Electron. Commun. Probab. **24** (2019), no. 15, 1–14.
https://doi.org/10.1214/19-ECP222
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

A spectral decomposition for a simple mutation model

Martin Möhle*

Abstract

We consider a population of N individuals. Each individual has a type belonging to some at most countable type space K. At each time step each individual of type $k \in K$ mutates to type $l \in K$ independently of the other individuals with probability $m_{k,l}$. It is shown that the associated empirical measure process is Markovian. For the two-type case $K = \{0,1\}$ we derive an explicit spectral decomposition for the transition matrix P of the Markov chain $Y = (Y_n)_{n\geq 0}$, where Y_n denotes the number of individuals of type 1 at time n. The result in particular shows that P has eigenvalues $(1 - m_{0,1} - m_{1,0})^i$, $i \in \{0,\ldots,N\}$. Applications to mean first passage times are provided.

Keywords: eigenvalues; eigenvectors; empirical measure process; finite Markov chain; first passage time; mixing time; mutation model; potential theory; product chain; random walk on the hypercube; spectral analysis.
AMS MSC 2010: Primary 60J10; 15A18, Secondary 60J45; 92D10.
Submitted to ECP on August 11, 2018, final version accepted on February 26, 2019.

1 Introduction and main results

Consider a population consisting of $N \in \mathbb{N} := \{1,2,\ldots\}$ individuals. It is assumed that each individual has a type belonging to some at most countable type space K. At each time step each individual of type $k \in K$ mutates its type independently of the other individuals to type $l \in K$ with given probability $m_{k,l} \in [0,1]$. We call the stochastic matrix $M := (m_{k,l})_{k,l\in K}$ the mutation matrix.

For $n \in \mathbb{N}_0 := \{0,1,\ldots\}$ and $r \in \{1,\ldots,N\}$ let $X_n^{(r)}$ denote the type of individual r at time step n. Clearly, $X := (X_n)_{n\in\mathbb{N}_0}$, defined via $X_n := (X_n^{(1)},\ldots,X_n^{(N)})$ for all $n \in \mathbb{N}_0$, is a homogeneous discrete-time Markov chain with state space K^N. Since for each $r \in \{1,\ldots,N\}$ the process $(X_n^{(r)})_{n\in\mathbb{N}_0}$ is a Markov chain and since the types of the N individuals evolve independently, the Markov chain X is a so-called product chain having n-step transition probabilities

$$\pi_{\mathbf{i},\mathbf{j}}^{(n)} \;:=\; \mathbb{P}(X_n = \mathbf{j}\,|\,X_0 = \mathbf{i}) \;=\; \prod_{r=1}^N \mathbb{P}(X_n^{(r)} = j_r\,|\,X_0^{(r)} = i_r) \;=\; \prod_{r=1}^N m_{i_r,j_r}^{(n)} \tag{1.1}$$

for all $n \in \mathbb{N}_0$ and all $\mathbf{i} = (i_1,\ldots,i_N), \mathbf{j} = (j_1,\ldots,j_N) \in K^N$, where $M^n = (m_{k,l}^{(n)})_{k,l\in K}$ denotes the n-step mutation matrix. The process $\eta := (\eta_n)_{n\in\mathbb{N}_0}$, defined via

$$\eta_n \;:=\; \sum_{r=1}^N \delta_{X_n^{(r)}}, \qquad n \in \mathbb{N}_0, \tag{1.2}$$

*Eberhard Karls Universität Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany.
E-mail: martin.moehle@uni-tuebingen.de

is called the empirical measure process of X. Let $\mathcal{K}$ denote the power set of K. Clearly, η has state space $\mathcal{M}$, the set of measures μ on $(K,\mathcal{K})$ with values in $\{0,\ldots,N\}$ and total mass $\mu(K)=N$. It is a priori not totally obvious that η is still Markovian, since functions of Markov chains are in general not Markovian anymore. Proposition 1.1 below shows that η still enjoys the Markov property and provides a formula for its n-step transition probabilities. The process of going from X to η is an example of what is called projection or lumping of Markov chains in the literature. The proof of Proposition 1.1 is provided in Section 4. In the following, for any measure $\mu\in\mathcal{M}$, we use the notation $\mu_k:=\mu(\{k\})$, $k\in K$.

Proposition 1.1. *The empirical measure process η of the product chain X is a homogeneous discrete-time Markov chain with state space $\mathcal{M}$, the set of measures μ on $(K,\mathcal{K})$ with values in $\{0,\ldots,N\}$ and total mass $\mu(K)=N$, and n-step transition probabilities*

$$p_{\nu,\mu}^{(n)} \;:=\; \mathbb{P}(\eta_n=\mu\,|\,\eta_0=\nu) \;=\; \sum_T\prod_{k\in K}\biggl(\nu_k!\biggl(\prod_{l\in K}\frac{(m_{k,l}^{(n)})^{t_{k,l}}}{t_{k,l}!}\biggr)\biggr),\qquad \nu,\mu\in\mathcal{M}, n\in\mathbb{N}_0,$$

where the sum $\sum_T$ extends over all $T=(t_{k,l})_{k,l\in K}\in\mathbb{N}_0^{K\times K}$ with marginals $\sum_{l\in K}t_{k,l}=\nu_k$, $k\in K$, and $\sum_{k\in K}t_{k,l}=\mu_l$, $l\in K$.

From now on we restrict to the two-type situation $K=\{0,1\}$ and write the mutation matrix $M=(m_{k,l})_{k,l\in\{0,1\}}$ as well in the form

$$M \;=\; \begin{pmatrix} 1-a & a\\ b & 1-b\end{pmatrix}$$

to avoid indices. The model has hence three parameters, the population size N and the two mutation probabilities $a=m_{0,1}$ and $b=m_{1,0}$. We exclude the two trivial cases $a=b=0$ and $a=b=1$. Note that the entries of

$$M^n \;=\; (m_{k,l}^{(n)})_{k,l\in\{0,1\}} \;=\; \begin{pmatrix} 1-a_n & a_n\\ b_n & 1-b_n\end{pmatrix}$$

are explicitly known, namely $a_n=m_{0,1}^{(n)}=a(a+b)^{-1}(1-(1-a-b)^n)$ and $b_n=m_{1,0}^{(n)}=b(a+b)^{-1}(1-(1-a-b)^n)$, $n\in\mathbb{N}_0$.

For the particular situation $a=b$ this model was studied by Scoppola [12] and in a preprint of Berger and Cerf [1]. We allow here for a general stochastic mutation matrix M. Nestoridi [10] studies a different but slightly related random walk on the hypercube $\{0,1\}^N$ which at each step flips a fixed number k of randomly chosen coordinates.

As in [1] we are interested in the stochastic process $Y:=(Y_n)_{n\in\mathbb{N}_0}$, defined via $Y_n:=\eta_n(\{1\})=\sum_{r=1}^N X_n^{(r)}=:\|X_n\|_1$ for all $n\in\mathbb{N}_0$. Note that Y_n counts the individuals of type 1 at time n. Corollary 1.2 below is a straightforward consequence of Proposition 1.1. In the following we use for the binomial probabilities the notation $\mathrm{B}(n,p,k):=\binom{n}{k}p^k(1-p)^{n-k}$, $n\in\mathbb{N}_0$, $p\in[0,1]$, $k\in\{0,\ldots,n\}$.

Corollary 1.2. *The process Y is a homogeneous discrete-time Markov chain with state space $S:=\{0,\ldots,N\}$ and n-step transition probabilities*

$$p_{i,j}^{(n)} \;:=\; \mathbb{P}(Y_n=j\,|\,Y_0=i) \;=\; \sum_{k=0}^{\min(i,j)}\mathrm{B}(i,1-b_n,k)\,\mathrm{B}(N-i,a_n,j-k),\quad i,j\in S, n\in\mathbb{N}_0, \tag{1.3}$$

where $a_n=m_{0,1}^{(n)}=a(a+b)^{-1}(1-(1-a-b)^n)$ and $b_n=m_{1,0}^{(n)}=b(a+b)^{-1}(1-(1-a-b)^n)$, $n\in\mathbb{N}_0$.

Remark 1.3. Let $n \in \mathbb{N}_0$ and $i \in S$. Corollary 1.2 shows that $p_{i,j}^{(n)} = \mathbb{P}(K_n + L_n = j)$, where K_n and L_n are independent random variables with K_n binomially distributed with parameters i and $1 - b_n$ and L_n binomially distributed with parameters $N - i$ and a_n. From $|1 - a - b| < 1$ we conclude that $\lim_{n\to\infty} a_n = a/(a + b)$ and $\lim_{n\to\infty} b_n = b/(a + b)$. Thus, $K_n \to K_\infty$ and $L_n \to L_\infty$ in distribution as $n \to \infty$, where K_∞ and L_∞ are independent with K_∞ binomially distributed with parameters i and $a/(a + b)$ and L_∞ binomially distributed with parameters $N - i$ and $a/(a + b)$. It follows that

$$\lim_{n\to\infty} p_{i,j}^{(n)} \; = \; \mathbb{P}(K_\infty + L_\infty = j) \; = \; \mathrm{B}\Big(N, \frac{a}{a+b}, j\Big) \; =: \; \varrho_j, \qquad i, j \in S. \tag{1.4}$$

The stationary distribution $(\varrho_j)_{j\in S}$ of Y is hence the binomial distribution with parameters N and $a/(a + b)$.

Let us now turn to spectral analysis. Spectral decompositions of transition matrices P for Ehrenfest type models [5] have been provided by Kac [8]. Ehrenfest type models belong to the class of nearest neighborhood models or local Markov chains allowing for jumps only to the nearest neighbors (or at least to close neighbors) with positive probability. In contrast, the Markov chain Y under consideration is non-local satisfying $p_{i,j} > 0$ for all $i, j \in S$. In this situation it is usually much harder to find the eigenvalues and corresponding eigenvectors of P. The literature on (examples of) such classes of full occupied matrices with explicitly known spectral decomposition is hence somewhat more sparse. One example from mathematical population genetics are transition matrices of the forward process of exchangeable Cannings population models [3, 4, 6]. Another well known example, being important in the theory of discrete Fourier transforms, are matrices of the form $P = (a_{(i+j) \bmod N+1})_{i,j\in\{0,\ldots,N\}}$ for some given sequence $(a_0, \ldots, a_N)$.

In the following it is assumed that $a \neq 0$ and $b \neq 0$ to avoid trivialities. In order to state the main result (Theorem 1.4 below), let us introduce the left lower triangular matrix $A := (a_{i,j})_{i,j\in S}$ and the right upper triangular matrix $B := (b_{i,j})_{i,j\in S}$ via

$$a_{i,j} \; := \; \binom{i}{j}\Big(-\frac{b}{a}\Big)^{i-j} \quad \text{and} \quad b_{i,j} \; := \; \binom{N-i}{j-i}\Big(\frac{a+b}{a}\Big)^{i}, \quad i, j \in S. \tag{1.5}$$

Note that A and B are non-singular with inverses A^{-1} and B^{-1} having entries

$$(A^{-1})_{i,j} \; = \; \binom{i}{j}\Big(\frac{b}{a}\Big)^{i-j} \quad \text{and} \quad (B^{-1})_{i,j} \; = \; (-1)^{j-i}\binom{N-i}{j-i}\Big(\frac{a}{a+b}\Big)^{j}, \quad i, j \in S. \tag{1.6}$$

The main result (Theorem 1.4) provides an explicit spectral decomposition for the transition matrix of the Markov chain Y. Its proof is provided in Section 4.

Theorem 1.4. *Assume that $a \neq 0$ and $b \neq 0$. Then the transition matrix $P = (p_{i,j})_{i,j\in S}$ of the Markov chain Y has a spectral decomposition of the form*

$$P \; = \; RDL \tag{1.7}$$

with $R := AB$ and $L := R^{-1} = B^{-1}A^{-1}$, where A and B are defined via (1.5) and D is the diagonal matrix with entries $d_{i,i} := (m_{1,1} - m_{0,1})^{N-i} = (1 - a - b)^{N-i}$, $i \in S$. In particular, P has eigenvalues $\lambda_i := d_{N-i,N-i} = (1 - a - b)^i$, $i \in S$.

Remark 1.5. The matrices $R = (r_{i,j})_{i,j\in S}$ and $L = (l_{i,j})_{i,j\in S}$ have entries

$$r_{i,j} \; = \; \sum_{k\in S} a_{i,k}b_{k,j} \; = \; \sum_{k=0}^{\min(i,j)} \binom{i}{k}\Big(-\frac{b}{a}\Big)^{i-k}\binom{N-k}{j-k}\Big(\frac{a+b}{a}\Big)^{k}, \quad i, j \in S, \tag{1.8}$$

and

$$\begin{aligned} l_{i,j} &= \sum_{k\in S}(B^{-1})_{i,k}(A^{-1})_{k,j} \\ &= \sum_{k=\max(i,j)}^{N}(-1)^{k-i}\binom{N-i}{k-i}\left(\frac{a}{a+b}\right)^{k}\binom{k}{j}\left(\frac{b}{a}\right)^{k-j}, \quad i,j\in S. \end{aligned} \tag{1.9}$$

The last column of R contains the (obvious) right eigenvector $(1,\ldots,1)^T$ to the eigenvalue $(1-a-b)^0=1$. The last row of L contains the probabilities $l_{N,j}=\mathrm{B}(N,a/(a+b),j)=\varrho_j$, $j\in S$, of the stationary distribution of Y. For $a=b$ the matrices R and L only depend on the population size N but not on the parameters a and b.

Lemma 5.2 in the appendix (Section 5) shows that R and L have horizontal generating functions

$$r_i(z) := \sum_{j\in S} r_{i,j}z^j = (z+1)^{N-i}\left(z-\frac{b}{a}\right)^{i}, \qquad i\in S, z\in\mathbb{C},$$

and

$$l_i(z) := \sum_{j\in S} l_{i,j}z^j = \left(\frac{a}{a+b}\right)^{N}\left(z+\frac{b}{a}\right)^{i}(1-z)^{N-i}, \qquad i\in S, z\in\mathbb{C}.$$

From

$$l_i(-z) = \left(\frac{a}{a+b}\right)^{N}\left(-z+\frac{b}{a}\right)^{i}(1+z)^{N-i} = \left(\frac{a}{a+b}\right)^{N}(-1)^i r_i(z), \quad i\in S, z\in\mathbb{C},$$

it follows that the entries of the matrices R and L are related via

$$l_{i,j} = (-1)^{i-j}\left(\frac{a}{a+b}\right)^{N} r_{i,j} \quad\text{and}\quad r_{i,j} = (-1)^{i-j}\left(\frac{a+b}{a}\right)^{N} l_{i,j}, \quad i,j\in S.$$

2 Applications to mean passage times

Spectral decompositions of transition matrices $P=(p_{i,j})_{i,j\in S}$ of Markov chains have many applications, in particular in the context of potential theory of Markov chains. For example, each spectral decomposition $P=RDL$ with $L=R^{-1}$ implies as well a spectral decomposition of the resolvent $R_\alpha := (I-\alpha P)^{-1}$, $\alpha\in(0,1)$, namely $R_\alpha=\sum_{n\ge 0}\alpha^n P^n = R(\sum_{n\ge 0}\alpha^n D^n)L = RD_\alpha L$, where D_α denotes the diagonal matrix with diagonal entries $\sum_{n\ge 0}\alpha^n d_{i,i}^n = 1/(1-\alpha d_{i,i})$, $i\in S$. For some more information on the resolvent we refer the reader to Norris [11, p. 146].

We provide here another application concerning mean passage times. Define $W := (w_{i,j})_{i,j\in S} := \lim_{n\to\infty} P^n$. Clearly, $WP=W$ and $W^2=W$ and, hence, $(P-W)^n = P^n - W$ for all $n\in\mathbb{N}$. Let $F=(f_{i,j})_{i,j\in S}$ denote the fundamental matrix associated with the Markov chain Y, i.e.

$$F := (I-P+W)^{-1} = \sum_{n\ge 0}(P-W)^n = I + \sum_{n\ge 1}(P^n - W).$$

For $i,j\in S$ let $\tau_{i,j} := \inf\{n\in\mathbb{N}\,|\,Y_0=i, Y_n=j\}$ denote the first time step $n\in\mathbb{N}$ when the chain Y reaches state j if started from state i. It is well known that $\mathbb{E}(\tau_{j,j})=1/\varrho_j$, $j\in S$, where $(\varrho_j)_{j\in S}$ denotes the stationary distribution (1.4) of Y. Now let $i\ne j$. Then $\tau_{i,j}$ has mean (see, for example, Grinstead and Snell [7, p. 459, Theorem 11.16])

$$\mathbb{E}(\tau_{i,j}) = \frac{f_{j,j}-f_{i,j}}{\varrho_j}.$$

Since, for all $i\neq j$,

$$\begin{aligned} f_{j,j}-f_{i,j} &= 1+\sum_{n\ge1}(p_{j,j}^{(n)}-w_{j,j})-\sum_{n\ge1}(p_{i,j}^{(n)}-w_{i,j})\\ &= 1+\sum_{n\ge1}(p_{j,j}^{(n)}-p_{i,j}^{(n)}+\underbrace{w_{i,j}-w_{j,j}}_{=\varrho_j-\varrho_j=0})\\ &= 1+\sum_{n\ge1}(p_{j,j}^{(n)}-p_{i,j}^{(n)}) \;=\; \sum_{n\ge0}(p_{j,j}^{(n)}-p_{i,j}^{(n)}), \end{aligned}$$

we can summaries these results in the form

$$\mathbb{E}(\tau_{i,j}) \;=\; \frac{\delta_{i,j}}{\varrho_j}+\frac{1}{\varrho_j}\sum_{n\ge0}(p_{j,j}^{(n)}-p_{i,j}^{(n)}),\qquad i,j\in S. \tag{2.1}$$

Explicit formulas for $p_{i,j}^{(n)}$ are available for the present example (see (1.3)). However, the calculation of the series on the right hand side in (2.1) is inconvenient. Fortunately, thanks to the spectral decomposition of P, this series can be expressed as a finite sum and can hence be calculated rather easily as follows. We have $r_{i,N}=1$ for all $i\in S$ and, hence,

$$\begin{aligned} \sum_{n\ge0}(p_{j,j}^{(n)}-p_{i,j}^{(n)}) &= \sum_{n\ge0}\sum_{k=0}^{N}(r_{j,k}\lambda_{N-k}^n l_{k,j}-r_{i,k}\lambda_{N-k}^n l_{k,j})\\ &= \sum_{n\ge0}\sum_{k=0}^{N-1}(r_{j,k}-r_{i,k})l_{k,j}\lambda_{N-k}^n \;=\; \sum_{k=0}^{N-1}\frac{(r_{j,k}-r_{i,k})l_{k,j}}{1-\lambda_{N-k}},\qquad i,j\in S. \end{aligned}$$

3 Final remarks and open problems

1. It is natural to conjecture that the transition matrix $\Pi=(\pi_{\mathbf{i},\mathbf{j}})_{\mathbf{i},\mathbf{j}\in\{0,1\}^N}$ of the original product Markov chain X has as well (as Y) the eigenvalues $\lambda_i:=(1-a-b)^i$, $i\in\{0,\ldots,N\}$, where the right and left eigenspaces to the eigenvalue λ_i have both dimension $\binom{N}{i}$, $i\in\{0,\ldots,N\}$. We leave the proof of this conjecture for future research.

2. Let us briefly come back to the multi-type model with (at most countable) type space K. For $n\in\mathbb{N}_0$ and $k\in K$ let $Y_n^{(k)}:=\eta_n(\{k\})=\sum_{r=1}^N 1_{\{X_n^{(r)}=k\}}$ denote the number of individuals of type k at time n. From Proposition 1.1 it follows that the process $Y=(Y_n)_{n\in\mathbb{N}_0}$, defined via $Y_n:=(Y_n^{(k)})_{k\in K}$ for all $n\in\mathbb{N}_0$, is a homogeneous discrete-time Markov chain with state space $S:=\{(i_k)_{k\in K}\in\mathbb{N}_0^K\,|\,\sum_{k\in K}i_k=N)$ and n-step transition probabilities

$$p_{i,j}^{(n)} \;=\; \mathbb{P}(Y_n=j\,|\,Y_0=i) \;=\; \sum_T\prod_{k\in K}\left(i_k!\left(\prod_{l\in K}\frac{(m_{k,l}^{(n)})^{t_{k,l}}}{t_{k,l}!}\right)\right) \tag{3.1}$$

for all $n\in\mathbb{N}_0$ and all $i=(i_k)_{k\in K}, j=(j_k)_{k\in K}\in S$, where the sum $\sum_T$ extends over all $T:=(t_{k,l})_{k,l\in K}\in\mathbb{N}_0^{K\times K}$ with marginals $\sum_{l\in K}t_{k,l}=i_k$, $k\in K$, and $\sum_{k\in K}t_{k,l}=j_l$, $l\in K$. Finding spectral decompositions of the transition matrices of X or Y for the multi-type model is a challenging open problem.

3. Let us finally provide some (mainly non-rigorous) information on the mixing time of the chain Y. Assume that $\mathbb{P}(Y_0=i)=1$ for some given fixed state $i\in S$. Let Y_∞ be binomially distributed with parameters N and $a/(a+b)$. Theorem 1.4 yields an exact

formula for the total variation distance

$$d(Y_n, Y_\infty) \;:=\; \frac{1}{2}\sum_{j=0}^N |\mathbb{P}(Y_n = j) - \mathbb{P}(Y_\infty = j)| \;=\; \frac{1}{2}\sum_{j=0}^N |p_{i,j}^{(n)} - \varrho_j| \tag{3.2}$$

$$= \;\frac{1}{2}\sum_{j=0}^N \Big|\sum_{k=0}^N r_{i,k} d_{k,k}^n l_{k,j} - \varrho_j\Big| \;=\; \frac{1}{2}\sum_{j=0}^N \Big|\sum_{k=0}^{N-1} r_{i,k} d_{k,k}^n l_{k,j}\Big|, \tag{3.3}$$

since $r_{i,N} = 1$, $d_{N,N} = \lambda_0 = 1$ and $l_{N,j} = \varrho_j$. In the following it is assumed for simplicity that $a + b < 1$ such that all eigenvalues $\lambda_i = (1 - a - b)^i$, $i \in S$, are strictly positive. Note that the spectral gap $1 - \lambda_1 = a + b$ does not depend on N. From $d_{k,k} = \lambda^{N-k}$ with $\lambda := \lambda_1 = 1 - a - b$ we conclude that

$$d(Y_n, Y_\infty) \;\sim\; \frac{1}{2}\sum_{j=0}^N |r_{i,N-1} d_{N-1,N-1}^n l_{N-1,j}| \;=\; \lambda^n c_{N,i}, \qquad n \to \infty, \tag{3.4}$$

with coefficients $c_{N,i} := \frac{1}{2}|r_{i,N-1}| \sum_{j=0}^N |l_{N-1,j}|$. The coefficients $c_{N,i}$ are explicitly known, since, by Theorem 1.4 and the remarks thereafter,

$$r_{i,N-1} \;=\; N - \frac{a+b}{a} i \quad \text{and} \quad l_{N-1,j} \;=\; \varrho_j\left(\frac{a}{b} - \frac{a+b}{b}\frac{j}{N}\right).$$

Fix $\varepsilon \in (0,1)$. Based on (3.4) it is reasonable to conjecture that a good approximation for the mixing time $t_{\mathrm{mix}}(\varepsilon) := \inf\{n \in \mathbb{N}_0 : d(Y_n, Y_\infty) < \varepsilon\}$ of the chain Y is

$$t_{\mathrm{mix}}(\varepsilon) \;\approx\; \left\lfloor \frac{\log c_{N,i} - \log\varepsilon}{-\log\lambda} \right\rfloor + 1. \tag{3.5}$$

Note however that this approximation is a non-rigorous result. For large N one may further simplify the approximation (3.5) as follows. We have

$$\sum_{j=0}^N |l_{N-1,j}| \;=\; \frac{a}{b}\sum_{j=0}^N \left|\frac{a+b}{a}\frac{j}{N} - 1\right|\varrho_j \;=\; \frac{a}{b}\mathbb{E}\left(\left|\frac{Y_\infty}{\mathbb{E}(Y_\infty)} - 1\right|\right) \;\sim\; \sqrt{\frac{2a}{\pi b N}}, \qquad N \to \infty,$$

where the last asymptotics follows from well known results for asymptotically normal random variables. Together with $|r_{i,N-1}| \sim N$ as $N \to \infty$ it follows that $c_{N,i} \sim \sqrt{(aN)/(2\pi b)}$ as $N \to \infty$. Thus, for large N, one can expect that, independent of the fixed initial state i,

$$t_{\mathrm{mix}}(\varepsilon) \;\approx\; \left\lfloor \frac{\frac{1}{2}\log\frac{aN}{2\pi b} - \log\varepsilon}{-\log\lambda} \right\rfloor + 1 \tag{3.6}$$

is a good approximation for the mixing time. Numerical comparisons with the exact value $t_{\mathrm{mix}}(\varepsilon)$, which can be computed via (3.2) using Corollary 1.2 or via (3.3) using Eqs. (1.8) and (1.9), show that both approximations (3.5) and (3.6) are rather sharp. Based on these approximations we conjecture that Y is rapidly mixing with $t_{\mathrm{mix}}(\varepsilon) = \Theta(\log N)$ as $N \to \infty$. We leave a rigorous proof of this conjecture as an open problem. Since Y is non-reversible, it is not straightforward to use the eigenvalues and eigenvectors to rigorously provide upper or lower bounds for the mixing time. For example, Wilson's lower bound formula (see, for example, [9, Theorem 13.28]) states that, for $1/2 < \lambda < 1$,

$$t_{\mathrm{mix}}(\varepsilon) \;\geq\; \frac{1}{2\log(1/\lambda)}\left(\log\left(\frac{(1-\lambda)(\Phi(i))^2}{2R}\right) + \log\left(\frac{1-\varepsilon}{\varepsilon}\right)\right),$$

where $\Phi := (\Phi(0), \ldots, \Phi(N))^T$ is a right eigenvector to the eigenvalue λ and $R := \max_{0 \leq i \leq N} \mathbb{E}_i((\Phi(Y_1) - \Phi(i))^2)$. In our situation, $\Phi(i) = r_{i,N-1} = N - \frac{a+b}{a} i$, $i \in \{0, \ldots, N\}$, and, hence, $\Phi(i) \sim N$ as $N \to \infty$. Straightforward calculations show that $R = O(N^2)$ as $N \to \infty$. Wilson's lower bound is hence bounded in N and therefore not helpful to prove the conjecture that $t_{\mathrm{mix}}(\varepsilon) = \Theta(\log N)$ as $N \to \infty$.

4 Proofs

In this section, proofs of Proposition 1.1, Corollary 1.2 and Theorem 1.4 are provided.

Proof. (of Proposition 1.1) We have $\eta_n = f \circ X_n$, where $f : K^N \to \mathcal{M}$ is defined via $f(k) := \sum_{r=1}^N \delta_{k_r}$ for all $k = (k_1, \ldots, k_N) \in K^N$. By a criterion of Burke and Rosenblatt [2], provided for convenience in Lemma 5.1 in the appendix, the process η is Markovian if, for every $\mathbf{i} = (i_1, \ldots, i_N) \in K^N$ and every $\mu \in \mathcal{M}$, the sum $\sum_{\mathbf{j} \in f^{-1}(\mu)} \pi_{\mathbf{i},\mathbf{j}}^{(n)}$ depends on $\mathbf{i}$ only via $f(\mathbf{i})$. In this case η has n-step transition probabilities $p_{\nu,\mu}^{(n)} = \sum_{\mathbf{j} \in f^{-1}(\mu)} \pi_{\mathbf{i},\mathbf{j}}^{(n)}$, where $\mathbf{i} \in f^{-1}(\nu)$ can be chosen arbitrarily. We therefore fix $n \in \mathbb{N}_0$, $\mathbf{i} = (i_1, \ldots, i_N) \in K^N$ and $\mu \in \mathcal{M}$, and focus on the sum $\sum_{\mathbf{j} \in f^{-1}(\mu)} \pi_{\mathbf{i},\mathbf{j}}^{(n)}$. Define $\nu := f(\mathbf{i})$. By (1.1),

$$\sum_{\mathbf{j} \in f^{-1}(\mu)} \pi_{\mathbf{i},\mathbf{j}}^{(n)} \;=\; \sum_{\substack{\mathbf{j} \in S \\ f(\mathbf{j}) = \mu}} \prod_{r=1}^N m_{i_r, j_r}^{(n)} \;=\; \sum_{\substack{\mathbf{j} \in S \\ f(\mathbf{j}) = \mu}} \prod_{k,l \in K} (m_{k,l}^{(n)})^{n_{k,l}(\mathbf{i},\mathbf{j})},$$

where $n_{k,l}(\mathbf{i},\mathbf{j}) := |\{r \in \{1, \ldots, N\} : i_r = k, j_r = l\}|$. The matrix $N(\mathbf{i},\mathbf{j}) := (n_{k,l}(\mathbf{i},\mathbf{j}))_{k,l \in K}$ has marginals

$$\sum_{l \in K} n_{k,l}(\mathbf{i},\mathbf{j}) \;=\; |\{r \in \{1, \ldots, N\} : i_r = k\}| \;=\; \sum_{r=1}^N \delta_{i_r}(\{k\}) \;=\; f(\mathbf{i})(\{k\}) \;=\; \nu_k, \quad k \in K,$$

and

$$\sum_{k \in K} n_{k,l}(\mathbf{i},\mathbf{j}) \;=\; |\{r \in \{1, \ldots, N\} : j_r = l\}| \;=\; \sum_{r=1}^N \delta_{j_r}(\{l\}) \;=\; f(\mathbf{j})(\{l\}) \;=\; \mu_l, \quad l \in K.$$

We therefore obtain

$$\sum_{\mathbf{j} \in f^{-1}(\mu)} \pi_{\mathbf{i},\mathbf{j}}^{(n)} \;=\; \sum_T \left(\prod_{k,l \in K} (m_{k,l}^{(n)})^{t_{k,l}} \right) \sum_{\substack{\mathbf{j} \in S \\ N(\mathbf{i},\mathbf{j}) = T}} 1,$$

where the sum $\sum_T$ extends over all $T = (t_{k,l})_{k,l \in K} \in \mathbb{N}_0^{K \times K}$ with marginals $\sum_{l \in K} t_{k,l} = \nu_k$, $k \in K$, and $\sum_{k \in K} t_{k,l} = \mu_l$, $l \in K$. Since there exist exactly $(\prod_{k \in K} \nu_k!)/(\prod_{k,l \in K} t_{k,l}!)$ vectors $\mathbf{j} = (j_1, \ldots, j_N) \in K^N$ satisfying $N(\mathbf{i},\mathbf{j}) = T$ we obtain

$$\sum_{\mathbf{j} \in f^{-1}(\mu)} \pi_{\mathbf{i},\mathbf{j}}^{(n)} \;=\; \sum_T \prod_{k \in K} \left(\nu_k! \left(\prod_{l \in K} \frac{(m_{k,l}^{(n)})^{t_{k,l}}}{t_{k,l}!} \right) \right).$$

The latter expression depends on $\mathbf{i}$ only via $f(\mathbf{i}) = \nu$. Now apply Lemma 5.1. □

Proof. (of Corollary 1.2) Let $K := \{0, 1\}$ and $\mathcal{K}$ the power set of K. The process $(Y_n)_{n \in \mathbb{N}_0}$ and the empirical measure process $\eta := (\eta_n)_{n \in \mathbb{N}_0}$ of X, defined via $\eta_n := \sum_{r=1}^N \delta_{X_n^{(r)}}$ for all $n \in \mathbb{N}_0$, are for $K = \{0, 1\}$ essentially the same object, since $\eta_n(\{0\}) = N - Y_n$ and $\eta_n(\{1\}) = Y_n$. It hence suffices to verify the statements for η instead of Y. By Proposition 1.1, η is a homogeneous, discrete-time Markov chain with state space $\mathcal{M}$, the set of measures μ on $(K, \mathcal{K})$ with values in $\{0, \ldots, N\}$ and total mass $\mu(K) = N$, and n-step transition probabilities

$$p_{\nu,\mu}^{(n)} \;=\; \sum_T \nu_0! \frac{(1 - a_n)^{t_{0,0}}}{t_{0,0}!} \frac{a_n^{t_{0,1}}}{t_{0,1}!} \nu_1! \frac{b_n^{t_{1,0}}}{t_{1,0}!} \frac{(1 - b_n)^{t_{1,1}}}{t_{1,1}!},$$

where the sum $\sum_T$ extends over all $T=(t_{k,l})_{k,l\in\{0,1\}}\in\mathbb{N}_0^{K\times K}$ with marginals $\sum_{l\in K}t_{k,l}=\nu_k$, $k\in K$, and $\sum_{k\in K}t_{k,l}=\mu_l$, $l\in K$. Using the notation $k:=t_{1,1}$, $i:=\nu_1$ and $j:=\mu_1$, this turns into

$$p_{\nu,\mu}^{(n)} \;=\; \sum_{k=0}^{\min(i,j)}(N-i)!\frac{(1-a_n)^{N-i-(j-k)}}{(N-i-(j-k))!}\frac{a_n^{j-k}}{(j-k)!}i!\frac{b_n^{i-k}}{(i-k)!}\frac{(1-b_n)^k}{k!},$$

which is equal to the right hand side in (1.3). □

Proof. (of Theorem 1.4) Two proofs of Theorem 1.4 are provided, both based on generating functions. The first proof shows that the generating functions of $P_{i,\cdot}$ and $(RDL)_{i,\cdot}$ coincide. This proof is relatively short but somewhat intransparent. In particular, the spectral decomposition, i.e. the matrices R, D and L, have to be known in advance. The second proof is more technical and hence longer, but has the advantage that a recipe is provided how to obtain, recursively over $l\in\{0,\ldots,N\}$, a right eigenvector to the eigenvalue λ_{N-l}. Hence, the matrices R, D and L do not need to be known in advance. The formulas for the entries of these three matrices pop up naturally while performing the calculations.

Proof 1. It suffices to verify that the generating functions of $P_{i,\cdot}$ and $(RDL)_{i,\cdot}$ coincide. From Proposition 1.2 it follows that the ith row of the transition matrix $P=(p_{i,j})_{i,j\in S}$ has probability generating function (pgf)

$$\sum_{j\in S}p_{i,j}s^j \;=\; (b+\bar{b}s)^i(\bar{a}+as)^{N-i},\quad i\in S, s\in\mathbb{R}, \tag{4.1}$$

where $\bar{a}:=1-a$ and $\bar{b}:=1-b$. On the other hand,

$$\sum_{j\in S}(RDL)_{i,j}s^j \;=\; \sum_{j\in S}s^j\sum_{m\in S}r_{i,m}d_{m,m}l_{m,j} \;=\; \sum_{m\in S}r_{i,m}d_{m,m}\sum_{j\in S}l_{m,j}s^j.$$

Plugging in the formula for $l_{m,j}$ (see (1.9)) and interchanging the sums over j and over k yields

$$\begin{aligned}
\sum_{j\in S}(RDL)_{i,j}s^j &= \sum_{m=0}^N r_{i,m}d_{m,m}\sum_{k=m}^N(-1)^{k-m}\binom{N-m}{k-m}\Big(\frac{a}{a+b}\Big)^k\sum_{j=0}^k\binom{k}{j}\Big(\frac{b}{a}\Big)^{k-j}s^j\\
&= \sum_{m=0}^N r_{i,m}d_{m,m}\sum_{k=m}^N(-1)^{k-m}\binom{N-m}{k-m}\Big(\frac{a}{a+b}\Big)^k\Big(s+\frac{b}{a}\Big)^k\\
&= \sum_{m=0}^N r_{i,m}d_{m,m}\sum_{k=m}^N(-1)^{k-m}\binom{N-m}{k-m}\Big(\frac{as+b}{a+b}\Big)^k\\
&= \sum_{m=0}^N r_{i,m}d_{m,m}\Big(\frac{as+b}{a+b}\Big)^m\Big(\frac{a-as}{a+b}\Big)^{N-m}.
\end{aligned}$$

Plugging in $d_{m,m}=(1-a-b)^{N-m}$ and the formula for $r_{i,m}$ (see (1.8)) and interchanging the sums over m and k yields

$$\begin{aligned}
&\sum_{j\in S}(RDL)_{i,j}s^j\\
&\quad= \sum_{k=0}^i\binom{i}{k}\Big(-\frac{b}{a}\Big)^{i-k}\Big(\frac{a+b}{a}\Big)^k\sum_{m=k}^N\binom{N-k}{m-k}\Big(\frac{as+b}{a+b}\Big)^m\Big(\frac{a-as}{a+b}(1-a-b)\Big)^{N-m}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{i} \binom{i}{k} \Big(-\frac{b}{a}\Big)^{i-k} \Big(\frac{a+b}{a}\Big)^{k} \Big(\frac{as+b}{a+b}\Big)^{k} \Big(\frac{as+b}{a+b} + \frac{(a-as)(1-a-b)}{a+b}\Big)^{N-k}\\
&= \sum_{k=0}^{i} \binom{i}{k} \Big(-\frac{b}{a}\Big)^{i-k} \Big(\frac{as+b}{a}\Big)^{k} (1-a+as)^{N-k}\\
&= (1-a+as)^{N-i} \sum_{k=0}^{i} \binom{i}{k} \Big(-\frac{b}{a}\Big)^{i-k} \Big(\frac{as+b}{a}\Big)^{k} (1-a+as)^{i-k}\\
&= (1-a+as)^{N-i} \Big(\frac{as+b}{a} - \frac{b}{a}(1-a+as)\Big)^{i} \;=\; (\bar{a}+as)^{N-i}(b+s\bar{b})^{i}.
\end{aligned}
$$

This expression coincides with (4.1) and the result is shown. □

Proof 2. This proof of Theorem 1.4 is somewhat technical in detail, but there is a straightforward approach behind the technical calculations. We therefore first describe the basic method and work out the details afterwards. We proceed as follows. As in the first proof the starting point is formula (4.1) for the pgf of the ith row of P. Now choose in (4.1) $s = s_0$ with $s_0 := -b/a$. From $b + \bar{b}s_0 = s_0(1-a-b)$ and $\bar{a} + as_0 = 1-a-b$ we conclude that

$$
\sum_{j\in S} p_{i,j} s_0^j \;=\; \big(s_0(1-a-b)\big)^i (1-a-b)^{N-i} \;=\; (1-a-b)^N s_0^i, \qquad i \in S.
$$

Thus, $\lambda_N = (1-a-b)^N$ is an eigenvalue of P with corresponding right eigenvector $x_0 = (x_{0,0}, \ldots, x_{0,N})^T$ having entries $x_{0,j} := s_0^j$, $j \in S$. Taking the derivative with respect to s in (4.1) it follows that (product rule)

$$
\begin{aligned}
\sum_{j\in S} p_{i,j} j s^{j-1} &= i(b+\bar{b}s)^{i-1}\bar{b}(\bar{a}+as)^{N-i} + (b+\bar{b}s)^i (N-i)(\bar{a}+as)^{N-i-1} a\\
&= (b+\bar{b}s)^{i-1}(\bar{a}+as)^{N-i-1}\big(i\bar{b}(\bar{a}+as) + (b+\bar{b}s)(N-i)a\big).
\end{aligned}
$$

Choosing again $s = s_0 = -b/a$ it follows that

$$
\begin{aligned}
\sum_{j\in S} p_{i,j} j s_0^{j-1} &= \big(s_0(1-a-b)\big)^{i-1} (1-a-b)^{N-i-1} \cdot\\
&\qquad \cdot \big(i\bar{b}(1-a-b) + s_0(1-a-b)(N-i)a\big)\\
&= (1-a-b)^{N-1} s_0^{i-1} \big(i\bar{b} + s_0(N-i)a\big)\\
&= \lambda_{N-1} s_0^{i-1}\big(i\bar{b} + ib + s_0 Na\big)\\
&= \lambda_{N-1} s_0^{i-1}(i + s_0 Na) \;=\; \lambda_{N-1}\big(i s_0^{i-1} + Na s_0^i\big).
\end{aligned}
$$

Adding to this equation the C-fold ($C \in \mathbb{R}$) of the equation $\sum_{j\in S} p_{i,j} s_0^j = \lambda_N s_0^i = \lambda_{N-1}(1-a-b)s_0^i$ yields

$$
\sum_{j\in S} p_{i,j}\big(j s_0^{j-1} + C s_0^j\big) \;=\; \lambda_{N-1}\big(i s_0^{i-1} + (Na + C(1-a-b))s_0^i\big).
$$

Choosing the constant C such that $C = Na + C(1-a-b)$, so $C = \frac{a}{a+b}N$, yields a right eigenvector x_1 to the eigenvalue λ_{N-1}, namely $x_1 = (x_{1,0}, \ldots, x_{1,N})^T$ with entries

$$
x_{1,j} \;:=\; j s_0^{j-1} + C s_0^j \;=\; j s_0^{j-1} + \frac{a}{a+b} N s_0^j, \qquad j \in S.
$$

The general method is now obvious. We differentiate (4.1), successively for $l = 1, 2, \ldots, N$, exactly l-times with respect to s and choose afterwards $s = s_0$. With some skill one obtains a right eigenvector x_l to the eigenvalue λ_{N-l}. We now carry out the details of this procedure.

Step 1. In this step a system of key equations is derived which will turn out be useful to find the eigenvectors of the transition matrix P. This step also explains how the matrix A pops up during the calculations. The lth derivative of (4.1) is (Leipniz rule)

$$\begin{aligned}\sum_{j\in S} p_{i,j}(j)_l s^{j-l} &= \sum_{m=0}^{l}\binom{l}{m}\frac{\partial^m}{\partial s^m}(b+\bar{b}s)^i\frac{\partial^{l-m}}{\partial s^{l-m}}(\bar{a}+as)^{N-i}\\ &= \sum_{m=0}^{l}\binom{l}{m}(i)_m(b+\bar{b}s)^{i-m}\bar{b}^m(N-i)_{l-m}(\bar{a}+as)^{N-i-(l-m)}a^{l-m}.\end{aligned}$$

Choosing $s=s_0=-b/a$ and noting that $b+\bar{b}s_0=s_0(1-a-b)$ and $\bar{a}+as_0=1-a-b$ it follows that

$$\sum_{j\in S} p_{i,j}(j)_l s_0^{j-l} \;=\; \lambda_{N-l}\sum_{m=0}^{l}\binom{l}{m}\bar{b}^m a^{l-m}s_0^{i-m}(i)_m(N-i)_{l-m} \;=:\; \lambda_{N-l}\Sigma.$$

In the following we would like to represent Σ as a linear combination of terms of the descending factorials $(i)_k$, $k\in\{0,\ldots,l\}$. In order to do this we rewrite the last factor $(N-i)_{l-m}$ by making use of the formula $(x+y)_n=\sum_{k=0}^{n}\binom{n}{k}(x)_k(y)_{n-k}$ as

$$\begin{aligned}(N-i)_{l-m} &= (-1)^{l-m}(i-m+l-N-1)_{l-m}\\ &= (-1)^{l-m}\sum_{k=m}^{l}\binom{l-m}{k-m}(i-m)_{k-m}(l-N-1)_{l-k}\\ &= \sum_{k=m}^{l}\binom{l-m}{k-m}(-1)^{k-m}(i-m)_{k-m}(N-k)_{l-k},\end{aligned}$$

since $(l-N-1)_{l-k}=(-1)^{l-k}(N-k)_{l-k}$. Plugging in this expression yields

$$\Sigma \;=\; \sum_{m=0}^{l}\binom{l}{m}\bar{b}^m a^{l-m}s_0^{i-m}(i)_m\sum_{k=m}^{l}\binom{l-m}{k-m}(-1)^{k-m}(i-m)_{k-m}(N-k)_{l-k}.$$

Using that $(i)_m(i-m)_{k-m}=(i)_k$ and $\binom{l}{m}\binom{l-m}{k-m}=\binom{l}{k}\binom{k}{m}$ and interchanging the two sums yields

$$\begin{aligned}\Sigma &= \sum_{k=0}^{l}\binom{l}{k}(N-k)_{l-k}(i)_k a^{l-k}s_0^{i-k}\underbrace{\sum_{m=0}^{k}\binom{k}{m}\bar{b}^m(-as_0)^{k-m}}_{=(\bar{b}-as_0)^k=1}\\ &= \sum_{k=0}^{l}\binom{l}{k}(N-k)_{l-k}(i)_k a^{l-k}s_0^{i-k} \;=\; l!\sum_{k=0}^{l}\binom{N-k}{l-k}a^{l-k}\binom{i}{k}s_0^{i-k}.\end{aligned}$$

Therefore,

$$\sum_{j\in S} p_{i,j}\binom{j}{l}s_0^{j-l} \;=\; \lambda_{N-l}\frac{\Sigma}{l!} \;=\; \lambda_{N-l}\sum_{k=0}^{l}\binom{N-k}{l-k}a^{l-k}\binom{i}{k}s_0^{i-k}.$$

Let $A:=(a_{i,j})_{i,j\in S}$ denote the matrix with entries $a_{i,j}:=\binom{i}{j}s_0^{i-j}$, $i,j\in S$. Then, the above equation takes the form

$$(PA)_{i,l} \;=\; \sum_{j\in S}p_{i,j}a_{j,l} \;=\; \lambda_{N-l}\sum_{k=0}^{l}\binom{N-k}{l-k}a^{l-k}a_{i,k}, \qquad i,l\in S. \tag{4.2}$$

We will see in the following step, that these are the key equations in order to find the eigenvectors of the transition matrix P.

Step 2. We will now use the system of equations (4.2) in order to derive, successively for $j = 0, \ldots, N$, a right eigenvector x_j to the eigenvalue $\lambda_{N-j} = (1-a-b)^{N-j}$. We make the ansatz that $x_j = (x_{j,0}, \ldots, x_{j,N})^T$ is of the form

$$x_{j,i} \;=\; \sum_{k=0}^{j} a_{i,k} b_{k,j}, \quad i \in S, \tag{4.3}$$

where the coefficients $b_{k,j}$ may depend on the parameters N, a and b of the model but not on the state i. Since eigenvectors can be scaled arbitrarily, we can in principle choose $b_{j,j}$ arbitrary (not equal to zero), but we shall see soon that $b_{j,j} := ((a+b)/a)^j$ is a convenient choice. Since x_j should be(come) a right eigenvector to the eigenvalue λ_{N-j}, the chain of equalities

$$\sum_{k=0}^{N} p_{i,k} x_{j,k} \;=\; \lambda_{N-j} x_{j,i} \;=\; \lambda_{N-j} \sum_{k=0}^{j} a_{i,k} b_{k,j} \tag{4.4}$$

should hold for all $i \in S$, where the last equality holds by the ansatz (4.3). On the other hand, by the ansatz (4.3) and the characteristic equations (4.2),

$$\begin{aligned}
\sum_{k=0}^{N} p_{i,k} x_{j,k} \;&=\; \sum_{k=0}^{N} p_{i,k} \sum_{l=0}^{j} a_{k,l} b_{l,j} \;=\; \sum_{l=0}^{j} b_{l,j} \sum_{k=0}^{N} p_{i,k} a_{k,l} \;=\; \sum_{l=0}^{j} b_{l,j} (PA)_{i,l} \\
&=\; \sum_{l=0}^{j} b_{l,j} \lambda_{N-l} \sum_{k=0}^{l} \binom{N-k}{l-k} a^{l-k} a_{i,k} \\
&=\; \lambda_{N-j} \sum_{k=0}^{j} a_{i,k} \sum_{l=k}^{j} b_{l,j} \lambda_{j-l} \binom{N-k}{l-k} a^{l-k},
\end{aligned} \tag{4.5}$$

since $\lambda_{N-l} = \lambda_{N-j}\lambda_{j-l}$. Since (4.4) and (4.5) coincide for all $i \in S$, the coefficients $b_{k,j}$ need to satisfy the system of equations

$$b_{k,j} \;=\; \sum_{l=k}^{j} b_{l,j} \lambda_{j-l} \binom{N-k}{l-k} a^{l-k}, \qquad j \in S, k \in \{0, \ldots, j\}.$$

We now solve this system of equations. Defining

$$c_{k,j} \;:=\; \frac{(N-j)!}{(N-k)!} b_{k,j}, \qquad j \in S, k \in \{0, \ldots, j\},$$

the system reduces to

$$c_{k,j} \;=\; \sum_{l=k}^{j} c_{l,j} \lambda_{j-l} \frac{a^{l-k}}{(l-k)!}, \qquad j \in S, k \in \{0, \ldots, j\}.$$

In particular, the coefficients $c_{k,j}$ do not depend on N and they satisfy for each fixed $j \in S$ the recursion

$$c_{k,j} \;=\; \frac{1}{1-\lambda_{j-k}} \sum_{l=k+1}^{j} c_{l,j} \lambda_{j-l} \frac{a^{l-k}}{(l-k)!}, \qquad k = j-1, j-2, \ldots, 1, 0,$$

with initial value $c_{j,j} = b_{j,j} = ((a+b)/a)^j$. We prove by backward induction on $k = j, j-1, \ldots, 1, 0$ that this recursion has the solution

$$c_{k,j} = \frac{1}{(j-k)!}\left(\frac{a+b}{a}\right)^k, \qquad j \in S, k \in \{0,\ldots,j\}.$$

For $k = j$ this holds since $c_{j,j} = ((a+b)/a)^j$. The induction step from $j, j-1, \ldots, k+1$ to k reads

$$\begin{aligned}
c_{k,j} &= \frac{1}{1-\lambda_{j-k}} \sum_{l=k+1}^{j} c_{l,j}\lambda_{j-l}\frac{a^{l-k}}{(l-k)!}\\
&= \frac{1}{1-\lambda_{j-k}} \sum_{l=k+1}^{j} \frac{1}{(j-l)!}\left(\frac{a+b}{a}\right)^l (1-a-b)^{j-l}\frac{a^{l-k}}{(l-k)!}\\
&= \frac{1}{(j-k)!}\left(\frac{a+b}{a}\right)^k \frac{1}{1-\lambda_{j-k}} \sum_{l=k+1}^{j} \binom{j-k}{j-l}(1-a-b)^{j-l}(a+b)^{l-k}\\
&= \frac{1}{(j-k)!}\left(\frac{a+b}{a}\right)^k \frac{1}{1-\lambda_{j-k}} \sum_{r=0}^{j-k-1} \binom{j-k}{r}(1-a-b)^{r}(a+b)^{j-k-r}\\
&= \frac{1}{(j-k)!}\left(\frac{a+b}{a}\right)^k \frac{1}{1-\lambda_{j-k}}(1-(1-a-b)^{j-k}) \;=\; \frac{1}{(j-k)!}\left(\frac{a+b}{a}\right)^k.
\end{aligned}$$

In summary, it is shown that $x_j := (x_{j,0}, \ldots, x_{j,N})^T$, defined via (4.3), with coefficients $a_{i,k} := \binom{i}{k} s_0^{i-k}$ and

$$b_{k,j} := \frac{(N-k)!}{(N-j)!}c_{k,j} = \frac{(N-k)!}{(N-j)!}\frac{1}{(j-k)!}\left(\frac{a+b}{a}\right)^k = \binom{N-k}{j-k}\left(\frac{a+b}{a}\right)^k$$

is a right eigenvector to the eigenvalue $\lambda_{N-j} = (1-a-b)^{N-j}$, $j \in S$.

Step 3. It is now straightforward to derive the desired spectral decomposition of the transition matrix P. Let $R = (r_{i,j})_{i,j\in S}$ denote the matrix, which contains in the jth column the right eigenvector x_j to the eigenvalue λ_{N-j}, i.e. $r_{i,j} := x_{j,i} = \sum_{k=0}^{j} a_{i,k}b_{k,j}$ or, in matrix notation, $R := AB$, where $A := (a_{i,j})_{i,j\in S}$ is the left lower triangular matrix and $B := (b_{i,j})_{i,j\in S}$ the right upper triangular matrix defined via (1.5). Note that A and B are non-singular with inverses (1.6). Since R is the matrix containing in the jth column the right eigenvector x_j to the eigenvalue λ_{N-j}, we have $PR = RD$, where D is the diagonal matrix with entries $d_{i,i} := \lambda_{N-i} = (1-a-b)^{N-i}$, $i \in S$. Multiplying from the right with $L := R^{-1}$ yields the spectral decomposition $P = RDL$. The proof is complete. □

5 Appendix

For convenience we record a criterion which ensures that a function of a Markov chain is still Markovian. The result is well known in the probability community (see, for example, Levin and Peres [9, Lemma 2.5]) and essentially goes back to Burke and Rosenblatt [2].

Lemma 5.1. *Let $X = (X_n)_{n\in\mathbb{N}_0}$ be a homogeneous discrete-time Markov chain with state space S and transition probabilities $\pi_{i,j}$, $i,j \in S$. Moreover, let $f : S \to S'$ be a surjective function in a space S'. Define $Y_n := f \circ X_n$ for all $n \in \mathbb{N}_0$. If, for every $i \in S$ and every $v \in S'$, the sum $\sum_{j\in f^{-1}(v)} \pi_{i,j}$ depends on i only via $f(i)$, then $Y = (Y_n)_{n\in\mathbb{N}_0}$ is a homogeneous discrete-time Markov chain with state space S'. In this case Y has n-step transition probabilities $p_{u,v}^{(n)} = \sum_{j\in f^{-1}(v)} \pi_{i,j}^{(n)}$, $n \in \mathbb{N}_0$, $u, v \in S'$, where $i \in f^{-1}(u)$ can be* *chosen arbitrarily. Here* $\pi^{(n)}_{i,j}$, $n \in \mathbb{N}_0$, $i,j \in S$, *denote the* n*-step transition probabilities of* X.

The following lemma provides the horizontal generating functions of the transformation matrices R and L from Theorem 1.4.

Lemma 5.2. *Let* R *and* L *be the matrices defined via* $R := AB$ *and* $L := R^{-1}$*, where* A *and* B *are defined via (1.5), and let* $r_i : \mathbb{C} \to \mathbb{C}$ *and* $l_i : \mathbb{C} \to \mathbb{C}$*,* $i \in S$*, denote the associated horizontal generating functions defined via* $r_i(z) := \sum_{j\in S} r_{i,j} z^j$ *and* $l_i(z) := \sum_{j\in S} l_{i,j} z^j$ *for all* $z \in \mathbb{C}$ *and* $i \in S$*. Then*

$$r_i(z) \;=\; (z+1)^{N-i}\Big(z-\frac{b}{a}\Big)^i \quad \text{and} \quad l_i(z) \;=\; \Big(\frac{a}{a+b}\Big)^N\Big(z+\frac{b}{a}\Big)^i(1-z)^{N-i}, \quad i \in S, z \in \mathbb{C}.$$

Proof. From (1.8) it follows that

$$\begin{aligned}
r_i(z) \;&:=\; \sum_{j=0}^N r_{i,j} z^j \;=\; \sum_{k=0}^i \binom{i}{k}\Big(-\frac{b}{a}\Big)^{i-k}\Big(\frac{a+b}{a}\Big)^k \sum_{j=k}^N \binom{N-k}{j-k} z^j\\
&=\; \sum_{k=0}^i \binom{i}{k}\Big(-\frac{b}{a}\Big)^{i-k}\Big(\frac{a+b}{a}\Big)^k z^k(z+1)^{N-k}\\
&=\; (z+1)^{N-i}\sum_{k=0}^i \binom{i}{k}\Big(\frac{a+b}{a}z\Big)^k\Big(-\frac{b}{a}(z+1)\Big)^{i-k}\\
&=\; (z+1)^{N-i}\Big(z-\frac{b}{a}\Big)^i, \qquad i \in S, z \in \mathbb{C}.
\end{aligned}$$

Similarly, we deduce from (1.9) that

$$\begin{aligned}
l_i(z) \;&:=\; \sum_{j=0}^N l_{i,j} z^j \;=\; \sum_{k=i}^N \binom{N-i}{k-i}\Big(\frac{a}{a+b}\Big)^k(-1)^{k-i}\sum_{j=0}^k \binom{k}{j}\Big(\frac{b}{a}\Big)^{k-j} z^j\\
&=\; \sum_{k=i}^N \binom{N-i}{k-i}\Big(\frac{a}{a+b}\Big)^k(-1)^{k-i}\Big(z+\frac{b}{a}\Big)^k\\
&=\; \Big(\frac{a}{a+b}\Big)^i\Big(z+\frac{b}{a}\Big)^i\sum_{k=i}^N \binom{N-i}{k-i}\Big(-\frac{a}{a+b}\Big(z+\frac{b}{a}\Big)\Big)^{k-i}\\
&=\; \Big(\frac{a}{a+b}\Big)^i\Big(z+\frac{b}{a}\Big)^i\Big(1-\frac{a}{a+b}\Big(z+\frac{b}{a}\Big)\Big)^{N-i}\\
&=\; \Big(\frac{a}{a+b}\Big)^i\Big(z+\frac{b}{a}\Big)^i\Big(\frac{a}{a+b}-\frac{a}{a+b}z\Big)^{N-i}\\
&=\; \Big(\frac{a}{a+b}\Big)^N\Big(z+\frac{b}{a}\Big)^i(1-z)^{N-i}, \qquad i \in S, z \in \mathbb{C},
\end{aligned}$$

which is the desired formula for the horizontal generating function l_i. ☐

References

[1] Berger, M. and Cerf, R.: A basic model of mutations, (2018), arXiv:1806.01212

[2] Burke, C. J. and Rosenblatt, M.: A Markovian function of a Markov chain. *Ann. Math. Stat.* **29**, (1958), 1112–1122. MR-0101557

[3] Cannings, C.: The latent roots of certain Markov chains arising in genetics: a new approach, I. Haploid models. *Adv. Appl. Probab.* **6**, (1974), 260–290. MR-0343949

[4] Cannings, C.: The latent roots of certain Markov chains arising in genetics: a new approach, II. Further haploid models. *Adv. Appl. Probab.* **7**, (1975), 264–282. MR-0371430

[5] Ehrenfest, P. and Ehrenfest, T.: Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem. *Phys. Z.* **8**, (1907), 311–314.

[6] Gladstien, K.: The characteristic values and vectors for a class of stochastic matrices arising in genetics. *SIAM J. Appl. Math.* **34**, (1978), 630–642. MR-0475977

[7] Grinstead, C. M. and Snell J. L.: *Introduction to Probability*. Second Revised Version. American Mathematical Society, 1997. MR-0362410

[8] Kac, M.: Random walk and the theory of Brownian motion. *Amer. Math. Monthly* **54**, (1947), 369–391. MR-0021262

[9] Levin, D. and Peres, Y.: *Markov Chains and Mixing Times*. American Mathematical Society, Providence, 2017. MR-3726904

[10] Nestoridi, E.: A non-local random walk on the hypercube. *Adv. Appl. Probab.* **49**, (2017), 1288–1299. MR-3732195

[11] Norris, J. R.: *Markov Chains*. Cambridge Studies in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 1998. MR-1600720

[12] Scoppola, B.: Exact solution for a class of random walks on the hypercube. *J. Stat. Phys.* **143**, (2011), 413–419. MR-2799945

Acknowledgments. The author thanks two anonymous referees for their constructive suggestions leading to a significant improvement of the manuscript.