

COMPUTABLE STRONGLY ERGODIC RATES OF CONVERGENCE FOR CONTINUOUS-TIME MARKOV CHAINS

YUANYUAN LIU^{✉1}, HANJUN ZHANG² and YIQIANG ZHAO³

(Received 5 November, 2007; revised 15 April, 2008)

Abstract

In this paper, we investigate computable lower bounds for the best strongly ergodic rate of convergence of the transient probability distribution to the stationary distribution for stochastically monotone continuous-time Markov chains and reversible continuous-time Markov chains, using a drift function and the expectation of the first hitting time on some state. We apply these results to birth–death processes, branching processes and population processes.

2000 *Mathematics subject classification*: 60J27, 60J80.

Keywords and phrases: strong ergodicity, convergence rate, birth and death process, branching process, population process.

1. Introduction

As a type of traditional ergodicity, strong ergodicity has been studied extensively for many years. The earliest example may be dated back to as early as the 1930s [8] for discrete-time Markov chains. Later, the work was extended to continuous-time Markov processes, and criteria involving drift functions and first hitting time moments were obtained; for a detailed description of these results, see, for example, [1, 9] and [12]. For strongly ergodic discrete-time Markov chains, the best (largest) strongly ergodic rate of convergence was obtained in [11] by using spectral conditions. There are also other studies on the estimate of the strongly ergodic rate for continuous-time Markov processes, for example, [7] for finite Markov chains by Nash inequality and [16, 17] for Markov processes by the coupling method.

¹School of Mathematics, Railway Campus, Central South University, Changsha, Hunan, 410075, PR China; e-mail: liuyy@csu.edu.cn.

²School of Mathematics and Computational Science, Xiangtan University, Xiangtan, 411105, PR China; e-mail: hjzhang001@gmail.com.

³School of Mathematics and Statistics, Carleton University, Ottawa, Ontario, K1S 5B6, Canada; e-mail: zhao@math.carleton.ca.

© Australian Mathematical Society 2008, Serial-fee code 0334-2700/08

Throughout this paper, we denote by \mathbb{R} (\mathbb{R}_+) the real (nonnegative) number set, \mathbb{Z}_+ the nonnegative integer set and \mathbb{N} the positive integer set. Let Φ_t be a right continuous and time-homogeneous Markov process on a Polish space E , endowed with the Borel field $\mathcal{B}(E)$. Denote by $P^t(x, \cdot)$ the transition function of Φ_t . Suppose that the process Φ_t is ergodic, that is, there exists the unique probability distribution π such that

$$\|P^t(x, \cdot) - \pi\| \rightarrow 0, \quad t \rightarrow \infty,$$

where $\|\mu\|$ denotes the usual total variation norm for a signed measure μ , that is, $\|\mu\| = \sup_{|g| \leq 1} |\mu(g)|$.

The process Φ_t is called strongly ergodic if there exists constants $\gamma \geq 0$ and $\lambda \geq 0$ such that

$$\sup_{x \in E} \|P^t(x, \cdot) - \pi\| \leq \gamma e^{-\lambda t} \tag{1.1}$$

for all $t \geq 0$. For an infinite E , without loss of generality, we can require $\gamma \geq 2$ since $\sup_{x \in E} \|P^t(x, \cdot) - \pi\| = 2$ when $t = 0$. Following Mao [17], define

$$\alpha(\gamma) = \sup \left\{ \lambda > 0 \mid \sup_{x \in E} \|P^t(x, \cdot) - \pi\| \leq \gamma e^{-\lambda t}, \forall t \geq 0 \right\} \tag{1.2}$$

and $\alpha = \alpha(\infty) = \lim_{\gamma \rightarrow \infty} \alpha(\gamma)$ to be the largest (best) strongly ergodic rate of convergence.

REMARK 1.1. It follows from the definition that $\alpha(\gamma)$ is increasing in γ , and reaches its maximal value at a finite value of γ .

For any set $A \in \mathcal{B}(X)$, let τ_A be the first hitting time on A , that is,

$$\tau_A = \inf\{t \geq 0 \mid X_t \in A\}.$$

We write $\tau_x = \tau_{\{x\}}$. In [17], a lower bound estimate that $\alpha \geq 1/(\sup_{x \in E} E_x[\tau_0])$ was obtained for both one-dimensional diffusion processes and birth–death processes by constructing reflection coupling and classical coupling, respectively. Moreover, an upper bound for a general Markov process was also given in [17] as follows.

PROPOSITION 1.2. *If Φ_t is strongly ergodic, then for any closed $A \subset E$ with $\pi(A) > 0$,*

$$\alpha(\gamma) \leq \left[\frac{2}{\pi(A)} \log \frac{\gamma}{\pi(A)} \right] \frac{1}{\sup_{x \in E} E_x[\tau_A]}.$$

REMARK 1.3. However, we note that the upper bound depends on γ and becomes unbounded as γ goes to infinity. This will not be very helpful in bounding the best strongly ergodic rate.

In this paper, we extend the work in [17] and previous literature results to two classes of important continuous-time Markov chains: stochastically monotone

continuous-time Markov chains (for example, [14]) and reversible continuous-time Markov chains (for example, [4, 5]), by investigating lower bounds on α in terms of drift functions and first hitting time moments.

For the remainder of this paper, let Φ_t be a continuous-time Markov chain on a countable state space E . Denote by $Q = (q_{ij})$ the q -matrix (transition rate matrix) and $P^t(i, j)$ the transition function. We assume that Q is totally stable, irreducible and regular. We write $q_i = -q_{ii} = \sum_{k \neq i} q_{ik}$ for any $i \in E$.

2. Stochastically monotone Markov chains

In the section, we refine computable lower bounds for strongly ergodic rates of convergence for stochastically monotone continuous-time Markov chains, using simple drift functions.

We first provide two equivalent drift conditions that will be used frequently later. (DC1) is in general favorable over (DC2) since it does not involve the parameter λ . Let k_0 be an arbitrary state in E .

(DC1) *Drift Condition 1.* There exists a bounded nonnegative solution y of the system

$$\begin{cases} \sum_{j \in E} q_{ij} y_j \leq -1, & i \neq k_0, \\ y_i = 0, & i = k_0. \end{cases} \tag{2.1}$$

(DC2) *Drift Condition 2.* For some $\lambda > 0$, there exists a bounded solution $V \geq 1$ of the system

$$\begin{cases} \sum_{j \in E} q_{ij} V_j \leq -\lambda V_i, & i \neq k_0, \\ V_i = 1, & i = k_0. \end{cases} \tag{2.2}$$

The equivalence is obvious: if (DC1) holds, then (DC2) holds for $V = y + 1$ and any $\lambda \leq 1/(\sup_{i \in E} y_i + 1)$; conversely, if (DC2) holds, then (DC1) holds for $y = \lambda(V - 1)$. Note that (DC1) is a well-known sufficient and necessary condition for strong ergodicity (see, for example, [1, Proposition 3.3 in Section 6.3]). We also note that if (DC2) holds, then $\lambda < q_i$ is satisfied automatically for any $i \neq k_0$ because $0 < \sum_{j \neq i} q_{ij} V_j \leq (q_i - \lambda)V_i$.

The following proposition is essentially the same as [19, Theorem 1] (see also [1, Lemma 1.5 in Section 6.1]) with some minor variations.

PROPOSITION 2.1.

- (i) *The sequence $\{E_i[\tau_{k_0}], i \in E\}$ is the minimal nonnegative solution of (2.1) and satisfies the system with equality.*
- (ii) *For some λ such that $0 < \lambda < q_i$ for any $i \neq k_0$, the sequence $\{E_i[e^{\lambda \tau_{k_0}}], i \in E\}$ is the minimal nonnegative solution of (2.2) and satisfies the system with equality.*

PROOF. (i) The first assertion is the same as that in [1, Lemma 1.5 in Section 6.1].

(ii) From [1, Lemma 1.5 in Section 6.1], we see that the sequence

$$\left\{ \frac{E_i[e^{\lambda\tau_{k_0}}] - 1}{\lambda}, i \in E \right\}$$

is the minimal nonnegative solution of the following system

$$\begin{cases} \sum_{j \in E} q_{ij}x_j \leq -\lambda x_i - 1, & i \neq k_0, \\ x_i = 0, & i = k_0. \end{cases} \tag{2.3}$$

Let $V_i = \lambda x_i + 1$, then (2.3) becomes (2.2), and (ii) follows. □

We now review the definitions of stochastic monotonicity. The chain Φ_t is said to be stochastically monotone if $\sum_{j \geq k} P^t(i, j)$ is an increasing function of i for any fixed $k \in \mathbb{Z}_+$ and $t \in \mathbb{R}_+$. It follows from [20, Theorem 3.1] that Φ_t is stochastically monotone if and only if Q is (stochastically) monotone, that is,

$$\sum_{j \geq k} q_{ij} \leq \sum_{j \geq k} q_{i+1,j}$$

for all $i, k \in \mathbb{Z}_+$ such that $k \neq i + 1$.

We let $E = \mathbb{Z}_+$ in this and the following section. We apply a result in [14] to derive the following lemma. In [14], a stochastically ordered continuous-time Markov process was considered and explicit exponential rates of convergence were found using the coupling method. Note that stochastic monotonicity is equivalent to stochastic ordering for continuous-time Markov chains.

LEMMA 2.2. *Assume that Q is monotone. If $\sup_{i \in E} E_i[e^{\lambda\tau_0}] < \infty$ for some $\lambda > 0$, then*

$$\sup_{i \in E} \|P^t(x, \cdot) - \pi(\cdot)\| \leq 4 \sup_{i \in E} E_i[e^{\beta\tau_0}]e^{-\beta t} \tag{2.4}$$

for any $0 < \beta \leq \lambda$ and $t \geq 0$. Furthermore, if V satisfies (DC2) for some $\lambda > 0$, then

$$\sup_{i \in E} E_i[e^{\lambda\tau_0}] \leq \sup_{i \in E} V_i < \infty.$$

PROOF. Since Q is monotone, the chain Φ_t is stochastically monotone (equivalently, stochastically ordered). If $\sup_{i \in E} E_i[e^{\lambda\tau_0}] < \infty$, then it follows from [14, Theorem 2.1] that

$$\|P^t(i, \cdot) - \pi(\cdot)\| \leq M_i e^{-\lambda t},$$

where $M_i \leq 2(E_i[e^{\lambda\tau_0}] + E_\pi[e^{\lambda\tau_0}])$. With the fact that

$$E_\pi[e^{\beta\tau_0}] = \sum_{j \in E} E_j[e^{\beta\tau_0}]\pi(j) \leq \sup_{i \in E} E_i[e^{\beta\tau_0}] \leq \sup_{i \in E} E_i[e^{\lambda\tau_0}]$$

for any $\beta \leq \lambda$, we obtain (2.4) immediately. The second assertion follows from Proposition 2.1 directly. □

ASSUMPTION 2.3. There exists a positive integer N satisfying $q_{n0} = 0$ for $n \geq N$.

Many important classes of q -matrices satisfy this assumption, for example, the birth–death q -matrix and the branching q -matrix that is studied in Section 3.

For a monotone matrix, the following lemma reveals the relationship between (DC1) and (DC2) in more detail.

LEMMA 2.4. Assume that Q is monotone. If (DC1) holds for an increasing function y , then (DC2) holds for any $\lambda < 1/(\sup_{i \in E} y_i)$ and the function V defined by

$$V_0 = 1, \quad V_i = cy_i, \quad i \geq 1, \tag{2.5}$$

where

$$c = \max \left\{ \sup_{i \in E} \frac{q_{i0}}{1 - \lambda y_i}, \frac{1}{y_1} \right\}.$$

Furthermore, if Q satisfies Assumption 2.3 and y is finally strictly increasing (that is, there exists some $M \in \mathbb{N}$ such that $y_{i+1} > y_i$ for $i \geq M$), then (DC2) holds also for $\lambda = 1/(\sup_{i \in E} y_i)$.

PROOF. Since Q is monotone, it implies that q_{i0} is decreasing for $i \geq 1$. By Proposition 2.1(i), we see

$$y_1 \geq E_1[\tau_0] \geq \frac{1}{q_1} > 0.$$

With the assumption that y is increasing, we know that

$$c \leq \max \left\{ \frac{q_{10}}{1 - \lambda \sup_{i \in E} y_i}, \frac{1}{y_1} \right\} < \infty$$

for any $\lambda < 1/(\sup_{i \in E} y_i)$.

Let $V_0 = 1$ and $V_i = cy_i, i \geq 1$, then $V_i \geq 1$ for all $i \geq 0$. Now, we check that V satisfies (DC2). For $i \geq 1$, since $c \geq (q_{i0})/(1 - \lambda y_i)$, $q_{i0} - c \leq -\lambda cy_i = -\lambda V_i$. By this inequality, (DC1) implies that

$$\sum_{j \in E} q_{ij} V_j = q_{i0} + \sum_{j \geq 1} q_{ij} cy_j \leq q_{i0} - c \leq -\lambda V_i, \quad i \geq 1.$$

Finally, if Assumption 2.3 holds for Q and y is finally strictly increasing, then

$$c = \max \left\{ \max_{1 \leq i \leq N} \frac{q_{i0}}{1 - \lambda y_i}, \frac{1}{y_1} \right\} < \infty$$

even for $\lambda = 1/(\sup_{i \in E} y_i)$, and therefore the conclusion holds. □

Based on Lemmas 2.2, 2.4 and the drift condition (DC1), we obtain the following result, which will be applied to stochastically monotone Markov chains.

THEOREM 2.5. *Let Q be a monotone q -matrix. If (DC1) holds for an increasing function y , then (1.1) holds for any $\lambda < 1/(\sup_{i \in E} y_i)$ and the constant $\gamma = 4 \sup_{i \in E} V_i$, where V is given by (2.5). Furthermore, if Q satisfies Assumption 2.3 and y is finally strictly increasing, then (1.1) holds also for $\lambda = 1/(\sup_{i \in E} y_i)$.*

COROLLARY 2.6. *Assume that Q is monotone. If $\sup_{i \in E} E_i[\tau_0] < \infty$, then*

$$\alpha \geq \frac{1}{\sup_{i \in E} E_i[\tau_0]}.$$

PROOF. Let $y_i = E_i[\tau_0]$, $i \geq 0$. Since the chain Φ_t is stochastically monotone, $y_i = E_i[\tau_0]$ is increasing of i for $i \geq 0$. By Proposition 2.1(i), we know that y satisfies (DC1) with equality. Hence, the assertion follows from Theorem 2.5. \square

3. Applications: stochastically monotone Markov chains

3.1. Birth–death processes Let Φ_t be a birth–death process with q -matrix Q given by

$$q_{i,i+1} = b_i, \quad i \in \mathbb{Z}_+; \quad q_{i,i-1} = a_i, \quad i \in \mathbb{N}; \quad q_{ij} = 0, \quad |i - j| \geq 2,$$

where $a_i > 0$ for $i \in \mathbb{N}$ and $b_i > 0$ for $i \in \mathbb{Z}_+$. It is well-known that Q is regular if and only if Q is conservative and

$$R_b := \sum_{n=1}^{\infty} \left(\frac{1}{b_n} + \frac{a_n}{b_n b_{n-1}} + \dots + \frac{a_n \dots a_2}{b_n \dots b_1} \right) = \infty.$$

Define the potential coefficients μ_i , $i \geq 0$, by

$$\mu_0 = 1, \quad \mu_i = \frac{b_0 b_1 \dots b_{i-1}}{a_1 a_2 \dots a_i}, \quad i \geq 1.$$

Assume that Φ_t is ergodic, that is,

$$\sum_{n=0}^{\infty} \frac{1}{\mu_n b_n} \sum_{k=0}^n \mu_k = \infty \quad \text{and} \quad \mu = \sum_{n=0}^{\infty} \mu_n < \infty.$$

Then, the invariant probability measure π exists satisfying $\pi_i = (\mu_i/\mu)$, $i \in \mathbb{Z}_+$.

Define

$$A = \sum_{i=0}^{\infty} \frac{1}{\mu_i b_i} \sum_{j=i+1}^{\infty} \mu_j, \quad \text{and} \quad S = \sum_{n=1}^{\infty} \left(\frac{1}{a_{n+1}} + \sum_{k=1}^n \frac{b_k \dots b_n}{a_k \dots a_{n+1}} \right). \quad (3.1)$$

It is easy to see that $A = S + (1/a_1)$. It was proved in [21] that the birth–death process is strongly ergodic if and only if $S < \infty$. This result can be also proved differently, for example, using an analytic method in [22] or using the coupling method in [16]. The strongly ergodic rate for this process was also considered in [17].

THEOREM 3.1. *Assume that the birth–death process Φ_t is strongly ergodic. Then (1.1) holds for any $\lambda \leq 1/A$ and $\gamma = 4cA$ with $c = \max\{1/y_1, a_1/(1 - \lambda y_1)\}$, where y_1 is given by (3.2).*

PROOF. It is easy to check that Q is monotone, so is Φ_t . From [22], we know that the function y can be defined by

$$y_0 = 0, \quad y_i = \sum_{k=0}^{i-1} \left(\frac{1}{a_{k+1}} + \sum_{j=k+1}^{\infty} \frac{b_{k+1} \cdots b_j}{a_{k+1} \cdots a_{j+1}} \right), \quad i \geq 1, \quad (3.2)$$

satisfying (DC1) with equality. Obviously, y_i is strictly increasing, Q satisfies Assumption 2.3 and $\sup_{i \in E} y_i = A$. Therefore, from Theorem 2.5 we have the assertion. \square

REMARK 3.2. Theorem 3.1 is an improvement of the result in [17], which shows that α assumes the value $(1/A)$ when $\gamma = 4Ac$. \square

3.2. Generalized Markov branching processes In this section, we consider a generalized Markov branching process with q -matrix Q given by

$$q_{ij} = \begin{cases} q_{ij}, & j > i = 0, \\ -\sum_{k=1}^{\infty} q_{0k}, & j = i = 0, \\ r_i p_{j-i+1}, & j \geq i - 1 \geq 0, j \neq i, \\ -r_i(1 - p_1), & j = i \geq 1, \\ 0, & \text{otherwise} \end{cases} \quad (3.3)$$

where $r_i \geq 0$. For $s \geq 0$, define $P(s) = \sum_{i=0}^{\infty} p_i s^i - s$. The ordinary ergodicity and exponential ergodicity of this class of branching process were investigated in [6], and the strong ergodicity was considered in [22]. However, estimating the convergence rate, as presented here, was not considered in [22]. From [6], we know that if $M_1 := \sum_{k=1}^{\infty} k p_k \leq 1$, then Q is regular, and the unique Q -process is denoted by Φ_t .

ASSUMPTION 3.3. The sequence r_i is increasing and $\sum_{j \geq k} q_{0j} \leq r_1 \sum_{j \geq k} p_j$ for any integer $k \geq 2$.

By the definition of the monotonicity of q -matrix, we know that Q is monotone if Q satisfies Assumption 3.3. Many branching processes satisfy this assumption, for example, $q_{0j} = p_j$ and $r_i = i^\theta, \theta > 0$.

THEOREM 3.4. *Assume that Q satisfies Assumption 3.3. If*

$$M_1 < 1 \quad \text{and} \quad \sum_{i=1}^{\infty} (1/r_i) < \infty,$$

then (1.1) holds for any

$$\lambda \leq \frac{p_0 - \Gamma}{\sum_{i=1}^{\infty} (1/r_i)} \quad \text{and} \quad \gamma = 4c \frac{1}{p_0 - \Gamma} \sum_{i=1}^{\infty} \frac{1}{r_i}$$

with

$$\Gamma = \sum_{k=1}^{\infty} kp_{k+1} \quad \text{and} \quad c = \max \left\{ \frac{1}{y_1}, \frac{r_1 p_0}{1 - \lambda y_1} \right\},$$

where y_1 is given by (3.4).

PROOF. Using the results in [20] under the Assumption 3.3 and $M_1 < 1$, we obtain that the branching process Φ_t is stochastically monotone. Let $\Gamma = \sum_{k=1}^{\infty} kp_{k+1}$, then $\Gamma = M_1 - 1 + p_0$. Hence, $M_1 < 1$ if and only if $\Gamma < p_0$. Construct a function y as follows:

$$y_0 = 0, \quad y_n = \frac{1}{p_0 - \Gamma} \sum_{i=1}^n \frac{1}{r_i}, \quad n \geq 1. \tag{3.4}$$

Since r_i is increasing, $y_i - y_{i-1}$ is decreasing for $i \geq 1$. For $i \geq 1$,

$$\begin{aligned} \sum_{j \in E} q_{ij} y_j + 1 &= \sum_{j \in E} q_{ij} (y_j - y_i) + 1 \\ &\leq r_i p_0 (y_{i-1} - y_i) + \sum_{k=1}^{\infty} kp_{k+1} r_i (y_{i+1} - y_i) + 1 \\ &= r_i p_0 (y_{i-1} - y_i) + r_i \Gamma (y_{i+1} - y_i) + 1 \\ &\leq 0. \end{aligned}$$

Thus we have checked that y satisfies (DC1). Obviously, Q satisfies Assumption 3.3 and y is strictly increasing, so the assertion follows from Theorem 2.5. \square

REMARK 3.5. In (3.3), let $r_i = i^\theta, \theta > 0$. From [3, Lemma 2.2], we know that

$$E_i[\tau_0] = \frac{1}{\Gamma(\theta)} \int_0^1 \frac{1 - x^i}{P(x)} (-\ln x)^{\theta-1} dx$$

for any $i \geq 1$, where $\Gamma(\theta)$ is the gamma function. Then Φ_t is strongly ergodic if and only if

$$D := \sup_{i \in E} E_i[\tau_0] = \frac{1}{\Gamma(\theta)} \int_0^1 \frac{dx}{P(x)(-\ln x)^{1-\theta}} < \infty.$$

Suppose that $D < \infty$ and that $\sum_{j \geq k} q_{0j} \leq r_1 \sum_{j \geq k} p_j$ for any $k \in \mathbb{N}$ with $k \geq 2$. Then it follows from Theorem 2.5 that (1.1) holds for any

$$\lambda \leq \frac{1}{D} \quad \text{and} \quad \gamma = \max \left\{ \frac{r_1 p_0}{1 - \lambda y_1}, \frac{1}{y_1} \right\}$$

with

$$y_1 = \frac{1}{\Gamma(\theta)} \int_0^1 \frac{1-x}{P(x)} (-\ln x)^{\theta-1} dx. \quad \square$$

4. Reversible Markov chains

The lower bound for the best rate α is investigated for the reversible Markov chains in this section. We do so by virtue of their skeleton chains. This idea is often used to extend the discrete-time results to their continuous-time counterparts. We first review the definition of the reversibility.

Suppose that the continuous-time chain Φ_t is ergodic with invariant probability measure (π_i) . We say that the Markov chain Φ_t is reversible if $\pi_i P^t(i, j) = \pi_j P^t(j, i)$ for all $i, j \in E$ and all $t \geq 0$. By [1, Propositions 1.2 and 1.6 in Section 7.1], we know that the process is reversible if and only if Q is symmetric (that is, $\pi_i q_{ij} = \pi_j q_{ji}$).

We call an ergodic discrete-time Markov chain reversible if $\pi_i P(i, j) = \pi_j P(j, i)$ for all $i, j \in E$, equivalently (see, for example, [2])

$$\sum_{i \in E} P f(i) g(i) \pi(i) = \sum_{i \in E} f(i) P g(i) \pi(i)$$

for all $f, g \in L^2(\pi)$. The chain Φ_n is called positive (for example, refer to [2]) if $\sum_{i \in E} P f(i) f(i) \pi(i) \geq 0$ for all $f \in L^2(\pi)$. In [2], the explicit geometric rates of convergence were obtained for a positive, reversible and geometrically ergodic discrete-time Markov chain on a general state space, which are stated in Proposition 4.3. It was pointed out in [2] that any skeleton chain Φ_{nh} of the reversible continuous-time chain Φ_t is positive. Hence, if Φ_t is reversible, then any skeleton chain of Φ_t is positive and reversible.

For the skeleton chain Φ_{nh} , define $\tau_i(h) = h \inf\{n \geq 1 \mid \Phi_{nh} = i\}$ to be the first hitting time on i and $P^h(i, j) = P[\Phi_h = j \mid \Phi_0 = i]$ to be the transition probability. In order to adopt the idea of skeleton chains, we need the following lemma.

LEMMA 4.1. *If $\sup_{i \in E} E_i[e^{r\tau_j}] < \infty$ for some $j \in E$ and $r > 0$, then there exists some $h_0 > 0$ such that $\sup_{i \in E} E_i[e^{r\tau_j^{(h)}}] < \infty$ for all h with $0 < h \leq h_0$.*

PROOF. Suppose that $\Phi_0 = i$. Let D_k be the k th sojourn time in j and W_k be the length of the interval between the k th exit from j and the next visit to j for the chain Φ_t . Note that the D_k are independent, and the W_k too. Also, the D_k are independent of the W_k . Moreover, D_k are identically exponentially distributed with the parameter $q_j > 0$.

Define

$$N = \min\{n \geq 1 \mid \text{the skeleton chain } \Phi_{nh} \text{ is in state } j \text{ during the interval } D_n\}.$$

Since $\tau_j(h) \leq \sum_{k=1}^{N-1} (D_k + W_k) + h + \tau_j$,

$$\begin{aligned}
 E_i[e^{r\tau_j(h)}] &\leq e^{rh} E_i[e^{r\tau_j}] \sum_{n=1}^{\infty} E_i \left[\exp \left[r \sum_{k=1}^{n-1} (D_k + W_k) \right] I_{\{N\}} \right] \\
 &\leq e^{rh} E_i[e^{r\tau_j}] \sum_{n=1}^{\infty} E_j \left[\exp \left[r \sum_{k=1}^{n-1} (D_k + W_k) \right] I_{\cap_{k=1}^{n-1} \{D_k \leq h\}} \right] \\
 &\leq e^{rh} E_i[e^{r\tau_j}] \sum_{n=1}^{\infty} E_j \left[\exp \left[r \left((n-1)h + \sum_{k=1}^{n-1} W_k \right) \right] I_{\cap_{k=1}^{n-1} \{D_k \leq h\}} \right] \\
 &= \sum_{n=1}^{\infty} e^{nrh} E_i[e^{r\tau_j}] E_j \left[\exp \left[r \sum_{k=1}^{n-1} W_k \right] (1 - e^{-q_j h})^{n-1} \right] \\
 &= E_i[e^{r\tau_j}] \sum_{n=1}^{\infty} e^{nrh} (E_j[e^{rW_1}])^{n-1} (1 - e^{-q_j h})^{n-1}. \tag{4.1}
 \end{aligned}$$

Note that

$$E_j[e^{rW_1}] = \sum_{k \neq j} \frac{q_{jk}}{q_j} E_k[e^{r\tau_j}] < \infty,$$

we have

$$\lim_{h \rightarrow 0} e^{rh} E_j[e^{rW_1}] (1 - e^{-q_j h}) = 0.$$

Thus, we can choose suitable h_0 such that

$$\sum_{n=1}^{\infty} e^{nrh} (E_j[e^{rW_1}])^{n-1} (1 - e^{-q_j h})^{n-1} < \infty \tag{4.2}$$

for all h with $0 < h \leq h_0$. Taking the supremum over i on both sides of (4.1), we have $\sup_{i \in E} E_i[e^{r\tau_j}(h)] < \infty$ for all h with $0 < h \leq h_0$. □

REMARK 4.2. Note that we can choose small enough h such that $\sup_{i \in E} E_i[e^{r\tau_j}(h)]$ sufficiently close to the value of $E_i[e^{r\tau_j}]$. □

The following proposition is a specific form of Theorem 1.3 in Baxendal [2], which is used in the proof to Lemma 4.4.

PROPOSITION 4.3 (Baxendal [2]). Let Φ_n be an irreducible, reversible and positive Markov chain on a countable state space E . Suppose that its transition probability matrix $P = (P(i, j))$ satisfies the following conditions:

- (i) $P(k_0, k_0) = \beta > 0$;
- (ii) there exists a finite function $\tilde{V} \geq 1$, and constants $\lambda < 1$ and $K < \infty$ satisfying

$$\sum_{j \in E} P(i, j) \tilde{V}_j \leq \begin{cases} \lambda \tilde{V}_i, & i \neq k_0, \\ K, & i = k_0. \end{cases}$$

Then whenever $\lambda < \rho < 1$,

$$\|P^n(i, \cdot) - \pi\|_{\tilde{V}} \leq M \tilde{V}(i) \rho^n$$

for all $i \in E$ and a positive constant M that depends only on β, λ, K and ρ , where $\|\mu\|_{\tilde{V}} = \sup_{|f| \leq \tilde{V}} |\mu(f)|$ is a stronger norm than the usual total variation norm.

LEMMA 4.4. *Let Φ_t be a reversible Markov chain. If $\sup_{i \in E} E_i[e^{r\tau_{k_0}}] < \infty$ for some $r > 0$, then (1.1) holds for any $\lambda < r$ with the constant γ well-defined in the proof to the lemma.*

PROOF. Since $\sup_{i \in E} E_i[e^{r\tau_{k_0}}] < \infty$, it follows from Lemma 4.1 that there exists some $h_0 > 0$ such that $\sup_{i \in E} E_i[e^{r\tau_{k_0}(h)}] < \infty$ for any h -skeleton chain with $0 < h \leq h_0$. For any fixed h with $0 < h \leq h_0$, let $\tilde{V}_{k_0} = 1$ and $\tilde{V}_i = E_i[e^{r\tau_{k_0}(h)}]$ for $i \neq k_0$, then by a similar argument as that in [15, proof of Theorem 5.1],

$$\sum_{j \in E} P^h(i, j) \tilde{V}_j = \begin{cases} e^{-rh} \tilde{V}_i, & i \neq k_0, \\ e^{-rh} E_{k_0}[e^{r\tau_{k_0}(h)}], & i = k_0. \end{cases}$$

For the continuous-time Markov chain Φ_t ,

$$P^h(k_0, k_0) \geq e^{-q_{k_0}h} > 0.$$

Since the skeleton chain Φ_{nh} is reversible and positive, by Proposition 4.3

$$\|P^{nh}(i, \cdot) - \pi\|_{\tilde{V}} \leq M \tilde{V}_i \rho^{nh}$$

for any ρ with $e^{-r} < \rho < 1$ and M depending only on r, q_{k_0}, ρ, h and $E_{k_0}[e^{r\tau_{k_0}(h)}]$. Hence, for any $t \in \mathbb{R}_+$ with $t = nh + s$ for some $n \in \mathbb{Z}_+$ and $s \in [0, h]$,

$$\begin{aligned} \|P^t(i, \cdot) - \pi\| &\leq \|P^{nh}(i, \cdot) - \pi\| \|P^s\| \\ &\leq \|P^{nh}(i, \cdot) - \pi\|_{\tilde{V}} \\ &\leq M \tilde{V}_i \rho^{nh+s} \rho^{-s} \\ &\leq M \tilde{V}_i \rho^{-h} \rho^t. \end{aligned}$$

Note that the value of h can be chosen small enough so that ρ^h is very close to 1. Thus,

$$\sup_{i \in E} \|P^t(i, \cdot) - \pi\| \leq M \sup_{i \in E} \tilde{V}_i \rho^{-h} \rho^t,$$

which completes the proof. □

Based on the above results, we can refine the lower bound for the reversible Markov chains.

THEOREM 4.5. *Let Φ_t be a reversible Markov chain. If $\sup_{i \in E} E_i[\tau_{k_0}] < \infty$ then the best rate $\alpha \geq 1/(\sup_{i \in E} E_i[\tau_{k_0}])$. Furthermore, if (DC1) holds for a function y , then $\alpha \geq 1/(\sup_{i \in E} y_i)$.*

PROOF. For a recurrent continuous-time Markov chain, by [10, Theorem 9.3.4]

$$E_i[(\tau_{k_0})^n] \leq n!(E_i[\tau_{k_0}])^n$$

for any $n \in \mathbb{Z}_+$. If $\sup_{i \in E} E_i[\tau_{k_0}] < \infty$, then

$$E_i[e^{r\tau_{k_0}}] = \sum_{n=0}^{\infty} \frac{r^n (E_i[(\tau_{k_0})^n])}{n!} \leq \sum_{n=0}^{\infty} \left(r \sup_{i \in E} E_i[\tau_{k_0}] \right)^n = \frac{1}{1 - r \sup_{i \in E} E_i[\tau_{k_0}]} < \infty$$

for any $r < 1/(\sup_{i \in E} E_i[\tau_{k_0}])$. From Lemma 4.4 we know that (1.1) holds for any $\lambda < 1/(\sup_{i \in E} E_i[\tau_{k_0}])$, which follows $\alpha \geq 1/(\sup_{i \in E} E_i[\tau_{k_0}])$.

Furthermore, if (DC1) holds, which implies from Proposition 2.1 that

$$\sup_{i \in E} E_i[\tau_{k_0}] \leq \sup_{i \in E} y_i < \infty,$$

then we have $\alpha \geq 1/(\sup_{i \in E} y_i)$ from the first part of the assertion. □

5. Examples: reversible Markov chains

5.1. Generalized birth–death processes Let Φ_t be a generalized birth–death process with q -matrix $Q = (q_{ij})$ specified by

$$q_{ij} = \begin{cases} h_j, & i = 0, \\ d_j, & j = 0, i \geq 1, \\ b_i, & j = i + 1, i \geq 1, \\ a_i, & j = i - 1, i \geq 2, \\ -(b_1 + d_1), & j = i = 1, \\ -(a_i + b_i + d_i), & j = i \geq 2, \\ 0, & \text{otherwise.} \end{cases} \tag{5.1}$$

Note that Q is conservative if and only if $h_0 = -\sum_{i=1}^{\infty} h_i$. It is known that Q is regular if and only if

$$R := \sum_{n=1}^{\infty} \left[\frac{1 + d_n}{b_n} + \frac{a_n(1 + d_{n-1})}{b_n b_{n-1}} + \dots + \frac{a_n \cdots a_2(1 + d_3)}{b_n \cdots b_2} + \frac{a_n \cdots a_2(1 + d_1)}{b_n \cdots b_2 b_1} \right] = \infty.$$

Denote by \hat{Q} the birth–death q -matrix obtained by deleting all elements in the first row and in the first column, and all d_i from the principal diagonal elements. The coefficients $\hat{\mu}_i$ are defined by

$$\hat{\mu}_1 = 1, \quad \hat{\mu}_i = \frac{b_1 b_2 \cdots b_{i-1}}{a_2 a_3 \cdots a_i}, \quad i \geq 2.$$

It is known that \hat{Q} is symmetric if and only if $\hat{\mu} := \sum_{i=1}^{\infty} \hat{\mu}_i < \infty$. For any positive real number $\hat{\mu}_0$, let $\mu_0 = \hat{\mu}_0$, $\mu_i = \hat{\mu}_i$ for $i \geq 1$ and $\mu = \sum_{i=0}^{\infty} \mu_i$. If $\hat{\mu} < \infty$, then Q is symmetric with (π_i) , where $\pi_i = \mu_i/\mu$ for $i \geq 0$, as its symmetrizing measure if and only if $h_i\pi_0 = d_i\pi_i$, $i \geq 1$.

Following the same notation as that in [22], define

$$q_1^{(0)} = d_1, \quad q_n^{(k)} = d_n, \quad k = 0, 1, \dots, n - 2, \quad q_n^{(n-1)} = d_n + a_n, \quad n \geq 2,$$

$$F_0^{(0)} = 1, \quad F_n^{(0)} = \frac{1}{b_n} \sum_{k=0}^{n-1} q_n^{(k)} F_k^{(0)}, \quad n \geq 1,$$

$$c_0 = 0, \quad c_n = \frac{1 + \sum_{k=0}^{n-1} q_n^{(k)} c_k}{q_{n,n+1}}, \quad n \geq 1.$$

We then have the following bounds for the best rate α .

THEOREM 5.1. *The Q -process Φ_t is strongly ergodic if and only if*

$$D := \sup_{k \in \mathbb{Z}_+} \sum_{n=0}^k (F_n^{(0)} c - c_n) < \infty, \tag{5.2}$$

where

$$c = \sup_{k \in \mathbb{Z}_+} \left[\frac{\sum_{n=0}^k c_n}{\sum_{n=0}^k F_n^{(0)}} \right].$$

Furthermore, if Φ_t is reversible, then the best rate $\alpha \geq (1/D)$.

PROOF. Note that the only difference between Q and a single-birth q -matrix (5.1) is in the first row. Therefore, for any state $i \neq 0$, the expectation $E_i[\tau_0]$ can be calculated according to the formula for the single-birth process, which is given, based on the [22, proof of Lemma 2.1], by

$$E_k[\tau_0] = \sum_{n=0}^k (F_n^{(0)} c - c_n).$$

It follows from Proposition 2.1 and (DC1) that Φ_t is strongly ergodic if and only if $D = \sup_{k \in E} E_k[\tau_0] < \infty$. The last assertion follows immediately from Theorem 4.1. \square

Condition (5.2) looks complicated. In the following, we provide a simpler sufficient condition for strong ergodicity by comparing it with a birth–death process.

THEOREM 5.2. *Let S and A be given by (3.1). If $S < \infty$, then Φ_t is strongly ergodic. Furthermore, if Φ_t is reversible, then the best rate $\alpha \geq 1/A$.*

PROOF. If $S < \infty$, then a bounded increasing solution to the equation in (DC1) for \hat{Q} can be constructed as in (3.2). Obviously, for $i \geq 1$

$$\sum_{j=0}^{\infty} q_{ij} y_j = -d_i y_i + [b_i y_{i+1} - (a_i + b_i) y_i + a_i y_{i-1}] \leq -1.$$

Thus, (y_i) is also a solution to the equation in (DC1) for Q . Hence, Φ_t is strongly ergodic. The last assertion follows immediately from Theorem 4.5. \square

5.2. Open Markov population processes Let $B = \mathbb{Z}_+$ and denote by B^m the set of all m -dimensional vectors $\mathbf{i} = (i_1, i_2, \dots, i_m)$ whose components $i_k \in B, 1 \leq k \leq m$, are nonnegative integers. Here \mathbf{i} is viewed as the population vector of a system consisting m colonies with i_k as the population of the k th colony. Given integers p and q , with $1 \leq p \neq q \leq m$, define the operators T_p^+, T_p^- and T_{pq} on B^m by

$$T_p^+(i_1, i_2, \dots, i_m) = (i_1, i_2, \dots, i_p + 1, \dots, i_m),$$

and provided $i_p \geq 1$,

$$T_p^-(i_1, i_2, \dots, i_d) = (i_1, i_2, \dots, i_p - 1, \dots, i_m)$$

and

$$T_{pq}(i_1, i_2, \dots, i_p, \dots, i_q, \dots, i_d) = (i_1, i_2, \dots, i_p - 1, \dots, i_q + 1, \dots, i_m).$$

DEFINITION 5.3. A continuous-time Markov chain Φ_t with the state space E a subset of B^m is called a Markov population process if its q -matrix Q is given by

$$q_{ij} = \begin{cases} \alpha_p(\mathbf{i}) & \text{if } \mathbf{j} = T_p^+(\mathbf{i}) \text{ for some } p, \\ \beta_p(\mathbf{i}) & \text{if } \mathbf{j} = T_p^-(\mathbf{i}) \text{ for some } p, \\ \gamma_{pq}(\mathbf{i}) & \text{if } \mathbf{j} = T_{pq}(\mathbf{i}) \text{ for some } p \neq q, \\ -\sum_{\mathbf{i} \neq \mathbf{j}} q_{ij} & \text{if } \mathbf{j} = \mathbf{i} \text{ (to keep } Q \text{ conservative),} \\ 0 & \text{otherwise,} \end{cases} \tag{5.3}$$

where the functions $\alpha_p(\mathbf{i}), \beta_p(\mathbf{i})$ and $\gamma_{pq}(\mathbf{i})$ are nonnegative with the latter two vanishing for $i_p = 0$. The Markov population process is called open if there exists a vector $\mathbf{I} = (I_1, I_2, \dots, I_m)$ where some or all of the components may be infinite, such that

$$E = \{\mathbf{i} \in B^m \mid i_1 \leq I_1, i_2 \leq I_2, \dots, i_m \leq I_m\}.$$

For any $n \geq 0$, denote by E_n the level subset $\{\mathbf{i} \in E \mid i_1 + i_2 + \dots + i_m = n\}$ of E . Here $\{E_n, n \in \mathbb{Z}_+\}$ is a partition of the space E such that $E_0 = \{\mathbf{0}\} = \{(0, 0, \dots, 0)\}$ and $\sum_{n=0}^\infty E_n = E$. Define

$$b_n = \max \left\{ \sum_{p=1}^m \alpha_p(\mathbf{i}) \mid \mathbf{i} \in E_n \right\} \quad \text{and} \quad a_n = \min \left\{ \sum_{p=1}^d \beta_p(\mathbf{i}) \mid \mathbf{i} \in E_n \right\}.$$

Consider a birth–death process with birth coefficients b_n and death coefficients a_n . Define R_b, A and S as in Section 3.1. The population process was studied by comparing it with the birth–death process (for example, [18]). By [1, Proposition 3.1 in Section 9.3], we know that if $R_b = \infty$, then Q is regular and Φ_t is unique. In [13], it was shown that under some conditions on Q , the chain Φ_t is reversible.

THEOREM 5.4. *If $S < \infty$, then the Q -process Φ_t is strongly ergodic. Moreover, if Φ_t is reversible, then $\alpha \geq 1/A$.*

PROOF. Since $S < +\infty$, the Q_b -process is strongly ergodic. As $\{y_n, n \in \mathbb{Z}_+\}$ defined by (3.2) is a bounded nonnegative and increasing solution to the following equation:

$$b_n(y_{n+1} - y_n) = a_n(y_n - y_{n-1}) - 1, \quad n \geq 1.$$

Then for $\mathbf{i} \in E_n, n \geq 1$,

$$\sum_{p=1}^m \alpha_p(\mathbf{i})(y_{n+1} - y_n) \leq \sum_{p=1}^m \beta_p(\mathbf{i})(y_n - y_{n-1}) - 1, \tag{5.4}$$

which, by letting $z_i = y_n$, leads to

$$\begin{aligned} & \sum_{p=1}^m \alpha_p(\mathbf{i})z_{T_p^+(\mathbf{i})} + \sum_{p=1}^m \beta_p(\mathbf{i})z_{T_p^-(\mathbf{i})} + \sum_{p=1}^m \sum_{q=1}^m \gamma_{pq}(\mathbf{i})z_{T_{pq}(\mathbf{i})} + 1 \\ & \leq \left(\sum_{p=1}^m \alpha_p(\mathbf{i}) + \sum_{p=1}^m \beta_p(\mathbf{i}) + \sum_{p=1}^m \sum_{q=1}^m \gamma_{pq}(\mathbf{i}) \right) z_{\mathbf{i}}, \end{aligned}$$

or

$$\sum_{\mathbf{j} \in B^m} q_{ij}z_{\mathbf{j}} \leq -1, \quad \mathbf{i} \neq \mathbf{0}. \tag{5.5}$$

Thus, we have checked that z with $z_{\mathbf{0}} = \mathbf{0}$ is a solution to (DC1). Since

$$\sup_{\mathbf{i} \in E} z_{\mathbf{i}} = \sup_{n \in \mathbb{Z}_+} y_n = A < \infty.$$

Here Φ_t is strongly ergodic. The last assertion follows from Theorem 4.5. □

Acknowledgements

This work was supported in part by NSERC/MITACS grants of Canada, National Natural Science Foundation of China (Grant No. 10671212), and the Australian Research Council Centre of Excellence for Mathematics and Statistics of Complex Systems.

References

- [1] W. J. Anderson, *Continuous-time Markov chains. An applications-oriented approach* (Springer-Verlag, New York, 1991).
- [2] P. H. Baxendale, "Renewal theory and computable convergence rates for geometrically ergodic Markov chains", *Ann. Appl. Probab.* **15** (2005) 700–738.
- [3] A. Y. Chen, "Ergodicity and stability generalized Markov branching processes with resurrection", *J. Appl. Probab.* **39** (2002) 786–803.

- [4] M. F. Chen, "Estimation of spectral gap for Markov chains", *Acta Math. Sin. New Ser.* **12** (1996) 337–360.
- [5] M. F. Chen, "Equivalence of exponential ergodicity and L^2 -exponential convergence for Markov chains", *Stochastic Process Appl.* **87** (2000) 281–297.
- [6] R. R. Chen, "An extended class of time-continuous branching processes", *J. Appl. Probab.* **34** (1997) 14–23.
- [7] P. Diaconis and L. Saloff-Coste, "Nash's inequality for finite Markov chains", *J. Theor. Probab.* **9** (1996) 459–510.
- [8] W. Doeblin, "Sur les propriétés asymptotiques du mouvement régis par certains types de chaînes simples", *Bull. Math. Soc. Roum. Sci.* **39** (1937) 57–115.
- [9] D. Down, S. P. Meyn and R. L. Tweedie, "Exponential and uniform ergodicity of Markov processes", *Ann. Probab.* **23** (1995) 1671–1691.
- [10] Z. T. Hou and Q. F. Guo, *Homogeneous denumerable Markov processes* (Springer, New York, 1988).
- [11] D. Isaacson and G. R. Luecke, "Strongly ergodic Markov chains and rates of convergence using spectral conditions", *Stochastic Process Appl.* **7** (1978) 113–121.
- [12] D. Isaacson and R. L. Tweedie, "Criteria for strong ergodicity for Markov chains", *J. Appl. Probab.* **15** (1978) 87–95.
- [13] J. F. C. Kingman, "Markov population processes", *J. Appl. Probab.* **6** (1969) 1–18.
- [14] R. B. Lund, S. P. Meyn and R. L. Tweedie, "Computable exponential convergence rates for stochastically ordered Markov processes", *Ann. Appl. Probab.* **6** (1996) 218–237.
- [15] R. B. Lund and R. L. Tweedie, "Geometric convergence rates for stochastically ordered Markov chains", *Math. Oper. Res.* **20** (1996) 182–194.
- [16] Y. H. Mao, "Strong ergodicity for Markov processes by coupling methods", *J. Appl. Probab.* **39** (2002) 839–852.
- [17] Y. H. Mao, "Convergence rates in strong ergodicity for Markov processes", *Stochastic Process Appl.* **116** (2006) 1964–1976.
- [18] G. E. H. Reuter, "Competition processes", *Proc. Fourth Berkeley Symp. Math. Statist. Probab.* **2** (1961) 421–430.
- [19] R. L. Tweedie, "Criteria for ergodicity, exponential ergodicity and strong ergodicity of Markov processes", *J. Appl. Probab.* **18** (1981) 122–130.
- [20] H. J. Zhang and A. Y. Chen, "Stochastic comparability and dual q -functions", *J. Math. Anal. Appl.* **234** (1999) 482–499.
- [21] H. J. Zhang, A. Y. Chen, X. Lin and Z. T. Hou, "Strong ergodicity of monotone transition functions", *Statist. Probab. Lett.* **55** (2001) 63–69.
- [22] Y. H. Zhang, "Strong ergodicity for single-birth processes", *J. Appl. Probab.* **38** (2001) 270–277.