

Quasi-stationary distributions for discrete-state models

Erik A. van Doorn^a and Philip K. Pollett^b

^a Department of Applied Mathematics, University of Twente

P.O. Box 217, 7500 AE Enschede, The Netherlands

E-mail: e.a.vandoorn@utwente.nl

^b Department of Mathematics, The University of Queensland

Qld 4072, Australia

E-mail: pkp@maths.uq.edu.au

13 January 2013

Abstract. This paper contains a survey of results related to quasi-stationary distributions, which arise in the setting of stochastic dynamical systems that eventually evanesce, and which may be useful in describing the long-term behaviour of such systems before evanescence. We are concerned mainly with continuous-time Markov chains over a finite or countably infinite state space, since these processes most often arise in applications, but will make reference to results for other processes where appropriate. Next to giving an historical account of the subject, we review the most important results on the existence and identification of quasi-stationary distributions for general Markov chains, and give special attention to birth-death processes and related models. The question of under what circumstances a quasi-stationary distribution, given its existence, is indeed a good descriptor of the long-term behaviour of a system before evanescence, is addressed as well. We conclude with a discussion of computational aspects, with more details given in a web appendix accompanying this paper.

Keywords: Applied probability, Markov processes

1 Introduction

Many biological systems are certain to “die out” eventually, yet appear to be stationary over any reasonable time scale. This phenomenon, termed *quasi stationarity*, is illustrated in Figure 1. Here, a model for the number $X(t)$ of occupied habitat patches in an n -patch metapopulation (population network) is simulated. For the parameter values given, the expected time to total extinction starting with one patch occupied is 4.7287×10^7 years, yet over the period of 250 years simulated, the number of patches occupied stabilises near 16. The notion of a *quasi-stationary distribution* has proved to be a useful tool in modelling this kind of behaviour. Figure 1 also shows the quasi-stationary distribution for this model. Notice that while the limiting distribution necessarily assigns all its probability mass to the extinction state 0, the quasi-stationary distribution assigns mass to states in a way that mirrors the quasi stationarity observed.

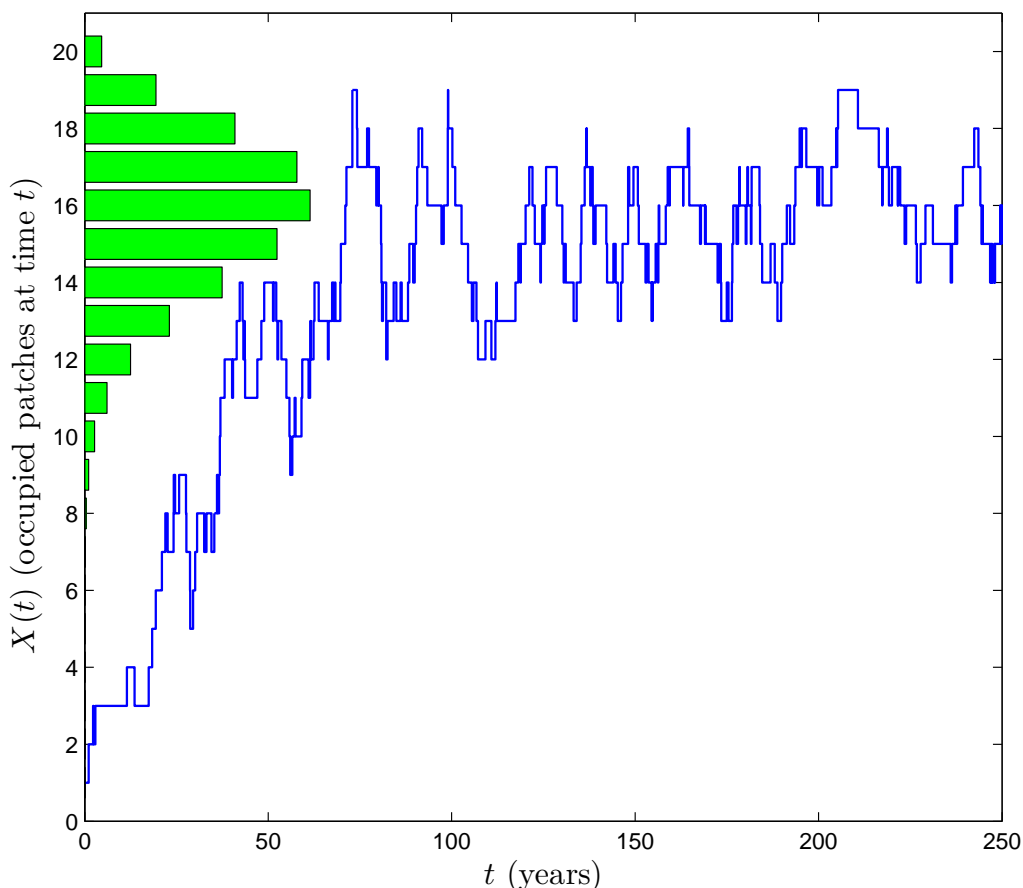


Fig 1. Simulation of a 20-patch metapopulation model with colonization rate $c = 0.1625$ and local extinction rate $e = 0.0325$, starting with one patch occupied. The quasi-stationary distribution is superimposed.

To make this notion more precise, think of an observer who at some time t is aware of the occupancy of some patches, yet cannot tell exactly which patches are occupied. What is the chance of there being precisely i patches occupied? If we were equipped with the full set of state probabilities $p_i(t) = \Pr(X(t) = i)$, $i \in \{0, 1, \dots, n\}$, we would evaluate the conditional probability $u_i(t) = \Pr(X(t) = i | X(t) \neq 0) = p_i(t)/(1 - p_0(t))$, for i in the set $S = \{1, \dots, n\}$ of transient states. Then, in view of the behaviour observed in our simulation, it would be natural for us to seek a distribution $\mathbf{u} = (u_i, i \in S)$ over S such that if $u_i(s) = u_i$ for a particular $s > 0$, then $u_i(t) = u_i$ for all $t > s$. Such a distribution is called a *stationary conditional distribution* or *quasi-stationary distribution*. Our key message is that \mathbf{u} can usually be determined from the transition rates of the process and that \mathbf{u} might then also be a *limiting conditional distribution* in that $u_i(t) \rightarrow u_i$ as $t \rightarrow \infty$, and thus be of use in modelling the long-term behaviour of the process. When the set S of transient states is finite, classical matrix theory is enough to establish the existence of a quasi-stationary distribution, which is unique when S is irreducible and admits a limiting conditional interpretation that is independent of initial conditions. When S is infinite the question even of the existence and then uniqueness of a quasi-stationary distribution is both subtle and interesting, and not yet fully resolved.

We shall be concerned here with continuous-time Markov chains over a finite or countably-infinite state space, since these processes most often arise in applications, but we will make reference to results for other processes where appropriate. We will review theoretical results on quasi-stationary distributions and limiting conditional distributions in Sections 3 and 4, giving special attention to birth-death processes and related models in Section 5. Computational aspects and some examples will be discussed in Section 6, with more detail on the numerical methods and guidance on how to implement these in MATLAB[®] given in the web appendix accompanying this paper. We start with some historical background in Section 2.

Our review is by no means exhaustive. For additional references on quasi-stationary distributions and related work, we refer the reader to the annotated bibliography maintained here:

<http://www.maths.uq.edu.au/~pkp/papers/qsds.html>

2 Modelling quasi stationarity

Yaglom [140] was the first to identify explicitly a limiting conditional distribution, establishing the existence of such for the subcritical Bienaymé-Galton-Watson branching process. However, the idea of a limiting conditional distribution goes back much further, at least to Wright [139, Page 111] in his

discussion of gene frequencies in finite populations:

“As time goes on, divergences in the frequencies of factors may be expected to increase more and more until at last some are either completely fixed or completely lost from the population. The distribution curve of gene frequencies should, however, approach a definite form if the genes which have been wholly fixed or lost are left out of consideration.”

The idea of “quasi stationarity” was crystalized by Bartlett [15, Page 38] and he later coined the term “quasi-stationary distribution” [16, Page 24]:

“It still may happen that the time to extinction is so long that it is still of more relevance to consider the effectivly ultimate distribution (called a ‘quasi-stationary’ distribution) of [the process] N .”

He went further giving details (in the context of an absorbing birth-death process) of one approach to modelling quasi stationarity whereby the process is imagined to be returned to state 1 (corresponding to one individual) at some small rate ϵ at the moment of extinction and the stationary distribution π^ϵ (if it exists) of the resulting “returned process” is used to model the long-term behaviour of the original process. Bartlett then argued that, under natural conditions, the dependence of π^ϵ on ϵ would be weak. Ewens [56, 57] exploited the idea for certain diffusion models that arise in population genetics, as well as their discrete-state analogues. Ewens returned his processes to an arbitrary state, and he coined the term *pseudo-transient* for the stationary distribution of the returned process. More generally, one might consider returning the process to the set S of transient states according to a probability measure \mathbf{m} over S , and then evaluating the stationary distribution $\pi^{\mathbf{m}}$ of the returned process. This was fleshed out by Darroch and Seneta [41, Section 2] in the context of discrete-time Markov chains, but they raised an objection that $\pi^{\mathbf{m}}$ depends on \mathbf{m} “to such an extent that it can be made into almost any distribution” over S by a suitable choice of \mathbf{m} . They described several other possibilities for a “quasi-stationary” distribution, many that (importantly) *do not* depend on the initial distribution, the most natural being those described in Section 1: the *stationary conditional* distribution (now usually termed *quasi-stationary distribution*) and the *limiting conditional distribution* (studied earlier by Mandl [93, 94]). These notions gained prominence in the literature and, following treatments for finite-state Markov chains by Darroch and Seneta [41, 42], there were significant early developments by Seneta and Vere-Jones [124, 135, 136] for countable-state chains that built on important work by Vere-Jones [134] on x -invariant measures and geometric ergodicity (but see also Albert [3] and Kingman [83]).

The notions of quasi-stationary distribution and pseudo-transient distribution can be reconciled, at least in our present context of countable-state Markov chains. Under mild assumptions (to be detailed later), the quasi-stationary distribution \mathbf{u} is unique and is a *fixed point* of the map $\mathbf{m} \mapsto \pi^{\mathbf{m}}$ (called the “return map”), that is, \mathbf{u} satisfies $\mathbf{u} = \pi^{\mathbf{u}}$ uniquely. Under these assumptions also, the return map is contractive, and so iteration leads us to \mathbf{u} (see, for example, Ferrari et al. [60]). Furthermore, $\pi^{\mathbf{m}}$ is expected to be “close” to \mathbf{u} for a range of return distributions \mathbf{m} , a statement that is made precise in Barbour and Pollett [14], thus assuaging to some extent the concerns of Darroch and Seneta. This has practical implications, for $\pi^{\mathbf{m}}$, interpreted as a “ratio of means” (its j -th component is the ratio of the expected time spent in state j and the expected time to extinction) [41, Section 3], can often be evaluated explicitly (see, for example, Artalejo and Lopez-Herrero [9]). Furthermore, since $\pi^{\mathbf{m}}$ is a *stationary distribution*, there is a range of efficient computational methods and accurate approximation (truncation) methods that can be exploited.

Bartlett [16, Page 25] mentioned a further, final, approach to modelling quasi stationarity whereby an approximating distribution of the state variable is identified, there (and typically) a Gaussian distribution, the quality of this approximation improving under some limiting regime (typically the size of the system increasing). The idea of using a Gaussian approximation for a Markovian state variable dates back at least to Kac [68, Page 384]. It was made concrete by Van Kampen [71] and, since then, “Van Kampen’s method” has become ubiquitous in biological and physical sciences literature. It was given a rigorous treatment as a diffusion approximation by Kurtz [86, 87] and Barbour [11, 12] (see also McNeil and Schach [95]) for the class of *density-dependent* Markovian models, the connection with quasi stationarity crystallized by Barbour [13]. Within this rigorous framework, one can not only identify the most appropriate approximating model, but delineate limiting conditions under which the approximation is faithful.

The nineteen sixties and seventies saw further developments in the theory of quasi-stationary distributions for countable-state Markov chains (for example Flaspohler [61], Tweedie [131]), as well as generalizations to semi-Markov processes (Arjas et al. [5], Cheong [30], Nummelin [104]) and Markov chains on a general state space (Tweedie [132]), and detailed results for generic models, for example, birth-death processes (Cavender [25], Good [63], Kesten [76]), random walks (Daley [38], Seneta [120]), queueing systems (Kyprianou [88]) and branching processes (Buiculescu [22], Green [65], Seneta and Vere-Jones [125]). Many of these early developments were influenced by ratio limit theory, which itself enjoyed significant growth during this period (see, for example, Papangelou [109]).

Our review is concerned with the most recent theoretical developments, within the context of continuous-time Markov chains and related generic models. For branching processes and related population processes there are excellent recent reviews by Lambert [89, Section 3] and Méléard and Villemonais [96]. For diffusions and other continuous-state processes, a good starting point is Steinsaltz and Evans [128], but see also Cattiaux et al. [24], Pinsky [111] and especially the recent book [35] by Collet et al., which aims at giving a unified approach towards quasi-stationarity for Markov chains, one-dimensional diffusions and dynamical systems. Whilst many issues remain unresolved, the theory has reached maturity, and the use of quasi-stationary distributions is now widespread, encompassing varied and contrasting areas of application, including cellular automata (Atman and Dickman [10]), complex systems (Collet et al. [36]), ecology (Day and Possingham [43], Gosselin [64], Gyllenberg and Sylvestrov [66], Kukhtin et al. [85], Pollett [114]) epidemics (Nåsell [100, 101, 102], Artalejo et al. [7, 8]), immunology (Stirk et al. [129]), medical decision making (Chan et al. [26]), physical chemistry (Dambrine and Moreau [39, 40], Oppenheim et al. [105], Pollett [113]), queues (Boucherie [19], Chen et al. [27], Kijima and Makimoto [81]), reliability (Kalpakam and Shahul-Hameed [69], Kalpakam [70], Li and Cao [90, 91], Pijnenburg et al. [110]), survival analysis (Aalen and Gjessing [1, 2], Steinsaltz and Evans [127]) and telecommunications (Evans [55], Ziedins [141]).

Concluding this section we observe that the concept of ageing is a common feature of many physical systems: the individual, or the system as a whole, moves irreversibly through a series of states before reaching a stable regime. For example, individual patients suffering a progressive disease move from lower to higher risk states, while an ecosystem, consisting of species that affect one another's ability to survive, will shed some species before a state of coexistence is reached. This phenomenon necessitates examining the theory of quasi-stationarity within the context of a *reducible* state space, an aspect that has hitherto received little attention in the literature, but will be discussed in this review in a finite setting.

3 Existence and identification

3.1 Introduction

This section contains an survey of results on the existence and identification of quasi-stationary and limiting conditional distributions. As announced we will focus on continuous-time Markov chains and consider finite and countably infinite state spaces separately in the Subsections 3.2 and 3.3, respectively. In the infinite setting we restrict ourselves to Markov chains for which the nonabsorbing

states constitute an irreducible class. We will briefly discuss quasi-stationarity for discrete-time Markov chains in Subsection 3.4. Some special cases of continuous-time Markov chains, to wit birth-death processes and birth-death processes with killing, are analysed in more detail in Section 5. For more information on Markov chains we refer the reader to, for example, Anderson [4] and Kijima [80].

3.2 Finite Markov chains

Preliminaries

We start off by introducing some notation and terminology, and deriving some basic results. Consider a continuous-time Markov chain $\mathcal{X} := \{X(t), t \geq 0\}$ on a state space $\{0\} \cup S$ consisting of an absorbing state 0 and a *finite* set of *transient* states $S := \{1, 2, \dots, n\}$. The generator of \mathcal{X} then takes the form

$$\begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{a}^T & Q \end{pmatrix}, \quad (1)$$

where $Q = (q_{ij})$ is the generator of the (sub)Markov chain on S and the vector $\mathbf{a} = (a_i, i \in S)$ of absorption (or *killing*) rates satisfies

$$\mathbf{a} = -\mathbf{1}Q^T \geq \mathbf{0}, \quad \mathbf{a} \neq \mathbf{0}. \quad (2)$$

Here $\mathbf{0}$ and $\mathbf{1}$ are row vectors of zeros and ones, respectively, superscript T denotes transposition, and \geq indicates componentwise inequality. Since all states in S are transient, state 0 is accessible from any state in S . Hence, whichever the initial state, the process will eventually escape from S into the absorbing state 0 with probability one.

We write \mathbb{P}_i for the probability measure of the process when the initial state is i , and \mathbb{E}_i for the expectation with respect to this measure. For any vector $\mathbf{u} = (u_i, i \in S)$ representing a distribution over S we let $a_{\mathbf{u}} := \sum_{i \in S} u_i a_i$ and $\mathbb{P}_{\mathbf{u}} := \sum_{i \in S} u_i \mathbb{P}_i$. We also write $P_{ij}(t) := \mathbb{P}_i(X(t) = j)$, and recall that the matrix $P(t) = (P_{ij}(t), i, j \in S)$ satisfies

$$P(t) = e^{Qt} := \sum_{k=0}^{\infty} \frac{Q^k}{k!} t^k, \quad t \geq 0 \quad (3)$$

(see, for example, [80, Section 4.6]).

We allow S to be reducible, so we suppose that S consists of communicating classes S_1, S_2, \dots, S_L , with $L \geq 1$, and let Q_k be the submatrix of Q corresponding to the states in S_k . We define a partial order on $\{S_1, S_2, \dots, S_L\}$ by writing $S_i \prec S_j$ when S_i is *accessible* from S_j , that is, when there exists a sequence of states k_0, k_1, \dots, k_ℓ , such that $k_0 \in S_j$, $k_\ell \in S_i$, and $q_{k_m k_{m+1}} > 0$ for every

$m \in \{0, 1, \dots, \ell - 1\}$. We assume in what follows that the states are labelled such that Q is in lower block-triangular form, so that we must have

$$S_i \prec S_j \implies i \leq j. \quad (4)$$

Considering that the matrices Q_k reside on the diagonal of Q , it is easily seen that the set of eigenvalues of Q is precisely the union of the sets of eigenvalues of the individual Q_k 's. It is well known (see, for example, Seneta [123, Theorem 2.6]) that the eigenvalue with maximal real part of Q_k , denoted by $-\alpha_k$, is unique, simple, and negative. Hence $\alpha := \min_k \alpha_k > 0$, and $-\alpha$ is the (possibly degenerate) eigenvalue of Q with maximal real part. The quantity α plays a crucial role in what follows and will be referred to as the *decay parameter* of \mathcal{X} .

A vector $\mathbf{u} = (u_i, i \in S)$ and, if appropriate, the probability distribution over S represented by \mathbf{u} , are called *x-invariant for Q* if

$$\sum_{i \in S} u_i q_{ij} = -x u_j, \quad j \in S, \quad (5)$$

that is, in the finite setting at hand, if \mathbf{u} is a left eigenvector of Q corresponding to the eigenvalue $-x$. The vector (or distribution) \mathbf{u} is called *x-invariant for P* if

$$\sum_{i \in S} u_i P_{ij}(t) = e^{-xt} u_j, \quad j \in S, \quad t \geq 0. \quad (6)$$

We recall that \mathbf{u} is a *quasi-stationary distribution* for \mathcal{X} if the distribution of $X(t)$, conditional on nonabsorption up to time t , is constant over time when \mathbf{u} is the initial distribution. That is, \mathbf{u} is a quasi-stationary distribution if, for all $t \geq 0$,

$$\mathbb{P}_{\mathbf{u}}(X(t) = j | T > t) = u_j, \quad j \in S, \quad (7)$$

where $T := \inf\{t \geq 0 : X(t) = 0\}$ is the *absorption time* (or *survival time*) of \mathcal{X} , the random variable representing the time at which absorption at 0 occurs. The notions of *x-invariant distribution* and *quasi-stationary distribution* are intimately related, as the next theorem shows. The theorem seems to be stated in its entirety only in a discrete-time setting (in [51]), so for the sake of completeness we also furnish a proof.

Theorem 1 Let $\mathbf{u} = (u_i, i \in S)$ represent a proper probability distribution over $S = \bigcup_{k=1}^L S_k$. Then the following statements are equivalent:

(a): \mathbf{u} is a quasi-stationary distribution for \mathcal{X} ,

- (b): \mathbf{u} is x -invariant for Q for some $x > 0$,
- (c): \mathbf{u} is x -invariant for P for some $x > 0$,
- (d): \mathbf{u} is α_k -invariant for Q for some $k \in \{1, 2, \dots, L\}$,
- (e): \mathbf{u} is α_k -invariant for P for some $k \in \{1, 2, \dots, L\}$.

Moreover, if \mathbf{u} is x -invariant for Q , then $x = a_{\mathbf{u}} := \sum_{i \in S} a_i u_i > 0$.

Proof The last claim is proven by summing (5) over $j \in S$ and noting that $x = 0$ would contradict the fact that all states in S are transient.

Since S is finite it follows readily from (3) that an x -invariant distribution for Q is also x -invariant for P . Conversely, taking derivatives in (6) and letting $t \rightarrow 0$ yields (5). So (5) and (6) are equivalent, and, as a consequence, (b) \iff (c) and (d) \iff (e). Moreover, a simple substitution shows (e) \implies (a). To prove (a) \implies (b), let \mathbf{u} be a quasi-stationary distribution. Then, evidently, $\mathbb{P}_{\mathbf{u}}(X(t) = j) = u_j \mathbb{P}_{\mathbf{u}}(T > t)$ for all $j \in S$ and $t \geq 0$, that is,

$$\sum_{i \in S} u_i P_{ij}(t) = u_j \left(1 - \sum_{i \in S} u_i P_{i0}(t) \right), \quad j \in S, t \geq 0.$$

Taking derivatives and letting $t \rightarrow 0$ subsequently shows that \mathbf{u} is $a_{\mathbf{u}}$ -invariant for Q . This establishes (b), since $a_{\mathbf{u}} > 0$ by the last claim.

Finally, we will show (b) \implies (d). So let $x > 0$ and assume that \mathbf{u} represents an x -invariant distribution for Q , that is, $\mathbf{u}Q = -x\mathbf{u}$. Recalling that Q is assumed to be in lower block-triangular form we decompose the vector $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_L)$ accordingly, and note that

$$\mathbf{u}_L Q_L = -x\mathbf{u}_L.$$

If $\mathbf{u}_L \neq \mathbf{0}$, then, by [123, Theorem 2.6] applied to the matrix Q_L , we have $x = \alpha_L$. On the other hand, if $\mathbf{u}_L = \mathbf{0}$ we must have

$$\mathbf{u}_{L-1} Q_{L-1} = -x\mathbf{u}_{L-1},$$

and we can repeat the argument. Thus proceeding we conclude that there must be a $k \in \{1, 2, \dots, L\}$ such that $x = \alpha_k$. This establishes (d) and completes the proof of the theorem. \square

Theorem 1 identifies all quasi-stationary distributions for \mathcal{X} . We call a distribution on S a *limiting conditional distribution* for \mathcal{X} if it is the limiting distribution as $t \rightarrow \infty$ of $X(t)$ conditional on survival

up to time t , that is,

$$\lim_{t \rightarrow \infty} \mathbb{P}_{\mathbf{w}}(X(t) = j | T > t), \quad j \in S, \quad (8)$$

for some initial distribution $\mathbf{w} = (w_i, i \in S)$. Vere-Jones [136, Theorem 2] has shown (in a more general setting) that if the limits (8) constitute a proper distribution, then this distribution must be a quasi-stationary distribution. Conversely, any quasi-stationary distribution is of course a limiting conditional distribution, so Theorem 1 also identifies all limiting conditional distributions. Evidently, what remains to be solved is the problem of identifying the quasi-stationary distribution (if any) that is the limiting conditional contribution for any given initial distribution.

Noting that $\mathbb{P}_{\mathbf{u}}(T > t) = \mathbb{P}_{\mathbf{u}}(X(t) \in S) = \sum_{j \in S} \sum_{i \in S} u_i P_{ij}(t)$, Theorem 1 immediately yields the following.

Corollary 2 If $\mathbf{u} = (u_i, i \in S)$ is a quasi-stationary distribution for \mathcal{X} over $S = \bigcup_{k=1}^L S_k$, then

$$\mathbb{P}_{\mathbf{u}}(T > t) = \exp(-a_{\mathbf{u}}t), \quad t \geq 0,$$

and $a_{\mathbf{u}} = \alpha_k$ for some $k \in \{1, 2, \dots, L\}$.

So the residual survival time conditional on survival up to some time t is exponentially distributed if the initial distribution is a quasi-stationary distribution. In what follows we are also interested in

$$\lim_{t \rightarrow \infty} \mathbb{P}_{\mathbf{w}}(T > t + s | T > t), \quad s \geq 0, \quad (9)$$

that is, in the limiting distribution as $t \rightarrow \infty$ of the residual survival time conditional on survival up to time t , for *any* initial distribution $\mathbf{w} = (w_i, i \in S)$.

Irreducible state space

Let us first assume that S is irreducible, that is, $L = 1$, and so $S_1 = S$ and $Q_1 = Q$. Hence $-\alpha$, the eigenvalue of Q with maximal real part, is unique, simple, and negative. It is well known (see, for example, [123, Theorem 2.6]) that the associated left and right eigenvectors $\mathbf{u} = (u_i, i \in S)$ and $\mathbf{v}^T = (v_i, i \in S)^T$ can be chosen strictly positive componentwise, and hence such that

$$\mathbf{u}\mathbf{1}^T = 1 \quad \text{and} \quad \mathbf{u}\mathbf{v}^T = 1. \quad (10)$$

Classical Markov-chain theory (see, for example, [80, Theorem A.7]) then tells us that the transition probabilities $P_{ij}(t)$ satisfy

$$e^{\alpha t} P_{ij}(t) = v_i u_j + o(e^{-\epsilon t}) \quad \text{as } t \rightarrow \infty, \quad i, j \in S, \quad (11)$$

for some $\varepsilon > 0$, which explains why α is called the decay parameter of \mathcal{X} .

Our definition of \mathbf{u} implies that the distribution represented by \mathbf{u} is α -invariant for Q , whence, by Theorem 1, \mathbf{u} is the unique quasi-stationary distribution for \mathcal{X} . So if we let \mathbf{u} be the initial distribution, then, conditional on survival up to time t , the distribution of $X(t)$ is constant over t , and, by Corollary 2, the remaining survival time has an exponential distribution with parameter α . Darroch and Seneta [42] have shown that similar results hold true in the limit as $t \rightarrow \infty$ when the initial distribution differs from \mathbf{u} . Namely, for any initial distribution \mathbf{w} one has

$$\lim_{t \rightarrow \infty} \mathbb{P}_{\mathbf{w}}(X(t) = j | T > t) = u_j, \quad j \in S. \quad (12)$$

and

$$\lim_{t \rightarrow \infty} \mathbb{P}_{\mathbf{w}}(T > t + s | T > t) = e^{-\alpha s}, \quad s \geq 0. \quad (13)$$

We summarize these results in a theorem.

Theorem 3 [42, Section 3] When all states in S communicate the Markov chain \mathcal{X} with decay parameter α has a unique quasi-stationary distribution $\mathbf{u} = (u_i, i \in S)$, which is the (unique, positive) solution of the system $\mathbf{u}Q = -\alpha\mathbf{u}$ and $\mathbf{u}\mathbf{1}^T = 1$. Moreover, for any initial distribution $\mathbf{w} = (w_i, i \in S)$ the limits (8) and (9) exist, and are given by (12) and (13), respectively, where $\mathbf{u} = (u_i, i \in S)$ is the quasi-stationary distribution of \mathcal{X} .

General state space

The situation is more complicated when $L \geq 1$, and we must introduce some more notation and terminology before we can state the results. We recall that $-\alpha_k < 0$ is the unique and simple eigenvalue of Q_k with maximal real part, and that $-\alpha = -\min_k \alpha_k < 0$ is the eigenvalue of Q with maximal real part. We let $I_\alpha := \{k : \alpha_k = \alpha\}$, so that $\text{card}(I_\alpha)$ is the algebraic multiplicity of the eigenvalue $-\alpha$, and write $S_{\min} := S_{\min I_\alpha}$. A class S_k will be called a *minimal class for α* if it is a minimal element in the set $\{S_j, j \in I_\alpha\}$ with respect to the partial order \prec , that is, for all $j \neq k$,

$$S_j \prec S_k \implies j \notin I_\alpha.$$

Letting m_α be the number of minimal classes for α , we have $m_\alpha \geq 1$, since S_{\min} is always a minimal class for α . Moreover, it is shown in [50, Section 6] that there are precisely m_α linearly independent, *nonnegative* vectors \mathbf{u} satisfying $\mathbf{u}Q = -\alpha\mathbf{u}$. (Hence, with g_α denoting the geometric multiplicity of the eigenvalue $-\alpha$, we have $m_\alpha \leq g_\alpha \leq \text{card}(I_\alpha)$.) It is also shown in [50, Section 6] that if

$\mathbf{u} = (u_i, i \in S)$ is a quasi-stationary distribution from which S_{\min} is accessible (by which we mean that there is a state i such that $u_i > 0$ and S_k is accessible from i), then \mathbf{u} must satisfy $\mathbf{u}Q = -\alpha\mathbf{u}$, and $u_i > 0$ if and only if state i is accessible from S_{\min} . So, in view of Theorem 1, we can generalize the first part of Theorem 3 as follows.

Theorem 4 [50, Theorem 10] The Markov chain \mathcal{X} with decay parameter α has a unique quasi-stationary distribution \mathbf{u} from which S_{\min} is accessible if and only if S_{\min} is the only minimal class for α , in which case \mathbf{u} is the (unique) nonnegative solution to the system $\mathbf{u}Q = -\alpha\mathbf{u}$ and $\mathbf{u}\mathbf{1}^T = 1$, and has a positive i th component if and only if state i is accessible from S_{\min} .

As shown in [50, Section 6] the second part of Theorem 3 can be generalized in the same spirit, leading to the next theorem.

Theorem 5 [50, Theorem 11] If the Markov chain \mathcal{X} with decay parameter α has an initial distribution \mathbf{w} such that S_{\min} is accessible, and if S_{\min} is the only minimal class for α , then the limits (8) and (9) exist and are given by (12) and (13), respectively, where \mathbf{u} is the unique quasi-stationary distribution of \mathcal{X} from which S_{\min} is accessible.

We illustrate the preceding results by an example. Suppose that the generator of \mathcal{X} is given by

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 2 & -2 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 1 - \gamma & \gamma & -1 & 0 \\ 0 & 0 & 1 & 1 & -2 \end{pmatrix}, \quad (14)$$

where $0 \leq \gamma \leq 1$. Evidently, we have $S_i = \{i\}$ for $i = 1, 2, 3$ and 4. Moreover, $\alpha = 1$ and $I_\alpha = \{2, 3\}$.

If $0 < \gamma \leq 1$, then S_2 is the only minimal class, so $m_\alpha = 1$. Hence, by Theorem 4 there is a unique quasi-stationary distribution from which state 2 is accessible, which is readily seen to be $(\frac{1}{2}, \frac{1}{2}, 0, 0)$. Observe that the restriction is relevant, since $(1, 0, 0, 0)$ is a quasi-stationary distribution too, but not one from which state 2 is accessible. Theorem 5 tells us that $(\frac{1}{2}, \frac{1}{2}, 0, 0)$ is the limiting conditional distribution for any initial distribution from which state 2 is accessible.

If $\gamma = 0$, then both S_2 and S_3 are minimal classes, so $m_\alpha = 2$. In this case there is no unique quasi-stationary distribution from which state 2 is accessible. In fact, it is easy to see that there are infinitely many such distributions. Also the limiting conditional distribution is not unique in this case, but depends on the initial distribution.

3.3 Infinite, irreducible Markov chains

The setting of this subsection is again that of a continuous-time Markov chain $\mathcal{X} := \{X(t), t \geq 0\}$ on a state space $\{0\} \cup S$ consisting of an absorbing state 0 and a set of transient states S . But now S is supposed to be countably infinite, so we set $S = \{1, 2, \dots\}$. We restrict ourselves to the case in which all states in S communicate. Note that absorption at 0 is not necessarily certain; we set $T := \inf\{t \geq 0 : X(t) = 0\} = \infty$ if absorption at 0 does not occur, so that $\mathbb{P}(T > t) = 1 - \mathbb{P}(T \leq t)$. As before our aim is to identify quasi-stationary and limiting conditional distributions.

Preliminaries

We use the notation of the previous subsection insofar as it extends to the infinite setting at hand, so we let $Q = (q_{ij})$ be the generator of the (sub)Markov chain on S and $\mathbf{a} = (a_i, i \in S)$ the vector of absorption (or killing) rates. We will assume that Q is stable and conservative, that is,

$$-q_{ii} = a_i + \sum_{j \in S, j \neq i} q_{ij} < \infty, \quad i \in S,$$

and, in addition, that \mathcal{X} is nonexplosive and hence uniquely determined by Q . Nonexplosiveness also implies that the events $\{T > t\}$ and $\{X(t) \in S\}$ are identical.

The result (11) cannot be extended in full to the setting at hand. However, Kingman [83] has shown (see also [4, Theorem 5.1.9]) that under our assumptions there exist strictly positive constants c_{ij} (with $c_{ii} = 1$) and a parameter $\alpha \geq 0$ such that

$$P_{ij}(t) \leq c_{ij} e^{-\alpha t}, \quad t \geq 0, \quad i, j \in S, \tag{15}$$

and

$$\alpha = - \lim_{t \rightarrow \infty} \frac{1}{t} \log P_{ij}(t), \quad i, j \in S. \tag{16}$$

Again we will refer to α as the *decay parameter* of \mathcal{X} . It is not difficult to see that α is the rate of convergence of the transition probabilities $P_{ij}(t)$ in the sense that

$$\alpha = \inf \left\{ a \geq 0 : \int_0^\infty e^{at} P_{ij}(t) dt = \infty \right\}, \quad i, j \in S. \tag{17}$$

The definitions given in the previous section for x -invariant and quasi-stationary distributions remain valid, but the relationships between these notions given in Theorem 1 allow only a partial extension. Namely, on combining results in Vere-Jones [136] and Pollett and Vere-Jones [117] we obtain the following.

Theorem 6 Let $\mathbf{u} = (u_i, i \in S)$ represent a proper probability distribution over S . Between the statements (a), (b) and (c) defined in Theorem 1 the following relationships exist:

(i) (a) \iff (c),

(ii) (c) \iff (b) and $x = a_{\mathbf{u}}$.

Moreover,

(iii) (c) $\implies 0 < x \leq \alpha$.

More precisely, statement (i) in the theorem is implied by [136, Theorem 2] (see also Nair and Pollett [99, Proposition 3.1]), and statement (ii) combines [136, Theorem 5] and [117, Corollary 1]. Finally, [136, Theorem 4] and [117, Theorem 2] together yield statement (iii).

Since \mathcal{X} is nonexplosive we have $\mathbb{P}_{\mathbf{u}}(T > t) = \mathbb{P}_{\mathbf{u}}(X(t) \in S)$. So, by summing (6) over all $j \in S$ and applying Theorem 6, we obtain the following corollary (see also Artalejo [6]).

Corollary 7 If $\mathbf{u} = (u_i, i \in S)$ is a quasi-stationary distribution over S for the Markov chain \mathcal{X} with decay parameter α , then

$$\mathbb{P}_{\mathbf{u}}(T > t) = \exp(-a_{\mathbf{u}}t), \quad t \geq 0, \tag{18}$$

and $0 < a_{\mathbf{u}} \leq \alpha$.

It is enlightening to point out some differences between Theorem 6 and Theorem 1 with $L = 1$, the corresponding result in a finite setting.

First note that in a finite setting an x -invariant distribution \mathbf{u} (for Q or for P) must satisfy $x = a_{\mathbf{u}} = \alpha$. In the infinite setting of Theorem 6 we have $0 < x = a_{\mathbf{u}} \leq \alpha$ if \mathbf{u} is x -invariant for P , and we can even have $x < a_{\mathbf{u}}$ if \mathbf{u} is only x -invariant for Q . Note that summing (5) over all $j \in S$ would yield $x = a_{\mathbf{u}}$ if the interchange of summation would be justified, but this is *not* the case in general since $\sum_{i \in S} u_i |q_{ii}|$ may be infinite. In Section 5 we will encounter Markov chains exhibiting the various possibilities, including chains having an x -invariant distribution for P (and hence a quasi-stationary distribution) for all x in the interval $(0, \alpha]$.

Secondly, observe that $\alpha > 0$ in a finite setting, but we may have $\alpha = 0$ in an infinite setting, indicating that the transition probabilities $P_{ij}(t)$, $i, j \in S$, tend to zero at a rate that is slower than exponential. If $\alpha > 0$ the Markov chain \mathcal{X} is called *exponentially transient*. Statements (i) and (iii) of Theorem 6 imply that exponential transience is necessary for the existence of a quasi-stationary distribution.

Thirdly, if \mathbf{u} is a quasi-stationary distribution, then, for all $j \in S$,

$$u_j \mathbb{P}_{\mathbf{u}}(T > t) = \mathbb{P}_{\mathbf{u}}(X(t) = j | T > t) \mathbb{P}_{\mathbf{u}}(T > t) = \mathbb{P}_{\mathbf{u}}(X(t) = j). \quad (19)$$

Since, under our assumptions, the right-hand side of (19) tends to zero as $t \rightarrow \infty$ (for any initial distribution \mathbf{u}), we must have $\mathbb{P}_{\mathbf{u}}(T > t) \rightarrow 0$ as $t \rightarrow \infty$, that is, absorption is certain. This is always true in a finite but not necessarily in an infinite setting, because there may be a drift to infinity. Moreover, if absorption is certain, then, for \mathbf{u} to be a quasi-stationary distribution, it is also necessary for $\mathbb{P}_{\mathbf{u}}(X(t) = j)$ and $\mathbb{P}_{\mathbf{u}}(T > t)$ to have the same rate of convergence. Again, this is true, for any initial distribution \mathbf{u} , in a finite, but not necessarily in an infinite setting.

The preceding observation motivates us to define α_0 as the rate of convergence to zero of $\mathbb{P}_i(T > t) = 1 - P_{i0}(t)$, that is,

$$\alpha_0 = \inf \left\{ a \geq 0 : \int_0^\infty e^{at} \mathbb{P}_i(T > t) dt = \infty \right\}, \quad i \in S, \quad (20)$$

or, equivalently,

$$\alpha_0 = \inf \{ a \geq 0 : \mathbb{E}_i(e^{aT}) = \infty \}, \quad i \in S. \quad (21)$$

It follows easily from the irreducibility of S that α_0 is independent of i . Moreover, since $P_{ii}(t) \leq \mathbb{P}_i(T > t) \leq 1$, we have

$$0 \leq \alpha_0 \leq \alpha, \quad (22)$$

where each inequality can be strict (Jacka and Roberts [67, Remark 3.1.4]). It will be useful to note that

$$S_{\text{abs}} \text{ is finite and absorption certain} \implies \alpha_0 = \alpha, \quad (23)$$

where $S_{\text{abs}} := \{i \in S : a_i > 0\}$ (see [67, Theorem 3.3.2 (iii)]).

Necessary conditions for existence

We have now gathered sufficient information to state some necessary conditions for the existence of a quasi-stationary distribution for \mathcal{X} .

Theorem 8 If there exists a quasi-stationary distribution \mathbf{u} for the absorbing Markov chain \mathcal{X} , then absorption is certain and $0 < a_{\mathbf{u}} \leq \alpha_0 \leq \alpha$.

Proof We concluded already from (19) that absorption must be certain if there exists a quasi-stationary distribution \mathbf{u} for \mathcal{X} . Moreover, by Corollary 7, $a_{\mathbf{u}} > 0$ and

$$\exp(-a_{\mathbf{u}}t) = \mathbb{P}_{\mathbf{u}}(T > t) \geq u_i \mathbb{P}_i(T > t), \quad i \in S, t \geq 0,$$

whence $a_{\mathbf{u}} \leq \alpha_0$. The theorem follows in view of (22). \square

Thus the condition $\alpha_0 > 0$ is stronger than exponential transience, and necessary for a quasi-stationary distribution to exist. Following [60, Page 515] we call a quasi-stationary \mathbf{u} *minimal* if $a_{\mathbf{u}} = \alpha_0$. Clearly, the Theorems 6 and 8 lead to another sufficient condition for $\alpha_0 = \alpha$ besides (23).

Corollary 9 If \mathbf{u} is an α -invariant quasi-stationary distribution for \mathcal{X} , then $\alpha_0 = \alpha$ and \mathbf{u} is a minimal quasi-stationary distribution.

In all examples of quasi-stationary distributions known to us equality of α_0 and α prevails, and we suspect that equality is necessary for the existence of a quasi-stationary distribution. (The Markov chain of [67, Remark 3.1.4] has $0 < \alpha_0 < \alpha$ but, in view of our Theorem 20, does not have a quasi-stationary distribution.) In fact, we conjecture the validity of a stronger statement, which can be verified in some specific settings (see the Theorems 11, 12, 16 and 20).

Conjecture 10 The existence of a quasi-stationary distribution for the Markov chain \mathcal{X} with decay parameter α implies the existence of an α -invariant quasi-stationary distribution.

Sufficient conditions for existence

The problem of finding a necessary *and* sufficient condition, stated directly in terms of Q , for the existence of a quasi-stationary distribution of the Markov chain \mathcal{X} is intriguing, but remains unsolved. We continue this subsection with a survey of sufficient conditions. In some cases the Markov chain \mathcal{X} is required to be *uniformizable*, which means that the matrix $Q = (q_{ij})$ of transition rates satisfies

$$-q_{ii} = a_i + \sum_{j \in S, j \neq i} q_{ij} \leq C,$$

for some constant C and all $i \in S$.

For *birth-death processes* it is known that exponential transience is necessary and sufficient for the existence of a quasi-stationary distribution when absorption at 0 is certain. Moreover, if the birth and death rates satisfy a certain condition that is weaker than uniformizability, then, for any number x in the interval $0 < x \leq \alpha$, there is a unique quasi-stationary distribution \mathbf{u} such that $a_{\mathbf{u}} = x$. We

postpone giving the details of these results to Subsection 5.1, but note at this point that Kijima [79, Theorem 3.3] has partly generalized these results to Markov chains that are uniformizable and *skip-free to the left*, that is, $q_{ij} = 0$ if $j < i - 1$. Namely, if $\alpha > 0$ and, for some x in the interval $0 < x \leq \alpha$, there is an x -invariant distribution for Q , then, for each y in the interval $x \leq y \leq \alpha$, there is a unique quasi-stationary distribution \mathbf{u} such that $a_{\mathbf{u}} = y$. Of course, (23) implies that $\alpha_0 = \alpha$ in this case.

More concrete existence results are available when the Markov chain \mathcal{X} satisfies *asymptotic remoteness*, that is,

$$\lim_{i \rightarrow \infty} \mathbb{P}_i(T \leq t) = 0 \quad \text{for all } t > 0, \tag{24}$$

or, in the terminology of Li and Li [92], the absorbing state 0 is a *Feller state*. (By [92, Proposition 2.1] it then follows in the setting at hand that \mathcal{X} is a *Feller process*, cf. [4, Section 1.5].) If asymptotic remoteness prevails then, by [60, Theorem 1.1], $\alpha_0 > 0$ is necessary and sufficient for the existence of a quasi-stationary distribution when absorption at 0 is certain. Moreover, [60, Theorem 4.1 and Proposition 5.1(a)] tell us that there exists an α_0 -invariant quasi-stationary distribution if \mathcal{X} is also uniformizable, while, by [60, Corollary 5.3], we must have $\alpha_0 = \alpha$ in that case. Summarizing we can state the following.

Theorem 11 [60] Let the absorbing Markov chain \mathcal{X} with decay parameter α be such that absorption is certain and $\alpha_0 > 0$. If asymptotic remoteness prevails, then there exists a quasi-stationary distribution for \mathcal{X} . If, in addition, \mathcal{X} is uniformizable, then $\alpha_0 = \alpha$ and there exists an α -invariant quasi-stationary distribution.

Since asymptotic remoteness prevails if the Markov chain \mathcal{X} is uniformizable and skip-free to the left, Theorem 11 guarantees the existence of an α -invariant quasi-stationary distribution in the setting of Kijima's paper [79].

Given $\alpha_0 > 0$ and certain absorption, asymptotic remoteness is sufficient but not necessary for the existence of a quasi-stationary distribution. Examples of Markov chains having a quasi-stationary distribution while (24) is violated are given by Pakes [108] and Bobrowski [18], but can also be found in the setting of birth-death processes (see Section 5).

Another approach towards obtaining sufficient conditions for the existence of a quasi-stationary distribution is to confine attention to α -recurrent Markov chains, which are Markov chains satisfying

$$\int_0^\infty e^{\alpha t} P_{ii}(t) dt = \infty \tag{25}$$

for some state $i \in S$ (and then for all states $i \in S$). A chain is called α -transient if it is not α -recurrent. For later reference we also note at this point that an α -recurrent process is said to be α -positive if for some state $i \in S$ (and then for all states $i \in S$)

$$\lim_{t \rightarrow \infty} e^{\alpha t} P_{ii}(t) > 0, \tag{26}$$

and α -null otherwise. (Note that a finite absorbing Markov chain is always α -positive, in view of (11).) It is shown in [83, Theorem 4] that if \mathcal{X} is α -recurrent, then there exists, up to normalization, a unique positive solution to the system (5) with $x = \alpha$. However, besides $\alpha_0 > 0$ and certain absorption, additional restrictions on Q are required to ensure summability, and hence the existence of a (unique) α -invariant quasi-stationary distribution, even if \mathcal{X} is α -positive (cf. [124, Page 414]). One such condition is given in the next theorem, which is inspired by the observations on Page 414 of [124] in a discrete-time setting (see also [67, Lemma 3.3.5]).

Theorem 12 Let the absorbing Markov chain \mathcal{X} with decay parameter α be α -recurrent and such that S_{abs} is finite, absorption is certain and $\alpha > 0$. Then $\alpha_0 = \alpha$ and there exists a unique α -invariant quasi-stationary distribution.

Proof From (23) we know already that $\alpha_0 = \alpha$. Let \mathbf{u} be the unique solution (up to a multiplicative constant) to the system (5) with $x = \alpha$. Then, by [112, Theorem 1 (iii)], \mathbf{u} also solves (6) with $x = \alpha$. Summation over $j \in S_{\text{abs}}$ and integration yields

$$\sum_{i \in S} u_i \int_0^\infty \left(\sum_{j \in S_{\text{abs}}} P_{ij}(t) \right) dt = \alpha^{-1} \sum_{j \in S_{\text{abs}}} u_j,$$

since S_{abs} is finite and $\alpha > 0$. The integral represents the expected sojourn time in S_{abs} when the initial state is i , so is finite and bounded below by $(\max_{j \in S_{\text{abs}}} |q_{jj}|)^{-1} > 0$. It follows that \mathbf{u} must be summable, and hence can be normalized to be a distribution, which, by Theorem 6, must then be a quasi-stationary distribution. \square

Since α -recurrence is usually difficult to verify, one might attempt to replace it by a condition which is stated directly in terms of Q . A result of this type is the continuous-time counterpart of Kesten's result [77, Theorem 2 and Page 657]. Insofar as it concerns quasi-stationary distributions, this result states that – besides finiteness of S_{abs} and certain absorption – uniformizability, certain restrictions on the number of nonzero downward rates, and a type of uniform irreducibility condition are sufficient for the existence of a unique α -invariant quasi-stationary distribution.

Limiting conditional distributions

As observed already in a finite setting, any quasi-stationary distribution \mathbf{u} for \mathcal{X} is a *limiting conditional distribution*, in the sense that by a suitable choice of \mathbf{w} (namely $\mathbf{w} = \mathbf{u}$) the limits (12) constitute a distribution represented by \mathbf{u} . Conversely, by [136, Theorem 2] (see also [108, Lemma 2.1]), any limiting conditional distribution must be a quasi-stationary distribution. In fact, restricting ourselves to initial distributions that are concentrated on a single state, we can state the following.

Theorem 13 Let the absorbing Markov chain \mathcal{X} with decay parameter α be such that, for some $i \in S$, the limits

$$u_j = \lim_{t \rightarrow \infty} \mathbb{P}_i(X(t) = j | T > t), \quad j \in S, \quad (27)$$

exist and constitute a proper distribution. Then absorption is certain, $\alpha = \alpha_0 > 0$, and $\mathbf{u} = (u_j, j \in S)$ is an α -invariant quasi-stationary distribution.

Proof If the limits (27) constitute a proper distribution, then, as noted above, \mathbf{u} is a quasi-stationary distribution, so that absorption must be certain and $a_{\mathbf{u}} > 0$. Moreover, we have $\alpha = \alpha_0$ by [67, Theorem 3.2.2 (ii)]. Finally, by [67, Lemma 4.1], \mathbf{u} is an α -invariant distribution, so that $\alpha = a_{\mathbf{u}} > 0$. \square

Remarks (i) Proposition 5.1(b) in [60] is similar to our Theorem 13, but contains the unnecessary requirement that \mathcal{X} be uniformizable.

(ii) If the existence of a quasi-stationary distribution for \mathcal{X} would imply the existence of the limits (27) as a proper distribution, then Theorem 13 would imply the validity of Conjecture 10.

The existence of the limits in (27) has been proven in various settings, usually more restrictive, however, than those required for the existence of a quasi-stationary distribution (see, for example, Seneta [120], Cheong [31], and Karlin and Tavaré [74]).

The last theorem of this section is a partial converse to Theorem 13 and gives a sufficient condition for the existence of the limits (8) and (9) when the initial distribution is concentrated on a single state. The theorem is implied by [136, Theorems 2 and 3] (the first part) and [67, Lemma 4.1] (the second part). Its discrete-time counterpart is contained in [124, Theorem 3.1].

Theorem 14 Let the absorbing Markov chain \mathcal{X} with decay parameter α be such that absorption is certain and $\alpha > 0$. If \mathcal{X} is α -positive and there exists a (unique) α -invariant quasi-stationary distribution $\mathbf{u} = (u_j, j \in S)$, then, for all $i \in S$,

$$\lim_{t \rightarrow \infty} \mathbb{P}_i(X(t) = j | T > t) = u_j, \quad j \in S, \quad (28)$$

and

$$\lim_{t \rightarrow \infty} \mathbb{P}_i(T > t + s | T > t) = e^{-\alpha s}, \quad s \geq 0. \quad (29)$$

Evidently, the statement of the theorem may be generalized to initial distributions with finite support.

We conclude this section with the observation that in the preceding theorems the condition of certain absorption can be relaxed, provided we work with the process “restricted to the event of absorption”. Indeed, this allows us to deal with

$$\lim_{t \rightarrow \infty} \mathbb{P}_i(X(t) = j | t < T < \infty), \quad j \in S,$$

via the dual chain $\tilde{p}_{ij}(t) = p_{ij}(t)e_j/e_i$, $i, j \in S$, where $e_i = \mathbb{P}_i(T < \infty)$ is the probability of absorption starting in state i (see Waugh [137]). Our irreducibility assumption ensures that $e_i > 0$ for all i .

3.4 Discrete-time Markov chains

Most results for continuous-time Markov chains have more or less obvious analogues for discrete-time Markov chains. In one respect the discrete-time setting is simpler, since the requirement of uniformizability that we have encountered in several results of the previous section has no bearing on discrete-time Markov chains. On the other hand, the phenomenon of periodicity may pose problems in a discrete-time setting, in particular when considering limiting conditional distributions. In this subsection we will briefly describe how discrete-time results may be obtained from continuous-time results, and give appropriate references for further details.

So let $\mathcal{Y} := \{Y_n, n = 0, 1, \dots\}$ be a discrete-time Markov chain taking values in a state space consisting of an absorbing state 0 and a finite or countably infinite set of *transient* states S . We denote the matrix of 1-step transition probabilities *within* S by $P := (p_{ij}, i, j \in S)$, and let p_{i0} , $i \in S$, be the 1-step absorption probabilities. The n -step transition probabilities are denoted by $P_{ij}(n)$, and the matrix of n -step transition probabilities *within* S by $P(n) := (P_{ij}(n), i, j \in S)$, so that $P_{ij}(1) = p_{ij}$, and $P(n) = P^n$. A vector $\mathbf{u} = (u_j, j \in S)$, or a proper probability distribution over S represented by \mathbf{u} , are called *x-invariant* for P if

$$\sum_{i \in S} u_i p_{ij} = x u_j, \quad j \in S. \quad (30)$$

By analogy with (7) and (8) the distribution \mathbf{u} is said to be a *quasi-stationary distribution* for \mathcal{Y} if, for all $n = 0, 1, \dots$,

$$\mathbb{P}_{\mathbf{u}}(Y(n) = j | T > n) = u_j, \quad j \in S, \quad (31)$$

and a *limiting conditional distribution* for \mathcal{Y} if, for some initial distribution \mathbf{w} ,

$$\lim_{n \rightarrow \infty} \mathbb{P}_{\mathbf{w}}(Y(n) = j \mid T > n) = u_j, \quad j \in S, \quad (32)$$

where $T := \inf\{n \geq 0 : Y(n) = 0\}$ is the *absorption time* and $\mathbb{P}_{\mathbf{w}}$ the probability measure of the process when the initial distribution is \mathbf{w} .

The case of a finite, irreducible and aperiodic state space S was analysed in the classic paper of Darroch and Seneta [41]. Their results have recently been generalized in [51] to a reducible setting, leading to discrete-time analogues of the Theorems 1, 4 and 5, with the restriction that the analogue of the latter requires an additional aperiodicity condition. We refer to [51] for details.

If S is countably infinite, irreducible and aperiodic there exists a real number ρ (the *decay parameter* of the Markov chain \mathcal{Y} in S) such that $0 < \rho \leq 1$ and, for each $i, j \in S$,

$$(P_{ij}(n))^{1/n} \rightarrow \rho, \quad (33)$$

as $n \rightarrow \infty$ [134]. (Note that if S is periodic the same is true provided the limit is taken through values of n for which $P_{ij}(n)$ is strictly positive [83].) The chain is said to be *geometrically transient* if $\rho < 1$. A link with the results of Subsection 3.3 is established by observing that for any $q > 0$ we can associate with \mathcal{Y} a continuous-time, uniformizable Markov chain \mathcal{X}_q on $\{0\} \cup S$ by letting $Q = (q_{ij})$ and $\mathbf{a} = (a_i, i \in S)$ such that

$$q_{ij} = q(p_{ij} - \delta_{ij}), \quad a_i = qp_{i0}, \quad i, j \in S,$$

where δ_{ij} is Kronecker's delta. Namely, the decay parameter α_q of \mathcal{X}_q is easily seen to satisfy $\alpha_q = q(1 - \rho)$. Moreover, a vector $\mathbf{u} = (u_i, i \in S)$ is x -invariant for Q if and only if \mathbf{u} is $(1 - x/q)$ -invariant for P , and \mathbf{u} is a quasi-stationary distribution for \mathcal{X}_q if and only if it is a quasi-stationary distribution for \mathcal{Y} .

These observations enable us to translate all results for continuous-time, uniformizable Markov chains in Subsection 3.3 to the discrete-time setting at hand, with the restriction that we have to impose aperiodicity of S in statements involving limiting conditional distributions. For details we refer to Coolen-Schrijner and van Doorn [37].

We finally note that the existence of the limits (32) as a bona fide distribution – and hence the existence of a quasi-stationary distribution – has been proven in various settings, usually more restrictive, however, than those required by results such as the discrete-time counterparts of the Theorems 11 and 12 (see, for example, Seneta and Vere-Jones [124], Daley [38], Pakes [107], Kijima [78], Kesten [77], van Doorn and Schrijner [52, 53], Ferrari et al. [59] and Moler et al. [98]).

4 Speed of convergence to quasi stationarity

In the previous section we have focused on the existence of quasi-stationary distributions, given the parameters that determine the dynamics of the process, and on the identification of a quasi-stationary distribution as the limiting conditional distribution for a given initial distribution. However, as noted in the first paragraph of Section 2, the existence of a limiting conditional distribution does not necessarily imply that the process exhibits the type of “quasi stationary” behaviour depicted in Figure 1, characterized by relatively fast convergence to the limiting conditional distribution, and eventual evanescence after a much longer time. In the present section we will focus on the circumstances under which such a “quasi stationary” scenario prevails, given the existence of a limiting conditional distribution.

As before our setting will be that of a continuous-time Markov chain \mathcal{X} on a finite or countably infinite state space $\{0\} \cup S$. For convenience we will assume that Q , the generator of the (sub)Markov chain on S , is irreducible, and that, for some initial state $i \in S$, the limits (27) exist and constitute a proper distribution (and hence a quasi-stationary distribution) $\mathbf{u} = (u_j, j \in S)$. Theorem 13, which is obviously valid in a finite setting too, then tells us that absorption at 0 must be certain and $\alpha = \alpha_0 > 0$. Our task will be to characterize the time-scale on which the conditional probabilities $\mathbb{P}_i(X(t) = j | T > t)$ converge to their limits u_j , and to relate it to the time-scale on which absorption takes place.

We characterize the time-scale on which absorption takes place by α_0 , the (common) rate of convergence to zero of the probabilities $\mathbb{P}_i(T > t) = 1 - P_{i0}(t)$, and the time-scale on which the conditional probabilities $\mathbb{P}_i(X(t) = j | T > t)$ converge to their limits $u_j, j \in S$, by

$$\beta := \inf\{\beta_{ij}, i, j \in S\}, \tag{34}$$

where β_{ij} is the rate of convergence of the conditional probability $\mathbb{P}_i(X(t) = j | T > t)$ to its limit u_j , that is,

$$\beta_{ij} := \inf \left\{ a \geq 0 : \int_0^\infty e^{at} |\mathbb{P}_i(X(t) = j | T > t) - u_j| dt = \infty \right\}, \quad i, j \in S. \tag{35}$$

First assuming that the state space of \mathcal{X} is finite, we know that $-\alpha$, the eigenvalue of Q with the largest real part, is real, simple and negative. It follows from classical matrix theory (see, for example, Schmidt [119, Theorem 1.3]) that if $-\alpha_2$ is an eigenvalue of Q with next largest real part (and largest multiplicity, if there are several such eigenvalues with different multiplicities), then the result (11) can

be stated more precisely as

$$e^{\alpha t} P_{ij}(t) = v_i u_j + \mathcal{O}(t^{m-1} e^{-\gamma t}) \quad \text{as } t \rightarrow \infty, \quad i, j \in S, \quad (36)$$

where m is the multiplicity of $-\alpha_2$ and

$$\gamma := \operatorname{Re}(\alpha_2) - \alpha > 0. \quad (37)$$

We will refer to γ as the *spectral gap* of Q (or \mathcal{X}). Since $\alpha_0 = \alpha$, it now follows readily that $\mathbb{P}_i(X(t) = j | T > t) - u_j = \mathcal{O}(t^{m-1} e^{-\gamma t})$, so that $\beta_{ij} \geq \gamma$ for all $i, j \in S$, and hence $\beta \geq \gamma > 0$. Perhaps examples can be constructed in which $\beta > \gamma$, but as a rule there will be some $i, j \in S$, such that the spectral expansion of $P_{ij}(t) - u_j \mathbb{P}_i(T > t)$ involves a term that is the product of $e^{-\alpha_2 t}$ and a nonzero polynomial in t , in which case $\beta = \gamma$.

The preceding observations lead to the conclusion that the type of “quasi stationary” behaviour depicted in Figure 1 occurs in the setting of a finite Markov chain only if the spectral gap γ is substantially larger than the decay parameter α . This fact was noted already in [41] in a discrete-time setting. (See [39, 40] for a similar conclusion in the setting of a finite birth-death process.)

The situation is more complicated when the state space $\{0\} \cup S$ of the Markov chain \mathcal{X} is countably infinite, all other circumstances being unchanged. Again α_0 (as defined in (20)) may be used to characterize the time-scale on which absorption takes place (and we still have $\alpha_0 = \alpha$), but we must resort to operator theory to determine the rate of convergence of $\mathbb{P}_i(X(t) = j | T > t) - u_j$, by extending the notion of *spectral gap* to Markov chains with a countably infinite state space. While the spectral gap for ergodic Markov chains, and in particular reversible ergodic Markov chains, has received quite some attention in the literature (see, for example, Chen [29] and the references therein), much less is known for absorbing Markov chains. However, if we restrict ourselves to birth-death processes (with killing), as we do in the next section, more detailed results can be obtained. Since the behaviour of birth-death processes is often indicative of what holds true in much greater generality, the results of the next section lead one to suspect that for a wide class of Markov chains α -positivity is necessary (but by no means sufficient) for $\gamma > 0$, and hence for the type of quasi-stationary behaviour depicted in Figure 1.

Concluding this section we note that Pollett and Roberts [116] explain quasi stationarity using a dynamical systems approach. They proved, under mild conditions, that the Kolmogorov forward equations always admit a *centre manifold* M consisting of points on the straight line in the unit $|S|$ -simplex connecting the quasi-stationary distribution \mathbf{u} with the degenerate limiting distribution $\boldsymbol{\pi}_0 = (1 \ \mathbf{0})$. If (and only if) we are dealing with the type of quasi-stationary behaviour depicted in

Figure 1, then the state-probability vector evolves exponentially quickly towards M and then slowly in a region near M towards its eventual limit π_0 .

5 Birth-death and related processes

In this section we will analyse in detail some special continuous-time Markov chains with a countably infinite, irreducible state space. So our setting is that of Subsection 3.3. We consider birth-death processes in Subsection 5.1 and birth-death processes *with killing* in Subsection 5.2.

5.1 Birth-death processes

Let the Markov chain \mathcal{X} of Subsection 3.3 now be a *birth-death process*, that is, the generator (1) of \mathcal{X} satisfies $q_{ij} = 0$ if $|i - j| > 1$. Since $S = \{1, 2, \dots\}$ is supposed to be irreducible we must have

$$\lambda_i := q_{i,i+1} > 0 \text{ and } \mu_{i+1} := q_{i+1,i} > 0, \quad i \in S.$$

We also require that $a_1 > 0$ and $a_i = 0$ for $i > 1$, so that absorption at state 0 (*killing*) can only occur via state 1. (This assumption will be relaxed in the next subsection.) The parameters λ_i and μ_i are the *birth rate* and *death rate*, respectively, in state i . In the literature the killing rate a_1 is usually referred to as the death rate in state 1 (and denoted by μ_1). Throughout this subsection the birth and death rates are assumed to satisfy

$$\sum_{n=1}^{\infty} \frac{1}{\lambda_n \pi_n} = \infty, \tag{38}$$

where

$$\pi_1 := 1 \text{ and } \pi_n := \frac{\lambda_1 \lambda_2 \dots \lambda_{n-1}}{\mu_2 \mu_3 \dots \mu_n}, \quad n > 1, \tag{39}$$

which is necessary and sufficient for absorption to be certain and, hence, sufficient for \mathcal{X} to be nonexplosive (see, for example, [4, Section 8.1]).

Karlin and McGregor [72] have shown that the transition probabilities $P_{ij}(t)$ can be represented as

$$P_{ij}(t) = \pi_j \int_0^{\infty} e^{-xt} Q_{i-1}(x) Q_{j-1}(x) \psi(dx), \quad t \geq 0, \quad i, j \in S. \tag{40}$$

Here $\{Q_n\}$ is a sequence of polynomials satisfying the recurrence relation

$$\begin{aligned} \lambda_n Q_n(x) &= (\lambda_n + \mu_n - x) Q_{n-1}(x) - \mu_n Q_{n-2}(x), \quad n > 1, \\ \lambda_1 Q_1(x) &= \lambda_1 + a_1 - x, \quad Q_0(x) = 1, \end{aligned} \tag{41}$$

and ψ – the *spectral measure* of \mathcal{X} – is the (unique) Borel measure of total mass 1 on the interval $(0, \infty)$ with respect to which the *birth-death polynomials* Q_n are orthogonal.

The next theorem shows how the decay parameters α, α_0 and β are related to $\text{supp}(\psi)$ – the support of the measure ψ , also referred to as the *spectrum* of \mathcal{X} – and more specifically to

$$\xi_1 := \inf \text{supp}(\psi) \quad \text{and} \quad \xi_2 := \inf \{ \text{supp}(\psi) \cap (\xi_1, \infty) \}. \quad (42)$$

The difference $\xi_2 - \xi_1$ is the *spectral gap* of \mathcal{X} . Not surprisingly, it can be interpreted as the limit as $n \rightarrow \infty$ of the spectral gap (as defined in (37)) of a suitably truncated birth-death process on $\{0, 1, 2, \dots, n\}$. Interestingly, ξ_1 and ξ_2 are also the limits as $n \rightarrow \infty$ of the smallest zero and second smallest zero, respectively, of the polynomial $Q_n(x)$, all of whose zeros are distinct, real and positive (see, for example, Chihara [32]).

Theorem 15 The birth-death process \mathcal{X} has $\alpha = \alpha_0 = \xi_1$ and $\beta = \xi_2 - \xi_1$.

Proof Evidently, (23) implies $\alpha = \alpha_0$. The representation for α was established by Callaert [23], and that for β in [45, Theorem 5]. \square

In the setting at hand the existence of a quasi-stationary distribution can be established under much weaker conditions than those of the Theorems 11 or 12. In fact, certain absorption and exponential transience, which are necessary conditions by Theorem 8, happen to be sufficient as well. Moreover, all quasi-stationary distributions can actually be identified. A crucial role is played by the series

$$\sum_{n=1}^{\infty} \frac{1}{\lambda_n \pi_n} \sum_{j=n+1}^{\infty} \pi_j, \quad (43)$$

which, by a result of Keilson's [75] (see also [4, Section 8.1] or [80, Section 5.1]), can be interpreted as the limit as $n \rightarrow \infty$ of the expected first passage time from state n to state 1.

Theorem 16 [45, Theorem 3.2] Let the birth-death process \mathcal{X} be such that absorption is certain and $\alpha > 0$. If the series (43) diverges, then there is a one-parameter family of quasi-stationary distributions $\{\mathbf{u}(x) = (u_i(x), i \in S), 0 < x \leq \alpha\}$ for \mathcal{X} , where

$$u_i(x) = \frac{x}{a_1} \pi_i Q_{i-1}(x), \quad i \in S. \quad (44)$$

If the series (43) converges, then there is precisely one quasi-stationary distribution for \mathcal{X} , namely $\mathbf{u}(\alpha)$.

It is enlightening to interpret $\mathbf{u}(x) = (u_i(x), i \in S)$ of (44) from the more general perspective of Theorem 6. Indeed, it is not difficult to see that $\mathbf{u}(x)$ is the unique x -invariant vector for Q satisfying $x = a_{\mathbf{u}(x)} (= a_1 u_1(x))$. But $\mathbf{u}(x)$ represents a proper distribution (if and) only if $x = \alpha$, or, $0 < x < \alpha$ and the series (43) diverges. If $0 < x < \alpha$ and the series (43) converges, it is possible to renormalize $\mathbf{u}(x)$ such that a proper distribution \mathbf{v} results, but then $x < a_{\mathbf{v}}$ (see [45] for details). In any case we have the following.

Corollary 17 If absorption is certain and $\alpha > 0$, then $\mathbf{u}(\alpha)$ is the unique α -invariant (and hence minimal) quasi-stationary distribution for \mathcal{X} .

Note that α -recurrence is not required for the existence of a quasi-stationary distribution, nor for uniqueness of the α -invariant quasi-stationary distribution. Theorem 16 also shows that asymptotic remoteness (see (24)) is not necessary for the existence of a quasi-stationary distribution, since asymptotic remoteness implies divergence of the series (43). Indeed, by Karlin and McGregor [73, Theorem 10] (or, for example, [89, Corollary 1.2.4.1]) we have

$$\mathbb{E}_i(T) = \frac{1}{a_1} \sum_{j=1}^{\infty} \pi_j + \sum_{n=1}^{i-1} \frac{1}{\lambda_n \pi_n} \sum_{j=n+1}^{\infty} \pi_j, \quad i \in S, \quad (45)$$

while $\mathbb{E}_i(T) \geq t \mathbb{P}_i(T > t)$ for all $t \geq 0$ by Markov's inequality. So if asymptotic remoteness prevails we must have $\mathbb{E}_i(T) \rightarrow \infty$ as $i \rightarrow \infty$, implying divergence of (43). (Actually, the converse holds true as well, cf. [108, Lemma 2.2]). As an aside we remark that divergence of (43) is also equivalent to the Kolmogorov forward equations having a unique solution (see, for example, [4, Section 8.1]), the prevailing situation in most practical models.

In summary, to establish the existence of a quasi-stationary distribution for a birth-death process we have to verify certain absorption and exponential transience. Given the birth and death rates verification of certain absorption is trivial through (38), but it is less straightforward to establish exponential transience. Some necessary and some sufficient conditions, which settle the problem for most processes encountered in practice, have been collected in [44]. One such result is

$$\alpha > 0 \implies \sum_{n=1}^{\infty} \pi_n < \infty, \quad (46)$$

which is implied by [73, Equation (9.19)] and the fact that $\alpha = \xi_1$. Also, a complete solution is given in [44] for processes having birth rates λ_i and death rates μ_i that are asymptotically rational functions of i .

Next turning to limiting conditional distributions the situation is again quite straightforward in the setting at hand, provided we restrict ourselves to initial distributions that are concentrated on a single state (or a finite set of states), cf. Theorem 13.

Theorem 18 [45, Theorem 4.1] Let the birth-death process \mathcal{X} be such that absorption is certain. Then, for all $i \in S$,

$$\lim_{t \rightarrow \infty} \mathbb{P}_i(X(t) = j | T > t) = \frac{\alpha}{a_1} \pi_j Q_{j-1}(\alpha), \quad j \in S, \quad (47)$$

so that these conditional limits constitute the α -invariant quasi-stationary distribution for \mathcal{X} if $\alpha > 0$.

In Section 4 we have emphasized that one should use the (minimal) quasi-stationary distribution to approximate the unconditional time-dependent distribution of an absorbing Markov chain only if the decay parameter α of the unconditional process is substantially smaller than the decay parameter β of the process conditioned on nonabsorption. Unfortunately, determining $\beta = \gamma$ (the spectral gap) is usually at least as difficult as determining $\alpha = \xi_1$ (the first point in the spectrum). However, some information on β can be obtained if α is known explicitly. Namely, Miclo [97] and Chen [28, Theorem 3.5] have presented a necessary and sufficient condition for the spectral gap $\xi_2 - \xi_1$ to be positive in the setting of a nonabsorbing, nonexplosive, ergodic birth-death process (in which case $\xi_1 = 0$). With the transformation technique employed, for instance, in [46, Section 5] this result can be translated into the setting at hand, yielding

$$\beta > 0 \iff \sup_i \left(\sum_{n=1}^i \frac{1}{\lambda_n \pi_n Q_{n-1}(\alpha) Q_n(\alpha)} \right) \left(\sum_{n=i+1}^{\infty} \pi_n Q_{n-1}^2(\alpha) \right) < \infty. \quad (48)$$

It follows in particular that

$$\beta > 0 \implies \sum_{n=1}^{\infty} \pi_n Q_{n-1}^2(\alpha) < \infty. \quad (49)$$

Indeed, by a classical result in the theory of orthogonal polynomials (see Shohat and Tamarkin [126, Corollary 2.6]), the conclusion in (49) is equivalent to the spectral measure ψ having a point mass at α , which is obviously necessary for the spectral gap to be positive. Interestingly, [47, Theorem 3.1] (which is a corrected version of [46, Theorem 5.1]) tells us that the conclusion is also equivalent to α -positivity of \mathcal{X} . In other words, an absorbing birth-death process with decay parameter α will have $\beta = 0$ if it is α -transient or α -null. As mentioned already in the previous section, we suspect this conclusion to be valid in much greater generality.

5.2 Birth-death processes with killing

We generalize the setting of the previous subsection by allowing absorption from any state $i \in S$ rather than only state 1, that is, we allow $a_i \geq 0$ for all $i \in S$. We will assume that absorption is certain, which, by van Doorn and Zeifman [54, Theorem 1], is now equivalent to

$$\sum_{n=1}^{\infty} \frac{1}{\lambda_n \pi_n} \sum_{j=1}^n a_j \pi_j = \infty. \quad (50)$$

Evidently, we must have $a_i > 0$ for at least one state $i \in S$. The representation (40) remains valid provided we redefine the polynomials Q_n by means of the recurrence relation

$$\begin{aligned} \lambda_n Q_n(x) &= (\lambda_n + \mu_n + a_n - x)Q_{n-1}(x) - \mu_n Q_{n-2}(x), \quad n > 1, \\ \lambda_1 Q_1(x) &= \lambda_1 + a_1 - x, \quad Q_0(x) = 1, \end{aligned} \quad (51)$$

which reduces to (41) in the specific setting of the previous subsection. The quantities π_n are as in (39), and the measure ψ is again the (unique) Borel measure of total mass 1 on the interval $(0, \infty)$ with respect to which the polynomials Q_n are orthogonal. Defining ξ_1 and ξ_2 as in (42) we still have $\alpha = \xi_1$ and $\beta = \xi_2 - \xi_1$, but we do not necessarily have $\alpha_0 = \alpha$ any longer. (See [37] for proofs and developments.) However, by (23), we do have $\alpha_0 = \alpha$ if $S_{\text{abs}} = \{i \in S : a_i > 0\}$ is finite, and under this condition Theorem 16 can be generalized as follows.

Theorem 19 [37, Theorems 6.5 and 6.6] Let the birth-death process with killing \mathcal{X} be such that S_{abs} is finite, absorption is certain, and $\alpha > 0$. If the series (43) diverges, then there is a one-parameter family of quasi-stationary distributions $\{\mathbf{u}(x) = (u_i(x), i \in S), 0 < x \leq \alpha\}$ for \mathcal{X} , where

$$u_i(x) = c(x) \pi_i Q_{i-1}(x), \quad i \in S, \quad (52)$$

and $c(x)$ is a normalizing constant. If the series (43) converges, then there is precisely one quasi-stationary distribution for \mathcal{X} , namely $\mathbf{u}(\alpha)$.

Without a restriction on S_{abs} the situation is less clear. Indeed, let $0 < x \leq \alpha$. It is easy to see that, up to a multiplicative constant, $(\pi_i Q_{i-1}(x), i \in S)$ is the unique x -invariant vector for Q . We also have $\pi_i Q_{i-1}(x) > 0$, as shown in [37]. So for $\mathbf{u}(x) = (u_i(x), i \in S)$ of (52) to be a quasi-stationary distribution it is, in view of Theorem 6, necessary and sufficient that $\sum_{i \in S} \pi_i Q_{i-1}(x)$ converges and $x = a_{\mathbf{u}}$, or equivalently,

$$x \sum_{i \in S} \pi_i Q_{i-1}(x) = \sum_{i \in S} a_i \pi_i Q_{i-1}(x) < \infty. \quad (53)$$

In view of the last statement of Theorem 19 the equality sign need not prevail, but even if it does the sums need not be bounded. Of course the latter can happen only if S_{abs} is infinite. Examples are given in [37] of birth-death processes with killing with certain absorption and $\alpha > 0$ for which

- (i) no quasi-stationary distribution exists, and
- (ii) an x -invariant quasi-stationary distributions exists if and only if $\gamma < x \leq \alpha$ for some $\gamma > 0$.

So the simple structure that prevails in the setting of pure birth-death processes is lost already in the setting of birth-death processes with killing as soon as we allow infinitely many positive killing rates. We can nevertheless give some sufficient conditions for the existence of a quasi-stationary distribution under these circumstances. The first result involves the condition (38), which can be interpreted in terms of the *unkilled* process – the process one obtains by setting all killing rates equal to 0 – as necessary and sufficient for recurrence (see [4, Section 8.1]).

Theorem 20 [49, Theorem 4.2] Let the birth-death process with killing \mathcal{X} be such that absorption is certain, the series (43) diverges, and $\alpha > \limsup_{i \rightarrow \infty} a_i$. Then there exists a quasi-stationary distribution for \mathcal{X} if and only if (38) is satisfied, in which case there is a one-parameter family of quasi-stationary distributions $\{\mathbf{u}(x) = (u_i(x), i \in S), 0 < x \leq \alpha\}$, with u_i given by (52).

We note that divergence of (43) is equivalent to (38) if S_{abs} is finite, so this theorem generalizes part of Theorem 19. Theorem 20 also generalizes [48, Theorem 2], in which only the existence of a quasi-stationary distribution is concluded (and the divergence of (43) as a condition is unjustifiedly omitted).

Our final result is another sufficient condition for the existence of a quasi-stationary distribution. It was first given in [48, Theorem 1] but again with the unjustified omission of the divergence of (43) as a condition.

Theorem 21 [49, Theorem 4.3] Let the birth-death process with killing \mathcal{X} be such that absorption is certain, the series (43) diverges, and $\alpha < \liminf_{i \rightarrow \infty} a_i$. Then $\{\mathbf{u}(\alpha) = (u_i(\alpha), i \in S)\}$, with u_i given by (52), constitutes a quasi-stationary distribution for \mathcal{X} .

6 Computational aspects

Suppose that $S = \{1, 2, \dots, n\}$, so that Q is an $n \times n$ matrix. Restricting our attention to the case where S is irreducible, we seek to compute the left eigenvector $\mathbf{u} = (u_i, i \in S)$ of Q corresponding to the eigenvalue, $-\alpha$, with maximal real part (being simple, real and strictly negative). Once normalized

so that $\mathbf{u}\mathbf{1}^T = 1$, \mathbf{u} is the unique quasi-stationary (and then limiting conditional) distribution. If S is reducible, then we would address an eigenvector problem within a restricted set of states (typically, when $-\alpha$ has geometric multiplicity one, we would determine \mathbf{u} over states that are accessible from the minimal class S_{\min} , and then put $u_j = 0$ whenever j is not accessible from S_{\min}).

Most standard mathematical software permits evaluation of all, or some, of the eigenvalues and/or eigenvectors of a square matrix, so it should be a routine task to evaluate \mathbf{u} . In the web appendix accompanying this paper we describe and compare a range of numerical methods, giving guidance on how to implement these in MATLAB[®], perhaps the most widely used package for scientific computing. To illustrate the application of these methods, we will consider two examples, the first being the metapopulation model referred to in Section 1 and second being an elaboration that accounts for a dynamic landscape.

Example 1 Let $X(t)$ be the number of occupied patches in a network consisting of a fixed number of patches n . Each occupied patch becomes empty at rate e and colonization of empty patches occurs at rate c/n for each occupied-unoccupied pair. Thus we have a birth-death process over a finite space $\{0\} \cup S$, where $S = \{1, 2, \dots, n\}$, with birth rates $\lambda_i = (c/n)i(n-i)$ and death rates $\mu_i = ei$. It is sometimes called the SIS (susceptible-infectious-susceptible) model, for it is often used to model the number of infectives in a population of fixed size n , with per-capita recovery rate e and per-proximate encounter transmission rate c . It is an example of Feller's [58] *stochastic logistic model*, and one of the earliest stochastic models for the spread of infections that do not confer any long lasting immunity, and where individuals become susceptible again after infection (Weiss and Dishon [138]). It appears, not only in ecology and epidemiology, but also in the propagation of rumours (Bartholomew [17]) and in chemical reaction kinetics (Oppenheim et al. [105]).

Clearly S is irreducible, and so there is a unique quasi-stationary distribution $\mathbf{u} = (u_i, i \in S)$ and this has a limiting conditional interpretation. Whilst it can be written down in terms of the coefficients (39) and the birth-death polynomials (41), by way of Theorems 16 and 18, neither these quantities, nor indeed the decay parameter α , can be exhibited explicitly. However, the quasi-stationary distribution is easily computed numerically. It is depicted in Figures 1 and 2 for the 20-patch model with $c = 0.1625$ and $e = 0.0325$. Figure 2 also depicts the pseudo-transient distribution when the initial distribution assigns all its mass to state 1, that is,

$$\pi_i = \pi_1 \frac{(n-1)!}{i(n-i)!} \left(\frac{c}{en}\right)^{i-1}, \quad i = 1, 2, \dots, n$$

(see Clancy and Pollett [34], Kryscio and Lefèvre [84], Nåsell [102]) and the Gaussian approximation

(for large n , $X(t)/n$ has an approximate normal $N(1 - e/c, e/c)$ distribution; see for example [115]). The remarkable closeness of the pseudo-transient distribution to the quasi-stationary distribution is explained in [14]. Further approximations to quasi-stationary distribution can be found in Nåsell [101, 102, 103], Ovaskainen [106] and Clancy and Mendy [33].

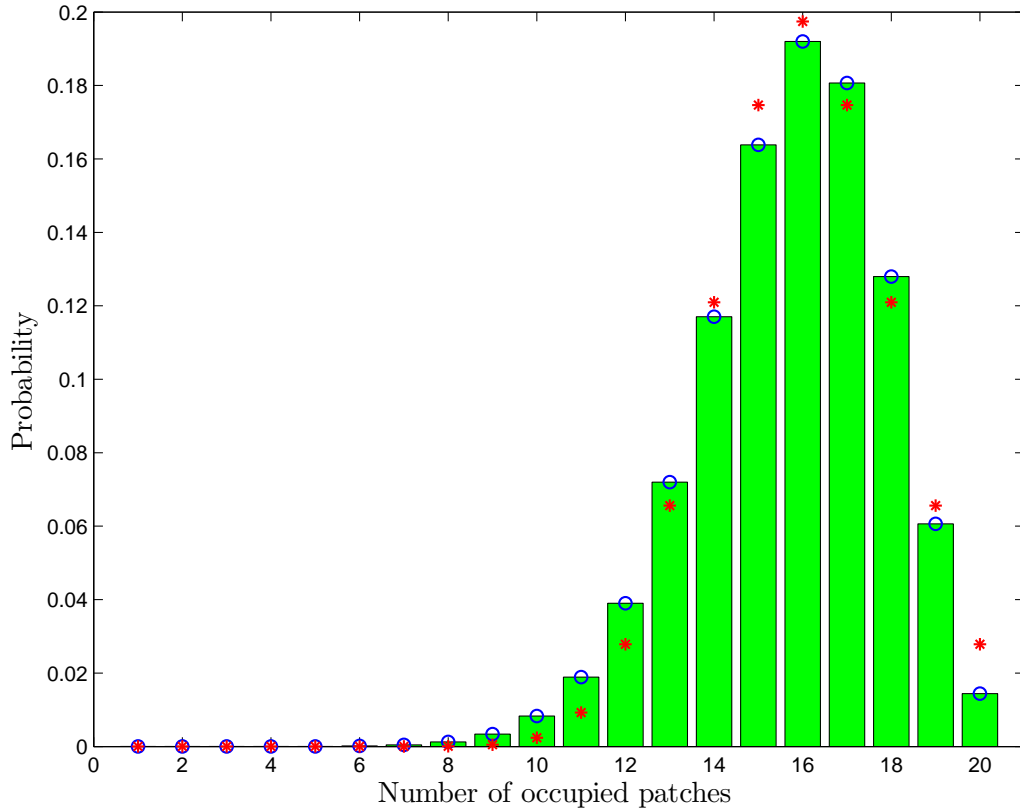


Fig 2. The quasi-stationary distribution (bars), the pseudo-transient distribution when the initial distribution assigns all its mass to state 1 (circles) and the Gaussian approximation (asterisks), for the 20-patch (SIS) metapopulation model with colonization rate $c = 0.1625$ and local extinction rate $e = 0.0325$.

Next we consider an elaboration of Example 1 where the number of patches available for occupancy is a stochastic variable (Ross [118]).

Example 2 The network consists of a fixed number of patches n , but only $Y(t)$ of these are suitable for occupancy. As before, $X(t)$ is the number of occupied patches, but now $0 \leq X(t) \leq Y(t)$. The

process $((X(t), Y(t)))$ is assumed to have non-zero transition rates

$$\begin{aligned} q((x, y); (x, y + 1)) &= r(n - y), \\ q((x, y); (x, y - 1)) &= d(y - x), \\ q((x, y); (x - 1, y - 1)) &= dx, \\ q((x, y); (x + 1, y)) &= (c/n)x(y - x), \\ q((x, y); (x - 1, y)) &= ex. \end{aligned}$$

The additional parameters d and r are, respectively, the disturbance rate (the per-capita rate at which patches suitable for occupancy become unsuitable) and the recovery rate (the per-capita rate at which unsuitable patches become suitable); notice that if an occupied patch is disturbed it also suffers local extinction. The state space consists of a set $S = \{(x, y) : 1 \leq x \leq y \leq n\}$ (irreducible) of transient states that correspond to at least one patch being occupied (“extant” states), and a set $A = \{(0, y) : 0 \leq y \leq n\}$ (irreducible) in which the process is eventually trapped.

Since S is irreducible and finite, there is a unique quasi-stationary distribution $\mathbf{u} = (u_{(x,y)}, (x, y) \in S)$ and this has a limiting conditional interpretation. Whilst it cannot be exhibited explicitly it can be computed numerically. Figure 3 plots the quasi-stationary distribution of the 100-patch metapopulation model with parameters $e = 0.1$, $c = 0.6$, $d = 0.1$ and $r = 0.5$, which was evaluated using sparse-matrix techniques; for details see the web appendix. We also estimated the quasi-stationary distribution using *one* iteration of the return map starting from the diffusion approximation of Ross [118, Page 797] (the stationary distribution of the approximating Ornstein-Uhlenbeck process) and obtained an excellent approximation.

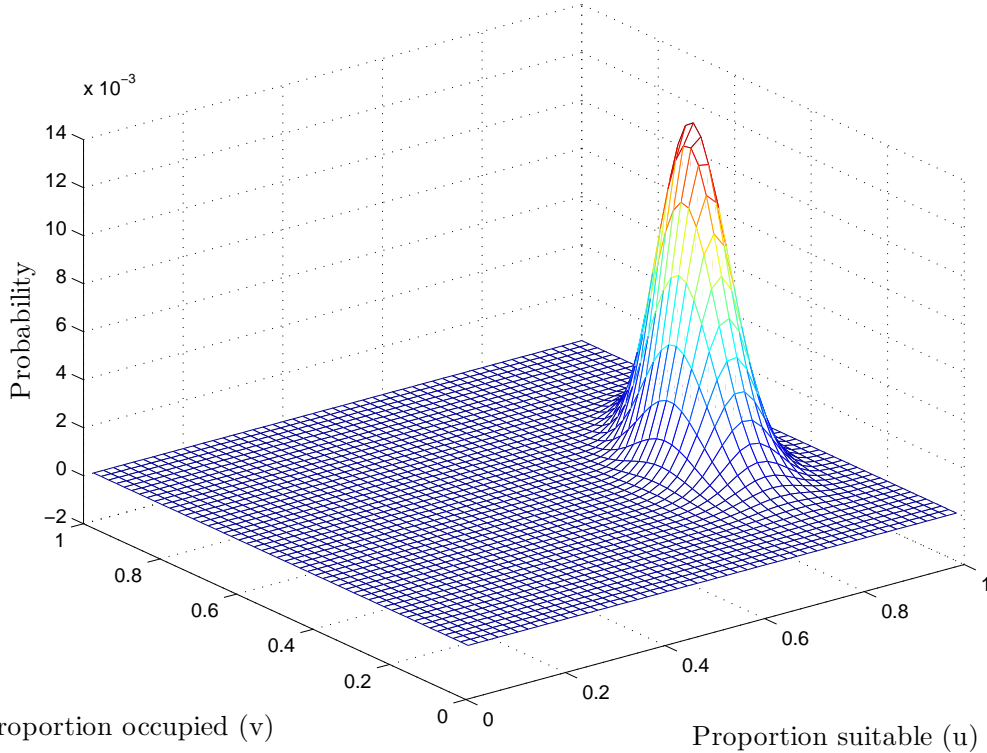


Fig 3. The quasi-stationary distribution of the 100-patch metapopulation model.

Finally we note that if the state space is infinite, then we can employ a truncation procedure whereby the infinite Q is approximated by a sequence $\{Q^{(n)}\}$ of irreducible finite square matrices (see for example Gibson and Seneta [62], Seneta [121, 122], Tweedie [130, 133]), anticipating that the corresponding sequence $\{\mathbf{u}^{(n)}\}$ of normalized left eigenvectors approximates the desired quasi-stationary distribution \mathbf{u} . When S is irreducible we can always construct an increasing sequence $\{S^{(n)}\}$ of *irreducible* finite subsets of S with limit S (see Breyer and Hart [21, Lemma 1]). Successive $\mathbf{u}^{(n)}$ would be evaluated as above. Ideally we would want truncations large enough to capture quasi stationarity of the process, and thus in choosing $\{S^{(n)}\}$ we might be led by the results of a simulation study or an analytical approximation. If the state space is multi-dimensional care would be needed to find an appropriate state-space enumeration that facilitates appending states as n increases; Brent's algorithm [20] provides one such method. If $Q^{(n)}$ is simply Q restricted to $S^{(n)}$ (and $\{S^{(n)}\}$ is an increasing sequence of irreducible finite subsets with limit S), then the corresponding sequence $\{\alpha^{(n)}\}$ of decay parameters will converge to α , the decay parameter of S [21, Lemma 2]. However, convergence of the corresponding sequence $\{\mathbf{u}^{(n)}\}$ of left eigenvectors is not guaranteed, let alone convergence to \mathbf{u} ; see [123, Section 7.3] and [21, Section 4] for further details. Whilst technical subtleties abound in the most general setting, truncation is known to work under natural conditions including, for example,

α -positivity, and is valid for a variety of standard models, including birth-death processes (Kijima and Seneta [82]) and branching processes [79].

Acknowledgement

The work of Phil Pollett is supported by the Australian Research Council (Discovery grants DP110101929 and DP120102398 and the Centre of Excellence for Mathematics and Statistics of Complex Systems). We are grateful to Tony Pakes for drawing our attention to early historical references and to the referees for their valuable comments.

References

- [1] Aalen, O.O. and Gjessing, H.K. (2001). Understanding the shape of the hazard rate: a process point of view. *Statist. Sci.* **16**, 1-22.
- [2] Aalen, O.O. and Gjessing, H.K. (2003). A look behind survival data: underlying processes and quasi-stationarity. In Lindqvist, B.H. and Doksum, K.A. (Eds.), *Mathematical and Statistical Methods in Reliability*, pp. 221-234. Singapore: World Scientific Publishing, Singapore.
- [3] Albert, E. (1963). Markov chains and λ -invariant measures. *J. Math. Anal. Appl.* **6**, 404-418.
- [4] Anderson, W.J. (1991). *Continuous-Time Markov Chains*. New York: Springer.
- [5] Arjas, E., Nummelin, E. and Tweedie, R.L. (1980). Semi-Markov processes on a general state space: α -theory and quasistationarity. *J. Austral. Math. Soc. Ser. A* **30**, 187-200.
- [6] Artalejo, J.R. (2012). On the time to extinction from quasi-stationarity: A unified approach. *Physica A* **391**, 4483-4486.
- [7] Artalejo, J.R., Economou, A. and Lopez-Herrero, M.J. (2010). On the number of recovered individuals in the *SIS* and *SIR* stochastic epidemic models. *Math. Biosc.* **228**, 45-55.
- [8] Artalejo, J.R., Economou, A. and Lopez-Herrero, M.J. (2010). The maximum number of infected individuals in *SIS* epidemic models: Computational techniques and quasi-stationary distributions. *J. Comput. Appl. Math.* **233**, 2563-2574.
- [9] Artalejo, J.R. and Lopez-Herrero, M.J. (2010). Quasi-stationary and ratio of expectations distributions: A comparative study. *J. Theor. Biology* **266**, 264-274.

- [10] Atman, A.P.F. and Dickman, R. (2002). Quasistationary distributions for the Domany-Kinzel stochastic cellular automaton. *Phys. Rev. E* **66**, 046135, 9 pages.
- [11] Barbour, A.D. (1972). The principle of the diffusion of arbitrary constants. *J. Appl. Probab.* **9**, 519-541.
- [12] Barbour, A.D. (1974). On a functional central limit theorem for Markov population processes. *Adv. Appl. Probab.* **6**, 21-39.
- [13] Barbour, A.D. (1976). Quasi-stationary distributions in Markov population processes. *Adv. Appl. Probab.* **8**, 296-314.
- [14] Barbour, A.D. and Pollett, P.K. (2010). Total variation approximation for quasi-equilibrium distributions. *J. Appl. Probab.* **47**, 934-946.
- [15] Bartlett, M.S. (1957). On theoretical models for competitive and predatory biological systems. *Biometrika* **44**, 27-42.
- [16] Bartlett, M.S. (1960). *Stochastic Population Models in Ecology and Epidemiology*. London: Methuen.
- [17] Bartholomew, D.J. (1976). Continuous time diffusion models with random duration of interest. *J. Math. Sociol.* **4**, 187-199.
- [18] Bobrowski, A. (2004). Quasi-stationary distributions of a pair of Markov chains related to time evolution of a DNA locus. *Adv. Appl. Probab.* **36**, 56-77.
- [19] Boucherie, R.J. (1997). On the quasi-stationary distribution for queueing networks with defective routing. *J. Austral. Math. Soc. B* **38**, 454-463.
- [20] Brent, R.P. (1973). Reducing the retrieval time of scatter storage techniques. *Communications of the ACM* **16**, 105-109.
- [21] Breyer, L.A. and Hart, A.G. (2000). Approximations of quasi-stationary distributions for Markov chains. *Math. Computer Modelling* **31**, 69-79.
- [22] Buiculescu, M. (1975). On quasi-stationary distributions for multi-type Galton-Watson processes. *J. Appl. Probab.* **12**, 60-68.

- [23] Callaert, H. (1974). On the rate of convergence in birth-and-death processes. *Bull. Soc. Math. Belg.* **26**, 173-184.
- [24] Cattiaux, P., Collet, P., Lambert, A., Martínez, S., Méléard, S. and San Martín, J. (2009). Quasi-stationary distributions and diffusion models in population dynamics. *Ann. Probab.* **37**, 1926-1969.
- [25] Cavender, J.A. (1978). Quasistationary distributions of birth-death processes. *Adv. Appl. Probab.* **10**, 570-586.
- [26] Chan, T., Pollett, P.K. and Weinstein, M.C. (2009). Quantitative risk stratification in Markov chains with limiting conditional distributions. *Medical Decision Making* **29**, 532-540.
- [27] Chen, A., Li, J., Hou, Z. and Ng, K.W. (2010). Decay properties and quasi-stationary distributions for stopped Markovian bulk-arrival and bulk-service queues. *Queueing Syst.* **66**, 275-311.
- [28] Chen, M.-F. (2000). Explicit bounds of the first eigenvalue. *Sci. China Ser. A* (English ed.) **43**, 1051-1059.
- [29] Chen, M.-F. (2005). *Eigenvalues, Inequalities, and Ergodic Theory*. london: Springer.
- [30] Cheong, C.K. (1970). Quasi-stationary distributions in semi-Markov processes. *J. Appl. Probab.* **7**, 388-399. [Correction: *J. Appl. Probab.* **7**, 788.]
- [31] Cheong, C.K. (1972). Quasi-stationary distributions for the continuous-time Galton-Watson process. *Bull. Soc. Math. Belgique* **24**, 343-350.
- [32] Chihara, T.S. (1978). *An Introduction to Orthogonal Polynomials*. New York: Gordon and Breach.
- [33] Clancy, D. and Mendy, S.T. (2011). Approximating the quasi-stationary distribution of the SIS model for endemic infection. *Methodol. Comput. Appl. Probab.* **13**, 603-618.
- [34] Clancy, D. and Pollett, P.K. (2003). A note on quasi-stationary distributions of birth-death processes and the SIS logistic epidemic. *J. Appl. Probab.* **40**, 821-825.
- [35] Collet, P., Martínez, S. and San Martín, J. (2013). *Quasi-Stationary Distributions*. Heidelberg: Springer.

- [36] Collet, P., Martínez, S. and Schmitt, B. (1996). Quasi-stationary distribution and Gibbs measure of expanding systems. *Nonlinear Phenomena and Complex Systems* **1**, 205-219.
- [37] Coolen-Schrijner, P. and van Doorn, E.A. (2006). Quasi-stationary distributions for birth-death processes with killing. *J. Appl. Math. Stochastic Anal.* **2006**, Article ID 84640, 15 pages.
- [38] Daley, D.J. (1969). Quasi-stationary behaviour of a left-continuous random walk. *Ann. Math. Statist.* **40**, 532-539.
- [39] Dambrine, S. and Moreau, M. (1981). Note on the stochastic theory of a self-catalytic chemical reaction, I. *Physica A* **106**, 559-573.
- [40] Dambrine, S. and Moreau, M. (1981). Note on the stochastic theory of a self-catalytic chemical reaction, II. *Physica A* **106**, 574-588.
- [41] Darroch, J.N. and Seneta, E. (1965). On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *J. Appl. Probab.* **2**, 88-100.
- [42] Darroch, J.N. and Seneta, E. (1967). On quasi-stationary distributions in absorbing continuous-time finite Markov chains. *J. Appl. Probab.* **4**, 192-196.
- [43] Day, J.R. and Possingham, H.P. (1995). A stochastic metapopulation model with variable patch size and position. *Theoret. Pop. Biol.* **48**, 333-360.
- [44] van Doorn, E.A. (1985). Conditions for exponential ergodicity and bounds for the decay parameter of a birth-death process. *Adv. Appl. Probab.* **17**, 514-530.
- [45] van Doorn, E.A. (1991). Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes. *Adv. Appl. Probab.* **23**, 683-700.
- [46] van Doorn, E.A. (2003). Birth-death processes and associated polynomials. *J. Comput. Appl. Math.* **153**, 497-506.
- [47] van Doorn, E.A. (2006). On the α -classification of birth-death and quasi-birth-death processes. *Stochastic Models* **22**, 411-421.
- [48] van Doorn, E.A. (2012). Conditions for the existence of quasi-stationary distributions for birth-death processes with killing. *Stochastic Process. Appl.* **122**, 2400-2410.

- [49] van Doorn, E.A. (2013). Weighted sums of orthogonal polynomials related to birth-death processes with killing. *Adv. Dyn. Syst. Appl.*, to appear.
- [50] van Doorn, E.A. and Pollett, P.K. (2008). Survival in a quasi-death process. *Linear Alg. Appl.* **429**, 776-791.
- [51] van Doorn, E.A. and Pollett, P.K. (2009). Quasi-stationary distributions for reducible absorbing Markov chains in discrete time. *Markov Process. Related Fields* **15**, 191-204.
- [52] van Doorn, E.A. and Schrijner, P. (1995). Geometric ergodicity and quasi-stationarity in discrete-time birth-death processes. *J. Austral. Math. Soc. (B)* **37**, 121-144.
- [53] van Doorn, E.A. and Schrijner, P. (1995). Ratio limits and limiting conditional distributions for discrete-time birth-death processes. *J. Math. Anal. Appl.* **190**, 263-284.
- [54] van Doorn, E.A. and Zeifman, A.I. (2005). Extinction probability in a birth-death process with killing. *J. Appl. Probab.* **42**, 185-198.
- [55] Evans, S.P. (1990). A quasi-stationary analysis of a virtual path in a B-ISDN network shared by services with very different characteristics. *Comput. Networks ISDN Systems* **20**, 391-399.
- [56] Ewens, W.J. (1963). The diffusion equation and a pseudo-distribution in genetics. *J. Roy. Statist. Soc., Ser B* **25**, 405-412.
- [57] Ewens, W.J. (1964). The pseudo-transient distribution and its uses in genetics. *J. Appl. Probab.* **1**, 141-156.
- [58] Feller, W. (1939). Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in wahrscheinlichkeitstheoretischer Behandlung, *Acta Biotheoretica* **5**, 11-40.
- [59] Ferrari, P.A., Kesten, H. and Martínez, S. (1996). R-positivity, quasi-stationary distributions and ratio limit theorems for a class of probabilistic automata. *Ann. Appl. Probab.* **6**, 577-616.
- [60] Ferrari, P.A., Kesten, H., Martínez, S. and Picco, S. (1995). Existence of quasistationary distributions. A renewal dynamical approach. *Ann. Probab.* **23**, 501-521.
- [61] Flaspohler, D.C. (1974). Quasi-stationary distributions for absorbing continuous-time denumerable Markov chains. *Ann. Inst. Statist. Math.* **26**, 351-356.

- [62] Gibson, D. and Seneta, E. (1987). Augmented truncations of infinite stochastic matrices. *J. Appl. Probab.* **24**, 600-608.
- [63] Good, P. (1968). The limiting behaviour of transient birth and death processes conditioned on survival. *J. Austral. Math. Soc. Ser. B* **8**, 716-722.
- [64] Gosselin, F. (1998). Reconciling theoretical approaches to stochastic patch-occupancy metapopulation models. *Bull. Math. Biol.* **60**, 955-971.
- [65] Green, P.J. (1977). Generalizing the Yaglom limit theorems. In *Recent developments in statistics (Proc. European Meeting Statisticians, Grenoble, 1976)* pp. 441-444). Amsterdam: North-Holland.
- [66] Gyllenberg, M. and Silvestrov, D.S. (1994). Quasi-stationary distributions of a stochastic metapopulation model. *J. Math. Biol.* **33**, 35-70.
- [67] Jacka, S.D. and Roberts, G.O. (1995). Weak convergence of conditioned processes on a countable state space. *J. Appl. Probab.* **32**, 902-916.
- [68] Kac, M. (1947). Random walk and the theory of Brownian motion. *Amer. Math. Monthly* **54**, 369-391.
- [69] Kalpakam, S. and Shahul Hameed, M.A. (1983). Quasi-stationary distribution of a two-unit warm-standby redundant system. *J. Appl. Probab.* **20**, 429-435.
- [70] Kalpakam, S. (1993). On the quasi-stationary distribution of the residual lifetime. *IEEE Trans. Reliab.* **42**, 623-624.
- [71] Van Kampen, N.G. (1961). A power series expansion of the master equation. *Canad. J. Phys.* **39**, 551-567.
- [72] Karlin, S. and McGregor, J.L. (1957). The differential equations of birth-and-death processes, and the Stieltjes moment problem. *Trans. Amer. Math. Soc.* **85**, 489-546.
- [73] Karlin, S. and McGregor, J.L. (1957). The classification of birth and death processes. *Trans. Amer. Math. Soc.* **86**, 366-400.
- [74] Karlin, S. and Tavaré, S. (1982). Linear birth and death processes with killing. *J. Appl. Probab.* **19**, 477-487.

- [75] Keilson, J. (1965). A review of transient behaviour in regular diffusion and birth-death processes. Part II. *J. Appl. Probab.* **2**, 405-428.
- [76] Kesten, H. (1970). Review of Good (1968). *Math. Rev.* **39**, 410.
- [77] Kesten, H. (1995). A ratio limit theorem for (sub)Markov chains on $\{0, 1, 2, \dots\}$ with bounded jumps. *Adv. Appl. Probab.* **27**, 652-691.
- [78] Kijima, M. (1992). On the existence of quasi-stationary distributions in denumerable R-transient Markov chains. *J. Appl. Probab.* **29**, 21-36. [Correction: *J. Appl. Probab.* **30**, 496.]
- [79] Kijima, M. (1993). Quasi-limiting distributions of Markov chains that are skip-free to the left in continuous time. *J. Appl. Probab.* **30**, 509-517.
- [80] Kijima, M. (1997). *Markov Processes for Stochastic Modeling*. London: Chapman & Hall.
- [81] Kijima, M. and Makimoto, N. (1998). Quasi-stationary distributions of Markov chains arising from queueing processes: A survey. In J.G. Shantikumar and U. Sumita (Eds.), *Applied Probability and Stochastic Processes*, pp. 277-311. Boston: Kluwer.
- [82] Kijima, M. and Seneta, E. (1991). Some results for quasistationary distributions of birth-death processes. *J. Appl. Probab.* **28**, 503-511.
- [83] Kingman, J.F.C. (1963). The exponential decay of Markov transition probabilities. *Proc. London Math. Soc.* **13**, 337-358.
- [84] Kryscio, R. and Lefèvre, C. (1989). On the extinction of the S-I-S stochastic logistic epidemic. *J. Appl. Probab.* **27**, 685-694.
- [85] Kukhtin, V.V., Kuzmenko, N.V. and Shramko, O.V. (1997). Green light as a possible pressing factor for oceanic phytoplankton near the base of the euphotic zone. *J. Theor. Biol.* **188**, 319-322.
- [86] Kurtz, T.G. (1970). Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Probab.* **7**, 49-58.
- [87] Kurtz, T.G. (1971). Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J. Appl. Probab.* **8**, 344-356.
- [88] Kyprianou, E.K. (1972). The quasi-stationary distributions of queues in heavy traffic. *J. Appl. Probab.* **9**, 821-831.

- [89] Lambert, A. (2008). Population dynamics and random genealogies. *Stochastic Models* **24**, suppl. 1, 45-163.
- [90] Li, W. and Cao, J. (1993). Limiting distributions of the residual lifetimes of several repairable systems. *Microelectron. Reliab.* **33**, 1069-1072.
- [91] Li, W. and Cao, J. (1993). The limiting distributions of the residual lifetimes of a Markov repairable system. *Reliab. Eng. System Safety* **41**, 103-105.
- [92] Li, Y and Li, J. (2009). Criteria for Feller transition functions. *J. Math. Anal. Appl.* **359**, 653-665.
- [93] Mandl, P. (1959). On the asymptotic behaviour of probabilities within groups of states of a homogeneous Markov chain. *Časopis Pěst. Mat.* **84**, 140-149 (in Russian).
- [94] Mandl, P. (1960). On the asymptotic behaviour of probabilities within groups of states of a homogeneous Markov process. *Časopis Pěst. Mat.* **85**, 448-455 (in Russian).
- [95] McNeil, D.R. and Schach, S. (1973). Central limit analogues for Markov population processes. *J. Roy. Statist. Soc. Ser. B* **35**, 1-23.
- [96] Méléard, S. and Villemonais, D. (2011). Quasi-stationary distributions and population processes. *Report No. 736*, Centre de Mathématiques appliquées, École Polytechnique CNRS, France.
- [97] Miclo, L. (1999). An example of application of discrete Hardy's inequalities. *Markov Process. Related Fields* **5**, 319-330.
- [98] Moler, J.A., Plo, F. and San Miguel, M. (2000). Minimal quasi-stationary distribution under null R -recurrence. *Test* **9**, 455-470.
- [99] Nair, N.G. and Pollett, P.K. (1993). On the relationship between μ -invariant measures and quasi-stationary distributions for continuous-time Markov chains. *Adv. Appl. Probab.* **25**, 82-102.
- [100] Nåsell, I. (1991). On the quasistationary distribution of the Ross Malaria model. *Math. Biosci.* **107**, 187-208.
- [101] Nåsell, I. (1996). The quasi-stationary distribution of the closed endemic SIS model. *Adv. Appl. Probab.* **28**, 895-932.
- [102] Nåsell, I. (1999). On the quasi-stationary distribution of the stochastic logistic epidemic. *Math. Biosci.* **156**, 21-40.

- [103] Nåsell, I. (2003). Moment closure and the stochastic logistic model. *Theoret. Pop. Biol.* **63**, 159-168.
- [104] Nummelin, E. (1976). Limit theorems for α -recurrent semi-Markov processes. *Adv. Appl. Probab.* **8**, 531-547.
- [105] Oppenheim, I., Schuler, K.K. and Weiss, G.H. (1977). Stochastic theory of nonlinear rate processes with multiple stationary states. *Physica A*, **88**, 191-214.
- [106] Ovaskainen, O. (2001). The quasistationary distribution of the stochastic logistic model. *J. Appl. Probab.* **38**, 898-907.
- [107] Pakes, A.G. (1984). The Galton-Watson process with killing. *Math. Biosci.* **69**, 171-188.
- [108] Pakes, A.G. (1995). Quasi-stationary laws for Markov processes: examples of an always proximate absorbing state. *Adv. Appl. Probab.* **27**, 120-145.
- [109] Papangelou, F. (1967). Strong ratio limits, R -recurrence and mixing properties of discrete parameter Markov processes. *Z. Wahrsch. Verw. Gebiete* **8**, 259-297.
- [110] Pijnenburg, M., Ravichandran, N. and Regterschot, G. (1993). Stochastic analysis of a dependent parallel system. *European J. Oper. Res.* **68**, 90-104.
- [111] Pinsky, R.G. (1985). On the convergence of diffusion processes conditioned to remain in a bounded region for a large time to limiting positive recurrent diffusion processes. *Ann. Probab.* **13**, 363-378.
- [112] Pollett, P.K. (1986). On the equivalence of μ -invariant measures for the minimal process and its q -matrix. *Stochastic Process. Appl.* **22**, 203-221.
- [113] Pollett, P.K. (1988). On the problem of evaluating quasistationary distributions for open reaction schemes. *J. Statist. Phys.* **53**, 1207-1215.
- [114] Pollett, P.K. (1989). Modelling quasi-stationary behaviour in metapopulations. *Math. Computers Simulat.* **48**, 393-405.
- [115] Pollett, P.K. (2001). Diffusion approximations for ecological models. In F. Ghassemi (Ed.), *Proceedings of the International Congress on Modelling and Simulation*, Vol. 2, Canberra, Australia, 2001, pp. 843-848. Modelling and Simulation Society of Australia and New Zealand.

- [116] Pollett, P.K. and Roberts, A.J. (1990). A description of the long-term behaviour of absorbing continuous-time Markov chains using a centre manifold. *Adv. Appl. Probab.* **22**, 111-128.
- [117] Pollett, P.K. and Vere-Jones, D. (1992). A note on evanescent processes. *Austral. J. Statist.* **34**, 531-536.
- [118] Ross, J.V. (2006). A stochastic metapopulation model accounting for habitat dynamics. *J. Math. Biol.* **52**, 788-806.
- [119] Schmidt, E.J.P.G. (1986). An alternative approach to canonical forms of matrices. *Amer. Math. Monthly* **93**, 176-184.
- [120] Seneta, E. (1966). Quasi-stationary behaviour in the random walk with continuous time. *Austral. J. Statist.* **8**, 92-98.
- [121] Seneta, E. (1967). Finite approximations to infinite non-negative matrices I. *Proc. Camb. Phil. Soc.* **63**, 983-992.
- [122] Seneta, E. (1968). Finite approximations to infinite non-negative matrices II: refinements and applications. *Proc. Camb. Phil. Soc.* **64**, 465-470.
- [123] Seneta, E. (1981). *Non-negative Matrices and Markov Chains*, Rev. ed. New York: Springer.
- [124] Seneta, E. and Vere-Jones, D. (1966). On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Probab.* **3**, 403-434.
- [125] Seneta, E. and Vere-Jones, D. (1968). On the asymptotic behaviour of subcritical branching processes with continuous state space. *Z. Wahrsch. Verw. Gebiete* **10**, 212-225.
- [126] Shohat, J.A. and Tamarkin, J.D. (1963). *The Problem of Moments*, Math. Surveys I. (Rev. ed.). Providence, R.I.: American Mathematical Society.
- [127] Steinsaltz, D. and Evans, S.N. (2004). Markov mortality models: implications of quasistationarity and varying initial distributions. *Theoret. Population Biol.* **65**, 319-337.
- [128] Steinsaltz, D. and Evans, S.N. (2007). Quasistationary distributions for one-dimensional diffusions with killing. *Trans. Amer. Math. Soc.* **359**, 1285-1324.
- [129] Stirk, E.R., Lythe, G., van den Berg, H.A., Hurst, G.A.D. and Molina-París, C. (2010). The limiting conditional probability distribution in a stochastic model of T cell repertoire maintenance. *Math. Biosci.* **224**, 74-86.

- [130] Tweedie, R.L. (1971). Truncation procedures for non-negative matrices. *J. Appl. Probab.* **8**, 311-320.
- [131] Tweedie, R.L. (1974). Some ergodic properties of the Feller minimal process. *Quart. J. Math. Oxford* **25**, 485-495.
- [132] Tweedie, R.L. (1974). Quasi-stationary distributions for Markov chains on a general state-space. *J. Appl. Probab.* **11**, 726-741.
- [133] Tweedie, R.L. (1998). Truncation approximations of invariant measures for Markov chains. *J. Appl. Probab.* **35**, 517-536.
- [134] Vere-Jones, D. (1962). Geometric ergodicity in denumerable Markov chains. *Quart. J. Math. Oxford* **13**, 7-28.
- [135] Vere-Jones, D. (1967). Ergodic properties of nonnegative matrices, I. *Pacific J. Math.* **22**, 361-386.
- [136] Vere-Jones, D. (1969). Some limit theorems for evanescent processes. *Austral. J. Statist.* **11**, 67-78.
- [137] Waugh, W.A.O'N. (1958). Conditioned Markov processes. *Biometrika* **45**, 241-249.
- [138] Weiss, G.H. and Dishon, M. (1971). On the asymptotic behaviour of the stochastic and deterministic models of an epidemic. *Math. Biosci.* **11**, 261-265.
- [139] Wright, S. (1931). Evolution in Mendelian populations. *Genetics* **16**, 97-159.
- [140] Yaglom, A.M. (1947). Certain limit theorems of the theory of branching processes. *Dokl. Acad. Nauk SSSR* **56**, 795-798 (in Russian).
- [141] Ziedins, I.B. (1987). Quasi-stationary distributions and one-dimensional circuit-switched networks. *J. Appl. Probab.* **24**, 965-977.