\pi}^c}\right) = \Lambda P \Lambda^{-1}$$

where $\Lambda = D_{\boldsymbol{\pi}^c}^{-1} H' D_{\boldsymbol{\pi}}$. The entries of Λ are $\Lambda(\widetilde{x},x) = \frac{\pi_x}{\pi^c_{\widetilde{x}}}\mathbf{1}(x \leq \widetilde{x})$, satisfying $\sum_{x \leq \widetilde{x}} \frac{\pi_x}{\pi^c_{\widetilde{x}}} = 1$. Thus Λ is stochastic as its row sums all sum to 1.

Suppose $X_0 \overset{d}{\sim} \boldsymbol{\pi}_0$, $\widetilde{X}_0 \overset{d}{\sim} \widetilde{\boldsymbol{\pi}}_0$ and assume the initial distributions $\boldsymbol{\pi}_0$ and $\widetilde{\boldsymbol{\pi}}_0$ are $\Lambda-$compatible, that is $\boldsymbol{\pi}_0' = \widetilde{\boldsymbol{\pi}}_0' \Lambda$. Using (18), $\widetilde{P}^n \Lambda = \Lambda P^n$ for all $n \geq 1$. This means $\boldsymbol{\pi}_n' = \widetilde{\boldsymbol{\pi}}_n' \Lambda$, for all $n \geq 1$, where $\pi_n(.) = \mathbb{P}_{\boldsymbol{\pi}_0}(X_n = \cdot)$, $\widetilde{\pi}_n(.) = \mathbb{P}_{\widetilde{\boldsymbol{\pi}}_0}\left(\widetilde{X}_n = \cdot\right)$ are the laws at time n of X_n and $\widetilde{X}_n$. Thus $\Lambda(\widetilde{x},x) = \mathbb{P}\left(X_n = x \mid \widetilde{X}_n = \widetilde{x}\right)$, for all $n \geq 1$.

(ii) The last row of Λ is given by $\Lambda(N-1,x) = \pi_x$ so that once $\widetilde{X}_n$ hits state $\{N-1\}$, the law of X_n is $\boldsymbol{\pi}$.

(iii) The first row of Λ is $(1,0,...,0)$ so that $\mathbf{e}_0' = \mathbf{e}_0' \Lambda$. Thus both chains X_n and $\widetilde{X}_n$ may consistently start in state $\{1\}$. $\triangle$

So $\widetilde{P}$ (as the algebraic composition of the Siegmund dual of P with a Doob-transform) can be obtained from P through a stochastic link Λ.

Intertwining and strong stationary time. The intertwining construction shows that the original positive recurrent birth and death chain X_n with transition matrix P may also be viewed as the output (through the stochastic link Λ) of a dual hidden Markov chain $\widetilde{X}_n$ governed by $\widetilde{P} = \Lambda P \Lambda^{-1}$ and absorbed in the single state $\{N-1\}$. This is a setup reminiscent of filtering theory with $\widetilde{X}_n$ the hidden process and X_n the observable. The peculiarity of the intertwining construction is that X_n is a Markov output which is itself Markov. We emphasize that there is so far no clear 'genetical' interpretation of the process governed by $\widetilde{P}$. Because the starting point process governed by P may be viewed as a Markovian output of the $\widetilde{P}-$process, this suggests that the latter process should have a very basic (still missing) meaning.

What we only get is that once $\widetilde{X}_n$ hits its absorbing state $\{N-1\}$, the random walk X_n is distributed like $\boldsymbol{\pi}$, provided both X_n and $\widetilde{X}_n$ were both started in $\{1\}$. This suggests that, given $X_0 = \widetilde{X}_0 = 1$, the time needed for X_n to enter its invariant probability measure should be related to the entrance time of the hidden process $\widetilde{X}_n$ into its absorbing state $\{N-1\}$.

4. Time to reach stationarity for the recurrent Moran model

We shall let

$$\widetilde{\tau}_{1,N-1} = \inf\left(n : \widetilde{X}_n = N-1 \mid \widetilde{X}_0 = 1\right) \tag{22}$$

be the first hitting time of $\{N-1\}$ of $\widetilde{X}_n$, starting from the state $\{1\}$. The random time $\widetilde{\tau}_{1,N-1}$ gives some information on the speed of convergence of the law of the original process X_n to its invariant measure. It is a strong stationary time in the sense of Diaconis and Fill, [7]. The facts (18, 19, 20, 21) indeed guarantee that $\widetilde{\tau}_{1,N-1}$ is a strong stationary time of X_n in the sense that $X_{\widetilde{\tau}_{1,N-1}} \overset{d}{\sim} \boldsymbol{\pi}$ and is independent of $\widetilde{\tau}_{1,N-1}$ (see [7] Theorems 2.4 and 2.17 or [16] Theorem 2.1). Equivalently (see [1], Prop. 3.2), it holds that:

$$\operatorname{sep}\left(\boldsymbol{\pi}_{n,1}, \boldsymbol{\pi}\right) \leq \mathbb{P}\left(\widetilde{\tau}_{1,N-1} > n\right) \leq \mathbb{E}\left(\widetilde{\tau}_{1,N-1}\right)/n \tag{23}$$

where $\pi_{n,1}(\cdot) = \mathbb{P}_1(X_n = \cdot)$ is the law of X_n started in $\{1\}$, $\boldsymbol{\pi}$ its invariant measure. In (23), the separation discrepancy is defined by:

$$\operatorname{sep}\left(\boldsymbol{\pi}_{n,1}, \boldsymbol{\pi}\right) \equiv \sup_y \left[1 - \pi_n(y)/\pi_y\right].$$

It satisfies $\operatorname{sep}(\boldsymbol{\pi}_{n,1}, \boldsymbol{\pi}) \geq \|\boldsymbol{\pi}_{n,1} - \boldsymbol{\pi}\|_{TV}$ where $\|\boldsymbol{\pi}_{n,1} - \boldsymbol{\pi}\|_{TV} = \frac{1}{2}\sum_y |\pi_{n,1}(y) - \pi_y|$ is the total variation distance between $\boldsymbol{\pi}_n$ and $\boldsymbol{\pi}$.

Furthermore, from (18, 21), there is a unique 'witness' state d for X_n, here with $d = N-1$, such that either

$$\mathbb{P}_1\left(\widetilde{X}_n = N-1\right) \equiv \widetilde{\pi}_{n,1}(N-1) = 0 \Rightarrow \pi_{n,1}(d) = 0$$

or

$$\widetilde{\pi}_{n,1}(N-1) > 0 \Rightarrow \pi_{n,1}(d) = \widetilde{\pi}_{n,1}(N-1)\,\pi_d > 0,$$

showing that this random time is stochastically the smallest since the first inequality in (23) turns out to be an equality (see Remark 2.39 of [7] and Proposition 13 of [22]).

In our context of the birth and death chain $\widetilde{X}_n$ absorbed in $\{N-1\}$, the probability generating function of $\widetilde{\tau}_{1,N-1} \geq N-2$ is explicitly given by, [24], [16]:

$$\mathbb{E}\left(z^{\widetilde{\tau}_{1,N-1}}\right) = \prod_{k=2}^{N-1} \frac{(1-t_k)\,z}{1-t_k z},\ z \in [0,1]\,, \tag{24}$$

where $0 < t_k < 1$, $k = 2, ..., N-1$ are the $N-2$ distinct eigenvalues of both $\widetilde{P}$ and P, avoiding $t_1 = 1$ described in (10). Because the eigenvalues t_k are positive, then $\widetilde{\tau}_{1,N-1} \stackrel{d}{=} \sum_{k=2}^{N-1} \tau_k$ where the τ_ks are independent with $\tau_k \stackrel{d}{\sim} \text{geom}(1-t_k)$, the geometric distribution with success parameter $1-t_k$ on the half-line $\{1, 2, ...\}$. The formula (24) also reads

$$\mathbb{E}\left(z^{\widetilde{\tau}_{1,N-1}}\right) = \sum_{l=2}^{N-1} A_l \frac{z\,(1-t_l)}{1-t_l z},$$

where

$$A_l = \prod_{k \neq l} \frac{1-t_k}{t_l - t_k}.$$

Therefore

$$\mathbb{P}\left(\widetilde{\tau}_{1,N-1} > n\right) = \sum_{l=2}^{N-1} \prod_{k \neq l} \frac{1-t_k}{t_l - t_k} t_l^n,\ n \geq N-1. \tag{25}$$

Thus, $t_2^{-n}\mathbb{P}\left(\widetilde{\tau}_{1,N-1} > n\right) \to_{n\uparrow\infty} \prod_{k=3}^{N-1} \frac{1-t_k}{t_2-t_k}$ and $\widetilde{\tau}_{1,N-1}$ has geometric tails with exponent t_2. We also have:

$$\mu_N \equiv \mathbb{E}\left(\widetilde{\tau}_{1,N-1}\right) = \sum_{k=2}^{N-1} (1-t_k)^{-1} \text{ and} \tag{26}$$

$$\sigma_N^2 \equiv \sigma^2\left(\widetilde{\tau}_{0,N}\right) = \sum_{k=2}^{N-1} t_k\,(1-t_k)^{-2} = \sum_{k=2}^{N-1} (1-t_k)^{-2} - \sum_{k=2}^{N-1} (1-t_k)^{-1}\,. \tag{27}$$

Note that since t_2 is the dominant eigenvalue

$$\sigma^2\left(\widetilde{\tau}_{1,N-1}\right) \leq \frac{\mathbb{E}\left(\widetilde{\tau}_{1,N-1}\right)}{1-t_2}. \tag{28}$$

We can summarize these results (in the spirit of [1], [7] and [16]) as follows:

Proposition 2. *The Siegmund dual for the recurrent Moran birth and death chain X_n exists. There exists a Markov chain $\widetilde{X}_n$, intertwined with X_n, with $\{N-1\}$ as an absorbing state and fully described in Proposition* 1. *The random time $\widetilde{\tau}_{1,N-1}$ is a fastest strong stationary time for X_n whose law is characterized either by (24) or (25) involving the spectrum (given in (10)), of either P or $\widetilde{P}$, the transition matrices governing the two processes.*

Conditions for a cutoff phenomenon. Because the eigenvalues t_k of P are known explicitly from (10), it is possible to compute $\mu_N = \mathbb{E}\left(\widetilde{\tau}_{1,N-1}\right)$ and $\sigma_N^2 = \sigma^2\left(\widetilde{\tau}_{1,N-1}\right)$ and decide whether or not

$$\mathbb{E}\left(\widetilde{\tau}_{1,N-1}\right) \to \infty \text{ and } \sigma^2\left(\frac{\widetilde{\tau}_{1,N-1}}{\mathbb{E}\left(\widetilde{\tau}_{1,N-1}\right)}\right) \to 0 \text{ as } N \to \infty, \tag{29}$$

meaning $\sigma_N/\mu_N \to 0$ as $N \to \infty$.

Would these conditions both hold, then $\frac{\widetilde{\tau}_{1,N-1}}{\mathbb{E}(\widetilde{\tau}_{1,N-1})} \to 1$ in probability and $\lfloor \mu_N/2 \rfloor$ would be expected to be a cutoff time for X_n started in $\{1\}$ in the sense ([6]) that: With $n_N(x) = \lfloor (\mu_N + x\sigma_N) \rfloor$, then

$$\left\| \boldsymbol{\pi}_{n_N(x),1} - \boldsymbol{\pi} \right\|_{TV} \underset{N\uparrow\infty}{\to} c(x)$$

where $c(x) \to_{x\uparrow\infty} 0$ and $c(x) \to_{x\uparrow-\infty} 1$. The chain would then exhibit a sharp cutoff: if $n < n_N(x)$, $\|\boldsymbol{\pi}_{n,1} - \boldsymbol{\pi}\|$ would keep close to its maximal value 1, whereas if $n > n_N(x)$ it would suddenly drop to a smaller value.

When there is cutoff, the expected mixing time (measured in terms of $\mathrm{sep}(\boldsymbol{\pi}_{n,1}, \boldsymbol{\pi})$) is μ_N whereas the spectral gap is $1 - t_2$, the product of the 2 of which tends to ∞. Recalling $\sigma^2(\widetilde{\tau}_{1,N-1}) \leq \frac{\mu_N}{1-t_2}$, then $\sigma^2(\widetilde{\tau}_{1,N-1}/\mu_N) = \mu_N^{-2}\sigma^2(\widetilde{\tau}_{1,N-1}) \leq 1/((1-t_2)\mu_N)$ and the condition

$$(1-t_2)\mu_N \to \infty$$

is a sufficient condition for $\sigma^2(\widetilde{\tau}_{1,N-1}/\mu_N) \to 0$. If this condition holds, the contribution of $\sum_{k=3}^{N-1}(1-t_k)^{-1}$ to μ_N dominates the lead term $(1-t_2)^{-1}$ (see [8] for recent developments and precisions).

For some birth and death models, it may happen that $\mu_N \to \infty$ and $\sigma_N/\mu_N \nrightarrow 0$ as $N \to \infty$. In such cases, there is no cutoff and this is precisely what happens for the recurrent Moran model under concern.

Computing the mean and variance of $\widetilde{\tau}_{1,N-1}$ for the recurrent Moran model. For the case of the recurrent Moran model, using (10): $1 - t_k = \frac{k(k+1)-2}{N^2-2}$ and the large N values of μ_N and σ_N^2 can be estimated. Using indeed the integral approximation

$$\mu_N \sim (N-2)\left(N^2-2\right)\int_0^1 \frac{dx}{(2+x(N-2))(3+x(N-2))-2}$$

(where we used the change of variables $\frac{k-2}{N-2} = x$ and $\mu_N \sim \alpha_N \Leftrightarrow \mu_N/\alpha_N \to 1$ as $N \to \infty$), with $c_1 \equiv \frac{2\log 2}{3}$, we easily get

$$\mu_N \sim N\int_0^1 \frac{dx}{(x+1/N)(x+4/N)} = c_1 N^2.$$

In this situation, the contribution of $\sum_{k=3}^{N-1}(1-t_k)^{-1}$ to μ_N contribute equally likely with the lead term $(1-t_2)^{-1}$. Note also that for this model, $(1-t_2)\mu_N \sim 4c_1 \nrightarrow \infty$ as $N \to \infty$ suggesting that there is no cutoff.

Using similar techniques indeed, with $c_2 \equiv \frac{5}{36} + \frac{4\log 2}{27}$, we get

$$\sigma_N^2 \sim N\int_0^1 \frac{dx}{(x+1/N)^2(x+4/N)^2} \sim c_2 N^4$$

showing that $\sigma^2\left(\frac{\widetilde{\tau}_{1,N-1}}{\mathbb{E}(\widetilde{\tau}_{1,N-1})}\right) \sim \frac{c_2}{c_1^2} \nrightarrow 0$ as $N \to \infty$.

Thus $\frac{\widetilde{\tau}_{1,N-1}}{\mathbb{E}(\widetilde{\tau}_{1,N-1})} \nrightarrow 1$ in probability and there is no cutoff phenomenon. In fact, in the recurrent Moran case, we have

Proposition 3.

$$\frac{\widetilde{\tau}_{1,N-1}}{\mathbb{E}\left(\widetilde{\tau}_{1,N-1}\right)} \overset{d}{\to} T \text{ as } N \to \infty$$

where the limit T is random with

$$T \overset{d}{=} \sum_{l\geq 1} \frac{1}{l\left(l+3\right)} E_l\left(c_1\right)$$

where the $E_l\left(c_1\right)$ are independent and identically distributed (iid) random variables, with common distribution $\mathbb{P}\left(E_1\left(c_1\right)>x\right)=e^{-c_1 x}$.

Proof: Indeed, from the large N approximation $t_k/\left(1-t_k\right) \sim \frac{N^2}{k(k+1)-2}$ and making use of (24), for all $t \in \mathbb{R}$, we get

$$\mathbb{E}\left(e^{it\frac{\widetilde{\tau}_{1,N-1}}{\mu_N}}\right) = e^{it\frac{N-2}{\mu_N}} \prod_{k=2}^{N-1} \frac{\left(1-t_k\right) e^{\frac{it}{\mu_N}}}{1-t_k e^{\frac{it}{\mu_N}}} \sim e^{\frac{it}{c_1 N}} \prod_{k=2}^{N-1} \frac{1}{1-\frac{t_k}{1-t_k}\frac{it}{c_1 N^2}}$$

$$\sim \prod_{l=1}^{N-2} \frac{1}{1-\frac{it}{c_1 l(l+3)}} \to \Phi\left(t\right),$$

where $\Phi\left(t\right)=\prod_{l=1}^{\infty}\frac{1}{1-\frac{i(t/c_1)}{l(l+3)}}$. $\Phi\left(t\right)$ is the characteristic function of the random variable

$$T = \sum_{l\geq 1} \frac{1}{l\left(l+3\right)} E_l\left(c_1\right),$$

where the $E_l\left(c_1\right)$ are iid random variables, with common distribution $\mathbb{P}\left(E_1\left(c_1\right)>x\right)=e^{-c_1 x}$. The limiting T is thus a weighted sum of iid exponentially-distributed random variables with quadratic weights. Such random variables are known to have all their integral moments finite (see [3]). For instance

$$\mathbb{E}\left(T\right) = \frac{1}{c_1}\sum_{l\geq 1}\frac{1}{l\left(l+3\right)} = \frac{1}{3c_1}\left(\sum_{l\geq 1}\frac{1}{l} - \sum_{l\geq 4}\frac{1}{l}\right) = \frac{11}{18c_1}. \quad \triangle$$

We conclude from (23) that as N is large

$$\text{sep}\left(\boldsymbol{\pi}_{n,1},\boldsymbol{\pi}\right) = \mathbb{P}\left(\widetilde{\tau}_{1,N-1}>n\right) \sim \overline{F}_T\left(\frac{n}{c_1 N^2}\right)$$

where $\overline{F}_T\left(x\right)=\mathbb{P}\left(T>x\right)$ is the complementary probability distribution of $T>0$. This shows that, for the recurrent neutral Moran model, the transition to stationarity occurs gradually after about $c_1 N^2$ steps. There is no abrupt entrance of this particular chain into the invariant measure (no cutoff phenomenon).

A plausible scaling limit. Let $Y_n = X_n - 1$ be a shifted version of the recurrent model with state-space $\{0,...,N-2\}$. First, we observe that Y_n is a Moran model with mutations (see [14] p. 107). Indeed, from (6, 8), the transition probabilities of Y_n read: $p_y = \theta_2^{-1}\frac{y+2}{N}\left(1-\frac{y+2}{N}\right)$, $q_y = \theta_2^{-1}\frac{y}{N}\left(1-\frac{y}{N}\right)$, $r_y = 1-\left(p_y+q_y\right)$. Recalling $\theta_2 = 1-\frac{2}{N^2-2}$, these may be recast as

$$p_y = \left(1-\frac{y}{N-2}\right)\pi\left(\frac{y}{N-2}\right)$$

$$q_y = \frac{y}{N-2}\left(1-\pi\left(\frac{y}{N-2}\right)\right)$$
$$r_y = \frac{y}{N-2}\pi\left(\frac{y}{N-2}\right)+\left(1-\frac{y}{N-2}\right)\left(1-\pi\left(\frac{y}{N-2}\right)\right),$$

where

$$\pi(u) = \mu_1(1-u)+(1-\mu_2)u$$

is an affine mutation mapping from $[0,1]$ to $[0,1]$, with mutation probabilities from type 2 to 1 and from type 1 to 2

$$\mu_1 = \frac{2(N-2)}{N^2-2} \sim \frac{2}{N} \text{ and } \mu_2 = \frac{2(N-1)}{N^2-2} \sim \frac{2}{N}.$$

Let now $\Delta > 0$ be a small constant. Let $t = n\Delta$ be a discretization of time and consider the scaled process $U_t = Y_{n\Delta}/(N-2) \in [0,1]$. Given $U_t = u$, assume

$$\begin{aligned} U_{t+\Delta} &= u \pm \frac{1}{N} \text{ with probability } \Delta \cdot p_{(N-2)u} \text{ (respectively } \Delta \cdot q_{(N-2)u}) \\ &= u \text{ with probability } 1-\Delta \cdot r_{(N-2)u}. \end{aligned}$$

Then $\mathbb{E}_u(U_{t+\Delta}-u) = \frac{\Delta}{N-2}\left(p_{(N-2)u}-q_{(N-2)u}\right) = \frac{\Delta}{N-2}(\pi(u)-u)$, with $\pi(u)-u \sim \frac{1}{N}(2-4u)$ for large N. Thus

$$\frac{1}{\Delta}\mathbb{E}_u(U_{t+\Delta}-u) \sim \frac{1}{N^2}(2-4u).$$

Similarly, $\mathbb{E}_u\left[(U_{t+\Delta}-u)^2\right] = \frac{\Delta}{(N-2)^2}\left(1-r_{(N-2)u}\right)$ leading to the dominant order in N to

$$\frac{1}{\Delta}\mathbb{E}_u\left[(U_{t+\Delta}-u)^2\right] \sim \frac{2}{N^2}u(1-u).$$

Following ([28], Chapter 2), this shows (by passing to the limit in Δ), that $U_t \in [0,1]$ obeys the Wright-Fisher-like diffusion approximation equation driven by Brownian motion (B_t)

$$dU_t = \frac{2}{N^2}(1-2U_t)\,dt + \frac{1}{N}\sqrt{2U_t(1-U_t)}dB_t.$$

Performing the time substitution $\tau = N^{-2}t$ and defining $V_\tau = U_{N^2\tau}$, V_τ obeys the diffusion equation

$$dV_\tau = 2(1-2V_\tau)\,d\tau + \sqrt{2V_\tau(1-V_\tau)}dB_\tau$$

with a unit time τ corresponding to N^2 generations in the discrete model. The above considerations on the time to reach stationarity of X_n suggests that V_τ enters gradually into its invariant measure (which is $6v(1-v)\,dv$) by time $O(1)$.

5. Other transition times for the recurrent Moran model

Consider the ergodic Moran birth and death Markov chain X_n on the state-space $\{1,..,N-1\}$. It is well-known that the mean return time to state $\{1\}$ is $\mathbb{E}(\tau_{1,1}) = \frac{1}{\pi_1}$. Thus, using (9)

$$\mathbb{E}(\tau_{1,1}) \sim \frac{N^2}{6} \text{ as } N \to \infty.$$

We wish now to estimate the time it takes for X_n to move from one end of the state-space to the other, that is from state $\{1\}$ to $\{N-1\}$. Let $\tau_{1,N-1}$ be this random time. We shall prove the following estimation of its mean value:

Proposition 4.

$$\mathbb{E}\left(\tau_{1,N-1}\right) \sim \frac{N^3}{6} \text{ as } N \to \infty.$$

Proof: Let indeed $\tau_{x,x+1}$ be the random time to first hit the state $\{x+1\}$ starting from the state $\{x\}$. Depending on whether the move starting in x is up, down or no move, $\tau_{x,x+1}$ is either 1 or $1+\tau^{'}_{x,x+1}$ (with $\tau^{'}_{x,x+1}$ a statistical copy of $\tau_{x,x+1}$) or $1+\tau_{x-1,x}+\tau^{'}_{x,x+1}$. If we let μ_x be the mean value of $\tau_{x,x+1}$, we thus get $\mu_x = p_x + r_x\left(1+\mu_x\right) + q_x\left(1+\mu_{x-1}+\mu_x\right)$, $x \geq 2$. This leads to the recurrence ($\mu_1 = 1/p_1$):

$$\mu_x = \frac{q_x}{p_x}\mu_{x-1} + \frac{1}{p_x},\ x \geq 2.$$

This recurrence can be solved to give

$$\mu_x = \frac{1}{p_x \pi_x}\sum_{y=1}^{x}\pi_y.$$

Thus

$$\mathbb{E}\left(\tau_{1,N-1}\right) = \sum_{x=1}^{N-2}\mu_x = \sum_{x=1}^{N-2}\frac{1}{p_x\pi_x}\sum_{y=1}^{x}\pi_y.$$

Looking at this sum formula, one expects that its leading term is $\frac{1}{p_{N-2}\pi_{N-2}}\sum_{y=1}^{N-2}\pi_y$ because this is where $\sum_{y=1}^{x}\pi_y$ is the largest and π_x the smallest. Looking at the expressions of p_x and π_x at $x = N-2$ in (6, 9), this crude estimation would lead to $\mathbb{E}\left(\tau_{1,N-1}\right) \sim N^3/12$ but this is not quite true because the other terms contribute equally likely.

In fact, the computations can be done explicitly. From (9), we have

$$\sum_{y=1}^{x}\pi_y = \frac{x\left(x+1\right)\left(3N-\left(2x+1\right)\right)}{N\left(N-1\right)\left(N+1\right)}$$

and so

$$\begin{aligned}
\mathbb{E}\left(\tau_{1,N-1}\right) &= \frac{\theta_2 N^2}{6}\sum_{x=1}^{N-2}\frac{3N-\left(2x+1\right)}{\left(N-x\right)\left(N-x-1\right)} \\
&= \frac{\theta_2 N^2}{6}\sum_{y=1}^{N-2}\frac{N+1+2y}{y\left(y+1\right)} = \frac{\theta_2 N^2}{6}\left(\sum_{y=1}^{N-2}\frac{N+1}{y} - \sum_{y=1}^{N-2}\frac{N-1}{y+1}\right) \\
&= \frac{\theta_2 N^2}{6}\left[\left(N+1\right)H_{N-2} - \left(N-1\right)\left(H_{N-1}-1\right)\right]
\end{aligned}$$

where H_N is the $N-$harmonic number. Thus

$$\begin{aligned}
\mathbb{E}\left(\tau_{1,N-1}\right) &= \frac{\theta_2 N^2}{6}\left[2H_{N-1} - \frac{N+1}{N-1} + N - 1\right] \\
&\sim \frac{N^3}{6} \text{ as } N \to \infty. \quad \triangle
\end{aligned}$$

References

[1] Aldous, D.; Diaconis, P. Strong uniform times and finite random walks. Adv. in Appl. Math. 8, no. 1, 69–97, (1987).

[2] Athreya, K. B.; Ney, P. E. *Branching processes.* Die Grundlehren der mathematischen Wissenschaften, Band 196. Springer-Verlag, New York-Heidelberg, 1972. xi+287 pp.

[3] Biane, P.; Pitman, J.; Yor, M. Probability laws related to the Jacobi theta and Riemann zeta functions, and Brownian excursions. Bull. Amer. Math. Soc. (N.S.) 38, (2001), no. 4, 435–465.

[4] Chen, J. Extremality of invariant measures and ergodicity of stochastic systems. J. Phys. A 35 (2002), no. 2, 229–238.

[5] Dette, H.; Fill, J. A.; Pitman, J.; Studden, W. J. Wall and Siegmund duality relations for birth and death chains with reflecting barrier. Dedicated to Murray Rosenblatt. J. Theoret. Probab. 10, (1997), no. 2, 349–374.

[6] Diaconis, P. The cutoff phenomenon in finite Markov chains. Proc. Natl. Acad. Sci. USA 36, (1996), 1659–1664.

[7] Diaconis, P.; Fill, J. A. Strong stationary times via a new form of duality. Ann. Probab. 18, no. 4, 1483–1522, (1990).

[8] Diaconis, P.; Saloff-Coste, L. Separation cut-offs for birth and death chains. Ann. Appl. Probab. 16, (2006), no. 4, 2098–2122.

[9] Diaconis, P.; Shahshahani, M. Time to reach stationarity in the Bernoulli-Laplace diffusion model. SIAM J. Math. Anal. 18 (1987), no. 1, 208–218.

[10] Donnelly, P. *Dual processes in population genetics.* Stochastic spatial processes (Heidelberg, 1984), 94–105, Lecture Notes in Math., 1212, Springer, Berlin, 1986.

[11] Donnelly, P.; Rodrigues, E. R. Convergence to stationarity in the Moran model. J. Appl. Probab. 37, (2000), no. 3, 705–717.

[12] Dunham, B. Fluctuation theory for Moran's genetics model. J. Math. Anal. Appl. 210, (1997), no. 2, 777–789.

[13] Dynkin, E. B. *Markov processes.* Vols. I, II. Translated with the authorization and assistance of the author by J. Fabius, V. Greenberg, A. Maitra, G. Majone. Die Grundlehren der Mathematischen Wissenschaften, Bände 121, 122 Academic Press Inc., Publishers, New York; Springer-Verlag, Berlin-Göttingen-Heidelberg 1965.

[14] Ewens, W. J. *Mathematical population genetics. I. Theoretical introduction.* Second edition. Interdisciplinary Applied Mathematics, 27. Springer-Verlag, New York, 2004.

[15] Feller, W. *An introduction to probability theory and its applications,* Vol. **1**. John Wiley and Sons, Third Edition, New York, 1968.

[16] Fill, J. A. The passage time distribution for a birth-and-death chain: Strong stationary duality gives a first stochastic proof. Journal of Theoretical Probability, Volume 22, Number 3, 543-557, (2009).

[17] Gillespie, J. H. *The Causes of Molecular Evolution.* New York and Oxford: Oxford University Press, 1991.

[18] Griffeath, D. Frank Spitzer's pioneering work on interacting particle systems. Ann. Probab. 21 (1993), no. 2, 608–621.

[19] Huillet, T. A Duality Approach to the Genealogies of Discrete Non-Neutral Wright-Fisher Models. Journal of Probability and Statistics, vol. 2009, Article ID 714701, 22 pages, (2009).

[20] Huillet, T.; Möhle, M. Duality and asymptotics for a class of nonneutral discrete Moran models. J. Appl. Probab. 46, (2009), no. 3, 866–893.

[21] Huillet, T. Reversing the drift of the Ehrenfest urn model and three conditionings. J. Phys. A 42 (2009), no. 34, 345005, 8 pp.

[22] Huillet, T.; Martínez, S. Duality and Intertwining for discrete Markov kernels: a relation and examples. arXiv hal-00401732, (2009).

[23] Karlin, S.; McGregor, J. On a genetics model of Moran. Proc. Cambridge Philos. Soc. 58, 299–311, (1962).

[24] Keilson, J. *Markov chain models - rarity and exponentiality.* Applied Mathematical Sciences, 28. Springer-Verlag, New York-Berlin, 1979.

[25] Lambert, A. Population dynamics and random genealogies. Stochastic Models 24, suppl. 1, 45–163, (2008).

[26] Lamperti, J.; Ney, P. Conditioned branching processes and their limiting diffusions. Teor. Verojatnost. i Primenen. 13, (1968), 126–137.
[27] Liggett, T. M. *Interacting particle systems.* Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 276. Springer-Verlag, New York, 1985.
[28] Maruyama, T. *Stochastic problems in population genetics.* Lecture Notes in Biomathematics, 17. Springer-Verlag, Berlin-New York, 1977.
[29] Möhle, M. The concept of duality and applications to Markov processes arising in neutral population genetics models. Bernoulli 5, (1999), no. 5, 761–777.
[30] Siegmund, D. The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes. Ann. Probability 4, (1976), no. 6, 914–924.
[31] Sudbury, A.; Lloyd, P. Quantum operators in classical probability theory. II. The concept of duality in interacting particle systems. Ann. Probab. 23, (1995), no. 4, 1816–1830.
[32] Spitzer, F. Interaction of Markov processes. Advances in Math. 5, (1970), 246–290.

Laboratoire de Physique Théorique et Modélisation, CNRS-UMR 8089 et Université de Cergy-Pontoise, 2 Avenue Adolphe Chauvin, 95302, Cergy-Pontoise, FRANCE, E-mail: Thierry.Huillet@u-cergy.fr