J. Appl. Prob. **46,** 296–307 (2009)
Printed in England

EXTINCTION TIMES IN MULTITYPE MARKOV BRANCHING PROCESSES

DOMINIK HEINZMANN,* *University of Zurich*

Abstract

In this paper, a distributional approximation to the time to extinction in a subcritical continuous-time Markov branching process is derived. A limit theorem for this distribution is established and the error in the approximation is quantified. The accuracy of the approximation is illustrated in an epidemiological example. Since Markov branching processes serve as approximations to nonlinear epidemic processes in the initial and final stages, our results can also be used to describe the time to extinction for such processes.

Keywords: Multitype branching process; extinction time; convergence rate

2000 Mathematics Subject Classification: Primary 60J80
Secondary 92D30

1. Introduction

This paper is concerned with approximating the time to extinction in a subcritical multitype Markov branching process, starting with many individuals. The argument is based on the classical exponential approximation to the extinction probabilities [1], [11]–[13], [17]. These approximations are then combined with the branching property to derive a Gumbel approximation. The bound on the error in the total variation distance is inversely proportional to a positive power of a weighted sum of the number of individuals of the different types. The power depends on the means and higher moments of the offspring distribution.

In infectious disease modeling, the initial and final stages of epidemic processes can often be approximated by suitable branching processes; see [18]. More recently, in [2]–[5], different constructions have been used to quantify the path accuracy of such approximations. These results can be combined with ours to derive corresponding statements about the extinction time in epidemic processes.

2. Equations for extinction probabilities

The notation is chosen with [1, p. 200], [11, p. 113], and [17, p. 77] as basic references. For $k < \infty$, set $\mathbf{Z}(t) = (Z_1(t), \dots, Z_k(t))$, where $Z_i(t)$ is the number of individuals of type i at time t. A type i individual has exponential lifetime with parameter a_i and rises at death j_i type i individuals, $1 \le i \le k$, with probability $p_i^{\boldsymbol{j}}$, where $\boldsymbol{j} = (j_1, \dots, j_k) \in \mathbb{Z}_+^k$, independent of everything that has happened up to this time. Assume that

$$R_{il} := \sum_{\boldsymbol{j} \in \mathbb{N}^k} p_i^{\boldsymbol{j}} j_l < \infty \quad \text{for } i \ge 1 \text{ and } l \le k. \tag{2.1}$$

Received 15 August 2008; revision received 4 November 2008.
* Postal address: Institute of Mathematics, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.
Email address: dominik.heinzmann@math.uzh.ch

Let $\|\cdot\|$ be the supremum norm, and let $\mathrm{P}_{\boldsymbol{I}}$ be a conditional distribution of the process at time t given $\boldsymbol{Z}(0) = \boldsymbol{I}$ for $\boldsymbol{I} = (I_1, \dots, I_k) \in \mathbb{Z}_+^k$. In particular, let P_i correspond to the case when $Z_i(0) = 1$ and $Z_m(0) = 0$, $m \neq i$. Let T be the extinction time of the process, and define the survival probability of the process when starting with a single type i individual as $q_i(t) := 1 - \mathrm{P}_i(T \le t) = 1 - \mathrm{P}_i(\boldsymbol{Z}(t) = \boldsymbol{0})$.

Then Equation (15.2) of [11, p. 114] implies that

$$\frac{1}{a_i}\frac{\mathrm{d}}{\mathrm{d}t}(1 - q_i(t)) = q_i(t) - \sum_{\boldsymbol{j}\in\mathbb{N}^k} p_i^{\boldsymbol{j}}\left(1 - \prod_{l=1}^{k}(1 - q_l(t))^{j_l}\right) \quad \text{for } 1 \le i \le k. \tag{2.2}$$

Now the relation $(1-x)^l \ge 1 - xl$ for $0 \le x \le 1$, together with the Bonferroni inequalities [7, p. 27], implies that

$$\sum_{\boldsymbol{j}\in\mathbb{N}^k} p_i^{\boldsymbol{j}}\left(1 - \prod_{l=1}^{k}(1 - q_l(t))^{j_l}\right) \le \sum_{\boldsymbol{j}\in\mathbb{N}^k} p_i^{\boldsymbol{j}}(\boldsymbol{j}^\top \boldsymbol{q}(t)) \tag{2.3}$$

and

$$\sum_{\boldsymbol{j}\in\mathbb{N}^k} p_i^{\boldsymbol{j}}\left(1 - \prod_{l=1}^{k}(1 - q_l(t))^{j_l}\right) \ge \sum_{\boldsymbol{j}\in\mathbb{N}^k} p_i^{\boldsymbol{j}} \max\{\boldsymbol{j}^\top \boldsymbol{q}(t) - F^{\boldsymbol{j}}(\boldsymbol{q}(t)), 0\}, \tag{2.4}$$

where $\boldsymbol{q}(t) = (q_1(t), \dots, q_k(t))$ and

$$F^{\boldsymbol{j}}(\boldsymbol{q}(t)) := \frac{1}{2}\sum_{\substack{l,l'=1\\ l\neq l'}}^{k} j_l j_{l'} q_l(t) q_{l'}(t) + \frac{1}{2}\sum_{l=1}^{k} j_l(j_l - 1)q_l^2(t) \le \frac{1}{2}(\boldsymbol{j}^\top \boldsymbol{q}(t))^2.$$

Using (2.3) and (2.4) in (2.2) and recalling (2.1) gives

$$\frac{1}{a_i}\frac{\mathrm{d}q_i(t)}{\mathrm{d}t} \le \{(\boldsymbol{R} - \boldsymbol{I})\boldsymbol{q}(t)\}_i \tag{2.5}$$

and

$$\frac{1}{a_i}\frac{\mathrm{d}q_i(t)}{\mathrm{d}t} = \{(\boldsymbol{R} - \boldsymbol{I})\boldsymbol{q}(t)\}_i - v_i(t), \tag{2.6}$$

where $v_i(t)$ summarizes all the nonlinear terms in $\boldsymbol{q}(t)$ (see Section 5 for an example) and satisfies

$$\begin{aligned} 0 &\le v_i(t) \\ &= \sum_{\boldsymbol{j}\in\mathbb{N}^k} p_i^{\boldsymbol{j}}\left(\sum_{l=1}^{k} j_l q_l(t) - 1 + \prod_{l=1}^{k}(1 - q_l(t))^{j_l}\right) \\ &\le \sum_{\boldsymbol{j}\in\mathbb{N}^k} p_i^{\boldsymbol{j}} \min\{F^{\boldsymbol{j}}(\boldsymbol{q}(t)), \boldsymbol{j}^\top \boldsymbol{q}(t)\}. \end{aligned} \tag{2.7}$$

Since (2.2) is nonlinear in $\boldsymbol{q}(t)$, it cannot in general be solved analytically. However, we will see, using Theorem 3.1, below, that the behavior of the solution $\boldsymbol{q}(t)$ can be approximated by that of

$$\frac{\mathrm{d}\boldsymbol{q}(t)}{\mathrm{d}t} = \{\boldsymbol{A}(\boldsymbol{R} - \boldsymbol{I})\}\boldsymbol{q}(t) =: \boldsymbol{B}\boldsymbol{q}(t),$$

https://doi.org/10.1239/jap/1238592131 Published online by Cambridge University Press

so long as $\|\boldsymbol{q}(t)\|$ is small and $\boldsymbol{A} := \operatorname{diag}\{a_1, \ldots, a_k\}$. The matrix $\boldsymbol{B} = \boldsymbol{A}(\boldsymbol{R} - \boldsymbol{I})$ has nonnegative elements off the diagonal, and is thus a Metzler–Leontief matrix [16, p. 40]. If $\boldsymbol{B}$ is irreducible [16, p. 15], the process $\boldsymbol{Z}(t)$ is irreducible [17, p. 99] and the following Perron–Fröbenius result [16, Theorem 2.5] applies.

Theorem 2.1. *Assume that* $\boldsymbol{B}$ *is a* $k \times k$ *irreducible matrix with nonnegative off-diagonal elements. Then there exists an eigenvalue* ω *such that*

(i) ω *is real;*

(ii) *there exists a unique (up to a constant factor) strictly positive left eigenvector* $\boldsymbol{f}_1$ *and a unique strictly positive right eigenvector* $\boldsymbol{b}_1$ *associated with* ω*;*

(iii) $\omega > \operatorname{Re}(\omega_i)$ *for any eigenvalue* $\omega_i \neq \omega$ *of* $\boldsymbol{B}$*;*

(iv) ω *is a simple root of the characteristic equation of* $\boldsymbol{B}$.

In what follows, it is assumed that the process is subcritical, i.e. $\omega < 0$. Define $r := -\omega$. The left eigenvector $\boldsymbol{f}_1$ can be used to derive an upper bound for $\boldsymbol{q}(t)$ $(t > 0)$.

Lemma 2.1. *Assume that* $\boldsymbol{f}_1^\top = (f_{11}, \ldots, f_{1k})$ *is such that* $\|\boldsymbol{f}_1\| = 1$. *Then*

$$q_i(t) \leq \mathrm{e}^{-rt}\left(\frac{\boldsymbol{f}_1^\top \boldsymbol{1}}{f_{1i}}\right) \quad \text{for } 1 \leq i \leq k,$$

where $\boldsymbol{1}$ *denotes a column vector of* 1*s.*

Proof. Theorem 2.1 implies that $\boldsymbol{f}_1$ has only positive entries and, hence, inequality (2.5) implies that

$$\frac{\mathrm{d}}{\mathrm{d}t}\{\boldsymbol{f}_1^\top \boldsymbol{q}(t)\} \leq \boldsymbol{f}_1^\top \boldsymbol{B} \boldsymbol{q}(t) = \omega \boldsymbol{f}_1^\top \boldsymbol{q}(t) = -r \boldsymbol{f}_1^\top \boldsymbol{q}(t).$$

Using Grönwall's lemma [10] yields

$$\boldsymbol{f}_1^\top \boldsymbol{q}(t) \leq \mathrm{e}^{-rt} \boldsymbol{f}_1^\top \boldsymbol{q}(0) = \mathrm{e}^{-rt} \boldsymbol{f}_1^\top \boldsymbol{1}.$$

The result follows immediately, since $\boldsymbol{f}_1$ and $\boldsymbol{q}(t)$ are both positive vectors.

The following useful lemma is proved by a standard argument.

Lemma 2.2. *Let* X *be a nonnegative random variable with* $\mathrm{E}(X) < \infty$, *and let* $d > 1$. *Then, for* $\delta \to 0$, $\mathrm{E}(X\delta \wedge (X\delta)^d) = o(\delta)$. *If, in addition,* $\mathrm{E}(X^\psi) < \infty$ *for some* $1 \leq \psi \leq d$ *then* $\mathrm{E}(X\delta \wedge (X\delta)^d) \leq 2\,\mathrm{E}(X\delta)^\psi = O(\delta^\psi)$.

Let J_i denote a random variable with $\mathrm{P}(J_i = \boldsymbol{j}) = p_i^{\boldsymbol{j}}$.

Theorem 2.2. *If* $\mathrm{E}(\|J_i\|) < \infty$ *for* $1 \leq i \leq k$ *then* $v_i(t) = o(\|\boldsymbol{q}(t)\|)$ *as* $t \to \infty$.

Proof. From (2.7), it follows that

$$0 \leq v_i(t) \leq \sum_{\boldsymbol{j} \in \mathbb{N}^k} p_i^{\boldsymbol{j}} \left(\frac{1}{2}(\boldsymbol{j}^\top \boldsymbol{q}(t))^2 \wedge (\boldsymbol{j}^\top \boldsymbol{q}(t))\right) = \mathrm{E}\left(\frac{1}{2}(J^\top \boldsymbol{q}(t))^2 \wedge (\boldsymbol{J}^\top \boldsymbol{q}(t))\right), \qquad (2.8)$$

where $\boldsymbol{J}^\top = (J_1, \ldots, J_k)$. Since $\|\boldsymbol{q}(t)\| \leq \sum_{i=1}^k |q_i(t)|$, Lemma 2.1 indicates that $\|\boldsymbol{q}(t)\| \to 0$ as $t \to \infty$, and, thus, Lemma 2.2 can be applied.

The following corollary gives a more specific asymptotic upper bound on $v_i(t)$, if the offspring distributions have a finite moment higher than the first.

https://doi.org/10.1239/jap/1238592131 Published online by Cambridge University Press

Corollary 2.1. *Suppose that* $\mathrm{E}(\|J_i\|^{1+\alpha}) < \infty$ *for some* $0 < \alpha \leq 1$ *and all* $1 \leq i \leq k$. *Then there exist constants* $c_i^* < \infty$ *such that* $0 \leq v_i(t) \leq c_i^* \|\boldsymbol{q}(t)\|^{1+\alpha}$, $1 \leq i \leq k$.

Proof. The proof follows immediately from (2.8) and Lemma 2.2.

3. Exponential limit behavior

The following result is the basis for approximating the survival time of the process, bounding the error in the exponential approximation to the extinction probabilities.

Theorem 3.1. *Assume that* $\mathrm{E}(\|J_i\|^{1+\alpha}) < \infty$ *for some* $0 < \alpha \leq 1$ *and all* $1 \leq i \leq k$. *If* $\boldsymbol{B}$ *is irreducible with largest eigenvalue* $-r < 0$, *the probability of survival* $q_i(t)$ *when starting with a single individual of type* i *satisfies*

$$q_i(t) = c_i \mathrm{e}^{-rt}(1 + o(\mathrm{e}^{-\gamma t})),$$

where $0 < \gamma < r$ *is given below and* $c_i/c_l = b_{1i}/b_{1l}$, *where* $\boldsymbol{b}_1$ *is the right eigenvector of* $\boldsymbol{B}$ *corresponding to the eigenvalue* $-r$.

Proof. Let $\boldsymbol{v}(t) := (v_1(t), \dots, v_k(t))$, and define $\boldsymbol{u}(t) := \mathrm{e}^{rt}\,\boldsymbol{q}(t)$. It follows from (2.6) that

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{u}(t) = \boldsymbol{C}\boldsymbol{u}(t) - \mathrm{e}^{rt}\boldsymbol{A}\boldsymbol{v}(t), \tag{3.1}$$

where the largest eigenvalue of $\boldsymbol{C} := (\boldsymbol{B} + r\boldsymbol{I})$ is 0. Let 0 and $\{\omega_j; 2 \leq j \leq k^*\}$ denote the eigenvalues corresponding to the $k^* \leq k$ Jordan blocks of $\boldsymbol{C}$, and denote by k_j, $2 \leq j \leq k^*$, their dimensions. The left eigenvector of $\boldsymbol{C}$ corresponding to the eigenvalue 0 is $\boldsymbol{f}_1^\top$; for $1 \leq m \leq k_j$ and $2 \leq j \leq k^*$, let $\boldsymbol{f}_{i,m}^\top$ denote the corresponding Jordan basis vectors, with $\|\boldsymbol{f}_{j,m}\| = 1$; set $-\beta_j = \mathrm{Re}(\omega_j)$, so that, for $2 \leq m \leq k_j$ and $2 \leq j \leq k^*$, $\boldsymbol{f}_{j,1}^\top \boldsymbol{C} = \omega_j \boldsymbol{f}_{j,1}^\top$ and $\boldsymbol{f}_{j,m}^\top \boldsymbol{C} = \omega_j \boldsymbol{f}_{j,m}^\top + \boldsymbol{f}_{j,m-1}^\top$.

Define $\boldsymbol{w}(t) := \boldsymbol{f}_1^\top \boldsymbol{A}\boldsymbol{v}(t)/(\boldsymbol{f}_1^\top \boldsymbol{q}(t))$. From Lemma 2.1 and Corollary 2.1, it is immediate that $\|\boldsymbol{w}(t)\| = O(\mathrm{e}^{-r\alpha t})$ and, hence, that $\int_s^\infty \|\boldsymbol{w}(t)\|\,\mathrm{d}t < \infty$.

Now, from (3.1),

$$\frac{\mathrm{d}}{\mathrm{d}t}\log(\boldsymbol{f}_1^\top \boldsymbol{u}(t)) = -\boldsymbol{w}(t),$$

and, hence,

$$\log(\boldsymbol{f}_1^\top \boldsymbol{q}(t)) + rt = \log(\boldsymbol{f}_1^\top \boldsymbol{q}(s)) + rs - \int_s^t \boldsymbol{w}(z)\,\mathrm{d}z.$$

By the Cauchy criterion and the integrability of $\|\boldsymbol{w}(z)\|$, it follows that

$$\lim_{t\to\infty}\{\log(\boldsymbol{f}_1^\top \boldsymbol{q}(t)) + rt\} =: \log h^*$$

exists and is finite, and, thus, using $\|\boldsymbol{w}(t)\| = O(\mathrm{e}^{-r\alpha t})$,

$$\boldsymbol{f}_1^\top \boldsymbol{q}(t) = h^* \mathrm{e}^{-rt}(1 + O(\mathrm{e}^{-r\alpha t}))$$

with $h^* > 0$.

For the remaining part of the argument, we refer to the theory of perturbed linear systems. Rewrite (3.1) as

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{u}(t) = [\boldsymbol{C} + \boldsymbol{D}(t)]\boldsymbol{u}(t), \tag{3.2}$$

where

$$\boldsymbol{D}(t) = -\frac{\boldsymbol{A}\boldsymbol{v}(t)\boldsymbol{q}(t)^\top}{\|\boldsymbol{q}(t)\|^2},$$

so that $\|\boldsymbol{D}(t)\| \le K^* \mathrm{e}^{-r\alpha t}$ with $K^* < \infty$. System (3.2) is a special case of the system in Theorem 2 of [14], from which it follows that, for any $\gamma < \min\{r\alpha, \beta_{[2]}\}$, where $-\beta_{[2]}$ is the second largest real part of any eigenvalue of $\boldsymbol{C}$, we have $|\boldsymbol{f}_{j,m}^\top \boldsymbol{u}(t)| = o(\mathrm{e}^{-\gamma t})$, $1 \le m \le k_j$ and $2 \le j \le k^*$.

Now the set of vectors $\{\boldsymbol{f}_1^\top, \boldsymbol{f}_{j,m}^\top;\ 1 \le m \le k_j,\ 2 \le j \le k^*\}$ constitutes a basis of $\mathbb{R}^d$. Let $\boldsymbol{x} \in \mathbb{R}^d$ have coefficients $(x_1, x_{j,m};\ 1 \le m \le k_j,\ 2 \le j \le k^*)$ with respect to this basis. Then

$$\boldsymbol{x}^\top(\mathrm{e}^{rt}\boldsymbol{q}(t)) = \bigg(x_1\boldsymbol{f}_1^\top + \sum_{j=2}^{k^*}\sum_{n=1}^{k_j} x_{j,n}\boldsymbol{f}_{j,n}^\top\bigg)\boldsymbol{u}(t) = x_1 h^* + o(\|\boldsymbol{x}\|\mathrm{e}^{-\gamma t}). \tag{3.3}$$

In particular, for $1 \le i \le k$, it follows that $q_i(t) = c_i \mathrm{e}^{-rt}(1 + o(\mathrm{e}^{-\gamma t}))$, where

$$c_i = (\boldsymbol{e}_i^\top \boldsymbol{b}_1)h^* = b_{1i}h^* > 0, \tag{3.4}$$

$\boldsymbol{e}_i$ is a column vector with 1 in the ith position and 0s elsewhere, and $\boldsymbol{b}_1$ is the right eigenvector of $\boldsymbol{B}$ corresponding to the eigenvalue $-r$ such that $\boldsymbol{f}_1^\top \boldsymbol{b}_1 = 1$.

Remark 3.1. The order of convergence is simplified for clarity in the statement of Theorem 3.1. For the case where $\boldsymbol{B}$ is diagonalizable, the exact formulation is as follows. If $-\beta_2$ is the second largest real part of an eigenvalue of $\boldsymbol{C}$ and if $r\alpha \ne \beta_2$, then $q_i(t) = c_i\mathrm{e}^{-rt}(1 + O(\mathrm{e}^{-\gamma t}))$, where $\gamma = \min\{r\alpha, \beta_2\}$. Otherwise, if $r\alpha = \beta_2$ then $q_i(t) = c_i\mathrm{e}^{-rt}(1 + O(t\mathrm{e}^{-r\alpha t}))$.

4. Time to extinction

If $\mathrm{E}(\|J_i\|^{1+\alpha}) < \infty$ for some $0 < \alpha \le 1$ and all $1 \le i \le k$, Theorem 3.1 implies that, as $t \to \infty$,

$$\mathrm{P}_{\boldsymbol{I}}(T > t) = 1 - \prod_{i=1}^{k}(\mathrm{P}_i(T \le t))^{I_i} = 1 - \prod_{i=1}^{k}(1 - q_i(t))^{I_i} \sim 1 - \prod_{i=1}^{k}(1 - c_i\mathrm{e}^{-rt})^{I_i}, \tag{4.1}$$

where $c_i > 0$ $(1 \le i \le k)$ and $\boldsymbol{I} = (I_1, \dots, I_k)$ with I_i the initial number of type i individuals. Define $C_{\boldsymbol{I}} := \sum_{j=1}^{k} I_j c_j$. The approximation error in (4.1) is controlled by the following result.

Lemma 4.1. *Suppose that* $\mathrm{E}(\|J_i\|^{1+\alpha}) < \infty$ *for some* $0 < \alpha \le 1$ *and all* $1 \le i \le k$*. Then, for any* γ *as in Theorem 3.1, there exist* $t_0, \nu_1 < \infty$*, not depending on* $\boldsymbol{I}$*, such that*

$$\bigg|\prod_{i=1}^{k}(1 - q_i(t))^{I_i} - \prod_{i=1}^{k}(1 - c_i\mathrm{e}^{-rt})^{I_i}\bigg| \le \nu_1 C_{\boldsymbol{I}} \exp\bigg(-\frac{1}{2}C_{\boldsymbol{I}}\mathrm{e}^{-rt}\bigg)\mathrm{e}^{-(r+\gamma)t}, \qquad t \ge t_0.$$

Proof. Denote the approximation error in (4.1) as $\varepsilon^{(1)}(t)$. Choose

$$t_1 \ge \frac{1}{r}\max_{1\le i\le k}(\log c_i)_+$$

such that $q_i(t) \ge \frac{1}{2}c_i\mathrm{e}^{-rt}$ for all i and $t \ge t_1$. Using

$$\bigg|\prod_{i=1}^{k}A_i - \prod_{i=1}^{k}B_i\bigg| \le \sum_{l=1}^{k}|A_l - B_l|\bigg(\prod_{i=1}^{l-1}|A_i|\bigg)\bigg(\prod_{i=l+1}^{k}|B_i|\bigg),$$

where $A_i = (1 - q_i(t))^{I_i}$ and $B_i = (1 - c_i \mathrm{e}^{-rt})^{I_i}$, it follows that

$$\varepsilon^{(1)}(t) \le \exp\biggl(-\frac{1}{2}C_{\boldsymbol{I}}\mathrm{e}^{-rt}\biggr)\sum_{i=1}^{k} I_i \biggl\{\frac{|1-q_i(t)-(1-c_i\mathrm{e}^{-rt})|}{\min\{1-q_i(t), 1-c_i\,\mathrm{e}^{-rt}\}}\biggr\}, \qquad t \ge t_1.$$

Determine t_2 such that $\min_{1\le i\le k}\{\min\{1-q_i(t), 1-c_i\mathrm{e}^{-rt}\}\} \ge \frac{1}{2}$ for $t \ge t_2$. From Theorem 3.1, we have $|q_i(t) - c_i\,\mathrm{e}^{-rt}| \le K^* c_i \mathrm{e}^{-(r+\gamma)t}$, $1 \le i \le k$, for some $K^* < \infty$. Hence, for all $t \ge t_0 := \max\{t_1, t_2\}$,

$$\varepsilon^{(1)}(t) \le 2C_{\boldsymbol{I}} \exp\bigl(-\tfrac{1}{2}C_{\boldsymbol{I}}\mathrm{e}^{-rt}\bigr)K^*\mathrm{e}^{-(r+\gamma)t}$$

for γ as in Theorem 3.1. This completes the proof.

A further approximation to the last term in (4.1) is

$$1 - \prod_{i=1}^{k}(1-c_i\mathrm{e}^{-rt})^{I_i} \sim 1 - \exp(-C_{\boldsymbol{I}}\mathrm{e}^{-rt}). \tag{4.2}$$

The approximation error in (4.2) can be bounded as follows (the proof is omitted).

Lemma 4.2. *We have*

$$\biggl|\prod_{i=1}^{k}(1-c_i\,\mathrm{e}^{-rt})^{I_i} - \exp(-C_{\boldsymbol{I}}\mathrm{e}^{-rt})\biggr| \le \nu_2 C_{\boldsymbol{I}} \exp(-C_{\boldsymbol{I}}\,\mathrm{e}^{-rt})\,\mathrm{e}^{-2rt}, \qquad t \ge t_2,$$

where t_2 is as for Lemma 4.1 and $\nu_2 = \max_{1\le i\le k} c_i < \infty$.

Remark 4.1. Lemmas 4.1 and 4.2 imply that

$$\varepsilon^{(1)}(t) \le \frac{\nu_1}{C_{\boldsymbol{I}}^{\gamma/r}}\max_{x>0}\{\mathrm{e}^{-x/2}x^{1+\gamma/r}\} = \frac{\nu_3}{C_{\boldsymbol{I}}^{\gamma/r}}, \qquad t \ge t_0,$$

$$\varepsilon^{(2)}(t) \le \frac{4\nu_2\mathrm{e}^{-2}}{C_{\boldsymbol{I}}} = \frac{\nu_4}{C_{\boldsymbol{I}}}, \qquad t \ge t_2,$$

with t_0, t_2, ν_1, ν_2, and γ as before.

Definition 4.1. Define the random variable $\tilde{T}_{\boldsymbol{I}}$ such that $\mathrm{P}(\tilde{T}_{\boldsymbol{I}} > t) = 1 - \exp(-C_{\boldsymbol{I}}\mathrm{e}^{-rt})$, where $C_{\boldsymbol{I}} = \sum_{i=1}^{k} I_i c_i$. The random variable $\tilde{T}_{\boldsymbol{I}}$ satisfies

$$\tilde{T}_{\boldsymbol{I}} = \frac{\log C_{\boldsymbol{I}}}{r} + \frac{1}{r}V,$$

where V has a Gumbel distribution.

Theorem 4.1. *Suppose that* $\mathrm{E}(\|J_i\|^{1+\alpha}) < \infty$ *for some* $0 < \alpha \le 1$ *and all* $1 \le i \le k$*. Then, for* $t \ge 0$ *and with* $\gamma < r\alpha$ *as in Theorem 3.1, there exists a constant* $\nu^* < \infty$ *such that*

$$|\mathrm{P}_{\boldsymbol{I}}(T > t) - \mathrm{P}(\tilde{T}_{\boldsymbol{I}} > t)| \le \frac{\nu^*}{C_{\boldsymbol{I}}^{\gamma/r}}.$$

Proof. Remark 4.1 implies that

$$|\mathrm{P}_{\boldsymbol{I}}(T > t) - \mathrm{P}(\tilde{T}_{\boldsymbol{I}} > t)| \le \frac{\nu_3}{C_{\boldsymbol{I}}^{\gamma/r}} + \frac{\nu_4}{C_{\boldsymbol{I}}}, \qquad t \ge t_0.$$

For $t \le t_0$, we have

$$0 \le \mathrm{P}_{\boldsymbol{I}}(T \le t) \le \mathrm{P}_{\boldsymbol{I}}(T \le t_0) \le \mathrm{P}(\tilde{T}_{\boldsymbol{I}} \le t_0) + \frac{\nu_3}{C_{\boldsymbol{I}}^{\gamma/r}} + \frac{\nu_4}{C_{\boldsymbol{I}}}$$

and $0 \le \mathrm{P}(\tilde{T}_{\boldsymbol{I}} \le t) \le \mathrm{P}(\tilde{T}_{\boldsymbol{I}} \le t_0) = \exp(-C_{\boldsymbol{I}}\mathrm{e}^{-rt_0})$, completing the proof.

Theorem 4.1 thus shows that

$$d_{\mathrm{K}}\biggl(\mathcal{L}\biggl(T - \frac{\log C_{\boldsymbol{I}}}{r} \biggm| \boldsymbol{Z}(0) = \boldsymbol{I}\biggr), \mathcal{L}\biggl(\frac{V}{r}\biggr)\biggr) = O(C_{\boldsymbol{I}}^{-\gamma/r}) \quad \text{as } \|I\| \to \infty,$$

where d_K denotes the Kolmogorov distance between the two distributions indicated by $\mathcal{L}$ and γ is as in Theorem 3.1.

We now strengthen the mode of convergence. Let $\tilde{f}_{\boldsymbol{I}}$ be the probability density function of $\tilde{T}_{\boldsymbol{I}}$, and let $f_{\boldsymbol{I}}$ be the probability density function of T under $\mathrm{P}_{\boldsymbol{I}}$.

Lemma 4.3. *Suppose that* $\mathrm{E}(\|J_i\|^{1+\alpha}) < \infty$ *for some* $0 < \alpha \le 1$ *and all* $1 \le i \le k$. *For all* $t \ge t_0$, *there exists a constant* $K < \infty$ *such that*

$$|f_{\boldsymbol{I}}(t) - \tilde{f}_{\boldsymbol{I}}(t)| \le K C_{\boldsymbol{I}} \mathrm{e}^{-(r+\gamma)t}(1 + C_{\boldsymbol{I}}\mathrm{e}^{-rt})\exp\bigl(-\tfrac{1}{2}C_{\boldsymbol{I}}\mathrm{e}^{-rt}\bigr),$$

where γ *is as in Theorem 3.1 and* t_0 *is as in Lemma 4.1.*

Proof. From (4.1) we know that $\mathrm{P}_{\boldsymbol{I}}(T \le t) = \prod_{i=1}^{k}(1 - q_i(t))^{I_i}$, and, thus,

$$f_{\boldsymbol{I}}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{P}_{\boldsymbol{I}}(T \le t) = \mathrm{P}_{\boldsymbol{I}}(T \le t)\sum_{i=1}^{k}\biggl(-\frac{I_i}{1 - q_i(t)}\frac{\mathrm{d}q_i(t)}{\mathrm{d}t}\biggr). \tag{4.3}$$

Furthermore,

$$\tilde{f}_{\boldsymbol{I}}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{P}(\tilde{T}_{\boldsymbol{I}} \le t) = \mathrm{P}(\tilde{T}_{\boldsymbol{I}} \le t) r C_{\boldsymbol{I}}\mathrm{e}^{-rt}. \tag{4.4}$$

Lemmas 4.1 and 4.2 imply that, for $t \ge t_0$,

$$|\mathrm{P}_{\boldsymbol{I}}(T \le t) - \mathrm{P}(\tilde{T}_{\boldsymbol{I}} \le t)| \le K_1 C_{\boldsymbol{I}} \exp\bigl(-\tfrac{1}{2}C_{\boldsymbol{I}}\mathrm{e}^{-rt}\bigr)\mathrm{e}^{-(r+\gamma)t} \quad \text{for some } K_1 < \infty. \tag{4.5}$$

Then, also, for $t \ge t_0$,

$$\begin{aligned}\biggl|\boldsymbol{I}^{\top}\frac{\mathrm{d}\boldsymbol{q}(t)}{\mathrm{d}t} - \sum_{i=1}^{k}\frac{I_i}{1 - q_i(t)}\frac{\mathrm{d}q_i(t)}{\mathrm{d}t}\biggr| &\le \sum_{i=1}^{k} I_i \biggl|\frac{\mathrm{d}q_i(t)}{\mathrm{d}t}\biggr|\biggl|\frac{1}{1 - q_i(t)} - 1\biggr| \\ &\le K_2 C_{\boldsymbol{I}}\mathrm{e}^{-2rt}\end{aligned} \tag{4.6}$$

with $K_2 < \infty$, since Lemma 2.1 and Corollary 2.1 imply that $|\boldsymbol{I}^{\top}\mathrm{d}\boldsymbol{q}(t)/\mathrm{d}t| \le K_3 C_{\boldsymbol{I}}\mathrm{e}^{-rt}$ with $K_3 < \infty$, and, for $t \ge t_0$, $1 - q_i(t) \ge \frac{1}{2}$ for $1 \le i \le k$, implying that

$$\biggl|\frac{1}{1 - q_i(t)} - 1\biggr| \le 2q_i(t) = O(\mathrm{e}^{-rt}).$$

https://doi.org/10.1239/jap/1238592131 Published online by Cambridge University Press

Now, since $\mathrm{d}\boldsymbol{q}(t)/\mathrm{d}t = \boldsymbol{B}\boldsymbol{q}(t) - \boldsymbol{A}\boldsymbol{v}(t)$ as in (2.6), we have

$$\left|\boldsymbol{I}^\top \frac{\mathrm{d}\boldsymbol{q}(t)}{\mathrm{d}t} + rC_{\boldsymbol{I}}\mathrm{e}^{-rt}\right| = |\boldsymbol{I}^\top \boldsymbol{B}\boldsymbol{q}(t) - \boldsymbol{I}^\top \boldsymbol{A}\boldsymbol{v}(t) + rC_{\boldsymbol{I}}\mathrm{e}^{-rt}| \le K_4 C_{\boldsymbol{I}}\mathrm{e}^{-(r+\gamma)t}. \tag{4.7}$$

The final inequality in (4.7) with $K_4 < \infty$ follows because

(a) Equation (3.3) implies that

$$|\boldsymbol{I}^\top \boldsymbol{B}\boldsymbol{q}(t) - \boldsymbol{I}^\top \boldsymbol{B}\boldsymbol{b}_1 h^* \mathrm{e}^{-rt}| \le K_5 C_{\boldsymbol{I}}\mathrm{e}^{-(r+\gamma)t} \quad \text{for } t \ge 0 \text{ and } K_5 < \infty;$$

(b) Equation (3.4) and the definition of $\boldsymbol{b}_1$ give $\boldsymbol{I}^\top \boldsymbol{B}\boldsymbol{b}_1 h^* = -C_{\boldsymbol{I}} r$; and

(c) Corollary 2.1 shows that

$$|\boldsymbol{I}^\top \boldsymbol{A}\boldsymbol{v}(t)| \le K_6 C_{\boldsymbol{I}}\mathrm{e}^{-r(1+\alpha)t} \quad \text{for } t \ge 0 \text{ and } K_6 < \infty.$$

Combining (4.6) and (4.7) thus gives

$$\left| rC_{\boldsymbol{I}}\mathrm{e}^{-rt} + \sum_{i=1}^{k} \frac{I_i}{1 - q_i(t)} \frac{\mathrm{d}q_i(t)}{\mathrm{d}t} \right| \le K_7 \mathrm{e}^{-(r+\gamma)t}. \tag{4.8}$$

Using (4.5) and (4.8), together with the triangle inequality now applied to the difference of (4.3) and (4.4) in the form

$$|A_1A_2 - B_1B_2| \le |A_1 - B_1||A_2 - B_2| + |B_2||A_1 - B_1| + |B_1||A_2 - B_2|,$$

we obtain the lemma.

Using Lemma 4.3, we can show that the distribution of T under $\mathrm{P}_{\boldsymbol{I}}$ can be well approximated by that of $\tilde{T}_{\boldsymbol{I}}$ in terms of probability densities and the total variation distance d_{TV}.

Theorem 4.2. *Suppose that* $\mathrm{E}(\|J_i\|^{1+\alpha}) < \infty$ *for some* $0 < \alpha \le 1$ *and all* $1 \le i \le k$. *Then there exist constants* $K_a, K_b < \infty$ *such that*

(i) $|f_{\boldsymbol{I}}(t) - \tilde{f}_{\boldsymbol{I}}(t)| \le K_a C_{\boldsymbol{I}}^{-\gamma/r}$, $t \ge 0$;

(ii) $d_{\mathrm{TV}}(\mathcal{L}(T \mid \boldsymbol{Z}(0) = \boldsymbol{I}), \mathcal{L}(\tilde{T}_{\boldsymbol{I}})) = \frac{1}{2}\int_0^\infty |f_{\boldsymbol{I}}(t) - \tilde{f}_{\boldsymbol{I}}(t)|\,\mathrm{d}t \le K_b C_{\boldsymbol{I}}^{-\gamma/r}$.

Proof. For $t \ge t_0$, part (i) follows from Lemma 4.3, since $x^{1+\gamma/r}(1+x)\mathrm{e}^{-x/2}$ is uniformly bounded in $x \ge 0$. For $t \le t_0$, we have

$$\tilde{f}_{\boldsymbol{I}}(t) \le C_{\boldsymbol{I}} r \exp(-C_{\boldsymbol{I}} \exp(-rt_0)) = O(C_{\boldsymbol{I}}^{-s}) \quad \text{for all } s > 0;$$

similarly, from (4.1) and (4.3), it can be shown that

$$f_{\boldsymbol{I}}(t) \le KC_{\boldsymbol{I}} \exp\biggl(-\sum_{i=1}^{k} d_i I_i\biggr)$$

with $d_i = -\log(1 - q_i(t_0)) > 0$ $(1 \le i \le k)$ and $K < \infty$, which is also of order $O(C_{\boldsymbol{I}}^{-s})$ for all s, completing the proof of part (i).

For part (ii), by Lemma 4.3,

$$\int_{t_0}^{\infty} |f_{\boldsymbol{I}}(t) - \tilde{f}_{\boldsymbol{I}}(t)|\,\mathrm{d}t \le K_c C_{\boldsymbol{I}}^{-\gamma/r}$$

for $K_c < \infty$ a constant. The remaining part is bounded using part (i).

5. Application

Theorem 4.1 is illustrated by a two-type model for parasitic resistance, in which the parasite can enter a resting phase during which it does not reproduce, but can be transmitted easily to a new host. An example is the transmission cycle of the parasitic protozoa Toxoplasma gondii [6] in the intermediate hosts, which are warm-blooded. One third of the world's human population is estimated to carry a Toxoplasma infection [15]. The growth rate of a parasite population within the intermediate host can be modeled by a two-type continuous-time Markov branching process. A parasite is of type 1 if it is in the active state and of type 2 if it is in the resting state. A type 1 parasite can either die at rate d_1, enter the resting state at rate r_1, or reproduce itself by binary splitting at rate ρ. A type 2 parasite can either die at rate d_2 or become active within the host by changing to the active state at rate r_2. The transmission of the parasite to another host is incorporated in the death event. All interevent times are exponentially distributed.

Let $Z_i = Z_i(t)$ $(i = 1, 2)$ be the number of type i parasites in a host at time $t \geq 0$. The transition scheme of the process is given in Table 1.

Let $a_1 := d_1 + r_1 + \rho$ and $a_2 := d_2 + r_2$ be the total rates of transition for type 1 and type 2 individuals, respectively. For $q_i(t)$ $(i = 1, 2)$, system (2.2) yields

$$\frac{\mathrm{d}\boldsymbol{q}(t)}{\mathrm{d}t} = \begin{pmatrix} (-a_1 + 2\rho) & r_1 \\ r_2 & -a_2 \end{pmatrix} \boldsymbol{q}(t) - \begin{pmatrix} \rho q_1(t)^2 \\ 0 \end{pmatrix} = \boldsymbol{B}\boldsymbol{q}(t) - \boldsymbol{v}(t). \tag{5.1}$$

Since $\|\boldsymbol{J}\| \leq 2$ for $\boldsymbol{Z}(0) = (1, 0)^\top$ and $\boldsymbol{Z}(0) = (0, 1)^\top$, we can take $\alpha = 1$.

Theorem 2.1 implies that $\boldsymbol{B}$ has a unique real largest eigenvalue $-r$, with corresponding positive left, $\boldsymbol{f}_1^\top$, and right, $\boldsymbol{b}_1$, eigenvectors, which are given by

$$-r = \frac{-(a_1 - 2\rho + a_2) + \sqrt{D}}{2},$$

$$\boldsymbol{f}_1^\top = \frac{1}{N_1}\left(\frac{a_2 + 2\rho - a_1 + \sqrt{D}}{2r_1}, 1\right),$$

$$\text{and} \quad \boldsymbol{b}_1^\top = \frac{1}{N_2}\left(\frac{a_2 + 2\rho - a_1 + \sqrt{D}}{2r_2}, 1\right),$$

where

$$D = ((a_1 - 2\rho + a_2)^2 - 4(a_2(a_1 - 2\rho) - r_1 r_2)),$$

N_1 and N_2 are appropriate constants such that $|\boldsymbol{f}_1| = 1$ and $\boldsymbol{f}_1^\top \boldsymbol{b}_1 = 1$.

The process $(\boldsymbol{Z}(t))_{t\geq 0}$ is subcritical if and only if (i) $a_2(a_1 - 2\rho) > r_1 r_2$ and (ii) $a_1 - 2\rho + a_2 > 0$. Let the model parameters be fixed as in Table 2 such that the process is subcritical.

TABLE 1.

Transition	Rate
$Z_1 \to Z_1 - 1,\ Z_2 \to Z_2$	$d_1 Z_1$
$Z_1 \to Z_1 - 1,\ Z_2 \to Z_2 + 1$	$r_1 Z_1$
$Z_1 \to Z_1 + 1,\ Z_2 \to Z_2$	ρZ_1
$Z_1 \to Z_1,\ Z_2 \to Z_2 - 1$	$d_2 Z_2$
$Z_1 \to Z_1 + 1,\ Z_2 \to Z_2 - 1$	$r_2 Z_2$

TABLE 2: Accuracy of the extinction time approximation of the two-type Markov branching process (Table 1). For given $\boldsymbol{I} = (I_1, I_2)$, the approximation $\tilde{T}_{\boldsymbol{I}}$ is compared to T (500 000 simulations) by computing the proportion of simulated values of T larger than or equal to the median $\tilde{m}$ of $\tilde{T}_{\boldsymbol{I}}$, and by calculating the proportion of simulated values of T falling into the interquartile range (IQR) defined as the interval $(\tilde{q}_1, \tilde{q}_3)$, where $\tilde{q}_1$ and $\tilde{q}_3$ are the first and third quartiles of the approximating distribution of $\tilde{T}_{\boldsymbol{I}}$. The results are displayed for $I_1 : I_2 = 5 : 1$ and $2 : 1$ and different values of I_1. The corresponding $C_{\boldsymbol{I}} = c_1 I_1 + c_2 I_2$ are also represented. The model parameters are set to $(d_1, r_1, \rho) = (1, 1, 0.5)$ and $(d_2, r_2) = (1.2, 0.8)$ such that $-r < 0$.

$I_1 = 5I_2$	10	50	100	500	1000	5000
$C_{\boldsymbol{I}}$	9.618	48.090	96.179	480.897	961.793	4808.966
$\mathrm{P}(\geq \tilde{m})$	0.516	0.503	0.501	0.501	0.498	0.500
IQR	0.529	0.506	0.504	0.501	0.500	0.500
$I_1 = 2I_2$	10	50	100	500	1000	5000
$C_{\boldsymbol{I}}$	11.342	56.711	113.422	567.111	1134.221	5671.105
$\mathrm{P}(\geq \tilde{m})$	0.514	0.503	0.502	0.499	0.500	0.500
IQR	0.525	0.504	0.503	0.501	0.499	0.500

Thus, $r = 0.821$, $\beta_2 = 1.857$, and $\alpha = 1$, and Remark 3.1 implies that

$$q_i(t) = c_i \mathrm{e}^{-rt}(1 + O(\mathrm{e}^{-rt})).$$

Furthermore, the Kolmogorov and the total variation distances between the distributions of T given $\boldsymbol{Z}(0) = \boldsymbol{I}$ and of $\tilde{T}_{\boldsymbol{I}}$ are both of order $O(C_{\boldsymbol{I}}^{-1})$. To compute $c_i = (\boldsymbol{e}_i^\top \boldsymbol{b}_1)h^*$ $(i = 1, 2)$, it is necessary to determine h^*, given by

$$\log h^* := \lim_{t \to \infty} \{\log(\boldsymbol{f}_1^\top \boldsymbol{q}(t)) + rt\}.$$

This entails the numerical solution of system (5.1) up to a sufficient large t. To increase the numerical stability, it is advisable to solve for $\mathrm{e}^{rt}\boldsymbol{q}(t)$ instead of $\boldsymbol{q}(t)$ by appropriately reformulating (5.1). To determine an appropriate t, the reformulated system is successively solved for $t \in \{10, 11, 12, \dots\}$, and the corresponding c_1 and c_2 are evaluated until the absolute differences of successive values of c_1 and c_2 are both smaller than some predefined level, 10^{-10} in our example, resulting in $c_1 = 0.847$ and $c_2 = 0.575$ at $t = 16$.

Given $\boldsymbol{Z}(0) = \boldsymbol{I} = (I_1, I_2)$, the distribution of $\tilde{T}_{\boldsymbol{I}}$ given in Definition 4.1 can be compared with the distribution of the true extinction time T, which has to be computed by simulation, since the exact result is inaccessible. For the simulation, the Markov chain (see Table 1) can be simulated by the classical Gillespie algorithm [8] or an improved version thereof [9].

Table 2 indicates a location and a scale measure for evaluating the approximation performance. The closeness of the probabilities $\mathrm{P}_{\boldsymbol{I}}(T > \tilde{m})$ and $\mathrm{P}_{\boldsymbol{I}}(\tilde{q}_1 < T < \tilde{q}_3)$ to their limiting values 0.5, where $\tilde{m}$, $\tilde{q}_1$, and $\tilde{q}_3$ are the median, the first, and the third quartiles of the approximating distribution of $\tilde{T}_I$, increases with higher values of $C_{\boldsymbol{I}}$, which is in line with the previous results. Figure 1 represents the density function of the approximated extinction time versus the true extinction time for different initial configurations $\boldsymbol{I} = (I_1, I_2)$. The density of the approximated distribution closely matches the distribution of the simulated times, supporting the results in this paper.

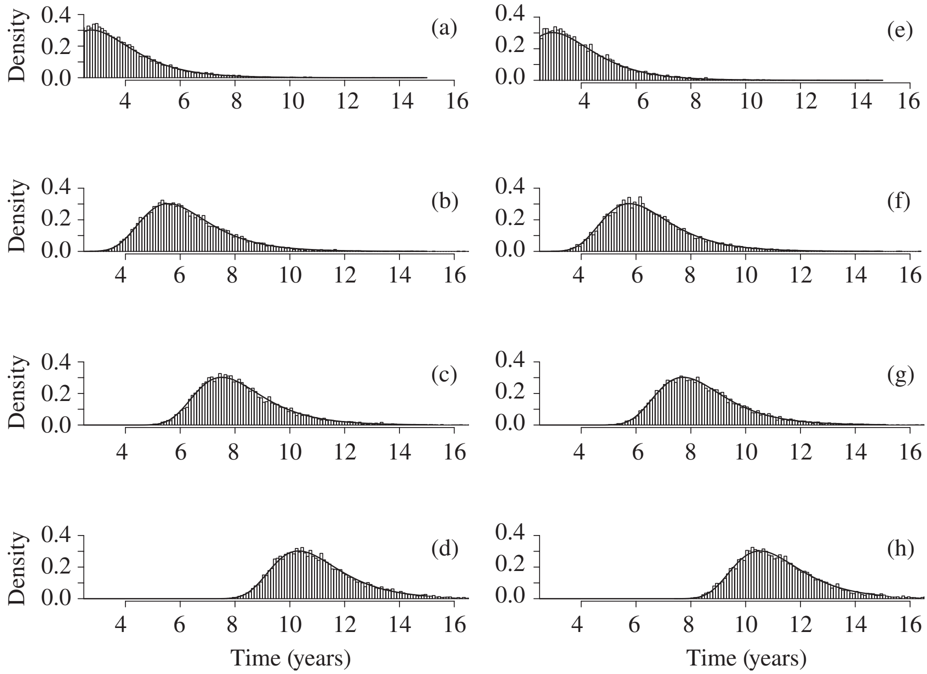

FIGURE 1: The density distribution of $\tilde{T}_{\boldsymbol{I}}$ (*solid line*) versus the simulated distribution of T (histogram of 10 000 simulations) for ratios $I_1 : I_2 = 5 : 1$ ((a)–(d)) and $2 : 1$ ((e)–(h)) with I_1 equal to 10 ((a) and (e)), 100 ((b) and (f)), 500 ((c) and (g)), and 5000 ((d) and (h)). The model parameters are fixed as in Table 2.

Acknowledgements

The author wishes to thank Andrew Barbour for many helpful suggestions and comments. The author also gratefully acknowledges the anonymous referee for suggestions that greatly improved the presentation. This work was supported by the Schweizerischer Nationalfonds (SNF), project number 107726.

References

[1] ATHREYA, K. B. AND NEY, P. E. (1972). *Branching Processes.* Springer, New York.
[2] BALL, F. (1983). The threshold behaviour of epidemic models. *J. Appl. Prob.* **20,** 227–241.
[3] BALL, F. AND DONNELLY, P. (1995). Strong approximations of epidemic models. *Stoch. Process. Appl.* **55,** 1–21.
[4] BARBOUR, A. D. (2007). Coupling a branching process to an infinite dimensional epidemic process. Preprint. Available at http://arxiv.org/abs/0710.3697v1.
[5] BARBOUR, A. D. AND UTEV, S. (2004). Approximating the Reed–Frost epidemic process. *Stoch. Process. Appl.* **113,** 173–197.
[6] ECKERT, J., FRIEDHOFF, K. T., ZAHNER, H. AND DEPLAZER, P. (eds) (2005). *Lehrbuch der Parasitologie für die Tiermedizin.* Enke Verlag, Stuttgart.
[7] GALAMBOS, J. AND SIMONELLI, I. (1996). *Bonferroni-type Inequalities with Applications.* Springer, New York.
[8] GILLESPIE, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *J. Chem. Phys.* **81,** 2340–2361.
[9] GILLESPIE, D. T. AND PETZOLD, L. R. (2003). Improved leap-size selection for accelerated stochastic simulation. *J. Chem. Phys.* **119,** 8229–8234.
[10] GRONWALL, T. H. (1919). Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *Ann. Math.* **20,** 292–296.
[11] HARRIS, T. E. (1963). *The Theory of Branching Processes.* Springer, Berlin.
[12] JAGERS, P. (1975). *Branching Processes with Biological Applications.* John Wiley, New York.

[13] Jagers, P., Klebaner, F. C. and Sagitov, S. (2007). On the path to extinction. *Proc. Nat. Acad. Sci. USA* **104,** 6107–6111.
[14] Levinson, N. (1948). The asymptotic nature of solutions of linear systems of differential equations. *Duke Math. J.* **15,** 111–126.
[15] Montoya, J. G. and Liesenfeld, O. (2004). Toxoplasmosis. *The Lancet* **363,** 1965–1976.
[16] Seneta, E. (1973). *Non-negative Matrices.* Halsted Press, New York.
[17] Sewastjanow, B. A. (1974). *Verzweigungsprozesse.* Akademie, Berlin.
[18] Whittle, P. (1955). The outcome of a stochastic epidemic—a note on Bailey's paper. *Biometrika* **42,** 116–122.