On a family of isospectral pure-birth processes

Laurent Miclo[*] and Chi Zhang[†]

[*] Institut de Mathématiques de Toulouse
Toulouse School of Economics
CNRS and University of Toulouse

[†] College of Mathematical Sciences
Ocean University of China

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.

Keywords: pure-birth processes, spectra of Markov processes, mixing rates, intertwining relations.

MSC2010: primary: 60J27, secondary: 60J35, 15A18, 37A25.

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 [2] of Levin, Peres and Wilmer. 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. [3, 4]. 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. The **Jordan spectrum** of a matrix $A \in \mathcal{M}(n)$ is the multi-set

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

where each (λ_k, n_k), $k \in [\![l]\!] := \{1, 2, ..., l\}$, corresponds to a Jordan block of size $n_k \times n_k$ associated to the eigenvalue λ_k. Since several of these blocks can be equal, $\sigma_{\mathrm{J}}(A)$ as to be seen as a multi-set. Note that $\sum_{k \in [\![l]\!]} n_k = n$. The traditional **geometric and algebraic spectra** of A are the multi-sets respectively defined by

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

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 $\lambda \in \mathbb{R}$.

The sets of the form $S := \{(\lambda_k, n_k) \in \mathbb{C} \times \mathbb{N} \ : \ k \in [\![l]\!]\}$ 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 $k \in [\![2, n]\!]$, 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 purpose 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 $[\![n]\!]$, $p := (p_0, p_1, ..., p_l)$ 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 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 [5] of Yang. 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 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 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 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 2. In the last section we will compare the mixing rates of the elements of this family, see Theorem 6.

2 An isospectral family

The purpose of this section is to prove Theorem 2.

We need some preparations.

Proposition 4 *Let $L \in \mathcal{S}_{\mathrm{r}}(n)$. Then the Jordan spectrum of $-L$ is a uni-spectral set.*

Proof

By irreducibility, the smallest eigenvalue of $-L$ is 0 and has algebraic multiplicity 1. So we only need to 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

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

from which we deduce

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

This means $f(i+1)$ can be determined by $f(1), \cdots, f(i)$. We must have $f(1) \neq 0$, otherwise (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 Horn and Johnson [1, Section 1.4].

■

Proposition 5 *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*

$$L_x \;:=\; \begin{pmatrix} -(\theta_2 - x) & \theta_2 - x & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\theta_3 & \theta_3 & \cdots & 0 & 0 & 0 \\ 0 & 0 & -\theta_4 & \theta_4 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & -\theta_{n-1} & \theta_{n-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\theta_n & \theta_n \\ u_2 & u_3 & u_4 & \cdots & u_{n-1} & u_n & -x \end{pmatrix} \tag{2}$$

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

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

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

(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) is S, by Proposition 4, we only need to prove the algebraic spectrum of $-L_x$ is $\{0, \theta_2, \cdots, \theta_n\}$. Define the diagonal matrix

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

and for any $a \in \mathbb{R}$ and $u := (u_2, ..., u_n) \in \mathbb{R}^{n-1}$,

$$
A(a, u) \;:=\; |aI + L_x| - |aI + L^*|
$$

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

$$
\forall\, a \in \mathbb{R}, \qquad A(a, u) \;=\; 0 \tag{3}
$$

then L_x and L^* have the same eigenvalues with the same corresponding algebraic multiplicities.

By splitting with respect to the last row of the first determinant, we get

$$
A(a, u) \;=\; \begin{vmatrix} a - (\theta_2 - x) & \theta_2 - x & \cdots & 0 & 0 \\ 0 & a - \theta_3 & \theta_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & a - \theta_n & \theta_n \\ u_2 & u_3 & \cdots & u_n & a - x \end{vmatrix} - \begin{vmatrix} a - \theta_2 & \theta_2 & \cdots & 0 & 0 \\ 0 & a - \theta_3 & \theta_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & a - \theta_n & \theta_n \\ 0 & 0 & \cdots & 0 & a \end{vmatrix}
$$

$$
=\; \begin{vmatrix} a - (\theta_2 - x) & \theta_2 - x & \cdots & 0 & 0 \\ 0 & a - \theta_3 & \theta_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & a - \theta_n & \theta_n \\ u_2 & u_3 & \cdots & u_n & -x \end{vmatrix} + \begin{vmatrix} a - (\theta_2 - x) & \theta_2 - x & \cdots & 0 & 0 \\ 0 & a - \theta_3 & \theta_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & a - \theta_n & \theta_n \\ 0 & 0 & \cdots & 0 & a \end{vmatrix}
$$

$$
-\; \begin{vmatrix} a - \theta_2 & \theta_2 & \cdots & 0 & 0 \\ 0 & a - \theta_3 & \theta_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & a - \theta_n & \theta_n \\ 0 & 0 & \cdots & 0 & a \end{vmatrix}
$$

$$
=\; \begin{vmatrix} x & -x & \cdots & 0 & 0 \\ 0 & a - \theta_3 & \theta_3 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & a - \theta_n & \theta_n \\ 0 & 0 & \cdots & 0 & a \end{vmatrix} + \begin{vmatrix} a - (\theta_2 - x) & \theta_2 - x & \cdots & 0 & 0 \\ 0 & a - \theta_3 & \theta_3 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & a - \theta_n & \theta_n \\ u_2 & u_3 & \cdots & u_n & -x \end{vmatrix}
$$

$$
=\; x(a - \theta_3) \cdots (a - \theta_n) a + (a - \theta_2 + x)(a - \theta_3) \cdots (a - \theta_n)(-x) + \sum_{i=2}^{n} u_i \cdot G_i(u)
$$

$$
=\; x \cdot (\theta_2 - x)(a - \theta_3) \cdots (a - \theta_n) + \sum_{i=2}^{n} u_i \cdot G_i(u) \tag{4}
$$

5

where

$$
\begin{cases}
\quad G_2(u) & := & (-1)^{n-1}(\theta_2 - x)\theta_3 \cdots \theta_n \\
\quad G_3(u) & := & (-1)^{n-2}(a - \theta_2 + x) \cdot \theta_3 \cdots \theta_n \\
\forall\, 4 \leqslant i \leqslant n, \quad G_i(u) & := & (-1)^{n+1-i}(a - \theta_2 + x)(a - \theta_3) \cdots (a - \theta_{i-1}) \cdot \theta_i \cdots \theta_n
\end{cases}
\tag{5}
$$

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

$$
A(a, u) \quad =: \quad F_{n-2}(u)a^{n-2} + \cdots + F_1(u)a + F_0(u)
\tag{6}
$$

where the $F_i(u)$, for $0 \leqslant i < n$, are polynomial functions of $u_1, \cdots, u_n$, and hence (3) is equivalent to

$$
F_{n-2}(u) = 0, \quad F_{n-3}(u) = 0, \quad \cdots, \quad F_0(u) = 0
\tag{7}
$$

Now assume that (7) holds. By (4)-(6), we have

$$
F_{n-2}(u) \quad = \quad x(\theta_2 - x) - \theta_n u_n
$$

and thus

$$
u_n \quad = \quad \frac{x(\theta_2 - x)}{\theta_n}
\tag{8}
$$

Further, by (4)-(6) and (8), we have

$$
F_{n-3}(u) \quad = \quad -x(\theta_2 - x)\sum_{i=3}^{n}\theta_i + u_n\theta_n\sum_{i=2}^{n-1}\theta_i' + u_{n-1}\theta_{n-1}\theta_n
$$
$$
= \quad -x(\theta_2 - x)(\theta_n - \theta_2 + x) + u_{n-1}\theta_{n-1}\theta_n
$$

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

$$
u_{n-1} \quad = \quad \frac{x(\theta_2 - x)}{\theta_{n-1}} \cdot \frac{(\theta_n - \theta_2 + x)}{\theta_n}.
\tag{9}
$$

Similarly,

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

from which we have

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

6

For the iteration argument, suppose that

$$u_{n-l} = \frac{x(\theta_2 - x)}{\theta_{n-l}} \prod_{k=n-l+1}^{n} \frac{\theta_k - \theta_2 + x}{\theta_k} \tag{10}$$

$$= \frac{x\theta_2'}{\theta_{n-l}'} \prod_{k=n-l+1}^{n} \frac{\theta_k' - \theta_2'}{\theta_k'} \tag{11}$$

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 (4)-(6), we have

$F_{n-m-2}(u)$

$$= (-1)^m x\theta_2' \sum_{3 \leqslant i_1 < \cdots < i_m \leqslant n} \theta_{i_1} \cdots \theta_{i_m} - (-1)^m u_n \theta_n' \sum_{2 \leqslant i_1 < \cdots < i_m < n} \theta_{i_1}' \cdots \theta_{i_m}' - (-1)^m u_{n-1} \prod_{k=n-1}^{n} \theta_k' \cdot \sum_{2 \leqslant i_1 < \cdots < i_{m-1} \leqslant n-2} \theta_{i_1}' \cdots \theta_{i_{m-1}}'$$

$$- \cdots - (-1)^m u_{n-m+1} \prod_{k=n-m+1}^{n} \theta_k' \cdot \sum_{i=2}^{n-m} \theta_i' - (-1)^m u_{n-m} \prod_{k=n-m}^{n} \theta_k'.$$

So

$$u_{n-m} \prod_{k=n-m}^{n} \theta_k'$$

$$= x\theta_2' \sum_{3 \leqslant i_1 < \cdots < i_m \leqslant n} \theta_{i_1} \cdots \theta_{i_m} - u_n \theta_n' \sum_{2 \leqslant i_1 < \cdots < i_m < n} \theta_{i_1}' \cdots \theta_{i_m}' - \sum_{j=n-m+1}^{n-1} \left[u_j \cdot \prod_{k=j}^{n} \theta_k' \cdot \sum_{2 \leqslant i_1 < \cdots < i_{j+m-n} < j} \theta_{i_1}' \cdots \theta_{i_{j+m-n}}' \right]$$

$$= x\theta_2' \left[\sum_{3 \leqslant i_1 < \cdots < i_m \leqslant n} \theta_{i_1} \cdots \theta_{i_m} - \sum_{2 \leqslant i_1 < \cdots < i_m < n} \theta_{i_1}' \cdots \theta_{i_m}' \right] - \sum_{j=n-m+1}^{n-1} \left[x\theta_2' \cdot \prod_{k=j+1}^{n} \left(\theta_k' - \theta_2' \right) \cdot \sum_{2 \leqslant i_1 < \cdots < i_{j+m-n} < j} \theta_{i_1}' \cdots \theta_{i_{j+m-n}}' \right]$$

$$= x\theta_2'(\theta_n - \theta_2') \left[\sum_{3 \leqslant i_1 < \cdots < i_{m-1} < n} \theta_{i_1}' \cdots \theta_{i_{m-1}'}' - \sum_{2 \leqslant i_1 < \cdots < i_{m-1} < n-2} \theta_{i_1}' \cdots \theta_{i_{m-1}}' \right]$$

$$- \sum_{j=n-m+1}^{n-2} \left[x\theta_2' \cdot \prod_{k=j+1}^{n} \left(\theta_k' - \theta_2' \right) \cdot \sum_{2 \leqslant i_1 < \cdots < i_{j+m-n} < j} \theta_{i_1}' \cdots \theta_{i_{j+m-n}}' \right]$$

$$= x\theta_2' \prod_{k=n-m+2}^{n} \left(\theta_k' - \theta_2' \right) \cdot \left[\sum_{i=3}^{n-m+1} \theta_i' - \sum_{i=2}^{n-m} \theta_i' \right]$$

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

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

which means the equality (10) holds for $l = m$.

Altogether, we have deduced that, if (7) holds, then (10) is true for $l \in [\![0, n - 2]\!]$. Conversely, if (10) holds for $l \in [\![0, n - 2]\!]$, then it is easy to check from above deduction that (7) is also true, which means L_x and L^* have the same eigenvalues and the same corresponding algebraic multiplicities.

It remains to see that $x = \sum_{i=2}^{n} u_i$. Actually, by (4), (5) and (7), we have

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

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

∎

Now we are ready to prove Theorem 2.

Proof of Theorem 2

By Proposition 4, we have $\sigma_{\mathrm{J}}(-\mathcal{S}_{\mathrm{r}}(n)) \subset \mathcal{U}_{\mathrm{r}}(n)$, while from Proposition 5, 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 5.

For $x \in (0, \theta_2)$, let L_x be the Markov generator defined in (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

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

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 6 *For any $x, y \in (0, \theta_2)$, we have*

$$\forall\; t \geqslant 0, \qquad \mathcal{E}_x(t) \;\leqslant\; \frac{\theta_2 + (x \vee y) - 2(x \wedge y)}{\theta_2 - (x \vee y)} \mathcal{E}_y(t)$$

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

Let us start by a general observation on the intertwining relations that can be abstractly deduced from Theorem 2.

Corollary 7 *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 [3] 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 5, which provides with an irreducible pure-birth Markov generator $\widetilde{L}$ with the same Jordan spectrum as L.

<div align="right">■</div>

Now let us study the similarity between any two Markov generators L_x and L_y as defined in (2). Define the $[\![n]\!] \times [\![n]\!]$ lower triangular matrix $\Lambda_{a,b}$ $(0 \leqslant a, b < \theta_2)$ as follows.

$$
\begin{cases}
\Lambda_{a,b}(1,1) & := \ 1 \\
\Lambda_{a,b}(2,1) & := \ \frac{b-a}{\theta_2-a} \\
\forall \ 2 \leqslant i \leqslant n, \quad \Lambda_{a,b}(i,i) & := \ \frac{\theta_2-b}{\theta_2-a} \\
\forall \ 1 \leqslant j < i \leqslant n, \quad \Lambda_{a,b}(i,j) & := \ \frac{\theta_2'}{\theta_{j+1}'}\left[\prod_{k=j+2}^{i} \frac{\theta_k-\theta_2'}{\theta_k}\right]\Lambda_{a,b}(2,1)
\end{cases}
\tag{12}
$$

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 8 *For any two irreducible pure-birth Markov generators L_x and L_y as defined in (2), with $0 < x < y < \theta_2$, there exist two invertible Markov matrices Λ and $\widetilde{\Lambda}$ such that*

$$
\begin{aligned}
L_x\Lambda &= \Lambda L_y & (13) \\
L_y\widetilde{\Lambda} &= \widetilde{\Lambda}L_x & (14)
\end{aligned}
$$

More precisely, we can take

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

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,

$$
\forall \ 1 \leqslant i, j \leqslant n, \qquad \pi_x(i,j) = \Lambda_{0,x}(n,j) \tag{15}
$$

and it is sufficient to take

$$
-\frac{x(\theta_2-y)}{\theta_2\theta_n}\prod_{k=3}^{n}\frac{\theta_k-\theta_2+x}{\theta_k} \ \leqslant \ \varepsilon \ \leqslant \ \frac{x(\theta_2-x)(\theta_2-y)}{(y-x)\theta_2\theta_n}\prod_{k=3}^{n}\frac{\theta_k-\theta_2+x}{\theta_k} \tag{16}
$$

Proof

Denote the rows of Λ by $\Lambda(i,\cdot)$, for $1 \leqslant i \leqslant n$. From (13), we have

$$
\forall \ 1 \leqslant i < n, \qquad -\theta_{i+1}''\Lambda(i,\cdot) + \theta_{i+1}''\Lambda(i+1,\cdot) = \Lambda(i,\cdot)L_y \tag{17}
$$

$$
\sum_{i=1}^{n-1}u_{i+1}(x)\Lambda(i,\cdot) - x\Lambda(n,\cdot) = \Lambda(n,\cdot)L_y \tag{18}
$$

where $\theta_2'' = \theta_2 - x$, $\theta_i'' = \theta_i$, for $3 \leqslant i \leqslant n$. By (17), we get

$$
\forall \ 1 \leqslant i < n, \qquad \Lambda(i+1,\cdot) = \theta_{i+1}''^{-1}\Lambda(i,\cdot)\left[\theta_{i+1}''I + L_y\right] \tag{19}
$$

9

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 (12). And it is not difficult to check that (12) also satisfies (18). Besides, it is easy to see that for all $j \leqslant i$, $\Lambda(i,j) > 0$ when $x < y$, and by (19), 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 (13).

Next, we prove that (14) and (15) hold. Actually, exchanging x and y in (13), we get

$$\Lambda_{y,x}L_x \;\;=\;\; L_y\Lambda_{y,x}. \tag{20}$$

Note that, when $y = 0$, (20) 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 (14) follows directly.

Finally, we prove that $\widetilde{\Lambda}$ is a Markov matrix under (16). 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 (16) 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 are not very useful to transfer information on convergence between the Markov processes generated by L_x and L_y. This is due to the component π_x which prevents the existence of an interweaving relation in the sense of [4]. Nevertheless, we believe (14) 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 9 *For any $x, y, z \in (0, \theta_2)$, we have*

$$\Lambda_{x,y}\Lambda_{y,z} \;\;=\;\; \Lambda_{x,z}$$

Particularly, we have

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

Proof

Since

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

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

$$\begin{cases} \Lambda(1,\cdot) \;\;=\;\; (1,0,\cdots,0) \\ L_x\Lambda \;\;=\;\; \Lambda L_z \end{cases} \tag{21}$$

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 (21), 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 10 *For any $a, b \in (0, \theta_2)$, we have*

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

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 6

Fix $x, y \in (0, \theta_2)$. First let us show that

$$\pi_x \Lambda_{x,y} \;=\; \pi_y \tag{22}$$

From (13), 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 (22), 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 (12). A shorter proof is to use again the intertwining relation (13): we get

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

11

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)

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

Taking into account the last assertion of Lemma 9, we write for any $\mu \in \mathcal{P}(\llbracket n \rrbracket)$,

$$
\begin{aligned}
\pi_x - \mu P_x(t) &= \pi_x\Lambda_{x,y}\Lambda_{y,x} - \mu P_x(t)\Lambda_{x,y}\Lambda_{y,x} \\
&= \pi_y\Lambda_{y,x} - \mu\Lambda_{x,y}P_y(t)\Lambda_{y,x} \\
&= (\pi_y - \nu P_y(t))\Lambda_{y,x}
\end{aligned}
\tag{23}
$$

where

$$\nu \;:=\; \mu\Lambda_{x,y}$$

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,

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

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 $\vertiii{\cdot}$, for any measure ν on $\llbracket n \rrbracket$ and any $\llbracket n \rrbracket \times \llbracket n \rrbracket$ matrix A, we have

$$\|\nu A\|_{\mathrm{tv}} \;\leqslant\; \vertiii{A}\,\|\nu\|_{\mathrm{tv}}$$

We deduce from (23)

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

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

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

Finally, we get

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

and taking the supremum over $\mu \in \mathcal{P}(\llbracket n \rrbracket)$,

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

The desired result follows from Lemma 10.

$\blacksquare$

References

[1] Roger A. Horn and Charles R. Johnson. *Matrix analysis.* Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.

[2] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times.* American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.

[3] Laurent Miclo. On the Markovian similarity. In *Séminaire de Probabilités XLIX*, volume 2215 of *Lecture Notes in Math.*, pages 375–403. Springer, Cham, 2018.

[4] Laurent Miclo and Pierre Patie. On interweaving relations. *arXiv e-prints*, October 2019.

[5] Xiang Qun Yang. *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. With a foreword by Zhen Ting Hou, Translated from the Chinese, With a foreword by David Kendall.

miclo@math.cnrs.fr

Institut de Mathéatiques de Toulouse
Université Paul Sabatier, 118, route de Narbonne
31062 Toulouse cedex 9, France

Toulouse School of Economics,
Manufacture des Tabacs, 21, Alléde Brienne
31015 Toulouse cedex 6, France

zhangchi@ouc.edu.cn

School of Mathematical Sciences
Ocean University of China, Qingdao
266100, People's Republic of China