STATIONARY MEASURES ON INFINITE GRAPHS

ALEXANDRE BARAVIERA, PEDRO DUARTE, AND MARIA JOANA TORRES

Abstract. We extend the theory of isospectral reductions of L. Bunimovich and B. Webb to infinite graphs, and describe an application of this extension to the problems of existence and approximation of stationary measures on infinite graphs.

Keywords: *Isospectral graph reduction, infinite graphs, stochastic operator, stationary measures*

2000 Mathematics Subject Classification: 05C63, 47A75, 47D07

Original publication:
Communications in Contemporary Mathematics, **24** (7), (2022), 2150025
https://doi.org/10.1142/S0219199721500255

1. Introduction

The spectral theory of graphs is an active area of research in modern mathematical physics (see, for example, [3, 8], and the references therein). This is due in part to problems concerning the theory itself but also to many applications found in biology, economics and social sciences, computer science and the theory of information and communication (see, for example [9, 14]).

The method of isospectral graph transformations introduced in [5, 4, 6] provides a way of understanding the interplay between the topology of a network (considered as a weighted graph) and its dynamics. More precisely, the authors introduce a concept of transformation of a graph (either by reduction or expansion) with the key property of preserving part of the spectrum of the graph's adjacency matrix. More recently in [11] the second and third authors proved that isospectral graph reductions also preserve the eigenvectors associated with the eigenvalues of the graph's weighted adjacency matrix. The results of this paper also explain how to reconstruct an eigenvector of the graph's adjacency matrix from an eigenvector of the reduced matrix. Because the spectral approach in [11] to isospectral graph reduction theory was based on *eigenvectors*, instead of *eigenvalues*, it was a natural question to ask about possible generalizations of this theory to infinite dimensions. Notice that in finite dimensions the eigenvalues are the zeros of the characteristic polynomial, a meaningless concept in infinite dimensions where eigenvalues correspond to the existence of non-trivial solutions of the eigenvalue equation. In [12] the isospectral reduction theory was generalized to infinite dimensions.

In this paper we provide a more general extension to infinite dimensions of the isospectral reduction theory (see Theorem 3). This new result opens the way to the main purpose of this work which is to prove the existence of stationary measures for certain stochastic processes defined on a class of infinite graphs (see Theorem 4). More precisely, we propose a method based on iterated isospectral reductions (see Section 6) to approximate the stationary measure (which is proved to exist) that converges in a super-exponential way, which is faster than the exponential speed of convergence of the classical Perron-Frobenius method.

The paper is organized as follows:

In Section 2 we describe the theory of isospectral graph reductions and the reduction statements for finite graphs. In Section 3 we extend the finite dimensional version of this theory, developed in [12], to reductions over arbitrary sets. In Section 4 we define Markov chains and introduce the class of birth-and-death processes to which our results apply. In Section 5 we introduce some auxiliary dynamical systems that play a key role in the proof of the main results leading to the existence of stationary measures for certain classes of infinite graphs. In Section 6 we introduce a class of tridiagonal stochastic infinite graphs and prove the existence of stationary measures for this class, providing a method based on iterated isospectral reductions to approximate the stationary measure. In Section 7 we apply previous results to the lattice $\mathbb{Z}^2$ and to the Bethe lattice. Finally, in Section 8 we describe an isospectral reduction-reconstruction algorithm and use a numerical example to compare its execution times with those of a standard algorithm.

2. Isospectral reduction theory on finite graphs

A key concept in Bunimovich-Webb's isospectral theory is that of a *structural set*. It specifies the class of subsets of the vertex set of a graph over which an isospectral reduction (expansion) is performed, allowing one to investigate how the structure of a graph is affected by an isospectral transformation. A key concept in [11] to perform the reconstruction of eigenvectors of the original matrix from the eigenvectors of the reduced matrix, recursively, is the *depth* of a vertex. The coordinates of the eigenvector on vertices of depth n are explicitly given (see item (3) of Theorem 1) in terms of the values of the eigenvector at vertices of depth $< n$.

Given a finite set V we denote by $\mathbb{C}^V$ the finite dimensional complex space of all functions $f\colon V \to \mathbb{C}$ which can be represented as vectors $(f(i))_{i\in V}$. Similarly, a vector $w \in \mathbb{C}^{V\times V}$ will be described as a $V \times V$ matrix and represented as a list $(w(i,j))_{i,j\in V}$.

Consider a *finite weighted graph*, i.e., a pair $G = (V, w)$ where the *vertex set* V is finite and $w\colon V \times V \to \mathbb{C}$ is any function (called the weight function of G). We make the convention that $j \to i$ is an edge of G iff $w(i, j) \neq 0$. [1] Denote by $\mathcal{A} = \mathcal{A}_w\colon \mathbb{C}^V \to \mathbb{C}^V$ the operator defined by the weighted adjacency matrix $(w(i,j))_{i,j\in V}$.

[1]Notice that for instance in [5, 4, 6, 11] the authors make the opposite convention.

A *path* $\gamma = (i_0, \ldots, i_p)$ in the graph $G = (V, w)$ is an ordered sequence of vertices $i_0, \ldots, i_p \in V$ such that $w(i_\ell, i_{\ell+1}) \neq 0$ for $0 \leq \ell \leq p-1$. The integer p is called the *length* of γ. If the vertices $i_0, \ldots, i_{p-1}$ are all distinct the path γ is called *simple*. If $i_0 = i_p$ then γ is called a *closed path*. A closed path of length 1 is called a *loop*. Finally, we call any simple closed path a *cycle*. If $S \subseteq V$ we will write $\bar{S} := V \setminus S$.

Definition 2.1. *Let $G = (V, w)$. A nonempty vertex set $S \subseteq V$ is a* structural set *for G if each cycle of G, that is not a loop, contains a vertex in S.*

Given a structural set S, a *branch of* (G, S) is any simple path $\beta = (i_0, i_1, \ldots, i_{p-1}, i_p)$ such that $i_1, \ldots, i_{p-1} \in \bar{S}$ and $i_0, i_p \in S$. Denote by $\mathcal{B}_{ij}$ the set of all branches that start in i and end in j. Define $\Sigma := \{w(i,i) : i \in \bar{S}\}$ and let $\lambda \in \mathbb{C} \setminus \Sigma$. For each branch $\beta = (i_0, i_1, \ldots, i_p)$ we define the λ-*weight of* β as follows:

$$w(\beta, \lambda) := w(i_0, i_1) \prod_{\ell=1}^{p-1} \frac{w(i_\ell, i_{\ell+1})}{\lambda - w(i_\ell, i_\ell)} .$$

Given $i, j \in S$ set

$$R_{S,\lambda}(i,j) := \sum_{\beta \in \mathcal{B}_{ij}} w(\beta, \lambda) .$$

The *reduced operator* $\mathcal{R}_S(\lambda) \colon \mathbb{C}^S \to \mathbb{C}^S$ is given by the matrix $(R_{S,\lambda}(i,j))_{i,j \in S}$.

We can also view the reduced operator $\mathcal{R}_S$ as an $S \times S$ matrix with entries in the field of rational functions $f(\lambda) = \frac{p(\lambda)}{q(\lambda)}$, where $p(\lambda)$ and $q(\lambda)$ are polynomials in a formal variable λ. We then define the *spectrum* of $\mathcal{R}_S(\lambda)$, denoted by $\mathrm{sp}(\mathcal{R}_S)$, to be

$$\mathrm{sp}(\mathcal{R}_S) := \{\lambda \in \mathbb{C} \setminus \Sigma \colon \det(\mathcal{R}_S(\lambda) - \lambda I) = 0\} .$$

Definition 2.2. *Assuming S is a structural set, the* depth *of a vertex is defined recursively as follows: a vertex $i \in S$ has depth 0 and a vertex $i \in \bar{S}$ has depth n iff i hasn't depth $< n$, and $w(i,j) \neq 0$ implies j has depth $< n$, for all $j \in V$.*

One can see that the depth of a vertex $i \in V$ is the length of the longest path in $\bar{S}$ from a vertex in S to i (all the nodes in this path, except the first one, must be in $\bar{S}$).

From now on, given $u \in \mathbb{C}^V$, we denote by u_S the restriction of u to S, i.e., $u_S = (u_i)_{i \in S}$. For finite graphs, results concerning isospectral reductions can be stated as follows.

Theorem 1. *Let S be a structural set of the graph $G = (V, w)$. Then for the associated operator $\mathcal{A} = \mathcal{A}_w$*

1. $\mathrm{sp}(\mathcal{A}) \setminus \Sigma = \mathrm{sp}(\mathcal{R}_S)$.
2. *If $\lambda_0 \in \mathbb{C} \setminus \Sigma$ is an eigenvalue of $\mathcal{A}$ and $u \in \mathbb{C}^V$ is an associated eigenvector, $\mathcal{A}\, u = \lambda_0\, u$, then $\mathcal{R}_S(\lambda_0)\, u_S = \lambda_0\, u_S$.*

(3) *If $\lambda_0 \in \mathbb{C} \setminus \Sigma$ is an eigenvalue of $\mathcal{R}_S(\lambda_0)$ and $v = (v_i)_{i\in S}$ is an associated eigenvector, $\mathcal{R}_S(\lambda_0)\, v = \lambda_0\, v$, then the following recursive relations*

$$\begin{cases} u_i = v_i \quad \text{for } \; i \in S_0 = S \\ \\ u_\ell = \displaystyle\sum_{j\in S_{n-1}} \frac{w(\ell,j)}{\lambda_0 - w(\ell,\ell)}\, u_j \quad \text{for all } \; \ell \in S_n \setminus S_{n-1} \end{cases}$$

uniquely determine an eigenvector u of $\mathcal{A}$ associated with λ_0, where S_n denotes the set of all vertices of depth $\leq n$.

Item (1) corresponds to a simplified [2] version of Bunimovich-Webb isospectral reduction theorem (see [5, Theorem 3.5.]), and items (2) and (3) correspond to [11, Theorem 1 and Proposition 2.1], respectively.

Theorem 1 was extended to infinite dimensions in [12, Theorem 3.12, Theorem 3.8, Theorem 3.18].

In [16] and [6, §1] the theory of isospectral reduction is extended to arbitrary sets not necessarily structural. More precisely, given a finite weighted graph $G = (V, w)$, consider an arbitrary set $S \subseteq V$. Recall that $\mathbb{C}^V$ denotes the space of all functions $f : V \to \mathbb{C}$ and identify $\mathbb{C}^S$, resp. $\mathbb{C}^{\bar S}$, as subspaces of $\mathbb{C}^V$ consisting of functions $f : V \to \mathbb{C}$ which vanish outside S, resp. $\bar S$. With these identifications we have $\mathbb{C}^V = \mathbb{C}^S \oplus \mathbb{C}^{\bar S}$. Let $\pi_S : \mathbb{C}^V \to \mathbb{C}^S$ and $\pi_{\bar S} : \mathbb{C}^V \to \mathbb{C}^{\bar S}$ be the canonical projections. We define the component operators of $\mathcal{A} = \mathcal{A}_w \colon \mathbb{C}^V \to \mathbb{C}^V$ by

$$\begin{aligned} &\mathcal{A}_{SS} \colon \mathbb{C}^S \to \mathbb{C}^S, &&\mathcal{A}_{SS} = \pi_S \circ \mathcal{A} \circ \pi_S, \\ &\mathcal{A}_{\bar S S} \colon \mathbb{C}^S \to \mathbb{C}^{\bar S}, &&\mathcal{A}_{\bar S S} = \pi_{\bar S} \circ \mathcal{A} \circ \pi_S, \\ &\mathcal{A}_{S \bar S} \colon \mathbb{C}^{\bar S} \to \mathbb{C}^S, &&\mathcal{A}_{S \bar S} = \pi_S \circ \mathcal{A} \circ \pi_{\bar S}, \\ &\mathcal{A}_{\bar S \bar S} \colon \mathbb{C}^{\bar S} \to \mathbb{C}^{\bar S}, &&\mathcal{A}_{\bar S \bar S} = \pi_{\bar S} \circ \mathcal{A} \circ \pi_{\bar S}. \end{aligned}$$

Next we introduce the *reduced operator* on $\mathbb{C}^S$, $\mathcal{R}(\lambda) = \mathcal{R}_S(\lambda) : \mathbb{C}^S \to \mathbb{C}^S$,

$$\mathcal{R}_S(\lambda) := \mathcal{A}_{SS} - \mathcal{A}_{S\bar S}\, (\mathcal{A}_{\bar S \bar S} - \lambda\, I)^{-1}\, \mathcal{A}_{\bar S S}$$

which is well defined for $\lambda \in \mathbb{C} \setminus \mathrm{sp}(\mathcal{A}_{\bar S \bar S})$. Note that the reduced operator $\mathcal{R}(\lambda)$ is the so called *Schur complement* of $\mathcal{A}_{\bar S \bar S} - \lambda\, I$ plus $\lambda\, I$.

Since we are in a finite dimensional context, all these operators and projections can naturally be represented by matrices. In particular, $\mathcal{A}_{SS}$, $\mathcal{A}_{\bar S S}$, $\mathcal{A}_{S\bar S}$ and $\mathcal{A}_{\bar S \bar S}$ can be regarded as submatrices of $\mathcal{A}$.

[2]This statement corresponds to [5, Corollary 3] where the adjacency matrix has complex entries.

A simplified [3] version of the isospectral reduction theorem in this setting (see [16, Theorem 2.1.], [6, Theorem 1.1.]), which is a consequence of Schur complement's determinant formula, can be stated as follows:

Theorem 2. *Let $G = (V, w)$, $\mathcal{A} = \mathcal{A}_w$ and $\emptyset \neq S \subsetneq V$. Then,*

$$\operatorname{sp}(\mathcal{A}) \setminus \operatorname{sp}(\mathcal{A}_{\bar{S}\bar{S}}) = \operatorname{sp}(\mathcal{R}_S).$$

Example 2.1. *Consider the graph $G = (V, w)$ given in Figure 1. Recall that $w(i, j)$ is the weight of the edge from j to i. The operator $\mathcal{A} = \mathcal{A}_w$ is given by the (column)-stochastic block tridiagonal matrix*

$$\mathcal{A} = \left[\begin{array}{c:cc:c} \frac{11}{12} & \frac{3}{4} & \frac{3}{4} & 0 \\ \hdashline \frac{1}{12} & 0 & \frac{1}{6} & 0 \\ 0 & \frac{1}{6} & 0 & \frac{3}{4} \\ \hdashline 0 & \frac{1}{12} & \frac{1}{12} & \frac{1}{4} \end{array}\right].$$

The set $S = \{i_1, i_4\}$ is not a structural set because the nonloop cycle i_2, i_3 does not contain a vertex in S. The component operators of $\mathcal{A}$ relative to S are given by

$$\mathcal{A}_{SS} = \begin{bmatrix} \frac{11}{12} & 0 \\ 0 & \frac{1}{4} \end{bmatrix}, \ \mathcal{A}_{\bar{S}S} = \begin{bmatrix} \frac{1}{12} & 0 \\ 0 & \frac{3}{4} \end{bmatrix}, \ \mathcal{A}_{S\bar{S}} = \begin{bmatrix} \frac{3}{4} & \frac{3}{4} \\ \frac{1}{12} & \frac{1}{12} \end{bmatrix}, \ \mathcal{A}_{\bar{S}\bar{S}} = \begin{bmatrix} 0 & \frac{1}{6} \\ \frac{1}{6} & 0 \end{bmatrix}.$$

Clearly, $\operatorname{sp}(\mathcal{A}_{\bar{S}\bar{S}}) = \{-\frac{1}{6}, \frac{1}{6}\}$ *and, therefore, the reduced operator $\mathcal{R}_S(\lambda)$ is well defined for all $\lambda \in \mathbb{C} \setminus \{-\frac{1}{6}, \frac{1}{6}\}$ by*

$$\mathcal{R}_S(\lambda) = \begin{bmatrix} \frac{11}{12} & 0 \\ 0 & \frac{1}{4} \end{bmatrix} - \begin{bmatrix} \frac{3}{4} & \frac{3}{4} \\ \frac{1}{12} & \frac{1}{12} \end{bmatrix} \cdot \begin{bmatrix} -\lambda & \frac{1}{6} \\ \frac{1}{6} & -\lambda \end{bmatrix}^{-1} \cdot \begin{bmatrix} \frac{1}{12} & 0 \\ 0 & \frac{3}{4} \end{bmatrix} = \begin{bmatrix} \frac{13-132\lambda}{24-144\lambda} & -\frac{27}{8-48\lambda} \\ \frac{1}{24(-1+6\lambda)} & \frac{1+12\lambda}{-8+48\lambda} \end{bmatrix}.$$

The spectrum of $\mathcal{A}$ is $\operatorname{sp}(\mathcal{A}) = \{1, \frac{5}{12}, -\frac{1}{6}, -\frac{1}{12}\}$ *while the spectrum of $\mathcal{R}_S(\lambda)$ is* $\operatorname{sp}(\mathcal{R}_S) = \{1, \frac{5}{12}, -\frac{1}{12}\} = \operatorname{sp}(\mathcal{A}) \setminus \operatorname{sp}(\mathcal{A}_{\bar{S}\bar{S}})$.

The cost of greater generality in this extension to arbitrary sets (not necessarily structural) is a less explicit reconstruction formula involving inverses which can only be represented as sums of infinite series (see [16, equation 15 on p. 157], [6, equation 5.6 on p. 137]). Moreover, from a graph theoretical perspective, the graph's path and cycle structure associated with structural sets provide a natural combinatorial interpretation which is lost in this extension.

[3] This statement corresponds to [16, Corollary 2.1.], [6, Corollary 1.1.] where the adjacency matrix has complex entries.

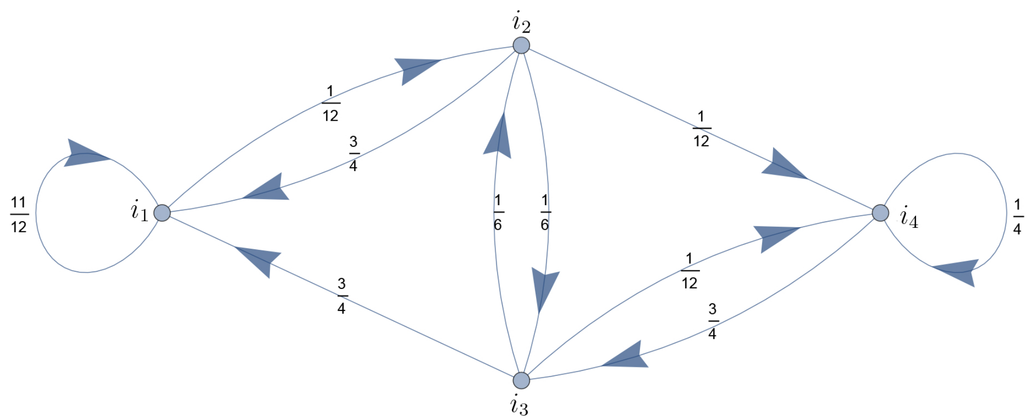

FIGURE 1. A stochastic finite graph.

3. Extension of Isospectral reduction theory to infinite dimension

In this section we extend the infinite dimension theory developed in [12] for structural sets to arbitrary sets. In particular, we extend Theorem 2 to infinite dimensions. This extension allows for possible applications to nonlinear problems which can be reduced to the analysis of infinite dimensional linear operators.

Let $\mathcal{E}$ be a Banach space with direct sum decomposition $\mathcal{E} = \mathcal{S} \oplus \bar{\mathcal{S}}$ into closed subspaces $\mathcal{S}$ and $\bar{\mathcal{S}}$. Let $\pi_{\mathcal{S}} : \mathcal{E} \to \mathcal{S}$ and $\pi_{\bar{\mathcal{S}}} : \mathcal{E} \to \bar{\mathcal{S}}$ be the canonical projections associated with this decomposition.

Let $\mathcal{T} : \mathcal{E} \to \mathcal{E}$ be a bounded operator. We define the component operators

$$\begin{aligned}
\mathcal{T}_{\mathcal{S}\mathcal{S}} &: \mathcal{S} \to \mathcal{S}, & \mathcal{T}_{\mathcal{S}\mathcal{S}} &= \pi_{\mathcal{S}} \circ \mathcal{T} \circ \pi_{\mathcal{S}}, \\
\mathcal{T}_{\bar{\mathcal{S}}\mathcal{S}} &: \mathcal{S} \to \bar{\mathcal{S}}, & \mathcal{T}_{\bar{\mathcal{S}}\mathcal{S}} &= \pi_{\bar{\mathcal{S}}} \circ \mathcal{T} \circ \pi_{\mathcal{S}}, \\
\mathcal{T}_{\mathcal{S}\bar{\mathcal{S}}} &: \bar{\mathcal{S}} \to \mathcal{S}, & \mathcal{T}_{\mathcal{S}\bar{\mathcal{S}}} &= \pi_{\mathcal{S}} \circ \mathcal{T} \circ \pi_{\bar{\mathcal{S}}}, \\
\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} &: \bar{\mathcal{S}} \to \bar{\mathcal{S}}, & \mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} &= \pi_{\bar{\mathcal{S}}} \circ \mathcal{T} \circ \pi_{\bar{\mathcal{S}}}.
\end{aligned}$$

We denote by $\mathrm{sp}(\mathcal{T})$ the *spectrum* of the operator $\mathcal{T}$, i.e.,

$$\mathrm{sp}(\mathcal{T}) := \{\lambda \in \mathbb{C} : \mathcal{T} - \lambda I \ \text{ is not invertible } \}\,.$$

Recall that a complex number $\lambda \in \mathbb{C}$ is called an *eigenvalue* of $\mathcal{T}$ if the equation $(\mathcal{T} - \lambda I)u = 0$ has non-zero solutions $u \in \mathcal{E}$, which are referred to as the *eigenvectors* of $\mathcal{T}$ associated with λ. The eigenvalues of $\mathcal{T}$ are the elements of the spectrum such that $\mathcal{T} - \lambda I$ fails to be injective. In infinite dimensions, the spectrum may also contain elements $\lambda \in \mathbb{C}$ such that $\mathcal{T} - \lambda I$ is injective but fails to be surjective.

Next we introduce the family of *reduced operators* on $\mathcal{S}$, $\mathcal{R}(\lambda) = \mathcal{R}_{\mathcal{T},\mathcal{S},\bar{\mathcal{S}}}(\lambda) : \mathcal{S} \to \mathcal{S}$,

$$\mathcal{R}(\lambda) := \mathcal{T}_{\mathcal{S}\mathcal{S}} - \mathcal{T}_{\mathcal{S}\bar{\mathcal{S}}}\,(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} - \lambda\, I)^{-1}\,\mathcal{T}_{\bar{\mathcal{S}}\mathcal{S}} \tag{3.1}$$

defined for $\lambda \in \mathbb{C} \setminus \mathrm{sp}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}})$.

Definition 3.1. *The* spectrum *of the family of operators $\mathcal{R}(\lambda)$ is the set*

$$\mathrm{sp}(\mathcal{R}) := \{\lambda \in \mathbb{C} : \mathcal{R}(\lambda) - \lambda I \quad \textit{is not invertible}\ \}\,.$$

Given $u \in \mathcal{E}$, we denote by $u_{\mathcal{S}}$ the projection $u_{\mathcal{S}} := \pi_{\mathcal{S}}(u)$. Similarly we write $u_{\bar{\mathcal{S}}} := \pi_{\bar{\mathcal{S}}}(u)$.

The following result extends theorems 1 and 2 to this infinite dimensional setting.

Theorem 3. *Let $\mathcal{T} : \mathcal{E} \to \mathcal{E}$ be a bounded operator on a Banach space $\mathcal{E}$ with a direct sum decomposition $\mathcal{E} = \mathcal{S} \oplus \bar{\mathcal{S}}$. Then*

(1) $\mathrm{sp}(\mathcal{T}) \setminus \mathrm{sp}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}}) = \mathrm{sp}(\mathcal{R})$.
(2) *Given $\lambda_0 \in \mathbb{C} \setminus \mathrm{sp}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}})$, λ_0 is an eigenvalue of $\mathcal{T}$ iff λ_0 is an eigenvalue of $\mathcal{R}(\lambda_0)$.*
(3) *If $\lambda_0 \in \mathbb{C} \setminus \mathrm{sp}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}})$ is an eigenvalue of $\mathcal{T}$ and $u \in \mathcal{E}$ is an associated eigenvector, $\mathcal{T}\,u = \lambda_0\,u$, then $\mathcal{R}(\lambda_0)\,u_{\mathcal{S}} = \lambda_0\,u_{\mathcal{S}}$.*
(4) *If $\lambda_0 \in \mathbb{C}\setminus\mathrm{sp}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}})$ is an eigenvalue of $\mathcal{R}(\lambda_0)$ and $v \in \mathcal{S}$ is an associated eigenvector, $\mathcal{R}(\lambda_0)\,v = \lambda_0\,v$, then the following relations*

$$\begin{cases} u_{\mathcal{S}} = v \in \mathcal{S} \\[2ex] u_{\bar{\mathcal{S}}} = -(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} - \lambda_0 I)^{-1}\,\mathcal{T}_{\bar{\mathcal{S}}\mathcal{S}}\,u_{\mathcal{S}} \in \bar{\mathcal{S}} \end{cases} \tag{3.2}$$

uniquely determine a reconstructed eigenvector $u = u_{\mathcal{S}} + u_{\bar{\mathcal{S}}} \in \mathcal{E}$ of $\mathcal{T}$, i.e., such that $\mathcal{T}\,u = \lambda_0\,u$.

Proof. First we prove item (3) and the direct implication in (2). Let $\lambda_0 \in \mathbb{C} \setminus \mathrm{sp}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}})$ be an eigenvalue of $\mathcal{T}$ and let $u = u_{\mathcal{S}} + u_{\bar{\mathcal{S}}} \in \mathcal{E}$ be an associated eigenvector. Since $\mathcal{T}\,u = \lambda_0\,u$, one has

$$\mathcal{T}_{\mathcal{S}\mathcal{S}}\,u_{\mathcal{S}} + \mathcal{T}_{\mathcal{S}\bar{\mathcal{S}}}\,u_{\bar{\mathcal{S}}} = \lambda_0\,u_{\mathcal{S}}, \tag{3.3}$$

and

$$\mathcal{T}_{\bar{\mathcal{S}}\mathcal{S}}\,u_{\mathcal{S}} + \mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}}\,u_{\bar{\mathcal{S}}} = \lambda_0\,u_{\bar{\mathcal{S}}}\,. \tag{3.4}$$

From equation (3.4) we get

$$u_{\bar{\mathcal{S}}} = -(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} - \lambda_0 I)^{-1}\,\mathcal{T}_{\bar{\mathcal{S}}\mathcal{S}}\,u_{\mathcal{S}}\,. \tag{3.5}$$

Observe that the operator $\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} - \lambda_0 I$ is invertible, given that $\lambda_0 \notin \mathrm{sp}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}})$. Since

$$(\mathcal{R}(\lambda_0) - \lambda_0 I)\,u_{\mathcal{S}} = \mathcal{T}_{\mathcal{S}\mathcal{S}}\,u_{\mathcal{S}} - \mathcal{T}_{\mathcal{S}\bar{\mathcal{S}}}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} - \lambda_0 I)^{-1}\mathcal{T}_{\bar{\mathcal{S}}\mathcal{S}}\,u_{\mathcal{S}} - \lambda_0\,u_{\mathcal{S}}\,,$$

using (3.5), we get

$$(\mathcal{R}(\lambda_0) - \lambda_0 I)\,u_{\mathcal{S}} = \mathcal{T}_{\mathcal{S}\mathcal{S}}\,u_{\mathcal{S}} + \mathcal{T}_{\mathcal{S}\bar{\mathcal{S}}}\,u_{\bar{\mathcal{S}}} - \lambda_0\,u_{\mathcal{S}}\,,$$

and then, using (3.3), we obtain that

$$(\mathcal{R}(\lambda_0) - \lambda_0 I)\, u_{\mathcal{S}} = 0\,.$$

This proves that λ_0 is an eigenvalue of $\mathcal{R}(\lambda_0)$ with $\mathcal{R}(\lambda_0)\, u_{\mathcal{S}} = \lambda_0\, u_{\mathcal{S}}$.

Next we prove (4) and the converse implication in (2). Let $\lambda_0 \in \mathbb{C} \setminus \mathrm{sp}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}})$ be an eigenvalue of $\mathcal{R}(\lambda_0)$ and $v \in \mathcal{S}$ be an associated eigenvector. Since $\mathcal{R}(\lambda_0)\, v = \lambda_0\, v$, one has

$$\mathcal{T}_{\mathcal{S}\mathcal{S}}\, v - \mathcal{T}_{\mathcal{S}\bar{\mathcal{S}}}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} - \lambda_0 I)^{-1}\mathcal{T}_{\bar{\mathcal{S}}\mathcal{S}}\, v = \lambda_0\, v\,. \tag{3.6}$$

Let $u = u_{\mathcal{S}} + u_{\bar{\mathcal{S}}} \in \mathcal{E}$ be defined by

$$\begin{cases} u_{\mathcal{S}} := v \\ \\ u_{\bar{\mathcal{S}}} := -(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} - \lambda_0 I)^{-1}\, \mathcal{T}_{\bar{\mathcal{S}}\mathcal{S}}\, u_{\mathcal{S}} \end{cases} \tag{3.7}$$

From the definition of $u_{\bar{\mathcal{S}}}$ we obtain

$$\mathcal{T}_{\bar{\mathcal{S}}\mathcal{S}}\, u_{\mathcal{S}} + \mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}}\, u_{\bar{\mathcal{S}}} = \lambda_0\, u_{\bar{\mathcal{S}}}\,. \tag{3.8}$$

On the other hand, using the definition of $u_{\mathcal{S}}$ and $u_{\bar{\mathcal{S}}}$ in equation (3.6), we obtain

$$\mathcal{T}_{\mathcal{S}\mathcal{S}}\, u_{\mathcal{S}} + \mathcal{T}_{\mathcal{S}\bar{\mathcal{S}}}\, u_{\bar{\mathcal{S}}} = \lambda_0\, u_{\mathcal{S}}\,. \tag{3.9}$$

Together, equations (3.8) and (3.9) are equivalent to $\mathcal{T}\, u = \lambda_0\, u$. This proves that λ_0 is an eigenvalue of $\mathcal{T}$ with associated reconstructed eigenvector $u \in \mathcal{E}$ defined by (3.7).

Finally we prove (1).

Given $\lambda_0 \notin \mathrm{sp}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}})$, assume that $\lambda_0 \notin \mathrm{sp}(\mathcal{R})$. We first prove that the operator $\mathcal{T} - \lambda_0\, I$ is surjective, i.e., that given $f = f_{\mathcal{S}} + f_{\bar{\mathcal{S}}} \in \mathcal{E}$ there exists $u = u_{\mathcal{S}} + u_{\bar{\mathcal{S}}} \in \mathcal{E}$ such that

$$(\mathcal{T} - \lambda_0 I)\, u = f\,,$$

i.e., such that

$$\begin{cases} (\mathcal{T}_{\mathcal{S}\mathcal{S}} - \lambda_0 I)\, u_{\mathcal{S}} + \mathcal{T}_{\mathcal{S}\bar{\mathcal{S}}}\, u_{\bar{\mathcal{S}}} = f_{\mathcal{S}} \\ \\ \mathcal{T}_{\bar{\mathcal{S}}\mathcal{S}}\, u_{\mathcal{S}} + (\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} - \lambda_0 I)\, u_{\bar{\mathcal{S}}} = f_{\bar{\mathcal{S}}}\,. \end{cases} \tag{3.10}$$

Since $\lambda_0 \notin \mathrm{sp}(\mathcal{R})$, the operator $\mathcal{R}(\lambda_0) - \lambda_0\, I$ is invertible, hence surjective. Thus, given $f_{\mathcal{S}} - \mathcal{T}_{\mathcal{S}\bar{\mathcal{S}}}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} - \lambda_0 I)^{-1}\, f_{\bar{\mathcal{S}}} \in \mathcal{S}$ there exists $u_{\mathcal{S}} \in \mathcal{S}$ such that

$$(\mathcal{R}(\lambda_0) - \lambda_0\, I)\, u_{\mathcal{S}} = f_{\mathcal{S}} - \mathcal{T}_{\mathcal{S}\bar{\mathcal{S}}}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} - \lambda_0 I)^{-1}\, f_{\bar{\mathcal{S}}}\,. \tag{3.11}$$

Let

$$u_{\bar{\mathcal{S}}} := (\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} - \lambda_0 I)^{-1}(f_{\bar{\mathcal{S}}} - \mathcal{T}_{\bar{\mathcal{S}}\mathcal{S}}\, u_{\mathcal{S}}) \in \bar{\mathcal{S}}\,. \tag{3.12}$$

Clearly, equation (3.12) defining $u_{\bar{\mathcal{S}}}$ is equivalent to the second equation in (3.10). On the other hand, equation (3.11) is equivalent to

$$\mathcal{T}_{\mathcal{S}\mathcal{S}}\, u_{\mathcal{S}} - \mathcal{T}_{\mathcal{S}\bar{\mathcal{S}}}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} - \lambda_0 I)^{-1}\mathcal{T}_{\bar{\mathcal{S}}\mathcal{S}}\, u_{\mathcal{S}} - \lambda_0\, u_{\mathcal{S}} = f_{\mathcal{S}} - \mathcal{T}_{\mathcal{S}\bar{\mathcal{S}}}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} - \lambda_0 I)^{-1} f_{\bar{\mathcal{S}}}\,,$$

which in turn is equivalent to

$$(\mathcal{T}_{\mathcal{S}\mathcal{S}} - \lambda_0 I)\, u_{\mathcal{S}} + \mathcal{T}_{\mathcal{S}\bar{\mathcal{S}}}\, \underbrace{(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} - \lambda_0 I)^{-1}(f_{\bar{\mathcal{S}}} - \mathcal{T}_{\bar{\mathcal{S}}\mathcal{S}}\, u_{\mathcal{S}})}_{u_{\bar{\mathcal{S}}}} = f_{\mathcal{S}}\,,$$

which at last, by (3.12), is equivalent to the first equation in (3.10). All together, we have proved that the operator $\mathcal{T} - \lambda_0 I$ is surjective. On the other hand, by (2), this operator must be injective. Hence $\mathcal{T} - \lambda_0 I$ is invertible, which implies that $\lambda_0 \notin \operatorname{sp}(\mathcal{T})$.

Conversely, assume that $\lambda_0 \notin \operatorname{sp}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}})$ but $\lambda_0 \in \operatorname{sp}(\mathcal{R})$, so that $\mathcal{R}(\lambda_0) - \lambda_0\, I$ is not invertible. By the Open Mapping Theorem this operator is either non-injective or non-surjective.

If $\mathcal{R}(\lambda_0) - \lambda_0\, I$ is not injective then λ_0 is an eigenvalue of $\mathcal{R}(\lambda_0)$ and by (4) it is also an eigenvalue of $\mathcal{T}$. Therefore $\lambda_0 \in \operatorname{sp}(\mathcal{T}) \setminus \operatorname{sp}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}})$.

Otherwise, if $\mathcal{R}(\lambda_0) - \lambda_0\, I$ is not surjective we can choose $f_{\mathcal{S}} \in \mathcal{S}$ such that the equation $(\mathcal{R}(\lambda_0) - \lambda_0\, I)\, u_{\mathcal{S}} = f_{\mathcal{S}}$ has no solutions $u_{\mathcal{S}} \in \mathcal{S}$. Take $f_{\bar{\mathcal{S}}} = 0$, let $f = f_{\mathcal{S}} = f_{\mathcal{S}} + f_{\bar{\mathcal{S}}}$ and consider the equation $(\mathcal{T} - \lambda_0\, I)\, u = f$. This equation may be written in the form (3.10). Any solution of (3.10) gives rise to a solution of (3.11), that is a solution of the equation $(\mathcal{R}(\lambda_0) - \lambda_0\, I)\, u_{\mathcal{S}} = f_{\mathcal{S}}$, which is not possible. Hence the operator $\mathcal{T} - \lambda_0\, I$ is non-invertible and, therefore, $\lambda_0 \in \operatorname{sp}(\mathcal{T}) \setminus \operatorname{sp}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}})$.

□

We introduce the family of *reconstruction operators* $\Phi_{\mathcal{T}} = \Phi_{\mathcal{T},\mathcal{S},\bar{\mathcal{S}}} \colon \mathbb{C} \setminus \operatorname{sp}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}}) \to \mathcal{L}(\mathcal{S}, \mathcal{E})$ that to each $\lambda \in \mathbb{C} \setminus \operatorname{sp}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}})$ and $v \in \mathcal{S}$ associates the unique function $u = \Phi_{\mathcal{T}}(\lambda)(v) \in \mathcal{E}$ defined by (3.2). This operator takes values in the space $\mathcal{L}(\mathcal{S}, \mathcal{E})$ of bounded linear maps $L : \mathcal{S} \to \mathcal{E}$. These reconstruction operators will play a key role in the statement and proof of the main theorem (Theorem 4 in Section 6).

In the infinite dimension theory developed in [12] for structural sets, the authors were able to reconstruct the eigenfunctions for the so called *structural sets of type B* (see [12, Definition 3.5]), through recursive relations (see (3.7) in [12]), which are much simpler than (3.2) in Theorem 3. Off course, when dealing with structural sets of type B, the two reconstruction formulas coincide.

4. Markov chains, birth and death processes

A *Markov chain* is a stochastic process $\{X_i : \Omega \to E\}_{i\in\mathbb{N}}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values on a finite or countable space E such that

$$\mathbb{P}(X_{n+1} = j \,|\, X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \mathbb{P}(X_{n+1} = j \,|\, X_n = i_n)\,,$$

that is, the state X_{n+1} of the system at time $n+1$ depends only on the state X_n at time n. When the probability $\mathbb{P}(X_{n+1} = j \,|\, X_n = i)$ is independent of n, the Markov chain is said to have *stationary transition probabilities*. A Markov chain with stationary transition probabilities can be described by a (column) stochastic matrix $P = (p_{ij})_{i,j\in E}$, where p_{ij} denotes the transition probability from j to i.

A *stationary probability measure* for a Markov chain is the common distribution of a stationary Markov chain with stationary probability transitions. It is represented by a vector $\pi = (\pi_j)_{j\in E}$ such that $\pi_j \geq 0$, $\sum_{j\in E} \pi_j = 1$, and $P\pi = \pi$.

One of the motivations for the main result in Section 6 is a class of random walks on the set of states $E = \{0, 1, 2, \ldots\}$ also known as the discrete time birth-and-death process, where the state space can be interpreted as the size of a certain population (for more details see [13]). It is interesting to remark that many problems of the real world can be modeled by birth-and-death processes: frequently cited examples include problems in evolutionary biology, ecology, population genetics, epidemiology and queuing theory (see [2, 7, 15] and references therein). The most simple example of this class can be defined as follows: the transition probability from the state j to the state i is given by p_{ij} where $p_{ii} = \delta_i$, $p_{i+1,i} = b_i$ and $p_{i,i+1} = c_{i+1}$ (where $i \in \{0, 1, \ldots\}$), satisfying

$$\delta_0 + b_0 = 1, \delta_i + b_i + c_i = 1 \quad \text{for any } i \in \{1, 2, \ldots\}$$

i.e.,

$$P = (p_{ij}) = \begin{bmatrix} \delta_0 & c_1 & 0 & \ldots \\ b_0 & \delta_1 & c_2 & \ldots \\ 0 & b_1 & \delta_2 & \ldots \\ 0 & 0 & b_2 & \ldots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

is a column stochastic tridiagonal matrix.

In the particular case where $b_i = b$ and $c_i = c$ for all i, with $b < c$, it is possible to obtain the stationary probability π explicitly as $\pi_i = \left(\frac{b}{c}\right)^i \pi_0$ for $i \in \{0, 1, \ldots\}$, just writing the equations for the components of π in a recursive way. This procedure can be generalized and the reader can easily check that the vector whose components are

$$\pi_j = \frac{b_0 b_1 \ldots b_{j-1}}{c_1 c_2 \ldots c_j} \pi_0 \quad \text{for } j \geq 1 \tag{4.1}$$

satisfy $P\pi = \pi$. Hence if

$$\sum_{j=0} \pi_j < +\infty$$

this vector can be normalized and then there exists an invariant probability for this process, say, a measure that is stationary for this stochastic process.

Inspired by this class of models one can consider the more general situation where each state $i \in \{0, 1, \ldots\}$ is replaced by a finite set of states Γ_i, but keeping the interactions among states stratified, in the sense that a point in Γ_i can only connect with other points in Γ_i, Γ_{i-1} or Γ_{i+1}. Rephrasing this idea, we consider a graph whose vertices are $\bigcup_{i\geq 0} \Gamma_i$ and whose edges (i, j) correspond to the following three cases (where $k \in \mathbb{N}$):

- $i, j \in \Gamma_k$
- $i \in \Gamma_k$ and $j \in \Gamma_{k+1}$
- $i \in \Gamma_{k+1}$ and $j \in \Gamma_k$

With this set of states, the transition probabilities are now described by matrices Δ_i, B_i and C_i (for the precise definition, see Section 6) instead of the real parameters δ_i, b_i and c_i. One can still consider the question of the existence (and uniqueness) of a stationary state; an adapted version of (4.1) seems meaningless (since the matrices C_i are not necessarily square, and even if they are square they do not need to be invertible), so an exact expression for the stationary probability is not known. Here we bypass this difficulty using isospectral reduction theory.

5. Auxiliary Dynamics

In this section we introduce some auxiliary dynamical systems which will play a key role in the proof of the main results leading to the existence of stationary measures for a class $\mathcal{W}$ of tridiagonal stochastic infinite graphs. The strategy is based on the fact that the isospectral reduction of a stochastic graph in $\mathcal{W}$ can be viewed as a graph in $\mathcal{W}$, which determines a nonlinear isospectral reducing map $\mathcal{R}\colon \mathcal{W} \to \mathcal{W}$. In the next section we introduce a projection $(\beta, \gamma) : \mathcal{W} \to \mathbb{R}^2$ that to each stochastic graph in $\mathcal{W}$ associates a pair of parameters $(\beta, \gamma) \in W \subset \mathbb{R}^2$ and a map $F : W \to W$ on the space W of parameters (β, γ). The projection $(\beta, \gamma) : \mathcal{W} \to W$ partially semi-conjugates the isospectral reducing map $\mathcal{R}\colon \mathcal{W} \to \mathcal{W}$ to $F : W \to W$ (see Proposition 6.4). This will provide a sort of Lyapunov function to control the dynamics of map $\mathcal{R}\colon \mathcal{W} \to \mathcal{W}$.

Consider the map $f : \mathbb{R}\backslash\{1\} \to \mathbb{R}$ defined by $f(x) := \frac{x^2}{(1-x)^2}$ (see Figure 2).

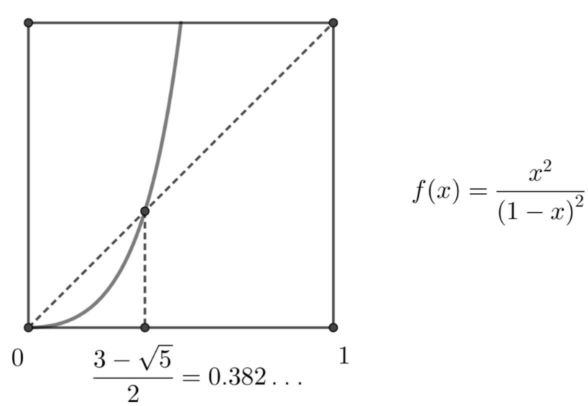

Figure 2. A trapping interval $[0, \frac{3-\sqrt{5}}{2}]$.

The map f has a super attractive fixed point at the origin, where $f(0) = f'(0) = 0$, and a repelling fixed point at $x = \frac{3-\sqrt{5}}{2}$. Hence we have:

Proposition 5.1. *For any* $0 \leq x < \frac{3-\sqrt{5}}{2}$, $\lim_{n\to\infty} f^n(x) = 0$, *with quadratic convergence.*

Definition 5.1. *Let* $W := \{(\beta, \gamma) \colon 0 \leq \beta < 1,\ 0 < \gamma < 1,\ \frac{\beta}{\gamma^2} < \frac{3-\sqrt{5}}{2}\}$ *and* $F : W \to \mathbb{R}^2$ *be the map*

$$F(\beta, \gamma) := \left(\frac{\beta^2}{\gamma}, \gamma - \frac{\beta}{\gamma} \right).$$

Next proposition shows that the function F induces a dynamical system on the space of parameters W. The dynamics of F will be used to control, by comparison, the dynamics of the isospectral reducing map $\mathcal{R} \colon \mathcal{W} \to \mathcal{W}$.

Proposition 5.2. *If* $(\beta, \gamma) \in W$ *and* $(\beta^*, \gamma^*) = F(\beta, \gamma)$ *then*

$$\frac{\beta^*}{(\gamma^*)^2} \leq f(\beta/\gamma^2) < \frac{3-\sqrt{5}}{2}.$$

In particular, $F(W) \subseteq W$.

Proof. Let $(\beta^*, \gamma^*) = F(\beta, \gamma)$. Then

$$\frac{\beta^*}{(\gamma^*)^2} = \frac{\beta^2}{\gamma} \, \frac{1}{(\gamma - \frac{\beta}{\gamma})^2} = \frac{\beta^2}{\gamma^3} \, \frac{1}{(1 - \frac{\beta}{\gamma^2})^2} \leq \frac{\beta^2}{\gamma^4} \, \frac{1}{(1 - \frac{\beta}{\gamma^2})^2} = f(\beta/\gamma^2) < \frac{3-\sqrt{5}}{2},$$

where the last inequality holds because of Proposition 5.1 and the observations that precede it. Since $(\beta, \gamma) \in W$ we have that $\beta < \gamma^2$ with $\gamma \in (0, 1)$ which implies that $\beta^2 < \gamma^4 < \gamma$. Hence $\beta^* \in [0, 1)$.

Because $\beta \in [0, 1)$ we have that $\beta/\gamma \geq 0$ which implies that $\gamma^* < 1$. On the other hand, $\gamma^* > 0$ is equivalent to $\beta < \gamma^2$, which holds because $(\beta, \gamma) \in W$. Hence $\gamma^* \in (0, 1)$. $\square$

By the previous proposition the map $F : W \to W$ defines a dynamical system on the set W (see Figure 3).

Definition 5.2. *Given two maps* $F : X \to X$ *and* $f : Y \to Y$ *where* $(Y, \leq)$ *is a partially ordered set, any function* $h : X \to Y$ *such that* $h \circ F \leq f \circ h$ *is called a partial semi-conjugacy between* F *and* f.

Consider the function $\varphi : W \to \mathbb{R}$, $\varphi(\beta, \gamma) := \beta/\gamma^2$.

Corollary 5.3. $\varphi \circ F \leq f \circ \varphi$, *i.e., the function* $\varphi : W \to \mathbb{R}$ *is a partial semi-conjugacy between the maps* $F : W \to W$ *and* $f : [0, \frac{3-\sqrt{5}}{2}] \to [0, \frac{3-\sqrt{5}}{2}]$.

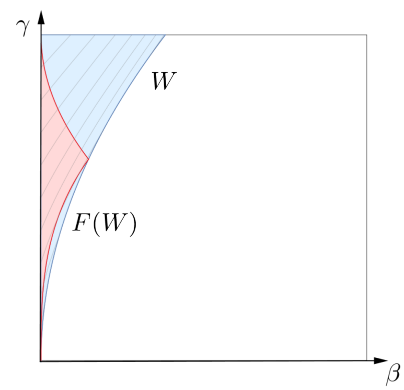

FIGURE 3. This picture represents the region W and its image $F(W)$.

Proof. Let $(\beta^*, \gamma^*) = F(\beta, \gamma)$. Then, using Proposition 5.2,

$$(\varphi \circ F)(\beta, \gamma) = \varphi(\beta^*, \gamma^*) = \frac{\beta^*}{(\gamma^*)^2} \leq f(\frac{\beta}{\gamma^2}) = (f \circ \varphi)(\beta, \gamma)\,.$$

□

Corollary 5.4. $\varphi \circ F^n \leq f^n \circ \varphi$, *for all* $n \in \mathbb{N}$.

Proposition 5.5. *For all* $\gamma \in (0,1)$, $(0,\gamma)$ *is a super-attractive fixed point of* $F : W \to W$ *with Jacobian matrix* $J_F(0,\gamma) = \begin{pmatrix} 0 & 0 \\ -\frac{1}{\gamma} & 1 \end{pmatrix}$.

Proof. Straightforward calculation. □

Finally we describe the dynamics of the control map $F\colon W \to W$.

Proposition 5.6. *For every* $(\beta, \gamma) \in W$ *there exists a unique fixed point* $(0, c) \in W$ *such that* $F^n(\beta, \gamma)$ *converges quadratically to* $(0, c)$.

Proof. Define $(\beta_n, \gamma_n) := F^n(\beta, \gamma)$. Then

$$c := \gamma \prod_{j=0}^{\infty} \left(1 - \frac{\beta_j}{\gamma_j^2}\right) \geq \gamma \prod_{j=0}^{\infty} \left(1 - f^j\left(\frac{\beta}{\gamma^2}\right)\right) > 0 \tag{5.1}$$

because, by Corollary 5.4 and Proposition 5.1, the $\frac{\beta_j}{\gamma_j^2} \leq f^j(\frac{\beta}{\gamma^2})$ decays quadratically to 0. By induction we get for all $n \geq 1$

$$\gamma_n = \gamma \prod_{j=0}^{n-1} \left(1 - \frac{\beta_j}{\gamma_j^2}\right).$$

Thus $c = \lim_{n\to\infty} \gamma_n$ with $c \leq \gamma_n$ and quadratic convergence.

Moreover, since $\beta_n \leq f^n(\beta/\gamma^2)$, the sequence β_n converges quadratically to 0. □

Remark 5.1. *The function $h : W \to \mathbb{R}$, $c = h(\beta, \gamma)$ defined in* (5.1) *is F-invariant. Its levels sets are the flow lines of the dynamics of F, depicted in Figure 3.*

6. Tridiagonal Stochastic Infinite Graphs

In this section we introduce a class of tridiagonal stochastic infinite graphs for which the existence and uniqueness of a stationary measure is proven. The proof exploits the dynamics of a nonlinear isospectral reduction operator, acting on some appropriate space of tridiagonal stochastic graphs, and provides a super-exponential numerical scheme to approximate the stationary measure.

Let $\mathbb{N}$ be the set of non-negative integers and let $\ell^1(\mathbb{N})$ be the Banach space of summable sequences of real numbers endowed with the norm $\|x\|_1 := \sum_{i\in\mathbb{N}} |x_i|$. Consider a stochastic operator $\mathcal{T} : \ell^1(\mathbb{N}) \to \ell^1(\mathbb{N})$ determined by a block tridiagonal matrix

$$\mathcal{T} = \begin{bmatrix} \Delta_0 & C_1 & 0 & 0 & \cdots \\ B_0 & \Delta_1 & C_2 & 0 & \cdots \\ 0 & B_1 & \Delta_2 & C_3 & \cdots \\ 0 & 0 & B_2 & \Delta_3 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix} \tag{6.1}$$

where $\{d_k\}_{k\geq 0}$ is a sequence of positive integers, $\Delta_k \in \mathrm{Mat}_{d_k\times d_k}(\mathbb{R})$, $B_k \in \mathrm{Mat}_{d_{k+1}\times d_k}(\mathbb{R})$ for $k \geq 0$ and $C_k \in \mathrm{Mat}_{d_{k-1}\times d_k}(\mathbb{R})$ for $k \geq 1$ are sub-stochastic matrices.

We say that a matrix $A \in \mathrm{Mat}_{d\times d'}(\mathbb{R})$ is *sub-stochastic*, resp. *strictly sub-stochastic*, if it has non-negative entries and the sum of each column is ≤ 1, resp. < 1. When the sum of each column is 1 we say that the matrix is *stochastic*. For an operator $\mathcal{T}\colon \ell^1(\mathbb{N}) \to \ell^1(\mathbb{N})$, represented by an infinite matrix $(t_{ij})_{i,j\in\mathbb{N}}$ with non-negative entries, we say that $\mathcal{T}$ is stochastic if $\sum_{i\in\mathbb{N}} t_{ij} = 1$, for all $j \in \mathbb{N}$.

Consider the associated weighted graph $G = (\mathbb{N}, \mathcal{T})$ which can be described as follows. Partition $\mathbb{N}$ as the union of a sequence of intervals $\{I_{d_0}, I_{d_1}, \ldots\}$, where each I_{d_k} is the integer interval with d_k elements defined by

$$I_{d_k} := \left[\sum_{j=0}^{k-1} d_j, -1 + \sum_{j=0}^{k} d_j \right].$$

Notice that $I_{d_0} = \{0, 1, \ldots, d_0-1\}$, $I_{d_1} = \{d_0, d_0+1, \ldots, d_0+d_1-1\}$, etc. Let $\Gamma_k = (I_{d_k}, \Delta_k)$ for $k \geq 0$, and consider the *graph sequence* $(\Gamma_k)_{k\geq 0}$. The graph G can be viewed as the union of the graphs Γ_k, where the matrix B_k describes the transitions from Γ_k to Γ_{k+1} ($k \geq 0$), the matrix C_k describes the transitions from Γ_k to Γ_{k-1} ($k \geq 1$) and Δ_k represents the internal transitions on Γ_k. No other transitions exist (see Figure 4).

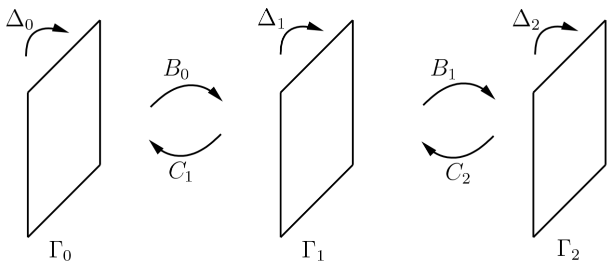

FIGURE 4. The first three subgraphs of the sequence $(\Gamma_k)_{k\geq 0}$.

Consider the structural set $S = I_{d_0} \cup I_{d_2} \cup \cdots$ which corresponds to the vertices of the graphs Γ_{2k}, $k \geq 0$.

Let $\mathcal{E} = \ell^1(\mathbb{N})$, $\mathcal{S} = \ell^1(S)$ and $\bar{\mathcal{S}} = \ell^1(\bar{S})$. The subspace $\mathcal{S}$, resp. $\bar{\mathcal{S}}$, will be regarded as a subspace of $\mathcal{E}$ by extending its elements as 0 outside $\mathcal{S}$, resp. $\bar{\mathcal{S}}$. With these identifications we have $\mathcal{E} = \mathcal{S} \oplus \bar{\mathcal{S}}$.

The *reduced operator* on $\mathcal{S}$ for $\lambda = 1$, $\mathcal{T}^* := \mathcal{R}_{\mathcal{T},\mathcal{S},\bar{\mathcal{S}}}(1) : \mathcal{S} \to \mathcal{S}$ (see (3.1))

$$\mathcal{T}^* = \mathcal{T}_{\mathcal{S}\mathcal{S}} - \mathcal{T}_{\mathcal{S}\bar{\mathcal{S}}} \left(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} - I\right)^{-1} \mathcal{T}_{\bar{\mathcal{S}}\mathcal{S}} \tag{6.2}$$

is well defined when $1 \notin \mathrm{sp}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}})$.

Making the canonical identification $\bar{S} \equiv \mathbb{N}$, the operator $\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}}$ is represented by the matrix

$$\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}} = \begin{pmatrix} \Delta_1 & 0 & 0 & \cdots \\ 0 & \Delta_3 & 0 & \cdots \\ 0 & 0 & \Delta_5 & \cdots \\ \vdots & \vdots & & \ddots \end{pmatrix}.$$

In the sequel we introduce a class of tridiagonal stochastic operators $\mathcal{W}$ such that $1 \notin \mathrm{sp}(\mathcal{T}_{\bar{\mathcal{S}}\bar{\mathcal{S}}})$ for all $\mathcal{T} \in \mathcal{W}$.

Remark 6.1. *The stochastic reductions defined by equation (6.2) were used in a finite dimensional setting in* [1] *where a method was introduced for predicting the formation or the detection of unobserved links in real-world networks, referred to as the method of* effective transitions*. This method relies on the theory of isospectral matrix reductions to compute the probability of eventually transitioning from one vertex to another in a (biased) random walk on the network.*

In the sequel we will be using the following notation: On the Euclidean space $\mathbb{R}^d$ we consider the sum norm

$$\|x\| := \sum_i |x_i|$$

and for a matrix $A = [a_{ij}]$ we consider the matrix norm

$$\|A\| := \max_j \sum_i |a_{ij}|.$$

Notice that these norms satisfy $\|A\,x\| \leq \|A\|\,\|x\|$. These relations also hold for infinite matrices with the obvious adaptations of the norm's definitions. Let $\mathbb{1} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}$ denote any array with all entries equal to 1. With this notation a matrix A with non-negative entries is stochastic, resp. sub-stochastic, if $\mathbb{1}\,A = \mathbb{1}$, resp. $\mathbb{1}\,A \leq \mathbb{1}$. The stochastic character of $\mathfrak{T}$ is equivalent to $\mathbb{1}\,\Delta_0 + \mathbb{1}\,B_0 = \mathbb{1}$ and $\mathbb{1}\,C_k + \mathbb{1}\,\Delta_k + \mathbb{1}\,B_k = \mathbb{1}$ for all $k \geq 1$.

The norm $\|\mathfrak{T}\|$ of the stochastic operator $\mathfrak{T}$ defined by (6.1) is clearly 1.

We assume that the operator $\mathfrak{T}$ defined by (6.1) and its block matrices satisfy

(a) all matrices Δ_k, B_k and C_k have non-negative entries,
(b) $\mathbb{1}\,\Delta_0 + \mathbb{1}\,B_0 = \mathbb{1}$,
(c) $\mathbb{1}\,C_k + \mathbb{1}\,\Delta_k + \mathbb{1}\,B_k = \mathbb{1}$, for all $k \geq 1$,
(d) $\|B_k\| \leq \beta$, for all $k \geq 0$,
(e) $\mathbb{1}\,C_k \geq \gamma\,\mathbb{1}$, for all $k \geq 1$,

for some constants $0 \leq \beta < 1$ and $0 < \gamma < 1$ such that

$$\frac{\beta}{\gamma^2} < \frac{3-\sqrt{5}}{2} \tag{6.3}$$

Let $\mathcal{W}$ denote the space of all operators $\mathfrak{T}$ satisfying (a)-(e) for some $(\beta,\gamma) \in W$. Notice the sequence of dimensions $\{d_k\}$ is not fixed in the definition of $\mathcal{W}$, which means that different operators in $\mathcal{W}$ may be associated to sequences of graphs with different dimensions.

Conditions (a)-(c) imply that $\mathfrak{T}$ is a stochastic operator. In particular, $\|\mathfrak{T}\| = 1$.

Condition (d) says that β is an upper bound for the column' sums of all matrices B_k, while condition (e) says that γ is a lower bound for every column' sum of all matrices C_k. In particular, $C_k \neq 0$, for all $k \geq 1$. Hence, we always have transitions from the graph Γ_k to the graph Γ_{k-1} for all $k \geq 1$. Conditions (c) and (e) ensure that all matrices Δ_k and B_k are strictly sub-stochastic for all $k \geq 1$. In particular $\mathfrak{T}_{\bar{\mathfrak{s}},\bar{\mathfrak{s}}}$ is a strictly substochastic operator with $\|\mathfrak{T}_{\bar{\mathfrak{s}},\bar{\mathfrak{s}}}\| < 1$. Hence the reduced operator (6.2) is well defined.

We can represent the operator $\mathfrak{T}$ by a stochastic matrix $(t_{ij})_{i,j\in\mathbb{N}}$ and consider the Markov chain with state space $\mathbb{N}$ in which the transition from j to i has probability t_{ij}. A fixed point $\mathfrak{T}q = q \in \ell^1(\mathbb{N})$ with $\sum_{j\in\mathbb{N}} q_j = 1$ and $q_j \geq 0$ for all $j \in \mathbb{N}$ is called a *stationary probability measure* of $\mathfrak{T}$. A stationary probability measure is called *ergodic* if it is an extremal point of the compact convex set of all stationary probability measures.

We say that j *leads* to i if there exists $m \geq 1$ such that $t_{ij}^m > 0$, where t_{ij}^m is the (i,j)-entry of $\mathfrak{T}^m$. Two states i and j *communicate* when i leads to j and j leads to i. The *recurrent set* is defined as the set of all states $i \in \mathbb{N}$ such that i leads to i. This set can be split into equivalence classes, each class being formed by states that communicate with each other. The set of all these classes is then partially ordered as follows: $C_1 \geq C_2$ if i_1 leads to i_2 for some $i_1 \in C_1$ and $i_2 \in C_2$. At the bottom of this hierarchy are the *essential* classes. More precisely, a class C is called essential if for every $i \in C$, if i leads to j then j leads to i.

Recall that a matrix $A \in \mathrm{Mat}_{d\times d}(\mathbb{R})$ with non-negative entries is *primitive* when there exists $n \geq 1$ such that all entries of A^n are strictly positive.

The next proposition establishes the uniqueness of the stationary probability measure for an operator $\mathcal{T} \in \mathcal{W}$. The less trivial issue of existence of a stationary probability measure will be dealt with latter (see Corollary 6.17).

Proposition 6.1. *If $\mathcal{T} \in \mathcal{W}$ and Δ_0 is primitive, then $\mathcal{T}$ has at most one stationary probability measure whose support contains I_{d_0}.*

Proof. From the Theory of Markov Chains (see [10, Theorem 5.7]) each ergodic stationary probability measure on $\mathbb{N}$ is associated with an essential class of states in $\mathbb{N}$.

Since Δ_0 is primitive all states in I_{d_0} are recurrent and communicate among themselves, i.e., they are contained in the same class. On the other hand, since $\mathbb{1}\, C_k \geq \gamma\, \mathbb{1}$ all states in $\mathbb{N}$ lead to a state in I_{d_0}. Hence all states in I_{d_0} are essential. Moreover every essential class must contain a state in I_{d_0}. Therefore there exists a unique essential class, which matches the support of a unique ergodic stationary measure. This implies that there is at most one stationary probability measure. $\square$

Next we describe the reduction $\mathcal{T}^*$ in (6.2) of an operator $\mathcal{T} \in \mathcal{W}$. We make the identification $S = \mathbb{N}$ which formally corresponds to defining $\mathcal{T}^* := \Psi \circ \mathcal{R}_{\mathcal{T},\ell^1(S),\ell^1(S)}(1) \circ \Psi^{-1}$, where $\Psi : \ell^1(S) \to \ell^1(\mathbb{N})$ is the bounded linear isomorphism $(\Psi u)(n) := u(2n)$.

Proposition 6.2. *Given $\mathcal{T} \in \mathcal{W}$ the reduced operator is*

$$\mathcal{T}^* = \begin{bmatrix} \Delta_0^* & C_1^* & 0 & 0 & \cdots \\ B_0^* & \Delta_1^* & C_2^* & 0 & \cdots \\ 0 & B_1^* & \Delta_2^* & 0 & \cdots \\ 0 & 0 & B_2^* & 0 & \cdots \\ \vdots & \vdots & \vdots & & \ddots \end{bmatrix}$$

with

$$\begin{aligned} \Delta_0^* &= \Delta_0 + C_1\, (I - \Delta_1)^{-1}\, B_0 \\ \Delta_k^* &= \Delta_{2k} + C_{2k+1}\, (I - \Delta_{2k+1})^{-1}\, B_{2k} + B_{2k-1}\, (I - \Delta_{2k-1})^{-1}\, C_{2k} \\ B_k^* &= B_{2k+1}\, (I - \Delta_{2k+1})^{-1}\, B_{2k} \\ C_k^* &= C_{2k-1}\, (I - \Delta_{2k-1})^{-1}\, C_{2k}\, . \end{aligned}$$

Proof. Follows from the definition of the reduced operator in (6.2). □

In the following proposition we note that isospectral reduction preserves the stochastic character of an operator $\mathcal{T} \in \mathcal{W}$.

Proposition 6.3. *If $\mathcal{T} \in \mathcal{W}$ then $\mathcal{T}^*$ is stochastic.*

Proof. Since $(I - \Delta_k)^{-1} = \sum_{n=0}^{\infty} \Delta_k^n$, all these matrices have non-negative entries. It follows that Δ_k^*, B_k^* and C_k^* have also non-negative entries. Using Proposition 6.2 and that $\mathbb{1}\, C_k + \mathbb{1}\, B_k = \mathbb{1} - \mathbb{1}\, \Delta_k = \mathbb{1}\, (I - \Delta_k)$ for all $k \geq 1$, we have that

$$\begin{aligned}
\mathbb{1}\, C_k^* + \mathbb{1}\, \Delta_k^* + \mathbb{1}\, B_k^* = \;& \mathbb{1}\, C_{2k-1}\, (I - \Delta_{2k-1})^{-1}\, C_{2k} + \mathbb{1}\, \Delta_{2k} + \mathbb{1}\, C_{2k+1}\, (I - \Delta_{2k+1})^{-1}\, B_{2k} \\
& + \mathbb{1}\, B_{2k-1}\, (I - \Delta_{2k-1})^{-1}\, C_{2k} + \mathbb{1}\, B_{2k+1}\, (I - \Delta_{2k+1})^{-1}\, B_{2k} \\
= \;& (\mathbb{1}\, C_{2k-1} + \mathbb{1}\, B_{2k-1})\, (I - \Delta_{2k-1})^{-1}\, C_{2k} + \mathbb{1}\, \Delta_{2k} \\
& + (\mathbb{1}\, C_{2k+1} + \mathbb{1}\, B_{2k+1})\, (I - \Delta_{2k+1})^{-1}\, B_{2k} \\
= \;& \mathbb{1}\, (I - \Delta_{2k-1})\, (I - \Delta_{2k-1})^{-1}\, C_{2k} + \mathbb{1}\, \Delta_{2k} \\
& + \mathbb{1}\, (I - \Delta_{2k+1})\, (I - \Delta_{2k+1})^{-1}\, B_{2k} \\
= \;& \mathbb{1}\, C_{2k} + \mathbb{1}\, \Delta_{2k} + \mathbb{1}\, B_{2k} = \mathbb{1}\,.
\end{aligned}$$

Analogously, we can check that $\mathbb{1}\, \Delta_0^* + \mathbb{1}\, B_0^* = \mathbb{1}$. □

We introduce a couple of measurements $\beta, \gamma : \mathcal{W} \to \mathbb{R}$ of an operator $\mathcal{T} \in \mathcal{W}$:

$$\beta(\mathcal{T}) := \inf\{b > 0 \colon\; \mathbb{1}\, B_k \leq b\mathbb{1},\, \forall k \geq 0\,\}$$
$$\gamma(\mathcal{T}) := \sup\{c > 0 \colon\, \mathbb{1}\, C_k \geq c\, \mathbb{1},\, \forall k \geq 1\}$$

By the definition of $\mathcal{W}$, these are the components of a joint function $(\hat{\beta}, \hat{\gamma}) : \mathcal{W} \to W$, where the set W was introduced in Definition 5.1.

We call *reducing map* to $\mathcal{R} \colon \mathcal{W} \to \mathcal{L}(\ell^1(\mathbb{N}), \ell^1(\mathbb{N}))$, defined by $\mathcal{R}(\mathcal{T}) := \mathcal{T}^*$. In the following proposition we prove that the reducing map takes values in $\mathcal{W}$, thus inducing a dynamical system $\mathcal{R} \colon \mathcal{W} \to \mathcal{W}$. We also introduce a partial order on the space $\mathcal{W}$ for which the operator $\mathcal{R} \colon \mathcal{W} \to \mathcal{W}$ is partially semi-conjugated to the maps $F \colon W \to W$ and $f \colon [0, \frac{3-\sqrt{5}}{2}] \to [0, \frac{3-\sqrt{5}}{2}]$. This will allow us to control the dynamics of the reduction procedure.

Proposition 6.4. *If $\mathcal{T} \in \mathcal{W}$ then*

$$\frac{\beta(\mathcal{T}^*)}{\gamma(\mathcal{T}^*)^2} \leq f\left(\frac{\beta(\mathcal{T})}{\gamma(\mathcal{T})^2}\right) < \frac{3-\sqrt{5}}{2}.$$

In particular $\mathcal{T}^* \in \mathcal{W}$.

Proof. Let $\beta = \hat\beta(\mathcal{T})$ and $\gamma = \hat\gamma(\mathcal{T})$. Because $\mathbb{1}\, C_k \geq \gamma\, \mathbb{1}$ we have

$$\|\Delta_k\| = \|\mathbb{1}\, \Delta_k\|_\infty \leq 1 - \gamma$$

and hence

$$\|(I - \Delta_k)^{-1}\| \leq \frac{1}{1 - \|\Delta_k\|} \leq \frac{1}{1-(1-\gamma)} = \frac{1}{\gamma}.$$

Since

$$B_k^* = B_{2k+1}\,(I - \Delta_{2k+1})^{-1}\, B_{2k}$$

we get

$$\|B_k^*\| \leq \frac{\beta^2}{\gamma}.$$

Thus $\beta^* := \beta(\mathcal{T}^*) \leq \frac{\beta^2}{\gamma}$.

Since

$$C_k^* = C_{2k-1}\,(I - \Delta_{2k-1})^{-1}\, C_{2k}$$

from (c) we get

$$\begin{aligned}
\mathbb{1}\, C_k^* &= \mathbb{1}\, C_{2k-1}\,(I - \Delta_{2k-1})^{-1}\, C_{2k} \\
&= ((\mathbb{1}\,(I - \Delta_{2k-1}) - \mathbb{1}\, B_{2k-1})\,(I - \Delta_{2k-1})^{-1}\, C_{2k} \\
&= \mathbb{1}\, C_{2k} - \mathbb{1}\, B_{2k-1}\,(I - \Delta_{2k-1})^{-1}\, C_{2k} \\
&\geq \gamma\, \mathbb{1} - \|\mathbb{1}\, B_{2k-1}\,(I - \Delta_{2k-1})^{-1}\, C_{2k}\|_\infty\, \mathbb{1}.
\end{aligned}$$

On the other hand

$$\begin{aligned}
\|\mathbb{1}\, B_{2k-1}\,(I - \Delta_{2k-1})^{-1}\, C_{2k}\|_\infty &= \|B_{2k-1}\,(I - \Delta_{2k-1})^{-1}\, C_{2k}\| \\
&\leq \|B_{2k-1}\|\, \|(I - \Delta_{2k-1})^{-1}\|\, \|C_{2k}\| \\
&\leq \frac{\beta}{\gamma}\,.
\end{aligned}$$

Thus

$$\mathbb{1}\, C_k^* \geq (\gamma - \frac{\beta}{\gamma})\, \mathbb{1}$$

and $\gamma^* := \gamma(\mathcal{T}^*) \geq \gamma - \frac{\beta}{\gamma}$. From the inequalities $\beta^* \leq \frac{\beta^2}{\gamma}$ and $\gamma^* \geq \gamma - \frac{\beta}{\gamma}$ and using Proposition 5.2, we deduce that

$$\frac{\beta(\mathcal{T}^*)}{\gamma(\mathcal{T}^*)^2} = \frac{\beta^*}{(\gamma^*)^2} \leq f\left(\frac{\beta(\mathcal{T})}{\gamma(\mathcal{T})^2}\right) < \frac{3-\sqrt{5}}{2}.$$

In particular, $\mathcal{T}^* \in \mathcal{W}$. $\square$

We denote by $\mathcal{R}^n$ the n-th iterate of the map $\mathcal{R}\colon \mathcal{W} \to \mathcal{W}$, i.e., $\mathcal{R}^n := \mathcal{R} \circ \cdots \circ \mathcal{R}$ (composition with n factors) for $n \geq 0$.

Consider the partial order relation on W defined by $(\beta_1, \gamma_1) \lesssim (\beta, \gamma)$ if $\beta_1 \leq \beta$ and $\gamma_1 \geq \gamma$.

Corollary 6.5. *$(\hat{\beta}, \hat{\gamma}) \circ \mathcal{R} \lesssim F \circ (\hat{\beta}, \hat{\gamma})$, i.e., the function $(\hat{\beta}, \hat{\gamma}) : \mathcal{W} \to W$ is a partial semi-conjugacy between $\mathcal{R} : \mathcal{W} \to \mathcal{W}$ and $F : W \to W$.*

Corollary 6.6. *Given $\mathcal{T} \in \mathcal{W}$, let $\mathcal{T}_n := \mathcal{R}^n(\mathcal{T})$, for $n \geq 0$. Then*
1. *$\beta_n := \beta(\mathcal{T}_n)$ decays quadratically to 0.*
2. *There exists $c = c(\mathcal{T}) > 0$ such that $\gamma_n := \gamma(\mathcal{T}_n) \geq c$ for all $n \geq 1$.*

Proof. Since $(\hat{\beta}, \hat{\gamma}) : \mathcal{W} \to W$ is a partial semi-conjugacy we have

$$(\beta_n, \gamma_n) = ((\hat{\beta}, \hat{\gamma}) \circ \mathcal{R}^n)(\mathcal{T}) \lesssim (F^n \circ (\hat{\beta}, \hat{\gamma}))(\mathcal{T}) = F^n(\beta_0, \gamma_0).$$

By Proposition 5.6, $(\beta_n^\sharp, \gamma_n^\sharp) := F^n(\beta_0, \gamma_0)$ converges quadratically to some fixed point $(0, c) \in W$ where $c = c(\mathcal{T}) > 0$. Hence $\beta_n \leq \beta_n^\sharp$ converges quadratically to 0. Moreover $\gamma_n \geq \gamma_n^\sharp$, where $\gamma_n^\sharp$ converges to c. Therefore $\gamma_n \geq c$. $\square$

We now show that $\mathcal{R}\colon \mathcal{W} \to \mathcal{W}$ has an attractor $\mathcal{W}_0$ consisting of all $\mathcal{T} \in \mathcal{W}$ such that $B_k = 0$ for all $k \geq 0$. Given $\mathcal{T} \in \mathcal{W}_0$ and assuming Δ_0 is primitive, arguing as in Proposition 6.1, the set I_{d_0} is the unique essential class. Because I_{d_0} is finite, in this case there exists a unique stationary probability measure supported in I_{d_0}.

Next define a projection $\Pi\colon \mathcal{W} \to \mathcal{W}_0$, $\mathcal{T} \mapsto \mathcal{T}^0$, by

$$\mathcal{T}^0 := \begin{cases} \Delta_k^0 = \Delta_k + \tilde{B}_k & \text{if } k \geq 0 \\ C_k^0 = C_k & \text{if } k \geq 1 \\ B_k^0 = 0 & \text{if } k \geq 0 \end{cases}, \tag{6.4}$$

where $\tilde{B}_k \in \mathrm{Mat}_{d_k \times d_k}(\mathbb{R})$ is the matrix with all rows equal to $\frac{1}{d_k} \mathbb{1} B_k$. Remark that if Δ_0 is primitive then so it is the matrix Δ_k^0.

The next proposition is a first step to prove that $\mathcal{W}_0$ is an attractor of $\mathcal{R}\colon \mathcal{W} \to \mathcal{W}$.

Proposition 6.7. *Given $\mathcal{T} \in \mathcal{W}$, let $\mathcal{T}_n = \mathcal{R}^n(\mathcal{T})$, $n \geq 0$, and consider the associated family of sub-stochastic matrices $(B_j^n, \Delta_j^n, C_{j+1}^n)_{j \geq 0}$. Then the sequence of matrices Δ_0^n converges to a stochastic matrix $\Delta^\infty \in \mathrm{Mat}_{d_0 \times d_0}(\mathbb{R})$.*

Moreover if Δ_0^0 is primitive then so is Δ^∞.

Proof. Since, by Proposition 6.2 and Corollary 6.6,

$$\|\Delta_0^{n+1} - \Delta_0^n\| = \|C_1^n\, (I - \Delta_1^n)^{-1}\, B_0^n\| \leq \|C_1^n\|\, \|(I - \Delta_1^n)^{-1}\|\, \|B_0^n\| \leq \frac{\beta_n}{\gamma_n} \leq \frac{\beta_n}{c}$$

and β_n converges quadratically to 0, the sequence Δ_0^n is Cauchy. Thus it converges to some matrix Δ^∞.

Because $\mathcal{T}_n$ is a stochastic operator we have for all $n \geq 0$

$$\mathbb{1}\,\Delta_0^n + \mathbb{1}\,B_0^n = \mathbb{1}.$$

Since B_0^n converges to 0 taking the limit as $n \to \infty$ we get $\mathbb{1}\,\Delta^\infty = \mathbb{1}$, which proves that Δ^∞ is a stochastic matrix.

Finally notice that $\Delta_0^{n+1} = \Delta_0^n + C_1^n\,(I - \Delta_1^n)^{-1}\,B_0^n \geq \Delta_0^n$ because the matrices C_1^n, $(I - \Delta_1^n)^{-1} = \sum_{j=0}^\infty (\Delta_1^n)^j$ and B_0^n have non-negative entries. Choose $k \in \mathbb{N}$ such that the matrix $(\Delta_0^0)^k$ has all its entries positive. The previous inequality implies by induction that $(\Delta_0^n)^k \geq (\Delta_0^0)^k$. Consequently, taking the limit we have $(\Delta^\infty)^k \geq (\Delta_0^0)^k$ which implies that Δ^∞ is primitive. □

To control the convergence (in our scheme) to the stationary measure, we need the following family of weighted norms. Given a parameter $\mu > 1$ we introduce a *seminorm* on the space of sequences $X = (X_n)_{n\in\mathbb{N}}$, where $X_n \in \mathbb{R}^{d_n}$ for each $n \in \mathbb{N}$, defined by

$$\|X\|_\mu := \sum_{n=1}^\infty \mu^n\,\|X_n\|_1.$$

Let $\ell^1_\mu(\mathbb{N})$ be the space of sequences $X = (X_n)_{n\in\mathbb{N}}$ such that $\|X\|_\mu < +\infty$.

Proposition 6.8. *Given $\mathcal{T} \in \mathcal{W}$, a number $\mu > 1$ and any norm $\|\cdot\|^*$ on $\mathbb{R}^{d_0}$ the operator $\mathcal{T} : \left(\ell^1_\mu(\mathbb{N}), \|\cdot\|^*_\mu\right) \to \left(\ell^1_\mu(\mathbb{N}), \|\cdot\|^*_\mu\right)$ is bounded, where $\|\cdot\|^*_\mu$ stands for the norm $\|X\|^*_\mu := \|X_0\|^* + \|X\|_\mu$.*

Proof. Analogous to that of the following Proposition 6.9. □

The projection Π defined at (6.4) contracts the norm $\|\cdot\|_\mu$.

Proposition 6.9. *Given $\mathcal{T} \in \mathcal{W}$ let $\Pi = \Pi(\mathcal{T})$. Taking $c = c(\mathcal{T}) \in (0,1)$ as in Corollary 6.6 and setting $\mu = 2/c$ then the truncated operator $\Pi : \left(\ell^1_\mu(\mathbb{N}), \|\cdot\|_\mu\right) \to \left(\ell^1_\mu(\mathbb{N}), \|\cdot\|_\mu\right)$ is a $(1 - \frac{c}{2})$-contraction, i.e., for all $X \in \ell^1_\mu(\mathbb{N})$,*

$$\|\Pi(X)\|_\mu \leq (1 - \frac{c}{2})\,\|X\|_\mu.$$

Proof. Let $\Pi = \Pi(\mathcal{T})$ and consider the associated sub-stochastic matrices $(\Delta_k, C_{k+1})_{k\geq 0}$.

If $X' = \Pi(X)$ then $\forall\, n \geq 0$

$$X'_n = \Delta_n\,X_n + C_{n+1}\,X_{n+1}$$

Hence

$$\begin{aligned}\|\Pi(X)\|_\mu = \|X'\|_\mu &= \sum_{n=1}^{\infty} \mu^n \,\|\Delta_n\, X_n + C_{n+1}\, X_{n+1}\|_1 \\ &\leq \sum_{n=1}^{\infty} \|\Delta_n\|\, \mu^n \,\|X_n\|_1 + \frac{1}{\mu}\,\|C_{n+1}\|\,\mu^{n+1}\,\|X_{n+1}\|_1 \\ &\leq (1 - c + \frac{1}{\mu})\,\|X\|_\mu\end{aligned}$$

Since

$$\mu = \frac{2}{c} \quad \Leftrightarrow \quad 1 - c + \frac{1}{\mu} = 1 - \frac{c}{2}$$

the claim follows. □

Given $\Pi = \Pi(\mathcal{T})$ with $\mathcal{T} \in \mathcal{W}$ as above consider the following Π-invariant direct sum decomposition into closed linear subspaces

$$\ell^1_\mu(\mathbb{N}) = E_0 \oplus E_1$$

where

$$\begin{aligned} E_0 &= \{X \in \ell^1_\mu(\mathbb{N})\colon X_n = 0,\, \forall n \geq 1\} \equiv \mathbb{R}^{d_0} \\ E_1 &= \{X \in \ell^1_\mu(\mathbb{N})\colon X_0 = 0\}. \end{aligned}$$

Proposition 6.10. *Given $\mathcal{T} \in \mathcal{W}$ let $\Pi = \Pi(\mathcal{T})$ and assume that the sub-stochastic matrix $\Delta_0 = \Delta_0(\mathcal{T})$ is primitive. Then there exists a Π-invariant direct sum decomposition into closed linear subspaces*

$$E_0 = H_0 \oplus \langle q_0 \rangle$$

where $\Pi(q_0) = q_0$ and H_0 is the space of vectors supported in I_{d_0} with zero total mass. Moreover $\Pi|_{H_0}\colon H_0 \to H_0$ is a contraction on the finite dimensional subspace $H_0 \simeq \mathbb{R}^{d_0-1}$. Furthermore, the contraction rate of $\Pi|_{H_0}$ depends only on $\Delta_0(\mathcal{T})$.

Proof. Note that if we make the identification $E_0 \equiv \mathbb{R}^{d_0}$ then the operator $\Pi|_{E_0}$ is represented by the primitive stochastic matrix $\Delta_0 = \Delta_0(\Pi)$.

Note also that $\Delta_0(\Pi)^k \geq \Delta_0(\mathcal{T})^k$ in a component-wise sense. □

Corollary 6.11. *Given $\mathcal{T} \in \mathcal{W}$ let $\Pi = \Pi(\mathcal{T})$ and assume that the sub-stochastic matrix $\Delta_0 = \Delta_0(\mathcal{T})$ is primitive. Then there exists an adapted norm $\|\cdot\|^*$ on E_0 and a Π-invariant direct sum decomposition into closed linear subspaces*

$$\ell^1_\mu(\mathbb{N}) = (E_1 \oplus H_0) \oplus \langle q_0 \rangle$$

where $\Pi(q_0) = q_0$ and $\Pi|_{E_1 \oplus H_0}\colon E_1 \oplus H_0 \to E_1 \oplus H_0$ is a contraction w.r.t. the norm $\|X\|^*_\mu := \|X_0\|^* + \|X\|_\mu$ on $\ell^1_\mu(\mathbb{N})$.

Proof. Follows from the previous two propositions adapting the norm on H_0. □

The following proposition shows that β controls the distance from $\mathcal{T} \in \mathcal{W}$ to $\mathcal{W}_0$.

Proposition 6.12. *Given $\mathcal{T} \in \mathcal{W}$ there is a constant $C < \infty$ depending only on $\Delta_0(\Pi(\mathcal{T}))$ such that $\|\mathcal{T} - \Pi(\mathcal{T})\|_\mu^* \leq 2\, C\, \beta$, where $\|\cdot\|_\mu^*$ denotes the operator norm w.r.t. the norm $\|\cdot\|_\mu^*$ on $\ell^1_\mu(\mathbb{N})$.*

Proof. Up to some constant $c_0 < \infty$ which depends only on the adapted norm $\|\cdot\|^*$ (hence depending only on $\Delta_0(\Pi(\mathcal{T}))$) we have that

$$\|\tilde{B}_0\, X_0\|^* = \frac{1}{d_0}\, \|B_0\, X_0\|^* \leq c_0\, \beta\, \|X_0\|^*.$$

Therefore writing $\Pi = \Pi(\mathcal{T})$

$$\begin{aligned}
\|\mathcal{T}(X) - \Pi(X)\|_\mu^* &= \|\tilde{B}_0 X_0\|^* + \sum_{n=1}^{\infty} \mu^n\, \|B_{n-1}\, X_{n-1} + \tilde{B}_n\, X_n\|_1 \\
&\leq c_0\, \beta\, \|X_0\|^* + \sum_{n=1}^{\infty} \beta\, \mu^n\, \|X_{n-1}\|_1 + \sum_{n=1}^{\infty} \beta\, \mu^n\, \|X_n\|_1 \\
&\leq c_0\, \beta\, \|X_0\|^* + \mu\, \beta\, \|X_0\|_1 + \mu\, \beta\, \|X\|_\mu + \beta\, \|X\|_\mu \\
&\leq c_1\, \beta\, \|X_0\|^* + (1+\mu)\, \beta\, \|X\|_\mu \leq C\, \beta\, \|X\|_\mu^*
\end{aligned}$$

for some appropriate constant $c_1 < \infty$ and where $C := \max\{c_1, 1+\mu\}$. $\square$

Given $\mathcal{T} \in \mathcal{W}$, making the canonical identification $S \equiv \mathbb{N}$, the *reconstruction operator* $\Phi_{\mathcal{T}} = \Phi_{\mathcal{T},\mathcal{S},\bar{\mathcal{S}}}(1)\colon \ell^1(S) \to \ell^1(\mathbb{N})$ becomes an operator on $\ell^1(\mathbb{N})$ which can be characterized as follows: Given $V = (V_k)_{k\in\mathbb{N}}$ and $U = (U_k)_{k\in\mathbb{N}}$ such that each U_k is a vector with d_k elements and each V_k is a vector with d_{2k} elements,

$$U = \Phi_{\mathcal{T}}(V) \;\Leftrightarrow\; \begin{cases} U_k \;= V_{\frac{k}{2}} & \text{if } \; k \text{ even} \\ \\ U_k \;= (I - \Delta_k)^{-1} C_{k+1} V_{\frac{k+1}{2}} + (I - \Delta_k)^{-1} B_{k-1} V_{\frac{k-1}{2}} & \text{if } \; k \text{ odd} \end{cases}.$$

Next we provide a bound for the norm of the reconstruction operator $\Phi_{\mathcal{T}}$.

Proposition 6.13. *If $\mathcal{T} \in \mathcal{W}$ then the operator $\Phi_{\mathcal{T}}$ has norm*

$$\|\Phi_{\mathcal{T}}\| \leq \frac{1+\gamma+\beta}{\gamma}.$$

Proof. Given $V = (V_k)_{k\in\mathbb{N}} \in \ell^1(\mathbb{N})$, we have that

$$\begin{aligned}\|\Phi_{\mathcal{J}}(V)\|_1 &= \sum_{k \text{ even}} \|V_{\frac{k}{2}}\|_1 + \sum_{k \text{ odd}} \|(I-\Delta_k)^{-1} C_{k+1} V_{\frac{k+1}{2}} + (I-\Delta_k)^{-1} B_{k-1} V_{\frac{k-1}{2}}\|_1 \\ &\le \sum_{k \text{ even}} \|V_{\frac{k}{2}}\|_1 + \sum_{k \text{ odd}} \Big(\|(I-\Delta_k)^{-1}\| \|C_{k+1}\| \|V_{\frac{k+1}{2}}\|_1 + \|(I-\Delta_k)^{-1}\| \|B_{k-1}\| \|V_{\frac{k-1}{2}}\|_1 \Big) \\ &\le \left(1 + \frac{1}{\gamma} + \frac{\beta}{\gamma}\right) \|V\|_1 = \frac{1+\gamma+\beta}{\gamma} \|V\|_1 \,.\end{aligned}$$

□

Consider the convex cone

$$\ell^1_+(\mathbb{N}) := \{x \in \ell^1(\mathbb{N}) \colon x_j \ge 0, \quad \forall j \in \mathbb{N}\}$$

and the ∞-dimensional simplex

$$\Delta^1(\mathbb{N}) := \{x \in \ell^1_+(\mathbb{N}) \colon \sum_{j\in\mathbb{N}} x_j = 1\}.$$

Let $\pi \colon \ell^1_+(\mathbb{N}) \to \Delta^1(\mathbb{N})$ be the canonical (nonlinear) projection $\pi(x) := x/\|x\|$.

Lemma 6.14. *For all $x, y \in \ell^1_+(\mathbb{N})$ with $\|x\|_1 \ge h$ and $\|y\|_1 \ge h$,*

$$\|\pi(x) - \pi(y)\|_1 \le 2\, h^{-1} \, \|x - y\|_1.$$

Proof. It is enough proving this proposition with $\|x\|_1 = h = 1$ and $\|y\|_1 \ge 1$. Notice that π is a map of class C^1 such that for any $x \in \Delta^1(\mathbb{N})$ and any $u \in \ell^1(\mathbb{N})$

$$D\pi_x(u) = u - (\sum_k u_k) x.$$

Thus

$$\|D\pi_x(u)\|_1 \le \|u\|_1 + (\sum_k |u_k|) \, \|x\|_1 = 2 \, \|u\|_1.$$

Hence, if $\pi(x) = x$ and $\|y\|_1 \ge 1$ then

$$\begin{aligned}\|\pi(x) - \pi(y)\|_1 &= \|x - \frac{y}{\sum_k y_k}\|_1 \le \|y - (\sum_k y_k)\, x\|_1 \\ &= \|y - x - (\sum_k y_k - x_k)\, x\|_1 = \|D\pi_x(y-x)\|_1 \\ &\le 2\, \|y - x\|_1.\end{aligned}$$

□

Next define the nonlinear map $\tilde{\Phi}_{\mathcal{J}} \colon \Delta^1(\mathbb{N}) \to \Delta^1(\mathbb{N})$ setting $\tilde{\Phi}_{\mathcal{J}} = \pi \circ \Phi_{\mathcal{J}}$.

Proposition 6.15. *Given $\mathcal{T} \in \mathcal{W}$, then for all $V, V' \in \Delta^1(\mathbb{N})$,*

$$\|\tilde{\Phi}_{\mathcal{T}}(V) - \tilde{\Phi}_{\mathcal{T}}(V')\|_1 \leq \frac{2\,(1+\gamma+\beta)}{\gamma}\,\|V - V'\|_1\,.$$

Proof. If $V \in \Delta^1(\mathbb{N})$ and $U = \Phi_{\mathcal{T}}(V)$ then $U_k = V_{\frac{k}{2}}$ when k is even while otherwise

$$U_k = (I - \Delta_k)^{-1} C_{k+1} V_{\frac{k+1}{2}} + (I - \Delta_k)^{-1} B_{k-1} V_{\frac{k-1}{2}}\,.$$

Therefore, when k is odd, because $\mathbb{1} C_k \geq \gamma \mathbb{1}$ and $\|I - \Delta_k\| \leq 1$, we have

$$\begin{aligned}
\|U_k\| &= \|(I - \Delta_k)^{-1} C_{k+1} V_{\frac{k+1}{2}} + (I - \Delta_k)^{-1} B_{k-1} V_{\frac{k-1}{2}}\| \\
&\geq \|(I - \Delta_k)^{-1} C_{k+1} V_{\frac{k+1}{2}}\| \\
&\geq \|(I - \Delta_k)(I - \Delta_k)^{-1} C_{k+1} V_{\frac{k+1}{2}}\| \\
&= \|C_{k+1} V_{\frac{k+1}{2}}\| \geq \gamma\,\|V_{\frac{k+1}{2}}\|\,.
\end{aligned}$$

Thus, since $\displaystyle\sum_{k \text{ even}} \|V_{\frac{k}{2}}\| = 1$ and $\displaystyle\sum_{k \text{ odd}} \|V_{\frac{k+1}{2}}\| = 1 - \|V_0\|$, one has

$$\|\Phi_{\mathcal{T}}(V)\| \geq 1 + \gamma\,(1 - \|V_0\|) \geq 1.$$

By Lemma 6.14 and Proposition 6.13, it follows that

$$\begin{aligned}
\|\tilde{\Phi}_{\mathcal{T}}(V) - \tilde{\Phi}_{\mathcal{T}}(V')\|_1 &\leq 2\,\|\Phi_{\mathcal{T}}(V) - \Phi_{\mathcal{T}}(V')\|_1 \\
&\leq 2\,\|\Phi_{\mathcal{T}}\|\,\|V - V'\|_1 \\
&\leq \frac{2\,(1+\gamma+\beta)}{\gamma}\,\|V - V'\|_1\,.
\end{aligned}$$

$\square$

The following lemma is an easy exercise.

Lemma 6.16. *Let (X, d) be a complete metric space and $T_j \colon X \to X$ be Lipschitz contractions, for $j = 1, 2$, such that* $\mathrm{Lip}(T_1) \leq \kappa < 1$. *If $x_j = T_j(x_j)$ is the unique fixed point of T_j, $j = 1, 2$, then*

$$d(x_1, x_2) \leq \frac{d_\infty(T_1, T_2)}{1 - \kappa},$$

where $d_\infty(T_1, T_2) := \sup_{x \in X} d(T_1(x), T_2(x))$.

We now state and prove our main theorem which establishes the existence and uniqueness of a stationary measure, also providing a method to approximate this stationary measure for any stochastic operator $\mathcal{T} \in \mathcal{W}$. The convergence of this method is quadratic, faster than any other method based on the Perron-Frobenius theorem.

Theorem 4. *Given $\mathcal{T} \in \mathcal{W}$ such that Δ_0 is primitive the stochastic operator $\mathcal{T}$ has a unique stationary probability measure $q \in \Delta^1(\mathbb{N})$. Moreover q can be well approximated in the following way: Set $\mathcal{T}_j = \mathcal{R}^j(\mathcal{T})$ for $0 \leq j \leq n$ and let $q_0 := (V_0, 0, 0, \ldots) \in \ell^1(\mathbb{N})$ be the unique fixed point of $\Pi(\mathcal{T}_n)$ in $\Delta^1(\mathbb{N})$, where V_0 is a vector with d_0 entries. Then*

$$\|q_n^* - q\|_1 \leq 2^{n+1} \left(\frac{1+\gamma+\beta}{\gamma}\right)^n \beta_n \quad \text{where} \quad q_n^* = (\tilde{\Phi}_{\mathcal{T}_0} \circ \tilde{\Phi}_{\mathcal{T}_1} \circ \cdots \circ \tilde{\Phi}_{\mathcal{T}_{n-1}})(q_0).$$

In particular the approximation error $\varepsilon_n := 2^{n+1} \left(\frac{1+\gamma+\beta}{\gamma}\right)^n \beta_n$ decays quadratically to 0.

Proof. Given $\mathcal{T} \in \mathcal{W}$ consider the stochastic matrix $\Delta^\infty = \Delta^\infty(\mathcal{T})$ in the previous proposition. Since Δ^∞ is primitive we can take an adapted norm $\|\cdot\|_0$ in $\mathbb{R}^{d_0}$ such that for some $\rho \in (0,1)$ we have that $\|\Delta^\infty x\|_0 \leq \rho\,\|x\|_0$ for all $x \in \mathbb{R}^{d_0}$ with $\sum_j x_j = 0$. Let $\mathcal{T}_n := \mathcal{R}^n(\mathcal{T})$, with associated sub-stochastic matrices $(B_j^n, \Delta_j^n, C_{j+1}^n)_{j\geq 0}$. For n large enough $\Delta_0^n \approx \Delta^\infty$. Hence we can assume, possibly replacing ρ by $\rho_1 \in (\rho, 1)$, that $\|\Delta_0^n x\|_0 \leq \rho\,\|x\|_0$ for all $x \in \mathbb{R}^{d_0}$ with $\sum_j x_j = 0$.

Then by Corollary 6.11, $\Pi(\mathcal{T}_n)$ is a good contraction on the space of probability measures w.r.t the distance $d_\mu^*(q,q') = \|q - q'\|_\mu^*$.

Let $q_n \in \Delta^1(\mathbb{N})$ be the true fixed point of $\mathcal{T}_n = \mathcal{R}^n(\mathcal{T})$. By Lemma 6.16 and Proposition 6.12, we conclude that

$$\|q_0 - q_n\|_1 \leq d_\mu^*(q_0, q_n) \lesssim \beta_n\,.$$

Finally, since by Theorem 3, $q := (\tilde{\Phi}_{\mathcal{T}_0} \circ \tilde{\Phi}_{\mathcal{T}_1} \circ \cdots \circ \tilde{\Phi}_{\mathcal{T}_{n-1}})(q_n)$ is an invariant probability measure of $\mathcal{T}$, using Proposition 6.15 one has

$$\begin{aligned} \|q_n^* - q\|_1 &\leq \left(2\,\frac{1+\gamma+\beta}{\gamma}\right)^n \|q_n - q_0\|_1 \\ &\lesssim 2^{n+1} \left(\frac{1+\gamma+\beta}{\gamma}\right)^n \beta_n\,. \end{aligned}$$

□

We now establish the existence of a stationary measure (the uniqueness was established in Proposition 6.1) under a slightly different assumption. Compare item (d) below and item (d) in the definition of the space $\mathcal{W}$.

Corollary 6.17. *Assume that the operator $\mathcal{T}$ defined in (6.1) and its associated matrices satisfy for some constant $\gamma > 0$*

(a) *all matrices Δ_k, B_k and C_k have non-negative entries,*
(b) $\mathbb{1}\,\Delta_0 + \mathbb{1}\,B_0 = \mathbb{1}$,
(c) $\mathbb{1}\,C_k + \mathbb{1}\,\Delta_k + \mathbb{1}\,B_k = \mathbb{1}$, *for all $k \geq 1$,*
(d) $\|B_k\| \leq \beta_k$, *for all $k \geq 0$, where the sequence $(\beta_k)_{k\in\mathbb{N}_0}$ converges to zero.*

(e) $\mathbb{1}\, C_k \geq \gamma\, \mathbb{1}$, *for all* $k \geq 1$.
(f) Δ_0 *is primitive.*

Then $\mathcal{T}$ *has a unique stationary probability measure.*

Proof. We just need to group the graphs $\Gamma_0, \Gamma_1, \ldots, \Gamma_k$ for some large enough k so that

$$\beta_k/\gamma^2 < (3 - \sqrt{5})/2.$$

□

7. Some examples

We want to exhibit here some examples where Corollary 6.17 can be applied. Let us fix the constant $0 < \gamma < 1$ and a sequence $\{\beta_k\}_{k\geq 0}$ that converges to zero.

7.1. **The $\mathbb{Z}^2$ lattice:** First consider the following undirected graph $G_{\mathbb{Z}^2}$ (with loops) whose vertices are the points of $\mathbb{Z}^2$ and the edges correspond to connections from (i,j) to (x,y) where $\|(x-i, y-j)\|_1 = 1$ and $\|(x_1,x_2)\|_1 := |x_1| + |x_2|$. Then we partition $\mathbb{Z}^2$ into the sets

$$\Gamma_k = \{(i,j) : \|(i,j)\|_1 = k\} \qquad \text{for any } k \geq 0\,.$$

Let $\mathcal{G}_{\mathbb{Z}^2}$ be the class of stochastic graphs satisfying (see Figure 5):

(1) Every graph in $\mathcal{G}_{\mathbb{Z}^2}$ is supported on $G_{\mathbb{Z}^2}$, in the sense that all non-zero probability transitions either correspond to loops or else edges of $G_{\mathbb{Z}^2}$.
(2) The sum of all probability transitions from $(i,j) \in \Gamma_k$ to any other vertex in Γ_{k+1} is always $\leq \beta_k$, for all $k \geq 0$.
(3) The sum of all probability transitions from $(i,j) \in \Gamma_k$ to any other vertex in Γ_{k-1} is always $\geq \gamma$, for all $k \geq 1$.
(4) The probability of remaining at $(0,0)$ is strictly positive.

The graphs of this class satisfy all the hypothesis of Corollary 6.17. Conditions (a)-(c) hold because these graphs are stochastic. Items (d) and (e) follow from conditions (2) and (3), respectively. Finally, by (4) the sub-stochastic matrix Δ_0 is clearly primitive.

7.2. **The Bethe lattice:** The Bethe lattice $\mathrm{B}(z)$ with coordination number z is the undirected graph where each vertex is connected to exactly z other vertices. We fix some site p and call it the origin; the set Γ_0 contains only this point. The origin is conected to z other points, that form the set Γ_1. Each point of Γ_1 is connected to z other points, one of them in Γ_0 and the others belonging to the set we call Γ_2. The procedure is repeated for the other Γ_k.

Let $\mathcal{G}_{\mathrm{B}(z)}$ be the class of stochastic graphs satisfying (see Figure 6):

(1) Every graph in $\mathcal{G}_{\mathrm{B}(z)}$ is supported on $\mathrm{B}(z)$, in the sense that all non-zero probability transitions either correspond to loops or else edges of $\mathrm{B}(z)$.

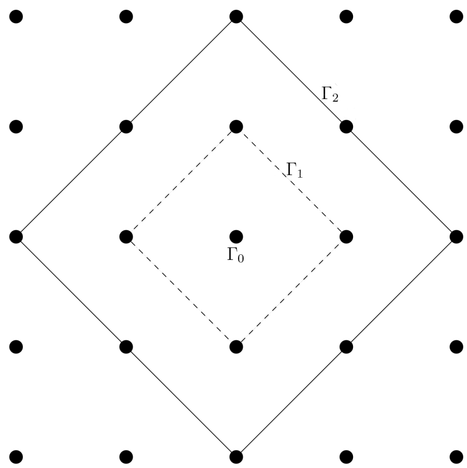

FIGURE 5. The first three subgraphs of the sequence $(\Gamma_k)_{k\geq 0}$.

(2) The sum of all probability transitions from $v \in \Gamma_k$ to any other vertex in Γ_{k+1} is always $\leq \beta_k$, for all $k \geq 0$.
(3) The sum of all probability transitions from $v \in \Gamma_k$ to any other vertex in Γ_{k-1} is always $\geq \gamma$, for all $k \geq 1$.
(4) The probability of remaining at the root p is strictly positive.

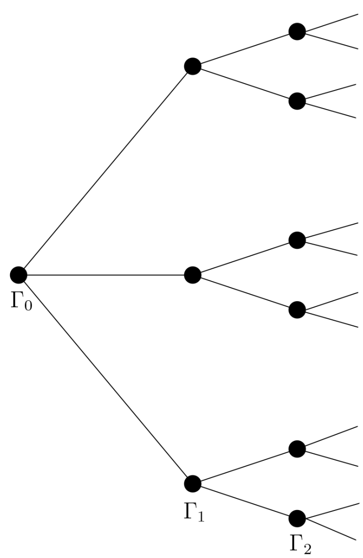

FIGURE 6. The first three subgraphs of the sequence $(\Gamma_k)_{k\geq 0}$ for the Bethe lattice when $z = 3$.

The graphs of this class satisfy all the hypothesis of Corollary 6.17. Conditions (a)-(c) hold because these graphs are stochastic. Items (d) and (e) follow from conditions (2) and (3), respectively. Finally, by (4) the sub-stochastic matrix Δ_0 is clearly primitive.

7.3. **Bounded range lattices.** Consider a graph G whose vertex set can be partitioned as a disjoint union $V = \dot{\bigcup}_{i\in\mathbb{N}} V_i$. We say that G is a bounded range lattice if there exists an integer constant K such that for all $i, j \in \mathbb{N}$, if some edge of G has endpoints in V_i and V_j then $|i - j| \leq K$.

For such graphs, defining the new partition

$$\Gamma_j = \bigcup_{jK\leq i\leq (j+1)K-1} V_i \,,$$

the interactions only connect sites lying in the same Γ_i or sites in Γ_i and Γ_j such that $|i-j| = 1$. Hence the infinite matrix describing the possible transitions of this partitioned graph is tridiagonal, showing that this operator can indeed be considered as a particular case of the one analysed in Section 6.

8. Computational algorithm

In this section we describe a numerical algorithm to approximate the stationary measure of a stochastic graph via an isospectral reduction algorithm implicit to the statement of Theorem 4.

Isospectral Reduction-Reconstruction Algorithm. An algorithm for approximating a stationary measure.

Input:

- $\{(B^n_j, \Delta^n_j, C^n_{j+1})\}_{j\geq 0}$ stochastic operator in $\mathcal{W}$,
- N number of isospectral reduction steps,
- M controls the size 2^{N+M} of the output approximation.

Output: $\{V^0_j\}_{0\leq j\leq 2^{N+M}}$ approximation of the stationary vector.

The reduction algorithm is based on the following recursive relations: for any $n \geq 0$

$$\begin{aligned}
\Delta^{n+1}_0 &:= \Delta^n_0 + C^n_1\,(I - \Delta^n_1)^{-1}\,B_0\\
\Delta^{n+1}_k &:= \Delta^n_{2k} + C^n_{2k+1}\,(I - \Delta^n_{2k+1})^{-1}\,B^n_{2k} + B^n_{2k-1}\,(I - \Delta^n_{2k-1})^{-1}\,C^n_{2k} \qquad (k \geq 1)\\
B^{n+1}_k &:= B^n_{2k+1}\,(I - \Delta^n_{2k+1})^{-1}\,B^n_{2k} \qquad (k \geq 1)\\
C^{n+1}_k &:= C^n_{2k-1}\,(I - \Delta^n_{2k-1})^{-1}\,C^n_{2k} \qquad (k \geq 0)\,.
\end{aligned}$$

Using these equations one computes $\{(B^N_j, \Delta^N_j, C^N_j)\}_{0\leq j\leq 2^M}$. Let $\mathcal{M}_{N,M}$ be the normalized stochastic tridiagonal operator obtained from these sub-stochastic blocks and compute the normalized eigenvector $\{V_j^N\}_{0\leq j\leq 2^M}$ of $\mathcal{M}_{N,M}$. Then using the following regressive recursion one computes, while $n \geq 0$,

$$V_k^{n-1} := V_{\frac{k}{2}}^n \qquad \text{if } k \text{ even},$$
$$V_k^{n-1} := (I-\Delta_k^{n-1})^{-1} C_{k+1}^{n-1} V_{\frac{k+1}{2}}^n + (I-\Delta_k^{n-1})^{-1} B_{k-1}^{n-1} V_{\frac{k-1}{2}}^n \qquad \text{if } k \text{ odd}.$$

The approximate normalized eigenvector of the original operator is the vector $\{V_j^0\}_{0\leq j\leq 2^{N+M}}$.

Bounds on the computational cost. Let $\mathcal{G}_{N,M}$ be a tree whose vertices are the symbols B_j^n, Δ_j^n and C_j^n, with $0 \leq n \leq N$ and $0 \leq j \leq 2^{N+M}$, that are used in the computation of $\{(B_j^N, \Delta_j^N, C_j^N)\}_{0\leq j\leq 2^M}$ through the above recursive equations. Each vertex B_k^n, Δ_k^n or C_k^n is connected to those on which its calculation through the previous recursive equations depends on. These edges determine a directed graph structure on $\mathcal{G}_{N,M}$.

We also define a tree $\mathcal{R}_{N,M}$ whose vertices are the symbols V_j^n, with $0 \leq n \leq N$ and $0 \leq j \leq 2^{N+M}$, that are used in the regressive recursive computation of $\{V_j^0\}_{0\leq j\leq 2^{N+M}}$. Each vertex V_k^n is connected to $V_{k/2}^{n+1}$ if k is even, or to both $V_{(k-1)/2}^{n+1}$ and $V_{(k+1)/2}^{n+1}$ when k is odd. These edges determine a directed graph structure on $\mathcal{R}_{N,M}$.

The roots of $\mathcal{G}_{N,M}$ are the output nodes $\{(B_j^N, \Delta_j^N, C_j^N)\}_{0\leq j\leq 2^M}$, while the end-leafs contain the input nodes $\{(B_j^0, \Delta_j^0, C_j^0)\}_{0\leq j\leq 2^{N+M}}$. Analogously, the roots of $\mathcal{R}_{N,M}$ are the output nodes $\{V_j^0\}_{0\leq j\leq 2^{N+M}}$, while the end-leafs contain the input nodes $\{V_j^N\}_{0\leq j\leq 2^M}$.

The number of nodes in $\mathcal{G}_{N,M}$ is equal to

$$\left|\mathcal{G}_{N,M}\right| = 2^M + 3\sum_{j=1}^{N}(2^{j+M} + 2^j - 1) + N + 1,$$

which can be bounded as follows:

$$3\,(2^N+1)(2^M-2) \leq \left|\mathcal{G}_{N,M}\right| \leq 3\,(2^N+1)(2^M-1)$$

for all $M \geq N \geq 1$. Similarly, the number of nodes in $\mathcal{R}_{N,M}$ is equal to

$$\left|\mathcal{R}_{N,M}\right| = N + 1 + \sum_{j=0}^{N} 2^{j+M} = N + 1 + 2^M\,(2^{N+1}-1).$$

These are good measurements of the computational effort required to determine, respectively, the isospectral reduced matrix $\mathcal{M}_{N,M}$ and the reconstruction from the normalized eigenvector $\{V_j^N\}_{0\leq j\leq 2^M}$. Both formulas were empirically verified for a large number of values of N and M, although it shouldn't be difficult to prove them rigorously.

Numerical comparison. The 16-digit machine precision of Mathematica impelled us to $N = 4$, since higher number of reduction steps would be wasted by computational errors.

In the following examples the sets Γ_k have the following dimensions: $\left|\Gamma_0\right| = 4$ and $\left|\Gamma_k\right| = 2$ for all $k \geq 1$. The matrix Δ_0 was taken to be $\Delta_0 := 0.9 \times \Delta$, where Δ is a 4×4

stochastic matrix with random entries. Analogously, B_0 was chosen as $B_0 := 0.1 \times B$, where B is a 2×4 stochastic matrix with random entries.

For $k \geq 1$ we define

$$\beta_k = \frac{1}{1600 + \log k}$$
$$\gamma_k = 10^{-4} + \sqrt{\frac{2\,\beta_k}{(3-\sqrt{5}) - 2/5}}.$$

The matrix C_1 is 4×2 with all entries equal to $\gamma_1/4$. For $k \geq 2$, the matrix C_k is 2×2 with all entries equal to $\gamma_k/2$. Then for all $k \geq 1$ we defined B_k to be 2×2 with all entries equal to $\beta_k/2$, and Δ_k to be 2×2 with all entries equal to $(1 - \beta_k - \gamma_k)/2$.

This family of sub-stochastic matrices determines an operator $\mathcal{T}$ in $\mathcal{W}$, to which we have applied the following algorithms:

- IRRA: Isospectral Reduction-Reconstruction Algorithm
- MEA: Mathematica Eigenvector Algorithm, executed running the command Eigenvector[A], where A stands for the normalized stochastic matrix determined by $\{(B_j^0, \Delta_j^0, C_j^0)\}_{0 \leq j \leq 2^{N+M}}$.

The table 1 presents execution times (in seconds) for different values of M.

Algorithm	$M = 4$	$M = 5$	$M = 6$	$M = 7$	$M = 8$
IRRA	0.151439	0.312129	0.723502	1.33863	2.53573
MEA	0.516156	2.6802	20.2746	100.531	737.443

Table 1. Execution times in Mathematica

The Mathematica code used in these experiments can be downloaded from the address `https://webpages.ciencias.ulisboa.pt/~pmduarte/Research/IsospectralAlgorithm/`.

Acknowledgments. The authors would like to thank the referee for the very useful comments and suggestions that significantly improved this paper. AB would like to thank the hospitality of Universidade de Lisboa. PD was supported by Fundação para a Ciência e a Tecnologia, through the project UID/MAT/04561/2013. MJT was partially financed by Portuguese Funds through FCT (Fundação para a Ciência e a Tecnologia) within the Projects UIDB/00013/2020 and UIDP/00013/2020. AB, PD and MJT were partially supported by the Project "New trends in Lyapunov exponents"(PTDC/MAT-PUR/29126/2017).

References

[1] B. Balls-Barker and B. Z. Webb, *Link prediction in networks using effective transitions*, Linear Algebra and its Applications, **599** (2020), 79–104.

[2] D. A. Bini, Matrix Structures in Queuing Models, *Exploiting Hidden Structure in Matrix Computations: Algorithms and Applications*, Lecture Notes in Mathematics, **2173**, 65–159, Springer Cham, 2016.

[3] A. E. Brouwer and W. H. Haemers, *Spectra of graphs*, Universitext, Springer, New York, 2012.

[4] L. A. Bunimovich and B. Z. Webb, *Isospectral graph reductions and improved estimates of matrices' spectra*, Linear Algebra and its Applications, **437** (2012), 1429–1457.

[5] L. A. Bunimovich and B. Z. Webb, *Isospectral graph transformations, spectral equivalence, and global stability of dynamical networks*, Nonlinearity **25** (2012), 211–254.

[6] L. A. Bunimovich and B. Z. Webb, *Isospectral transformations*, Springer Monographs in Mathematics, Springer, New York, 2014.

[7] F. W. Crawford and M. A. Suchard, *Transition probabilities for general birth-death processes with applications in ecology, genetics, and evolution*, Journal of Mathematical Biology, **65** (2012), 553–580.

[8] D. Cvetković, P. Rowlinson and S. Simić, *An introduction to the theory of graph spectra*, London Mathematical Society Student Texts, 75, Cambridge University Press, Cambridge, 2010.

[9] D. Cvetković and S. Simić, *Graph spectra in computer science*, Linear Algebra and its Applications, **434**(6) (2011), 1545–1562.

[10] J. L. Doob, *Stochastic Processes*, Wiley, New York, 1953.

[11] P. Duarte and M. J. Torres, *Eigenvectors of isospectral graph transformations*, Linear Algebra and its Applications, **474** (2015), 110–123.

[12] P. Duarte and M. J. Torres, *Isospectral Reduction in Infinite Graphs*, Journal of Spectral Theory, to appear.

[13] S. Karlin and H. M. Taylor, *An Introduction to Stochastic Modeling*, Academic Press, 1998.

[14] P. Van Mieghem, *Graph spectra for complex networks*, Cambridge University Press, Cambridge, 2011.

[15] A. S. Novozhilov, G. P. Karev and E. V. Koonin, *Biological applications of the theory of birth-and-death processes*, Brief Bioinform., **7** (2006), 70–85.

[16] F. G. Vasquez and B. Z. Webb, *Pseudospectra of isospectrally reduced matrices*, Numerical Linear Algebra with Applications **22** (2015), 145–174.

Instituto de Matemática e Estatística, UFRGS, Av. Bento Gonçalves 9500, 91500, Porto Alegre, RS, Brazil.

Email address: `baravi@mat.ufrgs.br`

CMAF, Departamento de Matemática, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal.

Email address: `pedromiguel.duarte@gmail.com`

CMAT and Departamento de Matemática, Universidade do Minho, Campus de Gualtar, 4700-057 Braga, Portugal

Email address: `jtorres@math.uminho.pt`