On finite interweaving relations

Laurent Miclo

Toulouse School of Economics
Institut de Mathématiques de Toulouse
CNRS and University of Toulouse

Abstract

An interweaving relation is a Markovian similarity-type relation between two Markov chains introducing a warming-up time after which their time-marginal distributions can be tightly compared (for different initial distributions). For non-transient Markov transition kernels on the same state space, these relations are shown to be equivalent to the usual similarity relation. Some bounds are deduced on corresponding warming-up times, when the eigenvalues are furthermore assumed to be real. When the eigenvalues are non-negative, the same approach enables us to construct strong stationary times for irreducible Markov chains through interweaving relations with model absorbed Markov chains, thus extending a result due to Matthews in the reversible situation.

Keywords: Intertwining relations, interweaving relations, finite state space transition kernels, generalised spectral decompositions, Jordan blocks, warming-up times, strong stationary times.

MSC2020: primary: 60J10, secondary: 60J22, 15B51, 15A20, 15A18.

Fundings: the grants ANR-17-EURE-0010 and AFOSR-22IOE016 are acknowledged.

1 Introduction

This paper investigates certain similarity-type relations between non-transient Markov kernels on the same finite state space. The interest of these relations is to introduce a warming-up (random) time after which the time-marginal distributions of corresponding Markov chains can be strongly related. They will also enable us to revisit a result of Matthews [3] on strong stationary times associated to reversible Markov chains and to extend it to the non-reversible setting, under the assumption that the eigenvalues of its Markov kernel are non-negative.

Let us begin by recalling the kind of relations we are interested in.

On a finite set V with cardinal $|V| \geqslant 2$, let be given two Markov transition matrices P and $\widetilde{P}$.

We say that a (**Markov**) **intertwining relation** from P to $\widetilde{P}$ holds through the link Λ, which is another Markov transition matrix on V, when

$$P\Lambda = \Lambda\widetilde{P} \tag{1}$$

(since here all the considered relations will be Markovian, from now on we drop the adjective "Markov" for them). Intertwining relations have a long history starting with the seminal paper of Rogers and Pitman [6]. The Markov kernels P and $\widetilde{P}$ are sometimes called **dual** and **primal**, see e.g. Diaconis and Fill [1].

When Λ is furthermore invertible, (1) is called a **faithful intertwining relation** from P to $\widetilde{P}$ through Λ.

We say that there is a **bi-intertwining relation** between P and $\widetilde{P}$, via the links Λ and $\widetilde{\Lambda}$, when in addition to (1) we have

$$\widetilde{P}\widetilde{\Lambda} = \widetilde{\Lambda}P \tag{2}$$

This relation is said to be a **faithful bi-intertwining relation** when furthermore Λ and $\widetilde{\Lambda}$ are invertible.

A bi-intertwining relation between P and $\widetilde{P}$, via the links Λ and $\widetilde{\Lambda}$, is said to be an **interweaving relation** when there exists a probability distribution $q = (q_n)_{n\in\mathbb{Z}_+}$ on $\mathbb{Z}_+$ so that

$$\Lambda\widetilde{\Lambda} = \sum_{n\in\mathbb{Z}_+} q_n P^n \tag{3}$$

(note that the r.h.s. is necessarily convergent). This notion was introduced in [5], where q is seen as the distribution of a warming-up time, independent of the underlying Markov chains, after which a lot of convergence to equilibrium informations can be transferred from the primal chain to the dual chain. This feature will appear again in Section 4 and 5 below, but in a slightly distorted way.

It is **bi-interweaving relation**, when there also exists a probability distribution $\widetilde{q} = (\widetilde{q}_n)_{n\in\mathbb{Z}_+}$ on $\mathbb{Z}_+$ so that

$$\widetilde{\Lambda}\Lambda = \sum_{n\in\mathbb{Z}_+} \widetilde{q}_n \widetilde{P}^n \tag{4}$$

These relations, interweaving and bi-interweaving, are said to be **faithful** when Λ and $\widetilde{\Lambda}$ are invertible.

Remark 1 When there is both a faithful bi-intertwining relation between P and $\widetilde{P}$ and an interweaving relation (3), then (4) is necessarily satisfied with $\widetilde{q} = q$, namely we also have a faithful bi-interweaving relation. Indeed, from (3) we deduce

$$\Lambda\widetilde{\Lambda}\Lambda = \sum_{n\in\mathbb{Z}_+} q_n P^n \Lambda$$

and via (1) we obtain

$$\Lambda\widetilde{\Lambda}\Lambda \;=\; \Lambda\sum_{n\in\mathbb{Z}_+} q_n\widetilde{P}^n$$

It remains to multiply on the left by Λ^{-1} to get (4) with $\widetilde{q} = q$. ▫

A bi-intertwining relation always holds between P and $\widetilde{P}$: it is sufficient to take $\Lambda = \widetilde{\pi}$ (meaning that all the rows of Λ coincide with $\widetilde{\pi}$) and $\widetilde{\Lambda} = \pi$, where π and $\widetilde{\pi}$ are invariant probability measures for P and $\widetilde{P}$ respectively (they always exists in the context of finite state space, but in general they are not unique and their supports are not the whole state space V).

It is proven in [4] that two non-transient Markov matrices P and $\widetilde{P}$ are similar if and only if there exists a faithful bi-intertwining relation between them (be careful, we changed the names given in [4]: there, a link was necessarily invertible, bi-intertwining corresponded to *mutual intertwining* and faithful bi-intertwining was called *Markov-similarity*). In fact the arguments of [4] contain an error that can be corrected following the approach of Section 3 below, showing the above mentioned result of [4] is indeed true.

In contrast with this result, we will show that non-transience and similarity of P and $\widetilde{P}$ are not sufficient to ensure the existence of a faithful-bi-intertwining relation between them. To give a natural necessary and sufficient for the existence of such a relation for non-transient kernels, denote by C_1, C_2, ..., C_ℓ (respectively $\widetilde{C}_1$, $\widetilde{C}_2$, ..., $\widetilde{C}_\ell$) the irreducible classes of P (resp. $\widetilde{P}$). They are in the same number $\ell \in \mathbb{N}$, because this is the (geometric and algebraic) multiplicity of the eigenvalue 1. For all $l \in [\![\ell]\!] \coloneqq \{1, 2, ..., \ell\}$, denote P_{C_l} (resp. $\widetilde{P}_{\widetilde{C}_l}$) the restriction of P (resp. $\widetilde{P}$) to C_l (resp. $\widetilde{C}_l$). Note that these matrices are Markovian and irreducible.

Our first contribution here prove the following characterisation of faithful bi-interweaving relations:

Theorem 2 *There exists a faithful bi-interweaving relation between P and $\widetilde{P}$ if and only if there exists a permutation $\sigma \in \mathcal{S}_\ell$ and a probability q on $\mathbb{Z}_+$ such that for any $l \in [\![\ell]\!]$, $|C_l| = |\widetilde{C}_{\sigma(l)}|$ and there is a faithful bi-interweaving relation between P_{C_l} and $\widetilde{P}_{\widetilde{C}_{\sigma(l)}}$ with the same probability $\widetilde{q} = q$. It can furthermore be imposed that q has a finite support.*

Thus faithful bi-interweaving relations give a more accurate account of the geometry of non-transient Markov matrices than faithful bi-intertwining relations. As seen in [4], the case of transient Markov matrices is more complicated and will not be considered here.

Theorem 2 also enables us to give an example of Markov matrices P and $\widetilde{P}$ satisfying a faithful bi-intertwining relation but no faithful bi-interweaving relation:

Example 3 Consider on $V \coloneqq [\![4]\!]$,

$$P \;\coloneqq\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{pmatrix} \quad\text{and}\quad \widetilde{P} \;\coloneqq\; \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix}$$

Both matrices are non-transient: for P the state space can be decomposed into the union of the irreducible classes which are $C_1 \coloneqq \{1\}$ and $C_2 \coloneqq \{2, 3, 4\}$, and for $\widetilde{P}$ the irreducible classes are $\widetilde{C}_1 \coloneqq \{1, 2\}$ and $\widetilde{C}_2 \coloneqq \{3, 4\}$. The common spectrum of P and $\widetilde{P}$ corresponds to the eigenvalues 1 with multiplicity 2 and 0 with multiplicity 2. To check it, write

$$P \;=\; \begin{pmatrix} 1 & 0 \\ 0 & J_3 \end{pmatrix} \quad\text{and}\quad \widetilde{P} \;=\; \begin{pmatrix} J_2 & 0 \\ 0 & J_2 \end{pmatrix}$$

where for any $n \in \mathbb{N}$, J_n is the $n \times n$ matrix whose entries are all equal to $1/n$. Note that for $n \geqslant 2$, the spectrum of J_n is 1 with multiplicity 1 (the corresponding eigenspace is the space of constant vectors) and 0 with multiplicity $n-1$ (the corresponding eigenspace is the space of vectors whose entries sum up to 0).

Since P and $\widetilde{P}$ are non-transient and similar, we deduce from [4] the existence of a faithful bi-intertwining relation between P and $\widetilde{P}$. Nevertheless it is impossible to find a permutation $\sigma \in \mathcal{S}_2$ such that the condition of Theorem 2 is satisfied, so there is no faithful bi-intertwining relation between P and $\widetilde{P}$.

□

The situation where both P and $\widetilde{P}$ are irreducible, in addition to the assumptions of Theorem 2 will play an important role in its proof. It also enables us to be more precise about the possible restrictions on the size of the support of q:

Proposition 4 *Assume that P and $\widetilde{P}$ are irreducible and similar. Then there exists a faithful bi-interweaving relation between them, with equal probability distribution $q = \widetilde{q}$ whose support contains at most $d+1$ points, where d is the common period of P and $\widetilde{P}$. Thus when P is aperiodic, there exists a faithful bi-interweaving relation between P and $\widetilde{P}$ with $q = \widetilde{q}$ having a support with at most two points. When in addition of aperiodicity, we assume that none of the eigenvalues of P vanishes, then there exists a faithful bi-interweaving relation between P and $\widetilde{P}$ with $q = \widetilde{q}$ a Dirac mass.*

An example of bound on the support of q will be given at the end of Section 3, at least when all the eigenvalues of P are real. Nevertheless this bound is certainly too universal to be relevant and it can probably be improved in particular situations, as the steps of its proof are sometimes quite coarse. For instance it could not be applied in the following degenerate situation.

The arguments of the proof of this result can be adapted to recover the following result due to Matthews [3]. Let P be an irreducible Markov kernel on V whose invariant probability is denoted π. It is unique and its support is V. Assume that π is reversible for P, so that P seen as an operator on $\mathbb{L}^2(\pi)$ is symmetric and thus diagonalisable. Denote its eigenvalues (with multiplicities) by

$$1 = \theta_1 > \theta_2 \geqslant \theta_3 \geqslant \cdots \geqslant \theta_{|V|} \geqslant -1 \tag{5}$$

where the strict inequality comes from irreducibility. Let $(\varphi_k)_{k\in[\![|V|]\!]}$ be an orthonormal basis of $\mathbb{L}^2(\pi)$ consisting of corresponding eigenvectors, where the orthogonality is possible due to reversibility.

Let $X \coloneqq (X(n))_{n\in\mathbb{Z}_+}$ be a Markov chain admitting P for transition kernel. Recall that a strong stationary time for X is a finite stopping time τ (with respect to the filtration generated by X and maybe some independent randomness) such that τ and X_τ are independent and X_τ is distributed according to π. For any integers $m \leqslant n$, we will denote $[\![m,n]\!] \coloneqq \{m, m+1, ..., n\}$ and we already used previously the shortcut $[\![n]\!] \coloneqq [\![1,n]\!]$ for any $n \in \mathbb{N}$.

Let μ_0 be the law of X_0 and consider the probability distribution $\widetilde{\mu}_0$ on $[\![|V|]\!]$ given by

$$\forall\ k \in [\![|V|]\!], \qquad \widetilde{\mu}_0(k) \ \coloneqq \ \begin{cases} \frac{|\mu_0[\varphi_k]|}{Z(\mu_0)} & \text{, if } k \geqslant 2 \\ 0 & \text{, if } k = 1 \end{cases} \tag{6}$$

with

$$Z(\mu_0) \ \coloneqq \ \sum_{l\in[\![|V|]\!]\setminus\{1\}} |\mu_0[\varphi_l]| \tag{7}$$

The definition (6) is not valid when $Z(\mu_0) = 0$, namely when $\mu_0[\varphi_k] = \pi[\varphi_k]$ for all $k \in [\![|V|]\!]$, i.e. $\mu_0 = \pi$. In this situation we take $\widetilde{\mu}_0 \coloneqq \delta_1$, and formally the following result enables to recover that 0 is then a strong stationary time.

For the next result, assume that the eigenvalues in (5) are non-negative, i.e. $\theta_{|V|} \geqslant 0$. Consider $(G_k)_{k\in[\![2,|V|]\!]}$ a family of independent geometric random variables of respective parameters $(\theta_k)_{k\in[\![2,|V|]\!]}$, namely

$$\forall\ k \in [\![2,|V|]\!],\ \forall\ n \in \mathbb{N}, \qquad \mathbb{P}[G_k = n] \ = \ \theta_k^{n-1}(1-\theta_k)$$

Construct a random variable $\mathcal{G}$ taking values in $\mathbb{Z}_+$ in the following way. First we sample an element K from $[\![|V|]\!]$ according to $\widetilde{\mu}_0$. If $K = 1$ we take $\mathcal{G} \coloneqq 0$, and otherwise we take $\mathcal{G} \coloneqq G_k$.

Theorem 5 (Matthews [3]) *Assume that P is irreducible, reversible and that its eigenvalues are all non-negative. Then there exists a strong stationary time for X which is stochastically dominated by*

$$\left\lceil \max\left\{\frac{\ln(Z(\mu_0)\,\|\varphi_k\|_\infty)}{\ln(1/\theta_k)}\ :\ k \in [\![2, |V|]\!]\right\}\right\rceil + \mathcal{G} \tag{8}$$

(where the ratio vanishes when $\theta_k = 0$). This random variable is itself stochastically dominated by $\left\lceil \frac{\ln(|V|/\pi_\wedge)}{\ln(1/\theta_2)} \right\rceil + G_2$, where $\pi_\wedge \coloneqq \min\{\pi(x)\ :\ x \in V\}$ and where G_2 is a geometric random variable of parameter θ_2.

Instead of assuming that the eigenvalues of P are non-negative, Matthews [3] stated his result for the Markov kernel P^2, whose eigenvalues are indeed non-negative. Our second goal here is to extend Theorem 5 in Theorem 14 of Section 5, by removing the assumption of reversibility, up to introducing in the bounds a factor including the condition number of the Gramian matrix associated to the (generalized) eigenvectors.

The plan of the paper is as follows. In the next section we show Proposition 4 under the additional assumption that both P and $\widetilde{P}$ are reversible, as this situation allows for a pedagogical exposure of the main arguments. The full Proposition 4 and Theorem 2 are proven in Section 3. Section 4 adapts the arguments of Section 2 to recover Theorem 5. The underlying idea is to replace $\widetilde{P}$ by a very simple absorbed Markov kernel, serving as a “model”. The random variable $\mathcal{G}$ comes from this model, while the first term of (8) corresponds to a warming-up time between P and this model. This approach is extended in Section 5, taking into account the arguments of the proof of Proposition 4, to remove the reversibility assumption. The final section extend these results to the continuous framework.

2 The reversible case

Here for the sake of clarity, we show Proposition 4 under the simplifying assumption that both P and $\widetilde{P}$ are reversible. The proof takes up the arguments of [4] for intertwining and modifies them to deal with interweaving.

More precisely, our purpose is to show the following result:

Proposition 6 *Assume that P and $\widetilde{P}$ are similar and that P and $\widetilde{P}$ are irreducible and reversible. Then there exists a faithful bi-interweaving relation between them, with $q = \widetilde{q}$ whose support contains at most three points. When P is aperiodic (and by consequence $\widetilde{P}$ too), we can find such a relation with $q = \widetilde{q}$ whose support contains at most two points. When in addition to aperiodicity, none of the common eigenvalues of P and $\widetilde{P}$ vanishes, we can furthermore impose that $q = \widetilde{q}$ is a Dirac mass.*

Before coming to the proof of this proposition, we modify the arguments of Lemma 6 in [4] to construct more general invertible links Λ and $\widetilde{\Lambda}$ from V to V for a faithful bi-intertwining relation between P and $\widetilde{P}$, than those considered there.

Since P is irreducible and reversible, as before the statement of Theorem 5, we denote π, $1 = \theta_1 > \theta_2 \geqslant \theta_3 \geqslant \cdots \geqslant \theta_{|V|} \geqslant -1$ and $(\varphi_k)_{k\in[\![|V|]\!]}$, respectively, the invariant probability, the ordered eigenvalues and a corresponding orthonormal basis of $\mathbb{L}^2(\pi)$ of eigenvectors.

The same holds for $\widetilde{P}$ with the same eigenvalues. We denote $(\widetilde{\varphi}_k)_{k\in[\![|V|]\!]}$ a corresponding orthonormal basis of $\mathbb{L}^2(\widetilde{\pi})$ of eigenvectors, where $\widetilde{\pi}$ is the reversible probability of $\widetilde{P}$. Without loss of generality, we assume that $\varphi_1 = \widetilde{\varphi}_1 = \mathbb{1}$ (the function always taking the value 1).

To any sequence $b \coloneqq (b_k)_{k\in[\![2,|V|]\!]}$ of real numbers, associate the operator A_b defined by

$$\forall\ k \in [\![|V|]\!], \qquad A_b[\widetilde{\varphi}_k] \ \coloneqq \ \begin{cases} b_k\varphi_k & , \text{if } k \geqslant 2 \\ 0 & , \text{if } k = 1 \end{cases} \tag{9}$$

Symmetrically, to any sequence $\widetilde{b} \coloneqq (\widetilde{b}_k)_{k\in[\![2,|V|]\!]}$ of real numbers, associate the operator $\widetilde{A}_{\widetilde{b}}$ defined by

$$\forall\ k \in [\![|V|]\!], \qquad \widetilde{A}_{\widetilde{b}}[\varphi_k] \ \coloneqq \ \begin{cases} \widetilde{b}_k\widetilde{\varphi}_k & , \text{if } k \geqslant 2 \\ 0 & , \text{if } k = 1 \end{cases}$$

Here are the corresponding matrices:

Lemma 7 *We have for any $x, y \in V$,*

$$\begin{aligned} A_b(x,y) &= \sum_{k\geqslant 2} b_k\varphi_k(x)\widetilde{\varphi}_k(y)\widetilde{\pi}(y) \\ \widetilde{A}_{\widetilde{b}}(x,y) &= \sum_{k\geqslant 2} \widetilde{b}_k\widetilde{\varphi}_k(x)\varphi_k(y)\pi(y) \end{aligned}$$

It follows that

$$\begin{aligned} |A_b(x,y)| &\leqslant \sqrt{\frac{\widetilde{\pi}(y)}{\pi(x)}} \max_{k\in[\![2,|V|]\!]} |b_k| \leqslant \frac{1}{\sqrt{\pi_\wedge\widetilde{\pi}_\wedge}} \max_{k\in[\![2,|V|]\!]} |b_k|\widetilde{\pi}(y) \\ |\widetilde{A}_{\widetilde{b}}(x,y)| &\leqslant \sqrt{\frac{\pi(y)}{\widetilde{\pi}(x)}} \max_{k\in[\![2,|V|]\!]} |\widetilde{b}_k| \leqslant \frac{1}{\sqrt{\pi_\wedge\widetilde{\pi}_\wedge}} \max_{k\in[\![2,|V|]\!]} |\widetilde{b}_k|\pi(y) \end{aligned}$$

where we recall that

$$\begin{aligned} \pi_\wedge &\coloneqq \min_{x\in V} \pi(x) \\ \widetilde{\pi}_\wedge &\coloneqq \min_{x\in V} \widetilde{\pi}(x) \end{aligned}$$

Proof

For any $b \coloneqq (b_k)_{k\in[\![2,|V|]\!]} \in \mathbb{R}^{[\![2,|V|]\!]}$, introduce the matrix A'_b whose entries are given by

$$\forall\ x, y \in V, \qquad A'_b(x,y) \ \coloneqq \ \sum_{k\geqslant 2} b_k\varphi_k(x)\widetilde{\varphi}_k(y)\widetilde{\pi}(y)$$

To show that A'_b is the matrix associated to the operator A_b, it is sufficient to check that for any $l \in [\![|V|]\!]$,

$$\forall\ x \in V, \qquad \sum_{y\in V} A'_b(x,y)\widetilde{\varphi}_l(y) \ = \ b_l\varphi_l(x)$$

with the convention that $b_1 = 0$.

By definition of A'_b, we compute

$$\begin{aligned} \sum_{y\in V} A'_b(x,y)\widetilde{\varphi}_l(y) &= \sum_{k\geqslant 2} b_k\varphi_k(x)\widetilde{\pi}[\widetilde{\varphi}_k\widetilde{\varphi}_l] \\ &= b_l\varphi_l(x) \end{aligned}$$

by orthonormality of the basis $(\widetilde{\varphi}_k)_{k\in[\![|V|]\!]}$ in $\mathbb{L}^2(\widetilde{\pi})$.

This shows the first announced equality.

The second equality is obtained by symmetry, exchanging the roles of P and $\widetilde{P}$.

To prove the bounds, for any $y \in V$, introduce the following decomposition of the indicator function $\mathbb{1}_y$ of y in the basis $(\widetilde{\varphi}_k)_{k\in[\![|V|]\!]}$:

$$\mathbb{1}_y(\cdot) \quad = \sum_{k\in[\![|V|]\!]} \widetilde{\alpha}_k(y)\widetilde{\varphi}_k \tag{10}$$

where by orthonormality, the coefficients $(\widetilde{\alpha}_k(y))_{k\in[\![|V|]\!]}$ are given by

$$\widetilde{\alpha}_k(y) \quad = \quad \widetilde{\pi}[\mathbb{1}_y\widetilde{\varphi}_k] \quad = \quad \widetilde{\pi}(y)\widetilde{\varphi}_k(y)$$

Applying (10) at the point y, we get

$$\begin{aligned} 1 &= \mathbb{1}_y(y) \\ &= \sum_{k\in[\![|V|]\!]} \widetilde{\alpha}_k(y)\widetilde{\varphi}_k(y) \\ &= \sum_{k\in[\![|V|]\!]} \widetilde{\varphi}_k^2(y)\widetilde{\pi}(y) \\ &= \widetilde{\pi}(y) + \widetilde{\pi}(y) \sum_{k\in[\![2,|V|]\!]} \widetilde{\varphi}_k^2(y) \end{aligned}$$

so that

$$\begin{aligned} \sum_{k\in[\![2,|V|]\!]} \widetilde{\varphi}_k^2(y) &= \frac{1}{\widetilde{\pi}(y)} - 1 \\ &\leqslant \frac{1}{\widetilde{\pi}(y)} \end{aligned}$$

Similarly, we get

$$\sum_{k\in[\![2,|V|]\!]} \varphi_k^2(x) \quad \leqslant \quad \frac{1}{\pi(x)} \tag{11}$$

Cauchy-Schwartz inequality now leads to

$$\begin{aligned} |A_b(x,y)| &\leqslant \max_{k\in[\![2,|V|]\!]} |b_k| \sum_{k\geqslant 2} |\varphi_k(x)||\widetilde{\varphi}_k(y)|\widetilde{\pi}(y) \\ &\leqslant \max_{k\in[\![2,|V|]\!]} |b_k| \sqrt{\sum_{k\geqslant 2} \varphi_k^2(x)} \sqrt{\sum_{k\geqslant 2} \widetilde{\varphi}_k^2(y)}\widetilde{\pi}(y) \\ &\leqslant \max_{k\in[\![2,|V|]\!]} |b_k| \frac{1}{\sqrt{\pi(x)\widetilde{\pi}(y)}}\widetilde{\pi}(y) \end{aligned}$$

and thus to the first announced bounds.

The second ones follow by symmetry. ■

We are interested in the operator

$$\Lambda_b \quad \coloneqq \quad \widetilde{\pi} + A_b \tag{12}$$

where again $\widetilde{\pi}$ is interpreted as the matrix whose rows are all equal to the probability $\widetilde{\pi}$. We check that

$$\forall\ k \in [\![|V|]\!], \qquad \Lambda_b[\widetilde{\varphi}_k] \quad \coloneqq \quad \begin{cases} b_k\varphi_k & \text{, if } k \geqslant 2 \\ \varphi_1 & \text{, if } k = 1 \end{cases}$$

due to the fact that for $k \in [\![2,|V|]\!]$, we have $\widetilde{\pi}[\widetilde{\varphi}_k] = \widetilde{\pi}[\widetilde{\varphi}_1\widetilde{\varphi}_k] = 0$ by orthogonality.

It implies the intertwining relation $P\Lambda_b = \Lambda_b\widetilde{P}$ and Λ_b is invertible as soon as all the entries of b are non-zero.

From the relation $\Lambda_b[\mathbb{1}] = \Lambda_b[\widetilde{\varphi}_1] = \varphi_1 = \mathbb{1}$, it appears that the row sums of Λ_b are all equal to 1.

Furthermore, all the entries of Λ_b will be non-negative as soon as

$$\forall\, x,y \in V, \qquad \widetilde{\pi}(y) - |A_b(x,y)| \;\geqslant\; 0$$

From Lemma 7, this is true when

$$\max_{k\in[\![2,|V|]\!]} |b_k| \;\leqslant\; \sqrt{\pi_\wedge\widetilde{\pi}_\wedge} \tag{13}$$

Since similar arguments are valid for $\widetilde{\Lambda}_{\widetilde{b}} \coloneqq \pi + \widetilde{A}_{\widetilde{b}}$, we get a faithful bi-intertwining relation between P and $\widetilde{P}$, with Λ_b and $\widetilde{\Lambda}_{\widetilde{b}}$ as links, by choosing any b and $\widetilde{b}$ with coordinates belonging to $[-\sqrt{\pi_\wedge\widetilde{\pi}_\wedge}, \sqrt{\pi_\wedge\widetilde{\pi}_\wedge}]\backslash\{0\}$.

With these preliminaries in hand, we can now come to the

Proof of Proposition 6

We use the links Λ_b and $\widetilde{\Lambda}_{\widetilde{b}}$ defined above and look for conditions on b and $\widetilde{b}$ so that (3) and (4) are satisfied with a probability $q = \widetilde{q}$ with minimal support.

Let us first assume that P is aperiodic, which is equivalent to the fact that in (5), we have $\theta_{|V|} > -1$.

Concerning (3), we have on one hand,

$$\forall\, k \in [\![|V|]\!], \qquad \Lambda_b\widetilde{\Lambda}_{\widetilde{b}}[\varphi_k] \;=\; \begin{cases} \varphi_1 & \text{, if } k=1 \\ \widetilde{b}_k b_k \varphi_k & \text{, if } k \geqslant 2 \end{cases}$$

and on the other hand, for a given probability $q \coloneqq (q_n)_{n\in\mathbb{Z}_+}$,

$$\forall\, k \in [\![|V|]\!], \qquad \sum_{n\in\mathbb{Z}_+} q_n P^n[\varphi_k] \;=\; \sum_{n\in\mathbb{Z}_+} q_n\theta_k^n\varphi_k$$

Note that for $k=1$, we have

$$\Lambda_b\widetilde{\Lambda}_{\widetilde{b}}[\varphi_1] \;=\; \sum_{n\in\mathbb{Z}_+} q_n P^n[\varphi_1]$$

since both terms are equal to $\mathbb{1}$.

Thus the desired equality (3) is equivalent to

$$\forall\, k \in [\![2,|V|]\!], \qquad \sum_{n\in\mathbb{Z}_+} q_n\theta_k^n \;=\; \widetilde{b}_k b_k \tag{14}$$

As alluded to at the end of the proposition, let us look for a probability $q = \delta_{n_0}$, the Dirac mass at some $n_0 \in \mathbb{Z}_+$. The above condition then writes

$$\forall\, k \in [\![2,|V|]\!], \qquad \theta_k^{n_0} \;=\; \widetilde{b}_k b_k$$

Consider $\zeta \coloneqq \max\{|\theta_k| \;:\; k \in [\![2,|V|]\!]\}$, we have $\zeta \in [0,1)$ by irreducibility, reversibility and aperiodicity. It follows that if we take

$$n_0 \;\coloneqq\; 1 + \left\lfloor \frac{\ln(\pi_\wedge\widetilde{\pi}_\wedge)}{\ln(\zeta)} \right\rfloor$$

(note that both logarithms in the integer part $\lfloor\cdot\rfloor$ are negative, since $\pi_\wedge, \widetilde{\pi}_\wedge < 1/2$, as $|V| \geqslant 2$), then (13) is satisfied as soon as we take

$$\forall\, k \in [\![2,|V|]\!], \qquad b_k \;\coloneqq\; \sqrt{|\theta_k^{n_0}|}$$

$$\forall\ k \in [\![2, |V|]\!], \qquad \widetilde{b}_k \ \coloneqq\ \sqrt{|\theta_k^{n_0}|}\mathrm{sign}(\theta_k^{n_0})$$

where $\mathrm{sign}(\cdot)$ is the sign mapping (with e.g. the convention that $\mathrm{sign}(0) = 1$).

Furthermore, when none of the eigenvalues θ_k, for $k \in [\![2, |V|]\!]$, vanishes, the entries of b and $\widetilde{b}$ are non-zero, so we get the wanted faithful interweaving relation (3) with the links Λ_b and $\widetilde{\Lambda}_{\widetilde{b}}$, and q a Dirac mass.

To get the wanted faithful bi-interweaving relation, with $\widetilde{q} = q$ a Dirac mass, we can proceed similarly, since we deduce from (4) the same equations for $\widetilde{q}$ as for q due to the isospectrality of P and $\widetilde{P}$, or we just rely on Remark 1.

When some of the eigenvalues θ_k, for $k \in [\![2, |V|]\!]$, vanish, we rather consider a probability of the form

$$q \ \coloneqq\ \frac{\pi_\wedge \widetilde{\pi}_\wedge}{2}\delta_0 + \left(1 - \frac{\pi_\wedge \widetilde{\pi}_\wedge}{2}\right)\delta_{n_1} \tag{15}$$

with

$$n_1 \ \coloneqq\ 1 + \left\lfloor \frac{\ln(\pi_\wedge \widetilde{\pi}_\wedge/4)}{\ln(\zeta)} \right\rfloor \tag{16}$$

Indeed, defining for any θ in the complex unit disk,

$$\begin{aligned} Q(\theta) \ &\coloneqq\ \sum_{n\in\mathbb{Z}_+} q_n \theta^n \qquad (17) \\ &=\ \frac{\pi_\wedge \widetilde{\pi}_\wedge}{2} + \left(1 - \frac{\pi_\wedge \widetilde{\pi}_\wedge}{2}\right)\theta^{n_1} \end{aligned}$$

we get for any $k \in [\![2, |V|]\!]$,

$$\frac{\pi_\wedge \widetilde{\pi}_\wedge}{2} - \zeta^{n_1} \ \leqslant\ Q(\theta_k) \ \leqslant\ \frac{\pi_\wedge \widetilde{\pi}_\wedge}{2} + \zeta^{n_1}$$

By choice of n_1, these bounds imply

$$\frac{\pi_\wedge \widetilde{\pi}_\wedge}{2} - \frac{\pi_\wedge \widetilde{\pi}_\wedge}{4} \ \leqslant\ Q(\theta_k) \ \leqslant\ \frac{\pi_\wedge \widetilde{\pi}_\wedge}{2} + \frac{\pi_\wedge \widetilde{\pi}_\wedge}{4}$$

i.e.

$$\frac{\pi_\wedge \widetilde{\pi}_\wedge}{4} \ \leqslant\ Q(\theta_k) \ \leqslant\ 3\frac{\pi_\wedge \widetilde{\pi}_\wedge}{4} \tag{18}$$

Thus considering

$$\forall\ k \in [\![2, |V|]\!], \qquad b_k \ =\ \widetilde{b}_k \ \coloneqq\ \sqrt{|Q(\theta_k)|} \tag{19}$$

we get the wanted faithful interweaving relation (3) with the links Λ_b and $\widetilde{\Lambda}_{\widetilde{b}}$.

Again, Remark 1 provides the wanted faithful bi-interweaving relation, with $\widetilde{q} = q$ supported by two points.

Let us now come to the situation where P is periodic, so that in (5) we have $\theta_{|V|-1} > \theta_{|V|} = -1$. Indeed, under the irreducibility and reversibility assumptions, the aperiodicity is equivalent to the existence of a (necessarily unique) eigenvalue -1, that is why both P and $\widetilde{P}$ are aperiodic together, when they have the same spectrum.

The previous considerations are still valid: it is sufficient to find b, $\widetilde{b}$ and q (with $\widetilde{q} = q$), so that (14) holds with (13) and

$$\min_{k\in[\![2,|V|]\!]} |b_k| \ >\ 0 \tag{20}$$

The only difference with the above arguments comes from $k = |V|$ in (14), namely

$$\sum_{n\in\mathbb{Z}_+} q_n(-1)^n \quad = \quad \widetilde{b}_k b_k$$

It leads us to replace (15) by

$$q \quad := \quad \frac{\pi_\wedge\widetilde{\pi}_\wedge}{2}\delta_0 + \left(1 - \frac{\pi_\wedge\widetilde{\pi}_\wedge}{2}\right)\frac{\delta_{n_1} + \delta_{n_1+1}}{2}$$

with n_1 still given by (16).

Indeed, (18) is still true for $k \in [\![2, |V| - 1]\!]$. For $k = |V|$, we get $Q(\theta_{|V|}) = \pi_\wedge\widetilde{\pi}_\wedge/2$. Thus taking again (19), we get (14) satisfied with (13) and (20). Furthermore the support of $q = \widetilde{q}$ only contains the three points $0, n_1$ and $n_1 + 1$. ■

3 The general case

Our purpose here is to show Proposition 4 and Theorem 2. Proposition 4 is the transposition to interweaving relations of Lemma 7 in [4] for intertwining relations. Unfortunately the proof of the latter is wrong, so we are to present new arguments that enable us to correct it.

Before coming to the proof of Proposition 4, we need some reminders from complex linear algebra. Recall that seen as a complex matrix, P is similar to a block matrix, whose blocks are of Jordan types (θ_1, γ_1), (θ_2, γ_2), ..., (θ_r, γ_r), where $\theta_1, \theta_2, ..., \theta_r \in \mathbb{C}$ are the eigenvalues of P (with geometric multiplicities) and $r \in \mathbb{N}$, $\gamma_1, \gamma_2, ..., \gamma_r \in \mathbb{N}$ satisfy $\gamma_1 + \gamma_2 + \cdots + \gamma_r = |V|$. Recall that a Jordan block of type (θ, n) is a $n \times n$ matrix whose diagonal entries are equal to θ, whose first above diagonal entries are equal to 1 and whose other entries vanish. The set $\{(\theta_k, \gamma_k) : k \in [\![r]\!]\}$ is a characteristic invariant for the complex similarity class of P and will be called the **characteristic set** of P. It is characterised by the existence of a complex basis $(\varphi_{(k,l)})_{(k,l)\in S}$ of $\mathbb{C}^V$, where $S := \{(k, l) : k \in [\![r]\!] \text{ and } l \in [\![\gamma_k]\!]\}$, such that

$$\forall\ (k, l) \in S, \qquad P[\varphi_{(k,l)}] \quad = \quad \theta_k\varphi_{(k,l)} + \varphi_{(k,l-1)} \tag{21}$$

where by convention, $\varphi_{(k,0)} = 0$ for all $k \in [\![r]\!]$.

But for our purpose, it is more advantageous to work with real functions. So let us decompose

$$S \quad = \quad S_\mathrm{r} \sqcup S_\mathrm{i}$$

with

$$\begin{aligned} S_\mathrm{r} &:= \{(k, l) \in S : \theta_k \in \mathbb{R}\} \\ S_\mathrm{i} &:= \{(k, l) \in S : \theta_k \notin \mathbb{R}\} \end{aligned}$$

There exists an involution of S_i, denoted $S_\mathrm{i} \ni (k, l) \mapsto (\bar{k}, \bar{l})$ such that

$$\forall\ (k, l) \in S_\mathrm{i}, \qquad \theta_{\bar{k}} = \bar{\theta}_k \quad \text{and} \quad \gamma_{\bar{k}} = \gamma_k$$

Let $R_\mathrm{i} \subset S_\mathrm{i}$ be such that $R_\mathrm{i} \ni (k, l) \mapsto (\bar{k}, \bar{l}) \in S_\mathrm{i}\backslash R_\mathrm{i}$ is a bijection. Consider $C_\mathrm{i} := R_i \times \{0, 1\}$ and define

$$C \quad := \quad S_\mathrm{r} \sqcup C_\mathrm{i}$$

We can find a basis $(\psi_c)_{c\in C}$ of $\mathbb{R}^V$ such that

$$\forall\ (k, l) \in S_\mathrm{r}, \qquad P[\psi_{(k,l)}] \quad = \quad \theta_k\psi_{(k,l)} + \psi_{(k,l-1)} \tag{22}$$

(again with the convention $\psi_{(k,0)} = 0$ for all $(k,1) \in S_{\mathrm{r}}$), and

$$\forall\ (k,l) \in R_{\mathrm{i}}, \qquad \begin{cases} P[\psi_{(k,l,0)}] &= \theta_{k,\mathrm{r}}\psi_{(k,l,0)} - \theta_{k,\mathrm{i}}\psi_{(k,l,1)} + \psi_{(k,l-1,0)} \\ P[\psi_{(k,l,1)}] &= \theta_{k,\mathrm{i}}\psi_{(k,l,0)} + \theta_{k,\mathrm{r}}\psi_{(k,l,1)} + \psi_{(k,l-1,1)} \end{cases} \tag{23}$$

where $\theta_{k,\mathrm{r}}$ and $\theta_{k,\mathrm{i}}$ are respectively the real and imaginary parts of θ_k (and $\psi_{(k,0,0)} = \psi_{(k,0,1)} = 0$ for all $(k,1) \in R_{\mathrm{i}}$).

Note that (23) is equivalent to

$$\forall\ (k,l) \in R_{\mathrm{i}}, \qquad P[\psi_{(k,l)}] \ = \ \theta_k\psi_{(k,l)} + \psi_{(k,l-1)} \tag{24}$$

where $\psi_{(k,l)} \in \mathbb{C}^V$ is given by

$$\forall\ (k,l) \in R_{\mathrm{i}}, \qquad \psi_{(k,l)} \ \coloneqq \ \psi_{(k,l,0)} + \mathrm{i}\,\psi_{(k,l,1)} \tag{25}$$

Observe that the conjugate functions $\bar{\psi}_{(k,l)}$, for $(k,l) \in R_{\mathrm{i}}$, play the same role for $\bar{\theta}_k$: $P[\bar{\psi}_{(k,l)}] = \bar{\theta}_k\bar{\psi}_{(k,l)} + \bar{\psi}_{(k,l-1)}$. An example of basis $(\varphi_{(k,l)})_{(k,l)\in S}$ satisfying (21) is given by

$$\forall\ (k,l) \in S, \qquad \varphi_{(k,l)} \ \coloneqq \ \begin{cases} \psi_{(k,l)} & \text{, if } (k,l) \in S_{\mathrm{r}} \sqcup R_{\mathrm{i}} \\ \bar{\psi}_{(\bar{k},\bar{l})} & \text{, if } (k,l) \in S_{\mathrm{i}} \backslash R_{\mathrm{i}} \end{cases}$$

Such a basis $(\psi_c)_{c\in C}$ will be said to be **adapted** to P.

These linear algebra considerations are valid for any real matrix P, let us now specify what can be said in addition for irreducible transition matrices. By irreducibility of P, 1 is an eigenvalue of multiplicity 1, so we can assume that $(\theta_1, \gamma_1) = (1,1)$ and $\psi_{(1,1)} = \mathbb{1}$. The irreducibility assumption also implies there is a unique invariant probability π for P and it gives a positive weight to every point of V. It can be assumed that all the eigenvectors ψ_c, for $c \in C$ are normalized in $\mathbb{L}^2(\pi)$, but in general they will not be orthogonal. The only orthogonality property is that of $\psi_{(1,1)}$ with the ψ_c, for $c \in C\backslash\{(1,1)\}$, namely

$$\forall\ c \in C\backslash\{(1,1)\}, \qquad \pi[\psi_c] = 0 \tag{26}$$

Indeed, for any $k \in [\![2,r]\!]$ such that $(k,1) \in S_{\mathrm{r}} \sqcup R_{\mathrm{i}}$, we have $P[\psi_{(k,1)}] = \theta_k\psi_{(k,1)}$ with $\theta_k \neq 1$ (where $\psi_{(k,l)}$ is given by (25) for $(k,1) \in R_{\mathrm{i}}$). Integrating the previous relation with respect to π, we obtain due to the invariance of π,

$$\pi[\psi_{(k,1)}] \ = \ \theta_k\pi[\psi_{(k,1)}]$$

so that $\pi[\psi_{(k,1)}] = 0$ (for $(k,1) \in R_{\mathrm{i}}$, this equality means that both $\pi[\psi_{(k,l,0)}] = 0$ and $\pi[\psi_{(k,l,1)}] = 0$). Next we show that

$$\pi[\psi_{(k,l)}] \ = \ 0 \tag{27}$$

by iteration on l, where $k \in [\![2,r]\!]$ is fixed as above. If (27) is true for some $l \in [\![\gamma_k - 1]\!]$, then integrating with respect to π the relation

$$P[\psi_{(k,l+1)}] \ = \ \theta_k\psi_{(k,l+1)} + \psi_{(k,l)} \tag{28}$$

we get $(1-\theta_k)\pi[\psi_{(k,l+1)}] = 0$, namely (27) with l replaced by $l+1$.

Let $(\widetilde{\psi}_c)_{c\in C}$ be a basis adapted to $\widetilde{P}$, with the same characteristic set $\{(\theta_k,\gamma_k) : k \in [\![r]\!]\}$ as P and the same index set C.

The next step is to construct the analogue of the operator A_b given in (9), for any family $b \coloneqq (b_c)_{c\in C\backslash\{(1,1)\}}$ of real numbers. We cannot proceed directly as in (9), i.e. define $A_b[\widetilde{\psi}_c] \coloneqq b_c\psi_c$ for all $c \in C\backslash\{(1,1)\}$, because it would not be compatible with the commutation relation

$$PA_b \ = \ A_b\widetilde{P} \tag{29}$$

Indeed applying the latter relation to $\widetilde{\psi}_{(k,l)}$ with $(k,l) \in S_{\mathrm{r}} \sqcup R_{\mathrm{i}} \backslash \{(1,1)\}$, we should have the equality

$$b_{(k,l)} P[\psi_{(k,l)}] \quad = \quad A_b[\theta_k \widetilde{\psi}_{(k,l)} + \widetilde{\psi}_{(k,l-1)}]$$

namely

$$b_{(k,l)}(\theta_k \psi_{(k,l)} + \psi_{(k,l-1)}) \quad = \quad b_{(k,l)} \theta_k \psi_{(k,l)} + b_{(k,l-1)} \psi_{(k,l-1)}$$

which implies that $b_{(k,l)} = b_{(k,l-1)}$ as soon as $l \geqslant 2$. But this equality leads to restrictive choices of b which are not in line with our purpose.

Fix the family $b \coloneqq (b_c)_{c \in C \backslash \{(1,1)\}}$ of real numbers.

We start by constructing A_b on the vector space generated by $(\widetilde{\psi}_{(k,l)})_{(k,l) \in S_{\mathrm{r}} \backslash \{(1,1)\}}$. Define for any $(k,l) \in S_{\mathrm{r}} \backslash \{(1,1)\}$,

$$A_b[\widetilde{\psi}_{(k,l)}] \quad \coloneqq \quad \sum_{j \in [\![l]\!]} b_{(k,l-j+1)} \psi_{(k,j)} \tag{30}$$

On the vector space generated by $(\widetilde{\psi}_c)_{c \in C_{\mathrm{i}}}$, we have to be more careful. By analogy with (30), we would like to define for any $(k,l) \in R_{\mathrm{i}}$,

$$A_b[\widetilde{\psi}_{(k,l)}] \quad \coloneqq \quad \sum_{j \in [\![l]\!]} b_{(k,l-j+1)} \psi_{(k,j)}$$

with $\psi_{(k,l)}$ is given by (25) and where

$$\forall\ j \in [\![\gamma_k]\!], \qquad b_{(k,j)} \quad \coloneqq \quad b_{(k,j,0)} + \mathrm{i}\, b_{(k,j,1)} \tag{31}$$

In "real" terms, this amounts to taking for any $(k,l) \in R_{\mathrm{i}}$

$$\begin{aligned} A_b[\widetilde{\psi}_{(k,l,0)}] \quad &\coloneqq \quad \sum_{j \in [\![l]\!]} b_{(k,l-j+1,0)} \psi_{(k,j,0)} - b_{(k,l-j+1,1)} \psi_{(k,j,1)} \\ A_b[\widetilde{\psi}_{(k,l,1)}] \quad &\coloneqq \quad \sum_{j \in [\![l]\!]} b_{(k,l-j+1,1)} \psi_{(k,j,0)} + b_{(k,l-j+1,0)} \psi_{(k,j,1)} \end{aligned}$$

It remains to define A_b on $\psi_{(1,1)}$. We just take $A_b[\psi_{(1,1)}] = 0$.

Lemma 8 *The operator A_b constructed above satisfies* (29).

Proof

It is sufficient to check that

$$\forall\ c \in C, \qquad P A_b[\widetilde{\psi}_c] \quad = \quad A_b \widetilde{P}[\widetilde{\psi}_c] \tag{32}$$

For $c = (1,1)$, this equality holds since both sides vanish, due to the fact that $\mathbb{1} = \widetilde{P}[\mathbb{1}]$.

We consider next the case $e = (k,l) \in S_{\mathrm{r}} \backslash \{(1,1)\}$. We compute on one hand,

$$\begin{aligned} P A_b[\widetilde{\psi}_{k,l}] \quad &= \quad P\left[\sum_{j \in [\![l]\!]} b_{(k,l-j+1)} \psi_{(k,j)}\right] \\ &= \quad \sum_{j \in [\![l]\!]} b_{(k,l-j+1)} P[\psi_{(k,j)}] \\ &= \quad \sum_{j \in [\![l]\!]} b_{(k,l-j+1)} (\theta_k \psi_{(k,j)} + \psi_{(k,j-1)}) \end{aligned}$$

$$= \quad \theta_k \sum_{j\in[\![l]\!]} b_{(k,l-j+1)}\psi_{(k,j)} + \sum_{j\in[\![l-1]\!]} b_{(k,l-j)}\psi_{(k,j)} \tag{33}$$

and on the other hand,

$$\begin{aligned} A_b\widetilde{P}[\widetilde{\psi}_{(k,l)}] &= A_b[\theta_k\widetilde{\psi}_{(k,l)} + \widetilde{\psi}_{(k,l-1)}] \\ &= \theta_k \sum_{j\in[\![l]\!]} b_{(k,l-j+1)}\psi_{(k,j)} + \sum_{j\in[\![l-1]\!]} b_{k,l-1-j+1}\psi_{(k,j)} \end{aligned}$$

which coincides with (33).

Let us now check (32) for $c \in C_{\mathrm{i}}$. Note that the above computations are equally valid for $\psi_{(k,l)}$ defined in (25), with $(k,l) \in R_{\mathrm{i}}$. Namely, we have

$$PA_b[\widetilde{\psi}_{(k,l,0)} + \mathrm{i}\,\widetilde{\psi}_{(k,l,1)}] \quad = \quad A_b\widetilde{P}[\widetilde{\psi}_{(k,l,0)} + \mathrm{i}\,\widetilde{\psi}_{(k,l,1)}]$$

Since $PA_b[\widetilde{\psi}_{(k,l,0)}]$, $PA_b[\widetilde{\psi}_{(k,l,1)}]$, $A_b\widetilde{P}[\widetilde{\psi}_{(k,l,0)}$ and $A_b\widetilde{P}[\widetilde{\psi}_{(k,l,1)}]$ are vectors with real entries, we deduce

$$\begin{aligned} PA_b[\widetilde{\psi}_{(k,l,0)}] &= A_b\widetilde{P}[\widetilde{\psi}_{(k,l,0)}] \\ PA_b[\widetilde{\psi}_{(k,l,1)}] &= A_b\widetilde{P}[\widetilde{\psi}_{(k,l,1)}] \end{aligned}$$

These identities are valid for all $(k,l) \in R_{\mathrm{i}}$, so that (32) is satisfied for all $c \in C_{\mathrm{i}}$. ■

Writing A_b in the bases $(\psi_c)_{c\in C}$ and $(\widetilde{\psi}_c)_{c\in C}$, we see that

$$\lim_{b\to 0} A_b \quad = \quad 0 \tag{34}$$

Introduce

$$R \quad \coloneqq \quad (S_{\mathrm{r}} \sqcup R_{\mathrm{i}})\setminus\{(1,1)\} \tag{35}$$

Taking into account the definition (31), we can see $b \in R^{C\setminus\{(1,1)\}}$ as the element $(b_{(k,l)})_{(k,l)\in R} \in \mathbb{R}^{S_{\mathrm{r}}\setminus\{(0,0)\}} \times \mathbb{C}^{R_{\mathrm{i}}}$. From (34) we can find $\eta > 0$ so that for any $b \coloneqq (b_{(k,l)})_{(k,l)\in R}$,

$$\max\{|b_{(k,l)}| \,:\, (k,l) \in R\} \leqslant \eta \quad \Rightarrow \quad \max\{|A_b(x,y)|/\widetilde{\pi}(y) \,:\, x,y \in V\} \leqslant 1 \tag{36}$$

A more quantitative description of $\eta > 0$, in the spirit of Lemma 7, will be given in Lemma 11 at the end of this section, under the assumption that all the eigenvalues are real. It is not needed in the proofs of Proposition 4 and Theorem 2, which are rather qualitative as long as no bound is asked on the support of q (i.e. not only on its cardinal).

Introduce

$$\mathcal{B} \quad \coloneqq \quad \left\{ b \in \mathbb{R}^{S_{\mathrm{r}}\setminus\{(1,1)\}} \times \mathbb{C}^{R_{\mathrm{i}}} \,:\, \max\{|b_{(k,l)}| \,:\, (k,l) \in R\} \leqslant \eta \right\} \tag{37}$$

For $b \in \mathcal{B}$, we consider the operator Λ_b given as in (12) by

$$\Lambda_b \quad \coloneqq \quad \widetilde{\pi} + A_b \tag{38}$$

where again $\widetilde{\pi}$ is interpreted as the matrix whose rows are all equal to the probability $\widetilde{\pi}$.

Taking into account that $P\widetilde{\pi} = \widetilde{\pi}\widetilde{P} = \widetilde{\pi}$ and Lemma 8, we get the intertwining relation $P\Lambda_b = \Lambda_b\widetilde{P}$. From the relation $\Lambda_b[\mathbb{1}] = \Lambda_b[\widetilde{\psi}_{(1,1)}] = \psi_{(1,1)} = \mathbb{1}$, it appears that the row sums of Λ_b are all equal to 1. Furthermore, all the entries of Λ_b will be non-negative as soon as

$$\forall\; x,y \in V, \qquad \widetilde{\pi}(y) - |A_b(x,y)| \quad \geqslant \quad 0$$

which is satisfied by definition of $\mathcal{B}$, see (36) and (37).

Thus Λ_b is a Markov kernel for $b \in \mathcal{B}$. In general it is not invertible, for instance for $b = 0$. Introduce

$$\mathcal{C} \quad := \quad \left\{b \in \mathcal{B} \,:\, \min\{|b_{(k,1)}| \,:\, k \in K\} > 0\right\} \tag{39}$$

where

$$K \quad := \quad \{k \in [\![r]\!] \,:\, (k,1) \in R\} \tag{40}$$

Its interest is:

Lemma 9 *The operator Λ_b is invertible for $b \in \mathcal{C}$.*

Proof

Expressed in the bases $(\widetilde{\psi}_{(k,l)})_{(k,l)\in\{(1,1)\}\sqcup R}$ and $(\psi_{(k,l)})_{(k,l)\in\{(1,1)\}\sqcup R}$ the matrix of Λ_b is block-diagonal. The block-matrix associated to $(1,1)$ is just 1. For $k \in K$, the block matrix associated to the Jordan block (k, γ_k) is the Toeplitz matrix

$$T_k \quad := \quad \begin{pmatrix} b_{k,1} & b_{k,2} & b_{k,3} & \cdots & b_{k,\gamma_k} \\ 0 & b_{k,1} & b_{k,2} & \ddots & b_{k,\gamma_k-1} \\ 0 & 0 & b_{k,1} & \ddots & b_{k,\gamma_k-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & b_{k,1} \end{pmatrix} \tag{41}$$

whose entries are real numbers if $k \in S_{\mathrm{r}}$, but some of the entries may be complex numbers which are not real for $k \in R_{\mathrm{i}}$. Whatever the case, this matrix is invertible if and only if $b_{k,1} \neq 0$. Thus in view of its definition (39), $\mathcal{C}$ exactly consists of the elements $b \in \mathcal{B}$ such that Λ_b is invertible. ■

Since similar arguments are valid for $\widetilde{\Lambda}_{\widetilde{b}} := \pi + \widetilde{A}_{\widetilde{b}}$ with similar definitions, we get a strong bi-intertwining relation between P and $\widetilde{P}$, with Λ_b and $\widetilde{\Lambda}_{\widetilde{b}}$ as links, by choosing any $b \in \mathcal{C}$ and $\widetilde{b} \in \widetilde{\mathcal{C}}$. This shows the validity of the statement of Lemma 7 in [4], although its proof is erroneous.

With these preliminaries in hand, we can now come to the

Proof of Proposition 4

As for Proposition 6, we first compute both $\Lambda_b\widetilde{\Lambda}_{\widetilde{b}}$ and $Q(P)$ for given $b \in \mathcal{C}$, $\widetilde{b} \in \widetilde{\mathcal{C}}$ and probability q with a finite support on $\mathbb{Z}_+$, where Q is the associated polynomial defined in (17).

Expressed in the basis $(\psi_{(k,l)})_{(k,l)\in\{(1,1)\}\sqcup R}$ (and in the intermediate basis $(\widetilde{\psi}_{(k,l)})_{(k,l)\in\{(1,1)\}\sqcup R}$ for the product $\Lambda_b\widetilde{\Lambda}_{\widetilde{b}}$), both $\Lambda_b\widetilde{\Lambda}_{\widetilde{b}}$ and $Q(P)$ have a block diagonal structure.

Note there is no problem for the one-dimensional Jordan block associated to $\theta_1 = 1$: we have

$$\Lambda_b\widetilde{\Lambda}_{\widetilde{b}}[\mathbb{1}] \quad = \quad \mathbb{1} \quad = \quad Q(P)[\mathbb{1}]$$

whatever the choice of $b \in \mathcal{C}$, $\widetilde{b} \in \widetilde{\mathcal{C}}$ and of the probability q.

Let us now fix $k \in K$ and consider the block matrices associated to the Jordan block (k, γ_k).

• For $\Lambda_b\widetilde{\Lambda}_{\widetilde{b}}$, the $\gamma_k \times \gamma_k$-block is $T_k\widetilde{T}_k$, where $\widetilde{T}_k$ is defined as in (41), but with respect to $\widetilde{b}$. Note that $T_k\widetilde{T}_k$ is a upper diagonal Toeplitz matrix determined by its first row which is the vector

$$\left(\sum_{j\in[\![l]\!]} b_{(k,j)}\widetilde{b}_{(k,\gamma_k-j+1)}\right)_{l\in[\![\gamma_k]\!]} \tag{42}$$

- For any $n \in \mathbb{Z}_+$, the $\gamma_k \times \gamma_k$-block of P^n is

$$(\theta_k I_k + N_k)^n \;=\; \sum_{m\in[\![n-\gamma_k+1,n]\!]} \binom{n}{m} \theta_k^m N_k^{n-m} \tag{43}$$

where I_k is the $\gamma_k \times \gamma_k$ identity matrix and N_k is the matrix whose first upper diagonal consists of 1's and the other entries vanish (i.e. $\theta_k I_k + N_k$ is the usual $\gamma_k \times \gamma_k$ Jordan block associated to the eigenvalue θ_k). In (43), we took into account that $N_k^{\gamma_k} = 0$.

The matrix in (43) is also upper diagonal Toeplitz and is determined by its first row which is the vector

$$\left(\binom{n}{l-1}\theta_k^{n-l+1}\right)_{l\in[\![\gamma_k]\!]} \tag{44}$$

(for $n = 0$, by convention this vector is $(1, 0, 0, \cdots, 0)$).

Thus (3) is satisfied with q the Dirac mass at $n \in \mathbb{Z}_+$, if and only if (42) and (44) coincide. We get the system of equations in $(b_{(k,l)})_{l\in[\![\gamma_k]\!]}$ and $(\widetilde{b}_{(k,l)})_{l\in[\![\gamma_k]\!]}$:

$$\left\{\begin{array}{rcl} \widetilde{b}_{(k,1)}b_{(k,1)} & = & \theta_k^n \\ \widetilde{b}_{(k,1)}b_{(k,2)} + \widetilde{b}_{(k,2)}b_{(k,1)} & = & n\theta_k^{n-1} \\ \widetilde{b}_{(k,1)}b_{(k,3)} + \widetilde{b}_{(k,2)}b_{(k,2)} + \widetilde{b}_{(k,3)}b_{(k,1)} & = & \frac{n(n-1)}{2}\theta_k^{n-2} \\ & \vdots & \end{array}\right. \tag{45}$$

Let us consider the case where $|\theta_k| \in (0, 1)$. Introduce the polar decomposition $\theta_k = \rho_k e^{\mathrm{i}\alpha_k}$, with $\rho_k \in (0, 1)$ and $\alpha_k \in [0, 2\pi)$. We look for a solution of the form $b_{(k,l)} = \rho_k^{n/2-l+1}\beta_{(k,l)}$ and $\widetilde{b}_{(k,l)} = \rho_k^{n/2-l+1}$ for $l \in [\![\gamma_k]\!]$, so we get the system of equations in $(\beta_{(k,l)})_{l\in[\![\gamma_k]\!]}$:

$$\left\{\begin{array}{rcl} \beta_{(k,1)} & = & e^{\mathrm{i}n\alpha_k} \\ \beta_{(k,2)} + \beta_{(k,1)} & = & ne^{\mathrm{i}(n-1)\alpha_k} \\ \beta_{(k,3)} + \beta_{(k,2)} + \beta_{(k,1)} & = & \frac{n(n-1)}{2}e^{\mathrm{i}(n-2)\alpha_k} \\ & \vdots & \end{array}\right. \tag{46}$$

which admits a unique solution. Note that if $k \in S_{\mathrm{r}}$, then $\alpha_k \in \{0, \pi\}$ and the solution $(\beta_{(k,l)})_{l\in[\![\gamma_k]\!]}$ is real valued. Otherwise for $k \in S_{\mathrm{i}}$, recall that $(\beta_{(k,l)})_{l\in[\![\gamma_k]\!]}$ has to be decomposed as in (31) to provide for the desired real coefficients $(\beta_{(k,l,0)})_{l\in[\![\gamma_k]\!]}$ and $(\beta_{(k,l,1)})_{l\in[\![\gamma_k]\!]}$. Furthermore an immediate iteration proves that

$$\forall\ l \in [\![\gamma_k]\!], \qquad |\beta_{(k,l)}| \;\leqslant\; \sum_{j\in[\![l-1]\!]} \binom{n}{j}$$

and it follows that

$$\max\{|b_{(k,l)}| \vee |\widetilde{b}_{(k,l)}| \,:\, l \in [\![\gamma_k]\!]\} \;\leqslant\; \sum_{j\in[\![\gamma_k-1]\!]} \binom{n}{j} \rho_k^{n/2-\gamma_k+1} \tag{47}$$

Note that the r.h.s. goes to zero as n goes to infinity. We can thus find $n_0(k) \in \mathbb{Z}_+$ large enough so that for $n \geqslant n_0(k)$ we have

$$\max\{|b_{(k,l)}| \vee |\widetilde{b}_{(k,l)}| \,:\, l \in [\![\gamma_k]\!]\} \;\leqslant\; \eta$$

Note furthermore that $b_{(k,1)} \neq 0$ and $\widetilde{b}_{(k,1)} \neq 0$.

It follows that if the eigenvalues θ_k, for $k \in K$ (or equivalently for $k \in [\![2, r]\!]$), do not vanish and have modulus strictly less than one, then we can find $b \in \mathcal{C}$ and $\widetilde{b} \in \widetilde{\mathcal{C}}$ so that $\Lambda_b \widetilde{\Lambda}_{\widetilde{b}} = P^{n_0}$ with $n_0 := \max\{n_0(k) : k \in K\}$. This exactly corresponds to the situation where P is aperiodic and does not admit zero as eigenvalue. Thus the last assertion of the proposition is shown.

Concerning the last-but-one (and not deducing it from the first one, to be more pedagogical), when P is aperiodic, we still have that all the eigenvalues, except $\theta_1 = 1$, have a modulus strictly smaller than 1, but some of the eigenvalues can be zero. Consider $k_0 \in K$ such that $\theta_{k_0} = 0$. Solving (45), we end up with either $b_{(k_0,1)} = 0$ or $\widetilde{b}_{(k_0,1)} = 0$, which is not convenient for our purpose. So as in the proof of Proposition 6, we rather look for q of the form $\zeta^n \delta_0 + (1 - \zeta^n)\delta_n$, where we take again

$$\zeta \quad := \quad \max\{|\theta_k| \,:\, k \in [\![2, r]\!]\} \tag{48}$$

Then for any $k \in K$, (45) transforms into

$$\left\{\begin{array}{rcl} \widetilde{b}_{(k,1)} b_{(k,1)} & = & \zeta^n + (1-\zeta^n)\theta_k^n \\ \widetilde{b}_{(k,1)} b_{(k,2)} + \widetilde{b}_{(k,2)} b_{(k,1)} & = & (1-\zeta^n) n \theta_k^{n-1} \\ \widetilde{b}_{(k,1)} b_{(k,3)} + \widetilde{b}_{(k,2)} b_{(k,2)} + \widetilde{b}_{(k,3)} b_{(k,1)} & = & (1-\zeta^n)\frac{n(n-1)}{2}\theta_k^{n-2} \\ & \vdots & \end{array}\right.$$

This system can be solved as before, in particular with

$$\begin{array}{rcl} \widetilde{b}_{(k,1)} & = & |\zeta^n + (1-\zeta^n)\theta_k^n|^{1/(2n)} \;\neq\; 0 \\ b_{(k,1)} & = & \widetilde{b}_{(k,1)} \dfrac{\zeta^n + (1-\zeta^n)\theta_k^n}{|\zeta^n + (1-\zeta^n)\theta_k^n|} \;\neq\; 0 \end{array}$$

and similarly to (47) we get

$$\max\{|b_{(k,l)}| \vee |\widetilde{b}_{(k,l)}| \,:\, l \in [\![\gamma_k]\!]\} \quad \leqslant \quad \sum_{j \in [\![\gamma_k - 1]\!]} \binom{n}{j} \zeta^{n/2 - \gamma_k + 1} \tag{49}$$

Since the r.h.s. converges to zero as n goes to infinity, we end up with the conclusion that we can find $b \in \mathcal{C}$, $\widetilde{b} \in \widetilde{\mathcal{C}}$ and essentially the same n_0 as above so that $\Lambda_b \widetilde{\Lambda}_{\widetilde{b}} = \zeta^{n_0} + (1 - \zeta^{n_0}) P^{n_0}$.

We now come to the case where some of the eigenvalues of P outside θ_1 have modulus 1. It is well-known that for a irreducible transition matrix, there exists $d \in \mathbb{N}$ called the period such that the eigenvalues of modulus 1 are of the form $e^{\mathrm{i}2\pi m/d}$ for $m \in [\![0, d-1]\!]$ and each of the latter have geometric multiplicity 1.

In this situation, we consider the probability

$$q \quad := \quad \zeta^{n_0}\delta_0 + (1-\zeta^{n_0})\frac{\delta_{n_0} + \delta_{n_0+1} + \cdots + \delta_{n_0+d-1}}{d}$$

where now

$$\zeta \quad := \quad \max\{|\theta_k| \,:\, k \in \mathcal{K}\} \tag{50}$$

with

$$\mathcal{K} \quad := \quad \{k \in [\![2, r]\!] \text{ and } |\theta_k| < 1\} \tag{51}$$

and as above

$$n_0 \quad := \quad \min\left\{ n \in \mathbb{Z}_+ \,:\, \forall\, k \in \mathcal{K}, \sum_{j \in [\![\gamma_k - 1]\!]} \binom{n}{j} |\theta_k|^{n/2 - \gamma_k + 1} \leqslant \eta \right\} \tag{52}$$

$$\leqslant \quad \min\left\{n\in\mathbb{Z}_+ \,:\, \Gamma\binom{n}{\Gamma-1}\zeta^{n/2-\Gamma+1}\leqslant \eta\right\} \tag{53}$$

where $\Gamma := \max\{\gamma_k \,:\, k\in[\![r]\!]\}$ is the largest dimension of the Jordan blocks.

Our goal is to find $b\in\mathcal{C}$ and $\widetilde{b}\in\widetilde{\mathcal{C}}$ such that for any $k\in K$,

$$\left\{\begin{array}{rcl} \widetilde{b}_{(k,1)}b_{(k,1)} &=& \zeta^{n_0}+(1-\zeta^{n_0})\theta_k^{n_0}\dfrac{1+\theta_k+\cdots+\theta_k^{d-1}}{d}\\ \widetilde{b}_{(k,1)}b_{(k,2)}+\widetilde{b}_{(k,2)}b_{(k,1)} &=& (1-\zeta^{n_0})n_0\theta_k^{n_0-1}\dfrac{1+\theta_k+\cdots+\theta_k^{d-1}}{d}\\ \widetilde{b}_{(k,1)}b_{(k,3)}+\widetilde{b}_{(k,2)}b_{(k,2)}+\widetilde{b}_{(k,3)}b_{(k,1)} &=& (1-\zeta^{n_0})\dfrac{n_0(n_0-1)}{2}\theta_k^{n_0-2}\dfrac{1+\theta_k+\cdots+\theta_k^{d-1}}{d}\\ &\vdots& \end{array}\right. \tag{54}$$

Note that if $k\in K$ is such that θ_k is an eigenvalue of modulus equal to 1, i.e. of the form $e^{\mathrm{i}2\pi m/d}$ with $m\in[\![d-1]\!]$ (recall that $1\notin K$, so that $m=0$ is not permitted), then

$$\begin{array}{rcl} 1+\theta_k+\cdots+\theta_k^{d-1} &=& \dfrac{1-e^{\mathrm{i}2\pi m}}{1-e^{\mathrm{i}2\pi m/d}}\\ &=& 0 \end{array}$$

so that the above system (54) reduces to

$$\left\{\begin{array}{rcl} \widetilde{b}_{(k,1)}b_{(k,1)} &=& \zeta^{n_0}\\ \widetilde{b}_{(k,1)}b_{(k,2)}+\widetilde{b}_{(k,2)}b_{(k,1)} &=& 0\\ \widetilde{b}_{(k,1)}b_{(k,3)}+\widetilde{b}_{(k,2)}b_{(k,2)}+\widetilde{b}_{(k,3)}b_{(k,1)} &=& 0\\ &\vdots& \end{array}\right.$$

which can be solved by taking $\widetilde{b}_{(k,1)}=b_{(k,1)}=\zeta^{n_0/2}$ and $\widetilde{b}_{(k,l)}=b_{(k,l)}=0$ for all $l\in[\![2,\gamma_k]\!]$.

For the $k\in K$ such that $|\theta_k|<1$, we can proceed as before, taking into account that

$$\left|\frac{1+\theta_k+\cdots+\theta_k^{d-1}}{d}\right| \quad\leqslant\quad 1$$

to construct $b\in\mathcal{C}$ and $\widetilde{b}\in\widetilde{\mathcal{C}}$ solving (54) and such that (49) holds (with (50) instead of (48)).

To sum up, we have constructed a strong bi-intertwining relation between P and $\widetilde{P}$ with a corresponding interweaving relation from P to $\widetilde{P}$, so we get a strong bi-interweaving relation between P and $\widetilde{P}$ with $\widetilde{q}=q$ as above according to Remark 1. ■

We can now proceed to the

Proof of Theorem 2

The reverse implication is obvious: assume that $\sigma\in\mathcal{S}_\ell$, the probability q on $\mathbb{Z}_+$ and the invertible links Λ_l (from C_l to $\widetilde{C}_{\sigma(l)}$) and $\widetilde{\Lambda}_l$ (from $\widetilde{C}_{\sigma(l)}$ to C_l), for $l\in[\![\ell]\!]$, are such that for any $l\in[\![\ell]\!]$, we have

$$\left\{\begin{array}{rcl} P_{C_l}\Lambda_l &=& \Lambda_l\widetilde{P}_{\widetilde{C}_{\sigma(l)}}\\ \widetilde{P}_{\widetilde{C}_{\sigma(l)}}\widetilde{\Lambda}_l &=& \widetilde{\Lambda}_lP_{C_l}\\ \Lambda_l\widetilde{\Lambda}_l &=& \sum_{n\in\mathbb{Z}_+}q_nP_{C_l}^n \end{array}\right. \tag{55}$$

(the corresponding relation (4) with $\widetilde{q}=q$ is a consequence of Remark 1).

Consider Σ a permutation of V such that $\Sigma(C_l) = \widetilde{C}_{\sigma(l)}$ for all $l \in [\![\ell]\!]$. Identify Σ with its $V \times V$ matrix $(\mathbb{1}_{y=\sigma(x)})_{(x,y)\in V\times V}$. Replacing $\widetilde{P}$ by $\Sigma\widetilde{P}\Sigma^{-1}$ (which amounts to "rename" the elements of V for $\widetilde{P}$), we can assume that $C_l = \widetilde{C}_{\sigma(l)}$ for all $l \in [\![\ell]\!]$. Ordering appropriately the elements of V, we have

$$P = \begin{pmatrix} P_{C_1} & 0 & \cdots & 0 \\ 0 & P_{C_2} & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & P_{C_\ell} \end{pmatrix} \quad \text{and} \quad \widetilde{P} = \begin{pmatrix} \widetilde{P}_{C_1} & 0 & \cdots & 0 \\ 0 & \widetilde{P}_{C_2} & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \widetilde{P}_{C_\ell} \end{pmatrix} \tag{56}$$

It remains to define the invertible links

$$\Lambda := \begin{pmatrix} \Lambda_1 & 0 & \cdots & 0 \\ 0 & \Lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \Lambda_\ell \end{pmatrix} \quad \text{and} \quad \widetilde{\Lambda} := \begin{pmatrix} \widetilde{\Lambda}_1 & 0 & \cdots & 0 \\ 0 & \widetilde{\Lambda}_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \widetilde{\Lambda}_\ell \end{pmatrix} \tag{57}$$

to get a strong bi-interweaving relation between P and $\widetilde{P}$ associated to the probability $\widetilde{q} = q$.

Conversely assume a strong bi-interweaving relation between P and $\widetilde{P}$ holds with respect to some invertible links Λ, $\widetilde{\Lambda}$ and to the probability $\widetilde{q} = q$.

Denote E and $\widetilde{E}$ the eigenspaces associated to the eigenvalue 1 respectively for P and $\widetilde{P}$. From the intertwining relation $P\Lambda = \Lambda\widetilde{P}$, we deduce that $\Lambda(\widetilde{E}) \subset E$ and in fact $\Lambda(\widetilde{E}) = E$ since Λ is invertible and $\dim(E) = \dim(\widetilde{E})$ by similarity of P and $\widetilde{P}$. As a vector space, E (resp. $\widetilde{E}$) is generated by the indicator functions $\mathbb{1}_{C_l}$ (resp. $\mathbb{1}_{\widetilde{C}_l}$), for $l \in [\![\ell]\!]$. Thus there exists matrices $M := (M_{k,l})_{k,l\in[\![\ell]\!]}$ and $\widetilde{M} := (\widetilde{M}_{k,l})_{k,l\in[\![\ell]\!]}$ so that for any $l \in [\![\ell]\!]$,

$$\begin{aligned} \Lambda[\mathbb{1}_{\widetilde{C}_l}] &= \sum_{k\in[\![\ell]\!]} M_{k,l}\mathbb{1}_{C_k} \\ \widetilde{\Lambda}[\mathbb{1}_{C_l}] &= \sum_{k\in[\![\ell]\!]} \widetilde{M}_{k,l}\mathbb{1}_{\widetilde{C}_k} \end{aligned}$$

From the fact that Λ and $\widetilde{\Lambda}$ are Markov matrices, we deduce that M and $\widetilde{M}$ are Markov matrices too. We also get for any $l \in [\![\ell]\!]$,

$$\Lambda\widetilde{\Lambda}[\mathbb{1}_{C_l}] = \sum_{k\in[\![\ell]\!]} (M\widetilde{M})_{k,l}\mathbb{1}_{C_k}$$

but from the interweaving relation, we have

$$\begin{aligned} \Lambda\widetilde{\Lambda}[\mathbb{1}_{C_l}] &= \sum_{n\in\mathbb{Z}_+} q_n P^n[\mathbb{1}_{C_l}] \\ &= \sum_{n\in\mathbb{Z}_+} q_n \mathbb{1}_{C_l} \\ &= \mathbb{1}_{C_l} \end{aligned}$$

We deduce that $M\widetilde{M}$ is the identity matrix. Since both M and $\widetilde{M}$ are Markov matrices, this is only possible, see Lemma 10 below, if there exists a permutation $\sigma \in \mathcal{S}_\ell$ such that M and $\widetilde{M}$ are the matrices respectively associated to σ and σ^{-1}:

$$\forall\ k, l \in [\![\ell]\!], \qquad \begin{cases} M_{k,l} = \mathbb{1}_{l=\sigma(k)} \\ \widetilde{M}_{k,l} = \mathbb{1}_{k=\sigma(l)} \end{cases} \tag{58}$$

For any $l \in [\![\ell]\!]$, the relation $\widetilde{\Lambda}[\mathbb{1}_{C_l}] = \mathbb{1}_{\widetilde{C}_{\sigma(l)}}$ and the invertibility of Λ, imply $|C_l| = |\widetilde{C}_{\sigma(l)}|$. Define Λ_l the $C_l \times \widetilde{C}_{\sigma(l)}$ restriction of Λ, which is a Markov transition matrix from C_l to $\widetilde{C}_{\sigma(l)}$. Similarly, let $\widetilde{\Lambda}_l$ be the $\widetilde{C}_{\sigma(l)} \times C_l$ restriction of $\widetilde{\Lambda}$. Up to the renaming transformations considered in the first part of this proof, we can assume that for any $l \in [\![\ell]\!]$, $\widetilde{C}_{\sigma(l)} = C_l$ and that both (56) and (57) hold. Expressing the bi-intertwining relation between P and $\widetilde{P}$ in this block-diagonal matrix form, we get the validity of (55) (with $\widetilde{C}_{\sigma(l)}$ replaced by C_l), which is the desired result.

The last assertion of Theorem 2 comes from the constructions of the probabilities q in the irreducible case. They can be made compatible for the P_{C_l} and $\widetilde{P}_{\widetilde{C}_{\sigma(l)}}$, for $l \in [\![\ell]\!]$, by considering a probability $q = \epsilon\delta_0 + (1-\epsilon)\mathcal{U}_{[\![n,n+d-1]\!]}$, where $\mathcal{U}_{[\![n,n+d-1]\!]}$ is the uniform distribution on $[\![n, n+d-1]\!]$, with $\epsilon \in (0,1)$ small enough, $n \in \mathbb{Z}_+$ large enough, and d the least common multiple of the periods of the P_{C_l}. ■

In the above proof we needed the following well-known result, given for completeness.

Lemma 10 *Assume that M and $\widetilde{M}$ are two Markov matrices on $[\![\ell]\!]$ such that $\widetilde{M}$ is the inverse of M. Then there exists a permutation $\sigma \in \mathcal{S}_\ell$ of the state space such that* (10) *holds.*

Proof

It is sufficient to show that for any $k \in [\![\ell]\!]$, there exist a unique $l \in [\![\ell]\!]$ such that $M(k,l) > 0$. Indeed, then we have $M(k,l) = 1$ and we define $\sigma(k) \coloneqq l$. The mapping σ constructed in this way is necessarily a permutation, otherwise M would not be invertible.

So by contradiction, assume there exist $k \in [\![\ell]\!]$ as well as $l_1 \neq l_2 \in [\![\ell]\!]$ with $M(k,l_1) > 0$ and $M(k,l_2) > 0$. Since

$$\sum_{l \in [\![\ell]\!]} M(k,l)\widetilde{M}(l,k) \quad = \quad 1$$

we deduce that we must have $\widetilde{M}(l_1,k) = 1 = \widetilde{M}(l_2,k) = 1$, otherwise the sum in the l.h.s. would be strictly less than 1. It follows that the row $\widetilde{M}(l_1,\cdot)$ and $\widetilde{M}(l_2,\cdot)$ are the Dirac mass at k and in particular we have $\widetilde{M}(l_1,\cdot) = \widetilde{M}(l_2,\cdot)$, in contradiction with the fact that $\widetilde{M}$ is invertible. ■

As promised after (36), let us present an estimate on the quantity η introduced there, under the assumption that all the eigenvalues are real. An investigation of the general case should be possible in a similar fashion, but we refrain from entering the corresponding more involved calculations. Indeed, they will not serve as an inspiring guide in Section 5, where only non-negative eigenvalues will be considered. Nevertheless, at the end of this section we will deduce an example of bounds that can given on the support of q in Proposition 4, namely on the warming-up time to pass from P to $\widetilde{P}$ and conversely, when all the eigenvalues are assumed to be real.

We need the Gramian matrices

$$R \;\coloneqq\; (\pi[\varphi_{(k,l)}\varphi_{(k',l')}])_{(k,l),(k',l')\in S} \quad \text{and} \quad \widetilde{R} \;\coloneqq\; (\widetilde{\pi}[\widetilde{\varphi}_{(k,l)}\widetilde{\varphi}_{(k',l')}])_{(k,l),(k',l')\in S}$$

where $(\varphi_{(k,l)})_{(k,l)\in S}$ and $(\widetilde{\varphi}_{(k,l)})_{(k,l)\in S}$ are bases adapted to the spectral structure of P and $\widetilde{P}$ respectively.

These matrices are positive definite. Let $\lambda_\vee \geqslant \lambda_\wedge > 0$ (respectively $\widetilde{\lambda}_\vee \geqslant \widetilde{\lambda}_\wedge > 0$) be the largest and the smallest eigenvalues of R (resp. $\widetilde{R}$).

Their interest comes from the following analogue of Lemma 7:

Lemma 11 *We have for any $x, y \in V$,*

$$\left|\frac{A_b(x,y)}{\widetilde{\pi}(y)}\right| \;\leqslant\; \Gamma\sqrt{\frac{\lambda_\vee}{\widetilde{\lambda}_\wedge}}\frac{1}{\sqrt{\pi(x)\widetilde{\pi}(y)}}\max\{|b_{(k,l)}| \,:\, (k,l) \in S_0\}$$

$$\left|\frac{\widetilde{A}_{\widetilde{b}}(x,y)}{\pi(y)}\right| \leqslant \Gamma\sqrt{\frac{\widetilde{\lambda}_\vee}{\lambda_\wedge}}\frac{1}{\sqrt{\pi(x)\widetilde{\pi}(y)}}\max\{|\widetilde{b}_{(k,l)}| \,:\, (k,l)\in S_0\}$$

where $S_0 \coloneqq S\backslash\{(1,1)\}$.

In particular, in (36) *we can take*

$$\eta \coloneqq \frac{1}{\Gamma}\sqrt{\frac{\widetilde{\lambda}_\wedge}{\lambda_\vee}}\sqrt{\pi_\wedge\widetilde{\pi}_\wedge}$$

Proof

We adapt the proof of Lemma 7. The entries of the matrix associated to A_b are given by

$$\forall\ x,y\in V,\qquad A_b(x,y) = \langle \mathbb{1}_x, A_b[\mathbb{1}_y]\rangle$$

where $\langle\cdot,\cdot\rangle$ is the usual scalar product in $\mathbb{R}^V$ (recall that $\mathbb{1}_x$ and $\mathbb{1}_y$ are the indicators function of x and y). Using integration with respect to π, this can be written

$$\forall\ x,y\in V,\qquad A_b(x,y) = \pi\left[\frac{\mathbb{1}_x}{\pi(x)}A_b[\mathbb{1}_y]\right]$$

or equivalently

$$\forall\ x,y\in V,\qquad \frac{A_b(x,y)}{\widetilde{\pi}(y)} = \pi\left[\frac{\mathbb{1}_x}{\pi(x)}A_b\left[\frac{\mathbb{1}_y}{\widetilde{\pi}(y)}\right]\right]$$

Introduce the following decompositions in the bases $(\widetilde{\varphi}_{(k,l)})_{(k,l)\in S}$ and $(\varphi_{(k,l)})_{(k,l)\in S}$:

$$\frac{\mathbb{1}_x}{\pi(x)}(\cdot) = \sum_{(k,l)\in S}\alpha_{(k,l)}(x)\varphi_{(k,l)}(\cdot) \tag{59}$$

$$\frac{\mathbb{1}_y}{\widetilde{\pi}(y)}(\cdot) = \sum_{(k,l)\in S}\widetilde{\alpha}_{(k,l)}(y)\widetilde{\varphi}_{(k,l)}(\cdot)$$

with some real coefficients $\alpha(x)\coloneqq(\alpha_{(k,l)}(x))_{(k,l)\in S}$ and $\widetilde{\alpha}(y)\coloneqq(\widetilde{\alpha}_{(k,l)}(y))_{(k,l)\in S}$.

We deduce

$$\begin{aligned}
\forall\ x,y\in V,\qquad \frac{A_b(x,y)}{\widetilde{\pi}(y)} &= \sum_{(k,l),(k',l')\in S}\alpha_{(k,l)}(x)\widetilde{\alpha}_{(k',l')}(y)\pi\left[\varphi_{(k,l)}A_b[\widetilde{\varphi}_{(k',l')}]\right]\\
&= \sum_{(k,l),(k',l')\in S_0}\alpha_{(k,l)}(x)\widetilde{\alpha}_{(k',l')}(y)\pi\left[\varphi_{(k,l)}A_b[\widetilde{\varphi}_{(k',l')}]\right]\\
&= \sum_{(k,l),(k',l')\in S_0}\alpha_{(k,l)}(x)\widetilde{\alpha}_{(k',l')}(y)\sum_{j\in[\![l']\!]}b_{(k',l'-j+1)}R_{(k,l),(k',j)}\\
&= \sum_{(k,l),(k',j)\in S_0}\alpha_{(k,l)}(x)\beta_{(k',j)}(y)R_{(k,l),(k',j)}
\end{aligned} \tag{60}$$

where we took into account the orthogonality of $\varphi_{(1,1)}$ with the other elements of the basis in the second equality and where $\beta_0(y)\coloneqq(\beta_{(k',j)}(y))_{(k',j)\in S_0}$ is defined by

$$\forall\ (k',j)\in S_0,\qquad \beta_{(k',j)}(y) \coloneqq \sum_{l'\in[\![\gamma_{k'}]\!]\,:\,j\in[\![l']\!]}\widetilde{\alpha}_{(k',l')}(y)b_{(k',l'-j+1)}$$

Multiplying (59) by $\varphi_{(k',l')}$ for any $(k',l')\in S$ and integrating with respect to π, we get

$$\varphi_{(k',l')}(x) = \sum_{(k,l)\in S}\alpha_{(k,l)}(x)R_{(k,l),(k',l')}$$

namely we have the vectorial equality

$$R\alpha(x) \;=\; (\varphi_{(k',l')}(x))_{(k',l')\in S} \;=\; \varphi(x)$$

i.e.

$$\alpha(x) \;=\; R^{-1}\varphi(x) \tag{61}$$

Note that we can write

$$R \;=\; \left(\begin{array}{cc} 1 & 0 \\ 0 & R_0 \end{array}\right) \tag{62}$$

with $R_0 \coloneqq (R_{(k,l),(k',l')})_{(k,l),(k',l')\in S_0}$. Furthermore, we have

$$R^{-1} \;=\; \left(\begin{array}{cc} 1 & 0 \\ 0 & R_0^{-1} \end{array}\right) \tag{63}$$

From (61), we deduce $\alpha_{(1,1)}(x) = 1$ and $\alpha_0(x) = R_0^{-1}\varphi_0(x)$, with $\alpha_0(x) \coloneqq (\alpha_{(k,l)}(x))_{(k,l)\in S_0}$ and $\varphi_0(x) \coloneqq (\varphi_{(k,l)}(x))_{(k,l)\in S_0}$.

Applying (59) at the point x, we get

$$\begin{aligned} \frac{1}{\pi(x)} &= \frac{\mathbb{1}_x(x)}{\pi(x)} \\ &= 1 + \langle \alpha_0(x), \varphi_0(x)\rangle_0 \\ &= 1 + \langle \alpha_0(x), R_0\alpha_0(x)\rangle_0 \end{aligned}$$

where $\langle\cdot,\cdot\rangle_0$ is the usual scalar product on $\mathbb{R}^{S_0}$.

It follows that

$$\begin{aligned} \frac{1}{\pi(x)} &\geqslant \langle \alpha_0(x), R_0\alpha_0(x)\rangle_0 \\ &\geqslant \lambda_\wedge \left\|\alpha_0(x)\right\|_0^2 \end{aligned} \tag{64}$$

(since $\lambda_\wedge$ is also the smallest eigenvalue of R_0).

Similarly, we have

$$\frac{1}{\widetilde{\pi}(y)} \;\geqslant\; \widetilde{\lambda}_\wedge \left\|\widetilde{\alpha}_0(x)\right\|_0^2 \tag{65}$$

Coming back to (60), the Cauchy-Schwartz' inequality implies

$$\left|\frac{A_b(x,y)}{\widetilde{\pi}(y)}\right| \;\leqslant\; \left\|\beta_0(y)\right\|_0 \left\|R_0\alpha_0(x)\right\|_0 \tag{66}$$

Let us deal with the last factor:

$$\begin{aligned} \left\|R_0\alpha_0(x)\right\|_0 &= \sqrt{\langle R_0\alpha_0(x), R_0\alpha_0(x)\rangle_0} \\ &= \sqrt{\left\langle \sqrt{R_0}\alpha_0(x), R_0\sqrt{R_0}\alpha_0(x)\right\rangle_0} \\ &\leqslant \sqrt{\lambda_\vee \left\langle \sqrt{R_0}\alpha_0(x), \sqrt{R_0}\alpha_0(x)\right\rangle_0} \\ &= \sqrt{\lambda_\vee \left\langle \alpha_0(x), R_0\alpha_0(x)\right\rangle_0} \end{aligned}$$

$$\leqslant \sqrt{\frac{\lambda_\vee}{\pi(x)}} \tag{67}$$

On the other hand, we can bound the square of first factor of the r.h.s. of (66) by

$$\begin{aligned}
\|\beta_0(y)\|_0^2 &= \sum_{(k,j)\in S_0} \beta^2_{(k,j)}(y) \\
&= \sum_{(k,j)\in S_0} \left(\sum_{l\in[\![\gamma_k]\!]\,:\,j\in[\![l]\!]} \widetilde{\alpha}_{(k,l)}(y) b_{(k,l-j)} \right)^2 \\
&\leqslant \sum_{(k,j)\in S_0} \sum_{l\in[\![\gamma_k]\!]\,:\,j\in[\![l]\!]} \widetilde{\alpha}^2_{(k,l)}(y) \sum_{l'\in[\![\gamma_k]\!]\,:\,j\in[\![l']\!]} b^2_{(k,l'-j+1)} \\
&\leqslant \max\left\{ \sum_{l'\in[\![\gamma_{k'}]\!]} b^2_{(k',l')} \,:\, k'\in[\![r]\!] \right\} \sum_{(k,j)\in S_0} \sum_{l\in[\![\gamma_k]\!]\,:\,j\in[\![l]\!]} \widetilde{\alpha}^2_{(k,l)}(y) \\
&\leqslant \max\left\{ \gamma_{k'} b^2_{(k',l')} \,:\, (k',l')\in S_0 \right\} \sum_{(k,l)\in S_0} \widetilde{\alpha}^2_{(k,l)}(y) \sum_{j\in[\![l]\!]} 1 \\
&\leqslant \max\left\{ \gamma_{k'} b^2_{(k',l')} \,:\, (k',l')\in S_0 \right\} \sum_{(k,l)\in S_0} l\widetilde{\alpha}^2_{(k,l)}(y) \\
&\leqslant \Gamma \max\left\{ \gamma_{k'} b^2_{(k',l')} \,:\, (k',l')\in S_0 \right\} \sum_{(k,l)\in S_0} \widetilde{\alpha}^2_{(k,l)}(y) \\
&\leqslant \Gamma^2 \max\left\{ b^2_{(k',l')} \,:\, (k',l')\in S_0 \right\} \|\widetilde{\alpha}_0(y)\|_0^2 \\
&\leqslant \Gamma^2 \max\left\{ b^2_{(k',l')} \,:\, (k',l')\in S_0 \right\} \frac{1}{\widetilde{\lambda}_\wedge \widetilde{\pi}(y)}
\end{aligned}$$

according to (65). This leads to the first announced bound. The second bound is obtained by symmetry. The last assertion about η follows at once. ■

To finish this section, we give an application for bounding the support of q in Proposition 4, when all the eigenvalues of P are real.

Coming back to (50), (51) and (52), it appears that the support of the constructed q is included into $[\![0, n_0]\!]$. Taking into account (53), the bound $\binom{n}{m} \leqslant \frac{n^m}{m!}$ valid for all $n, m \in \mathbb{Z}_+$, and Lemma 11, we get $n_0 \leqslant \bar{n}_0$ with

$$\bar{n}_0 \;\coloneqq\; \min\left\{ n\in\mathbb{Z}_+ \,:\, \forall\, k\in\mathcal{K},\; n^{\Gamma-1}\zeta^{n/2} \leqslant \frac{(\Gamma-1)!}{\Gamma^2}\sqrt{\frac{\widetilde{\lambda}_\wedge}{\lambda_\vee}}\sqrt{\pi_\wedge\widetilde{\pi}_\wedge}\zeta^{\Gamma-1} \right\}$$

4 Matthews result

Our purpose here is to show Theorem 5 of Matthews [3] by interpreting it as a degenerate version of Proposition 4 where $\widetilde{P}$ is an absorbed Markov chain.

Let P be an irreducible and reversible transition matrix on V and recall the notations introduced before Theorem 5. We assume that the eigenvalues of P are non-negative.

Consider the state space $\widetilde{V} \coloneqq [\![|V|]\!]$ endowed with the transition kernel $\widetilde{P}$ defined by

$$\forall\, k,l\in\widetilde{V}, \qquad \widetilde{P}(k,l) \;\coloneqq\; \begin{cases} 1 & \text{, if } k=l=1 \\ \theta_k & \text{, if } k=l\geqslant 2 \\ 1-\theta_k & \text{, if } k\geqslant 2 \text{ and } l=1 \\ 0 & \text{, otherwise} \end{cases}$$

The corresponding Markov chains are absorbed at 1. Since $\widetilde{P}$ is lower diagonal, its eigenvalues are given by the entries of the diagonal, namely are exactly those of P. Furthermore, for any $k \in [\![2, |V|]\!]$, an eigenvector associated to θ_k for $\widetilde{P}$ is $\widetilde{\varphi}_k := \mathbb{1}_k$. As usual we take $\widetilde{\varphi}_1 = \mathbb{1}$.

We say that $\widetilde{P}$ is a **simple model** for P.

Let $X := (X(n))_{n\in\mathbb{Z}_+}$ be a Markov chain as in Theorem 5, namely with transition matrix P and initial distribution μ_0, which is fixed from now on. Up to multiplying some of the eigenfunctions by -1, we can assume that

$$\forall\ k \in [\![|V|]\!], \qquad \mu_0[\varphi_k] \;\geqslant\; 0 \tag{68}$$

(of course this is automatically satisfied for $\varphi_1 = \mathbb{1}$). In particular for the quantity defined in (7), we have

$$Z(\mu_0) \;=\; \sum_{l\in[\![|V|]\!]} \mu_0[\varphi_l]$$

as mentioned in the introduction $Z(\mu_0) = 0$ if and only $\mu_0 = \pi$, which is also the only case where in (71) below we have $n_0 = 0$. From now on we assume that $Z(\mu_0) > 0$.

Consider the $V \times \widetilde{V}$ matrix Λ defined by

$$\forall\ x \in V,\ \forall\ k \in \widetilde{V}, \qquad \Lambda(x,k) \;:=\; \begin{cases} \frac{\varphi_k(x)}{Z(\mu_0)} & \text{, if } k \geqslant 2 \\ 0 & \text{, if } k = 1 \end{cases}$$

Contrary to the previous sections, Λ is not a transition matrix, since some of its entries are negative. Nevertheless it has two interesting properties. First we check that

$$\mu_0\Lambda \;=\; \widetilde{\mu}_0 \tag{69}$$

the probability on $\widetilde{V}$ defined in (6).

Secondly, we have

$$\forall\ k \in [\![|V|]\!], \qquad \Lambda[\widetilde{\varphi}_k] \;=\; \begin{cases} \frac{1}{Z(\mu_0)}\varphi_k & \text{, if } k \geqslant 2 \\ \varphi_1 & \text{, if } k = 1 \end{cases} \tag{70}$$

In contrast, we look for a "true" transition kernel $\widetilde{\Lambda}$ from $\widetilde{V}$ to V verifying the properties of the following lemma. Define

$$n_0 \;:=\; \left\lceil \max\left\{ \frac{\ln(Z(\mu_0)\,\|\varphi_k\|_\infty)}{\ln(1/\theta_k)} \;:\; k \in [\![2, |V|]\!] \right\} \right\rceil \tag{71}$$

(again the ratio vanishes when $\theta_k = 0$).

Lemma 12 *There exist a transition kernel $\widetilde{\Lambda}$ from $\widetilde{V}$ to V and a probability $q := (q_n)_{n\in\mathbb{Z}_+}$ on $[\![0, n_0]\!]$ such that* (1) *and* (3) *are satisfied.*

Before proving Lemma 12, let us show how it implies Theorem 5:

Proof of Theorem 5

Consider $\widetilde{X} := (\widetilde{X}(n))_{n\in\mathbb{Z}_+}$ and $Y := (Y(n))_{n\in\mathbb{Z}_+}$, respectively a Markov chain with transition kernel $\widetilde{P}$ and $\widetilde{\mu}_0$ as initial distribution and a Markov chain with transition kernel P and $\nu_0 := \widetilde{\mu}_0\widetilde{\Lambda}$ as initial distribution.

Due to (1) and $\nu_0 = \widetilde{\mu}_0\widetilde{\Lambda}$, Diaconis and Fill [1] provide a coupling of $\widetilde{X}$ and Y such that we have for any $n \in \mathbb{Z}_+$,

$$\mathcal{L}(\widetilde{X}([\![0,n]\!])|Y) \;=\; \mathcal{L}(\widetilde{X}([\![0,n]\!])|Y([\![0,n]\!])) \tag{72}$$

$$\mathcal{L}(Y(n)|\widetilde{X}([\![0,n]\!])) \quad = \quad \widetilde{\Lambda}(\widetilde{X}(n),\cdot) \tag{73}$$

(where the various $\mathcal{L}(\cdot|\cdot)$ stand for conditional distributions).

From the first relation, we deduce that any stopping time relative to $\widetilde{X}$ is also a stopping time relative to Y. The second relation, which can be seen as a probabilistic version of (1), is still valid when n is replaced by a stopping time for $\widetilde{X}$. It leads us to introduce the stopping time

$$\widetilde{\tau} \quad \coloneqq \quad \inf\{n \in \mathbb{Z}_+ \,:\, \widetilde{X}(n) = 1\}$$

which is finite a.s., since $1-\theta_k > 0$ for any $k \in [\![2,|V|]\!]$.

From (1) and the fact that 1 is absorbing for $\widetilde{X}$, we deduce that $\widetilde{\Lambda}(1,\cdot)$ is invariant for P, namely $\widetilde{\Lambda}(1,\cdot) = \pi$. It follows that $Y(\widetilde{\tau})$ is distributed according to $\pi = \widetilde{\Lambda}(\widetilde{X}(\widetilde{\tau}),\cdot)$. Furthermore, from $\mathcal{L}(Y(\widetilde{\tau})|\widetilde{X}([\![0,\widetilde{\tau}]\!])) = \widetilde{\Lambda}(\widetilde{X}(\widetilde{\tau}),\cdot) = \pi$, we deduce that $Y(\widetilde{\tau})$ is independent from $\widetilde{\tau}$, since $\widetilde{\tau}$ is measurable with respect to $\widetilde{X}([\![0,\widetilde{\tau}]\!])$ (and maybe to some additional independent randomness). Thus $\widetilde{\tau}$ is a strong stationary time for Y. For more details about these classical assertions, see Diaconis and Fill [1].

The extreme simplicity of $\widetilde{P}$ shows that $\widetilde{\tau}$ is distributed as the random variable $\mathcal{G}$ described above the statement of Theorem 5.

Consider $\widehat{\tau}$ a time independent from X and distributed according to the probability q appearing in Lemma 12.

From (69) and (3), we deduce that Y has the same law as $(X(\widehat{\tau}+n))_{n\in\mathbb{Z}_+}$. It leads us to define $\tau \coloneqq \widehat{\tau}+\widetilde{\tau}$, since we get that $X(\tau)$ is distributed according to π. To see that τ is a strong stationary time for X, it remains to check that τ and $X(\tau)$ are independent. So let be given two functions $f\,:\,V\to\mathbb{R}_+$ and $g\,:\,\mathbb{Z}_+\to\mathbb{R}_+$. We compute

$$\begin{aligned}
\mathbb{E}[f(X_\tau)g(\tau)] &= \mathbb{E}[f(Y(\widetilde{\tau}))g(\widehat{\tau}+\widetilde{\tau})]\\
&= \sum_{n\in[\![0,n_0]\!]} q_n\mathbb{E}[f(Y(\widetilde{\tau}))g(n+\widetilde{\tau})]\\
&= \sum_{n\in[\![0,n_0]\!]} q_n\mathbb{E}[f(Y(\widetilde{\tau}))]\mathbb{E}[g(n+\widetilde{\tau})]\\
&= \mathbb{E}[f(Y(\widetilde{\tau}))]\sum_{n\in[\![0,n_0]\!]} q_n\mathbb{E}[g(n+\widetilde{\tau})]\\
&= \mathbb{E}[f(Y(\widetilde{\tau}))]\mathbb{E}[g(\tau)]\\
&= \mathbb{E}[f(X(\tau)]\mathbb{E}[g(\tau)]
\end{aligned}$$

where in the third equality we used the independence of $Y(\tau)$ and τ.

Since the support of q is included into $[\![0,n_0]\!]$, τ is stochastically dominated by $n_0+\mathcal{G}$, showing the first assertion of Theorem 5.

For the second assertion, note that for any $k \in [\![2,|V|]\!]$, we have $\|\varphi_k\|_\infty \leqslant 1/\sqrt{\pi_\wedge}$ (use either $\pi[\varphi_k^2]=1$ or (11)) and $0\leqslant\theta_k\leqslant\theta_2$. We deduce, on one hand, that

$$\begin{aligned}
Z(\mu_0) &\leqslant |V|\max\{\|\varphi_k\|_\infty \,:\, k\in[\![|V|]\!]\setminus\{1\}\}\\
&= \frac{|V|}{\sqrt{\pi_\wedge}}
\end{aligned}$$

and

$$n_0 \quad \leqslant \quad \left\lceil \frac{\ln(|V|/\pi_\wedge)}{\ln(1/\theta_2)} \right\rceil$$

On the other hand, it is clear that $\mathcal{G}$ is stochastically dominated by a geometric random variable of parameter θ_2. ■

Let us now come to the

Proof of Lemma 12

The calculations are inspired by those of Lemma 7.

Let be given a family $\widetilde{b} \coloneqq (\widetilde{b}_k)_{k\in[\![|V|]\!]}$ with $\widetilde{b}_1 = 1$. We look for an operator $\widetilde{\Lambda}_{\widetilde{b}}$ which is such that

$$\forall\, l \in [\![|V|]\!], \qquad \widetilde{\Lambda}_{\widetilde{b}}[\varphi_l] \;=\; \widetilde{b}_l\widetilde{\varphi}_l \tag{74}$$

which ensures the commutativity property (1). Let us check that $\widetilde{\Lambda}_{\widetilde{b}}$ is a transition kernel for appropriate choices of $\widetilde{b}$.

The associated matrix $(\widetilde{\Lambda}_{\widetilde{b}}(k,x))_{k\in\widetilde{V},\, x\in V}$ is such that

$$\forall\, k \in \widetilde{V},\, \forall\, x \in V, \qquad \frac{\widetilde{\Lambda}_{\widetilde{b}}(k,x)}{\pi(x)} \;=\; \left\langle \mathbb{1}_k, \widetilde{\Lambda}_{\widetilde{b}}\left[\frac{\mathbb{1}_x}{\pi(x)}\right]\right\rangle$$

where $\langle\cdot,\cdot\rangle$ is the usual scalar product in $\mathbb{R}^{\widetilde{V}}$.

Let us decompose

$$\frac{\mathbb{1}_x}{\pi(x)}(\cdot) \;=\; \sum_{l\in[\![|V|]\!]} \alpha_l(x)\varphi_l(\cdot) \tag{75}$$

with some real coefficients $\alpha(x) \coloneqq (\alpha_l(x))_{l\in[\![|V|]\!]}$.

We deduce

$$\forall\, k \in \widetilde{V},\, \forall\, x \in V, \qquad \frac{\widetilde{\Lambda}_{\widetilde{b}}(k,x)}{\pi(x)} \;=\; \sum_{l\in[\![|V|]\!]} \alpha_l(x)\widetilde{b}_l\,\langle \mathbb{1}_k, \widetilde{\varphi}_l\rangle$$

On one hand, we compute for any $k, l \in \widetilde{V}$,

$$\langle \mathbb{1}_k, \widetilde{\varphi}_l\rangle \;=\; \begin{cases} 1 & \text{, if } l = 1 \\ \delta_{k,l} & \text{, if } l \geqslant 2 \end{cases}$$

where $\delta_{k,l}$ is the Kronecker symbol.

On the other hand, multiplying (75) by φ_j, for $j \in [\![|V|]\!]$ and integrating with respect to π, we get

$$\begin{aligned}
\varphi_j(x) &= \sum_{l\in[\![|V|]\!]} \alpha_l(x)\pi[\varphi_j\varphi_l] \\
&= \sum_{l\in[\![|V|]\!]} \alpha_l(x)\delta_{j,l} \\
&= \alpha_j(x)
\end{aligned}$$

Thus we get,

$$\begin{aligned}
\forall\, k \in \widetilde{V},\, \forall\, x \in V, \qquad \frac{\widetilde{\Lambda}_{\widetilde{b}}(k,x)}{\pi(x)} &= \begin{cases} \alpha_1(x)\widetilde{b}_1 & \text{, if } k = 1 \\ \alpha_1(x)\widetilde{b}_1 + \alpha_k(x)\widetilde{b}_k & \text{, if } k \geqslant 2 \end{cases} \\
&= \begin{cases} 1 & \text{, if } k = 1 \\ 1 + \varphi_k(x)\widetilde{b}_k & \text{, if } k \geqslant 2 \end{cases}
\end{aligned}$$

We deduce that the entries of $\widetilde{\Lambda}$ are non-negative if and only if

$$\forall\, k \in \widetilde{V}\backslash\{1\},\, \forall\, x \in V, \qquad 1 + \varphi_k(x)\widetilde{b}_k \;\geqslant\; 0 \tag{76}$$

In this case, $\widetilde{\Lambda}_{\widetilde{b}}$ is a transition kernel, since $\widetilde{\Lambda}_{\widetilde{b}}[\mathbb{1}] = \widetilde{\Lambda}_{\widetilde{b}}[\varphi_1] = \widetilde{\varphi}_1 = \mathbb{1}$.

A simple sufficient condition ensuring (76) is

$$\forall\ k \in \widetilde{V}\setminus\{1\}, \qquad |\widetilde{b}_k| \ \leqslant\ \frac{1}{\|\varphi_k\|_\infty} \tag{77}$$

Let us compute $\Lambda\widetilde{\Lambda}_{\widetilde{b}}$. From (70) and (74) we get

$$\forall\ k \in [\![|V|]\!], \qquad \Lambda\widetilde{\Lambda}_{\widetilde{b}}[\varphi_k] \ =\ \begin{cases} \frac{\widetilde{b}_k}{Z(\mu_0)}\varphi_k & \text{, if } k \geqslant 2 \\ \varphi_1 & \text{if, } k=1 \end{cases}$$

Thus (3) is satisfied if and only if

$$\forall\ k \in [\![|V|]\!]\setminus\{1\}, \qquad \frac{\widetilde{b}_k}{Z(\mu_0)} \ =\ \sum_{n\in\mathbb{Z}_+} q_n\theta_k^n$$

Let us look for a probability q of the form δ_{n_0} for some $n_0 \in \mathbb{Z}_+$ with $\widetilde{b}$ satisfying (77) (contrary to the proof of Proposition 6, we do not need here that the entries of $\widetilde{b}$ do not vanish). Considering n_0 given in (71) we check that for any $k \in [\![|V|]\!]\setminus\{1\}$,

$$\begin{aligned} |\widetilde{b}_k| &= Z(\mu_0)\theta_k^{n_0} \\ &\leqslant Z(\mu_0)\theta_k^{\frac{\ln(Z(\mu_0)\|\varphi_k\|_\infty)}{\ln(1/\theta_k)}} \\ &\leqslant \frac{1}{\|\varphi_k\|_\infty} \end{aligned}$$

The Markov kernel $\widetilde{\Lambda} = \widetilde{\Lambda}_{\widetilde{b}}$ and the probability $q = \delta_{n_0}$ provide us with the desired properties. ■

5 Markov kernels with non-negative eigenvalues

Our purpose here is to extend Theorem 5 of Matthews [3] to all Markov kernels whose eigenvalues are non-negative. In particular we will introduce degenerate models for them. It would be interesting to extend the results presented here to any finite irreducible Markov kernel, but we are missing simple models for negative and complex eigenvalues. We hope this challenge will trigger research in this direction, as it also related to the understanding of transition kernel complex eigenvalues.

Let P be an irreducible transition matrix on V whose eigenvalues are non-negative. Recall the notations introduced before (21), in particular the eigenvalues are given by

$$1 = \theta_1 > \theta_2 \geqslant \theta_3 \geqslant \cdots \geqslant \theta_r \geqslant 0$$

Introduce the state space $\widetilde{V} \coloneqq S$, the characteristic set of P, endowed with the transition kernel $\widetilde{P}$ defined by

$$\forall\ (k,l),(k',l') \in \widetilde{V}, \qquad \widetilde{P}((k,l),(k',l')) \ \coloneqq\ \begin{cases} 1 & \text{, if } (k,l)=(k',l')=(1,1) \\ \theta_k & \text{, if } k=k' \geqslant 2 \text{ and } l=l' \\ 1-\theta_k & \text{, if } k=k' \geqslant 2 \text{ and } l'=l-1 \geqslant 1 \\ 1-\theta_k & \text{, if } k \geqslant 2,\ l=1 \text{ and } (k',l')=(1,1) \\ 0 & \text{, otherwise} \end{cases}$$

The associated graph looks like a star, with $(1,1)$ as central point to which are converging $r-1$ rays of respective lengths γ_2, ..., γ_r. The corresponding Markov chains are absorbed at $(1,1)$.

By removing $1-\theta_k$ times the first row to the rows $(k,1)$, $(k,2)$, ..., (k,γ_k), for any $k \in [\![2,r]\!]$, we transform $\widetilde{P}$ into a block diagonal matrix whose blocks are exactly the Jordan blocks of P. Thus P and $\widetilde{P}$ have the same characteristic set S.

For any $(k,l) \in S$, with $k \in [\![2,r]\!]$, a **generalized eigenvector** associated to θ_k for $\widetilde{P}$ is $\widetilde{\varphi}_{(k,l)} \coloneqq \mathbb{1}_{(k,l)}$, in the sense that

$$\widetilde{P}[\widetilde{\varphi}_{(k,l)}] \quad = \quad \theta_k \widetilde{\varphi}_{(k,l)} + \widetilde{\varphi}_{(k,l-1)}$$

where by convention, $\widetilde{\varphi}_{(k,0)} = 0$ for all $k \in [\![2,r]\!]$.

As usual we take $\widetilde{\varphi}_{(1,1)} = \mathbb{1}$.

We say again that $\widetilde{P}$ is a **simple model** for P.

Let $X \coloneqq (X(n))_{n\in\mathbb{Z}_+}$ be a Markov chain with transition matrix P and initial distribution μ_0, which is fixed from now on. We will need the following technical result replacing (68):

Lemma 13 *The adapted basis $(\varphi_{(k,l)})_{(k,l)\in S}$ can be modified into another adapted basis $(\varphi'_{(k,l)})_{(k,l)\in S}$ so that in addition to keeping $\varphi'_{(1,1)} = \mathbb{1}$, we have*

$$\forall\ (k,l) \in S, \qquad \mu_0[\varphi'_{(k,l)}] \quad \geqslant \quad 0$$

Proof

Fix $k \in [\![2,r]\!]$, we show by iteration on $l \in [\![1,\gamma_k]\!]$ that we can change the generalized the family of vectors $(\varphi_{(k,j)})_{j\in[\![l]\!]}$ into $(\varphi'_{(k,j)})_{j\in[\![l]\!]}$, so that

$$\forall\ j \in [\![l]\!], \qquad \mu_0[\varphi'_{(k,j)}] \quad \geqslant \quad 0$$

while keeping the relations

$$\forall\ j \in [\![l]\!], \qquad \widetilde{P}[\varphi'_{(k,j)}] \quad = \quad \theta_k \varphi'_{(k,j)} + \varphi'_{(k,j-1)}$$

(with $\varphi'_{(k,0)} = 0$).

For $l = 1$, this is clear: if $\mu_0[\varphi_{(k,1)}] \geqslant 0$, we just take $\varphi'_{(k,1)} \coloneqq \varphi_{(k,1)}$ and otherwise we carry out the replacement $\varphi'_{(k,1)} \coloneqq -\varphi_{(k,1)}$.

Assume the iteration is true for some $l \in [\![\gamma_k]\!]$ with $l < \gamma_k$. Let us change the family $(\varphi'_{(k,j)})_{j\in[\![l+1]\!]}$, with $\varphi'_{(k,l+1)} \coloneqq \varphi_{(k,l+1)}$ into $(\varphi''_{(k,j)})_{j\in[\![l+1]\!]}$ with the desired property. Note that if $\mu_0[\varphi'_{(k,l+1)}] \geqslant 0$, it is sufficient to keep the same sequence: $(\varphi''_{(k,j)})_{j\in[\![l+1]\!]} \coloneqq (\varphi'_{(k,j)})_{j\in[\![l+1]\!]}$. So let us assume that $\mu_0[\varphi'_{(k,l+1)}] < 0$.

We consider two cases.

• When for any $j \in [\![l]\!]$, $\mu_0[\varphi'_{(k,j)}] = 0$ we just carry out the replacement $(\varphi''_{(k,j)})_{j\in[\![l+1]\!]} \coloneqq (-\varphi'_{(k,j)})_{j\in[\![l+1]\!]}$.

• Otherwise consider the first $m \in [\![l]\!]$ such that $\mu_0[\varphi'_{(k,m)}] > 0$. For $a \geqslant 0$ consider

$$(\varphi''_{(k,j)})_{j\in[\![l+1]\!]} \quad \coloneqq \quad (\varphi'_{(k,j)} + a\varphi'_{(k,j+m-l-1)})_{j\in[\![l+1]\!]}$$

(with the convention that for any $u \leqslant 0$, $\varphi'_{(k,u)} = 0$). We check that

$$\forall\ j \in [\![l+1]\!], \qquad \widetilde{P}[\varphi''_{(k,j)}] \quad = \quad \theta_k \varphi''_{(k,j)} + \varphi''_{(k,j-1)}$$

so $(\varphi''_{(k,j)})_{j\in[\![l+1]\!]}$ still consists of generalized eigenvectors.

By the iteration assumption and since $a \geqslant 0$, we have $\mu_0[\varphi''_{(k,j)}] \geqslant 0$ for any $j \in [\![l]\!]$. Taking furthermore $a \geqslant -\mu_0[\varphi'_{(k,l+1)}]/\mu_0[\varphi'_{(k,m)}] > 0$, we also get $\mu_0[\varphi''_{(k,l+1)}] \geqslant 0$ as wanted. ■

From now on, we assume the adapted basis $(\varphi_{(k,l)})_{(k,l)\in S}$ satisfies

$$\forall\ (k,l)\in S, \qquad \mu_0[\varphi_{(k,l)}] \ \geqslant\ 0$$

We introduce the probability $\widetilde{\mu}_0$ on $\widetilde{V}$ via

$$\forall\ (k,l)\in\widetilde{V}, \qquad \widetilde{\mu}_0((k,l)) \ :=\ \begin{cases} \frac{\mu_0[\varphi_{(k,l)}]}{Z(\mu_0)} & \text{, if } k\geqslant 2 \\ 0 & \text{, if } (k,l)=(1,1) \end{cases} \tag{78}$$

where the normalizing factor is given by

$$Z(\mu_0) := \sum_{(k',l')\in S\setminus\{(1,1)\}} \mu_0[\varphi_{(k',l')}] \tag{79}$$

Consider $\widetilde{X} := (\widetilde{X}(n))_{n\in\mathbb{Z}_+}$ a Markov chain with transition matrix $\widetilde{P}$ and initial distribution $\widetilde{\mu}_0$. Define

$$\mathcal{G} \ :=\ \inf\{n\in\mathbb{Z}_+\ :\ \widetilde{X}(n)=(1,1)\}$$

The law of $\mathcal{G}$ is a mixture of convolutions of geometric law of parameters the eigenvalues of P.

As in (62), consider $S_0 := S\setminus\{(1,1)\}$ and the Gramian matrix defined by

$$\forall\ (k',l'),(k'',l'')\in S_0 \qquad R_{(k',l'),(k'',l'')} \ :=\ \pi[\varphi_{(k',l')}\varphi_{(k'',l'')}]$$

where π is the invariant probability associated to P.

Recall (see the sentence after (63)) that for any $x\in V$, $\varphi_0(x) := (\varphi_{(k,l)}(x))_{(k,l)\in S_0}$ and define $\alpha_0(x) := (\alpha_{(k,l)}(x))_{(k,l)\in S_0}$ by $\alpha_0(x) = R_0^{-1}\varphi_0(x)$. Introduce the quantities

$$M \ :=\ Z(\mu_0)\max\left\{\max_{k\in[\![2,r]\!]}\sum_{l\in[\![\gamma_k]\!]}|\alpha_{(l,k)}(x)|\ :\ x\in V\right\} \tag{80}$$

$$n_0 \ :=\ \min\left\{n\in\mathbb{Z}_+\ :\ \max_{(k,l)\in S\setminus\{(1,1)\}}\binom{n}{l-1}\theta_k^{n+1-l}\leqslant\frac{1}{M}\right\} \tag{81}$$

(recall that $\pi_\wedge := \min\{\pi(x)\ :\ x\in V\}$).

Equally recall the quantities $\lambda_\wedge>0$ and Γ described before and inside Lemma 11.

Define

$$\bar{M} \ :=\ \frac{\sqrt{|V|\Gamma}}{\pi_\wedge}\sqrt{\frac{\lambda_\vee}{\lambda_\wedge}}$$

$$\bar{n}_0 \ :=\ (2(\Gamma-1))\vee\min\left\{n\in\mathbb{Z}_+\ :\ \binom{n}{\Gamma-1}\theta_2^{n+1-\Gamma}\leqslant\frac{1}{\bar{M}}\right\} \tag{82}$$

Here is the generalization of Theorem 5 that we will prove here:

Theorem 14 *Assume that P is irreducible and that its eigenvalues are all non-negative. Then there exists a strong stationary time for X which is stochastically dominated by*

$$n_0+\mathcal{G} \tag{83}$$

This random variable is itself stochastically dominated by $\bar{n}_0+\mathcal{H}_2$, where $\mathcal{H}_2$ is the convolution of Γ independent geometric random variables of parameter θ_2.

The arguments adapt the proof of Theorem 5, taking into account the considerations of Section 3, in particular estimates such as those of Lemma 11.

Consider the $V \times \widetilde{V}$ matrix Λ defined by

$$\forall\ x \in V, \forall\ (k,l) \in \widetilde{V}, \qquad \Lambda(x,(k,l)) \;\coloneqq\; \begin{cases} \frac{\varphi_{(k,l)}(x)}{Z(\mu_0)} & \text{, if } k \geqslant 2 \\ 0 & \text{, if } (k,l) = (1,1) \end{cases} \tag{84}$$

As in the previous section, Λ is not a transition matrix, since some of its entries are negative. Nevertheless it has the same two interesting properties. First we have

$$\mu_0 \Lambda \;=\; \widetilde{\mu}_0$$

the probability on $\widetilde{V}$ defined in (78).

Secondly, we have

$$\forall\ (k,l) \in \widetilde{V}, \qquad \Lambda[\widetilde{\varphi}_{(k,l)}] \;=\; \begin{cases} \frac{1}{Z(\mu_0)} \varphi_{(k,l)} & \text{, if } k \geqslant 2 \\ \varphi_{(1,1)} & \text{, if } (k,l) = (1,1) \end{cases} \tag{85}$$

Nevertheless, we look for a "true" transition kernel $\widetilde{\Lambda}$ from $\widetilde{V}$ to V verifying the properties of the following lemma.

Lemma 15 *There exist a transition kernel $\widetilde{\Lambda}$ from $\widetilde{V}$ to V and a probability $q \coloneqq (q_n)_{n \in \mathbb{Z}_+}$ on $[\![0, n_0]\!]$ such that* (1) *and* (3) *are satisfied.*

Proof

The calculations are inspired by those of Lemma 7.

Let be given a real-valued family $\widetilde{b} \coloneqq (\widetilde{b}_{(k,l)})_{(k,l) \in \widetilde{V}}$ with $\widetilde{b}_{(1,1)} = 1$. We look for an operator $\widetilde{\Lambda}_{\widetilde{b}}$ which is such that

$$\forall\ (k,l) \in \widetilde{V}, \qquad \widetilde{\Lambda}_{\widetilde{b}}[\varphi_{(k,l)}] \;=\; \sum_{j \in [\![l]\!]} \widetilde{b}_{(k,l-j+1)} \widetilde{\varphi}_{(k,j)} \tag{86}$$

which ensures the commutativity property (1), see Lemma 8. Let us check that $\widetilde{\Lambda}_{\widetilde{b}}$ is a transition kernel for appropriate choices of $\widetilde{b}$.

The associated matrix $(\widetilde{\Lambda}_{\widetilde{b}}((k,l),x))_{(k,l) \in \widetilde{V},\, x \in V}$ is such that

$$\forall\ (k,l) \in \widetilde{V}, \forall\ x \in V, \qquad \frac{\widetilde{\Lambda}_{\widetilde{b}}((k,l),x)}{\pi(x)} \;=\; \left\langle \mathbb{1}_{(k,l)}, \widetilde{\Lambda}_{\widetilde{b}} \left[\frac{\mathbb{1}_x}{\pi(x)} \right] \right\rangle$$

where we recall that $\langle \cdot, \cdot \rangle$ is the usual scalar product in $\mathbb{R}^{\widetilde{V}}$.

Let us decompose

$$\frac{\mathbb{1}_x}{\pi(x)} \;=\; \sum_{(k',l') \in \widetilde{V}} \alpha_{(k',l')}(x) \varphi_{(k',l')}$$

with some real coefficients $\alpha(x) \coloneqq (\alpha_{(k',l')}(x))_{(k',l') \in \widetilde{V}}$.

We deduce

$$\forall\ (k,l) \in \widetilde{V}, \forall\ x \in V, \qquad \frac{\widetilde{\Lambda}_{\widetilde{b}}((k,l),x)}{\pi(x)} \;=\; \sum_{(k',l') \in \widetilde{V}} \sum_{j \in [\![l']\!]} \alpha_{(k',l')}(x) \widetilde{b}_{(k',l'-j+1)} \left\langle \mathbb{1}_{(k,l)}, \widetilde{\varphi}_{(k',j)} \right\rangle$$

We compute for any (k,l), $(k',j)\in\widetilde{V}$,

$$\left\langle \mathbb{1}_{(k,l)},\widetilde{\varphi}_{(k',j)}\right\rangle \;=\; \begin{cases} 1 & \text{, if } (k',j)=(1,1)\\ \delta_{(k,l),(k',j)} & \text{, if } (k',j)\in\widetilde{V}\setminus\{(1,1)\}\end{cases}$$

where $\delta_{(k,l),(k',j)}$ is the Kronecker symbol, now respectively to the couples (k,l) and (k',j).

It follows that for any $(k,l)\in\widetilde{V}$ and $x\in V$,

$$\begin{aligned}\frac{\widetilde{\Lambda}_{\widetilde{b}}((k,l),x)}{\pi(x)} &= \begin{cases}\alpha_{(1,1)}(x)\widetilde{b}_{(1,1)} & \text{, if } (k,l)=(1,1)\\ \alpha_{(1,1)}(x)\widetilde{b}_{(1,1)}+\sum_{l'\in\llbracket l,\gamma_k\rrbracket}\alpha_{(k,l')}(x)\widetilde{b}_{(k,l'-l+1)} & \text{, if } (k,l)\in\widetilde{V}\setminus\{(1,1)\}\end{cases}\\ &= \begin{cases}\alpha_{(1,1)}(x) & \text{, if } (k,l)=(1,1)\\ \alpha_{(1,1)}(x)+\sum_{l'\in\llbracket l,\gamma_k\rrbracket}\alpha_{(k,l')}(x)\widetilde{b}_{(k,l'-l+1)} & \text{, if } (k,l)\in\widetilde{V}\setminus\{(1,1)\}\end{cases}\end{aligned}$$

Recall that the family of coefficients $\alpha(x)$ has been computed in (61), which is still valid here, with $R\coloneqq(R_{(k',l'),(k'',l'')})_{(k',l'),(k'',l'')\in\widetilde{V}}$ the Gramian matrix defined by

$$\forall\ (k',l'),(k'',l'')\in\widetilde{V},\qquad R_{(k',l'),(k'',l'')} \;\coloneqq\; \pi[\varphi_{(k',l')}\varphi_{(k'',l'')}]$$

The link with the matrix R_0 mentioned before the statement of Theorem 14 comes from (62). In particular, as observed after (63), we have $\alpha_{(1,1)}(x)=1$. Thus we get,

$$\frac{\widetilde{\Lambda}_{\widetilde{b}}((k,l),x)}{\pi(x)} \;=\; \begin{cases}1 & \text{, if } (k,l)=(1,1)\\ 1+\sum_{l'\in\llbracket l,\gamma_k\rrbracket}\alpha_{(k,l')}(x)\widetilde{b}_{(k,l'-l+1)} & \text{, if } (k,l)\in\widetilde{V}\setminus\{(1,1)\}\end{cases}$$

We deduce that the entries of $\widetilde{\Lambda}$ are non-negative if and only if

$$\forall\ (k,l)\in\widetilde{V}\setminus\{(1,1)\},\ \forall\ x\in V,\qquad \sum_{l'\in\llbracket l,\gamma_k\rrbracket}\alpha_{(k,l')}(x)\widetilde{b}_{(k,l'-l+1)} \;\geqslant\; -1 \tag{87}$$

In this case, $\widetilde{\Lambda}_{\widetilde{b}}$ is a transition kernel, since $\widetilde{\Lambda}_{\widetilde{b}}[\mathbb{1}]=\widetilde{\Lambda}_{\widetilde{b}}[\varphi_{(1,1)}]=\widetilde{\varphi}_{(1,1)}=\mathbb{1}$.

A simple sufficient condition ensuring (87) is

$$(k,l)\in\widetilde{V}\setminus\{(1,1)\},\qquad |\widetilde{b}_{(k,l)}| \;\leqslant\; \frac{1}{\max\left\{\max_{k\in\llbracket 2,r\rrbracket}\sum_{l\in\llbracket\gamma_k\rrbracket}|\alpha_{(l,k)}(x)|\ :\ x\in V\right\}} \tag{88}$$

Let us compute $\Lambda\widetilde{\Lambda}_{\widetilde{b}}$. From (85) and (86) we get

$$\forall\ (k,l)\in\widetilde{V},\qquad \Lambda\widetilde{\Lambda}_{\widetilde{b}}[\varphi_{(k,l)}] \;=\; \begin{cases}\sum_{j\in\llbracket l\rrbracket}\frac{1}{Z(\mu_0)}\widetilde{b}_{(k,l-j+1)}\varphi_{(k,j)} & \text{, if } k\geqslant 2\\ \varphi_{(1,1)}\end{cases}$$

Writing P in the adapted basis $(\varphi_{(k,l)})_{(k,l)\in S}$, it appears that for any $n\in\mathbb{Z}_+$, we have

$$\forall\ (k,l)\in S,\qquad P^n[\varphi_{(k,l)}] \;=\; \sum_{j\in\llbracket l\rrbracket}\binom{n}{l-j}\theta_k^{n+j-l}\varphi_{(k,j)} \tag{89}$$

It follows that (3) is satisfied with $q=\delta_{n_0}$ for some $n_0\in\mathbb{Z}_+$, if and only if

$$\forall\ (k,l)\in S,\qquad \widetilde{b}_{(k,l)} \;=\; Z(\mu_0)\binom{n_0}{l-1}\theta_k^{n_0+1-l}$$

(when $\theta_k = 0$, this requires that $n_0 \geqslant l-1$, we then rather take $n_0 \geqslant l$ so that $\widetilde{b}_{(k,l)} = 0$, this is not a problem here since we are not looking for an invertible link, explaining why we do not consider a probability of the form $q = q_0\delta_0 + (1-q_0)\delta_{n_0}$, contrary to Section 3).

Since we want (88) to be satisfied, we take n_0 defined in (81) and $q_0 = 1/M$, with M given in (80). Indeed, we check that for any $(k,l) \in S\backslash\{(1,1)\}$,

$$
\begin{aligned}
|\widetilde{b}_{(k,l)}| &= Z(\mu_0)\binom{n_0}{l-1}\theta_k^{n_0+1-l} \\
&\leqslant \frac{Z(\mu_0)}{M} \\
&\leqslant \frac{1}{\max\left\{\max_{k\in[\![2,r]\!]}\sum_{l\in[\![\gamma_k]\!]}|\alpha_{(l,k)}(x)| \ : \ x \in V\right\}}
\end{aligned}
$$

The Markov kernel $\widetilde{\Lambda} = \widetilde{\Lambda}_{\widetilde{b}}$ and the probability $q = \delta_{n_0}$ satisfy the desired properties. ■

We can now come to the

Proof of Theorem 14

The first assertion is shown in exactly the same way as in the first part of the proof of Theorem 5, with $\widetilde{X} \coloneqq (\widetilde{X}(n))_{n\in\mathbb{Z}_+}$ and $Y \coloneqq (Y(n))_{n\in\mathbb{Z}_+}$, Markov chains with transition kernels and initial distributions respectively given by $\widetilde{P}$ and $\widetilde{\mu}_0$ and P and $\nu_0 \coloneqq \widetilde{\mu}_0\widetilde{\Lambda}$.

Concerning the second assertion, note that for any $x \in V$ and $k \in [\![2,r]\!]$, the Cauchy-Schwartz inequality implies

$$
\begin{aligned}
\sum_{l\in[\![\gamma_k]\!]}|\alpha_{(l,k)}(x)| &\leqslant \sqrt{\gamma_k}\sqrt{\sum_{l\in[\![\gamma_k]\!]}\alpha^2_{(l,k)}(x)} \\
&\leqslant \sqrt{\Gamma}\sqrt{\sum_{(k',l')\in S\backslash\{(1,1)\}}\alpha^2_{(l',k')}(x)} \\
&= \sqrt{\Gamma}\,\|\alpha_0(x)\|_0 \\
&\leqslant \sqrt{\frac{\Gamma}{\pi(x)\lambda_\wedge}} \\
&\leqslant \sqrt{\frac{\Gamma}{\pi_\wedge\lambda_\wedge}}
\end{aligned}
$$

where (61) was taken into account.

To bound above $Z(\mu_0)$ independently from μ_0, also use the Cauchy-Schwartz inequality: write

$$
\begin{aligned}
Z(\mu_0) &\leqslant \sqrt{|V|}\sqrt{\mu_0\left[\sum_{(k,l)\in S\backslash\{(1,1)\}}\varphi^2_{(k,l)}\right]} \\
&= \sqrt{|V|}\sqrt{\mu_0\left[\|\varphi_0\|^2\right]} \\
&= \sqrt{|V|}\sqrt{\mu_0\left[\|R_0\alpha_0\|^2\right]}
\end{aligned}
$$

where we used that $\varphi_0 = R\alpha_0$, see the sentence after (63). Taking into account (67), we deduce

$$
Z(\mu_0) \leqslant \sqrt{|V|}\sqrt{\frac{\lambda_\vee}{\pi_\wedge}}
$$

This proves that $\bar{M} \geqslant M$ and by consequence $\bar{n}_0 \geqslant n_0$, the first term $2(\Gamma - 1)$ in (82) ensuring that for $n \geqslant 2(\Gamma - 1)$, we have

$$\forall\, l \in [\![\Gamma]\!], \qquad \binom{n}{l-1} \;\leqslant\; \binom{n}{\Gamma-1}$$

Moreover $\mathcal{G}$ is clearly stochastically dominated by $\mathcal{H}_2$, so the desired result follows. ■

6 On continuous time

Here we mention how to adapt to the continuous time setting the previous results obtained for discrete time.

Instead of transition kernels on the finite set V, we now work with Markov generators on V, namely matrices $L := (L(x,y))_{x,y \in V}$ whose off-diagonal entries are non-negative and whose row sums vanish. For such a matrix L, we can find $a \geqslant 0$ and a transition kernel Q_a such that $L = a(Q_a - I)$, where I is the $V \times V$ identity matrix. This decomposition is not unique as there is one for any $a \geqslant a_0$, where

$$a_0 \;\coloneqq\; \max\{|L(x,x)| \,:\, x \in V\}$$

since for positive $a \geqslant a_0$, $\frac{L}{a} + I$ is a Markov kernel.

Given two Markov generators L and $\widetilde{L}$, the notions of corresponding intertwining relation, faithful intertwining relation, bi-intertwining relation and faithful bi-intertwining relation are defined exactly as in the introduction for their transition kernel counterpart. We can even directly relate them: let $a \geqslant 0$ large enough so that we can write

$$L \;=\; a(Q_a - I) \quad \text{and} \quad \widetilde{L} \;=\; a(\widetilde{Q}_a - I) \tag{90}$$

where Q_a and $\widetilde{Q}_a$ are transition kernels. Then the above relations for L and $\widetilde{L}$ are equivalent to the same relations for Q_a and $\widetilde{Q}_a$, with the same links Λ and $\widetilde{\Lambda}$.

The notion of interweaving relation has to be slightly modified, replacing (3) by the existence of a probability q on $\mathbb{R}_+$ such that

$$\Lambda\widetilde{\Lambda} \;=\; \int_{\mathbb{R}_+} \exp(tL)\, q(dt) \tag{91}$$

The notions of faithful interweaving, bi-interweaving, faithful bi-interweaving relations follow accordingly.

Nevertheless, it is no longer so easy to relate interweaving relations for L and $\widetilde{L}$ and those for Q_a and $\widetilde{Q}_a$ appearing in (90). So instead of trying to extend the discrete time results to the continuous time via writings such as (90), we are to go straight back to the proofs, as it is quite simple to adapt them. Below we present the continuous time statements and we just indicate the main modifications that have to been brought to their proofs.

The analogue of Proposition 4 is:

Proposition 16 *Assume that the Markov generators L and $\widetilde{L}$ are irreducible and similar. Then there exists a faithful bi-interweaving relation between them, with equal probability distribution $q = \widetilde{q}$ which can be taken to be a Dirac mass.*

The construction of the links is identical to that given in Section 3. With the notations defined there, they are of the form Λ_b and $\widetilde{\Lambda}_{\widetilde{b}}$ for families of real numbers $b \coloneqq (b_c)_{c \in C\setminus\{(1,1)\}}$ and $\widetilde{b} \coloneqq (\widetilde{b}_c)_{c \in C\setminus\{(1,1)\}}$ belonging to the set $\mathcal{B}$ described in (37), using the number η defined in (36).

A first (little) difference pops up when we try to check (91), with $q = \delta_{t_0}$ for some $t_0 \geqslant 0$, namely we look for families b and $\widetilde{b}$ such that

$$\Lambda_b \widetilde{\Lambda}_{\widetilde{b}} = \exp(t_0 L)$$

As in Section 3, we verify this equality on an adapted basis $(\varphi_{(k,l)})_{(k,l)\in S}$, i.e. such that

$$\forall\ (k,l) \in S, \qquad L[\varphi_{(k,l)}] = -\lambda_k \varphi_{(k,l)} + \varphi_{(k,l-1)}$$

where by convention, $\varphi_{(k,0)} = 0$ for all $k \in [\![r]\!]$. Note that we then have

$$\forall\ (k,l) \in S, \qquad \exp(t_0 L)[\varphi_{(k,l)}] = \exp(-t_0\lambda_k)\left[\varphi_{(k,l)} + t_0\varphi_{(k,l-1)} + \cdots + \frac{t_0^{l-1}}{(l-1)!}\varphi_{(k,1)}\right] \tag{92}$$

It follows that, for any $k \in K$ (the set K was introduced in (40), using the set R defined in (35)), (45) has to be replaced by the system of equations

$$\left\{\begin{array}{rcl} \widetilde{b}_{(k,1)} b_{(k,1)} &=& \exp(-t_0\lambda_k) \\ \widetilde{b}_{(k,1)} b_{(k,2)} + \widetilde{b}_{(k,2)} b_{(k,1)} &=& t_0 \exp(-t_0\lambda_k) \\ \widetilde{b}_{(k,1)} b_{(k,3)} + \widetilde{b}_{(k,2)} b_{(k,2)} + \widetilde{b}_{(k,3)} b_{(k,1)} &=& \frac{t_0^2}{2}\exp(-t_0\lambda_k) \\ &\vdots& \end{array}\right. \tag{93}$$

Looking for a solution of the form $b_{(k,l)} = \exp(-\Re(\lambda_k)(n/2-l+1))\beta_{(k,l)}$ and $\widetilde{b}_{(k,l)} = \exp(-\Re(\lambda_k)(n/2-l+1))$ for $l \in [\![\gamma_k]\!]$, we end up with the following system replacing (46)

$$\left\{\begin{array}{rcl} \beta_{(k,1)} &=& e^{-\mathrm{i}t_0\alpha_k} \\ \beta_{(k,2)} + \beta_{(k,1)} &=& t_0 e^{-\mathrm{i}t_0\alpha_k} \\ \beta_{(k,3)} + \beta_{(k,2)} + \beta_{(k,1)} &=& \frac{t_0^2}{2} e^{-\mathrm{i}t_0\alpha_k} \\ &\vdots& \end{array}\right.$$

with $\alpha_k = \Im(\lambda_k)$. This system admits a unique solution, which satisfies

$$\forall\ l \in [\![\gamma_k]\!], \qquad |\beta_{(k,l)}| \leqslant \sum_{j\in[\![l-1]\!]} \frac{t_0^j}{j!}$$

and we get

$$\max\{|b_{(k,l)}| \vee |\widetilde{b}_{(k,l)}| \,:\, l \in [\![\gamma_k]\!]\} \leqslant \sum_{j\in[\![l-1]\!]} \frac{t_0^j}{j!}\exp(-t_0\Re(\lambda_k))$$

Since all the eigenvalues have a positive real part, except for the eigenvalue 0, it follows that for t_0 large enough, the constructed families b and $\widetilde{b}$ solution of (93) belong to $\mathcal{B}$. This ends the proof of Proposition 16 with $q = \widetilde{q} = \delta_{t_0}$. When all the eigenvalues are assumed to be real, it is possible to get estimates on t_0, as it was done at the end of Section 3.

For the equivalent of Theorem 2, consider L and $\widetilde{L}$ two non-transient Markov generators. We denote by C_1, C_2, ..., C_ℓ (respectively $\widetilde{C}_1$, $\widetilde{C}_2$, ..., $\widetilde{C}_\ell$) the irreducible classes of L (resp. $\widetilde{L}$). They are in the same number $\ell \in \mathbb{N}$ and they are also the irreducible classes of Q_a and $\widetilde{Q}_a$ appearing in (90). For all $l \in [\![\ell]\!] \coloneqq \{1,2,...,\ell\}$, denote L_{C_l} (resp. $\widetilde{L}_{\widetilde{C}_l}$) the restriction of L (resp. $\widetilde{L}$) to C_l (resp. $\widetilde{C}_l$). Note that these matrices are irreducible Markov generators.

Theorem 17 *There exists a faithful bi-interweaving relation between L and $\widetilde{L}$ if and only if there exists a permutation $\sigma \in \mathcal{S}_\ell$ and a probability q on $\mathbb{R}_+$ such that for any $l \in [\![\ell]\!]$, $|C_l| = |\widetilde{C}_{\sigma(l)}|$ and there is a faithful bi-interweaving relation between L_{C_l} and $\widetilde{L}_{\widetilde{C}_{\sigma(l)}}$ with the same probability $\widetilde{q} = q$. It can furthermore be imposed that q is a Dirac mass.*

The proof is identical to that of Theorem 2, since it mainly consists in manipulations of the links. The last assertion comes from the fact that in Proposition 16, any Dirac mass δ_{t_0} with t_0 large enough is allowed, we can thus choose one common t_0 for all the L_{C_l} and $\widetilde{L}_{\widetilde{C}_{\sigma(l)}}$ for $l \in [\![\ell]\!]$.

For the analogue of Theorem 14, we need to introduce corresponding notations. Let L be an irreducible Markov generator whose eigenvalues are real. They are necessarily non-positive, zero being one of them with (algebraic) multiplicity 1. The eigenvalues of $-L$ are denoted

$$0 = \lambda_1 < \lambda_2 \leqslant \lambda_3 \leqslant \cdots \leqslant \lambda_r$$

and to each of the λ_k, $k \in [\![r]\!]$, is associated a Jordan block of size γ_k (so that $\gamma_1 = 1$ and $\sum_{k\in[\![r]\!]} \gamma_k = |V|$). Consider $S := \{(k, l) \,:\, k \in [\![r]\!], l \in [\![\gamma_k]\!]\}$ and let $(\varphi_{(k,l)})_{(k,l)\in S}$ be an adapted basis, namely satisfying

$$\forall\ (k, l) \in S, \qquad L[\varphi_{(k,l)}] \;=\; -\lambda_k \varphi_{(k,l)} + \varphi_{(k,l-1)}$$

where by convention, $\varphi_{(k,0)} = 0$ for all $k \in [\![r]\!]$. As usual, we assume that $\varphi_{(1,1)} = \mathbb{1}$.

Let $X := (X(t))_{t\in\mathbb{R}_+}$ be a Markov process with Markov generator L and initial distribution μ_0, which is fixed from now on. Lemma 13 is still valid so we assume that

$$\forall\ (k, l) \in S, \qquad \mu_0[\varphi_{(k,l)}] \;\geqslant\; 0$$

As in Section 5, we see S as a state space on which we introduce the probability $\widetilde{\mu}_0$ given by

$$\forall\ (k, l) \in S, \qquad \widetilde{\mu}_0((k, l)) \;:=\; \begin{cases} \frac{\mu_0[\varphi_{(k,l)}]}{Z(\mu_0)} & \text{, if } k \geqslant 2 \\ 0 & \text{, if } (k, l) = (1, 1) \end{cases}$$

with $Z(\mu_0) := \sum_{(k,l)\in S\setminus\{1,1)\}} \mu_0[\varphi_{(k,l)}]$.

We furthermore endow S with the simple model Markov generator $\widetilde{L}$ given by

$$\forall\ (k, l) \neq (k', l') \in S, \qquad \widetilde{L}((k, l), (k', l')) \;:=\; \begin{cases} \lambda_k & \text{, if } k = k' \geqslant 2 \text{ and } l' = l - 1 \geqslant 1 \\ \lambda_k & \text{, if } k \geqslant 2,\ l = 1 \text{ and } (k', l') = (1, 1) \\ 0 & \text{, otherwise} \end{cases}$$

Consider $\widetilde{X} := (\widetilde{X}(t))_{t\in\mathbb{R}_+}$ a Markov process with generator $\widetilde{L}$ and initial distribution $\widetilde{\mu}_0$. It ends up being absorbed at $(1, 1)$ after following one of the $r - 1$ rays of the underlying graph. We denote $\mathcal{G}$ the absorption time:

$$\mathcal{G} \;:=\; \inf\{t \in \mathbb{R}_+ \,:\, \widetilde{X}(t) = (1, 1)\}$$

whose law is a mixture of gamma distributions whose scale parameters are (some of) the $1/\lambda_k$, for $k \in [\![2, r]\!]$.

As in (62), consider $S_0 := S\setminus\{(1, 1)\}$ and the Gramian matrix defined by

$$\forall\ (k', l'), (k'', l'') \in S_0 \qquad R_{(k',l'),(k'',l'')} \;:=\; \pi[\varphi_{(k',l')}\varphi_{(k'',l'')}]$$

where π is the invariant probability associated to L.

Recall (see the sentence after (63)) that for any $x \in V$, $\varphi_0(x) \coloneqq (\varphi_{(k,l)}(x))_{(k,l)\in S_0}$ and define $\alpha_0(x) \coloneqq (\alpha_{(k,l)}(x))_{(k,l)\in S_0}$ by $\alpha_0(x) = R_0^{-1}\varphi_0(x)$. Introduce the quantities

$$M \quad \coloneqq \quad Z(\mu_0) \max\left\{ \max_{k\in[\![2,r]\!]} \sum_{l\in[\![\gamma_k]\!]} |\alpha_{(l,k)}(x)| \,:\, x \in V \right\} \tag{94}$$

$$t_0 \quad \coloneqq \quad \min\left\{ t \in \mathbb{R}_+ \,:\, \max_{(k,l)\in S\setminus\{(1,1)\}} \frac{t^{l-1}}{(l-1)!} \exp(-\lambda_k t) \leqslant \frac{1}{M} \right\} \tag{95}$$

(recall that $\pi_\wedge \coloneqq \min\{\pi(x) \,:\, x \in V\}$).

Equally recall the quantities $\lambda_\wedge > 0$ and Γ described before and inside Lemma 11.

Define

$$\bar{M} \quad \coloneqq \quad \frac{\sqrt{|V|\Gamma}}{\pi_\wedge}\sqrt{\frac{\lambda_\vee}{\lambda_\wedge}}$$

$$\bar{t}_0 \quad \coloneqq \quad (\Gamma - 1) \vee \min\left\{ t \in \mathbb{R}_+ \,:\, \frac{t^{\Gamma-1}}{(\Gamma-1)!} \exp(-\lambda_2 t) \leqslant \frac{1}{\bar{M}} \right\} \tag{96}$$

Here is the analogue of Theorem 14 for continuous time:

Theorem 18 *Assume that L is irreducible and that its eigenvalues are all real. Then there exists a strong stationary time for X which is stochastically dominated by*

$$t_0 + \mathcal{G}$$

This random variable is itself stochastically dominated by $\bar{t}_0 + \mathcal{H}_2$, where $\mathcal{H}_2$ is the convolution of a gamma distribution of shape Γ and scale $1/\lambda_2$.

The underlying discrete time considerations of Diaconis and Fill [1] (see the first part of the proof of Theorem 5) have to be replaced by their continuous time analogues of Fill [2]. In addition, the construction from the adapted bases of the links Λ and $\widetilde{\Lambda}_{\widetilde{b}}$, given a real-valued family $\widetilde{b} \coloneqq (\widetilde{b}_{(k,l)})_{(k,l)\in S}$ with $\widetilde{b}_{(1,1)} = 1$ and satisfying (88), is done as in (84) and (86). To get the interweaving relation, we have to replace (89) by (92), which leads to

$$\forall\ (k,l) \in S, \qquad \widetilde{b}_{(k,l)} \quad = \quad Z(\mu_0)\frac{t_0^{l-1}}{(l-1)!}\exp(-t_0\lambda_k)$$

The choice of t_0 through (94) and (95) ensures us again that (88) is satisfied.

Finally the last assertion of Theorem 18 is proven in exactly the same way as that of Theorem 5, the first term $\Gamma - 1$ in (96) ensuring that for $t \geqslant \Gamma - 1$,

$$\forall\ l \in [\![\Gamma]\!], \qquad \frac{t^{l-1}}{(l-1)!} \quad \leqslant \quad \frac{t^{\Gamma-1}}{(\Gamma-1)!}$$

References

[1] Persi Diaconis and James Allen Fill. Strong stationary times via a new form of duality. *Ann. Probab.*, 18(4):1483–1522, 1990.

[2] James Allen Fill. Strong stationary duality for continuous-time Markov chains. I. Theory. *J. Theoret. Probab.*, 5(1):45–70, 1992.

[3] Peter Matthews. Strong stationary times and eigenvalues. *J. Appl. Probab.*, 29(1):228–233, 1992.

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

[5] Laurent Miclo and Pierre Patie. On interweaving relations. *J. Funct. Anal.*, 280(3):54, 2021. Id/No 108816.

[6] L. Chris G. Rogers and Jim W. Pitman. Markov functions. *Ann. Probab.*, 9(4):573–582, 1981.

miclo@math.cnrs.fr

Toulouse School of Economics,
1, Esplanade de l'université,
31080 Toulouse cedex 6, France.
Institut de Mathématiques de Toulouse,
Université Paul Sabatier, 118, route de Narbonne,
31062 Toulouse cedex 9, France.