Applied Probability Trust (28 August 2020)

STOCHASTIC ORDERING FOR BIRTH-DEATH PROCESSES WITH KILLING

SHOOU-REN HSIAU,* *National Changhua University of Education*

MAY-RU CHEN,** *National Sun Yat-sen University*

YI-CHING YAO,*** *Academia Sinica*

Abstract

We consider a birth-death process with killing where transitions from state i may go to either state $i-1$ or state $i+1$ or an absorbing state (killing). Stochastic ordering results on the killing time are derived. In particular, if the killing rate in state i is monotone in i, then the distribution of the killing time with initial state i is stochastically monotone in i. This result is a consequence of the following one for a non-negative tri-diagonal matrix M: If the row sums of M are monotone, so are the row sums of M^n for all $n \geq 2$.

Keywords: Absorbing time, uniformizable chain, birth-death process with catastrophes, tri-diagonal matrix.

2000 Mathematics Subject Classification: Primary 60E15

Secondary 60J10; 60J27

* Email address: srhsiau@cc.ncue.edu.tw

* Postal address: Department of Mathematics, National Changhua University of Education, No. 1, Jin-De Rd., Changhua 500, Taiwan, R.O.C.

** Email address: mayru@faculty.nsysu.edu.tw

** Postal address: Department of Applied Math., National Sun Yat-sen University, 70 Lien-hai Rd., Kaohsiung 804, Taiwan, R.O.C.

*** Email address: yao@stat.sinica.edu.tw

*** Postal address: Institute of Statistical Science, Academia Sinica, No. 128 Academia Road, Section 2, Nankang, Taipei 11529, Taiwan, R.O.C.

1. Introduction

Motivated by genetic studies, Karlin and Tavaré [11] introduced the class of continuous-time birth-death processes with killing (or CBDKs for short) and obtained explicit results on the killing time for the case of linear CBDKs. (Earlier Puri [13] also discussed related processes.) Later, van Doorn and Zeifman [17, 18] considered more general CBDKs and derived the eventual killing probability among other things. (See also [1, Chapter 9] and [2, 3, 4, 5, 6, 16] for work on the closely related topic of birth-death processes with catastrophes.)

A CBDK $\{X_t\}$ is a continuous-time Markov chain with state space $\{0, 1, \dots\} \cup \{e\}$ where e is an absorbing state and transitions from state i ($\neq e$) may go to either $i-1$ (if $i > 0$) or $i+1$ or e with respective transition rates μ_i (death rate), λ_i (birth rate), γ_i (killing rate). We assume that $\{X_t\}$ is non-explosive. See Equations (2) and (3) of [17] for conditions on λ_i, μ_i and γ_i under which $\{X_t\}$ is non-explosive.

Let $T_e := \inf\{t \geq 0 : X_t = e\}$, the killing time. (By convention, $\inf \emptyset := \infty$.) The distribution of T_e and the probability of $T_e < \infty$ (the eventual killing probability) are of primary interest in the study of CBDKs. Denote by $\mathcal{L}_i(T_e)$ the distribution of T_e given $X_0 = i$. We write $\mathcal{L}_i(T_e) \leq_d \mathcal{L}_j(T_e)$ (or equivalently, $\mathcal{L}_j(T_e) \geq_d \mathcal{L}_i(T_e)$) if $\mathcal{L}_i(T_e)$ is stochastically smaller than $\mathcal{L}_j(T_e)$, i.e.

$$P(T_e > t \mid X_0 = i) \leq P(T_e > t \mid X_0 = j) \text{ for all } t > 0.$$

CBDKs have found applications in medicine and biology. Nagylaki [12] considered a stochastic model for a progressive chronic disease. For a patient with the disease, it is assumed that there is a useful prognostic indicator for the course of the disease and the survival of the patient. This indicator may be modeled as a birth-death process with killing. (In [12], the indicator is modeled as a pure birth process with killing under the assumption that the condition of the patient cannot improve.) For example, a patient in a coma is given a score $i \in \{3, \dots, 15\}$ (on the Glasgow Coma Scale) indicating the state of the patient's consciousness. (A lower score corresponds to a more severe condition.) With the absorbing state e denoting the patient's death, the killing time T_e is the patient's survival time. A natural question is whether the survival time of the patient with score i is stochastically nondecreasing in i, i.e. $\mathcal{L}_i(T_e) \leq_d \mathcal{L}_j(T_e)$ for $i < j$. Theorem 1 in the next section answers this question in the affirmative provided the killing rate γ_i is nonincreasing in i, regardless of the values of λ_i and μ_i. In another application of CBDKs, Hadeler [7] considered a situation where a host carries a finite number of parasites. Let X_t' denote the parasite population size within the host at time t, which may be modeled as a birth-death process. Let T denote the death time of the host, and define the process $\{X_t\}$ by $X_t = X_t'$ for $t < T$ and $X_t = e$ for $t \geq T$. Hadeler [7] proposed to model $\{X_t\}$ as a CBDK. Note that $T = T_e = \inf\{t \geq 0 : X_t = e\}$ is the survival time of the host. Again, Theorem 1 shows that the survival time of the host carrying i parasites is stochastically nonincreasing in i if the killing rate γ_i is nondecreasing in i.

The rest of this paper is organized as follows. Section 2 states the main results Theorem 1 for CBDKs and Theorem 2 for the discrete-time counterpart of CBDKs. The proofs of both theorems rely on Lemma 1 concerning the monotonicity of the row sums of powers of non-negative tri-diagonal matrices which is of independent interest. Section 3 presents numerical examples. The proofs of Theorems 1 and 2 and Lemma 1 are relegated to Section 4.

2. Main results

For a non-negative matrix $M = (M_{i,j})$, let $(M)_{i,+}$ denote the ith row sum of M, i.e. $(M)_{i,+} = \sum_j M_{i,j}$. Before presenting the main results Theorems 1 and 2, we state Lemma 1 below which is needed for the proofs of Theorems 1 and 2. All the proofs are given in Section 4.

Lemma 1. *Let $\mathbb{S} = \{0, 1, \dots, I\}$ (for some integer $I > 0$) or $\mathbb{S} = \{0, 1, \dots\}$ or $\mathbb{S} = \mathbb{Z} := \{0, \pm 1, \dots\}$. Suppose $M = (M_{i,j})_{i,j \in \mathbb{S}}$ is a non-negative tri-diagonal matrix.*

(i) *If, for some i^*, $(M)_{i,+} \leq (M)_{j,+}$ for all $i \leq i^*$ and $j > i^*$, then we have*

$$(M^n)_{i^*,+} \;\leq\; (M^n)_{i^*+1,+} \qquad \text{for all} \quad n \geq 2 \,.$$

(ii) *If, for some i^*, $(M)_{i,+} \geq (M)_{j,+}$ for all $i \leq i^*$ and $j > i^*$, then we have*

$$(M^n)_{i^*,+} \;\geq\; (M^n)_{i^*+1,+} \qquad \text{for all} \quad n \geq 2 \,.$$

(iii) *If the row sums of M are monotone, so are the row sums of M^n for all $n \geq 2$.*

Theorem 1. *Suppose $\{X_t\}$ is a non-explosive CBDK.*

(i) *If, for some* $i^* \in \{0, 1. \dots\}$, *the killing rates satisfy* $\gamma_i \leq \gamma_j$ *for all* $i \leq i^*$ *and* $j > i^*$, *then we have* $\mathcal{L}_{i^*}(T_e) \geq_d \mathcal{L}_{i^*+1}(T_e)$.

(ii) *If, for some* i^*, $\gamma_i \geq \gamma_j$ *for all* $i \leq i^*$ *and* $j > i^*$, *then we have* $\mathcal{L}_{i^*}(T_e) \leq_d \mathcal{L}_{i^*+1}(T_e)$.

(iii) *If* $\gamma_i \leq \gamma_{i+1}$ *for all* $i \geq 0$, *then we have* $\mathcal{L}_i(T_e) \geq_d \mathcal{L}_{i+1}(T_e)$ *for all* $i \geq 0$.

(iv) *If* $\gamma_i \geq \gamma_{i+1}$ *for all* $i \geq 0$, *then we have* $\mathcal{L}_i(T_e) \leq_d \mathcal{L}_{i+1}(T_e)$ *for all* $i \geq 0$.

Theorem 1 has a discrete-time analogue concerning a Markov chain $\{Y_n\}$ with state space $\{0, 1, \dots\} \cup \{e\}$ in discrete time $n = 0, 1, \dots$. Specifically, $\{Y_n\}$ has transition probabilities satisfying

$$\sum_{j=\max\{0,i-1\}}^{i+1} P(Y_{n+1} = j \mid Y_n = i) + P(Y_{n+1} = e \mid Y_n = i) = 1\ ,\ \ i = 0, 1 \dots, \tag{1}$$

and $P(Y_{n+1} = e \mid Y_n = e) = 1$. (Note that $P(Y_{n+1} = i \mid Y_n = i) > 0$ is allowed.) We refer to $\{Y_n\}$ as a discrete-time birth-death process with killing (or DBDK for short). Let $T_e^d := \inf\{n \geq 0 : Y_n = e\}$, which is the discrete-time counterpart of T_e. (By convention, $\inf \emptyset := \infty$.)

Theorem 2. *Let* $p_{i,e} := P(Y_{n+1} = e \mid Y_n = i)$, $i = 0, 1, \dots$.

(i) *If, for some* i^*, $p_{i,e} \leq p_{j,e}$ *for all* $i \leq i^*$ *and* $j > i^*$, *then we have* $\mathcal{L}_{i^*}(T_e^d) \geq_d \mathcal{L}_{i^*+1}(T_e^d)$.

(ii) *If, for some* i^*, $p_{i,e} \geq p_{j,e}$ *for all* $i \leq i^*$ *and* $j > i^*$, *then we have* $\mathcal{L}_{i^*}(T_e^d) \leq_d \mathcal{L}_{i^*+1}(T_e^d)$.

(iii) *If* $p_{i,e} \leq p_{i+1,e}$ *for all* $i \geq 0$, *then we have* $\mathcal{L}_i(T_e^d) \geq_d \mathcal{L}_{i+1}(T_e^d)$ *for all* $i \geq 0$.

(iv) *If* $p_{i,e} \geq p_{i+1,e}$ *for all* $i \geq 0$, *then we have* $\mathcal{L}_i(T_e^d) \leq_d \mathcal{L}_{i+1}(T_e^d)$ *for all* $i \geq 0$.

Remark 1. By Theorem 1, we have $\mathcal{L}_i(T_e)$ stochastically monotone in i provided the killing rate γ_i is monotone in i, regardless of the birth and death rates λ_i and μ_i. The fact that the process $\{X_t\}$ is skip-free to the left and to the right (apart from jumping into the absorbing state e) is crucial for Theorem 1 to hold. In a different context, Irle and Gani [10] and Irle [9] obtained level-crossing stochastic ordering results for Markov chains and semi-Markov processes which are skip-free to the right.

Remark 2. He and Chavoushi [8] considered queueing systems with customer interjections in which two parameters (denoted by η_I and η_C) are introduced to describe the interjection behavior. They investigated the effects of the two parameters on the customer's waiting time, and derived some monotonicity properties of the distribution of the waiting time and its mean and variance in terms of the parameters. (The proof of their Lemma 2.2 contains a special case of our Lemma 1.) In particular, with $W_n(\eta_I, \eta_C)$ denoting the waiting time of a customer initially in position n in the queue, they showed that $W_n(\eta_I, \eta_C) \leq_d W_n(\eta_I', \eta_C')$ if $\eta_I \leq \eta_I'$ and $\eta_C \leq \eta_C'$. Note that two different pairs of parameter values (η_I, η_C) and (η_I', η_C') correspond to two different transition rate matrices (with the same state space). In contrast, our results are concerned with a single transition rate matrix with an absorbing state and compare the absorbing time distributions when the Markov chain starts from different states.

3. Numerical examples

In this section we illustrate Theorems 1 and 2 with numerical examples. We first consider a DBDK $\{Y_n\}$ with transition probabilities given by

$$\begin{aligned}
P(Y_{n+1} = e \mid Y_n = i) &= 0.01, \quad i = 0, 1, 2;\\
P(Y_{n+1} = e \mid Y_n = i) &= 0.02, \quad i \geq 3;\\
P(Y_{n+1} = 1 \mid Y_n = 0) &= 0.99 = 1 - P(Y_{n+1} = e \mid Y_n = 0);\\
P(Y_{n+1} = i+1 \mid Y_n = i) &= 0.5, \quad i \geq 1;\\
P(Y_{n+1} = i-1 \mid Y_n = i) &= 0.5 - P(Y_{n+1} = e \mid Y_n = i), \quad i \geq 1.
\end{aligned}$$

Figure 1 plots the survival probabilities

$$P(T_e^d > n \mid Y_0 = i) = (M^n)_{i,+} = \sum_{j=0}^{\infty} M_{i,j}^n, \quad \text{for } 0 \leq n \leq 160 \text{ and } i = 0, 1, 2, 3,$$

where the matrix $M = (M_{i,j})$ is given by $M_{i,j} = P(Y_1 = j \mid Y_0 = i)$ for $i, j = 0, 1, \ldots$. It shows that $P(T_e^d > n \mid Y_0 = i)$ decreases as i increases as expected in view of Theorem 2(iii).

For a linear CBDK $\{X_t\}$ with $\mu_i = ai$, $\lambda_i = bi + \theta$, and $\gamma_i = ci$ where a, b, c and θ are positive constants, Karlin and Tavaré [11] derived the explicit formula

$$P(T_e > t \mid X_0 = i) = e^{-\theta(1-v_0)t} \Big[\frac{v_0(v_1 - 1) + v_1\sigma_t(1 - v_0)}{v_1 - 1 + \sigma_t(1 - v_0)}\Big]^i \Big[\frac{v_1 - 1 + \sigma_t(1 - v_0)}{v_1 - v_0}\Big]^{-\theta/b},$$

where $0 < v_0 < 1 < v_1$ are the two roots of the equation $bx^2 - (a + b + c)x + a = 0$ and $\sigma_t = e^{-b(v_1 - v_0)t}$. Since γ_i is increasing in i, $P(T_e > t \mid X_0 = i)$ should be decreasing in i by

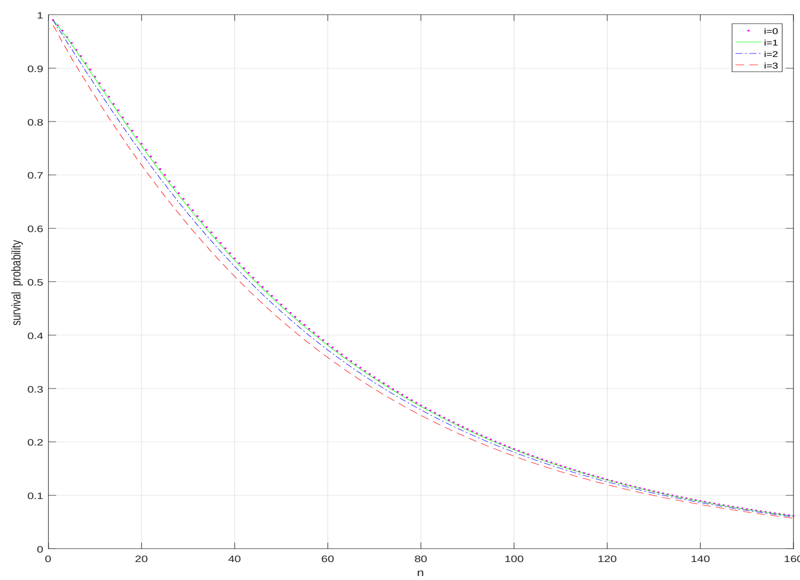

FIGURE 1: The survival probabilities $P(T_e^d > n \mid Y_0 = i)$, $i = 0, 1, 2, 3$.

Theorem 1(iii). Noting that $v_0(v_1 - 1) + v_1\sigma_t(1 - v_0) < v_1 - 1 + \sigma_t(1 - v_0)$ (since $0 < \sigma_t < 1$), we have $P(T_e > t \mid X_0 = i)$ decays geometrically in i. Figure 2 plots $P(T_e > t \mid X_0 = i)$ for $0 < t < 5$ and $i = 0, 2, 4, 6, 8$ with $a = b = c = \theta = 1$.

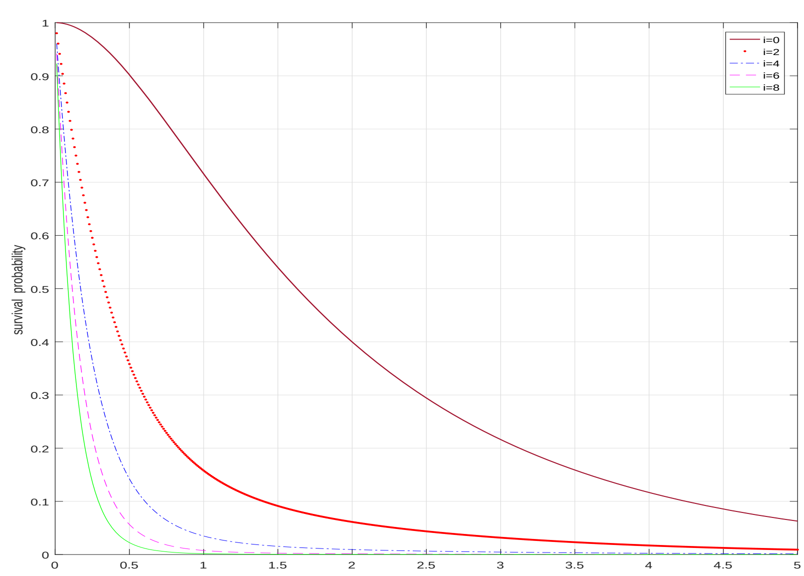

FIGURE 2: The survival probabilities $P(T_e > t \mid X_0 = i)$, $i = 0, 2, 4, 6, 8$.

While no explicit formula for $P(T_e > t \mid X_0 = i)$ is available for general CBDKs, we may use the technique of uniformization (*cf.* [14, Section 5.10] and [15, Section 6.7]) to compute $P(T_e > t \mid X_0 = i)$ if $\{X_t\}$ is uniformizable (i.e. $\sup_i(\lambda_i + \mu_i + \gamma_i) \leq C < \infty$ for some constant $C > 0$). Specifically, with $\mu_0 := 0$, let $\{Z_n\}$ be a DBDK with transition probabilities satisfying $P(Z_{n+1} = e \mid Z_n = e) = 1$, and for $i \geq 0$

$$
\begin{aligned}
&P(Z_{n+1} = i-1 \mid Z_n = i) = \frac{\mu_i}{C}\ , \quad P(Z_{n+1} = i+1 \mid Z_n = i) = \frac{\lambda_i}{C},\\
&P(Z_{n+1} = e \mid Z_n = i) = \frac{\gamma_i}{C}\ , \quad P(Z_{n+1} = i \mid Z_n = i) = 1 - \frac{\mu_i + \lambda_i + \gamma_i}{C}.
\end{aligned}
$$

Let $\{N(t)\}$ be a Poisson process of constant rate C, independent of $\{Z_n\}$. Then we have the following result :

$$
\textit{Given } X_0 = Z_0, \textit{ the two processes } \{X_t\} \textit{ and } \{Z_{N(t)}\} \textit{ have the same distribution.} \tag{2}
$$

In other words, the CBDK $\{X_t\}$ may be constructed by placing the transitions of the DBDK $\{Z_n\}$ at Poisson arrival epochs. It follows from (2) that

$$
\begin{aligned}
P(T_e > t \mid X_0 = i) &= P(X_t \neq e \mid X_0 = i)\\
&= P(Z_{N(t)} \neq e \mid Z_0 = i)\\
&= \sum_{n=0}^{\infty} P(Z_n \neq e\ , N(t) = n \mid Z_0 = i)\\
&= \sum_{n=0}^{\infty} P(N(t) = n) P(Z_n \neq e \mid Z_0 = i)\\
&= \sum_{n=0}^{\infty} \frac{e^{-Ct}(Ct)^n}{n!} (M^{*n})_{i,+},
\end{aligned}
$$

where the matrix $M^* = (M^*_{i,j})$ is given by $M^*_{i,j} = P(Z_1 = j \mid Z_0 = i)$. So $P(T_e > t \mid X_0 = i)$ may be approximated by $\sum_{n=0}^{L(t)} \frac{e^{-Ct}(Ct)^n}{n!}(M^{*n})_{i,+}$ for sufficiently large integer $L(t)$ (depending on t). The distributional equivalence of $\{X_t\}$ and $\{Z_{N(t)}\}$ in (2) will be called for in the proof of Theorem 1 in the next section.

4. Proofs of Theorems 1 and 2 and Lemma 1

To prove Theorem 2, let $M = (M_{i,j})_{i,j\in\{0,1,\dots\}}$ be the transition matrix of $\{Y_n\}$ restricted to the states $0, 1, \dots$. That is,

$$M_{i,j} = P(Y_{n+1} = j \mid Y_n = i)\ , \quad i, j = 0, 1, \dots .$$

By (1), we have that $M_{i,j} = 0$ for $|i - j| > 1$ and that

$$(M)_{i,+} = \sum_{j=\max\{0,i-1\}}^{i+1} M_{i,j} = 1 - P(Y_{n+1} = e \mid Y_n = i) = 1 - p_{i,e}\ . \tag{3}$$

Moreover,

$$P(T_e^d > n \mid Y_0 = i) = \sum_{j=0}^{\infty} P(Y_n = j \mid Y_0 = i) = (M^n)_{i,+}\ . \tag{4}$$

In view of (3) and (4), Theorem 2 follows immediately from Lemma 1. To prove Theorem 1, one may approximate the CBDK $\{X_t\}$ by a sequence of DBDKs via time-discretization and argue that T_e is the limit of the corresponding T_e^d's. Instead of taking this approach, we first consider uniformizable $\{X_t\}$ (i.e. $\sup_i\ (\lambda_i+\mu_i+\gamma_i) < \infty$) for which the associated (embedded) discrete-time Markov chain is a DBDK, so that the desired results follow from Theorem 2. The general CBDK $\{X_t\}$ is then approximated by a sequence of uniformizable CBDKs. The details are given in the proof below.

Proof of Theorem 1. We first prove part (i) under the additional assumption that $\{X_t\}$ is uniformizable, i.e. $\sup_i\ (\lambda_i + \mu_i + \gamma_i) \le C < \infty$ for some $C > 0$. Let $\{Z_n\}$ and $\{N(t)\}$ be, respectively, the DBDK and (independent) Poisson process of constant rate C as described in the last paragraph of Section 3. Letting $p_{i,e}^{(Z)} := P(Z_{n+1} = e \mid Z_n = i) = \gamma_i/C$ and $T_e^{d(Z)} := \inf\{n \ge 0 : Z_n = e\}$, we have that $\{Z_n\}$ is a DBDK with $p_{i,e}^{(Z)} \le p_{j,e}^{(Z)}$ for $i \le i^*$ and $j > i^*$, so that by Theorem 2(i),

$$\begin{aligned} P(Z_n \neq e \mid Z_0 = i^*) &= P(T_e^{d(Z)} > n \mid Z_0 = i^*) \\ &\ge P(T_e^{d(Z)} > n \mid Z_0 = i^* + 1) = P(Z_n \neq e \mid Z_0 = i^* + 1)\ . \end{aligned} \tag{5}$$

Consequently, for $t > 0$,

$$\begin{aligned} P(T_e > t \mid X_0 = i^*) &= P(Z_{N(t)} \neq e \mid Z_0 = i^*) \ \ (\text{by } (2)) \\ &= \sum_{n=0}^{\infty} P(N(t) = n) P(Z_n \neq e \mid Z_0 = i^*) \\ &\geq \sum_{n=0}^{\infty} P(N(t) = n) P(Z_n \neq e \mid Z_0 = i^* + 1) \ \ (\text{by } (5)) \\ &= P(Z_{N(t)} \neq e \mid Z_0 = i^* + 1) \\ &= P(T_e > t \mid X_0 = i^* + 1) \ . \end{aligned}$$

This proves part (i) for uniformizable $\{X_t\}$.

For general (non-explosive) $\{X_t\}$, to prove $P(T_e > t_0 \mid X_0 = i^*) \geq P(T_e > t_0 \mid X_0 = i^* + 1)$ for any (fixed) $t_0 > 0$, let

$$a_k := P(T_k \leq t_0 \mid X_0 = i^*) \ \text{ and } \ b_k := P(T_k \leq t_0 \mid X_0 = i^* + 1) \ , \tag{6}$$

where $T_k := \inf\{t \geq 0 : X_t = k\}$. For X_t to reach a (large) state k starting from i^* or $i^* + 1$, at least $k - i^* - 1$ transitions are required. Since $\{X_t\}$ is non-explosive, we have

$$\lim_{k\to\infty} a_k = 0 \ \text{ and } \ \lim_{k\to\infty} b_k = 0 \ . \tag{7}$$

For each $k > i^*$, let $\{X_t^{(k)}\}$ be a CBDK whose birth, death and killing rates in state i are given by $\lambda_i^{(k)} := \lambda_{i\wedge k}$, $\mu_i^{(k)} := \mu_{i\wedge k}$, $\gamma_i^{(k)} := \gamma_{i\wedge k}$ where $i \wedge k := \min\{i, k\}$. Clearly, $\{X_t^{(k)}\}$ is uniformizable with the killing rates satisfying $\gamma_i^{(k)} \leq \gamma_j^{(k)}$ for $i \leq i^*$ and $j > i^*$. Denote by $\mathcal{L}_i(T_k)$ ($\mathcal{L}_i(T_e \wedge T_k)$, *resp.*) the distribution of T_k ($T_e \wedge T_k$, *resp.*) given $X_0 = i$. Similarly, denote by $\mathcal{L}_i(T_k^{(k)})$ ($\mathcal{L}_i(T_e^{(k)} \wedge T_k^{(k)})$, *resp.*) the distribution of $T_k^{(k)}$ ($T_e^{(k)} \wedge T_k^{(k)}$, *resp.*) given $X_0^{(k)} = i$, where $T_r^{(k)} := \inf\{t \geq 0 : X_t^{(k)} = r\}$ for $r = k,\ e$.

For $k > i^*$, given $X_0 = i^*$ or $i^* + 1$, the distributions of T_k and $T_e \wedge T_k$ depend only on the values of $\lambda_i, \mu_i, \gamma_i$ for $i < k$. Similarly, given $X_0^{(k)} = i^*$ or $i^* + 1$, the distributions of $T_k^{(k)}$ and $T_e^{(k)} \wedge T_k^{(k)}$ depend only on the values of $\lambda_i^{(k)}, \mu_i^{(k)}, \gamma_i^{(k)}$ for $i < k$. Since $\lambda_i = \lambda_i^{(k)}, \mu_i = \mu_i^{(k)}$ and $\gamma_i = \gamma_i^{(k)}$ for $i \leq k$, we have that

$$\mathcal{L}_i(T_k) = \mathcal{L}_i(T_k^{(k)}) \ \text{ and } \ \mathcal{L}_i(T_e \wedge T_k) = \mathcal{L}_i(T_e^{(k)} \wedge T_k^{(k)}) \ \text{ for } \ i = i^*, i^* + 1 \ \text{ and } \ k > i^*,$$

implying that for $k > i^*$,

$$P(T_k > t_0 \mid X_0 = i) = P(T_k^{(k)} > t_0 \mid X_0^{(k)} = i) \; , \quad i = i^*, i^*+1 \tag{8}$$

$$P(T_e \wedge T_k > t_0 \mid X_0 = i) = P(T_e^{(k)} \wedge T_k^{(k)} > t_0 \mid X_0^{(k)} = i) \; , \quad i = i^*, i^*+1 \; . \tag{9}$$

Since $\{X_t^{(k)}\}$ is uniformizable with $\gamma_i^{(k)} \le \gamma_j^{(k)}$ for $i \le i^*$ and $j > i^*$, we have shown that

$$P(T_e^{(k)} > t_0 \mid X_0^{(k)} = i^*) \ge P(T_e^{(k)} > t_0 \mid X_0^{(k)} = i^*+1) \;\; \text{for} \;\; k > i^*. \tag{10}$$

Also for $k > i^*$ and $i = i^*, i^*+1$,

$$\begin{aligned} P(T_e > t_0 \mid X_0 = i) &\ge P(T_e \wedge T_k > t_0 \mid X_0 = i) \\ &= P(T_e^{(k)} \wedge T_k^{(k)} > t_0 \mid X_0^{(k)} = i) \;\; (\text{by} \;\; (9)) \; , \end{aligned}$$

and

$$\begin{aligned} P(T_e > t_0 \mid X_0 = i) &\le P(T_e \wedge T_k > t_0 \mid X_0 = i) + P(T_k \le t_0 \mid X_0 = i) \\ &= P(T_e^{(k)} \wedge T_k^{(k)} > t_0 \mid X_0^{(k)} = i) + P(T_k \le t_0 \mid X_0 = i) \;\; (\text{by} \;\; (9)) \\ &\le P(T_e^{(k)} \wedge T_k^{(k)} > t_0 \mid X_0^{(k)} = i) + a_k \vee b_k \; , \end{aligned}$$

where a_k and b_k are as defined in (6) and $a_k \vee b_k := \max\{a_k, b_k\}$. Consequently, for $k > i^*$,

$$0 \le P(T_e > t_0 \mid X_0 = i) - P(T_e^{(k)} \wedge T_k^{(k)} > t_0 \mid X_0^{(k)} = i) \le a_k \vee b_k \; , \qquad i = i^*, i^*+1. \tag{11}$$

Furthermore, for $k > i^*$ and $i = i^*, i^*+1$,

$$\begin{aligned} &0 \le P(T_e^{(k)} > t_0 \mid X_0^{(k)} = i) - P(T_e^{(k)} \wedge T_k^{(k)} > t_0 \mid X_0^{(k)} = i) \\ &\;\; \le P(T_k^{(k)} \le t_0 \mid X_0^{(k)} = i) \\ &\;\; = P(T_k \le t_0 \mid X_0 = i) \;\; (\text{by} \;\; (8)) \\ &\;\; \le a_k \vee b_k \; , \end{aligned}$$

which together with (11) implies that for $k > i^*$ and $i = i^*, i^*+1$,

$$|P(T_e > t_0 \mid X_0 = i) - P(T_e^{(k)} > t_0 \mid X_0^{(k)} = i)| \le a_k \vee b_k \; . \tag{12}$$

It follows from (7), (10) and (12) that

$$P(T_e > t_0 \mid X_0 = i^*) \geq P(T_e > t_0 \mid X_0 = i^* + 1),$$

completing the proof of part (i).

Part (ii) can be proved by using a similar argument and invoking Theorem 2(ii). Parts (iii) and (iv) follow immediately from parts (i) and (ii), respectively.

Proof of Lemma 1. Note that part (ii) follows from part (i) by reversing the ordering of the rows and that of the columns, and that part (iii) is a consequence of parts (i) and (ii). It remains to prove part (i). It suffices to deal only with the case $\mathbb{S} = \mathbb{Z}$, since the cases $\mathbb{S} = \{0, 1, \dots, I\}$ and $\mathbb{S} = \{0, 1, \dots\}$ can be treated as special cases of $\mathbb{S} = \mathbb{Z}$. More precisely, for $M = (M_{i,j})_{i,j \in \{0,1,\dots,I\}}$ satisfying $(M)_{i,+} \leq (M)_{j,+}$ for $i \leq i^* < j$, define $\widetilde{M} = (\widetilde{M}_{i,j})_{i,j \in \mathbb{Z}}$ by

$$\widetilde{M}_{i,j} = M_{i,j} \mathbf{1}_{\{0 \leq i,j \leq I\}} + \ (M)_{i^*,+} \ \ \delta_{i,j} \mathbf{1}_{\{i<0\}} + \ (M)_{i^*+1,+} \ \ \delta_{i,j} \mathbf{1}_{\{i>I\}} \ ,$$

where $\mathbf{1}_A$ denotes the indicator function of a set A and $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ otherwise. We have $(\widetilde{M})_{i,+} \leq (\widetilde{M})_{j,+}$ for $i \leq i^* < j$. Moreover, $(\widetilde{M}^n)_{i,+} = (M^n)_{i,+}$ for $i = 0, 1, \dots, I$.

We now prove part (i) for $\mathbb{S} = \mathbb{Z}$. Without loss of generality, we assume $i^* = 0$, so that the tri-diagonal matrix M satisfies $M_{i,j} \geq 0$ for $i, j \in \mathbb{Z}$, $M_{i,j} = 0$ if $|i - j| > 1$, and $(M)_{i,+} \leq (M)_{j,+}$ for $i \leq 0 < j$. To show that $(M^n)_{0,+} \leq (M^n)_{1,+}$ for any (fixed) $n \geq 2$, note that $(M^n)_{0,+}$ and $(M^n)_{1,+}$ do not depend on the values of $M_{i,i-1}$, $M_{i,i}$ and $M_{i,i+1}$ for $i \leq -n$ or $i \geq n+1$. Thus $(M^n)_{0,+} = (\overline{M}^n)_{0,+}$ and $(M^n)_{1,+} = (\overline{M}^n)_{1,+}$, where $\overline{M} = (\overline{M}_{i,j})$ is defined by

$$\overline{M}_{i,j} = M_{i,j} \mathbf{1}_{\{-n<i\leq n\}} + \ (M)_{0,+} \ \ \delta_{i,j} \mathbf{1}_{\{i\leq -n\}} + \ (M)_{1,+} \ \ \delta_{i,j} \mathbf{1}_{\{i>n\}}.$$

Note that $\overline{M}$ has bounded row sums and satisfies $(\overline{M})_{i,+} \leq (\overline{M})_{j,+}$ for $i \leq 0 < j$. If we can show part (i) of the theorem for non-negative tri-diagonal matrices with bounded row sums, then we have

$$(M^n)_{0,+} = (\overline{M}^n)_{0,+} \leq (\overline{M}^n)_{1,+} = (M^n)_{1,+} \ .$$

So it suffices to establish part (i) with M having bounded row sums. We may further assume that the row sums of M are bounded by 1.

To show that $(M^n)_{0,+} \ \le \ (M^n)_{1,+}$ for $n \ge 2$, we introduce a Markov chain $\{X_n : n = 0, 1, \ldots\}$ with state space $\mathbb{Z} \cup \{e\}$ and transition probabilities given by

$$\begin{aligned} P(X_{n+1} = j \mid X_n = i) &= M_{i,j} \ , \qquad i, j \in \mathbb{Z} \ , \\ P(X_{n+1} = e \mid X_n = i) &= 1 - (M)_{i,+} \ , \qquad i \in \mathbb{Z} \ , \\ P(X_{n+1} = e \mid X_n = e) &= 1 \ . \end{aligned}$$

Note that the state e is absorbing. Let $T_1 = \inf\{n \ge 0 : X_n = 1\}$ and $T_e = \inf\{n \ge 0 : X_n = e\}$. (In Theorem 1, T_e is defined with respect to the continuous-time process $\{X_t\}$. Here the same notation T_e is used with respect to the discrete-time process $\{X_n\}$. This proof does not involve the continuous-time process $\{X_t\}$.)

We write $P_i(\ \cdot\ ) = P(\ \cdot \mid X_0 = i)$, and claim that for $n = 1, 2, \ldots$, and $j = 1, 2, \ldots$,

$$P_0(T_1 \ge n,\ T_e > n) \ + \ \sum_{\ell=1}^{n-1} P_0(T_1 = \ell)\ P_j(T_e > n - \ell) \ \le \ P_j(T_e > n). \tag{13}$$

Note that for $1 \le \ell < n$,

$$\begin{aligned} P_0(T_1 = \ell,\ T_e > n) &= P(T_1 = \ell,\ T_e > n \mid X_0 = 0) \\ &= P(T_1 = \ell \mid X_0 = 0)\ P(T_e > n \mid T_1 = \ell,\ X_0 = 0) \\ &= P_0(T_1 = \ell)\ P(T_e > n \mid X_\ell = 1) \\ &= P_0(T_1 = \ell)\ P(T_e > n - \ell \mid X_0 = 1) \\ &= P_0(T_1 = \ell)\ P_1(T_e > n - \ell)\,. \end{aligned} \tag{14}$$

By (14), the left-hand side of (13) with $j = 1$ equals

$$\begin{aligned} & P_0(T_1 \ge n, T_e > n) \ + \ \sum_{\ell=1}^{n-1} P_0(T_1 = \ell)\ P_1(T_e > n - \ell) \\ &= P_0(T_1 \ge n,\ T_e > n) \ + \ \sum_{\ell=1}^{n-1} P_0(T_1 = \ell,\ T_e > n) \\ &= P_0(T_e > n) \ . \end{aligned}$$

Thus the inequality (13) with $j = 1$ is equivalent to

$$P_0(T_e > n) \ \le \ P_1(T_e > n) \ . \tag{15}$$

Since $P_i(T_e > n) \;=\; P(X_n \in \mathbb{Z} \mid X_0 = i) \;=\; (M^n)_{i,+}$, (15) is equivalent to $(M^n)_{0,+} \;\leq\; (M^n)_{1,+}$.

We now prove (13) by induction on n. For $n = 1$, the left-hand side of (13) equals

$$P_0(T_1 \geq 1,\ T_e > 1) = P_0(T_e > 1) = (M)_{0,+},$$

while the right-hand side equals $P_j(T_e > 1) = (M)_{j,+}$. Thus the inequality (13) with $n = 1$ follows from the assumption that $(M)_{0,+} \leq (M)_{j,+}$ for $j > 0$.

For $m \geq 1$, suppose (13) holds for $n = 1, \ldots, m$ and $j = 1, 2, \ldots$. In particular, the induction hypothesis implies (*cf.* (15)) that

$$P_0(T_e > \ell) \;\leq\; P_1(T_e > \ell) \;, \qquad \ell = 1, \ldots, m \;. \tag{16}$$

We need to show that for $j = 1, 2, \ldots$,

$$P_0(T_1 \geq m+1,\ T_e > m+1) \;+\; \sum_{\ell=1}^{m} P_0(T_1 = \ell)\ P_j(T_e > m+1-\ell) \;\leq\; P_j(T_e > m+1) \;. \tag{17}$$

Note that for $j = 1, 2, \ldots,$

$$P_0(T_1 \geq m,\ T_e > m) \;+\; \sum_{\ell=1}^{m-1} P_0(T_1 = \ell)\ P_{j-1}(T_e > m-\ell) \;\leq\; P_{j-1}(T_e > m) \;, \tag{18}$$

$$P_0(T_1 \geq m,\ T_e > m) \;+\; \sum_{\ell=1}^{m-1} P_0(T_1 = \ell)\ P_j(T_e > m-\ell) \;\leq\; P_j(T_e > m) \;, \tag{19}$$

$$P_0(T_1 \geq m,\ T_e > m) \;+\; \sum_{\ell=1}^{m-1} P_0(T_1 = \ell)\ P_{j+1}(T_e > m-\ell) \;\leq\; P_{j+1}(T_e > m) \;. \tag{20}$$

Except for $j = 1$ in (18), (18)–(20) follow immediately from the induction hypothesis. By (16), $P_0(T_e > m-\ell) \;\leq\; P_1(T_e > m-\ell)$ for $\ell = 1, \ldots, m-1$, so that the left-hand side of (18) with $j = 1$ equals

$$\begin{aligned}
&P_0(T_1 \geq m,\ T_e > m) \;+\; \sum_{\ell=1}^{m-1} P_0(T_1 = \ell)\ P_0(T_e > m-\ell) \\
&\quad\leq\; P_0(T_1 \geq m,\ T_e > m) \;+\; \sum_{\ell=1}^{m-1} P_0(T_1 = \ell)\ P_1(T_e > m-\ell) \\
&\quad=\; P_0(T_1 \geq m,\ T_e > m) \;+\; \sum_{\ell=1}^{m-1} P_0(T_1 = \ell,\ T_e > m) \;\; \text{(by (14))} \\
&\quad=\; P_0(T_e > m) \;,
\end{aligned}$$

establishing (18) for $j = 1$.

Note that for $\ell = m$,

$$\begin{aligned}
&P_0(T_1 \wedge T_e \ \geq m+1) \ + \ P_0(T_1 = \ell) \ P_{j-1}(T_e > m - \ell) \\
&\quad = \ P_0(T_1 \wedge T_e \ \geq m+1) \ + \ P_0(T_1 = m) \ P_{j-1}(T_e > 0) \\
&\quad = \ P_0(T_1 \wedge T_e \ \geq m+1) \ + \ P_0(T_1 = m) \\
&\quad = \ P_0(T_1 \geq m, \ T_e > m) \ ,
\end{aligned}$$

so that (18) is equivalent to

$$P_0(T_1 \wedge T_e \ \geq m+1) \ + \ \sum_{\ell=1}^{m} P_0(T_1 = \ell) \ P_{j-1}(T_e > m - \ell) \ \leq \ P_{j-1}(T_e > m) \ . \qquad (21)$$

Similarly, (19) and (20) are, respectively, equivalent to

$$P_0(T_1 \wedge T_e \ \geq m+1) \ + \ \sum_{\ell=1}^{m} P_0(T_1 = \ell) \ P_{j}(T_e > m - \ell) \ \leq \ P_{j}(T_e > m) \ , \qquad (22)$$

$$P_0(T_1 \wedge T_e \ \geq m+1) \ + \ \sum_{\ell=1}^{m} P_0(T_1 = \ell) \ P_{j+1}(T_e > m - \ell) \ \leq \ P_{j+1}(T_e > m) \ . \qquad (23)$$

Letting $a_j = M_{j,j-1}$, $b_j = M_{j,j}$ and $c_j = M_{j,j+1}$, we have

$$\begin{aligned}
P_j(T_e > m+1) \ &= \ P(T_e > m+1 \mid X_0 = j) \\
&= \ P(X_1 \neq e, \ T_e > m+1 \mid X_0 = j) \\
&= \ \sum_{r=j-1}^{j+1} P(X_1 = r, \ T_e > m+1 \mid X_0 = j) \\
&= \ a_j P(T_e > m+1 \mid X_1 = j-1) \ + \ b_j P(T_e > m+1 \mid X_1 = j) \\
&\qquad + \ c_j P(T_e > m+1 \mid X_1 = j+1) \\
&= \ a_j P_{j-1}(T_e > m) \ + \ b_j P_j(T_e > m) \ + \ c_j P_{j+1}(T_e > m) \ . \qquad (24)
\end{aligned}$$

Similarly, for $1 \leq \ell \leq m$,

$$P_j(T_e > m+1-\ell) = a_j P_{j-1}(T_e > m-\ell) + b_j P_j(T_e > m-\ell) + c_j P_{j+1}(T_e > m-\ell) \ . \qquad (25)$$

In view of (24) and (25), it follows from (21)–(23) that

$$\begin{aligned}
(a_j + b_j + c_j) P_0(T_1 \wedge T_e \ \geq m+1) + \sum_{\ell=1}^{m} P_0(T_1 = \ell) P_j(T_e > m+1-\ell) \\
\leq \ P_j(T_e > m+1). \qquad (26)
\end{aligned}$$

Note that given $X_0 = 0$, $\{T_1 \geq m+1, T_e > m+1\} = \{X_1 < 1, \dots, X_m < 1, X_{m+1} \neq e\}$ and

$$\{X_1 < 1, \dots, X_m < 1\} = \{X_\ell \in \{0, -1, -2, \dots\},\ \ell = 1, \dots, m\} = \{T_1 \wedge T_e \geq m+1\}.$$

We have

$$\begin{aligned}
& P_0(T_1 \geq m+1,\ T_e > m+1) \\
& = P(X_1 < 1, \dots, X_m < 1,\ X_{m+1} \neq e \mid X_0 = 0) \\
& = P(X_1 < 1, \dots, X_m < 1 \mid X_0 = 0)\ P(X_{m+1} \neq e \mid X_0 = 0,\ X_1 < 1, \dots, X_m < 1) \\
& = P_0(T_1 \wedge T_e\ \geq m+1)\ P(X_{m+1} \neq e \mid X_0 = 0,\ X_1 < 1, \dots, X_m < 1) \\
& \leq\ P_0(T_1 \wedge T_e\ \geq m+1)\ (a_j + b_j + c_j)\ . \qquad (27)
\end{aligned}$$

The above inequality follows since

$$\begin{aligned}
& P(X_{m+1} \neq e \mid X_0 = 0,\ X_1 < 1, \dots, X_m < 1) \\
& \quad = \sum_{i \leq 0} P(X_m = i, X_{m+1} \neq e \mid X_0 = 0,\ X_1 < 1, \dots, X_m < 1) \\
& \quad = \sum_{i \leq 0} P(X_m = i \mid X_0 = 0,\ X_1 < 1, \dots, X_m < 1)\ P(X_{m+1} \neq e \mid X_m = i) \\
& \quad = \sum_{i \leq 0} P(X_m = i \mid X_0 = 0,\ X_1 < 1, \dots, X_m < 1)\ (a_i + b_i + c_i) \\
& \quad \leq \sum_{i \leq 0} P(X_m = i \mid X_0 = 0,\ X_1 < 1, \dots, X_m < 1)\ (a_j + b_j + c_j) \\
& \quad = a_j + b_j + c_j\ ,
\end{aligned}$$

where the inequality is due to the assumption that $a_i + b_i + c_i \leq a_j + b_j + c_j$ for $i \leq 0 < j$.

Finally (17) follows from (26) and (27). The proof is complete.

Acknowledgment. The authors are grateful to the Editor and two reviewers for their helpful comments leading to an improved presentation. The authors also gratefully acknowledge support by the Ministry of Science and Technology, Taiwan, ROC.

References

[1] Anderson, W.J. (1991). *Continuous-Time Markov Chains*. Springer, New York.

[2] Brockwell, P.J. (1985). The extinction time of a birth, death and catastrophe process and a related diffusion model. *Adv. Appl. Probab.* **17**, 42–52.

[3] Brockwell, P.J. (1986). The extinction time of a general birth and death process with catastrophes. *J. Appl. Probab.* **23**, 851–858.

[4] Chao, X. and Zheng, Y. (2003). Transient analysis of immigration birth-death processes with total catastrophes. *Probab. Engin. Inform. Sci.* **17**, 83–116.

[5] Chen, A., Zhang, H., Liu, K. and Rennolls, K. (2004). Birth-death processes with disasters and instantaneous resurrection. *Adv. Appl. Probab.* **36**, 267–292.

[6] Di Crescenzo, A., Giorno, V., Nobile, A.G. and Ricciardi, L.M. (2008). A note on birth-death processes with catastrophes. *Statist. Probab. Lett.* **78**, 2248–2257.

[7] Hadeler, K.P. (1986). Birth and death processes with killing and applications to parasitic infections. In *Stochastic Spatial Processes* (P. Tautu, ed.), Lecture Notes in Mathematics **1212**, 175–186, Springer-Verlag, Heidelberg.

[8] He, Q.-M. and Chavoushi, A.A. (2013). Analysis of queueing systems with customer interjections. *Queueing Systems* **73**, 79–104.

[9] Irle, A. (2003). Stochastic ordering for continuous-time processes. *J. Appl. Probab.* **40**, 361–375.

[10] Irle, A. and Gani, J. (2001). The detection of words and an ordering for Markov chains. *J. Appl. Probab.* **38A**, 66-77.

[11] Karlin, S. and Tavaré, S. (1982). Linear birth and death processes with killing. *J. Appl. Probab.* **19**, 477–487.

[12] Nagylaki, T. (2005). A stochastic model for a progressive chronic disease. *J. Math. Biology* **51**, 268–280.

[13] Puri, P.S. (1972). A method for studying the integral functionals of stochastic processes with applications III. *Proc. Sixth Berkeley Symp. Math. Stat. Prob.* Vol. III, 481–500, UCLA Press.

[14] Resnick, S.I. (1992). *Adventures in Stochastic Processes.* Birkhäuser, Boston, MA.

[15] Ross, S.M. (2007). *Introduction to Probabilty Models* (9th edition). Academic Press, New York.

[16] Stirzaker, D. (2006). Processes with catastrophes. *Math. Scientist* **31**, 107–118.

[17] Van Doorn, E.A. and Zeifman, A.I. (2005a). Extinction probability in a birth-death process with killing. *J. Appl. Probab.* **42**, 185–198.

[18] Van Doorn, E.A. and Zeifman, A.I. (2005b). Birth-death processes with killing. *Statist. Probab. Lett.* **72**, 33–42.