

## QUASISTATIONARITY OF CONTINUOUS-TIME MARKOV CHAINS WITH POSITIVE DRIFT

PAULINE COOLEN-SCHRIJNER<sup>1</sup>, ANDREW HART<sup>2</sup> and PHIL POLLETT<sup>3</sup>

(Received 27 March 1998)

### Abstract

We shall study continuous-time Markov chains on the nonnegative integers which are both irreducible and transient, and which exhibit discernible stationarity before drift to infinity “sets in”. We will show how this ‘quasi’ stationary behaviour can be modelled using a limiting conditional distribution: specifically, the limiting state probabilities conditional on not having left 0 for the last time. By way of a *dual* chain, obtained by killing the original process on last exit from 0, we invoke the theory of quasistationarity for *absorbing* Markov chains. We prove that the conditioned state probabilities of the original chain are equal to the state probabilities of its dual conditioned on non-absorption, thus allowing us to establish the simultaneous existence and then equivalence, of their limiting conditional distributions. Although a limiting conditional distribution for the dual chain is always a quasistationary distribution in the usual sense, a similar statement is not possible for the original chain.

### 1. Quasistationarity and limiting conditional distributions

The phenomenon which we shall study is depicted in Figure 1. This illustrates a simulation of a population which suffers an explosion after an initial period of stationarity (a detailed description of the model will be given in Section 7). We will introduce the notion of a limiting conditional distribution to describe the apparent *quasistationary* behaviour of this and other more general transient Markov chains.

Let  $(X_t, t \geq 0)$  be a continuous-time Markov chain over a denumerable state space which, for simplicity, we shall take as  $S = \{0, 1, \dots\}$ . Let  $Q = (q_{ij}, i, j \in S)$  be the  $q$ -matrix of transition rates, assumed to be stable and conservative, so that  $q_{ij}$

<sup>1</sup>Department of Mathematical Sciences, University of Durham, Science Laboratories, South Road, Durham DH1 3LE, UK.

<sup>2</sup>School of Mathematical Sciences, Queensland University of Technology, GPO Box 2434, Brisbane, Qld 4001, Australia.

<sup>3</sup>Department of Mathematics, The University of Queensland, Qld 4072, Australia.

© Australian Mathematical Society 2000, Serial-fee code 0334-2700/00

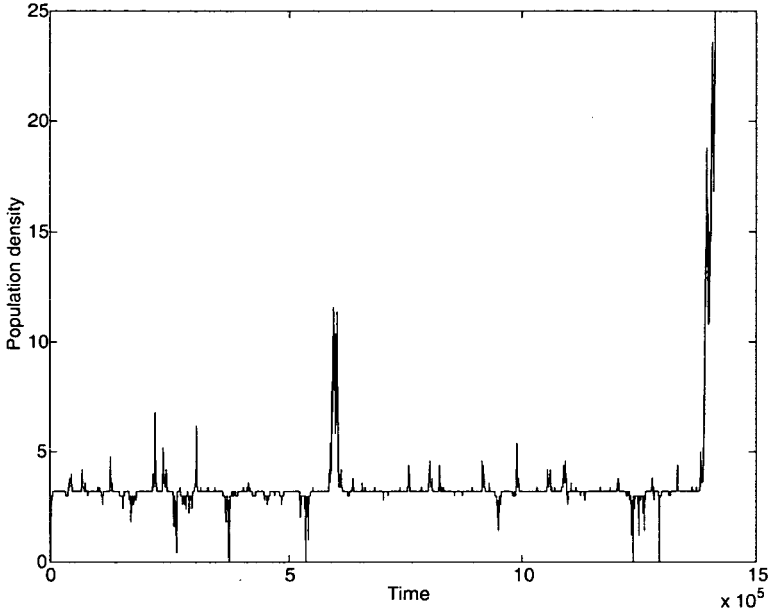


FIGURE 1. Simulation of a population process.

( $\geq 0$ ), for  $j \neq i$ , represents the transition rate from state  $i$  to state  $j$  and  $q_{ii} = -q_i$ , where  $q_i = \sum_{j \neq i} q_{ij}$  ( $< \infty$ ) represents the transition rate out of state  $i$ . Let  $P(t) = (p_{ij}(t), i, j \in S)$ , where

$$p_{ij}(t) = \Pr(X_t = j | X_0 = i),$$

and suppose that  $P$  is a  $Q$ -transition function, that is,  $p_{ij}'(0+) = q_{ij}$ . It will not be necessary to assume that the transition function  $P$  is determined uniquely by  $Q$ , but we will assume that  $P$  is honest ( $\sum_{j \in S} p_{ij}(t) = 1$ ). We further assume *only* that  $P$  is *irreducible*, that is,  $p_{ij}(t) > 0$  for all  $i, j \in S$  and *transient*, that is,

$$\int_0^\infty p_{ij}(t) dt < \infty, \quad i, j \in S.$$

Under this latter condition, the process will visit each state, in particular state 0, at most a finite number of times.

We shall examine the limiting behaviour conditional on recurrence to state 0:

DEFINITION 1. A probability distribution  $r = (r_j, j \in S)$  is called a *limiting conditional distribution* for the process if

$$\lim_{t \rightarrow \infty} \Pr(X_t = j | X_0 = i, X_{t+s} = 0 \text{ for some } s > 0) = r_j, \quad j \in S. \quad (1)$$

This is the definition introduced by van Doorn and Schrijner [23], who used it to study quasistationarity of random walks and other discrete-time Markov chains which are skip-free to the left. Their results were extended to cover birth-death processes in [10] and, more recently, general discrete-time Markov chains in [3].

Since the conditional probability can be evaluated as

$$\Pr(X_t = j | X_0 = i, X_{t+s} = 0 \text{ for some } s > 0) = \frac{p_{ij}(t)b_j}{\sum_{k \in S} p_{ik}(t)b_k}, \tag{2}$$

where  $b_i = \Pr(X_t = 0 \text{ for some } t > 0 | X_0 = i)$ , it should in principle be possible to determine the limit (1). To this end, we shall define a dual chain in terms of the hitting probabilities ( $b_i, i \in S$ ).

### 2. The dual chain

First we note that  $b_0 = 1$  and (by irreducibility and transience) that  $0 < b_i < 1, i \in S$ . Next, let  $b_i(t) = \Pr(X_s = 0 \text{ for some } s > t | X_0 = i)$  and observe that  $b_i(0) = b_i$ . By Lemma 4.1 of [10] we have that

$$b_i(t) = \frac{\int_t^\infty p_{i0}(u)du}{\int_0^\infty p_{00}(u)du}, \quad i \in S, t \geq 0.$$

(Although [10] deals only with birth-death processes, the proof of Lemma 4.1 holds for general Markov chains.) By conditioning on  $X_t$  we see that  $b_i(t)$  can also be evaluated as

$$b_i(t) = \sum_{j \in S} p_{ij}(t)b_j, \quad i \in S, t > 0.$$

It follows immediately that  $b = (b_i, i \in S)$  is a *subinvariant vector* for  $P$ : in particular,

$$\sum_{j \in S} p_{ij}(t)b_j = b_i - \frac{1}{u_0} \int_0^t p_{i0}(u)du \quad (< b_i \text{ for } t > 0), \tag{3}$$

where  $u_0 = \int_0^\infty p_{00}(u)du (< \infty)$ . Hence,  $b$  is also *subinvariant* for  $Q$  in that  $\sum_{j \in S} q_{ij}b_j \leq 0, i \in S$  (see for example [16, Theorem 3.1]). We can be more precise:

LEMMA 1.

$$\sum_{j \in S} q_{ij}b_j = \begin{cases} 0, & \text{if } i \geq 1, \\ -1/u_0, & \text{if } i = 0. \end{cases}$$

PROOF. Let  $u_i = \int_0^\infty p_{i0}(u)du$ , so that  $b_i = u_i/u_0$ . Since  $Q$  is conservative,  $P$  satisfies the *backward equations*

$$p_{ij}'(t) = \sum_{k \in S} q_{ik}p_{kj}(t), \quad i, j \in S, t > 0, \tag{4}$$

and hence, by [1, Proposition 2.1.1], it also satisfies the *backward integral equations*; in particular,

$$p_{i0}(t) = \delta_{i0}e^{-q_i t} + \int_0^t e^{-q_i s} \sum_{k \in S: k \neq i} q_{ik}p_{k0}(t-s) ds, \quad i \in S, t > 0.$$

Integrating this from 0 to  $\infty$ , we get

$$\begin{aligned} u_i &= \int_0^\infty \delta_{i0}e^{-q_i t} dt + \int_0^\infty \int_0^t e^{-q_i s} \sum_{k \in S: k \neq i} q_{ik}p_{k0}(t-s) ds dt \\ &= \frac{\delta_{i0}}{q_i} + \int_0^\infty e^{-q_i s} \sum_{k \in S: k \neq i} q_{ik} \int_0^\infty p_{k0}(u) du ds = \frac{\delta_{i0}}{q_i} + \frac{1}{q_i} \sum_{k \in S: k \neq i} q_{ik}u_k. \end{aligned}$$

Rearranging this gives the result.

Now define  $\bar{P}(t) = (\bar{p}_{ij}(t), i, j \in S)$  and  $\bar{Q} = (\bar{q}_{ij}, i, j \in S)$  by

$$\bar{p}_{ij}(t) = \frac{p_{ij}(t)b_j}{b_i}, \quad i, j \in S, \quad \text{and} \quad \bar{q}_{ij} = \frac{q_{ij}b_j}{b_i}, \quad i, j \in S. \tag{5}$$

Clearly  $\bar{Q}$  is a  $q$ -matrix and  $\bar{P}$  is a (standard) transition function; it is a  $\bar{Q}$ -transition function because  $\bar{p}_{ij}'(0+) = \bar{q}_{ij}$ . However, we would like  $\bar{Q}$  to be conservative and  $\bar{P}$  to be honest. They are not;  $\bar{P}$  is dishonest with  $\sum_{j \in S} \bar{p}_{ij}(t) = b_i(t)/b_i (< 1$  for  $t > 0)$  and  $\bar{Q}$  is non-conservative with

$$\sum_{j \in S} \bar{q}_{ij} = \begin{cases} 0, & \text{if } i \geq 1, \\ -1/u_0, & \text{if } i = 0. \end{cases}$$

(Recall that  $b_0 = 1$ .) We can remedy this by appending, to the irreducible class  $S$ , an absorbing state  $\theta$  and extending the definition of  $\bar{P}$  and  $\bar{Q}$  to  $\bar{S} := S \cup \{\theta\}$  as follows:

$$\bar{p}_{\theta j}(t) = \delta_{\theta j}, \quad j \in \bar{S}, \quad \bar{p}_{i\theta}(t) = 1 - \frac{b_i(t)}{b_i}, \quad i \in S, \tag{6}$$

and

$$\bar{q}_{\theta j} = 0, \quad j \in \bar{S}, \quad \bar{q}_{i\theta} = \frac{\delta_{i0}}{u_0}, \quad i \in S. \tag{7}$$

In this way  $\bar{Q}$  will be conservative and  $\bar{P}$  will be honest.  $\bar{P}$  will remain a  $\bar{Q}$ -transition function since, for  $i \in S$ ,

$$\begin{aligned} \lim_{t \downarrow 0} \frac{\bar{p}_{i\theta}(t)}{t} &= \lim_{t \downarrow 0} \frac{1}{t} \left( 1 - \frac{b_i(t)}{b_i} \right) = \frac{1}{b_i u_0} \lim_{t \downarrow 0} \frac{1}{t} \int_0^t p_{i\theta}(u) du \\ &= \frac{1}{b_i u_0} \lim_{t \downarrow 0} p_{i\theta}(t) = \frac{\delta_{i\theta}}{b_i u_0} = \bar{q}_{i\theta}. \end{aligned}$$

In order to evaluate limiting conditional distributions, we shall only draw on analytical properties of the dual transition function. However, we note that the dual has a pleasing probabilistic interpretation. Let  $L$  be the last exit time from state 0, that is,  $L(\omega) = \sup\{t \geq 0 : X_t(\omega) = 0\}$ , and define the dual process  $(\bar{X}_t, t \geq 0)$  by

$$\bar{X}_t(\omega) = \begin{cases} X_t(\omega), & \text{if } t < L(\omega), \\ \theta, & \text{if } t \geq L(\omega); \end{cases}$$

the dual process is thus the original process *killed* at time  $L$ . Since  $L$  is a *co-optional time* with  $\Pr(L > 0 | X_0 = i) = b_i$ , it follows, from [15, Theorem 2.1], that  $\bar{X}_t$  is Markovian with transition function  $\bar{P}$  and, if  $X_0 = i$ , its initial distribution is given by

$$\Pr(\bar{X}_0 = j) = \begin{cases} b_i, & \text{if } j = i, \\ 1 - b_i, & \text{if } j = \theta, \\ 0, & \text{otherwise.} \end{cases}$$

Notice that  $\bar{X}_t$  is absorbed with probability 1:

$$\bar{a}_i := \lim_{t \rightarrow \infty} \bar{p}_{i\theta}(t) = \lim_{t \rightarrow \infty} \left( 1 - \frac{b_i(t)}{b_i} \right) = 1, \quad i \in S.$$

### 3. The existence of limiting conditional distributions

By the definition of  $\bar{P}$ , we see immediately that, for every  $i, j \in S$ ,

$$\frac{p_{ij}(t)b_j}{\sum_{k \in S} p_{ik}(t)b_k} = \frac{\bar{p}_{ij}(t)}{\sum_{k \in S} \bar{p}_{ik}(t)}. \tag{8}$$

Thus, in view of (2), the existence or otherwise of a limiting conditional distribution can be established *via* the dual transition function using the analogous results for absorbing chains.

Here is one such result, which will be familiar to readers conversant with the classical (Vere-Jones) theory of quasistationary distributions (see for example [1, Proposition 5.2.10]):

**THEOREM 1.** *Suppose that  $\bar{P}$  is  $\lambda$ -positive recurrent and let  $\bar{m} = (\bar{m}_j, j \in S)$  be the (essentially unique)  $\lambda$ -invariant measure for  $\bar{P}$ . If  $\bar{m}$  is finite, that is,  $\sum_{i \in S} \bar{m}_i < \infty$ , then*

$$\lim_{t \rightarrow \infty} \frac{\bar{p}_{ij}(t)}{\sum_{k \in S} \bar{p}_{ik}(t)} = \frac{\bar{m}_j}{\sum_{k \in S} \bar{m}_k}, \quad i, j \in S.$$

In order to employ this result in the present context we will need to understand the term “ $\lambda$ -invariant measure” and the classification “ $\lambda$ -positive recurrent”. We will find that the dual and the original chain have the *same* classification and that there is a one-to-one correspondence between their  $\lambda$ -invariant measures. This makes the application of Theorem 1 quite straightforward.

First, there always exists a number  $\lambda \geq 0$ , called the *decay parameter* (of  $S$ ), with the property that, for each pair of states  $i, j \in S$ ,

$$\lim_{t \rightarrow \infty} -\frac{1}{t} \log p_{ij}(t) = \lambda.$$

Indeed, this is true of *any* transition function  $P$  and  $S$  could be any irreducible class. Clearly,  $P$  and  $\bar{P}$  have the *same* decay parameter (recall that  $S$  is an irreducible class for both – it is the whole state space for  $P$ ).  $P$  is then said to be *exponentially ergodic* if  $\lambda > 0$ . This is because  $p_{ij}(t) = O(e^{-\lambda t})$  for each  $i, j \in S$ . Clearly  $P$  is exponentially ergodic if and only if  $\bar{P}$  is exponentially ergodic.

Next,  $P$  is  $\lambda$ -transient or  $\lambda$ -recurrent according as

$$\int_0^\infty e^{\lambda t} p_{ij}(t) dt$$

is finite or infinite for some (and then all) pairs  $i, j \in S$  and, if  $P$  is  $\lambda$ -recurrent, it is  $\lambda$ -positive recurrent or  $\lambda$ -null recurrent according as  $\lim_{t \rightarrow \infty} e^{\lambda t} p_{ij}(t)$  is strictly positive or equal to 0 for some (and then all) pairs  $i, j \in S$ . Clearly, then,  $P$  and  $\bar{P}$  have the same  $\lambda$ -classification.

Finally, a collection of strictly positive numbers  $m = (m_j, j \in S)$  is called a  $\lambda$ -subinvariant measure for  $P$  if, for all  $t > 0$ ,

$$\sum_{i \in S} m_i p_{ij}(t) \leq e^{-\lambda t} m_j, \quad j \in S,$$

and  $\lambda$ -invariant if, for all  $t > 0$ ,

$$\sum_{i \in S} m_i p_{ij}(t) = e^{-\lambda t} m_j, \quad j \in S.$$

Now if, for any given collection  $m = (m_j, j \in S)$ , we define  $\bar{m} = (\bar{m}_j, j \in S)$  by  $\bar{m}_j = m_j b_j$ , then it is easy to prove that  $m$  is  $\lambda$ -(sub)invariant for  $P$  if and only if  $\bar{m}$  is  $\lambda$ -(sub)invariant for  $\bar{P}$ .

Thus, in the context of Theorem 1, we have, in particular, that  $P$  is  $\lambda$ -positive recurrent if and only if  $\bar{P}$  is  $\lambda$ -positive recurrent and, that  $m$  is a  $\lambda$ -invariant measure for  $P$  if and only if  $\bar{m}$  is a  $\lambda$ -invariant measure for  $\bar{P}$ . It is a remarkable fact that the premise  $\sum_{j \in S} \bar{m}_j < \infty$  holds in the present context whenever  $P$  is exponentially ergodic:

LEMMA 2. *Let  $m$  be a  $\lambda$ -subinvariant measure for  $P$ . Then  $\sum_{j \in S} m_j b_j < \infty$  if  $\lambda > 0$ . If  $m$  is  $\lambda$ -invariant, the converse is also true and  $\sum_{i \in S} m_i b_i = m_0 / (u_0 \lambda)$ .*

PROOF. Suppose that  $\lambda > 0$ . Then

$$\begin{aligned} \sum_{i \in S} m_i b_i &= \frac{1}{u_0} \sum_{i \in S} m_i \int_0^\infty p_{i0}(s) ds = \frac{1}{u_0} \int_0^\infty \sum_{i \in S} m_i p_{i0}(s) ds \\ &\leq \frac{1}{u_0} \int_0^\infty m_0 e^{-\lambda s} ds = \frac{m_0}{u_0 \lambda} < \infty. \end{aligned} \tag{9}$$

Now suppose that  $m$  is  $\lambda$ -invariant and that  $\sum_{i \in S} m_i b_i < \infty$ . If we set  $v_i(t) = \int_0^t p_{i0}(s) ds$ , then from (3) we have

$$b_i = \frac{v_i(t)}{u_0} + \sum_{j \in S} p_{ij}(t) b_j, \quad i \in S, t > 0.$$

Now, fix  $t > 0$  and suppose that  $\lambda = 0$ . Then

$$\begin{aligned} \sum_{i \in S} m_i b_i &= \frac{1}{u_0} \sum_{i \in S} m_i v_i(t) + \sum_{i \in S} m_i \sum_{j \in S} p_{ij}(t) b_j \\ &= \frac{1}{u_0} \sum_{i \in S} m_i \int_0^t p_{i0}(s) ds + \sum_{j \in S} \left( \sum_{i \in S} m_i p_{ij}(t) \right) b_j \\ &= \frac{1}{u_0} \int_0^t \sum_{i \in S} m_i p_{i0}(s) ds + \sum_{j \in S} m_j b_j \\ &= \frac{m_0 t}{u_0} + \sum_{j \in S} m_j b_j. \end{aligned}$$

Since  $\sum_{i \in S} m_i b_i < \infty$ , this gives  $m_0 t / u_0 = 0$ , which is a contradiction; hence  $\lambda > 0$ . It also gives equality in (9) and hence  $\sum_{i \in S} m_i b_i = m_0 / (u_0 \lambda)$ . This completes the proof of Lemma 2.

Using Theorem 1, Lemma 2 and both (2) and (8), we arrive at the following result:

**THEOREM 2.** *Suppose that  $\lambda > 0$  and that  $P$  is  $\lambda$ -positive recurrent. Let  $m = (m_j, j \in S)$  be the essentially unique  $\lambda$ -invariant measure for  $P$ . Then the limiting conditional distribution exists: for each  $i \in S$ ,*

$$\lim_{t \rightarrow \infty} \Pr(X_t = j | X_0 = i, X_{t+s} = 0 \text{ for some } s > 0) = \frac{m_j b_j}{\sum_{k \in S} m_k b_k}, \quad j \in S.$$

One drawback of Theorem 2 is that the transition function is seldom at our disposal. However, if we confine our attention to the minimal transition function we can obtain our  $\lambda$ -invariant measure directly from  $Q$ . So, let  $F(t) = (f_{ij}(t), i, j \in S)$  be the minimal  $Q$ -transition function, that is, the minimal solution to the backward equations (4). If  $m = (m_j, j \in S)$  is a  $\lambda$ -invariant measure for  $F$ , then, by [20, Proposition 2],  $m$  is also a  $\lambda$ -invariant measure for  $Q$ ; that is,

$$\sum_{i \in S} m_i q_{ij} = -\lambda m_j, \quad j \in S.$$

The converse statement is more delicate. We always have (by Fatou’s Lemma) that  $m$  is  $\lambda$ -subinvariant for  $F$  if  $m$  is  $\lambda$ -invariant for  $Q$ , but under what conditions is  $m$  a  $\lambda$ -invariant measure for  $F$ ? Certainly  $F$  being  $\lambda$ -recurrent is a sufficient condition (again see [20, Proposition 2]). Necessary and sufficient conditions are given by Pollett [17, 18] in terms of a “reverse”  $q$ -matrix, but these are usually difficult to check. Much simpler conditions hold for absorbing chains under the “natural” premise that the  $\lambda$ -invariant measure is finite (see [16] and [19]); the following is not the most general statement available, but it will be enough for our purposes.

**THEOREM 3.** *Suppose that  $\lambda > 0$  and let  $\bar{m} = (\bar{m}_j, j \in S)$  be a  $\lambda$ -invariant measure for  $\bar{Q}$  such that  $\sum_{i \in S} \bar{m}_i < \infty$ . Then  $\bar{m}$  is a  $\lambda$ -invariant measure for  $\bar{F}$  if and only if*

$$\lambda \sum_{i \in S} \bar{m}_i \bar{a}_i = \sum_{i \in S} \bar{m}_i \bar{q}_{i\theta}.$$

Theorem 3 is stated for convenience in terms of  $\bar{F}$  and  $\bar{Q}$ . Of course, it holds for any chain whose state space  $\bar{S}$  consists of an irreducible class  $S$  and an absorbing state  $\theta$  and whose absorption probabilities  $(\bar{a}_i, i \in S)$  are strictly positive. For the particular  $\bar{F}$  and  $\bar{Q}$  under consideration, namely those constructed from  $F$  and  $Q$  by way of (5), (6) and (7), we have already noted that  $\bar{a}_i = 1$  for every  $i \in S$ . It is also easy to prove, using say the backward integral recurrence ([1, expression (2.2.10)]), that  $\bar{F}$  is the minimal  $\bar{Q}$ -transition function if and only if  $F$  is the minimal  $Q$ -transition function. Further, if  $m$  is a  $\lambda$ -invariant measure for  $Q$  then  $\bar{m}$  is a  $\lambda$ -invariant measure for  $\bar{Q}$ , where, as before,  $\bar{m}_j = m_j b_j$  and, since  $m$  will also be  $\lambda$ -subinvariant for  $F$  it follows, from Lemma 2, that  $\sum_{i \in S} m_i b_i < \infty$ . Thus, we have the following result:



**THEOREM 4.** *Suppose that  $\lambda > 0$  and let  $m = (m_j, j \in S)$  be a  $\lambda$ -invariant measure for  $Q$ . Then  $\sum_{i \in S} m_i b_i < \infty$  and  $m$  is a  $\lambda$ -invariant measure for  $F$  if and only if*

$$\sum_{i \in S} m_i b_i = \frac{m_0}{u_0 \lambda},$$

whence if  $F$  is  $\lambda$ -positive the limiting conditional distribution exists: for each  $i \in S$ ,

$$\lim_{t \rightarrow \infty} \Pr(X_t = j | X_0 = i, X_{t+s} = 0 \text{ for some } s > 0) = \frac{m_j b_j}{\sum_{k \in S} m_k b_k}, \quad j \in S. \quad (10)$$

Our programme leading to Theorems 2 and 4 has been one of applying results from the theory of absorbing chains to the dual chain and then expressing the conditions which emerge in terms of the original chain. Further results can be obtained in this manner. For example, under asymptotic remoteness of  $\theta$  for  $\bar{F}$ , that is,  $\bar{f}_{i\theta}(t) \rightarrow 0$  as  $i \rightarrow \infty$ , a finite  $\lambda$ -invariant measure exists for  $\bar{F}$  if (and only if)  $\lambda > 0$ . Hence, under this condition, exponential ergodicity is enough to ensure that there exists a  $\lambda$ -invariant measure  $m$  such that  $\sum_{i \in S} m_i b_i < \infty$ . However, the asymptotic remoteness condition, once restated in terms of the original process, amounts to  $\lim_{i \rightarrow \infty} b_i(t)/b_i = 1$  and so, since this will not usually be easy to check, it is perhaps better to deal directly with  $\bar{F}$  and appeal to say [1, Theorem 1.5.7]. Other results can be obtained by applying Kesten’s conditions [8] to  $\bar{Q}$  and thus directly establish the existence of the limiting conditional distribution (1); again there is no obvious benefit in expressing these conditions in terms of  $Q$ .

The form of the dual  $q$ -matrix allows us to obtain results for special processes. In particular, since  $\bar{q}_{i\theta} > 0$  only if  $i = 0$ , a direct jump to the absorbing state is possible only from state 0. For example, the dual of an irreducible birth-death process is an absorbing birth-death process. In the next section we shall obtain a result on the existence of limiting conditional distributions for an irreducible birth-death process by directly applying van Doorn’s results [22] to its dual.

#### 4. Birth-death processes

Consider an irreducible birth-death process  $(X_t, t \geq 0)$  on  $S = \{0, 1, \dots\}$ . Its  $q$ -matrix is given by

$$q_{ij} = \begin{cases} \lambda_i, & \text{if } j = i + 1, \\ -(\lambda_i + \mu_i), & \text{if } j = i, \\ \mu_i, & \text{if } j = i - 1, \\ 0, & \text{otherwise,} \end{cases}$$

where  $\lambda_i > 0, i \geq 0, \mu_i > 0, i \geq 1$  and  $\mu_0 = 0$ . Define series  $A$  and  $C$  by

$$A = \sum_{i=0}^{\infty} \frac{1}{\lambda_i \pi_i} \quad \text{and} \quad C = \sum_{i=0}^{\infty} \frac{1}{\lambda_i \pi_i} \sum_{j=0}^i \pi_j,$$

where  $\pi_0 = 1$  and, for  $i \geq 1$ ,

$$\pi_i = \prod_{k=1}^i \frac{\lambda_{k-1}}{\mu_k}.$$

Assume that  $C = \infty$  ( $F$  honest) and that  $A < \infty$ . Then  $F$  is the unique  $Q$ -transition function. It is both irreducible and transient and the hitting probabilities are given by

$$b_i = \frac{1}{A} \sum_{k=i}^{\infty} \frac{1}{\lambda_k \pi_k}.$$

The classical theory of birth-death processes involves constructing a sequence of (orthogonal) polynomials  $(\phi_i(\cdot), i \in S)$ , where  $\phi_i : \mathbb{R} \rightarrow \mathbb{R}$ , by  $\phi_0(x) = 1, \lambda_0 \phi_1(x) = \lambda_0 - x$  and, for  $i \geq 1$ ,

$$\lambda_i \phi_{i+1}(x) = (\lambda_i + \mu_i - x) \phi_i(x) - \mu_i \phi_{i-1}(x).$$

If we let  $m_i = \pi_i \phi_i(\lambda), i \in S$ , where  $\lambda$  is the decay parameter of  $S$ , then  $m$  will be the essentially unique  $\lambda$ -invariant measure for  $Q$ .

The dual process is clearly an absorbing birth-death process with  $S$  an irreducible class and  $\theta$  an absorbing state which is accessible from  $S$ ; we get (in an obvious notation):

$$\begin{aligned} \bar{\lambda}_\theta &= 0 & \text{and} & \quad \bar{\lambda}_i = \lambda_i b_{i+1} / b_i, \quad i \geq 0, \\ \bar{\mu}_0 &= \lambda_0(1 - b_1) & \text{and} & \quad \bar{\mu}_i = \mu_i b_{i-1} / b_i, \quad i \geq 1, \\ \bar{\pi}_\theta &= 1 & \text{and} & \quad \bar{\pi}_i = b_i^2 \pi_i, \quad i \geq 0, \\ \bar{\phi}_\theta(x) &= 0 & \text{and} & \quad \bar{\phi}_i(x) = \phi_i(x) / b_i, \quad i \geq 0 \\ & & \text{and} & \quad \bar{m}_i = \bar{\pi}_i \bar{\phi}_i(\lambda) = b_i \pi_i \phi_i(\lambda), \quad i \geq 0. \end{aligned}$$

Since  $\bar{a}_i = 1$ , we may apply Theorem 4.1 of [22] to deduce that if  $\lambda > 0$  then

$$\lim_{t \rightarrow \infty} \frac{\bar{f}_{ij}(t)}{\sum_{k \in S} \bar{f}_{ik}(t)} = \bar{\mu}_0^{-1} \lambda \bar{\pi}_j \bar{\phi}_j(\lambda) = \frac{\lambda b_j \pi_j \phi_j(\lambda)}{\lambda_0(1 - b_1)}, \quad i, j \in S,$$

while if  $\lambda = 0$  then the limit is identically 0. Thus, using (2) and (8), we arrive at the following result for an irreducible, transient birth-death process (cf [10, Theorem 4.4]):

**THEOREM 5.** *If  $\lambda > 0$  the limiting conditional distribution exists: for each  $i \in S$ ,*

$$\lim_{t \rightarrow \infty} \Pr(X_t = j | X_0 = i, X_{t+s} = 0 \text{ for some } s > 0) = \frac{\lambda b_j \pi_j \phi_j(\lambda)}{\lambda_0(1 - b_1)}, \quad j \in S.$$

### 5. Quasistationary distributions

For irreducible Markov chains the existence of a *limiting* distribution corresponds to the existence of a *stationary* distribution, or, equivalently, the chain is positive recurrent if and only if its transition function admits a finite invariant measure (see for example [1, Theorem 1.6 and Proposition 1.7]). There is no direct analogy which characterizes limiting conditional distributions for absorbing chains. However, in cases where the chain is absorbed with probability 1, a limiting conditional distribution is also a *quasistationary distribution* in that the state probabilities at time  $t$ , conditional on non-absorption by  $t$ , do not depend on  $t$ , but the converse is not generally true (see [24]).

On this basis it is tempting to conjecture that if a limiting conditional distribution  $r = (r_j, j \in S)$  exists in the sense of (1), then it is also a quasistationary distribution in the following sense: for all  $t > 0$  and  $j \in S$ ,

$$\Pr(X_t = j | X_{t+s} = 0 \text{ for some } s > 0) = r_j. \tag{11}$$

However, this can never be true. We will establish that there is no initial distribution for which the conditional probability in (11) does not depend on  $t$ . If we denote this conditional probability by  $r_j(t)$  and let  $\nu = (\nu_j, j \in S)$  be the initial distribution of the chain, then, from (2), we have that

$$r_j(t) = \frac{p_j(t)b_j}{\sum_{k \in S} p_k(t)b_k}, \tag{12}$$

where  $p_j(t) = \sum_{i \in S} \nu_i p_{ij}(t)$ . On letting  $t \rightarrow 0$  we find that

$$r_j(0+) = \frac{\nu_j b_j}{\sum_{k \in S} \nu_k b_k}.$$

This can be justified as follows. First, since  $p_{ij}(t) \leq 1$ , we have  $p_j(0+) = \nu_j$  by the dominated convergence theorem. Next, using (3), we observe that the denominator of (12) can be written as

$$\sum_{k \in S} \sum_{i \in S} \nu_i p_{ik}(t)b_k = \sum_{i \in S} \nu_i b_i(t) = \sum_{i \in S} \nu_i b_i - \frac{1}{u_0} \int_0^t p_0(u)du$$

and so its limit near 0 is  $\sum_{i \in S} \nu_i b_i$ .

Therefore, if  $r = (r_j, j \in S)$  is to be a quasistationary distribution in the sense of (11), we must have that

$$\frac{p_j(t)b_j}{\sum_{k \in S} p_k(t)b_k} = \frac{\nu_j b_j}{\sum_{k \in S} \nu_k b_k},$$

and so we may write  $p_j(t) = g(t)v_j$ , where  $0 < g(t) < 1$ . A standard argument (see, for example, [16, proof of Proposition 3.1]) then shows that  $g(t) = e^{-\mu t}$  for some  $\mu > 0$ . We have proved that, for some  $\mu > 0$ ,  $v$  is a  $\mu$ -invariant measure for  $P$ :

$$\sum_{i \in S} v_i p_{ij}(t) = e^{-\mu t} v_j, \quad j \in S.$$

However, since  $P$  is honest it follows, from [16, Lemma 3.1], that when  $\mu > 0$  there can be no finite  $\mu$ -invariant measures for  $P$ . This is a contradiction and we deduce that  $r_j(t)$  must depend on  $t$ . Hence, there can be no quasistationary distribution in the sense of (11).

### 6. Numerical issues

The natural way to evaluate the limiting conditional distribution (10), in situations where it is known to exist, is to evaluate a suitably normalized  $\lambda$ -invariant measure of the  $q$ -matrix restricted to  $S_n = \{0, 1, \dots, n\}$  in the hope that the resulting sequence converges to the limiting conditional distribution as  $n$  gets large. However, one must tread carefully.

Define a truncated  $q$ -matrix  ${}_{(n)}Q$  by  ${}_{(n)}Q = ({}_{(n)}q_{ij}, i, j \in S_n)$ , where  ${}_{(n)}q_{ij} = q_{ij}$  and let  ${}_{(n)}F$  be the minimal  ${}_{(n)}Q$ -transition function. The first thing to notice is that  $S_n$  might not be irreducible for  ${}_{(n)}F$ . However, we can always choose an increasing sequence of finite, irreducible subsets of  $S$  which approximate  $S$  ([2, Lemma 1]). For simplicity, we shall therefore suppose that  $S_n$  is irreducible.

Next, let  $\lambda_n$  be the decay parameter of  $S_n$ . By [14] (see also [4]), all the eigenvalues of  ${}_{(n)}Q$  have negative real parts and  $-\lambda_n$  is the one with maximal real part. Moreover, it has multiplicity 1 and both the corresponding left and right eigenvectors have positive entries. This left eigenvector,  ${}_{(n)}m = ({}_{(n)}m_i, i \in S_n)$ , is of course the unique  $\lambda_n$ -invariant measure for  ${}_{(n)}Q$ :

$$\sum_{i \in S_n} {}_{(n)}m_i {}_{(n)}q_{ij} = -\lambda_n {}_{(n)}m_j, \quad j \in S_n.$$

Thus, provided  $n$  is not too large,  ${}_{(n)}m$  can be evaluated using any of the standard techniques for evaluating eigenvectors, such as inverse iteration (see for example [5]). The difficulty here is that whilst it is always true that  $\lambda_n \downarrow \lambda$  ([2, Lemma 2]), the sequence  $\{ {}_{(n)}m \}$  might not converge, in any sense, to a  $\lambda$ -invariant measure for  $Q$  (see [6, Example 5.1]). Another related issue is the normalization of  ${}_{(n)}m$ . In view of (10), it would be natural to define a sequence  ${}_{(n)}r = ({}_{(n)}r_i, i \in S_n)$  by

$${}_{(n)}r_i = \frac{{}_{(n)}m_i {}_{(n)}b_i}{\sum_{k \in S_n} {}_{(n)}m_k {}_{(n)}b_k},$$

where  $({}_n b_i, i \in S_n)$  are the hitting probabilities of state 0 for the truncated chain, and use this to approximate the limiting conditional distribution. (Note that the act of truncating  $Q$  to  $S_n$  effectively means that states  $n + 1, n + 2, \dots$  are aggregated into a single absorbing state and so  $({}_n b_i$  is the probability that the chain reaches 0 *before* it is absorbed there.) However, evaluating the product  $({}_n m_i ({}_n b_i$  may result in a loss of precision if  $({}_n m_i$  is large and  $({}_n b_i$  is small, or vice versa. This frequently occurs in cases when  $({}_n b_i$  decays quickly as  $i$  increases, in which case  $({}_n r$  may exhibit a mode near  $i = 0$  when none exists.

There are several ways to overcome this problem. Perhaps the most effective is to first evaluate the hitting probabilities  $(b_i, i \in S)$ , then calculate the dual  $q$ -matrix and work with this. Now only *ratios* of the  $b_i$ 's (usually of the form  $b_{j+l}/b_j, l \neq 0$ , with  $|l|$  small) need to be evaluated. In some cases the hitting probabilities can be obtained analytically (see below), but in most cases they must be estimated using a truncation procedure: a procedure which *always* works. First observe that, under the conditions we have imposed, the hitting probabilities form the unique, bounded solution to

$$\sum_{j \in S} q_{ij} b_j = 0, \quad i \neq 0, \quad b_0 = 1,$$

(see [1, page 298]). It is therefore natural to consider the corresponding system of equations for the truncated chain:

$$\sum_{j \in S_n} ({}_n q_{ij} ({}_n b_j = 0, \quad i \neq 0, \quad ({}_n b_0 = 1. \tag{13}$$

Indeed, it is easy to see that the hitting probabilities of the truncated chain form the unique solution to (13). More importantly, it can be shown that, for each  $i \in S, ({}_n b_i \uparrow b_i$  as  $n \rightarrow \infty$  ([6, Theorem 5.6]) and thus we have an effective means of estimating  $(b_i, i \in S)$ .

Once  $\bar{Q}$  is determined, we can approximate the limiting conditional distribution using truncations of  $\bar{Q}$ . Define  $({}_n \bar{Q}$  by  $({}_n \bar{Q} = ({}_n \bar{q}_{ij}, i, j \in S_n)$ , where  $({}_n \bar{q}_{ij} = \bar{q}_{ij}$  and let  $({}_n \bar{m} = ({}_n \bar{m}_i, i \in S_n)$  be the unique  $\lambda_n$ -invariant measure for  $({}_n \bar{Q}$ :

$$\sum_{i \in S_n} ({}_n \bar{m}_i ({}_n \bar{q}_{ij} = -\lambda_n ({}_n \bar{m}_j, \quad j \in S_n.$$

Note that  $\lambda_n$  is also the decay parameter of  $S_n$  for the truncated dual chain; this is because  $b_i ({}_n \bar{q}_{ij} = ({}_n q_{ij} b_j, i, j \in S_n$  and hence (in an obvious notation for the minimal transition functions)  $b_i ({}_n \bar{f}_{ij}(t) = ({}_n f_{ij}(t) b_j$ , using standard methods (see [18, proof of Lemma 3.3]). Note also that the minimal  $({}_n \bar{Q}$ -chain has two absorbing states: the absorbing state  $\theta$ , inherited from  $\bar{Q}$  and the one resulting from truncation, this being the aggregation of states  $n + 1, n + 2, \dots$ . We may now redefine the

sequence  ${}^{(n)}r = ({}^{(n)}r_i, i \in S_n)$  by

$${}^{(n)}r_i = \frac{{}^{(n)}\bar{m}_i}{\sum_{k \in S_n} {}^{(n)}\bar{m}_k},$$

and use this to approximate the limiting conditional distribution. However, although there is a growing collection of results, covering a variety of circumstances, which deal with *invariant measures* (see [7, 21]), there are presently few results on truncation procedures for estimating  $\lambda$ -invariant measures. One important exception is the class of Markov chains which are skip-free to the left; these include birth-death processes and branching processes. For these chains it is possible to show, using the methods of Kijima [9], that  ${}^{(n)}r$  converges weakly to the normalized  $\lambda$ -invariant measure for  $\bar{Q}$ , this being the limiting conditional distribution (10) (cf [6, Theorem 5.7]).

### 7. A density-dependent population model

To illustrate our results we shall use the following population model. Let  $f : [0, \infty) \rightarrow [0, \infty)$ ,  $V > 0$  and  $0 < \alpha < 1$  and consider the birth-death process over  $S = \{0, 1, \dots\}$  obtained by setting  $\lambda_0 = q_0$  and, for  $n \geq 1$ ,  $\lambda_n = \alpha q_n$  and  $\mu_n = (1 - \alpha)q_n$ , where  $q_n = Vf(n/V)$ . We will assume that  $f(n/V) > 0$  for all  $n \geq 0$ , so that the process is irreducible and that  $\alpha > 1/2$ , a condition which guarantees that the process is transient. Since the jump chain is a simple random walk with positive drift, the hitting probabilities are given by  $b_i = \rho^i, i \geq 0$ , where  $\rho = 1/\alpha - 1$ .

We have defined a *density-dependent* Markov chain (see for example Kurtz [11–13]) with constant jump probabilities. If  $X_t$  is the number in the population at time  $t$ , we can think of  $V$  as the area of the habitat and hence  $X_t/V$  as being the population density. It can be shown, using the methods of Kurtz [11], that as  $V$  becomes large the *density process*  $(x_t^{(V)}, t \geq 0)$ , where  $x_t^{(V)} = X_t/V$ , becomes more “deterministic”: if  $x_0^{(V)} \rightarrow x_0$ , then  $x_t^{(V)}$  converges (uniformly in probability over finite time intervals) to the unique solution  $(x(t), t \geq 0)$  of the differential equation

$$\frac{dx}{dt} = F(x), \quad x(0) = x_0, \tag{14}$$

where  $F(x) = \alpha f(x) - (1 - \alpha)f(x) = (2\alpha - 1)f(x)$ . For example, if we choose  $f$  to be

$$f(x) = \frac{(\pi - x)^2}{(\pi + x)^3},$$

then (14) has a semi-stable equilibrium point at  $\pi$  (see Figure 2). Trajectories of (14) with  $x(0) \leq \pi$  approach  $\pi$  as  $t \rightarrow \infty$ ; otherwise they diverge. We chose an irrational

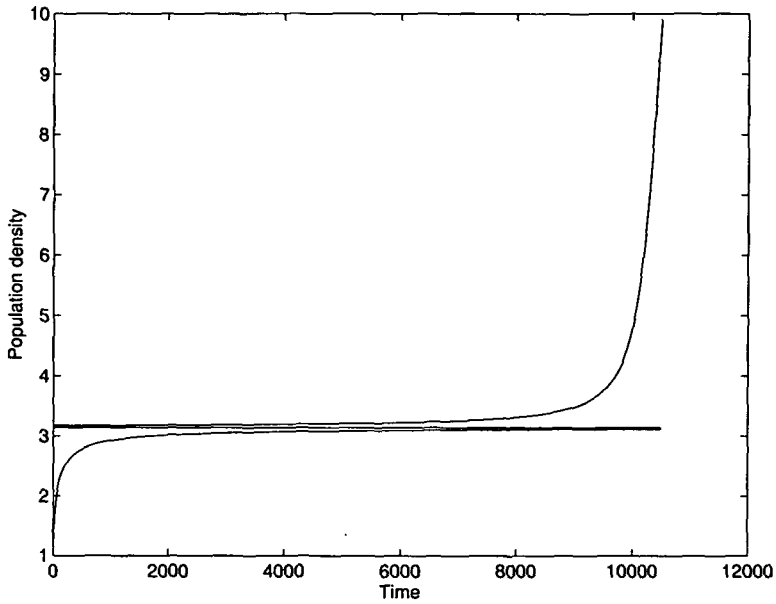


FIGURE 2. Solutions to (14) with  $f(x) = (\pi - x)^2/(\pi + x)^3$  and with boundary conditions  $x(0) < \pi$ ,  $x(0) = \pi$  and  $x(0) > \pi$  ( $\alpha = 0.501$ ).

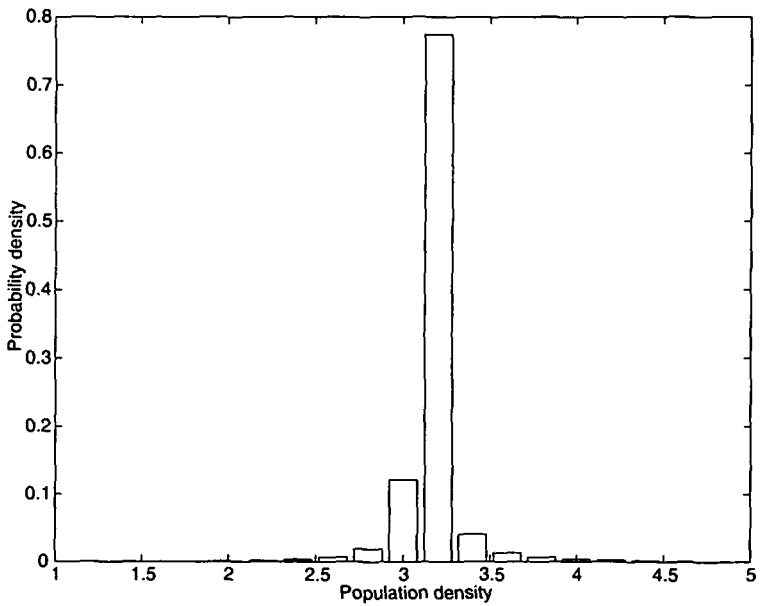


FIGURE 3. Limiting conditional distribution for the density process with  $f(x) = (\pi - x)^2/(\pi + x)^3$ ,  $V = 5$  and  $\alpha = 0.501$ .

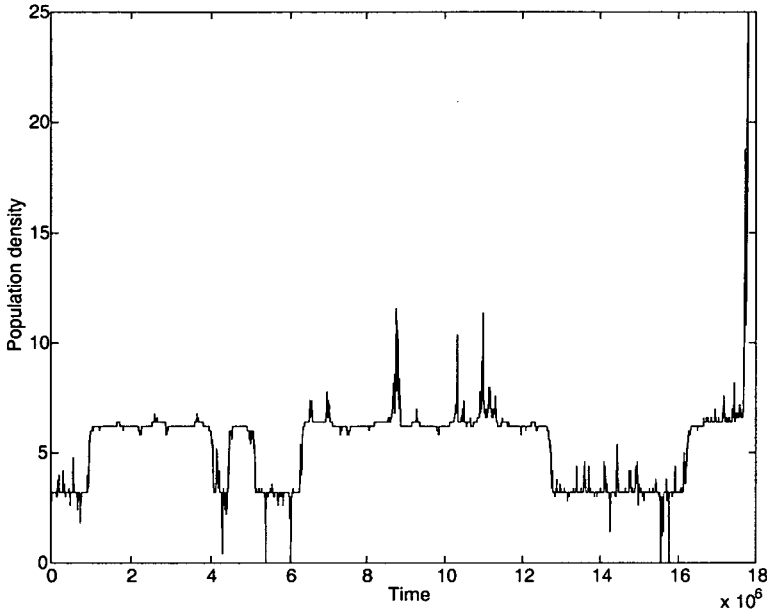


FIGURE 4. Simulation of the density process with  $f(x) = (\pi - x)^2(2\pi - x)^2/(\pi + x)^5$ ,  $V = 5$  and  $\alpha = 0.501$ .

number  $\pi$  as our equilibrium point because then  $q_n > 0$ , no matter what the value of  $V$  and, in particular, so the equilibrium point would not be an absorbing state for the density process. We also chose  $f$  to satisfy  $f(x) \rightarrow 0$  as  $x \rightarrow \infty$  in order that  $q_n$  would be small when  $n$  is large.

Figure 1 (above) shows a sample path of the density process with  $V = 5$  and  $\alpha = 0.501$ , starting at 0 and lasting for 7000 events. Notice that the density process appears to fluctuate, in turn, about each of the three deterministic trajectories illustrated in Figure 2, ‘switching’ from one to another when in close proximity. The limiting conditional distribution is useful in modelling these fluctuations. Figure 3 shows the (approximate) limiting conditional distribution of the density process evaluated using the same parameter values  $V = 5$  and  $\alpha = 0.501$ .

The limiting conditional distribution was estimated using the methods described in Section 6. The dual  $q$ -matrix was constructed using the hitting probabilities  $b_i = \rho^i$ ,  $i \geq 0$  and then a truncation procedure was applied. Since the dual is a birth-death process,  $({}_n)r$  is guaranteed to converge weakly to the limiting conditional distribution as  $n$  gets large. The particular distribution shown in Figure 3 was evaluated using a truncation size of  $n = 30$  (corresponding to a maximum population density of  $n/V = 6$ ); it was plotted only for values of the density between 1 and 5.

By varying  $f$ , our population model can incorporate a range of behaviours. For



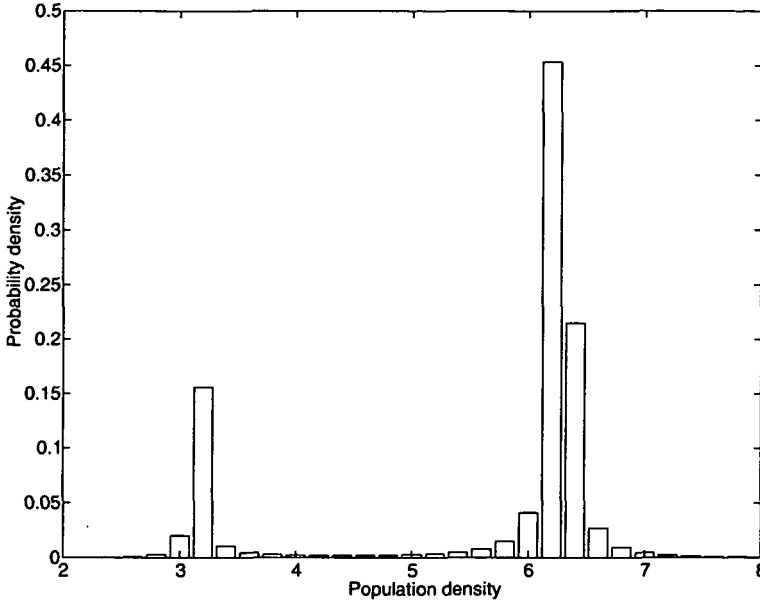


FIGURE 5. Limiting conditional distribution for the density process with  $f(x) = (\pi - x)^2(2\pi - x)^2/(\pi + x)^5$ ,  $V = 5$  and  $\alpha = 0.501$ .

example, if we had chosen  $f$  to be

$$f(x) = \frac{(\pi - x)^2(2\pi - x)^2}{(\pi + x)^5},$$

so that, as before,  $f(n/V) > 0$  for all  $n \geq 0$  and  $f(x) \rightarrow 0$  as  $x \rightarrow \infty$ , we would get *two* equilibria for the limiting deterministic model, one at  $\pi$  and the other at  $2\pi$ . A sample path of the density process for this model is illustrated in Figure 4 using the same parameter values as used previously.

Notice that the process fluctuates between the deterministic equilibria before eventually drifting to  $\infty$ . The approximate limiting conditional distribution illustrated in Figure 5 was evaluated using a truncation size of  $n = 50$  and plotted for values of the density between 2 and 8.

### Acknowledgements

We are grateful to Laird Breyer for helpful conversations on this work and, in particular, for drawing our attention to the paper by Meyer, Smithe and Walsh [15]. The research of P. Coolen-Schrijner was supported by TMR grant ERBFMBICT950067 of the European Community and was partly carried out while visiting the University

of Queensland. The research of A. Hart and P. Pollett was supported by the Australian Research Council (Grant No. A69130032).

## References

- [1] W. J. Anderson, *Continuous-Time Markov Chains: An Applications-Oriented Approach* (Springer, New York, 1991).
- [2] L. A. Breyer and A. G. Hart, "Approximations of quasistationary distributions for Markov chains", *Math. Computer Modelling* **25** (1999) to appear.
- [3] P. Coolen-Schrijner and P. K. Pollett, "Quasi-stationarity of discrete-time Markov chains with drift to infinity", *Methodol. Comput. Appl. Probab.* **1** (1999) 81–96.
- [4] J. N. Darroch and E. Seneta, "On quasi-stationary distributions in absorbing continuous-time finite Markov chains", *J. Appl. Probab.* **4** (1967) 192–196.
- [5] H. G. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed. (Johns Hopkins University Press, Baltimore, 1993).
- [6] A. G. Hart, "Quasistationary distributions for continuous-time Markov chains", Ph. D. Thesis, Department of Mathematics, The University of Queensland, 1997.
- [7] A. G. Hart and R. L. Tweedie, "Convergence of invariant measures of truncation approximations to Markov processes" (1998) submitted for publication.
- [8] H. Kesten, "A ratio limit theorem for (sub) Markov chains on  $\{1, 2, 3, \dots\}$  with bounded jumps", *Adv. Appl. Probab.* **27** (1995) 652–691.
- [9] M. Kijima, "Quasi-limiting distributions of Markov chains that are skip-free to the left in continuous-time", *J. Appl. Probab.* **30** (1993) 509–517.
- [10] M. Kijima, M. G. Nair, P. K. Pollett and E. van Doorn, "Limiting conditional distributions for birth-death processes", *Adv. Appl. Probab.* **29** (1997) 185–204.
- [11] T. G. Kurtz, "Solutions of ordinary differential equations as limits of pure jump Markov processes", *J. Appl. Probab.* **7** (1970) 49–58.
- [12] T. G. Kurtz, "Limit theorems for sequences of jump Markov processes approximating ordinary differential processes", *J. Appl. Probab.* **8** (1971) 344–356.
- [13] T. G. Kurtz, "Limit theorems and diffusion approximations for density dependent Markov chains", *Math. Prog. Study* **5** (1976) 67–78.
- [14] P. Mandl, "On the asymptotic behaviour of probabilities within groups of states of a homogeneous Markov processes" (in Czech), *Časopis Pěst. Mat.* **85** (1960) 448–456.
- [15] P. A. Meyer, R. T. Smythe and J. B. Walsh, "Birth and death of Markov processes", in *Proc. 6th Berkeley Symp. Math. Statist. Probab.* III, (1971), 295–305.
- [16] M. G. Nair and P. K. Pollett, "On the relationship between  $\mu$ -invariant measures and quasistationary distributions for continuous-time Markov chains", *Adv. Appl. Probab.* **25** (1993) 82–102.
- [17] P. K. Pollett, "On the equivalence of  $\mu$ -invariant measures for the minimal process and its  $q$ -matrix", *Stochastic Process. Appl.* **22** (1986) 203–221.
- [18] P. K. Pollett, "Reversibility, invariance and  $\mu$ -invariance", *Adv. Appl. Probab.* **20** (1988) 600–621.
- [19] P. K. Pollett and D. Vere-Jones, "A note on evanescent processes", *Austral. J. Statist.* **34** (1992) 531–536.
- [20] R. L. Tweedie, "Some ergodic properties of the Feller minimal process", *Quart. J. Math. Oxford* **25** (1974) 485–495.
- [21] R. L. Tweedie, "Truncation approximations of invariant measures for Markov chains", *J. Appl. Probab.* **35** (1999) 517–536.

- [22] E. A. van Doorn, "Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes", *Adv. Appl. Probab.* **23** (1991) 683–700.
- [23] E. A. van Doorn and P. Schrijner, "Limit theorems for discrete-time Markov chains on the nonnegative integers conditioned on recurrence to zero", *Stochastic Models* **14** (1996) 77–102.
- [24] D. Vere-Jones, "Some limit theorems for evanescent processes", *Austral. J. Statist.* **11** (1969) 67–78.