ALEA, Lat. Am. J. Probab. Math. Stat. 18, 1759–1771 (2021)
DOI: 10.30757/ALEA.v18-65

On A Family of Isospectral Pure-Birth Processes

Laurent Miclo and Chi Zhang

Institut de Mathématiques de Toulouse and Toulouse School of Economics,
CNRS and University of Toulouse, 1 Esplanade de l'université, 31080 Toulouse cedex 06, France
E-mail address: miclo@math.cnrs.fr
URL: https://perso.math.univ-toulouse.fr/miclo/

College of Mathematical Sciences, Ocean University of China,
Qingdao 266100, People's Republic of China
E-mail address: zhangchi@ouc.edu.cn
URL: http://eweb.ouc.edu.cn/ms/2019/0504/c4086a239038/page.htm

Abstract. It is shown that every spectrum of a finite irreducible Markov generator whose eigenvalues are real and of geometric multiplicity 1 can be obtained as the spectrum of an irreducible pure-birth Markov process with jumps from the right-most boundary to all the other points. A whole isospectral family of such processes is exhibited and their mixing rates are compared.

1. Introduction

The investigation of the spectra of finite Markov generators was first motivated by the quest of quantitative bounds on the convergence to equilibrium of the corresponding processes, see for instance the reference book Levin et al. (2009). There is also a classification reason: what are the possible spectra of Markov generators?, and for such a given spectrum, is there a simple representative Markov process? This structural question is related to the previous motivation, as ergodic finite isospectral Markov generators can be intertwined and under certain circumstances this relation enables good transfers of information about the speed of convergence, see e.g. Miclo (2018); Miclo and Patie (2021). Our goal here goes in this general direction, by providing a simple family of Markov generators for real spectra with geometric multiplicity 1.

Let us start by recalling some general definitions. For fixed $n \in \mathbb{N}$, let $\mathcal{M}(n)$ be the space of all $n \times n$ matrices with real entries. Then for any fixed $A \in \mathcal{M}(n)$, there exists a complex, non-singular and $n \times n$ matrix S such that $A = S^{-1}JS$, where J is the Jordan canonical form, i.e.

Received by the editors January 7th, 2021; accepted July 16th, 2021.
2010 *Mathematics Subject Classification.* 60J27, 60J35, 15A18, 37A25.
Key words and phrases. Pure-birth processes, spectra of Markov processes, mixing rates, intertwining relations.
Laurent Miclo has been supported by the grant ANR-17-EURE-0010. Chi Zhang (Corresponding author) has been supported by National Natural Science Foundation of China (Grant No. 11501531, 11801406, 11801556) and China State Foundation for Studying Abroad.

$$J \quad \coloneqq \quad \begin{pmatrix} J_{n_1}(\lambda_1) & & & 0 \\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & J_{n_l}(\lambda_l) \end{pmatrix}.$$

Here $n_k \in \mathbb{N}$ ($k \in [\![l]\!] \coloneqq \{1, 2, ..., l\}$), $\sum_{k\in[\![l]\!]} n_k = n$, the Jordan block $J_{n_k}(\lambda_k)$ is given by the $n_k \times n_k$ matrix

$$J_{n_k}(\lambda_k) \quad \coloneqq \quad \begin{pmatrix} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \lambda_k & 1 \\ 0 & & & \lambda_k \end{pmatrix}$$

and the eigenvalues $\lambda_k \in \mathbb{C}$ ($k \in [\![l]\!]$), while λ_k ($k \in [\![l]\!]$) are not necessarily distinct.

The **Jordan spectrum** of matrix A is defined as the multi-set

$$\sigma_{\mathrm{J}}(A) \quad \coloneqq \quad \{(\lambda_k, n_k) \,:\, k \in [\![l]\!]\},$$

where each (λ_k, n_k) corresponds to the Jordan block $J_{n_k}(\lambda_k)$. The traditional **geometric and algebraic spectra** of A are the multi-sets respectively defined by

$$\begin{aligned} \sigma_{\mathrm{g}}(A) &\coloneqq \{\lambda_k \,:\, k \in [\![l]\!]\}, \\ \sigma_{\mathrm{a}}(A) &\coloneqq \{\lambda_k\,[n_k] \,:\, k \in [\![l]\!]\}, \end{aligned}$$

namely each Jordan block corresponding to (λ_k, n_k), for $k \in [\![l]\!]$, brings a geometric (respectively algebraic) multiplicity of 1 (resp. n_k) to the eigenvalue λ_k of A.

Denote by $\mathcal{M}_{\mathrm{r}}(n)$ the set of matrices $A \in \mathcal{M}(n)$ whose spectrum is real, i.e. for all $(\lambda, m) \in \sigma_{\mathrm{J}}(A)$, we have $\lambda \in \mathbb{R}$.

Sets of the form $S \coloneqq \{(\lambda_k, n_k) \in \mathbb{C} \times \mathbb{N} \,:\, k \in [\![l]\!]\}$ are said to be **uni-spectral** when they satisfy the following conditions:

- $\sum_{k\in[\![l]\!]} n_k = n$ and in particular $l \in [\![n]\!]$,
- $(\lambda_1, n_1) = (0, 1)$,
- for all $k \in [\![2, l]\!]$, the real part of λ_k is positive,
- for all $k \neq k' \in [\![l]\!]$, we have $\lambda_k \neq \lambda_{k'}$.

The uni-spectral set S will be interpreted as a multi-set whose elements all have multiplicity 1 (here multiplicity is understood in the sense of multi-sets, when S will correspond to a spectrum, the multiplicity will thus stand for the geometrical one). The uni-spectral set S is said to be real when all the λ_k are real (and thus $\lambda_k > 0$ for all $k \in [\![2, n]\!]$). Note that without lost of generality, we can and will assume that the elements of a real uni-spectral set S are indexed so that $k < k' \in [\![l]\!]$ implies $\lambda_k < \lambda_{k'}$.

The set of (respectively real) uni-spectral multi-sets will be denoted $\mathcal{U}(n)$ (resp. $\mathcal{U}_{\mathrm{r}}(n)$).

One of the main purposes of this paper is to see that each uni-spectral set can be seen as the spectrum of a very simple Markov process, and even of an interesting family of them.

More precisely, we are interested in the following subclasses of $\mathcal{M}(n)$. First let $\mathcal{S}(n)$ be the set of $L \in \mathcal{M}(n)$ which are **irreducible skip-free Markov generators**, namely which satisfy the following properties:

- **Markov generator**: off-diagonal entries of L are non-negative and the row sums are null.
- **Irreducibility**: for any $x \neq y \in [\![n]\!]$, there exists a path in $[\![n]\!]$, $p \coloneqq (p_0, p_1, ..., p_l)$ with $l \in \mathbb{N}$, such that $p_0 = x$, $p_l = y$ and for any $k \in [\![l]\!]$, $L(p_{k-1}, p_k) > 0$.
- **Skip-free**: for all $x, y \in [\![n]\!]$ with $y > x + 1$, we have $L(x, y) = 0$.

The second subclass we will study is $\mathcal{T}(n)$, the set of $L \in \mathcal{S}(n)$ which are **irreducible pure-birth Markov generators**, namely which satisfy the additional property:

- **Pure-birth**: for any $x \in [\![n-1]\!]$ and $y \notin \{x, x+1\}$, we have $L(x,y) = 0$.

Remark 1.1. When the pure-birth condition is replaced by the following birth-death property:

- **Birth-death**: for any $x \in [\![n-1]\!]$ and $y \notin \{x-1, x, x+1\}$, we have $L(x,y) = 0$.

we get the class of **irreducible birth-death Markov generators**, which is intermediary between $\mathcal{S}(n)$ and $\mathcal{T}(n)$, and which plays an important role in the constructive theory of denumerable Markov processes, see the book Yang (1990). Irreducible birth-death Markov generators are used as approximations (as n goes to infinity) of general birth and death processes on $\mathbb{Z}_+$. Note nevertheless that the Markov generators from these three classes can jump from their right-most boundary n to any other point in the state space.

Here we will be interested in the following subclasses,

$$\begin{aligned} \mathcal{S}_{\mathrm{r}}(n) &\coloneqq \mathcal{S}(n) \cap \mathcal{M}_{\mathrm{r}}(n), \\ \mathcal{T}_{\mathrm{r}}(n) &\coloneqq \mathcal{T}(n) \cap \mathcal{M}_{\mathrm{r}}(n), \end{aligned}$$

and more precisely into their spectral properties, so denote

$$\begin{aligned} \sigma_{\mathrm{J}}(-\mathcal{S}_{\mathrm{r}}(n)) &\coloneqq \{\sigma_{\mathrm{J}}(-L) \,:\, L \in \mathcal{S}_{\mathrm{r}}(n)\}, \\ \sigma_{\mathrm{J}}(-\mathcal{T}_{\mathrm{r}}(n)) &\coloneqq \{\sigma_{\mathrm{J}}(-L) \,:\, L \in \mathcal{T}_{\mathrm{r}}(n)\}. \end{aligned}$$

The interest of these classes is:

Theorem 1.2. *For any fixed $n \in \mathbb{N}$, we have*

$$\sigma_{\mathrm{J}}(-\mathcal{S}_{\mathrm{r}}(n)) = \sigma_{\mathrm{J}}(-\mathcal{T}_{\mathrm{r}}(n)) = \mathcal{U}_{\mathrm{r}}(n)$$

We believe the same result is true without the requirement that the eigenvalues are real, but it would require a better understanding of the significance of complex eigenvalues:

Conjecture 1.3. *For any fixed $n \in \mathbb{N}$, we have*

$$\sigma_{\mathrm{J}}(-\mathcal{S}(n)) = \sigma_{\mathrm{J}}(-\mathcal{T}(n)) = \mathcal{U}(n).$$

Theorem 1.2 is a first step in the direction of this challenging conjecture.

In the following section, we will be exhibiting an interesting isospectral family from $\mathcal{T}(n)$ whose spectrum is a given uni-spectral set, this will be the important step in the proof of Theorem 1.2. In the last section we will compare the mixing rates of the elements of this family, see Theorem 3.1.

2. An isospectral family

The purpose of this section is to prove Theorem 1.2.

We need some preparations.

Proposition 2.1. *Let $L \in \mathcal{S}_{\mathrm{r}}(n)$. Then the Jordan spectrum of $-L$ is a uni-spectral set.*

Proof: We first prove the smallest eigenvalue λ_0 of $-L$ is 0 and has algebraic multiplicity 1. For this, let

$$q = \max_{1 \leqslant i \leqslant n} \{-L(i,i)\}, \quad B = L + q \cdot I,$$

where I is the $n \times n$ identity matrix. Then B is nonnegative and irreducible and Theorem 1.5 of Seneta (2006) can be applied to B: it admits an eigenvalue $\lambda^B_{\max}$ which is maximal in absolute value and whose algebraic multiplicity is 1. According to Corollary 1 page 8 of Seneta (2006), we have $\lambda^B_{\max} = q$ since the row sums of B are all equal to q. The spectral decomposition of L is obtained from that of B by a mere translation of the spectrum by q and the above assertions about $\lambda_0 = -\lambda^B_{\max} + q = 0$ follow.

Next we prove the geometric multiplicity of any eigenvalue of $-L$ is 1. Let λ be an eigenvalue of $-L$ and $f = (f(i))_{1\leqslant i,j\leqslant n}$ be a corresponding non-zero eigenvector. We have

$$\forall\ 1 \leqslant i < n, \qquad -L[f](i) \;=\; \sum_{j=1}^{i+1} -L(i,j)f(j) \;=\; \lambda f(i) \quad ,$$

from which we deduce

$$f(i+1) \;=\; \frac{1}{-L(i,i+1)}\left[\lambda f(i) + \sum_{j=1}^{i} L(i,j)f(j)\right]. \tag{2.1}$$

This means $f(i+1)$ can be determined by $f(1), \cdots, f(i)$. We must have $f(1) \neq 0$, otherwise (2.1) implies $f \equiv 0$. It follows that $\{cf : c \in \mathbb{R}\backslash\{0\}\}$ is the set of all the eigenvectors associated to λ. Therefore the dimension of the eigenspace related to λ is 1, which means the geometric multiplicity of λ is 1. See Section 1.4 of Horn and Johnson (1990). $\square$

Proposition 2.2. *Given $S = \{(\lambda_k, n_k) \in \mathbb{R}_+ \times \mathbb{N} : k \in [\![l]\!]\} \in \mathcal{U}_{\mathrm{r}}(n)$, we denote $0 = \theta_1 < \theta_2 = \cdots = \theta_{n_2+1} < \theta_{n_2+2} = \cdots = \theta_{n_2+n_3+1} < \cdots \leqslant \theta_n$ the elements of the multiset $\{\lambda_k[n_k] : k \in [\![l]\!]\}$. There exists a family of irreducible pure-birth Markov generators $(L_x)_{x\in(0,\theta_2)}$, such that the Jordan spectra of $-L_x$ are equal to S for all $x \in (0,\theta_2)$. More precisely, the isospectral family $(L_x)_{x\in(0,\theta_2)}$ is given by*

$$L_x \;\coloneqq\; \begin{pmatrix} -(\theta_2 - x) & \theta_2 - x & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\theta_3 & \theta_3 & \cdots & 0 & 0 & 0 \\ 0 & 0 & -\theta_4 & \theta_4 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & -\theta_{n-1} & \theta_{n-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\theta_n & \theta_n \\ u_2 & u_3 & u_4 & \cdots & u_{n-1} & u_n & -x \end{pmatrix}, \tag{2.2}$$

where for any $0 < x < \theta_2$,

$$u_2 \;\coloneqq\; x\prod_{k=3}^{n}\frac{\theta_k - \theta_2 + x}{\theta_k} > 0, \tag{2.3}$$

$$\forall\ 3 \leqslant i \leqslant n, \qquad u_i \;\coloneqq\; \frac{x(\theta_2 - x)}{\theta_i}\prod_{k=i+1}^{n}\frac{\theta_k - \theta_2 + x}{\theta_k} \;>\; 0 \tag{2.4}$$

(with the usual convention that for $i = n$, the empty product is equal to 1, note that $x = \sum_{i=2}^n u_i$).

Proof: For any fixed $x \in (0,\theta_2)$, to prove that the Jordan spectrum of $-L_x$ in (2.2) is S, by Proposition 2.1, we only need to prove the algebraic spectrum of $-L_x$ is $\{0,\theta_2,\cdots,\theta_n\}$. Define the diagonal matrix

$$L^* \;\coloneqq\; \begin{pmatrix} -\theta_2 & & & & \\ & -\theta_3 & & 0 & \\ & & \ddots & & \\ & 0 & & -\theta_n & \\ & & & & 0 \end{pmatrix}$$

and for any $s := (s_2, \ldots, s_n) \in \mathbb{R}^{n-1}$,

$$L_{x,s} := \begin{pmatrix} -(\theta_2 - x) & \theta_2 - x & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\theta_3 & \theta_3 & \cdots & 0 & 0 & 0 \\ 0 & 0 & -\theta_4 & \theta_4 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & -\theta_{n-1} & \theta_{n-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\theta_n & \theta_n \\ s_2 & s_3 & s_4 & \cdots & s_{n-1} & s_n & -x \end{pmatrix}.$$

Furthermore, for any $s \in \mathbb{R}^{n-1}$, define

$$A(a,s) := |aI + L_{x,s}| - |aI + L^*|.$$

If for some given $s \in \mathbb{R}^{n-1}$, we have

$$\forall\ a \in \mathbb{R}, \qquad A(a,s) = 0, \tag{2.5}$$

then $L_{x,s}$ and L^* have the same eigenvalues with the same corresponding algebraic multiplicities. Actually, we will prove that (2.5) is equivalent to $s_i = u_i$ $(2 \leqslant i \leqslant n)$, where u_i is given in (2.3). For this, split with respect to the last row of the first determinant and we get

$$\begin{aligned} A(a,s) &= \begin{vmatrix} a-(\theta_2-x) & \theta_2 - x & \cdots & 0 & 0 \\ 0 & a-\theta_3 & \theta_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & a-\theta_n & \theta_n \\ s_2 & s_3 & \cdots & s_n & a-x \end{vmatrix} - \begin{vmatrix} a-\theta_2 & \theta_2 & \cdots & 0 & 0 \\ 0 & a-\theta_3 & \theta_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & a-\theta_n & \theta_n \\ 0 & 0 & \cdots & 0 & a \end{vmatrix} \\ &= \begin{vmatrix} a-(\theta_2-x) & \theta_2 - x & \cdots & 0 & 0 \\ 0 & a-\theta_3 & \theta_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & a-\theta_n & \theta_n \\ s_2 & s_3 & \cdots & s_n & -x \end{vmatrix} + \begin{vmatrix} a-(\theta_2-x) & \theta_2 - x & \cdots & 0 & 0 \\ 0 & a-\theta_3 & \theta_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & a-\theta_n & \theta_n \\ 0 & 0 & \cdots & 0 & a \end{vmatrix} \\ &\quad - \begin{vmatrix} a-\theta_2 & \theta_2 & \cdots & 0 & 0 \\ 0 & a-\theta_3 & \theta_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & a-\theta_n & \theta_n \\ 0 & 0 & \cdots & 0 & a \end{vmatrix} \\ &= \begin{vmatrix} x & -x & \cdots & 0 & 0 \\ 0 & a-\theta_3 & \theta_3 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & a-\theta_n & \theta_n \\ 0 & 0 & \cdots & 0 & a \end{vmatrix} + \begin{vmatrix} a-(\theta_2-x) & \theta_2 - x & \cdots & 0 & 0 \\ 0 & a-\theta_3 & \theta_3 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & a-\theta_n & \theta_n \\ s_2 & s_3 & \cdots & s_n & -x \end{vmatrix}, \end{aligned}$$

where we used that the second and third matrices in the second equality have all rows equal except for the first one. It follows that

$$
\begin{aligned}
A(a,s) &= x(a-\theta_3)\cdots(a-\theta_n)a+(a-\theta_2+x)(a-\theta_3)\cdots(a-\theta_n)(-x)+\sum_{i=2}^{n}s_i\cdot G_i(s)\\
&= x\cdot(\theta_2-x)(a-\theta_3)\cdots(a-\theta_n)+\sum_{i=2}^{n}s_i\cdot G_i(s), \qquad (2.6)
\end{aligned}
$$

where

$$
\begin{cases}
G_2(s) \coloneqq (-1)^{n-1}(\theta_2-x)\theta_3\cdots\theta_n\\
G_3(s) \coloneqq (-1)^{n-2}(a-\theta_2+x)\cdot\theta_3\cdots\theta_n\\
\forall\ 4\leqslant i\leqslant n,\quad G_i(s) \coloneqq (-1)^{n+1-i}(a-\theta_2+x)(a-\theta_3)\cdots(a-\theta_{i-1})\cdot\theta_i\cdots\theta_n.
\end{cases} \qquad (2.7)
$$

So $A(a,s)$ can be written as

$$
A(a,s) \eqqcolon F_{n-2}(s)a^{n-2}+\cdots+F_1(s)a+F_0(s), \qquad (2.8)
$$

where the $F_i(s)$, for $0\leqslant i<n$, are polynomial functions of $s_1,\cdots,s_n$, and hence (2.5) is equivalent to

$$
F_{n-2}(s)=0,\quad F_{n-3}(s)=0,\quad \cdots,\quad F_0(s)=0. \qquad (2.9)
$$

Now assume that (2.9) holds. By (2.6)-(2.8), we have

$$
F_{n-2}(s) = x(\theta_2-x)-\theta_n s_n
$$

and thus

$$
s_n = \frac{x(\theta_2-x)}{\theta_n}. \qquad (2.10)
$$

Further, by (2.6)-(2.8) and (2.10), we have

$$
\begin{aligned}
F_{n-3}(s) &= -x(\theta_2-x)\sum_{i=3}^{n}\theta_i+s_n\theta_n\sum_{i=2}^{n-1}\theta_i^{'}+s_{n-1}\theta_{n-1}\theta_n,\\
&= -x(\theta_2-x)\left(\theta_n-\theta_2+x\right)+s_{n-1}\theta_{n-1}\theta_n,
\end{aligned}
$$

where $\theta_2^{'}=\theta_2-x,\ \theta_i^{'}=\theta_i$, for $3\leqslant i\leqslant n$, and thus

$$
s_{n-1} = \frac{x(\theta_2-x)}{\theta_{n-1}}\cdot\frac{(\theta_n-\theta_2+x)}{\theta_n}. \qquad (2.11)
$$

Similarly,

$$
\begin{aligned}
F_{n-4}(s) &= x(\theta_2-x)\sum_{3\leqslant i_1<i_2\leqslant n}\theta_{i_1}\theta_{i_2}-s_n\theta_n\sum_{2\leqslant i_1<i_2<n}\theta_{i_1}^{'}\theta_{i_2}^{'}-s_{n-1}\prod_{k=n-1}^{n}\theta_k\cdot\sum_{i=2}^{n-2}\theta_i^{'}-s_{n-2}\prod_{k=n-2}^{n}\theta_k\\
&= x(\theta_2-x)\left[\sum_{3\leqslant i_1<i_2\leqslant n}\theta_{i_1}\theta_{i_2}-\sum_{2\leqslant i_1<i_2<n}\theta_{i_1}^{'}\theta_{i_2}^{'}-(\theta_n-\theta_2+x)\sum_{i=2}^{n-2}\theta_i^{'}\right]-s_{n-2}\prod_{k=n-2}^{n}\theta_k\\
&= x(\theta_2-x)\left[(\theta_n-\theta_2^{'})\sum_{i=3}^{n-1}\theta_i-(\theta_n-\theta_2+x)\sum_{i=2}^{n-2}\theta_i^{'}\right]-s_{n-2}\prod_{k=n-2}^{n}\theta_k\\
&= x(\theta_2-x)\left(\theta_n-\theta_2+x\right)\left(\theta_{n-1}-\theta_2+x\right)-s_{n-2}\prod_{k=n-2}^{n}\theta_k,
\end{aligned}
$$

from which we have

$$s_{n-2} = \frac{x(\theta_2 - x)}{\theta_{n-2}} \prod_{k=n-1}^{n} \frac{\theta_k - \theta_2 + x}{\theta_k}.$$

For the iteration argument, suppose that

$$\begin{aligned} s_{n-l} &= \frac{x(\theta_2 - x)}{\theta_{n-l}} \prod_{k=n-l+1}^{n} \frac{\theta_k - \theta_2 + x}{\theta_k} \\ &= \frac{x\theta_2'}{\theta_{n-l}'} \prod_{k=n-l+1}^{n} \frac{\theta_k' - \theta_2'}{\theta_k'} \end{aligned} \tag{2.12}$$

for $l = 0, 1 \cdots, m-1 < n-2$, where $\prod_{k=n+1}^{n} \frac{\theta_k' - \theta_2'}{\theta_k'} := 1$. We will prove that it also holds for $l = m$. In fact, by (2.6)-(2.8), we have

$$\begin{aligned} &F_{n-m-2}(s) \\ &= (-1)^m x\theta_2' \sum_{3\leqslant i_1<\cdots<i_m\leqslant n} \theta_{i_1}\cdots\theta_{i_m} - (-1)^m s_n\theta_n' \sum_{2\leqslant i_1<\cdots<i_m<n} \theta_{i_1}'\cdots\theta_{i_m}' - (-1)^m s_{n-1} \prod_{k=n-1}^{n}\theta_k' \cdot \sum_{2\leqslant i_1<\cdots<i_{m-1}\leqslant n-2} \theta_{i_1}'\cdots\theta_{i_{m-1}}' \\ &\qquad - \cdots - (-1)^m s_{n-m+1} \prod_{k=n-m+1}^{n} \theta_k' \cdot \sum_{i=2}^{n-m} \theta_i' - (-1)^m s_{n-m} \prod_{k=n-m}^{n} \theta_k'. \end{aligned}$$

So

$$\begin{aligned} &s_{n-m} \prod_{k=n-m}^{n} \theta_k' \\ &= x\theta_2' \sum_{3\leqslant i_1<\cdots<i_m\leqslant n} \theta_{i_1}\cdots\theta_{i_m} - s_n\theta_n' \sum_{2\leqslant i_1<\cdots<i_m<n} \theta_{i_1}'\cdots\theta_{i_m}' - \sum_{j=n-m+1}^{n-1} \left[s_j \cdot \prod_{k=j}^{n} \theta_k' \cdot \sum_{2\leqslant i_1<\cdots<i_{j+m-n}<j} \theta_{i_1}'\cdots\theta_{i_{j+m-n}}' \right] \\ &= x\theta_2' \left[\sum_{3\leqslant i_1<\cdots<i_m\leqslant n} \theta_{i_1}\cdots\theta_{i_m} - \sum_{2\leqslant i_1<\cdots<i_m<n} \theta_{i_1}'\cdots\theta_{i_m}' \right] - \sum_{j=n-m+1}^{n-1} \left[x\theta_2' \cdot \prod_{k=j+1}^{n} \left(\theta_k' - \theta_2'\right) \cdot \sum_{2\leqslant i_1<\cdots<i_{j+m-n}<j} \theta_{i_1}'\cdots\theta_{i_{j+m-n}}' \right] \\ &= x\theta_2'(\theta_n - \theta_2') \left[\sum_{3\leqslant i_1<\cdots<i_{m-1}<n} \theta_{i_1}'\cdots\theta_{i_{m-1}'} - \sum_{2\leqslant i_1<\cdots<i_{m-1}<n-2} \theta_{i_1}'\cdots\theta_{i_{m-1}}' \right] \\ &\qquad\qquad - \sum_{j=n-m+1}^{n-2} \left[x\theta_2' \cdot \prod_{k=j+1}^{n} \left(\theta_k' - \theta_2'\right) \cdot \sum_{2\leqslant i_1<\cdots<i_{j+m-n}<j} \theta_{i_1}'\cdots\theta_{i_{j+m-n}}' \right] \\ &= x\theta_2' \prod_{k=n-m+2}^{n} \left(\theta_k' - \theta_2'\right) \cdot \left[\sum_{i=3}^{n-m+1} \theta_i' - \sum_{i=2}^{n-m} \theta_i' \right] \\ &= x\theta_2' \prod_{k=n-m+1}^{n} \left(\theta_k' - \theta_2'\right) \end{aligned}$$

$$= \quad x\theta_2' \prod_{k=n-m+1}^{n} (\theta_k - \theta_2 + x)\,,$$

which means the equality (2.12) holds for $l = m$.

Altogether, we have deduced that, if (2.9) holds, then (2.12) is true for $l \in [\![0, n-2]\!]$. Conversely, if (2.12) holds for $l \in [\![0, n-2]\!]$, then it is easy to check from above deduction that (2.9) is also true. That is, for fixed $x \in (0, \theta_2)$, (2.5) is equivalent to $s_i = u_i$ for any $2 \leqslant i \leqslant n$, which means L_x in (2.2) and L^* have the same eigenvalues and the same corresponding algebraic multiplicities.

It remains to see that $x = \sum_{i=2}^n s_i$. Actually, by (2.6), (2.7) and (2.9), we have

$$0 \;=\; F_0(s) \;=\; \left[x \cdot (-1)^{n-2} + \sum_{i=2}^{n} s_i \cdot (-1)^{n-1} \right] (\theta_2 - x)\theta_3 \cdots \theta_n,$$

from which we get the equality. This completes the proof. □

Now we are ready to prove Theorem 1.2.

Proof of Theorem 1.2: By Proposition 2.1, we have $\sigma_{\mathrm{J}}(-\mathcal{S}_{\mathrm{r}}(n)) \subset \mathcal{U}_{\mathrm{r}}(n)$, while from Proposition 2.2, we have $\mathcal{U}_{\mathrm{r}}(n) \subset \sigma_{\mathrm{J}}(-\mathcal{T}_{\mathrm{r}}(n))$. So $\sigma_{\mathrm{J}}(-\mathcal{S}_{\mathrm{r}}(n)) \subset \sigma_{\mathrm{J}}(-\mathcal{T}_{\mathrm{r}}(n))$. Besides, it is easy to see that $\sigma_{\mathrm{J}}(-\mathcal{T}_{\mathrm{r}}(n)) \subset \sigma_{\mathrm{J}}(-\mathcal{S}_{\mathrm{r}}(n))$, which implies $\sigma_{\mathrm{J}}(-\mathcal{T}_{\mathrm{r}}(n)) = \sigma_{\mathrm{J}}(-\mathcal{S}_{\mathrm{r}}(n)) = \mathcal{U}_{\mathrm{r}}(n)$. □

3. Comparison of the speed of convergence

It is interesting to compare the mixing rates between the elements of the family described in Proposition 2.2.

For $x \in (0, \theta_2)$, let L_x be the Markov generator defined in (2.2) and consider π_x the associated invariant measure on $[\![n]\!]$, as well as $(P_x(t))_{t \geqslant 0}$ the corresponding semi-group. For any $x \in (0, \theta_2)$ and $t \geqslant 0$, $P_x(t)$ is just the $[\![n]\!] \times [\![n]\!]$ matrix $\exp(tL_x)$. Any probability measure μ on $[\![n]\!]$ is seen as a row vector and $\mu P_x(t)$ then stands for the law of the position at time t of a Markov process starting with μ as initial distribution and whose generator is L_x. For large $t \geqslant 0$, this law converges toward π_x and let us evaluate its mixing rate through

$$\begin{aligned} \mathcal{E}_x(t) \quad &\coloneqq \quad \max\{\|\pi_x - \mu P_x(t)\|_{\mathrm{tv}} \,:\, \mu \in \mathcal{P}([\![n]\!])\} \\ &= \quad \max\{\|\pi_x - \delta_z P_x(t)\|_{\mathrm{tv}} \,:\, z \in [\![n]\!]\}, \end{aligned}$$

where $\|\cdot\|_{\mathrm{tv}}$ stands for the total variation norm, $\mathcal{P}([\![n]\!])$ is the convex set of probability measures on $[\![n]\!]$ and δ_z is the Dirac mass at $z \in [\![n]\!]$. The last equality is a consequence of the fact that the Dirac masses are exactly the extreme points of $\mathcal{P}([\![n]\!])$.

Our goal here is to prove the following comparison:

Theorem 3.1. *For any $x, y \in (0, \theta_2)$, we have*

$$\forall\ t \geqslant 0, \qquad \mathcal{E}_x(t) \quad \leqslant \quad \eta_{x,y}\mathcal{E}_y(t)$$

with

$$\begin{aligned} \eta_{x,y} \quad &\coloneqq \quad \frac{\theta_2 + (x \vee y) - 2(x \wedge y)}{\theta_2 - (x \vee y)} \\ &= \quad 1 + 2\frac{|y - x|}{\theta_2 - (x \vee y)} \quad \geqslant \quad 1 \end{aligned}$$

Note that when $x = y$, the inequality is indeed and equality.

For $x \in (0, \theta_2)$ and $\epsilon \in (0, 2]$, define the mixing time

$$\tau_x(\epsilon) \quad \coloneqq \quad \inf\{t \geqslant 0 \,:\, \mathcal{E}_x(t) \leqslant \epsilon\}.$$

The following result is immediately deduced from Theorem 3.1:

Corollary 3.2. *For any $x, y \in (0, \theta_2)$ and $\epsilon \in (0, 2]$, we have*

$$\tau_x(\epsilon) \leqslant \tau_y(\epsilon/\eta_{x,y}).$$

Note for a given irreducible finite Markov process such as the one generated by $L_{\theta_2/2}$, it is always possible to find $0 < A \leqslant A'$ (depending on $L_{\theta_2/2}$, i.e. on $\theta_2 \leqslant \theta_3 \leqslant \cdots \leqslant \theta_n$) such that

$$\forall\ \epsilon \in (0, 2], \qquad A \ln(1/\epsilon) \leqslant \tau_{\theta_2/2}(\epsilon) \leqslant A' \ln(1/\epsilon)$$

(the first bound is obtained by considering that the process has not jumped before a time of order $\ln(1/\epsilon)$ and the second bound is deduced from exponential convergence in total variation).

It then follows from Corollary 3.2 that uniformly in $x \in (0, \theta_2)$ not very close to θ_2, the mixing times $\tau_x(\epsilon)$ and $\tau_{\theta_2/2}(\epsilon)$ are of the same order as ϵ goes to 0_+.

To go in the direction of Theorem 3.1, let us start by a general observation on the intertwining relations that can be abstractly deduced from Theorem 1.2, even if it will not be directly useful in the sequel.

Lemma 3.3. *Consider L an irreducible Markov generator on $[\![n]\!]$ whose Jordan spectrum belongs to $\mathcal{U}_{\mathrm{r}}(n)$. Then there exist $\widetilde{L} \in \mathcal{T}_{\mathrm{r}}(n)$ and two invertible Markov matrices Λ and $\widetilde{\Lambda}$ such that*

$$\begin{aligned} L\Lambda &= \Lambda\widetilde{L}, \\ \widetilde{L}\widetilde{\Lambda} &= \widetilde{\Lambda}L. \end{aligned}$$

Proof: It was seen in Miclo (2018) that two irreducible Markov generators with the same Jordan spectrum can be intertwined via invertible Markov kernels. To apply this result, it is sufficient to use Proposition 2.2, which provides with an irreducible pure-birth Markov generator $\widetilde{L}$ with the same Jordan spectrum as L. □

Remark 3.4. To relate the speeds of convergence of the Markov processes generated by L and $\widetilde{L}$, one needs quantitative informations on Λ and $\widetilde{\Lambda}$, see e.g. Miclo and Patie (2021). We will not explore further this direction here.

Now let us study the similarity between any two Markov generators L_x and L_y as defined in (2.2). Define the $[\![n]\!] \times [\![n]\!]$ lower triangular matrix $\Lambda_{a,b}$ $(0 \leqslant a, b < \theta_2)$ as follows.

$$\left\{ \begin{array}{rrcl} & \Lambda_{a,b}(1,1) & := & 1 \\ & \Lambda_{a,b}(2,1) & := & \dfrac{b-a}{\theta_2 - a} \\ \forall\ 2 \leqslant i \leqslant n, & \Lambda_{a,b}(i,i) & := & \dfrac{\theta_2 - b}{\theta_2 - a} \\ \forall\ 1 \leqslant j < i \leqslant n, & \Lambda_{a,b}(i,j) & := & \dfrac{\theta_2'}{\theta_{j+1}'} \left[\displaystyle\prod_{k=j+2}^{i} \frac{\theta_k - \theta_2'}{\theta_k} \right] \Lambda_{a,b}(2,1) \end{array} \right. \tag{3.1}$$

where $\theta_2' = \theta_2 - b$, $\theta_i' = \theta_i$, for $3 \leqslant i \leqslant n$, with the convention $\prod_{k=i+1}^{i} \frac{\theta_k - \theta_2'}{\theta_k} = 1$. Then we have the following intertwining results.

Proposition 3.5. *For any two irreducible pure-birth Markov generators L_x and L_y as defined in (2.2), with $0 < x < y < \theta_2$, there exist two invertible Markov matrices Λ and $\widetilde{\Lambda}$ such that*

$$L_x\Lambda = \Lambda L_y, \tag{3.2}$$

$$L_y\widetilde{\Lambda} = \widetilde{\Lambda}L_x. \tag{3.3}$$

More precisely, we can take

$$\begin{aligned}\Lambda &= \Lambda_{x,y}\\ \widetilde{\Lambda} &= \widetilde{\Lambda}_{x,y} \coloneqq (1+\varepsilon)^{-1}\left[\pi_x + \varepsilon\Lambda_{y,x}\right]\end{aligned}$$

where, by slightly abusing notations, π_x stands for the matrix whose rows are equal to π_x, the stationary distribution of L_x, and ε is a constant with absolute value small enough. In fact,

$$\forall\ 1 \leqslant j \leqslant n, \qquad \pi_x(j) = \Lambda_{0,x}(n,j) \tag{3.4}$$

and it is sufficient to take

$$-C \leqslant \varepsilon \leqslant C \cdot \frac{(\theta_2 - x)}{(y-x)}, \tag{3.5}$$

where

$$C = \frac{x(\theta_2 - y)}{\theta_2\theta_n}\prod_{k=3}^{n}\frac{\theta_k - \theta_2 + x}{\theta_k}.$$

Proof: Denote the rows of Λ by $\Lambda(i,\cdot)$, for $1 \leqslant i \leqslant n$. From (3.2), we have

$$\begin{aligned}\forall\ 1 \leqslant i < n, \qquad -\theta''_{i+1}\Lambda(i,\cdot) + \theta''_{i+1}\Lambda(i+1,\cdot) &= \Lambda(i,\cdot)L_y, &(3.6)\\ \sum_{i=1}^{n-1} u_{i+1}(x)\Lambda(i,\cdot) - x\Lambda(n,\cdot) &= \Lambda(n,\cdot)L_y, &(3.7)\end{aligned}$$

where $\theta''_2 = \theta_2 - x$, $\theta''_i = \theta_i$, for $3 \leqslant i \leqslant n$. By (3.6), we get

$$\forall\ 1 \leqslant i < n, \qquad \Lambda(i+1,\cdot) = \theta''^{-1}_{i+1}\Lambda(i,\cdot)\left[\theta''_{i+1}I + L_y\right]. \tag{3.8}$$

Let us choose $\Lambda(1,\cdot) = (1,0,\cdots,0)$. Then for any $2 \leqslant i \leqslant n$, $\Lambda(i,\cdot)$ can be derived from a step by step calculation, and this exactly leads to (3.1). And it is not difficult to check that (3.1) also satisfies (3.7). Besides, it is easy to see that for all $j \leqslant i$, $\Lambda(i,j) > 0$ when $x < y$, and by (3.8), we have

$$\forall\ 1 \leqslant i < n, \qquad \Lambda(i+1,\cdot)\mathbf{1} = \Lambda(i,\cdot)\mathbf{1} = \Lambda(1,\cdot)\mathbf{1} = 1.$$

So, $\Lambda = \Lambda_{x,y}$ is a Markov matrix and satisfies (3.2).

Next, we prove that (3.3) and (3.4) hold. Actually, exchanging x and y in (3.2), we get

$$\Lambda_{y,x}L_x = L_y\Lambda_{y,x}. \tag{3.9}$$

Note that, when $y = 0$, (3.9) is also true and $\Lambda_{0,x}$ is also a Markov matrix. In this case, the n-th line of L_0 is zero and we get

$$\Lambda_{0,x}(n,\cdot)L_x = 0.$$

So, the n-th row $\Lambda_{0,x}(n,\cdot)$ is just the stationary distribution of L_x, and hence (3.3) follows directly.

Finally, we prove that $\widetilde{\Lambda}$ is a Markov matrix under (3.5). In fact, similar to $\Lambda_{x,y}$, it holds $\Lambda_{y,x}\mathbf{1} = \mathbf{1}$ and thus $\widetilde{\Lambda}\mathbf{1} = \mathbf{1}$. Furthermore, it is easy to check that (3.5) implies $\widetilde{\Lambda}(i,j) > 0$, for all $j \leqslant i$. This completes the proof.

■

Despite the kernels $\widetilde{\Lambda}_{x,y}$ are Markov for $x < y$, they do not lead to an interweaving relation in the sense of Miclo and Patie (2021). Indeed, such a relation would require that we can write

$$\Lambda_{x,y}\widetilde{\Lambda}_{x,y} = \int_0^\infty \exp(tL_x)\,\nu(dt)$$

for a probability ν on $\mathbb{R}_+$. This is not possible, due to the component π_x in $\widetilde{\Lambda}_{x,y}$ (which induces that ν should give a positive weight to $+\infty$). Nevertheless, we believe (3.3) holds with more appropriate

Markov kernels $\widetilde{\Lambda}$ than $\widetilde{\Lambda}_{x,y}$, so that an interweaving relation could indeed be worked out. But here, instead of using $\widetilde{\Lambda}_{x,y}$, we will only take into account the kernels $\Lambda_{x,y}$, even for $x > y$, when they are no longer Markovian.

Here is a semi-group type property satisfied by the family $(\Lambda_{x,y})_{x,y\in(0,\theta_2)}$.

Lemma 3.6. *For any $x, y, z \in (0, \theta_2)$, we have*

$$\Lambda_{x,y}\Lambda_{y,z} \quad = \quad \Lambda_{x,z}.$$

Particularly, we have

$$\Lambda_{x,y}\Lambda_{y,x} \quad = \quad \Lambda_{x,x} \;=\; I.$$

Proof: Since

$$\begin{aligned} L_x\Lambda_{x,y} &= \Lambda_{x,y}L_y \\ L_y\Lambda_{y,z} &= \Lambda_{y,z}L_z, \end{aligned}$$

it holds that

$$L_x\Lambda_{x,y}\Lambda_{y,z} \;=\; \Lambda_{x,y}L_y\Lambda_{y,z} \;=\; \Lambda_{x,y}\Lambda_{y,z}L_z.$$

Note that if Λ is a lower triangular matrix and

$$\begin{cases} \Lambda(1,\cdot) &= (1,0,\cdots,0), \\ L_x\Lambda &= \Lambda L_z, \end{cases} \tag{3.10}$$

then $\Lambda(i,\cdot)$, for $2 \leqslant i \leqslant n$, can be determined by a step by step calculation. It is easy to see $\Lambda_{x,y}\Lambda_{y,z}$ and $\Lambda_{x,z}$ are both lower triangular matrices and satisfy (3.10), so they are equal.

■

Define the matrix norm $\|\cdot\|$ via

$$\forall\ A \in \mathcal{M}(n), \qquad \|A\| \;:=\; \max_i \sum_j |A(i,j)|.$$

Another useful observation is:

Lemma 3.7. *For any $a, b \in (0, \theta_2)$, we have*

$$\|\Lambda_{a,b}\| \quad \leqslant \quad \begin{cases} 1 & , \textit{ if } a \leqslant b \\ \frac{\theta_2+a-2b}{\theta_2-a} & , \textit{ otherwise} \end{cases}$$

Proof: • When $a \leqslant b$, since $\Lambda_{a,b}$ is Markovian, it is clear that $\|\Lambda_{a,b}\| = 1$.

• When $a > b$, since

$$\begin{aligned} \forall\ j < i, \qquad \Lambda_{a,b}(i,j) &< 0; \\ \forall\ i, \qquad \Lambda_{a,b}(i,i) &> 0; \\ \sum_j \Lambda_{a,b}(i,j) &= 1, \end{aligned}$$

we have

$$\begin{aligned} \|\Lambda_{a,b}\| &= \max\left\{\Lambda_{a,b}(i,i) - \sum_{j\neq i}\Lambda_{a,b}(i,j) \;:\; 1 \leqslant i \leqslant n\right\} \\ &= \max\{2\Lambda_{a,b}(i,i) - 1 \;:\; 1 \leqslant i \leqslant n\} \\ &= 2\Lambda_{a,b}(2,2) - 1 \\ &= \frac{\theta_2 + a - 2b}{\theta_2 - a}. \end{aligned}$$

□

We can now come to the

Proof of Theorem 3.1: Fix $x, y \in (0, \theta_2)$. First let us show that

$$\pi_x \Lambda_{x,y} \quad = \quad \pi_y. \tag{3.11}$$

From (3.2), we deduce that

$$\begin{aligned} \pi_x \Lambda_{x,y} L_y &= \pi_x L_x \Lambda_{x,y} \\ &= 0, \end{aligned}$$

since π_x is the only probability measure such that $\pi_x L_x = 0$. It is also the only measure μ on $[\![n]\!]$ with $\mu(\mathbb{1}) = 1$ (where $\mathbb{1}$ is the function always taking the value 1 on $[\![n]\!]$) such that $\mu L_x = 0$. So to get (3.11), it remains to check that $\pi_x \Lambda_{x,y}(\mathbb{1}) = 1$. This is clear when $x \leqslant y$, since $\Lambda_{x,y}$ is Markovian. It is in fact always true, as it can be directly checked on (3.1). A shorter proof is to use again the intertwining relation (3.2): we get

$$\begin{aligned} L_x \Lambda_{x,y} \mathbb{1} &= \Lambda_{x,y} L_y \mathbb{1} \\ &= 0. \end{aligned}$$

So by irreducibility of L_x, the vector $\Lambda_{x,y}\mathbb{1}$ is constant. Its first value is $\Lambda_{x,y}(1,1) = 1$. It follows that $\Lambda_{x,y}\mathbb{1} = \mathbb{1}$ and $\pi_x \Lambda_{x,y}(\mathbb{1}) = 1$.

Note that the intertwining relation implies (in fact is equivalent to)

$$\forall\ t \geqslant 0, \qquad P_x(t)\Lambda_{x,y} \quad = \quad \Lambda_{x,y} P_y(t).$$

Taking into account the last assertion of Lemma 3.6, we write for any $\mu \in \mathcal{P}([\![n]\!])$,

$$\begin{aligned} \pi_x - \mu P_x(t) &= \pi_x \Lambda_{x,y} \Lambda_{y,x} - \mu P_x(t) \Lambda_{x,y} \Lambda_{y,x} \\ &= \pi_y \Lambda_{y,x} - \mu \Lambda_{x,y} P_y(t) \Lambda_{y,x} \\ &= (\pi_y - \nu P_y(t))\Lambda_{y,x} \end{aligned} \tag{3.12}$$

where

$$\nu \quad \coloneqq \quad \mu \Lambda_{x,y}.$$

When $x \leqslant y$, ν is a probability measure, otherwise it may only be a (signed) measure, but at least it satisfies $\nu(\mathbb{1}) = \mu(\Lambda_{x,y}\mathbb{1}) = \mu(\mathbb{1}) = 1$. Consider the decomposition of ν into its non-negative and non-positive parts: $\nu = \nu_+ - \nu_-$. Recall that ν_+ and ν_- are two non-negative measures and that $\|\nu\|_{\mathrm{tv}} = \nu_+(\mathbb{1}) + \nu_-(\mathbb{1})$. We have $\nu_+(\mathbb{1}) - \nu_-(\mathbb{1}) = 1$ and by consequence,

$$\pi_y - \nu P_y(t) \quad = \quad \nu_+(\mathbb{1})\left(\pi_y - \frac{\nu_+}{\nu_+(\mathbb{1})} P_y(t)\right) - \nu_-(\mathbb{1})\left(\pi_y - \frac{\nu_-}{\nu_-(\mathbb{1})} P_y(t)\right).$$

It follows that

$$\begin{aligned} \|\pi_y - \nu P_y(t)\|_{\mathrm{tv}} &\leqslant \nu_+(\mathbb{1}) \left\|\pi_y - \frac{\nu_-}{\nu_-(\mathbb{1})} P_y(t)\right\|_{\mathrm{tv}} + \nu_-(\mathbb{1}) \left\|\pi_y - \frac{\nu_+}{\nu_+(\mathbb{1})} P_y(t)\right\|_{\mathrm{tv}} \\ &\leqslant \nu_+(\mathbb{1})\mathcal{E}_y(t) + \nu_-(\mathbb{1})\mathcal{E}_y(t) \\ &= \|\nu\|_{\mathrm{tv}}\, \mathcal{E}_y(t) \end{aligned}$$

From the definition of the matrix norm $|\!|\!| \cdot |\!|\!|$, for any measure ν on $[\![n]\!]$ and any $[\![n]\!] \times [\![n]\!]$ matrix A, we have

$$\|\nu A\|_{\mathrm{tv}} \quad \leqslant \quad |\!|\!| A |\!|\!|\ \|\nu\|_{\mathrm{tv}}.$$

We deduce from (3.12)

$$\begin{aligned} \|\pi_x - \mu P_x(t)\|_{\mathrm{tv}} &\leqslant |\!|\!| \Lambda_{y,x} |\!|\!|\ \|\pi_y - \mu P_y(t)\|_{\mathrm{tv}} \\ &\leqslant |\!|\!| \Lambda_{y,x} |\!|\!|\ \|\nu\|_{\mathrm{tv}}\, \mathcal{E}_y(t). \end{aligned}$$

Recall that $\nu = \mu\Lambda_{x,y}$, so that

$$\begin{aligned} \|\nu\|_{\mathrm{tv}} &\leqslant \|\Lambda_{x,y}\| \, \|\mu\|_{\mathrm{tv}} \\ &= \|\Lambda_{x,y}\|. \end{aligned}$$

Finally, we get

$$\|\pi_x - \mu P_x(t)\|_{\mathrm{tv}} \leqslant \|\Lambda_{y,x}\| \|\Lambda_{x,y}\| \mathcal{E}_y(t),$$

and taking the supremum over $\mu \in \mathcal{P}([\![n]\!])$,

$$\mathcal{E}_x(t) \leqslant \|\Lambda_{y,x}\| \|\Lambda_{x,y}\| \mathcal{E}_y(t).$$

The desired result follows from Lemma 3.7. □

Acknowledgements

The authors thank an anonymous referee for his or her useful comments which have helped to improve the quality of the manuscript.

References

Horn, R. A. and Johnson, C. R. *Matrix analysis*. Cambridge University Press, Cambridge (1990). ISBN 0-521-38632-2. MR1084815.

Levin, D. A., Peres, Y., and Wilmer, E. L. *Markov chains and mixing times*. American Mathematical Society, Providence, RI (2009). ISBN 978-0-8218-4739-8. MR2466937.

Miclo, L. On the Markovian similarity. In *Séminaire de Probabilités XLIX*, volume 2215 of *Lecture Notes in Math.*, pp. 375–403. Springer, Cham (2018). MR3838869.

Miclo, L. and Patie, P. On interweaving relations. *J. Funct. Anal.*, **280** (3), Paper No. 108816, 53 (2021). MR4170787.

Seneta, E. *Non-negative matrices and Markov chains*. Springer Series in Statistics. Springer, New York (2006). ISBN 978-0387-29765-1; 0-387-29765-0. MR2209438.

Yang, X. Q. *The construction theory of denumerable Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. Hunan Science and Technology Publishing House, Changsha; John Wiley & Sons, Ltd., Chichester (1990). ISBN 0-471-92490-3. MR1119552.