STATIONARY MEASURES OF CONTINUOUS-TIME MARKOV CHAINS WITH APPLICATIONS TO STOCHASTIC REACTION NETWORKS

MADS CHR. HANSEN,* ** *University of Copenhagen*
CARSTEN WIUF,* *** *University of Copenhagen*
CHUANG XU ⓘD,**** *University of Hawai'i at Mānoa*

Abstract

We study continuous-time Markov chains on the nonnegative integers under mild regularity conditions (in particular, the set of jump vectors is finite and both forward and backward jumps are possible). Based on the so-called flux balance equation, we derive an iterative formula for calculating stationary measures. Specifically, a stationary measure $\pi(x)$ evaluated at $x \in \mathbb{N}_0$ is represented as a linear combination of a few generating terms, similarly to the characterization of a stationary measure of a birth–death process, where there is only one generating term, $\pi(0)$. The coefficients of the linear combination are recursively determined in terms of the transition rates of the Markov chain. For the class of Markov chains we consider, there is always at least one stationary measure (up to a scaling constant). We give various results pertaining to uniqueness and nonuniqueness of stationary measures, and show that the dimension of the linear space of signed invariant measures is at most the number of generating terms. A minimization problem is constructed in order to compute stationary measures numerically. Moreover, a heuristic linear approximation scheme is suggested for the same purpose by first approximating the generating terms. The correctness of the linear approximation scheme is justified in some special cases. Furthermore, a decomposition of the state space into different types of states (open and closed irreducible classes, and trapping, escaping and neutral states) is presented. Applications to stochastic reaction networks are well illustrated.

Keywords: recurrence; transience; explosivity; stationary distribution; signed invariant measure

2010 Mathematics Subject Classification: Primary 60J
 Secondary 60J27; 60J28

1. Introduction

Stochastic models of interacting particle systems are often formulated in terms of continuous-time Markov chains (CTMCs) on a discrete state space. These models find application in population genetics, epidemiology, and biochemistry, where the particles are known

as *species*. A natural and accessible framework for representing interactions between species is a stochastic reaction network (SRN), where the underlying graph captures the possible jumps between states and the interactions between species. In the case where the reaction network consists of a single species, it is referred to as a *one-species reaction network*. Such networks frequently arise in various applications, ranging from SIS (susceptible-infected-susceptible) models in epidemiology [24] to bursty chemical processes (for example, in gene regulation) [25] and the Schlögl model [11]. One often focuses on examining the long-term dynamic behavior of the system, which can be captured by its corresponding limiting or stationary distribution, provided it exists. Therefore, characterizing the structure of such distributions is of great interest.

Stochastic models of reaction networks, in general, are highly nonlinear, posing challenges for analytical approaches. Indeed, the characterization of stationary distributions remain largely incomplete [26], except for specific cases such as mono-molecular (linear) reaction networks [16], complex balanced reaction networks [2], and when the associated stochastic process is a birth–death process [4]. To obtain statistical information, it is common to resort to stochastic simulation algorithms [9, 12], running the Markov chain numerous times. However, this approach can be computationally intensive, rendering the analysis of complex reaction networks infeasible [20]. Furthermore, it fails to provide analytical insights into the system.

We investigate one-species reaction networks on the nonnegative integers $\mathbb{N}_0$, and present an analytic characterization of stationary measures for general CTMCs, subject to mild and natural regularity conditions (in particular, the set of jump vectors is finite and both forward and backwards jumps are possible); see Proposition 4.1. Furthermore, we provide a decomposition of the state space into different types of states: neutral, trapping, and escaping states, and positive irreducible components (PICs) and quasi-irreducible components (QICs) (Proposition 3.1). Based on this characterization, we provide an iterative formula to calculate $\pi(x)$, $x \in \mathbb{N}_0$, for any stationary measure π, not limited to stationary distributions, in terms of a few generating terms (Theorem 4.1); similarly to the characterization of the stationary distribution/measure of a birth–death process with one generating term $\pi(0)$. The iterative formula is derived from the general flow balance equation [17].

Moreover, we show that the linear subspace of signed invariant measures has dimension at most the number of generating terms and that each signed invariant measure is given by the iterative formula and a vector of generating terms (Theorem 6.1). We use [15] to argue that there always exists a stationary measure and give conditions for uniqueness and nonuniqueness (Corollary 5.3, Corollary 5.4, Theorem 6.3, Lemma 6.3). Furthermore, we demonstrate by example that there might be two or more linearly independent stationary measures (Example 8.5). As birth–death processes have a single generating term, then there cannot be a signed invariant measure taking values with both signs.

Finally, we demonstrate how the iterative formula can be used to approximate a stationary measure. Two methods are discussed: convex optimization (Theorem 7.1) and a heuristic linear approximation scheme (Lemma 7.1). We establish conditions under which the linear approximation scheme is correct, and provide simulation-based illustrations to support the findings. Furthermore, we observe that even when the aforementioned conditions are not met, the linear approximation scheme still produces satisfactory results. In particular, it allows us to recover stationary measures in certain cases. This approach offers an alternative to truncation approaches [14, 19] and forward-in-time simulation techniques such as Gillespie's algorithm [12].

2. Preliminaries

2.1. Notation

Let $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$, $\mathbb{R}$ be the nonnegative real numbers, the positive real numbers, and the real numbers, respectively, $\mathbb{Z}$ the integers, $\mathbb{N}$ the natural numbers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ the nonnegative integers. For $m, n \in \mathbb{N}_0$, let $\mathbb{R}^{m \times n}$ denote the set of $m \times n$ matrices over $\mathbb{R}$. The sign of $x \in \mathbb{R}$ is defined as $\mathrm{sgn}(x) = 1$ if $x > 0$, $\mathrm{sgn}(x) = 0$ if $x = 0$, and $\mathrm{sgn}(x) = -1$ if $x < 0$. We use $\lceil \cdot \rceil$ to denote the ceiling function, $\lfloor \cdot \rfloor$ to denote the floor function, and $||\cdot||_p$ to denote the p-norm. Furthermore, let $\not{\Vdash}_B \colon D \to \{0, 1\}$ denote the indicator function of a set $B \subseteq D$, where D is the domain.

2.2. Markov Chains

We define a class of CTMCs on $\mathbb{N}_0$ in terms of a finite set of jump vectors and nonnegative transition functions. The setting can be extended to CTMCs on $\mathbb{N}_0^d$ for $d > 1$ and to infinite sets of jump vectors. Let $\Omega \subseteq \mathbb{Z} \backslash \{0\}$ be a finite set and $\mathcal{F} = \{\lambda_\omega \colon \omega \in \Omega\}$ a set of nonnegative transition functions on $\mathbb{N}_0$,

$$\lambda_\omega \colon \mathbb{N}_0 \to \mathbb{R}_{\geq 0}, \quad \omega \in \Omega.$$

The notation is standard in reaction network literature [3], where we find our primary source of applications. For convenience, we let

$$\lambda_\omega(k) = 0 \quad \text{for } k < 0 \quad \text{and} \quad \lambda_\omega \equiv 0 \quad \text{for } \omega \notin \Omega. \tag{2.1}$$

The transition functions define a Q matrix $Q = (q_{x,y})_{x,y \in \mathbb{N}_0}$ with $q_{x,y} = \lambda_{y-x}(x)$, $x, y \in \mathbb{N}_0$, and subsequently, a class of CTMCs $(Y_t)_{t \geq 0}$ on $\mathbb{N}_0$ by assigning an initial state $Y_0 \in \mathbb{N}_0$. Since Ω is finite, there are at most finitely many jumps from any $x \in \mathbb{N}_0$. For convenience, we identify the class of CTMCs with $(\Omega, \mathcal{F})$.

A state $y \in \mathbb{N}_0$ is *reachable* from $x \in \mathbb{N}_0$ if there exists a sequence of states $x^{(0)}, \ldots, x^{(m)}$, such that $x = x^{(0)}$, $y = x^{(m)}$, and $\lambda_{\omega^{(i)}}(x^{(i)}) > 0$ with $\omega^{(i)} = x^{(i+1)} - x^{(i)} \in \Omega$, $i = 0, \ldots, m-1$. It is *accessible* if it is reachable from some other state. The state is *absorbing* if no other states can be reached from it. A set $A \subseteq \mathbb{N}_0$ is a *communicating class* of $(\Omega, \mathcal{F})$ if any two states in A can be reached from one another, and the set cannot be extended while preserving this property. A state $x \in \mathbb{N}_0$ is a *neutral state* of $(\Omega, \mathcal{F})$ if it is an absorbing state not accessible from any other state, a *trapping state* of $(\Omega, \mathcal{F})$ if it is an absorbing state accessible from some other state, and an *escaping state* of $(\Omega, \mathcal{F})$ if it forms its own communicating class and some other state is accessible from it. A set $A \subseteq \mathbb{N}_0$ is a PIC of $(\Omega, \mathcal{F})$ if it is a nonsingleton closed communicating class, and a QIC of $(\Omega, \mathcal{F})$ if it is a nonsingleton open communicating class.

Let N, T, E, P, and Q be the (possibly empty) set of all neutral states, trapping states, escaping states, PICs, and QICs of $(\Omega, \mathcal{F})$, respectively. Each state has a unique type; hence, N, T, E, P, and Q form a decomposition of the state space into disjoint sets.

A *nonzero* measure π on a closed irreducible component $A \subseteq \mathbb{N}_0$ of $(\Omega, \mathcal{F})$ is a *stationary measure* of $(\Omega, \mathcal{F})$ if π is invariant for the Q matrix, that is, if π is a nonnegative equilibrium of the *master equation* [13]

$$0 = \sum_{\omega \in \Omega} \lambda_\omega(x - \omega)\pi(x - \omega) - \sum_{\omega \in \Omega} \lambda_\omega(x)\pi(x), \quad x \in A, \tag{2.2}$$

and a *stationary distribution* if it is a stationary measure and $\sum_{x \in A} \pi(x) = 1$. Furthermore, we say π is *unique* if it is *unique up to a scaling constant*. We might leave out 'on A' and just say π is a stationary measure of $(\Omega, \mathcal{F})$, when the context is clear.

2.3. Stochastic reaction networks

For clarity, we only introduce one-species SRNs and not SRNs in generality [3], as one-species SRNs are our primary application area. A one-species SRN is a finite labelled digraph $(\mathcal{C}, \mathcal{R})$ with an associated CTMC on $\mathbb{N}_0$. The elements of $\mathcal{R}$ are *reactions*, denoted by $y \xrightarrow{\eta} y'$, where $y, y' \in \mathcal{C}$ are *complexes* and the label is a nonnegative intensity function on $\mathbb{N}_0$. Each complex is an integer multiple of the species S, that is, nS for some n. In examples, we simply draw the reactions as in the following example with $\mathcal{C} = \{0, S, 2S, 3S\}$, and four reactions:

$$4S \xrightarrow{\eta_1} 2S \underset{\eta_3}{\overset{\eta_2}{\rightleftharpoons}} 0 \xrightarrow{\eta_4} 6S. \tag{2.3}$$

For convenience, we represent the complexes as elements of $\mathbb{N}_0$ via the natural embedding, nS $\mapsto n$, and number the reactions as in the example above. The associated stochastic process $(X_t)_{t \geq 0}$ can be given as

$$X_t = X_0 + \sum_{k=1}^{\#\mathcal{R}} (y'_k - y_k) Y_k \left(\int_0^t \eta_k(X_s) \, ds \right),$$

where Y_k are independent unit-rate Poisson processes and $\eta_k \colon \mathbb{N}_0 \to [0, \infty)$ are intensity functions [3, 10, 23]. By varying the initial species count X_0, a whole family of Markov chains is associated with the SRN.

A common choice of intensity functions is that of stochastic *mass-action kinetics* [3],

$$\eta_k(x) = \kappa_k \frac{x!}{(x - y_k)!}, \qquad x \in \mathbb{N}_0,$$

where $\kappa_k > 0$ is the *reaction rate constant* of the kth reaction, and the combinatorial factor is the number of ways y_k molecules can be chosen out of x molecules (with order). This intensity function satisfies the stoichiometric admissibility condition

$$\eta_k(x) > 0 \quad \Leftrightarrow \quad x \geq y_k$$

($\geq$ is taken componentwise). Thus, every reaction can only fire when the copy numbers of the species in the current state are no fewer than those of the source complex.

We assume mass-action kinetics in many examples below and label the reactions with their corresponding reaction rate constants, rather than the intensity functions. To bring SRNs into the setting of the previous section, we define

$$\Omega = \{y'_k - y_k \mid y_k \to y'_k \in \mathcal{R}\},$$

$$\lambda_\omega(x) = \sum_{k=1}^{\#\mathcal{R}} 1_{\{\omega\}}(y'_k - y_k) \eta_k(x), \qquad x \in \mathbb{N}_0, \; \omega \in \Omega.$$

3. A classification result

In this section we classify the state space $\mathbb{N}_0$ into different types of states. In particular, we are interested in characterizing the PICs in connection with studying stationary measures. Our primary goal is to understand the class of one-species SRNs, which we study by introducing a larger class of Markov chains embracing SRNs.

We assume throughout that a class of CTMCs associated with $(\Omega, \mathcal{F})$ is given. Define

$$\Omega_+ := \{\omega \in \Omega \colon \operatorname{sgn}(\omega) = 1\}, \qquad \Omega_- := \{\omega \in \Omega \colon \operatorname{sgn}(\omega) = -1\}$$

as the sets of positive and negative jumps, respectively. To avoid trivialities and for regularity, we make the following assumptions.

(A1) Ω is finite, $\Omega_- \neq \varnothing$ and $\Omega_+ \neq \varnothing$.

(A2) For $\omega \in \Omega$, there exists $\mathsf{i}_\omega \in \mathbb{N}_0$ such that $\lambda_\omega(x) > 0$ if and only if $x \geq \mathsf{i}_\omega$.

Then, i_ω is the smallest state for which a jump of size ω can occur (i is for input). If either of Ω_- and Ω_+ is empty, then the chain is either a pure death or a pure birth process. Assumption (A1) and (A2) are fulfilled for stochastic mass-action kinetics.

For the classification, we need some further quantities. Let $\mathsf{o}_\omega := \mathsf{i}_\omega + \omega$ be the smallest possible 'output' state after applying a jump of size ω, and let

$$\mathsf{i} := \min_{\omega \in \Omega} \mathsf{i}_\omega, \qquad \mathsf{i}_+ := \min_{\omega \in \Omega_+} \mathsf{i}_\omega, \qquad \mathsf{o} := \min_{\omega \in \Omega} \mathsf{o}_\omega, \qquad \mathsf{o}_- := \min_{\omega \in \Omega_-} \mathsf{o}_\omega.$$

Any state $x < \mathsf{i}$ is a trapping or neutral state (no jumps can occur from one of these states), and i_+ is the smallest state for which a forward jump can occur. Similarly, o is the smallest state that can be reached from any other state and o_- is the smallest state that can be reached by a backward jump.

In the example given in (2.3), all jumps are multiples of 2, that is, the Markov chain is on $2\mathbb{N}_0$ or $2\mathbb{N}_0 + 1$, depending on the initial state. Generally, the number of infinitely large irreducible classes is $\omega_* = \gcd(\Omega)$, the greatest *positive* common divisor of $\omega \in \Omega$ ($\omega_* = 2$ in (2.3)). The following classification result is a consequence of [27, Corollary 3.15].

Proposition 3.1. *Assume that (A1) and (A2) hold. Then*

$$\mathsf{N} = \{0, \dots, \min\{\mathsf{i}, \mathsf{o}\} - 1\}, \qquad \mathsf{T} = \{\mathsf{o}, \dots, \mathsf{i} - 1\}, \qquad \mathsf{E} = \{\mathsf{i}, \dots, \max\{\mathsf{i}_+, \mathsf{o}_-\} - 1\}.$$

Furthermore, the following assertions hold.

(i) *If $\#\mathsf{T} = 0$ then $\mathsf{Q} = \varnothing$, and $\mathsf{P}_s = \omega_*\mathbb{N}_0 + s$, $s = \mathsf{o}_-, \dots, \mathsf{o}_- + \omega_* - 1$ are the PICs.*

(ii) *If $\#\mathsf{T} \geq \omega_*$ then $\mathsf{P} = \varnothing$, and $\mathsf{Q}_s = \omega_*\mathbb{N}_0 + s$, $s = \mathsf{i}_+, \dots, \mathsf{i}_+ + \omega_* - 1$ are the QICs.*

(iii) *If $0 < \#\mathsf{T} < \omega_*$ then $\mathsf{P}_s = \omega_*\mathbb{N}_0 + s$, $s = \mathsf{i}_+, \dots, \mathsf{o}_- + \omega_* - 1$ are the PICs, and $\mathsf{Q}_s = \omega_*\mathbb{N}_0 + s$, $s = \mathsf{o}_- + \omega_*, \dots, \mathsf{i}_+ + \omega_* - 1$ are the QICs.*

In either case, there are ω_* PICs and QICs in total. When PICs exist, these are indexed by $s = \max\{\mathsf{i}_+, \mathsf{o}_-\}, \dots, \mathsf{o}_- + \omega_* - 1$, and $\mathsf{P} \neq \emptyset$ if and only if $\mathsf{i}_+ < \mathsf{o}_- + \omega_*$.

Proof. We apply [27, Corollary 3.15]. To translate the notation of the current paper to that of [27, Corollary 3.15], we set $c = 0$, $L_c = \mathbb{N}_0$, $c_* = \max\{\mathsf{i}_+, \mathsf{o}_-\}$, $\omega^* = \omega_*$, and $\omega^{**} = \omega_*$. Then, the expressions of N, T, and E naturally follow. As in [27], we define the following sets:

$$\Sigma_0^+ = \left\{ 1 + v - \left\lfloor \frac{v}{\omega_*} \right\rfloor \omega_* \colon v \in \mathsf{T} \right\},$$

$$\Sigma_0^- = \left\{ 1 + v - \left\lfloor \frac{v}{\omega^*} \right\rfloor \omega_* \colon v \in \{\mathsf{i}, \dots, \mathsf{o} + \omega_* - 1\} \setminus \mathsf{T} \right\}.$$

Since, for $v \in \mathbb{N}_0$,

$$1 \leq 1 + v - \left\lfloor \frac{v}{\omega_*} \right\rfloor \omega_* \equiv 1 + v \bmod \omega_* < 1 + \omega_*,$$

we have $\Sigma_0^+ \cap \Sigma_0^- = \emptyset$, $\Sigma_0^+ \cup \Sigma_0^- = \{1, \dots, \omega_*\}$, and $\#\Sigma_0^+ = \min\{\#T, \omega^*\}$. If $o < i$, these conclusions follow easily; if $o \geq i$ then $\Sigma_0^+ = \emptyset$, and the conclusions follow.

According to [27, Corollary 3.15], it follows that

$$\omega_* \left(\mathbb{N}_0 + \left\lceil \frac{c_* - (v-1)}{\omega_*} \right\rceil \right) + v - 1 = \begin{cases} P_0^{(v)}, & v \in \Sigma_0^-, \\ Q_0^{(v)}, & v \in \Sigma_0^+, \end{cases} \tag{3.1}$$

are the disjoint PICs and the disjoint QICs of $(\Omega, \mathcal{F})$, respectively, in the terminology of [27]. Consequently, as $\#(N \cup T \cup E) = c_*$,

$$\bigcup_{v \in \Sigma_0^- \cup \Sigma_0^+} (P_0^{(v)} \cup Q_0^{(v)}) = \mathbb{N}_0 \setminus (N \cup T \cup E) = \mathbb{N}_0 + c_*.$$

Since, for $v \in \mathbb{Z}$,

$$c_* \leq \omega_* \left\lceil \frac{c_* - (v-1)}{\omega_*} \right\rceil + v - 1 < \omega_* + c_*,$$

then we might state (3.1) as

$$P_0^{(v)} = (\omega_* \mathbb{Z} + v - 1) \cap (\mathbb{N}_0 + c_*) \quad \text{for } v \in \Sigma_0^-,$$
$$Q_0^{(v)} = (\omega_* \mathbb{Z} + v - 1) \cap (\mathbb{N}_0 + c_*) \quad \text{for } v \in \Sigma_0^+.$$

We show that the expressions given for $P_0^{(v)}, Q_0^{(v)}$ correspond to those given for P_s, Q_s in the three cases, by suitable renaming of the indices. First, note that $T = \varnothing$ if and only if $o \geq i$. From $o \leq o_- < i_- := \min_{\omega \in \Omega_-} i_\omega$ and $o \geq i$, we have $i = i_+ \leq o_-$, which yields $c_* = o_-$. Consequently, $\Sigma_0^+ = \emptyset$ and $Q = \varnothing$. This proves the expression for P_s in (i).

Otherwise, if $o < i$ then $o < i \leq i_+ < o_+ := \min_{\omega \in \Omega_+} o_\omega$, which implies that $o = o_- < i_+$. Hence, $c_* = i_+$. If $\#T \geq \omega_*$ then $P = \varnothing$, which proves the expression for Q_s in (ii). It remains to prove the last case. If $0 < \#T < \omega_*$ then

$$P_0^{(v+1)} = (\omega_* \mathbb{Z} + v) \cap (\mathbb{N}_0 + c_*) \quad \text{for } v = i, \dots, o + \omega_* - 1.$$

If $i = i_+$ then using the above equation and $o = o_-$, the expression for P_s in (iii) follows directly, and the remaining irreducible classes must be QICs. Finally, we show that $i < i_+$ is impossible, which concludes the proof. Assume oppositely that $i < i_+$. Then, $i = i_-$, $T = \{o_-, \dots, i_{-1}\}$, and $i_- \in E$. This implies that one can jump from state i_- (the smallest state for which a backward jump can be made) leftwards to a state $x \leq i_- \omega_* < o_-$. The latter inequality comes from $0 < \#T = i - o < \omega_*$ and $o = o_-$. However, this implies that $x \in N$, which is impossible.

The total number of PICs and QICs follows from $\Sigma_0^+ \cup \Sigma_0^- = \{1, \dots, \omega_*\}$. The indexation follows from $c_* = \max\{i_+, o_-\}$ in the two case (i) and (iii). Also, the inequality $i_+ < o_- + \omega_*$ follows straightforwardly in these two cases. It remains to check that it is not fulfilled in case (ii). In that case, $\#T = i - o \geq \omega_*$ by assumption, hence, $i_+ \geq i \geq o + \omega_* = o_- + \omega_*$, and the conclusion follows. $\qquad \square$

In either of the three cases of the proposition, the index s is the smallest element of the corresponding class (PIC or QIC). The role of assumption (A2) in Proposition 3.1, together with assumption (A1), is to ensure the nonsingleton irreducible classes are infinitely large. If the assumption fails there could be nonsingleton finite irreducible classes, even when $\Omega_+ \neq \emptyset$.

A stationary distribution exists trivially on each element of $\mathsf{N} \cup \mathsf{T}$. If the chain starts in E, it will eventually be trapped into a closed class, either into T or P, unless absorption is not certain in which case it might be trapped in Q. If absorption is certain it will eventually leave Q. We give two examples of CTMCs on $\mathbb{N}_0$ showing how the chain might be trapped.

Example 3.1.

(i) *The reaction network* $\mathrm{S} \longrightarrow 0,\ 2\mathrm{S} \rightleftharpoons 3\mathrm{S}$ *has* $\Omega = \{1, -1\}$, $\omega_* = 1$, $\mathsf{i}_1 = 2$, $\mathsf{i}_{-1} = 1$ *(note that there are two reactions with the jump size* -1*). Hence,* $\mathsf{i} = 1$, $\mathsf{i}_+ = 2$, *and* $\mathsf{o} = \mathsf{o}_- = 0$. *It follows from Proposition 3.1 that* $\mathsf{N} = \varnothing$, $\mathsf{T} = \{0\}$, $\mathsf{E} = \{1\}$, $\mathsf{P} = \varnothing$, *and* $\mathsf{Q} = \mathbb{N}_0 + 2$. *There is only one infinite class, which is a QIC. From the escaping state, one can only jump backward to the trapping state. The escaping state is reached from* Q.

(ii) *The reaction network* $\mathrm{S} \longrightarrow 2\mathrm{S},\ 2\mathrm{S} \rightleftharpoons 3\mathrm{S}$ *has* $\Omega = \{1, -1\}$, $\omega_* = 1$, $\mathsf{i}_1 = 1$, $\mathsf{i}_{-1} = 3$ *(note that there are two reactions with the jump size 1). Hence,* $\mathsf{i} = 1$, $\mathsf{i}_+ = 1$ *and* $\mathsf{o} = \mathsf{o}_- = 2$. *It follows from Proposition 3.1 that* $\mathsf{N} = \{0\}$, $\mathsf{T} = \varnothing$, $\mathsf{E} = \{1\}$, $\mathsf{P} = \mathbb{N}_0 + 2$, *and* $\mathsf{Q} = \varnothing$. *There is only one infinite class, which is a PIC. From the escaping state, one can only jump forward to this PIC.*

4. Characterization of stationary measures

We present an identity for stationary measures based on the flux balance equation [17]; see also [29]. It provides means to express any term of a stationary measure as a linear combination with real coefficients of the generating terms. The coefficients are determined recursively, provided (A1) and (A2) are fulfilled.

To smooth the presentation, we assume without loss of generality that

(A3) $\omega_* = 1$, $s = 0$ and $\mathsf{P}_0 = \mathbb{N}_0$.

Hence, we assume that $\mathbb{N}_0$ is a PIC and remove the index s from the notation for convenience. For general $(\Omega, \mathcal{F})$ with $\omega_* \geq 1$ and $s \in \{\max\{\mathsf{i}_+, \mathsf{o}_-\}, \ldots, \mathsf{o}_- + \omega_* - 1\}$, we translate the Markov chain to one fulfilling (A3) for each s by defining $(\Omega_*, \mathcal{F}_s)$ by $\Omega_* = \Omega\omega_*^{-1}$ (elementwise multiplication) and $\mathcal{F}_s = \{\lambda_\omega^s : \omega \in \Omega_*\}$ with $\lambda_\omega^s(x) = \lambda_{\omega\omega_*}(\omega_* x + s)$, $x \in \mathbb{N}_0$. Hence, there is no loss in assuming that (A3) holds. We assume that (A1)–(A3) hold throughout Section 4 unless otherwise stated.

Let π be any stationary measure of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$. Define ω_+, ω_- to be the largest positive and negative jump sizes, respectively,

$$\omega_+ = \max\ \Omega_+, \qquad \omega_- = \min\ \Omega_-. \tag{4.1}$$

We show that $\pi(x)$ can be expressed as a linear combination with real coefficients of the *generating terms* $\pi(L), \ldots, \pi(U)$, where

$$L = \mathsf{i}_{\omega_-} + \omega_- = \mathsf{o}_{\omega_-}, \qquad U = \mathsf{i}_{\omega_-} - 1, \tag{4.2}$$

are the *lower* and *upper* numbers, respectively. Hence, as a sum of $U - L + 1 = -\omega_-$ terms. No backward jumps of size ω_- can occur from any state $x \leq U$, and L is the smallest output state of a jump of size ω_-.

Example 4.1.

(i) *Consider the reaction network,* $0 \rightleftharpoons 2\mathrm{S},\ 5\mathrm{S} \longrightarrow \mathrm{S}$. *In this case,* $\omega_* = 2$, $\mathsf{i}_+ = \mathsf{o}_- = 0$, *and (A3) does not apply. In fact,* $s = 0, 1$ *with PICs* $2\mathbb{N}_0$ *and* $2\mathbb{N}_0 + 1$, *respectively. After*

translation, we find that $L_0 = 1$, $U_0 = 2$, and $L_1 = 0$, $U_1 = 1$, where the index refers to $s = 0$, 1. Thus, the lower and upper numbers are not the same for each class.

(ii) *The reaction network, $0 \rightleftharpoons S$, $nS \rightleftharpoons (n+2)S$ has $\omega_* = 1$, $i_+ = o_- = 0$. Hence, $s = 0$ with PIC $\mathbb{N}_0$, and (A3) applies. We find that $L = n$, $U = n + 1$. Thus, the lower and upper numbers might be arbitrarily large depending on the SRN.*

Before presenting the main results, we study the following example.

Example 4.2. *Recall Example 4.1(ii) with mass-action kinetics, $n = 2$, and reactions*

$$0 \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} S, \qquad 2S \underset{\kappa_4}{\overset{\kappa_3}{\rightleftharpoons}} 4S.$$

According to [28, Theorem 7] this SRN is positive recurrent on $\mathbb{N}_0$ for all $\kappa_1, \ldots, \kappa_4 > 0$ and, hence, it has a unique stationary distribution. We have $L = 2$, $U = 3$.

By rewriting the master equation (2.2), we obtain

$$\pi(0) = \frac{\kappa_2}{\kappa_1}\pi(1), \quad \pi(1) = \frac{2\kappa_2}{\kappa_1}\pi(2),$$

$$\pi(4) = \sum_{i=1}^{3}\frac{\eta_i(2)}{\eta_4(4)}\pi(2) - \frac{\eta_1(1)}{\eta_4(4)}\pi(1) - \frac{\eta_2(3)}{\eta_4(4)}\pi(3),$$

$$\pi(\ell) = \sum_{i=1}^{4}\frac{\eta_i(\ell-2)}{\eta_4(\ell)}\pi(\ell-2) - \frac{\eta_1(\ell-3)}{\eta_4(\ell)}\pi(\ell-3) - \frac{\eta_2(\ell-1)}{\eta_4(\ell)}\pi(\ell-1) - \frac{\eta_3(\ell-4)}{\eta_4(\ell)}\pi(\ell-4),$$

the latter for $\ell > U + 1 = 4$. There is not an equation that expresses $\pi(3)$ in terms of the lower states $\ell < 3$ as the state 3 can only be reached from 2, 4, and 5. Consequently, $\pi(0)$ and $\pi(1)$ can be found from $\pi(2)$ and $\pi(3)$, but not vice versa, and $\pi(\ell)$, $\ell \geq U + 1 = 4$ is determined recursively from $\pi(2)$ and $\pi(3)$ using the last equations above, say, $\pi(\ell) = \gamma_2(\ell)\pi(2) + \gamma_3(\ell)\pi(3)$, where the coefficients $\gamma_2(\ell)$, $\gamma_3(\ell)$ depend on the intensity functions. For $\ell = 0$, 1, these follow from the first equations (see also Theorem 4.1 below),

$$\gamma_2(0) = \frac{2\kappa_2^2}{\kappa_1^2}, \qquad \gamma_3(0) = 0, \qquad \gamma_2(1) = \frac{2\kappa_2}{\kappa_1}, \qquad \gamma_3(1) = 0,$$

while for $\ell = 2$, 3, we obviously have $\gamma_j(\ell) = \delta_{j,\ell}$, where $\delta_{\ell,j}$ is the Kronecker symbol. For $\ell > 3$, the coefficients are given recursively; see Theorem 4.1 for the general expression. A recursion for $\pi(\ell)$ cannot be given in terms of $\pi(0), \ldots, \pi(1)$ alone.

First, we focus on the real coefficients of the linear combination. From Example 4.2 we learn that the coefficients take different forms depending on the state, which is reflected in the definition of $\gamma_j(\ell)$ below. For convenience, rows and columns of a matrix are indexed from 0.

By the definition of ω_+ and ω_-, any state is reachable from at most $\omega_+ - \omega_-$ other states in one jump. For this reason, let

$$m_* = \omega_+ - \omega_- 1 \geq 1 \tag{4.3}$$

(as $\omega_+ \geq 1$, $-\omega_- \geq 1$), and define the functions

$$\gamma_j \colon \mathbb{Z} \to \mathbb{R}, \quad L \leq j \leq U,$$

by

$$
\gamma_j(\ell) = \begin{cases}
0 & \text{for } \ell < 0, \\
G_{\ell, j-L} & \text{for } \ell = 0, \ldots, L-1, \ j < U, \\
0 & \text{for } \ell = 0, \ldots, L-1, \ j = U, \\
\delta_{\ell, j} & \text{for } \ell = L, \ldots, U, \\
\displaystyle\sum_{k=1}^{m_*} \gamma_j(\ell - k) f_k(\ell) & \text{for } \ell > U,
\end{cases}
\tag{4.4}
$$

where the functions $f_k \colon [U+1, \infty) \cap \mathbb{N}_0 \to \mathbb{R}$ are defined by

$$
f_k(\ell) = -\frac{c_k(\ell - k)}{c_0(\ell)} \qquad \text{for } \ell > U, \ k = 0, \ldots, m_*,
\tag{4.5}
$$

$$
c_k(\ell) = \operatorname{sgn}\left(\omega_- + k + \frac{1}{2}\right) \sum_{\omega \in B_k} \lambda_\omega(\ell), \qquad \ell \in \mathbb{Z}, \ k = 0, \ldots, m_*,
\tag{4.6}
$$

$$
B_k = \left\{ \omega \in \Omega \ \middle| \ k'(\omega - k') > 0, \ \text{with } k' = \omega_- + k + \tfrac{1}{2} \right\}
\tag{4.7}
$$

(note that $f_0(\ell) = -1$ is not used in the definition of $\gamma_j(\ell)$, but appears in the proof of the Proposition 4.1 below), and $G = -(H_1)^{-1} H_2$ is an $L \times (U - L)$ matrix defined from the $L \times U$ matrix $H = (H_1 \ H_2)$ with entries

$$
H_{m,n} = \delta_{m,n} - \frac{\lambda_{(m-n)}(n)}{\sum_{\omega \in \Omega} \lambda_\omega(m)}, \qquad m = 0, \ldots, L-1, \ n = 0, \ldots, U-1,
\tag{4.8}
$$

where H_1 is $L \times L$ dimensional and H_2 is $L \times (U - L)$ dimensional. Note that H is the empty matrix if $L = 0$ or $U = L$. In (4.8) we adopt the convention stated in (2.1). In particular, $H_{m,m} = 1$ for $0 \le m \le L - 1$. By definition, $(I_L \ - G)$ is the row reduced echelon form of H by Gaussian elimination, where I_L is the $L \times L$ identity matrix.

The functions $c_k \colon \mathbb{Z} \to \mathbb{R}$ and the sets B_k come out of the flux balance equation, that is, an equivalent formulation of the master equation [17]. We use it in (4.10) in the proof of Proposition 4.1. For each $x \in \mathbb{N}_0$, it provides an identity between two positive sums, one over terms evaluated in at most ω_+ states with values $\ell < x$, and one over terms evaluated in at most $-\omega_- + 1$ states with values $\ell \ge x$. The function $f_k(\ell)$ is well defined for $\ell > U$ if (A1) and (A2) are fulfilled. In that case, $c_0(\ell) < 0$; see Lemma A.1. The matrix H is well defined with invertible H_1 under the assumptions (A1) and (A2), provided $L > 0$ and $U > L$; see Lemma A.2.

Proposition 4.1 and Theorem 4.1 below do not require uniqueness of the stationary measure.

Proposition 4.1. *A nonzero measure π on $\mathbb{N}_0$ is a stationary measure of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$ if and only if*

$$
\sum_{k=0}^{m_*} \pi(\ell - k) c_k(\ell - k) = 0 \quad \text{for } \ell > U - L.
\tag{4.9}
$$

Here, $\pi(\ell) = 0$ for $\ell < 0$ for convenience, c_k is defined as in (4.6) and m_* is defined as in (4.12).

Proof. We recall an identity related to the flux balance equation [17] for stationary measures for a CTMC on the nonnegative integers [29, Theorem 3.3], which is a consequence of the master equation (2.2): a nonnegative measure (probability measure) π on $\mathbb{N}_0$ is a stationary measure (stationary distribution) of $(\Omega, \mathcal{F})$ if and only if

$$
-\sum_{j=\omega_-+1}^{0} \pi(x-j) \sum_{\omega \in A_j} \lambda_\omega(x-j) + \sum_{j=1}^{\omega_+} \pi(x-j) \sum_{\omega \in A_j} \lambda_\omega(x-j) = 0, \quad x \in \mathbb{Z},
\tag{4.10}
$$

where $\#\Omega < \infty$ is used, and the domain of π as well as λ_ω is extended from $\mathbb{N}_0$ to $\mathbb{Z}$ (that is, $\pi(x) = 0$, $\lambda_\omega(x) = 0$ for $x \in \mathbb{Z} \setminus \mathbb{N}_0$). Furthermore, the sets A_j are defined by

$$A_j = \begin{cases} \{\omega \in \Omega : j > \omega\} & \text{if } j \in \{\omega_-, \ldots, 0\}, \\ \{\omega \in \Omega : j \le \omega\} & \text{if } j \in \{1, \ldots, \omega_+ + 1\}. \end{cases}$$

If $x \le 0$ then all terms in both double sums in (4.10) are zero. In fact, (4.10) for x is equivalent to the master equation (2.2) for $x - 1$.

Assume that π is a stationary measure of $(\Omega, \mathcal{F})$. Let $x = \ell + (\omega_- + 1)$, $\ell \in \mathbb{Z}$, and let $j = k + (\omega_- + 1)$ with $j \in \{\omega_- + 1, \ldots, \omega_+\}$. Then, $x - j = \ell - k$, and it follows that (4.10) is equivalent to

$$\sum_{k=0}^{m_*} \text{sgn}\left(\omega_- + k + \frac{1}{2}\right) \pi(\ell - k) \sum_{\omega \in A_{k+\omega_-+1}} \lambda_\omega(\ell - k) = 0, \tag{4.11}$$

where $\text{sgn}(\omega_- + k + 1/2) = 1$ if $\omega_- + k \ge 0$, and -1 otherwise. It implies that

$$(k + \omega_- + 1) > \omega \quad \Longleftrightarrow \quad \left(k + \omega_- + \tfrac{1}{2}\right) > \omega,$$
$$(k + \omega_- + 1) \le \omega \quad \Longleftrightarrow \quad \left(k + \omega_- + \tfrac{1}{2}\right) < \omega.$$

Hence, it also follows from the definition (4.7) of B_k that

$$A_{k+\omega_-+1} = B_k, \quad k = 0, \ldots, m_*.$$

Thus, (4.11) is equivalent to

$$\sum_{k=0}^{m_*} \pi(\ell - k) c_k(\ell - k) = 0, \qquad \ell \in \mathbb{N}_0, \tag{4.12}$$

using the definition of $c_k(\ell)$. Since $\ell = x - (\omega_- + 1) = x + U - L$ and (4.10) is trivial for $x \le 0$, then (4.12) is also trivial for $\ell = 0, \ldots, U - L$ (all terms are zero). This proves the bi-implication and the proof is completed. $\qquad \square$

For $0 \le \ell \le U - L$, all m_* terms in (4.9) are zero, hence, the identity is a triviality. We express π in terms of the generating terms, $\pi(L), \ldots, \pi(U)$.

Theorem 4.1. *A nonzero measure π on $\mathbb{N}_0$ is a stationary measure of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$ if and only if*

$$\pi(\ell) = \sum_{j=L}^{U} \pi(j)\gamma_j(\ell) \quad \text{for } \ell \in \mathbb{N}_0, \tag{4.13}$$

where γ_j is defined as in (4.4).

Proof. Assume that π is a stationary measure. The proof is by induction. We first prove the induction step. Assume that (4.13) holds for $\ell - 1 \ge U$ and all ℓ' such that $\ell - 1 \ge \ell' \ge 0$. Then, from Proposition 4.1, (4.4), (4.5), (4.6), and the induction hypothesis, we have

$$\pi(\ell) = -\sum_{k=1}^{m_*} \pi(\ell - k) \frac{c_k(\ell - k)}{c_0(\ell)} = -\sum_{k=1}^{m_*} \sum_{j=L}^{U} \pi(j)\gamma_j(\ell - k) \frac{c_k(\ell - k)}{c_0(\ell)} = \sum_{j=L}^{U} \pi(j)\gamma_j(\ell).$$

We next turn to the induction basis. For $\ell = L, \ldots, U$, (4.13) holds trivially as $\gamma_j(\ell) = \delta_{\ell,j}$ in this case. It remains to prove (4.13) for $\ell = 0, \ldots, L-1$. Since π is a stationary measure, it fulfills the master equation (2.2) for $(\Omega, \mathcal{F})$. By rewriting this, we obtain

$$\pi(\ell) = \frac{\sum_{\omega \in \Omega} \lambda_\omega(\ell - \omega) \pi(\ell - \omega)}{\sum_{\omega \in \Omega} \lambda_\omega(\ell)}, \qquad \ell \in \mathbb{N}_0.$$

The denominator is never zero because for $\ell \geq 0$, it holds that $\lambda_\omega(\ell) > 0$ for at least one $\omega \in \Omega$ (otherwise zero is a trapping state).

In particular, for $\ell = 0, \ldots, L-1$, using (4.2), it holds that $\ell - \omega_- < i_{\omega_-}$. Hence, $\lambda_{\omega_-}(\ell - \omega_-) = 0$, and

$$\pi(\ell) = \frac{\sum_{\omega \in \Omega \setminus \{\omega_-\}} \lambda_\omega(\ell - \omega) \pi(\ell - \omega)}{\sum_{\omega \in \Omega} \lambda_\omega(\ell)}, \qquad \ell = 0, \ldots, L-1.$$

Now, defining $n = \ell - \omega$, we have $n < L - 1 - \omega_- = U$ for $\omega \in \Omega \setminus \{\omega_-\}$, using $U - L = -(\omega_- + 1)$; see (4.2). Hence,

$$\pi(\ell) = \frac{\sum_{n=0}^{U-1} \lambda_{\ell-n}(n) \pi(n)}{\sum_{\omega \in \Omega} \lambda_\omega(\ell)}, \qquad \ell = 0, \ldots, L-1,$$

with the convention in (2.1). Evoking the definition of H in (4.8), this equation may be written as

$$H \begin{pmatrix} \pi(0) \\ \vdots \\ \pi(U-1) \end{pmatrix} = 0.$$

Recall that $G = -(H_1)^{-1} H_2$ with $H = (H_1 \ H_2)$. Noting that $\gamma_U(\ell) = 0$, $\ell = 0, \ldots, L-1$, yields that (4.13) is fulfilled with $\gamma_j(\ell) = G_{\ell, j-L}$, $\ell = 0, \ldots, L-1$, and $j = L, \ldots, U-1$, and the proof of the first part is concluded.

For the reverse part, we note that for $0 \leq x - 1 \leq L - 1$, the argument above is 'if and only if': $\pi(x-1) = \sum_{j=L}^{U} \pi(j) \gamma_j(x-1)$ if and only if the master equation (2.2) is fulfilled for $x - 1$. As noted in the proof of Proposition 4.1, this is equivalent to (4.10) being fulfilled for x, which in turn is equivalent to (4.3) being fulfilled for $\ell = x + U - L$ (the equation is replicated here),

$$\sum_{k=0}^{m_*} \pi(\ell - k) c_k(\ell - k) = 0. \tag{4.14}$$

As $0 \leq x - 1 \leq L - 1$, then (4.14) holds for $\ell = U - L + 1, \ldots, U$.

For $\ell > U$, we have, using (4.4) and (4.5),

$$\sum_{k=0}^{m_*} \pi(\ell - k) c_k(\ell - k) = \sum_{k=0}^{m_*} \sum_{j=L}^{U} \pi(j) \gamma_j(\ell - k) c_k(\ell - k)$$

$$= \sum_{j=L}^{U} \pi(j) \sum_{k=0}^{m_*} \gamma_j(\ell - k) c_k(\ell - k)$$

$$= \sum_{j=L}^{U} \pi(j) \gamma_j(\ell) c_0(\ell) + \sum_{j=L}^{U} \pi(j) \sum_{k=1}^{m_*} \gamma_j(\ell - k) c_k(\ell - k)$$

$$= 0;$$

hence, (4.14) holds for all $\ell > U - L$. According to Proposition 4.1, π is then a stationary measure of $(\Omega, \mathcal{F})$. $\square$

For completeness, we apply Proposition 4.1 to Example 4.2. The SRN has $\omega_- = -2$ and $\omega_+ = 2$, such that $m_* = \omega_+ + -\omega_- 1 = 3$. Equation (4.14) for $1 = U - L < \ell \leq U = 3$ becomes

$$-\kappa_2 \pi(1) + \kappa_1 \pi(0) = 0, \qquad -2\kappa_2 \pi(2) + \kappa_1 \pi(1) = 0,$$

in agreement with the equations found in Example 4.2.

5. Skip-free Markov chains

5.1. Downwardly skip-free Markov chains

An explicit iterative formula can be derived from Theorem 4.1 for downwardly skip-free Markov chains.

Corollary 5.1. *Assume that $\omega_- = -1$, and let $\pi(0) > 0$. Then, π is a stationary measure of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$ if and only if*

$$\pi(\ell) = \pi(0)\gamma_0(\ell) \quad \text{for } \ell \in \mathbb{N}_0,$$

where

$$\gamma_0(0) = 1, \qquad \gamma_0(\ell) = \sum_{k=1}^{\omega_+} \gamma_0(\ell - k) f_k(\ell), \quad \ell > 0,$$

and f_k is as defined in (4.5). Consequently, there exists a unique stationary measure of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$. Furthermore, if π is a stationary distribution then $\pi(0) = (\sum_{\ell=0}^{\infty} \gamma_0(\ell))^{-1}$.

Proof. Since $\omega_- = -1$, then $L = U$; see (4.2). Moreover, $i_{\omega_-} = 1$ as otherwise the state zero could not be reached. Hence, $L = U = i_{\omega_-} - 1 = 0$ from (4.2). Consequently, $\pi(\ell) = \pi(0)\gamma_0(\ell)$, $\ell \in \mathbb{N}_0$, from Theorem 4.1. Furthermore, $m_* = \omega_+ - \omega_- 1 = \omega_+$ such that the expression for $\gamma_0(\ell)$, $\ell \in \mathbb{N}_0$, follows from (4.4). Positivity of $\gamma(\ell)$, $\ell \in \mathbb{N}_0$, follows from Theorem 6.1. If π is a stationary distribution then $1 = \sum_{\ell=0}^{\infty} \pi(\ell) = \pi(0) \sum_{\ell=0}^{\infty} \gamma_0(\ell)$, and the conclusion follows. $\square$

Corollary 5.1 leads to the classical birth–death process characterization.

Corollary 5.2. *Assume that $\Omega = \{-1, 1\}$. A measure π on $\mathbb{N}_0$ is a stationary measure of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$ if and only if*

$$\pi(\ell) = \pi(0) \prod_{j=1}^{\ell} \frac{\lambda_1(j-1)}{\lambda_{-1}(j)} \quad \text{for } \ell > 0,$$

where $\pi(0) = (1 + \sum_{\ell=1}^{\infty} \prod_{j=1}^{\ell} \lambda_1(j-1)/\lambda_{-1}(j))^{-1}$ in the case of a stationary distribution.

Proof. In particular, the process is downwardly skip free. Furthermore, $\omega_+ = 1$, such that for $\ell > 0$, we have, from Corollary 5.1, (4.5) and (4.6),

$$\gamma_0(\ell) = \sum_{k=1}^{\omega_+} \gamma_0(\ell - k) f_k(\ell) = \gamma_0(\ell - 1) f_1(\ell) = \prod_{j=1}^{\ell} f_1(j) = (-1)^{\ell} \frac{\prod_{j=1}^{\ell} c_1(j-1)}{\prod_{j=1}^{\ell} c_0(j)}.$$

By definition of B_k and $c_k(\ell)$, $k = 0, 1$ ($m_* = 1$), we have $B_0 = \{-1\}$, $B_1 = \{1\}$,

$$c_0(\ell) = -\lambda_{-1}(\ell), \qquad c_1(\ell) = \lambda_1(\ell).$$

By insertion, this yields

$$\gamma_0(\ell) = \prod_{j=1}^{\ell} \frac{\lambda_1(j-1)}{\lambda_{-1}(j)}, \quad \ell > 0,$$

and the statement follows from Corollary 5.1. □

5.2. Upwardly skip-free Markov chains

In general, we are not able to give conditions for when an upwardly skip-free Markov chain has a unique stationary measure (up to a scalar) and when it has more than one or even infinitely many linearly independent stationary measures. However, in a particular case, if there is not a unique stationary measure, we establish all stationary measures as an explicit one-parameter family of measures (Corollary 5.3). If the transition functions are polynomial then we show a stationary measure is always unique (Corollary 5.4). Thus, we need nonpolynomial transition rates for nonuniqueness and give one such example in Example 8.5.

Proposition 5.1. *Assume that $\omega_- = -2$ and $\omega_+ = 1$, hence, $U = L + 1$. A nonzero measure π on $\mathbb{N}_0$ is a stationary measure of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$ if and only if*

$$h(x)(\phi(x+1) + 1) = \phi(x)\phi(x+1), \quad x \geq U + 2, \tag{5.1}$$

where

$$\phi(x) = \frac{\pi(x-1)}{\pi(x)} \frac{\lambda_{-1}(x-1) + \lambda_{-2}(x-1)}{\lambda_{-2}(x)},$$
$$h(x) = \frac{(\lambda_{-1}(x-1) + \lambda_{-2}(x-1))(\lambda_{-1}(x) + \lambda_{-2}(x))}{\lambda_1(x-1)\lambda_{-2}(x)}, \tag{5.2}$$

for $x \geq U + 2$, and

$$\pi(x) = \frac{\pi(x+1)(\lambda_{-1}(x+1) + \lambda_{-2}(x+1)) + \pi(x+2)\lambda_{-2}(x+2)}{\lambda_1(x)}, \quad 0 \leq x \leq U, \tag{5.3}$$

with $\lambda_{-1} \equiv 0$ for convenience if there are no jumps of size -1. If this is the case then

$$\pi(x) = \pi(U+1) \prod_{j=U+1}^{x-1} \frac{\lambda_{-1}(j) + \lambda_{-2}(j)}{\lambda_{-2}(j+1)\phi(j+1)}, \quad x \geq U + 2. \tag{5.4}$$

Proof. Recall the master equation for a stationary measure π, in the form of (4.11):

$$\pi(x)(\lambda_{-1}(x) + \lambda_{-2}(x)) + \pi(x+1)\lambda_{-2}(x+1) = \pi(x-1)\lambda_1(x-1), \quad x \geq 1. \tag{5.5}$$

Define ϕ and h as in the statement for $x \geq U + 2$ (if $x = U + 1$ then $\phi(x)$, $h(x)$ might be zero; if $x < U + 1$ then division by zero occurs as $U + 1 = i_{\omega_-}$). Dividing $\pi(x+1)\lambda_{-2}(x+1)$ on both sides of (5.5) yields

$$\frac{\pi(x)}{\pi(x+1)} \frac{\lambda_{-1}(x) + \lambda_{-2}(x)}{\lambda_{-2}(x+1)} + 1 = \frac{\pi(x-1)}{\pi(x)} \frac{\lambda_{-1}(x-1) + \lambda_{-2}(x-1)}{\lambda_{-2}(x)}$$
$$\times \frac{\pi(x)}{\pi(x+1)} \frac{\lambda_{-1}(x) + \lambda_{-2}(x)}{\lambda_{-2}(x+1)}$$
$$\times \frac{\lambda_1(x-1)\lambda_{-2}(x)}{(\lambda_{-1}(x-1) + \lambda_{-2}(x-1))(\lambda_{-1}(x) + \lambda_{-2}(x))},$$

and the identity (5.1) follows. By rewriting the master equation, then (5.3) follows. The calculations can be done in reverse order yielding the bi-implication. Equation (5.4) follows by induction.   □

Lemma 5.1. *Let $h\colon [N, \infty) \cap \mathbb{N}_0 \to (0, \infty)$ and $\phi\colon [N, \infty) \cap \mathbb{N}_0 \to \mathbb{R}$ be functions with $N \in \mathbb{N}_0$, and such that*

$$h(x)(\phi(x+1)+1) = \phi(x)\phi(x+1), \quad x \geq N. \tag{5.6}$$

Let ϕ^* be a positive number. Then, ϕ is a positive function with $\phi(N) = \phi^*$, if and only if

$$\Psi_2(N) \leq \phi^* \leq \Psi_1(N),$$

where $\Psi_1(x) = \lim_{k\to\infty} \psi(x, 2k-1)$, $\Psi_2(x) = \lim_{k\to\infty} \psi(x, 2k)$, and $\psi(x, k)$ is determined recursively by

$$\psi(x, k) = h(x)\left(1 + \frac{1}{\psi(x+1, k-1)}\right), \quad x \geq N, \ k \geq 1, \tag{5.7}$$

with $\psi(x, 0) = h(x)$.

Proof. Let $\psi(x, k)$ be as in the statement and ϕ a positive function fulfilling (5.6). Note that $\psi(x, k)$ is the kth convergent of a generalized continued fraction, hence, $\psi(x, 2k)$ is increasing in $k \geq 0$ and $\psi(x, 2k+1)$ is decreasing in $k \geq 0$. Indeed, this follows from Theorem 4 (monotonicity of even and odd convergents) of [18]. See also the proof of Corollary 5.3.

By induction, we show that

$$\psi(x, 2k) < \phi(x) < \psi(x, 2k+1), \quad x \geq N, \ k \geq 0,$$

from which it follows that

$$\Psi_2(x) = \lim_{k\to\infty} \psi(x, 2k) \leq \phi(x) \leq \Psi_1(x) = \lim_{k\to\infty} \psi(x, 2k-1), \quad x \geq N. \tag{5.8}$$

For the base case, it follows from (5.6) that $\psi(x, 0) = h(x) < \phi(x)$ for $x \geq N$, and thus,

$$\psi(x, 1) = h(x)\left(1 + \frac{1}{\psi(x+1, 0)}\right) > h(x)\left(1 + \frac{1}{\phi(x+1)}\right) = \phi(x).$$

For the induction step, assume that $\psi(x, 2k) < \phi(x) < \psi(x, 2k+1)$ for $x \geq N$ and some $k \geq 0$. Then, using (5.7),

$$\psi(x, 2k+2) = h(x)\left(1 + \frac{1}{\psi(x+1, 2k+1)}\right) < h(x)\left(1 + \frac{1}{\phi(x+1)}\right) = \phi(x),$$

$$\psi(x, 2k+3) = h(x)\left(1 + \frac{1}{\psi(x+1, 2k+2)}\right) > h(x)\left(1 + \frac{1}{\phi(x+1)}\right) = \phi(x).$$

If $\phi(N) = \phi^*$ then the first implication follows from (5.8). For the reverse implication, assume that $\Psi_2(N) \leq \phi(N) \leq \Psi_1(N)$. Note from (5.6) that $\phi(x+1)$ is positive only if $h(x) < \phi(x)$. We showed that $\Psi_2(N) \leq \phi(N) \leq \Psi_1(N)$ implies $\Psi_2(x) \leq \phi(x) \leq \Psi_1(x)$ for all $x \geq N$, hence, also $h(x) = \psi(x, 0) < \Psi_2(x) \leq \phi(x)$ for all $x \geq N$, and we are done. Letting $k \to \infty$ in (5.7) yields

$$\Psi_1(x) = h(x)\left(1 + \frac{1}{\Psi_2(x+1)}\right), \qquad \Psi_2(x) = h(x)\left(1 + \frac{1}{\Psi_1(x+1)}\right), \quad x \geq N.$$

Hence, it follows from the induction hypothesis that

$$\Psi_2(x+1) = \frac{1}{(\Psi_1(x)/h(x)) - 1} \leq \frac{1}{(\phi(x)/h(x)) - 1} = \phi(x+1) \leq \frac{1}{(\Psi_2(x)/h(x)) - 1} = \Psi_1(x+1),$$

using (5.6), and the claim follows. $\qquad\square$

Below, we give a general condition for uniqueness and show that uniqueness does not always hold by example. In Section 6 we give concrete cases where uniqueness applies.

Corollary 5.3. *Assume that $\omega_- = -2$ and $\omega_+ = 1$. Choose $\pi^* > 0$ and $\Psi_2(U+2) \leq \phi^* \leq \Psi_1(U+2)$, where Ψ_1, Ψ_2 are as in Lemma 5.1 and h as in (5.2). Let ϕ be a solution to (5.6) in Lemma 5.1 with $\phi(U+2) = \phi^*$. Then, (5.4) and (5.3) define a stationary measure π of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$ by setting $\pi(U+1) = \pi^*$.*

The measure is unique if and only if $\Psi_2(U+2) = \Psi_1(U+2)$, if and only if the following sum is divergent for $x = U + 2$,

$$H(x) = \sum_{n=0}^{\infty} \left(h(x+2n+1) \prod_{i=0}^{n} \frac{h(x+2i)}{h(x+2i+1)} + \prod_{i=0}^{n} \frac{h(x+2i+1)}{h(x+2i)} \right)$$

$$= \sum_{n=0}^{\infty} \frac{(\lambda_{-1}(x+2n) + \lambda_{-2}(x+2n))(\lambda_{-1}(x-1) + \lambda_{-2}(x-1))}{\lambda_1(x+2n)\lambda_{-2}(x+2n+1)} \frac{1}{Q_n(x)}$$

$$+ \sum_{n=0}^{\infty} \frac{\lambda_{-1}(x+2n+1) + \lambda_{-2}(x+2n+1)}{\lambda_{-1}(x-1) + \lambda_{-2}(x-1)} Q_n(x),$$

where

$$Q_n(x) = \prod_{i=0}^{n} \frac{\lambda_1(x+2i-1)\lambda_{-2}(x+2i)}{\lambda_1(x+2i)\lambda_{-2}(x+2i+1)}.$$

Proof. The first part of the proof is an application of Proposition 5.1 and Lemma 5.1 with $N = U + 2$. The last bi-implication follows by noting that $\psi(x, k)$ in Lemma 5.1 is the kth convergent of a generalized continued fraction,

$$b_0 + \cfrac{c_1}{b_1 + \cfrac{c_2}{b_2 + \cfrac{c_3}{b_3 + \cdots}}} = h(x)(1 + \cfrac{1}{h(x+1)(1 + \cfrac{1}{h(x+2)(1 + \cfrac{1}{h(x+3)(1 + \cdots}}})}$$

that is, $b_n = c_{n+1} = h(x+n)$, $n \geq 0$. By transformation, this generalized continued fraction is equivalent to a (standard) continued fraction,

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}$$

with

$$a_{2n} = h(x+2n+1) \prod_{i=0}^{n} \frac{h(x+2i)}{h(x+2i+1)}, \qquad a_{2n+1} = \prod_{i=0}^{n} \frac{h(x+2i+1)}{h(x+2i)}, \qquad n \geq 0.$$

By [18, Theorem 10], the continued fraction converges if and only if $\sum_{n=0} a_n = \infty$, which proves the bi-implication, noting that by the first part of the proof it is sufficient to check $x = U + 2$, and using the concrete form of $h(x)$ in (5.2). □

We give a concrete example of a nonunique Markov chain in Example 8.5 in Section 8.

Corollary 5.4. *Assume that $\omega_+ = 1$, $\omega_- = -2$, and that $\lambda_\omega(x)$, $\omega \in \Omega$, is a polynomial for large x. Then, there is a unique stationary measure of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$.*

Proof. We prove that the first series in $H(x)$ in Corollary 5.3 diverges for $x \geq U + 2$, and hence, $H(x) = \infty$. Let $m_1 = \deg \lambda_1(x)$, $m_2 = \deg \lambda_{-2}(x)$. We first provide asymptotics of $Q_n(x)$ for large n. By the Euler–Maclaurin formula with $x \geq U + 2$,

$$
\begin{aligned}
Q_n(x) &= \exp\left(\sum_{i=0}^{n} \log\left(1 - \frac{\lambda_1(x + 2i - 1)}{\lambda_1(x + 2i)}\right) + \log\left(1 - \frac{\lambda_{-2}(x + 2i)}{\lambda_{-2}(x + 2i + 1)}\right) \right) \\
&= \exp\left(\sum_{i=0}^{n} \log\left(1 - \frac{m_1}{x + 2i} + O((x + 2i)^{-2})\right) + \log\left(1 - \frac{m_2}{x + 2i} + O((x + 2i)^{-2})\right) \right) \\
&= \exp\left(\sum_{i=0}^{n} -\frac{m_1}{x + 2i} + O((x + 2i)^{-2}) - \frac{m_2}{x + 2i} + O((x + 2i)^{-2}) \right) \\
&= \exp\left(-m_1 \log(x + 2n) - m_2 \log(x + 2n) + O(1)\right) \\
&= C(x, 2n)(x + 2n)^{-(m_1 + m_2)},
\end{aligned}
$$

where $0 < C_1(x) < C(x, 2n) < C_2(x) < \infty$ for all n, and $C_1(x)$, $C_2(x)$ are positive constants depending on x. This implies that the terms in the first series of $H(x)$ are bounded uniformly in n from below, i.e.

$$
\liminf_{n \to \infty} \frac{\lambda_{-2}(x + 2n)\lambda_{-2}(x - 1)}{\lambda_1(x + 2n)\lambda_{-2}(x + 2n + 1)} \frac{1}{Q_n(x)} > 0,
$$

and hence, the first series in $H(x)$ diverges. □

6. Invariant vectors and existence of stationary measures

In the previous section we focused on nonzero measures on $\mathbb{N}_0$ and conditions to ensure stationarity. However, the requirement of a nonzero measure is not essential. In fact, the conditions of Proposition 4.1 and Theorem 4.1 characterize equally any real-valued sequence that solves the master equation (2.2). Henceforth, we refer to the vector $(v_L, \ldots, v_U) \in \mathbb{R}^{U-L+1}$ as a *generator* and the sequence as an *invariant vector*. This implies that the linear subspace in $\ell(\mathbb{R})$ of invariant vectors is $(U - L + 1)$-dimensional. Such a vector might or might not be a *signed invariant measure* depending on whether the positive or the negative part of the vector has a finite 1-norm.

We assume that (A1)–(A3) hold throughout Section 6. If the CTMC is recurrent (positive or null) then it is well known that there exists a unique stationary measure. For transient CTMCs, including explosive chains, there also exists a nonzero stationary measure in the setting considered.

Proposition 6.1. *There exists a nonzero stationary measure of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$.*

Proof. It follows from [15, Corollary] and [23, Theorem 3.5.1], noting that the set of states with nonzero transition rates to any given state $x \in \mathbb{N}_0$ is finite, in fact $\leq \#\Omega$. □

Lemma 6.1. *Let π be a nonzero invariant measure of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$ such that $\pi(x) \geq 0$ for all $x \in \mathbb{N}_0$. Then, $\pi(x) > 0$ for all $x \in \mathbb{N}_0$. In fact, π is a nonzero stationary measure of $(\Omega, \mathcal{F})$.*

Proof. Assume that $\pi(x) = 0$. By rewriting the master equation (2.2), we obtain

$$\pi(x) = \frac{1}{\sum_{\omega \in \Omega} \lambda_\omega(x)} \sum_{\omega \in \Omega} \lambda_\omega(x - \omega)\pi(x - \omega) = 0.$$

Let x be reachable from $y \in \mathbb{N}_0$ in k steps. If $k = 1$ then it follows from above that $y = x - \omega$ for some $\omega \in \Omega$ and $\pi(y) = 0$. If $k > 1$ then $\pi(y) = 0$ by induction in the number of steps. Since $\mathbb{N}_0$ is irreducible by assumption, then any state can be reached from any other state in a finite number of steps, and $\pi(x) = 0$ for all $x \in \mathbb{N}_0$. However, this contradicts the measure is nonzero. Hence, $\pi(x) > 0$ for all $x \in \mathbb{N}_0$. $\square$

Theorem 6.1. *The sequences $\gamma_L, \ldots, \gamma_U$ form a maximal set of linearly independent invariant vectors in $\ell(\mathbb{R})$. Moreover, γ_j has positive and negative terms for all $j = L, \ldots, U$ if and only if $\omega_- < -1$, that is, if and only if $L \neq U$. If $L = U$ then γ_L has all terms positive. In any case, $\gamma(n) = (\gamma_L(n), \ldots, \gamma_U(n)) \neq 0$ for $n \in \mathbb{N}_0$.*

Proof. The former part follows immediately from Theorem 4.1, since $(\gamma_j(L), \ldots, \gamma_j(U))$ is the $(j - L + 1)$th unit vector of $\mathbb{R}^{U-L+1}$, for $j = L, \ldots, U$, by definition. The latter part follows from Lemma 6.1 and the fact that $\gamma_j(j) = 1$ and $\gamma_j(\ell) = 0$ for $\ell \in \{L, \ldots, U\} \setminus \{j\}$. If $L = U$ then the linear space is one-dimensional and positivity of all $\gamma_L(\ell)$, $\ell \in \mathbb{N}_0$, follows from Lemma 6.1 and the existence of a stationary measure; see Proposition 6.1. The equivalence between $\omega_- < 1$ and $L \neq U$ follows from (4.2). If $\gamma(n) = 0$ then $\pi(n) = 0$ for any invariant vector π, contradicting the existence of a stationary measure. $\square$

Theorem 6.2. *Let $G \subseteq \mathbb{R}^{U-L+1}$ be the set of generators of stationary measures of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$. Then, $G \subseteq \mathbb{R}_{>0}^{U-L+1}$ is a positive convex cone. Let $G_0 = G \cup \{(0, \ldots, 0)\}$ be the set including the zero vector, which generates the null measure. Then, G_0 is a closed set. The set $\{v \in G_0 \colon \|v\|_1 = 1\}$ is closed and convex.*

Proof. Positivity follows from Lemma 6.1. It is straightforward to see that G is a positive convex cone. Assume that $v^{(m)} \in G_0$, $m \in \mathbb{N}_0$, and $v^{(m)} \to v = (v_L, \ldots, v_U) \notin G_0$ as $m \to \infty$. Using Lemma 6.1, there exists $\ell \in \mathbb{N}_0$, such that $\sum_{j=L}^{U} v_j \gamma_j(\ell) < 0$. Then, there also exists $m \in \mathbb{N}_0$ such that $\sum_{j=L}^{U} v_j^{(m)} \gamma_j(\ell) < 0$ with $v^{(m)} = (v_L^{(m)}, \ldots, v_U^{(m)})$, contradicting that $v^{(m)} \in G_0$. Hence, G_0 is closed. The last statement is immediate from the previous. $\square$

Part of Theorem 6.2 might be found in [8, Theorem 1.4.6]. In general, we do not have uniqueness of a stationary measure of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$, unless in the case of recurrent CTMCs. For downwardly skip-free processes, we have $L = U = 0$, hence, the space of signed invariant measures is one-dimensional and uniqueness follows, that is, there does not exist a proper signed invariant measure taking values with both signs.

We end the section with a few results on uniqueness of stationary measures. For this, we need the following lemma.

Lemma 6.2. *Let $v = (v_L, \ldots, v_U) \in \mathbb{R}_{\geq 0}^{U-L+1}$ be a nonzero generator and assume that $\omega_+ = 1$. Let*

$$v(\ell) = \sum_{j=L}^{U} v_j \gamma_j(\ell), \quad \ell \in \mathbb{N}_0.$$

If, either

$$v(\ell) \geq 0, \quad \ell = n - (U - L), \ldots, n, \tag{6.1}$$

for some $n \geq U - L$, or

$$v(\ell) = 0, \quad \ell = n - (U - L - 1), \ldots, n, \tag{6.2}$$

for some $n \geq U - L - 1$, then

$$v(\ell) \geq 0, \quad \ell = 0, \ldots, n. \tag{6.3}$$

Proof. Since $\omega_+ = 1$, then $m_* = \omega_+ - \omega_- 1 = -\omega_- = U - L + 1$. We use $-\omega_-$ rather than $U - L + 1$ in the proof for convenience. From Lemma A.1, we have $c_k(\ell) \leq 0$ for $\ell \in \mathbb{N}_0$ and $0 \leq k < -\omega_-$, and $c_{-\omega_-}(\ell) > 0$ for $\ell \in \mathbb{N}_0$. The latter follows from $\Omega_+ = \{\omega_+\}$ and $i_{\omega_+} = 0$ in this case. Furthermore, since v defines an invariant vector, then from (4.9),

$$0 = \sum_{k=0}^{-\omega_- 1} v(n-k)c_k(n-k) + v(n+\omega_-)c_{-\omega_-}(n+\omega_-). \tag{6.4}$$

Assume that (6.1) holds. If $n = U - L$ then there is nothing to show. Hence, assume that $n \geq U - L + 1$. By assumption the sum in (6.4) is nonpositive, while the last term is nonnegative and $c_{-\omega_-}(n + \omega_- + 1) > 0$ as $n + \omega_- = n - (U - L + 1) \geq 0$. Hence, $v(n + \omega_-) \geq 0$. Continue recursively for $n := n - 1$ until $n + \omega_- = 0$. Note that $v(\ell) = v_\ell \geq 0$ for $\ell = L, \ldots, U$, in agreement with the conclusion.

Assume that (6.2) holds. If $n = U - L - 1$ then there is nothing to prove. For $n \geq U - L$, we aim to show (6.1) holds from which the conclusion follows. By simplification of (6.4),

$$-v(n - (U - L))c_{U-L}(n - (U - L)) = v(n + \omega_-)c_{-\omega_-}(n + \omega_-).$$

If $c_{U-L}(n - (U - L)) < 0$ then either $v(n - (U - L))$ and $v(n + \omega_-)$ take the same sign or are zero. Consequently, (a) $v(\ell) \geq 0$ for all $\ell = n + \omega_-, \ldots, n$ or (b) $v(\ell) \leq 0$ for all $\ell = n + \omega_-, \ldots, n$. Similarly, if $c_{U-L}(n - (U - L)) = 0$ then $v(n + \omega_-) = 0$. Consequently, (a) or (b) holds in this case too. If (a) holds then (6.3) holds. If (b) holds then (6.3) holds with reverse inequality by applying an argument similar to the nonnegative case. However, $v(\ell) = v_\ell \geq 0$, $\ell = L, \ldots, U$, and at least one of them is strictly larger than zero, by assumption. Hence, the negative sign cannot apply and the claim holds. $\qquad\square$

For $n \geq U$, define the $(U - L + 1) \times (U - L + 1)$ matrix by

$$A(n) = \begin{pmatrix} \dfrac{\gamma(n - (U - L - 1))^T}{\|\gamma(n - (U - L - 1))\|_1} \\ \vdots \\ \dfrac{\gamma(n - 1)^T}{\|\gamma(n - 1)\|_1} \\ \dfrac{\gamma(n)^T}{\|\gamma(n)\|_1} \\ \mathbf{1}^T \end{pmatrix},$$

where $\gamma(\ell) = (\gamma_L(\ell), \ldots, \gamma_U(\ell))$, $\mathbf{1} = (1, \ldots, 1)$, with '$T$' denoting transpose. This matrix is well defined by Theorem 6.1. The rows of $A(n)$, except the last one, has 1-norm one, and all entries are between -1 and 1.

Theorem 6.3. *Assume that there exists a strictly increasing subsequence $(n_k)_{k \in \mathbb{N}_0}$, such that $A(n_k) \to A$ as $k \to \infty$, with $\det(A) \neq 0$. Then the following hold.*

1. *There is at most one stationary measure of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$, say π, with the property*

$$\lim_{k \to \infty} \frac{\pi(n_k)}{\|\gamma(n_k)\|_1} = 0. \tag{6.5}$$

2. *If $\omega_+ = 1$ and π is a unique stationary measure of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$, then (6.5) holds.*

Proof.

(i) Let

$$\sigma(n) = \left(\frac{\pi(n - (U - L - 1))}{\|\gamma(n - (U - L - 1))\|_1}, \dots, \frac{\pi(n)}{\|\gamma(n)\|_1} \right),$$

and let $\pi_* = (\pi(L), \dots, \pi(U))$ be the generator of the stationary measure π. We have

$$A(n_k)\pi_* = \begin{pmatrix} \sigma(n_k) \\ 1 \end{pmatrix} \to A\pi_* = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}$$

as $k \to \infty$, where $\mathbf{0}$ is the zero vector of length $U - L$. Since A is invertible, then

$$\pi_* = A^{-1} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}.$$

Consequently, as this holds for any stationary measure with the property (6.5), then π is unique, up to a scalar.

(ii) According to Proposition 7.1 below, $A(n)$ is invertible and there is a unique nonnegative (componentwise) solution to

$$A(n)v^{(n)} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}$$

for all $n \geq U$. It follows that

$$Av^{n_k} = (A - A(n_k))v^{(n_k)} + A(n_k)v^{n_k} = (A - A(n_k))v^{(n_k)} + \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \to \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}$$

as $k \to \infty$, since $\|v^{(n_k)}\|_1 = 1$. Define

$$v = A^{-1} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix},$$

then v is nonnegative and $\|v\|_1 = 1$, since $v^{(n_k)}$ is nonnegative, $\|v^{(n_k)}\|_1 = 1$ and $v^{(n_k)} \to v$ as $k \to \infty$. We aim to show that v is the generator of the unique stationary measure π.

Define an invariant vector v^k, for each $k \in \mathbb{N}_0$, by

$$v^k(\ell) = \sum_{j=L}^{U} v_j^{(n_k)} \gamma_j(\ell), \quad \ell \in \mathbb{N}_0.$$

We have $v^k(\ell) = 0$ for all $\ell = n_k - (U - L - 1), \dots, n_k$. Hence, by Lemma 6.2, $v^k(\ell) \geq 0$ for $\ell = 0, \dots, n_k$. Fix $\ell \in \mathbb{N}_0$. Then for all large k such that $n_k \geq \ell$, we have $v^k(\ell) \geq 0$ and

$$v^k(\ell) \to v(\ell) = \sum_{j=L}^{U} v_j \gamma_j(\ell) \quad \text{for } k \to \infty.$$

Hence, $\nu(\ell) \geq 0$ for all $\ell \in \mathbb{N}_0$. Consequently, using Lemma 6.1, ν is a stationary measure and by uniqueness, it holds that $\nu = \pi$, up to a scaling constant.   □

To state the next result, we introduce some notation. Let

$$I_j^+ = \{n \in \mathbb{N}_0 \colon \gamma_j(n) > 0\}, \qquad I_j^- = \{n \in \mathbb{N}_0 \colon \gamma_j(n) < 0\}, \quad j = L, \ldots, U.$$

If $L = U$ then $I_L^+ = \mathbb{N}_0$ and $I_L^- = \emptyset$. If $L \neq U$ then by the definition of γ_j, $I_j^+ \neq \emptyset$ and $I_j^- \neq \emptyset$ (Theorem 6.1). In general, it follows from Theorem 4.1, Proposition 6.1, and Lemma 6.1, that $I_i^- \subseteq \cup_{i \neq j} I_i^+$, $\cup_{i=L}^U I_i^+ = \mathbb{N}_0$, and $\cap_{i=L}^U I_i^- = \emptyset$. In particular, for $U - L = 1$, that is, $\omega_- = -2$, then $I_U^- \subseteq I_L^+$ and $I_L^- \subseteq I_U^+$. Below, in the proof of Lemma 6.3, we show these four sets are infinite. Hence, we may define

$$r_1 = -\limsup_{n \in I_L^-} \frac{\gamma_L(n)}{\gamma_U(n)}, \qquad r_2 = -\liminf_{n \in I_U^-} \frac{\gamma_L(n)}{\gamma_U(n)}. \tag{6.6}$$

Lemma 6.3. *Assume that $\omega_- = -2$, that is, $U = L + 1$. It holds that $0 < r_1 \leq r_2 < \infty$. A nonzero measure π is a stationary measure if and only if*

$$\frac{\pi(U)}{\pi(L)} \in [r_1, r_2]. \tag{6.7}$$

Furthermore, a stationary measure π is unique if and only if $r_1 = r_2$, if and only if

$$\lim_{n \in I_L^- \cup I_U^-} \frac{\pi(n)}{\|\gamma(n)\|_1} = 0. \tag{6.8}$$

If this is the case then limes superior and limes inferior in (6.6) may be replaced by limes.

Proof. Let π be a stationary measure, which exists by Proposition 6.1. Then $\pi(n) = \pi(L)\gamma_L(n) + \pi(U)\gamma_U(n) > 0$, which implies that

$$\frac{\pi(U)}{\pi(L)} > -\frac{\gamma_L(n)}{\gamma_U(n)} \quad \text{for all } n \in I_L^- \quad \text{and} \quad \frac{\pi(U)}{\pi(L)} < -\frac{\gamma_L(n)}{\gamma_U(n)} \quad \text{for all } n \in I_U^-.$$

By taking supremum and infimum, respectively, this further implies that

$$\widetilde{r}_1 = \sup\left\{-\frac{\gamma_L(n)}{\gamma_U(n)} \colon n \in I_L^-\right\} \leq \inf\left\{-\frac{\gamma_L(n)}{\gamma_U(n)} \colon n \in I_U^-\right\} = \widetilde{r}_2,$$

and that $\pi(U)/\pi(L)$ is in the interval $[\widetilde{r}_1, \widetilde{r}_2]$. Note that $\widetilde{r}_2 < \infty$ and $\widetilde{r}_1 > 0$, since I_U^- and I_L^- are both nonempty, using Theorem 6.1. For the reverse conclusion, for any invariant vector π such that (6.7) holds, we have, using Theorem 4.1,

$$\pi(n) = \pi(L)\gamma_L(n) + \pi(U)\gamma_U(n) \geq 0 \quad \text{for all } n \in \mathbb{N}_0,$$

which implies that π is a stationary measure; see Lemma 6.1.

Assume that either $\widetilde{r}_1$ or $\widetilde{r}_2$ is attainable for some $n \in \mathbb{N}_0$. Then, there exists a stationary measure π such that

$$\frac{\pi(U)}{\pi(L)} = -\frac{\gamma_L(n)}{\gamma_U(n)},$$

that is, $\pi(n) = \pi(L)\gamma_L(n) + \pi(U)\gamma_U(n) = 0$, contradicting the positivity of π; see Lemma 6.1. Hence, I_L^- and I_U^- both contain infinitely many elements, and since neither $\widetilde{r}_1$ nor $\widetilde{r}_2$ are attainable, then sup and inf can be replaced by lim sup and lim inf, respectively, to obtain r_1 and r_2 in (6.6). Hence, also I_L^+ and I_U^+ are infinite sets, since $I_U^- \subseteq I_L^+$ and $I_L^- \subseteq I_U^+$.

For the second part, the bi-implication with $r_1 = r_2$ is straightforward. Assume that π is a stationary measure, such that $\pi(L), \pi(U) > 0$. Note that

$$g_1 = \liminf_{n \in I_L^-} \frac{\gamma_U(n)}{\|\gamma(n)\|_1} > 0, \qquad g_2 = \liminf_{n \in I_U^-} \frac{\gamma_L(n)}{\|\gamma(n)\|_1} > 0,$$

as otherwise $\pi(n) = \pi(L)\gamma_L(n) + \pi(U)\gamma_U(n) < 0$ for some large n. For $n \in I_L^-$,

$$\frac{\pi(n)}{\|\gamma(n)\|_1} = \pi(L)\frac{\gamma_L(n)}{\|\gamma(n)\|_1} + \pi(U)\frac{\gamma_U(n)}{\|\gamma(n)\|_1} = \left(\frac{\gamma_L(n)}{\gamma_U(n)} + \frac{\pi(U)}{\pi(L)}\right)\frac{\pi(L)\gamma_U(n)}{\|\gamma(n)\|_1} \tag{6.9}$$
$$\geq \pi(L)(g_1 - \epsilon)\left(\frac{\gamma_L(n)}{\gamma_U(n)} + \frac{\pi(U)}{\pi(L)}\right) \geq 0$$

for some small $\epsilon > 0$ and large n. Similarly, for $n \in I_U^-$,

$$\frac{\pi(n)}{\|\gamma(n)\|_1} = \pi(L)\frac{\gamma_L(n)}{\|\gamma(n)\|_1} + \pi(U)\frac{\gamma_U(n)}{\|\gamma(n)\|_1} = \left(\frac{\pi(L)}{\pi(U)} + \frac{\gamma_U(n)}{\gamma_L(n)}\right)\frac{\pi(U)\gamma_L(n)}{\|\gamma(n)\|_1} \tag{6.10}$$
$$\geq \pi(U)(g_2 - \epsilon)\left(\frac{\pi(L)}{\pi(U)} + \frac{\gamma_U(n)}{\gamma_L(n)}\right) \geq 0$$

for $\epsilon > 0$ and large n. Taking lim sup over $n \in I_L^-$ in (6.9) and lim sup over $n \in I_U^-$ in (6.10), using (6.8), yields

$$r_1 = -\limsup_{n \in I_L^-} \frac{\gamma_L(n)}{\gamma_U(n)} = \frac{\pi(U)}{\pi(L)}, \qquad \frac{1}{r_2} = -\limsup_{n \in I_U^-} \frac{\gamma_U(n)}{\gamma_L(n)} = \frac{\pi(L)}{\pi(U)},$$

or with lim sup replaced by lim inf. Hence, (6.6) holds with lim sup and lim inf replaced by lim. It follows that $r_1 = r_2$ and that π is unique. For the reverse statement, change the first inequality sign $\geq$ in (6.9) and (6.10) to $\leq$, and $(g_1 - \epsilon)$ and $(g_2 - \epsilon)$ to 1, and the conclusion follows. $\qquad\square$

7. Convex constrained optimization

When the Markov chain is sufficiently complex, an analytical expression for a stationary measure may not exist. In fact, this seems to be the rare case. If an analytical expression is not available, one may find estimates of the relative sizes of $\pi(L), \dots \pi(U)$, which in turn determines $\pi(\ell)$, $\ell \geq 0$, up to a scaling constant, by Theorem 4.1. If π is a stationary distribution, this constant may then be found numerically by calculating the first n probabilities $\pi(\ell)$, $\ell = 0, \dots, n$, for some n, and renormalizing to obtain a proper distribution. Here, we examine how one may proceed in practice.

Theorem 7.1. *Assume that (A1)–(A3) hold and, for $n \geq 0$, let*

$$K_n = \left\{v \in \mathbb{R}_{\geq 0}^{U-L+1} : \sum_{j=L}^{U} v_j \gamma_j(\ell) \geq 0, \ \ell = 0, \dots, n, \ \|v\|_1 = 1\right\}. \tag{7.1}$$

Then, $K_n \neq \emptyset$, $K_n \supseteq K_{n+1}$, $n \geq 0$, and $\cap_{n=1}^{\infty} K_n \subseteq \mathbb{R}_{>0}^{U-L+1}$ is nonempty and consists of all generators of nonzero stationary measures of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$, appropriately normalized. In fact, $G_0 = \cap_{n=1}^{\infty} K_n$ with G_0 as in Theorem 6.2.

Furthermore, there is a unique minimizer $v^{(n)}$ to the following constraint quadratic optimization problem:

$$\min_{v \in K_n} \|v\|_2^2. \tag{7.2}$$

Moreover, the limit $v^* = \lim_{n \to \infty} v^{(n)} \in \cap_{n=1}^{\infty} K_n$ exists and equals the generator of a stationary measure π of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$, that is, $v^* = (\pi(L), \ldots, \pi(U))$.

Proof. The sets K_n are nonempty by Proposition 6.1 and obviously nonincreasing: $K_n \supseteq K_{n+1}$. Hence, since all K_ns have a common element, the intersection $\cap_{n=1}^{\infty} K_n$ is also nonempty. Lemma 6.1 rules out the fact that the intersection contains boundary elements of $\mathbb{R}_{\geq 0}^{U-L+1}$.

Furthermore, the sets K_n are nonempty, closed, and convex. Since $\|\cdot\|_2^2$ is strictly convex, there exists a unique minimizer $v^{(n)} \in K_n$ for the constrained optimization problem (7.2)[6, p.137 or Example 8.1 on p.447].

Since the sets K_n are nonincreasing, then $v^{(n)}$ is nondecreasing in ℓ_2-norm: $\|v^{(n)}\|_2^2 \leq \|v^{(n+1)}\|_2^2$ for all $n \geq 1$. Furthermore, since the sets K_n are closed subsets of the simplex and hence compact, any sequence $v^{(n)}$, $n \geq 1$, of minimizers has a converging subsequence $v^{(n_k)}$, $k \geq 1$, say $v^* = \lim_{k \to \infty} v^{(n_k)} \in \cap_{n=1}^{\infty} K_n$ (the intersection is closed). To show uniqueness, suppose that there is another converging subsequence $v^{(m_k)}$, $k \geq 1$, such that $\widetilde{v}^* = \lim_{k \to \infty} v^{(m_k)} \in \cap_{n=1}^{\infty} K_n$ and $v^* \neq \widetilde{v}^*$. Then, it follows that $\|v^*\|_2^2 = \|\widetilde{v}^*\|_2^2$, since the norm is nondecreasing along the full sequence. By convexity of K_n, the intersection $\cap_{n=1}^{\infty} K_n$ is convex and $v_\alpha^* = \alpha v^* + (1-\alpha)\widetilde{v}^* \in \cap_{n=1}^{\infty} K_n$ for $\alpha \in (0, 1)$. By strict convexity of the norm and $v^* \neq \widetilde{v}^*$, then

$$\|v_\alpha^*\|_2^2 < \alpha \|v^*\|_2^2 + (1-\alpha)\|\widetilde{v}^*\|_2^2 = \|v^*\|_2^2, \quad \alpha \in (0, 1). \tag{7.3}$$

Let $v_\alpha^{(k)} = \alpha v^{(n_k)} + (1-\alpha)v^{(m_k)}$. By convexity and monotonicity of K_n, we have

$$v_\alpha^{(k)} \in K_{\min\{n_k, m_k\}} \quad \text{for } k \geq 1,$$
$$v_\alpha^{(k)} \to v_\alpha^* \in \cap_{n=1}^{\infty} K_n \quad \text{for } k \to \infty.$$

By assumption, $\|v_\alpha^{(k)}\|_2^2 \geq \min\{\|v^{(n_k)}\|_2^2, \|v^{(m_k)}\|_2^2\}$. Hence, $\|v_\alpha^*\|_2^2 \geq \|v^*\|_2^2$, contradicting (7.3). Consequently, $v^* = \widetilde{v}^*$.

Since $v^* \in \cap_{n=0}^{\infty} K_n$, then $v^* = (\pi(L), \ldots, \pi(U))$ for some nonzero stationary measure π of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$. □

If the process is downwardly skip free then $L = U$ and $\pi(L)$ might be set to 1. Consequently, $\pi(\ell)$, $\ell \geq 0$, can be found recursively from (4.13). Hence, it only makes sense to apply the optimization scheme for $L < U$.

The quadratic minimizing function is chosen out of convenience to identify a unique element of the set K_n. Any strictly convex function could be used for this. If there exists a unique stationary measure then one might choose a random element of K_n as any sequence of elements in K_n, $n \in \mathbb{N}$, which eventually converges to the unique element of $\cap_{n=1}^{\infty} K_n$. If there is more than one stationary measure, different measures might in principle be found by varying the convex function. In the case two different stationary measures are found, then any linear combination with positive coefficients is also a stationary measure.

In practice, the convex constrained optimization approach outlined in Theorem 7.1 often fails for (not so) large n; see the example given in Section 8. This is primarily because the coefficients $\gamma_j(\ell)$ become exponentially large with alternating signs, and because numerical evaluation of close to zero probabilities might return close to zero negative values, hence violating the nonnegativity constraint of the convex constrained optimization problem. The numerical difficulties in verifying the inequalities are nonnegative are most severe for large n, in particular, if $\pi(n)$ vanishes for large n. To face the problems mentioned above, we investigate an alternative approach to the optimization problem.

Lemma 7.1. *Assume that (A1)-(A3). Define the sets M_n, $n \geq U$, by*

$$M_n = \left\{ v \in \mathbb{R}^{U-L+1}_{\geq 0} : \sum_{j=L}^{U} v_j \gamma_j(\ell) = 0 \text{ for } \ell = n - (U-L) + 1, \ldots, n, \ \|v\|_1 = 1 \right\}.$$

If $M_n \neq \emptyset$ then there is a unique minimizer $w^{(n)}$ to the following constraint quadratic optimization problem:

$$\min_{v \in M_n} \|v\|_2^2.$$

Moreover, if $\omega_+ = 1$ then M_n is a singleton set, and $M_n \subseteq K_n$ for $n \geq U$, where K_n is as in (7.1). Furthermore, if there exists a unique stationary measure π of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$, then $w^* = \lim_{n \to \infty} w^{(n)} \in \cap_{n=1}^{\infty} K_n$ exists. In particular, w^* equals the generator of π, that is, $w^* = (\pi(L), \ldots, \pi(U))$, appropriately normalized.

Proof. Existence of the minimizer follows similarly to the proof of Theorem 7.1. If $\omega_+ = 1$ then it follows from Proposition 7.1 below that M_n is a singleton set for $n \geq U$. It follows from Lemma 6.2 that $M_n \subseteq K_n$. Since $w^{(n)} \in M_n \subseteq K_n$, and $\cap_{n=1}^{\infty} K_n$ contains the generator of the unique stationary measure π only, then $v^* = \lim_{n \to \infty} w^{(n)}$ exists and equals the generator of π. □

We refer to the optimization problem outlined in Lemma 7.1 as the linear approximation scheme. For $\omega_+ = 1$, a solution $v^{(n)}$ to the linear approximation scheme automatically fulfills $v(\ell) = \sum_{j=L}^{U} v_j^{(n)} \gamma_j(\ell) \geq 0$ for all $\ell = 0, \ldots, n$. In general, these inequalities need to be validated.

Proposition 7.1. *If $\omega_+ = 1$ then M_n, $n \geq U$, is a singleton set.*

Proof. For $n \geq U$, let G(n) be the $(U-L+1) \times (U-L+1)$ matrix,

$$G(n) = \begin{pmatrix} \gamma_L(n-(U-L-1)) & \gamma_{L+1}(n-(U-L-1)) & \cdots & \gamma_U(n-(U-L-1)) \\ \vdots & \vdots & & \vdots \\ \gamma_L(n-1) & \gamma_{L+1}(n-1) & \cdots & \gamma_U(n-1) \\ \gamma_L(n) & \gamma_{L+1}(n) & \cdots & \gamma_U(n) \\ 1 & 1 & \cdots & 1 \end{pmatrix},$$

and let $c_i(n)$ be the cofactor of G(n) corresponding to the $(U-L+1)$th row and ith column, for $i = 1, \ldots, U-L+1$. Then, $\det(G(n)) = \sum_{i=1}^{U-L+1} c_i(n)$, and there exists a unique solution $v^{(n)}$ to $G(n)v = e_{U-L+1}$, where e_{U-L+1} is the $(U-L+1)$th unit vector in $\mathbb{R}^{U-L+1}$, if and only if $\det(G(n)) \neq 0$. If this is the case, then by Cramer's rule,

$$v_i^{(n)} = \frac{c_i(n)}{\sum_{j=1}^{U-L+1} c_j(n)} \quad \text{for } i = 1, \ldots, U-L+1,$$

and hence, $v^{(n)} \in \mathbb{R}_{\geq 0}^{U-L+1}$ if and only if all cofactors have the same sign or are zero. Hence, we aim to show at least one cofactor is nonzero and that all nonzero cofactors have the same sign. In the following, for convenience, we say the elements of a sequence $a_1, \ldots, a_m$ have the same sign if all nonzero elements of the sequence have the same sign.

For $n \geq U$, define the $(U - L + 2) \times (U - L + 1)$ matrix

$$
\Gamma(n) = \begin{pmatrix}
\gamma_L(n - (U - L)) & \gamma_{L+1}(n - (U - L)) & \cdots & \gamma_U(n - (U - L)) \\
\vdots & \vdots & & \vdots \\
\gamma_L(n - 1) & \gamma_{L+1}(n - 1) & \cdots & \gamma_U(n - 1) \\
\gamma_L(n) & \gamma_{L+1}(n) & \cdots & \gamma_U(n) \\
1 & 1 & \cdots & 1
\end{pmatrix},
$$

and the $(U - L + 1) \times (U - L + 1)$ matrices $\Gamma^\ell(n)$, $\ell = 0, \ldots, U - L$, by removing row $m = U - L + 1 - \ell$ of $\Gamma(n)$ (that is, the $(\ell + 1)$th row counting from the bottom). For notational convenience, noting that the columns of $\Gamma(n)$ take a similar form, we write these matrices as

$$
\Gamma^\ell(n) = \begin{pmatrix}
\gamma_j(n - (U - L)) \\
\vdots \\
\gamma_j(n - (\ell + 1)) \\
\gamma_j(n - (\ell - 1)) \\
\vdots \\
\gamma_j(n) \\
1
\end{pmatrix}, \quad \ell = 0, \ldots, U - L.
$$

Note that

$$
G(n) = \Gamma^{U-L}(n) \quad \text{and} \quad \Gamma^0(n) = \Gamma^{U-L}(n - 1). \tag{7.4}
$$

Let $\Gamma_i^\ell(n)$ be the $(U - L) \times (U - L)$ matrix obtained by removing the bottom row and the ith column from $\Gamma^\ell(n)$. Hence, the cofactor $C_i^\ell(n)$ of $\Gamma^\ell(n)$ corresponding to the $(U - L + 1)$th row and ith column is

$$
C_i^\ell(n) = (-1)^{U-L+1+i} \det(\Gamma_i^\ell(n)) \quad \text{and} \quad C_i^{U-L}(n) = c_i(n), \tag{7.5}
$$

$i = 1, \ldots, U - L + 1$. By induction, we show that the signs of $C_i^\ell(n)$, $i = 1, \ldots, U - L + 1$, are the same, potentially with some cofactors being zero, but at least one being nonzero.

Induction basis: for $n = U$, we have

$$
\Gamma(U) = \begin{pmatrix}
1 & 0 & \cdots & 0 & 0 \\
0 & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & & \vdots & \vdots \\
0 & 0 & \cdots & 1 & 0 \\
0 & 0 & \cdots & 0 & 1 \\
1 & 1 & \cdots & 1 & 1
\end{pmatrix},
$$

and it follows by tedious calculation that

$$
C_{U-L+1-\ell}^\ell(U) = (-1)^\ell, \qquad C_i^\ell(U) = 0 \quad \text{for } \ell \neq U - L + 1 - i,
$$

$i = 1, \ldots, U - L + 1$ and $\ell = 0, \ldots, U - L$. It follows that the C_i^ℓs have the same sign for ℓ fixed and all $i = 1, \ldots, U - L + 1$ (all cofactors are zero, except for $i = U - L + 1 - \ell$).

We postulate that the nonzero elements fulfill

$$\operatorname{sign}(C_i^\ell(n)) = (-1)^{(n-U)(U-L)+\ell} \quad \text{for } n \geq U,$$

and $\ell = 0, \ldots, U - L$, $i = 1, \ldots, U - L + 1$. The hypothesis holds for $n = U$.

Induction step: Assume that the statement is correct for some $n \geq U$. Using $m_* = \omega_+ - \omega_- 1 = U - L + 1$ and (4.4), we obtain, for $\ell = 1, \ldots, U - L$ (excluding $\ell = 0$),

$$\Gamma^\ell(n+1) = \begin{pmatrix} \gamma_j(n+1-(U-L)) \\ \vdots \\ \gamma_j(n+1-(\ell+1)) \\ \gamma_j(n+1-(\ell-1)) \\ \vdots \\ \gamma_j(n) \\ \gamma_j(n+1) \\ 1 \end{pmatrix} = \begin{pmatrix} \gamma_j(n+1-(U-L)) \\ \vdots \\ \gamma_j(n+1-(\ell+1)) \\ \gamma_j(n+1-(\ell-1)) \\ \vdots \\ \gamma_j(n) \\ \sum_{k=1}^{U-L+1} \gamma_j(n+1-k)f_k(n+1) \\ 1 \end{pmatrix}.$$

Hence, using the linearity of the determinant, we obtain, for $0 < \ell \leq U - L$ (excluding $\ell = 0$),

$$\det\left(\Gamma_i^\ell(n+1)\right) = \begin{vmatrix} \gamma_j(n+1-(U-L)) \\ \vdots \\ \gamma_j(n+1-(\ell+1)) \\ \gamma_j(n+1-(\ell-1)) \\ \vdots \\ \gamma_j(n) \\ \gamma_j(n+1-\ell)f_\ell(n+1) \end{vmatrix} + \begin{vmatrix} \gamma_j(n+1-(U-L)) \\ \vdots \\ \gamma_j(n+1-(\ell+1)) \\ \gamma_j(n+1-(\ell-1)) \\ \vdots \\ \gamma_j(n) \\ \gamma_j(n+1-(U-L+1))f_{U-L+1}(n+1) \end{vmatrix}$$

$$= f_\ell(n+1)\begin{vmatrix} \gamma_j(n+1-(U-L)) \\ \vdots \\ \gamma_j(n+1-(\ell+1)) \\ \gamma_j(n+1-(\ell-1)) \\ \vdots \\ \gamma_j(n) \\ \gamma_j(n+1-\ell) \end{vmatrix} + f_{U-L+1}(n+1)\begin{vmatrix} \gamma_j(n+1-(U-L)) \\ \vdots \\ \gamma_j(n+1-(\ell+1)) \\ \gamma_j(n+1-(\ell-1)) \\ \vdots \\ \gamma_j(n) \\ \gamma_j(n+1-(U-L+1)) \end{vmatrix}$$

$$= f_\ell(n+1)(-1)^{\ell-1}\begin{vmatrix} \gamma_j(n-(U-L-1)) \\ \vdots \\ \gamma_j(n) \end{vmatrix}$$

$$+f_{U-L+1}(n+1)(-1)^{U-L-1}\begin{vmatrix} \gamma_j(n-(U-L)) \\ \vdots \\ \gamma_j(n-\ell) \\ \gamma_j(n-(\ell-2)) \\ \vdots \\ \gamma_j(n) \end{vmatrix}$$

$$=f_\ell(n+1)(-1)^{\ell-1}\det(\Gamma_i^{U-L}(n))+f_{U-L+1}(n+1)(-1)^{U-L-1}\det(\Gamma_i^{\ell-1}(n)),$$

where the remaining terms from the linear expansion of the determinant are zero. In the computation of the determinant above, we abuse $\gamma_j(n+1-k)$ for the row vector with the ith coordinate deleted. For $\ell=0$, then using (7.4),

$$\det(\Gamma_i^0(n+1))=\det(\Gamma_i^{U-L}(n)).$$

The above conclusions result in the following for the sign of the cofactors, using (7.5):

$$C_i^\ell(n+1))=f_\ell(n+1)(-1)^{\ell-1}C_i^{U-L}(n) \tag{7.6}$$
$$+f_{U-L+1}(n+1)(-1)^{U-L-1}C_i^{\ell-1}(n),\quad 0<\ell\le U-L$$
$$C_i^0(n+1)=C_i^{U-L}(n). \tag{7.7}$$

We recall some properties of $f_\ell(n)$. According to Lemma A.1(vi), $f_{U-L+1}(n)>0$ for $n\ge U+1$, using $i_+=0$ (otherwise zero is a trapping state) and $-\omega_-=U-L+1$. For $0\le\ell<-\omega_-$, we have $\operatorname{sgn}(\omega_-+\ell+1/2)=-1$, and hence, $f_\ell(n)\le 0$ for $n\ge U+1$, according to Lemma A.1(vii) and 6(viii) in the Appendix. Consequently, the sign of the two terms in (7.6) are

$$(-1)(-1)^{\ell-1}(-1)^{(U-L)(n-U)+(U-L)}=(-1)^{(U-L)(n+1-U)+\ell},$$
$$(+1)(-1)^{U-L+1}(-1)^{(U-L)(n-U)+\ell-1}=(-1)^{(U-L)(n+1-U)+\ell};$$

hence, the sign of $C_i^\ell(n+1)$ corroborates the induction hypothesis. The sign of the term in (7.7) is

$$(-1)^{U-L}(-1)^{(U-L)(n-U)}=(-1)^{(U-L)(n+1-U)+0},$$

again in agreement with the induction hypothesis.

It remains to show that at least one cofactor is nonzero, that is, $C_i^{U-L}(n)\ne 0$ for at least one $1\le i\le U-L+1$ and $n\ge U$. Let $a_{n,\ell}=\sum_{i=1}^{U-L+1}|C_i^\ell(n)|$. From (7.4) and (7.6), we have

$$a_{n+1,\ell}=|f_\ell(n+1)|a_{n,U-L}+|f_{U-L+1}(n+1)|a_{n,\ell-1},\quad 1\le\ell\le U-L, \tag{7.8}$$
$$a_{n+1,0}=a_{n,U-L},$$

for $n\ge U$. We show by induction that $a_{n,\ell}\ne 0$ for $n\ge U$ and $0\le\ell\le U-L$. Hence, the desired conclusion follows. For $n=U$, we have $a_{U,\ell}=1$ for all $\ell=0,\ldots,U-L$. Assume that $a_{n,\ell}\ne 0$ for $\ell=0,\ldots,U-L$ and some $n\ge U$. Since $f_{U-L+1}(n+1)>0$ for $n+1\ge U+1$, then it follows from (7.8) that $a_{n+1,\ell}\ne 0$ for $\ell=0,\ldots,U-L$. The proof is completed. □

8. Examples

To end, we present some examples using the linear approximation scheme and the convex constrained optimization approach. We use the criteria in [28, Theorem 7] to check whether

a CTMC with polynomial transition rates is positive recurrent, null recurrent, transient and nonexplosive, or explosive. These properties hold for either all PICs or none, provided the transition rate functions are polynomials for large state values, as in mass-action kinetics [28, Theorem 7].

We have made two implementations of the numerical methods. One in R and one in Python (only mass-action kinetics) with codes available on request. In all examples, the code runs in a few seconds on a standard laptop. The two codes agree when run on the same example. We have not aimed to optimize for speed. In Figures 1–5, 'State x' refers to the state of the original Markov chain, and 'Index n' refers to the translated Markov chain, the index of, for example, $\gamma_j^s(n)$. The index s refers to the irreducibility class in Proposition 7.1.

**Example 8.1.** *Consider the SRN with mass-action kinetics,*

$$\text{S} \xrightarrow{\kappa_1} 2\text{S} \underset{\kappa_4}{\overset{\kappa_2}{\rightleftharpoons}} 3\text{S} \xrightarrow{\kappa_3} \text{S}.$$
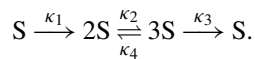
We have $\omega_+ = 1$ and $\omega_- = -2$ with $s = 1$ (zero is a neutral state), and $L_1 = 0$, $U_1 = 1$. Furthermore, a unique stationary distribution exists since the SRN is positive recurrent for all positive rate constants [28, Theorem 7], and $\pi_1(x) = \pi(1 + x)$. As the reaction network is weakly reversible (the reaction graph consists of a finite union of disjoint strongly connected components), then it is complex balanced for specific choices of rate constants, yielding a Poisson stationary distribution [2]. This is the case if and only if $\kappa_1(\kappa_3 + \kappa_4)^2 = \kappa_2^2 \kappa_3$.

Here, we focus on the stability of the numerical approximations using the linear approximation scheme and convex constrained optimization for a single set of parameters, $\kappa_1 = 40$, $\kappa_2 = 22$, $\kappa_3 = \kappa_4 = 1$; see Figure 4. Convex constrained optimization fails for $n > 18$ in (7.2) due to exponentially increasing $\gamma_j^1(\ell)$ values with alternating signs. In contrast, the linear approximation scheme is quite robust and returns accurate estimates for the generating terms $\pi_1(0)$, $\pi_1(1)$ ($= \pi(1), \pi(2)$), even for $n = 70$. However, in this situation, inaccurate and negative probability values for large state values are returned; see Figure 4. The estimated values of $\pi(32)$ and $\pi(33)$ are zero to the precision of the computer and the first negative estimate is $\pi(34) = -8.4 \cdot 10^{-13}$. From then on, the estimated probabilities increase in absolute value. The estimated generating terms for the convex constrained optimization problem with $n = 18$ and the linear approximation scheme with $n = 25$ deviate on the seventh decimal point only. In the latter case, the estimates remain unchanged for $25 \leq n \leq 70$ for up to seven decimal points, despite the fact that negative probabilities are found for large n.

It is remarkable that, for $n = 70$ with $\gamma_j^1(n)$ of the order $e^{50} \approx 10^{22}$, we still numerically find that M_n is a singleton set, as postulated in Proposition 7.1, despite the fact that the solution gives rise to instabilities in calculating the probabilities. Also the numerical computations confirm that the limit in (6.8) in Lemma 6.3 is zero, as $\gamma_j^1(n)$ increases beyond bound.

The following example demonstrates that both the linear approximation scheme and the convex optimization approach can be efficient in practice for ergodic CTMCs.

Example 8.2. *For the SRN with mass-action kinetics*

$$\text{S} \xrightarrow{\kappa_1} 3\text{S} \xrightarrow{\kappa_2} 2\text{S} \xrightarrow{\kappa_3} 4\text{S} \xrightarrow{\kappa_4} \text{S}, \tag{8.1}$$
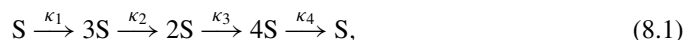
we obtain $\omega_+ = 2$, $\omega_- = -3$, and $s = 1$, $L_1 = 0$, $U_1 = 2$, such that there is one PIC with state space $\mathbb{N}$. Despite the fact that Proposition 7.1 does not apply (as $\omega_+ \neq 1$), numerically we find that M_n is a singleton set. Using the linear approximation scheme or the convex optimization
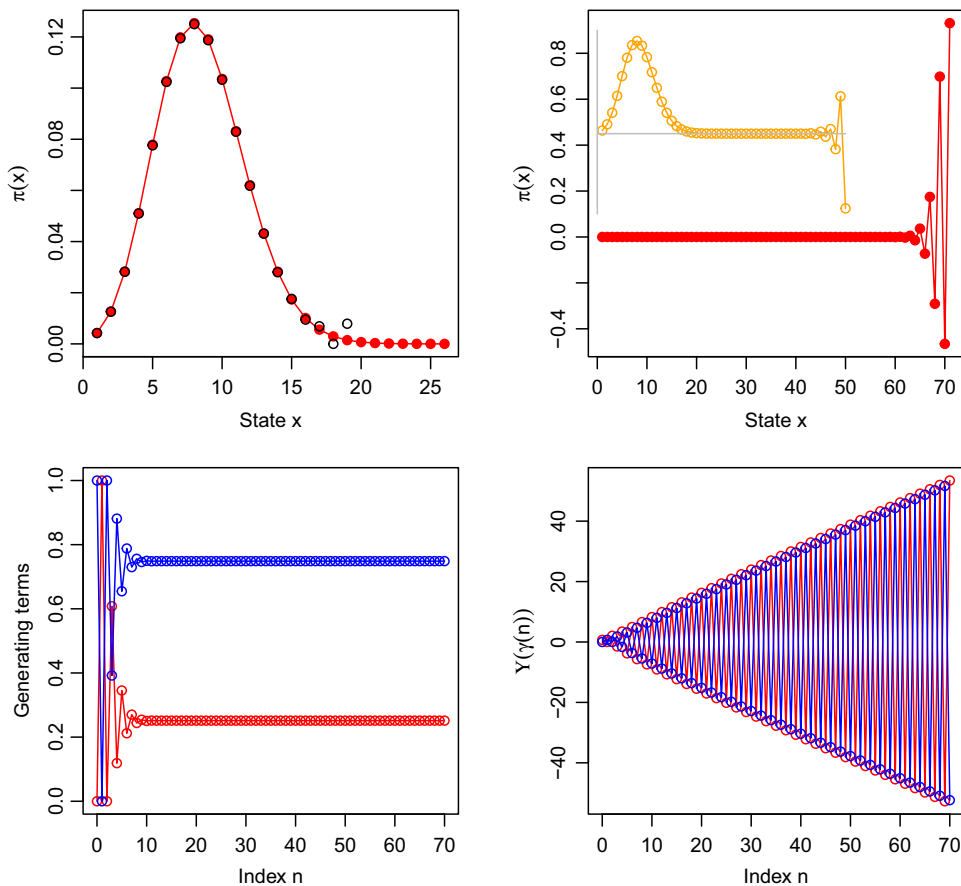
FIGURE 1. (a) The stationary distribution calculated using the linear approximation scheme with $n = 25$ (red dots) and convex constrained optimization with $n = 18$ (black circles). The latter results in wrong probability estimates for the states $x = 18$ and $x = 19$. (b) The stationary distribution calculated using the linear approximation scheme with $n = 70$ (red dots). The orange subfigure is a blow up of the first 50 states for $n = 70$, normalized to their sum, indicating that the correct form of the distribution is retrieved even in the presence of instability, and the onset of instabilities. (c) Approximate values of $\pi_1(L_1)$ and $\pi_1(U_1)$ as a function of n found using the linear approximation scheme. Convergence is very fast. (d) The values of $\gamma_{L_1}^1(n)$ and $\gamma_{U_1}^1(n)$ for $n = 0, \ldots, 70$. The coefficients are plotted as $\Upsilon(x) = \log(x + 1)$ for $x \geq 0$ and $\Upsilon(x) = -\log(-x + 1)$ for $x \leq 0$. Dots are connected by lines for convenience.

approach, we obtain a rather quick convergence; see Figure 2. In this case, the coefficients $\gamma_j^1(\ell)$, $j = 0, 1, 2$, decrease fast towards zero, taking both signs. Both algorithms run efficiently even for large n as the coefficients vanish. Figure 2(d) shows $\gamma(n)/\|\gamma(n)\|_1$. There appears to be a periodicity of $U_1 - L_1 + 1 = 3$, demonstrating numerically that the matrices $A(3n)$, $A(3n + 1)$, and $A(3n + 2)$, $n \in \mathbb{N}_0$, each converges as $n \to \infty$; see Theorem 6.3. The generator recovered from either of the three sequences $A(3n)$, $A(3n + 1)$, and $A(3n + 2)$ agree to high accuracy, and agree with the generator found using the linear approximation scheme.
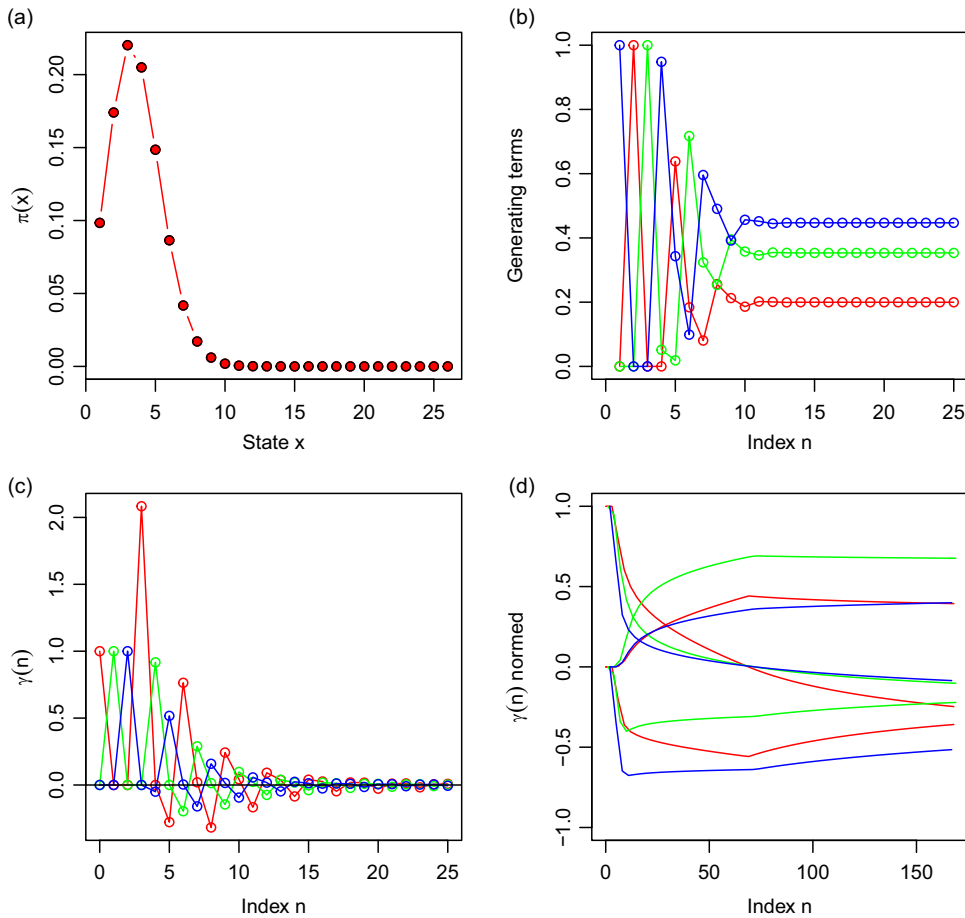
FIGURE 2. (a) Stationary distribution for (8.1) with $\kappa_1 = 50$, $\kappa_3 = 5$, and $\kappa_2 = \kappa_4 = 1$. The linear approximation scheme is shown in red, while the convex constrained optimization scheme is overlaid in black. Dots are connected by lines for convenience. (b,c) Convergence of the generating terms is fast as $\gamma_j^s(\ell)$ decreases fast to zero with ℓ becoming large. (d) Shown here is $\gamma(3n)/\|\gamma(3n)\|_1$ (red), $\gamma(3n+1)/\|\gamma(3n+1)\|_1$ (green), and $\gamma(3n+2)/\|\gamma(3n+2)\|_1$ (blue). The numerical computations clearly indicate periodicity. Note the fact that despite convergence has not been achieved for the simulated values of n, the generators recovered from the three series $A(3n)$, $A(3n+1)$, $A(3n+2)$ agree with those found from the linear approximation scheme; see Theorem 6.3. Dots are replaced by lines for visual reasons.

Although, in theory, it seems to make sense to use the linear approximation scheme *only* for stationary measures for which $\pi(n)/\|\gamma(n)\|$ is vanishing, in practice, it seems that the linear approximation scheme still captures the feature of a stationary measure with nonvanishing $\pi(n)/\|\gamma(n)\|_1$ decently, when the states are notr too large.

Example 8.3. *Computing an unknown distribution. We give an example of a mass-action SRN,*

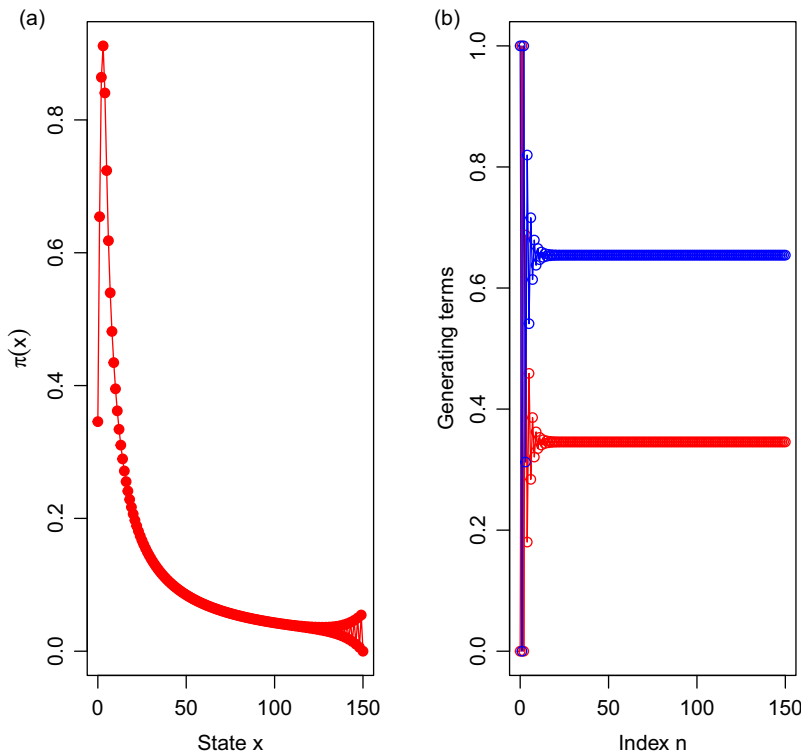$$0 \xrightarrow{10} S \xrightarrow{12} 2S \xrightarrow{1} 6S, \qquad 2S \xrightarrow{2} 0,$$
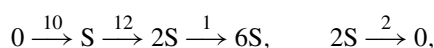
FIGURE 3. (a) The stationary measure computed with the linear approximation scheme and $n = 150$. For large states, significant errors occur in the estimation. (b) The generating terms.

which is null recurrent by the criterion in [28, Theorem 7], and hence, there exists a unique stationary measure due to [23, Theorem 3.5.1] and [7, Theorem 1]. In this case, $\omega_1 = -2$ and $\omega_+ = 4$. Furthermore, $s = 0$ and $L_0 = 0$, $U_0 = 1$. We apply the linear approximation scheme to the SRN with $n = 150$ and find that M_n is a singleton set, despite the fact that $\omega_+ \neq 1$; see Proposition 7.1. For large states, the point measures are considerably far from zero; see Figure 3. Moreover, instabilities occur. The inaccuracies in the values are due to small inaccuracies in the estimation of the generating terms and the large coefficients $\gamma_j(\ell)$ that increase exponentially.

We know from Corollary 5.4 that there exists a unique stationary measure for CTMCs with polynomial transition rates if $\omega_- = -2$ and $\omega_+ = 1$, it remains to see if such a stationary measure is finite. With the aid of our numerical scheme, we might be able to infer this information in some scenarios.

Example 8.4. *Computing an unknown measure. Consider the following SRN,*

$$10S \xrightarrow{5000} 12S \xrightarrow{10} 13S \xrightarrow{1} 16S, \qquad 13S \xrightarrow{1} 10S.$$
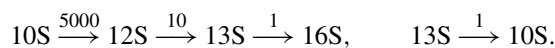
It is explosive by the criterion in [28, Theorem 7]. We have $s = 10$, $\omega_- = -3$, $\omega_+ = 3$, $L_{10} = 0$, and $U_{10} = 2$. The linear approximation scheme retrieves what appears to be a stationary distribution; see Figure 4. The numerical computations confirm that the limit in (6.5) in Theorem 6.3 is zero, pointing to the stationary distribution being unique.
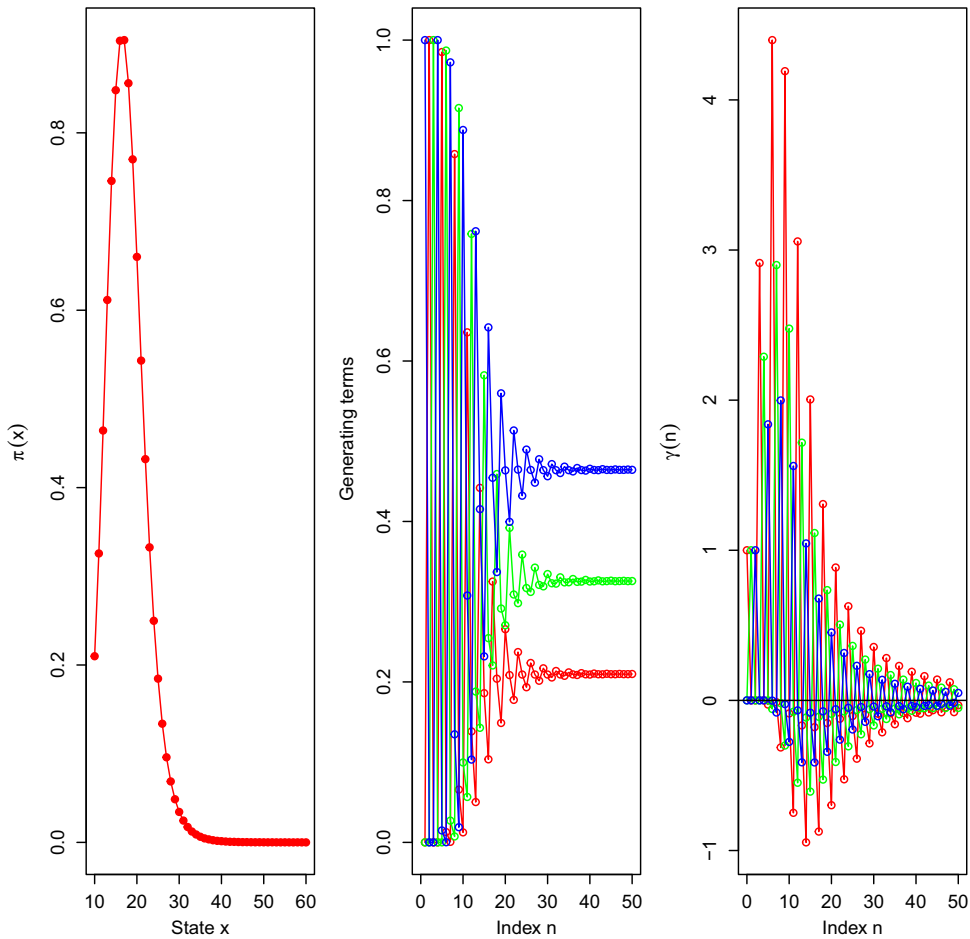
FIGURE 4. The stationary measure of an explosive SRN, the generating terms, and the coefficients $\gamma_j^s(\ell)$.

We end with an example for which there exists more than one stationary measure.

Example 8.5. *Consider the SRN with reactions* $0 \xrightarrow{\lambda_1} S$, $2S \xrightarrow{\lambda_{-2}} 0$, *and*

$$\lambda_1(0) = \lambda_1(1) = 1, \qquad \lambda_1(x) = \left\lfloor \frac{x}{2} \right\rfloor^2, \quad x \geq 2,$$

$$\lambda_{-2}(0) = \lambda_{-2}(1) = 0, \qquad \lambda_{-2}(x) = \left\lfloor \frac{x+1}{2} \right\rfloor^{-2}, \quad x \geq 2.$$

Then, (A1)–(A3) are fulfilled with $\omega_- = -2$, $\omega_+ = 1$, $s = 0$, $L_0 = 0$, $U_0 = 1$ *(so* $\pi_0 = \pi$*). From Corollary 5.3 with* $x = U_0 + 2 = 3$,

$$Q_n(3) = \prod_{i=0}^{n} \frac{\lambda_1(2+2i)\lambda_{-2}(3+2i)}{\lambda_1(3+2i)\lambda_{-2}(4+2i)} = 1,$$

$$H(3) = \sum_{n=0}^{\infty} \frac{1}{(n+1)^2} + \frac{1}{(n+2)^2} = -1 + \sum_{n=1}^{\infty} \frac{2}{n^2} = \frac{\pi^2}{3} - 1 < \infty.$$
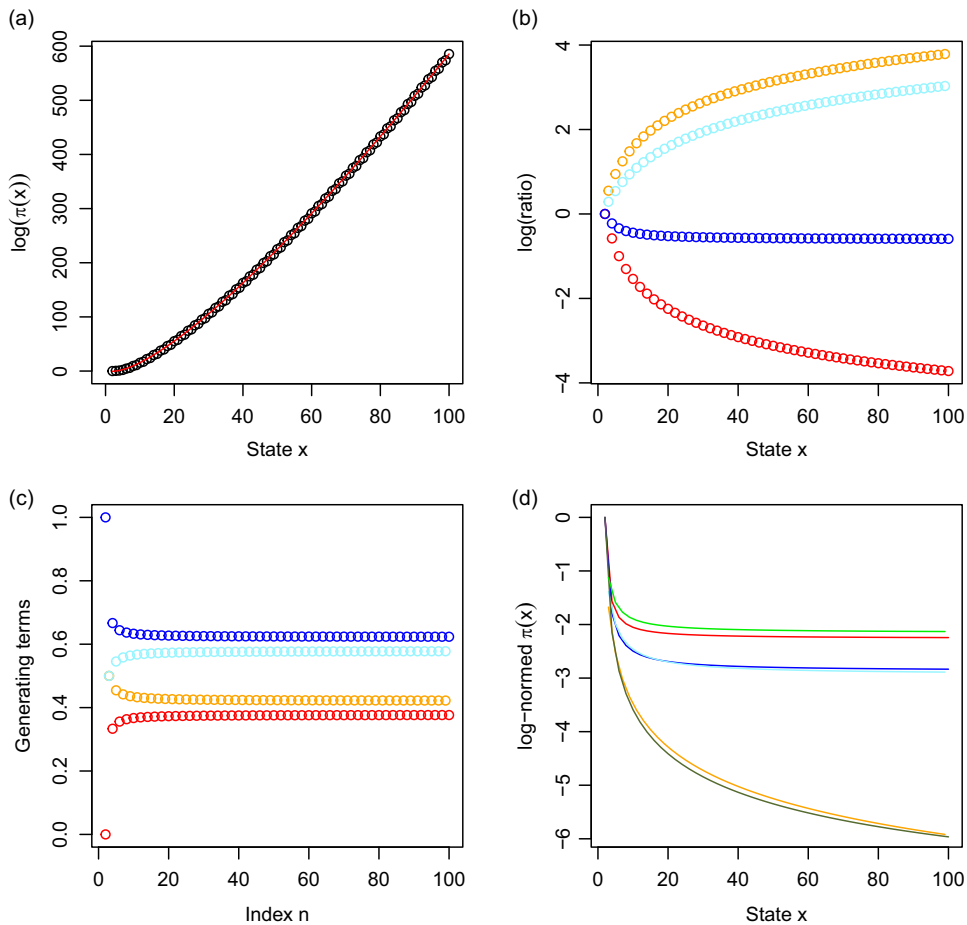
FIGURE 5. (a) The logarithm of the stationary measure with $\phi^* = 2.67$. This value is in the upper end of the estimated interval $[\Psi_2(3), \Psi_1(3)]$. The red line is the curve fitted to the log measure of the even states, using regression analysis. The stationary measures retrieved using the linear approximation scheme in Section 7 are visually identical to the plotted measure, not shown. (b) The difference in log $(\pi(x))$ between the measure with $\phi^* = 2.00$ (in the middle part of the interval) to that with $\phi^* = 2.67$ (blue: even states; light blue: odd states), and the log measure between the measure with $\phi^* = 1.54$ (in the lower end of the interval) to that with $\phi^* = 2.67$ (red: even states; orange: odd states). An alternating pattern emerges. (c) The generating terms estimated for different values of n, with red (even states)/orange (odd states) showing $\pi(L_0)$ and blue (even states)/light blue (odd states) showing $\pi(U_0)$. (d) The normed measure, $\pi(n)/\|\gamma(n)\|_1$ for different values of ϕ^*: 2.67 (red: even states; orange: odd states), 2.00 (blue: even states; light blue: odd states) and 1.54 (olive: even states; green: odd states). Dots are replaced by lines for visual reasons.

Consequently, there is not a unique stationary measure of $(\Omega, \mathcal{F})$ on $\mathbb{N}_0$. Numerical computations suggest that $[\Psi_2(3), \Psi_1(3)] \approx [1.5351, 2.6791]$, using (5.7) with $k = 700$. See Corollary 5.3 for a definition of Ψ_1, Ψ_2.

In the following, we explore the stationary measures using the theory of Section 5.2 and compare it to results obtained from the linear approximation scheme. For any stationary

measure π, by definition of $\phi(x)$ and $\phi(x) < h(x)$, we have, for $x \geq 3$,

$$\frac{\pi(x)}{\pi(x-1)} = \frac{\lambda_{-2}(x-1)}{\lambda_{-2}(x)} \frac{1}{\phi(x)} > \frac{\lambda_{-2}(x-1)}{\lambda_{-2}(x)} \frac{1}{h(x)} = \left(\left\lfloor \frac{x-1}{2} \right\rfloor \left\lfloor \frac{x+1}{2} \right\rfloor \right)^2,$$

hence,

$$\pi(x) > C \frac{(x!)^4}{16^x}, \quad x \geq 0,$$

for some constant $C > 0$, and π is not a distribution. This is shown in Figure 5(a) for the logarithm of $\pi(x)$ with $\phi^* = 2.67$, using Corollary 5.3. The red line is a fitted curve to $\log(\pi(x))$ for the *even* states: $\log(\pi(2x)) \approx 3.504 + 2.009x \log(x) - 3.429x$, $x \geq 2$. The errors between true and predicted values are numerically smaller than 0.1 for all even states. Hence, the measure appears to grow super-exponentially fast. The function $\log(\pi(x))$ for the odd states grows at a comparable rate as that for the even states, but not with the same regularity.

Additionally, we computed the difference in $\log(\pi(x))$ for different values of ϕ^*, showing again a distinction between odd and even states; see Figure 5(b).

We used the linear approximation scheme to estimate the generating terms (which we know are not uniquely given in this case). Also here, an alternating pattern emerges with even indices producing the generator $(\pi(L_0), \pi(U_0)) \approx (0.3764, 0.6235)$ ($n = 100$), while odd indices produce the generator $\approx (0.4222, 0.5777)$ ($n = 101$). For each n, a unique solution is found. Computing the corresponding $\phi^* = \phi(U_0 + 2)$ yields another approximation to $[\Psi_2(3), \Psi_1(3)]$, namely $[1.5240, 2.7161]$, which is slightly larger than the previous approximation, $[1.5351, 2.6791]$. By nature of the latter estimate, the first coordinate is increasing in n, while the second is decreasing in n, hence, the latter smaller interval is closer to the true interval $[\Psi_2(3), \Psi_1(3)]$ than the former. Figure 5(c) also shows the estimated generating terms for different values of n, providing a band for $\pi(L_0)$ and $\pi(U_0)$ for which stationary measures exist, in agreement with Lemma 6.3.

Finally, we compute the ratio of $\pi(n)$ to $\|\gamma(n)\|_1$ for different values of ϕ^*, and observe that there appears to be one behavior for odd states and one for even. While one cannot infer the large state behavior of the ratios in Figure 5(d) from the figure, it is excluded by nonuniqueness of the measures, that they both converge to zero; see Lemma 6.3.

Example 8.6. *Consider the SRN with non-mass-action kinetics*

$$0 \underset{\lambda_{-1}}{\overset{\lambda_1}{\rightleftharpoons}} S, \qquad 2S \overset{\lambda_{-2}}{\longrightarrow} 0,$$

with transition rates given by

$$\lambda_1(x) = 2^x, \quad x \geq 0,$$

$$\lambda_{-1}(0) = 0, \qquad \lambda_{-1}(1) = 2, \qquad \lambda_{-1}(x) = 2 \cdot 4^{x-1} - 2^x, \quad x \geq 2,$$

$$\lambda_{-2}(0) = \lambda_{-2}(1) = 0, \qquad \lambda_{-2}(x) = 2 \cdot 4^{x-1}, \quad x \geq 2.$$

For this SRN, $\omega_- = -2$, $\omega_+ = 1$, $s = 0$, $L_0 = 0$, and $U_0 = 1$. Hence, the conditions of Proposition 7.1 are satisfied. The underlying CTMC is from [22], where it is shown that $v(x) = (-1/2)^{x+1}$ for $x \in \mathbb{N}_0$ is a signed invariant measure. This measure has generator $(v(0), v(1)) = (-1/2, 1/4)$. The space of signed invariant measures is $U_0 - L_0 + 1 = 2$-dimensional, and a second linearly independent signed invariant measure has generator $(1,0)$. On the other hand, by the Foster–Lyapunov criterion [21], the process is positive recurrent, and hence, admits a stationary distribution. Numerical computations show that the stationary

distribution is concentrated on the first few states $x = 0, \ldots, 3$. The uniqueness of the stationary distribution is also confirmed by Corollary 5.3 in that $H(U + 2) = H(3) = \infty$ by a simple calculation.

Appendix A.

Lemma A.1. *Assume that (A1)–(A3) hold. Then the following assertions hold.*

(i) $B_0 = \{\omega_-\}$, $\omega_- \in B_k$ for $0 \leq k < -\omega_-$, and $\omega_+ \in B_k$ for $\omega_- \leq k \leq m_*$.

(ii) $c_0(\ell) < 0$ for $\ell > U$ and $c_0(\ell) = 0$ for $\ell \leq U$.

(iii) $c_k(\ell) > 0$ for $-\omega_- \leq k \leq m_*$ and $\ell \geq \mathsf{i}_{\omega_+}$. Generally, $c_k(\ell) \geq 0$, $\ell \in \mathbb{N}_0$.

(iv) $c_k(\ell) < 0$ for $0 \leq k < -\omega_-$ and $\ell \geq \mathsf{i}_{\omega_-} = U + 1$. Generally, $c_k(\ell) \leq 0$, $\ell \in \mathbb{N}_0$.

(v)   If $c_k(\ell) > 0$ then $c_k(\ell + 1) > 0$, and similarly, if $c_k(\ell) < 0$ then $c_k(\ell + 1) < 0$ for $k \in \{0, \ldots, m_*\}$ and $\ell \in \mathbb{N}_0$.

(vi) $f_k(\ell) > 0$ for $-\omega_- \leq k \leq m_*$ and $\ell - k \geq \mathsf{i}_{\omega_+}$.

(vii) $f_k(\ell) < 0$ for $0 \leq k < -\omega_-$ and $\ell - k \geq \mathsf{i}_{\omega_-} = U + 1$.

(viii) If $f_k(\ell) > 0$ then $f_k(\ell + 1) > 0$, and similarly, if $f_k(\ell) < 0$ then $f_k(\ell + 1) < 0$ for $k \in \{0, \ldots, m_*\}$ and $\ell > U$.

Proof. (i) Since $\omega_- \leq -1$, we have from (4.7),

$$B_0 = \left\{ \omega \in \Omega \mid \left(\omega_- + \tfrac{1}{2} \right)\left(\omega - \omega_- \tfrac{1}{2} \right) > 0 \right\} = \{ \omega \in \Omega \mid \omega \leq \omega_- \} = \{\omega_-\}.$$

For $-\omega_- \leq k \leq m_*$, we have $k' = \omega_- + k + 1/2 > 0$ and $\omega - k' \geq \omega - (\omega_- + m_* + 1/2) = \omega - \omega_+ + 1/2 > 0$ if $\omega = \omega_+$. Hence, $\omega_+ \in B_k$. For $0 \leq k < -\omega_-$, we have $k' = \omega_- + k + 1/2 < 0$ and $\omega - k' \leq \omega - (\omega_- + 1/2) < 0$ if $\omega = \omega_-$. Hence, $\omega_- \in B_k$. (ii) Since $\mathrm{sgn}(\omega_- + 1/2) = -1$, then for $\ell > U = \mathsf{i}_{\omega_-} - 1$, we have $c_0(\ell) = -\lambda_{\omega_-}(\ell) < 0$; see (4.6). Likewise, for $\ell \leq U$. (iii, iv) The sign follows similarly to the proof of (ii). Using (i), the conditions on ℓ ensure that $\lambda_{\omega_+}(\ell) > 0$ and $\lambda_{\omega_-}(\ell) > 0$, respectively, yielding the conclusions. (v) It follows from (A2). (vi)–(viii) Similar to (iii) and (iv) using (4.5) and (4.7). $\square$

For a matrix $D = (d_{ij}) \in \mathbb{R}^{n \times n}$, the jth column, $j = 1, \ldots, n$, is *weakly diagonally dominant* (WDD) if $|d_{jj}| \geq \sum_{i \neq j} |d_{ij}|$ and *strictly diagonally dominant* (SDD) if $|d_{jj}| > \sum_{i \neq j} |d_{ij}|$. In particular, a matrix is WDD if all columns are WDD. A WDD matrix is *weakly chain diagonally dominant* (WCDD) if for each WDD column, say j, which is not SDD, there exists an SDD column k such that there is a directed path from vertex k to vertex j in the associated digraph. Every WCDD matrix is nonsingular [5].

Lemma A.2. *Assume that (A1)–(A3) hold. Then, the row echelon form G of H, as defined in (4.8), exists.*

Proof. Recall that

$$H_{m,n} = \delta_{m,n} - \frac{\lambda_{(m-n)}(n)}{\sum_{\omega \in \Omega} \lambda_\omega(m)}, \quad m = 0, \ldots, L - 1, \ n = 0, \ldots, U - 1.$$

Note that if $\omega < -m$ then $m < \mathsf{i}_\omega$ and, hence, $\lambda_\omega(m) = 0$. Therefore, we might replace $\sum_{\omega \in \Omega} \lambda_\omega(m)$ by $\sum_{k=-m}^{\infty} \lambda_k(m)$ in $H_{n,m}$.

Define $a_{m,n} = \lambda_{(m-n)}(n)$ for $n, m \in \mathbb{Z}$. Then the row echelon form $-H$ restricted to its first L columns is invertible, that is, if

$$
\begin{pmatrix}
1 & -\dfrac{a_{0,1}}{\sum_{k=0}^{\infty} a_{k,0}} & \cdots & -\dfrac{a_{0,L-1}}{\sum_{k=0}^{\infty} a_{k,0}} \\[2mm]
-\dfrac{a_{1,0}}{\sum_{k=0}^{\infty} a_{k,1}} & 1 & \cdots & -\dfrac{a_{1,L-1}}{\sum_{k=0}^{\infty} a_{k,1}} \\[2mm]
\vdots & \vdots & \ddots & \vdots \\[2mm]
-\dfrac{a_{L-1,0}}{\sum_{k=0}^{\infty} a_{k,L-1}} & -\dfrac{a_{L-1,1}}{\sum_{k=0}^{\infty} a_{k,L-1}} & \cdots & 1
\end{pmatrix}
$$

is invertible. Multiplying the above matrix by the invertible diagonal matrix

$$
\mathrm{diag}\left(\sum_{k=0}^{\infty} a_{k,0}, \sum_{k=0}^{\infty} a_{k,1}, \ldots, \sum_{k=0}^{\infty} a_{k,L-1} \right)
$$

on the left side and its inverse on the right side gives a column diagonally dominant matrix

$$
A = \begin{pmatrix}
1 & -\dfrac{a_{0,1}}{\sum_{k=0}^{\infty} a_{k,1}} & \cdots & -\dfrac{a_{0,L-1}}{\sum_{k=0}^{\infty} a_{k,L-1}} \\[2mm]
-\dfrac{a_{1,0}}{\sum_{k=0}^{\infty} a_{k,0}} & 1 & \cdots & -\dfrac{a_{1,L-1}}{\sum_{k=0}^{\infty} a_{k,L-1}} \\[2mm]
\vdots & \vdots & \ddots & \vdots \\[2mm]
-\dfrac{a_{L-1,0}}{\sum_{k=0}^{\infty} a_{k,0}} & -\dfrac{a_{L-1,1}}{\sum_{k=0}^{\infty} a_{k,1}} & \cdots & 1
\end{pmatrix} .
$$

Note that $\sum_{k=1}^{\infty} a_{k,0} = \sum_{k=1}^{\infty} \lambda_k(0) > 0$. Furthermore, by (**A2**) the following property holds: $a_{m,n} > 0$ implies that $a_{m+1,n+1} > 0$ for any $m, n \in \mathbb{Z}$, or by contraposition, $a_{m+1,n+1} = 0$ implies that $a_{m,n} = 0$ for any $m, n \in \mathbb{Z}$. Hence,

$$
\sum_{k=1}^{\infty} a_{k+L-1,L-1} = \sum_{k=L}^{\infty} a_{k,L-1} > 0,
$$

and therefore,

$$
\sum_{k=0}^{L-1} a_{k,L-1} < \sum_{k=0}^{\infty} a_{k,L-1}.
$$

Consequently, the Lth column sum of A is positive, implying A is a WDD matrix with a SDD column L.

By Lemma A.3, A is invertible, and hence, the row reduced echelon form exists. $\qquad\square$

Lemma A.3. *Let $n \in \mathbb{N}$ and assume that D is an $n \times n$ matrix, such that*

$$
D = \begin{pmatrix}
1 & -\dfrac{d_{0,1}}{\sum_{k=0}^{\infty} d_{k,1}} & \cdots & -\dfrac{d_{0,n-1}}{\sum_{k=0}^{\infty} d_{k,n-1}} \\[2mm]
-\dfrac{d_{1,0}}{\sum_{k=0}^{\infty} d_{k,0}} & 1 & \cdots & -\dfrac{d_{1,n-1}}{\sum_{k=0}^{\infty} d_{k,n-1}} \\[2mm]
\vdots & \vdots & \ddots & \vdots \\[2mm]
-\dfrac{d_{n-1,0}}{\sum_{k=0}^{\infty} d_{k,0}} & -\dfrac{d_{n-1,1}}{\sum_{k=0}^{\infty} d_{k,1}} & \cdots & 1
\end{pmatrix} ,
$$

where $d_{i,j} \geq 0$ for $i \in \mathbb{N}_0$, $j = 0, \ldots, n-1$, $\sum_{i=1}^{\infty} d_{i,0} > 0$, and $d_{i+1,j+1} = 0$ implies that $d_{i,j} = 0$ for any $i \in \mathbb{N}_0$, $j = 0, \ldots, n-2$. Then D is WCDD, and thus, nonsingular.

Proof. Number rows and columns from 0 to $n-1$. If $n=1$ then the statement is trivial. Hence, assume that $n>1$.

Fact 1. If column $j>0$ sums to zero then column $j-1$ sums to zero. Indeed, that column j sums to zero is equivalent to $\sum_{i=n}^{\infty} d_{i,j}=0$, which by the property of $d_{i,j}$, implies that

$$\sum_{i=n}^{\infty} d_{i-1,j-1}=0. \tag{A.1}$$

Hence, also $\sum_{i=n}^{\infty} d_{i,j-1}=0$ and column $j-1$ sums to zero. Consequently, if column j is WDD, but not SDD, then all columns $0,\ldots,j$ are WDD, but not SDD.

Fact 2. If column $j>0$ sums to zero then from (A.1) it holds that $d_{n-1,j-1}=0$. Inductively, using in addition fact 1, $d_{i,k}=0$ for $i-k\geq n-j$, corresponding to a lower left triangular corner of size j. $\qquad\square$

Acknowledgements

Funding information

Competing interests

The authors declare no competing interests.

References

[1] ABRAMOWITZ, M. AND STEGUN, I.A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, volume 55 of *Applied Mathematics Series*. National Bureau of Standards, 1972.

[2] ANDERSON, D.F, CRACIUN, G. AND KURTZ, T.G. Product-form stationary distributions for deficiency zero chemical reaction networks. *Bulletin Mathematical Biology*, **72**:1947–1970, 2010.

[3] ANDERSON, D.F. AND KURTZ, T.G. *Stochastic Analysis of Biochemical Systems*, volume 1.2 of *Mathematical Biosciences Institute Lecture Series*. Springer International Publishing, Switzerland, 2015.

[4] ANDERSON, W.J. *Continuous-Time Markov Chains: An Applications–Oriented Approach. Springer Series in Statistics: Probability and its Applications*. Springer-Verlag, New York, 1991.

[5] AZIMZADEH, P. A fast and stable test to check if a weakly diagonally dominant matrix is a nonsingular M-matrix. *Mathematics of Computation*, **88**(316):783–800, 2019.

[6] BOYD, S.P. AND VANDENBERGHE, L. *Convex Optimization*. Cambridge University Press, 2004.

[7] DERMAN, C. Some contributions to the theory of denumerable Markov chains. *Transactions of the American Mathematical Society*, **79**(2):541–555, 1955.

[8] DOUC, R., MOULINES, E., PRIOURET, P. AND SOULIER, P. *Markov Chains. Springer Series in Operations Research and Financial Engineering*. Springer 2018.

[9] ERBAN, R., CHAPMAN, S.J. AND MAINI, P.K. A practical guide to stochastic simulations of reaction-diffusion processes. arXiv:0704.1908, 2017.

[10] ETHIER, S.N. AND KURTZ, T.G. *Markov Processes. Wiley Series in Probability and Mathematical Statistics*. Wiley, 1986.

[11] FALK, J., MENDLER, M. AND DROSSEL, B. A minimal model of burst-noise induced bistability. *PloS one*, **12**(4):e0176410, 2017.

[12] GILLESPIE, D.T. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, **81**:2340–2361, 1977.

[13] GILLESPIE, D.T. A rigorous derivation of the chemical master equation. *Physica A: Stat. Mech. Appl.*, **188**:404–425, 1992.

[14] GUPTA, A., MIKELSON, J. AND KHAMMASH, M. A finite state projection algorithm for the stationary solution of the chemical master equation. *The Journal of Chemical Physics*, **147**(15):154101, 2017.

[15] HARRIS, T.E. Transient Markov chains with stationary measures. *Proceedings of the American Mathematical Society*, **8**(5):937–942, 1957.

[16] JAHNKE, T. AND HUISINGA, W. Solving the chemical master equation for monomolecular reaction systems analytically. *J. Math. Biol.*, **54**:1–26, 2007.

[17] KELLY, F.P. *Reversibility and Stochastic Networks. Applied Stochastic Methods Series*. Cambridge University Press, 2011.

[18] KHINCHIN, A. Ya. *Continued Fractions*. Dover Publications, revised edition, 1997.

[19] KUNTZ, J., THOMAS, P., STAN, GUY-BART AND BARAHONA, M. Stationary distributions of continuous-time Markov chains: a review of theory and truncation-based approximations. *SIAM Review*, **63**(1):3–64, 2021.

[20] LÓPEZ-CAAMAL,F. AND MARQUEZ-LAGO,T. T. Exact probability distributions of selected species in stochastic chemical reaction networks. *Bulletin of Mathematical Biology*, **76**:2334–2361, 2014.

[21] MEYN, S. P. AND TWEEDIE, R. L. *Markov Chains and Stochastic Stability*. Cambridge University Press, 2nd edition, 2009.

[22] MILLER Jr, R. G. Stationary equations in continuous-time Markov chains. *Transactions of the American Mathematical Society*, **109**(1):35–44, 1963.

[23] NORRIS,J.R. *Markov Chains. Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, 2009.

[24] PASTOR-SATORRAS, R., CASTELLANO, C., VAN MIEGHEM, P. AND VESPIGNANI, A. Epidemic processes in complex networks. *Rev. Mod. Phys.*, **87**:925–979, 2015.

[25] SHAHREZAEI, V. AND SWAIN, P.S. Analytical distributions for stochastic gene expression. *PNAS*, **105**:17256–17261, 2008.

[26] WHITTLE, P. *Systems in Stochastic Equilibrium*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Inc., Chichester, 1986.

[27] XU, C., HANSEN, M.C. AND WIUF, C. Structural classification of continuous time Markov chains with applications. *Stochastics*, **94**:1003–1030, 2022.

[28] XU, C, HANSEN, M.C. AND WIUF, C. Full classification of dynamics for one-dimensional continuous-time Markov chains with polynomial transition rates. *Advances in Applied Probability*, **55**:321–355, 2023.

[29] XU, C, HANSEN, M.C. AND WIUF, C. The asymptotic tails of limit distributions of continuous-time Markov chains. *Advances in Applied Probability*, **56**:693–734, 2024.