Transient features for Markovian binary trees

Sophie Hautphenne
Université Libre de Bruxelles
Boulevard du Triomphe
1050 Bruxelles, Belgium
shautphe@ulb.ac.be

Guy Latouche
Université Libre de Bruxelles
Boulevard du Triomphe
1050 Bruxelles, Belgium
latouche@ulb.ac.be

Marie-Ange Remiche
Université Libre de Bruxelles
Boulevard du Triomphe
1050 Bruxelles, Belgium
mremiche@ulb.ac.be

ABSTRACT

We analyze a particular class of continuous-time multi-type branching processes, named Markovian binary trees (MBTs), and we focus on transient measures, such as the distribution of the population size at a finite time, the distribution of the time until extinction, and the distribution of the total family size until a given time, as well as the total size until extinction. Our results mainly are formulated as differential equations for probability generating functions and expressions for the factorial moments. They are applied to a demographic comparison of three countries.

Categories and Subject Descriptors

G.3 [**Probability and Statistics**]: Markov Processes; G.1 [**Numerical Analysis**]

General Terms

Algorithms, Theory

Keywords

Branching processes; matrix analytic methods; transient measures

1. INTRODUCTION

Markovian binary trees (MBTs) form a particular class of continuous-time multi-type branching processes [2], in which the life of each individual is controlled by a transient Markovian arrival process (MAP) [11]. An individual may give birth to a child at different epochs, and each child's life is controlled by an independent replica of the same MAP. Eventually, some event in the MAP causes the death of the individual.

These branching processes have been investigated under the name Markovian Binary Trees (MBTs) in Bean *et al.* [3], and in Hautphenne *et al.* [8, 7].

SMCTOOLS 2009, October 19, Pisa, Italy

DOI 10.4108/ICST.VALUETOOLS2009.7428

The most fundamental characteristic of an MBT is its probability of becoming extinct. Bean *et al.* [3], Hautphenne *et al.* [8, 7], and Hautphenne and Van Houdt [9] present algorithms to compute the extinction probability of an MBT, and give their probabilistic interpretations. Here, we focus on some transient, more detailed, aspects of MBTs: the size of the population alive at a given time, the length of time until an MBT becomes extinct, and the total progeny, that is, the total number of descendants of the individual who starts the process.

The paper is organized as follows. In Section 2, we define the Transient Markovian arrival process and the Markovian binary tree, and we recall some matrix derivative properties, which are needed to obtain expressions for various factorial moments. All vectors are supposed to be column vectors, unless otherwise specified.

In Section 3, we analyze the distribution of the population size at some finite time t. Its generating function satisfies a set of Kolmogorov backward and forward differential equations [2]. We use another argument here, referring to the evolution of the MBT over time. We also obtain the factorial moments as solutions of recursive differential equations.

In Section 4, the distribution of the time until extinction of an MBT is obtained as the solution of a differential equation, and we construct an approximation for the tail of the distribution, which we then use to evaluate the mean time until extinction.

We consider in Section 5 the total progeny size, up to some finite time t, and its limit as t goes to infinity. The asymptotic distribution is completely characterized, and in both cases all the factorial moments are recursively determined.

Finally, we present in the last section some numerical illustrations of the different transient measure on an example in the field of human demography, applying the MBT model to the evolution of female families in several countries. This application is presented in more detail in a forthcoming paper [6].

2. BACKGROUND AND DEFINITIONS

2.1 Transient Markovian arrival processes

A transient Markovian arrival process (MAP) is a two-dimensional continuous-time Markovian process

$$\{(M(x), \varphi(x)) : x \in \mathbb{R}^+\}$$

on the state space $\mathbb{N} \times \{0, 1, \ldots, n\}$, where $n \in \mathbb{N}_0$ is supposed to be finite. The process $\varphi(x)$ is a continuous-time Markov chain and is called the process of phase. Some transitions between phases are associated with an *event*. The random variable $M(x)$ counts the number of events which occur up to time x.

The Markovian arrival process is said to be transient (Latouche *et al.* [11]) if the phase 0 is absorbing so that, if $\varphi(x) = 0$ for some x, then there are no more transitions afterward. We assume that the phases 1 to n are all transient, so that, with probability one, the MAP eventually stops after a finite number of transitions. The MAP is characterized by

- two $n \times n$ matrices D_0 and D_1 of phase transition rates respectively without an associated event (these are called hidden transitions) and at an event epoch (these are called observable transitions). They are such that for all $i, j = 1, 2, \ldots, n$, $(D_0)_{ij} \geq 0$ if $i \neq j$ and $(D_0)_{ii} < 0$, $(D_1)_{ij} \geq 0$;
- a non negative n-vector $\boldsymbol{d}$ of transition rates to the absorbing phase 0.

And an initial probability vector which does not play a role here. The two matrices and the vector are such that $D_0 \mathbf{1} + D_1 \mathbf{1} + \boldsymbol{d} = \mathbf{0}$.

2.2 Markovian binary trees

A Markovian binary tree (MBT) is the visualization in time of a branching process which starts with one individual whose lifetime is controlled by an $n-$phases transient MAP $(D_0, D_1, \boldsymbol{d})$. Each observable transition of the MAP corresponds to the birth of one child. Upon transition to the absorbing phase of its MAP, an individual dies. After the birth of a child, the parent continues its life and the child's lifetime is controlled by a new MAP, with the same characteristics, and independent of the others.

An individual which is in phase i may give birth to a child in phase j, and make a transition to phase k; this occurs at the rate $B_{i,jk}$. The matrix B with entries $B_{i,jk}$ is called the birth rate matrix. The matrices B and D_1 are related by $D_1 = B\,(\mathbf{1} \otimes I)$. An individual in phase i makes a transition to the absorbing phase at the rate d_i.

Hidden phase transitions are not associated with an event. An individual which is in phase i may thus make a transition to another phase j; this occurs at the rate $(D_0)_{ij}$.

With this definition, the lifetime of an individual is phase-type distributed PH$(\boldsymbol{\alpha}, D)$, with $D = D_0 + D_1$ [10]. As the set of phase-type distributions is dense in the field of all positive-valued distributions, that is, it can be used to approximate any positive valued distribution, an MBT may thus approximate any Bellman-Harris branching process [5].

There are different ways to interpret an MBT as a multi-type branching process. For instance, we see the phases of the MAP as the types of the branching process. In this way, a type i particle either dies without offspring, which corresponds to the ending of the MAP in the phase i and occurs at the rate d_i, or it dies and gives birth to a type $j \neq i$ particle, which corresponds to a hidden phase change in the MAP and occurs at the rate $(D_0)_{ij}$, or it dies and gives birth to two particles, one of type j and one of type k, which corresponds to an observable event in the MAP and occurs at the rate $B_{i,jk}$.

An MBT eventually becomes extinct if and only if the initial individual dies without any offspring or if it produces a child and both the child and the parent processes eventually become extinct. Let $\boldsymbol{q}$ denote the extinction probability of an MBT, given its initial phase. This vector is the minimal nonnegative solution of the quadratic matrix equation

$$\boldsymbol{x} = (-D_0)^{-1}\,\boldsymbol{d} + (-D_0)^{-1}\,B\,(\boldsymbol{x} \otimes \boldsymbol{x}). \tag{1}$$

The MBT is called *sub-critical*, *supercritical* or *critical* if the eigenvalue of maximal real part of the matrix

$$\Omega = D_0 + B\,(\mathbf{1} \oplus \mathbf{1}) \tag{2}$$

is strictly less than zero, strictly greater than zero, or equal to zero, respectively (Athreya and Ney [2, Chapter 5]). In the sub-critical and critical cases, $\boldsymbol{q} = \mathbf{1}$, while in the supercritical case $\boldsymbol{q} \leq \mathbf{1}$, $\boldsymbol{q} \neq \mathbf{1}$. We assume, without much loss of generality, that all individuals have a finite lifetime with probability 1, this ensures that $q_i > 0$ for all i.

The Newton algorithm [7, 9] is the most efficient algorithm to compute the extinction probability.

2.3 Moments and matrix derivatives

Let $\boldsymbol{s}$ be a vector of size n, and let $d/d\boldsymbol{s}^T$ be a line vector of derivative operators (d/ds_i). The rules below are inspired from the matrix derivatives rules described by MacRae [12]; we particularize them to the derivative with respect to a vector $\boldsymbol{s}^T$.

Definition 1. If Y is a $p \times q$ matrix whose entries are function of the n-vector $\boldsymbol{s}$, then, the derivative of Y with respect to $\boldsymbol{s}^T$ is defined to be a $p \times nq$ matrix of partial derivatives, $dY/d\boldsymbol{s}^T$, which we write as

$$\frac{dY}{d\boldsymbol{s}^T} = Y \otimes d/d\boldsymbol{s}^T$$

with the understanding that $Y_{ij} \cdot d/ds_k = \partial Y_{ij}/\partial s_k$.

Observe that $d\boldsymbol{s}/d\boldsymbol{s}^T = I_n$, where I_n denotes the identity matrix of size n.

The general theorems below will be used in the determination of the factorial moments ([12]).

THEOREM 1 (SUM RULE). *Let Y and Z be matrix functions of $\boldsymbol{s}$, such that their sum is defined. Then,*

$$d(Y + Z)/d\boldsymbol{s}^T = dY/d\boldsymbol{s}^T + dZ/d\boldsymbol{s}^T.$$

THEOREM 2 (PRODUCT RULE). *Let Y and Z be matrix functions of $\boldsymbol{s}$, such that their product is defined. Then,*

$$d(YZ)/d\boldsymbol{s}^T = (dY/d\boldsymbol{s}^T)\,(Z \otimes I_n) + Y\,(dZ/d\boldsymbol{s}^T).$$

To give the rule for derivatives of Kronecker products, we need to introduce the permuted identity matrix $I_{(m,n)}$.

Definition 2. The permuted identity matrix $I_{(m,n)}$ is a square matrix of order mn partitioned into $m \times n$ sub-matrices such that the (i, j)th sub-matrix has a 1 in its (j, i)th position and zeros elsewhere.

One verifies by direct examination that $I_{(m,1)} = I_{(1,m)} = I_m$, $I_{(m,n)} = I^T_{(n,m)}$, and $I_{(m,n)} \cdot I_{(n,m)} = I_{mn}$. Furthermore, the permuted identity matrix may be used to reverse the order of a Kronecker product: if A is $m \times n$, and B is $p \times q$, then $B \otimes A = I_{(m,p)}\,(A \otimes B)\,I_{(q,n)}$.

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7428
http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7428

THEOREM 3 (KRONECKER PRODUCT RULE). *Let Y and Z be matrices of dimensions $s \times t$ and $p \times q$, respectively. We have*

$$\begin{aligned} & d(Y \otimes Z)/d\boldsymbol{s}^T \\ &= (Y \otimes dZ/d\boldsymbol{s}^T) + I_{(p,s)}\,(Z \otimes dY/d\boldsymbol{s}^T)(I_{(t,q)} \otimes I_n) \\ &= (Y \otimes dZ/d\boldsymbol{s}^T) + (dY/d\boldsymbol{s}^T \otimes Z)\,I_{(q,tn)}\,(I_{(t,q)} \otimes I_n). \end{aligned}$$

3. POPULATION SIZE AT TIME T

In this section, we investigate the distribution of the MBT at some finite time t. Let $\boldsymbol{X}(t) = (X_1(t), X_2(t), \ldots, X_n(t))^T$ denote the population size vector at time t, where $X_i(t)$ is the number of individuals in phase i at time t. Let φ_0 denote the initial phase of the MBT. Define the generating function vector $\boldsymbol{F}(\boldsymbol{s}, t)$, conditional on the phase of the first individual, with

$$F_i(\boldsymbol{s}, t) = \sum_{\boldsymbol{k} \geq \boldsymbol{0}} \mathrm{P}[\boldsymbol{X}(t) = \boldsymbol{k}\,|\varphi_0 = i]\,\boldsymbol{s}^{\boldsymbol{k}},$$

$\boldsymbol{s}^{\boldsymbol{k}} = s_1^{k_1}\, s_2^{k_2} \cdots s_n^{k_n}$, $|s_i| < 1$ for all i, and $\boldsymbol{k} \geq \boldsymbol{0}$ must be understood as $k_i \geq 0$ for all i. In the sequel, we shall write $\boldsymbol{F}(\boldsymbol{s}, t) = \sum_{\boldsymbol{k}\geq\boldsymbol{0}} \mathrm{P}_{\varphi_0}[\boldsymbol{X}(t) = \boldsymbol{k}]\,\boldsymbol{s}^{\boldsymbol{k}}$.

THEOREM 4. *The forward and backward Kolmogorov equations for the generating function $\boldsymbol{F}(\boldsymbol{s}, t)$ of an MBT are*

$$\frac{\partial}{\partial t}\boldsymbol{F}(\boldsymbol{s}, t) - \frac{\partial}{\partial \boldsymbol{s}^T}\boldsymbol{F}(\boldsymbol{s}, t) \cdot \boldsymbol{a}(\boldsymbol{s}) = \boldsymbol{0}, \quad (3)$$

where $\partial/\partial\boldsymbol{s}^T \boldsymbol{F}(\boldsymbol{s}, t) = \boldsymbol{F}(\boldsymbol{s}, t) \otimes \partial/\partial\boldsymbol{s}^T$, and

$$\frac{\partial}{\partial t}\boldsymbol{F}(\boldsymbol{s}, t) = \boldsymbol{a}(\boldsymbol{F}(\boldsymbol{s}, t)), \quad (4)$$

with $\boldsymbol{F}(\boldsymbol{s}, 0) = \boldsymbol{s}$ and $\boldsymbol{a}(\boldsymbol{x}) = \boldsymbol{d} + D_0\,\boldsymbol{x} + B\,(\boldsymbol{x} \otimes \boldsymbol{x})$.

As mentioned in [2, V.7], the equations follow from the Kolmogorov equations for the Markov process $\boldsymbol{X}(t)$, but (4) may also be justified by an argument based on the dynamics of the MBT. In order to illustrate the difference in approaches, we give the justification for (3) on the basis of the Kolmogorov equations, assuming that $n = 2$ in order to simplify the presentation.

The forward Kolmogorov equations for $\boldsymbol{X}(t)$ are

$$\begin{aligned} & \frac{d}{dt}\mathrm{P}_{\varphi_0}[\boldsymbol{X}(t) = (0,0)] \quad (5) \\ &= d_1\,\mathrm{P}_{\varphi_0}[\boldsymbol{X}(t) = (1,0)] + d_2\,\mathrm{P}_{\varphi_0}[\boldsymbol{X}(t) = (0,1)], \end{aligned}$$

and

$$\begin{aligned} & \frac{d}{dt}\mathrm{P}_{\varphi_0}[\boldsymbol{X}(t) = (k_1,k_2)] = \quad (6) \\ & \quad [k_1\,(B_{1,12} + B_{1,21}) + (k_2 - 1)\,B_{2,22}] \\ & \quad \times\mathrm{P}_{\varphi_0}[\boldsymbol{X}(t) = (k_1, k_2 - 1)] \\ &+ [k_2\,(B_{2,12} + B_{2,21}) + (k_1 - 1)\,B_{1,11}] \\ & \quad \times\mathrm{P}_{\varphi_0}[\boldsymbol{X}(t) = (k_1 - 1, k_2)] \\ &+ (k_1 + 1)\,B_{1,22}\,\mathrm{P}_{\varphi_0}[\boldsymbol{X}(t) = (k_1 + 1, k_2 - 2)] \\ &+ (k_2 + 1)\,B_{2,11}\,\mathrm{P}_{\varphi_0}[\boldsymbol{X}(t) = (k_1 - 2, k_2 + 1)] \\ &+ (k_1 + 1)\,d_1\,\mathrm{P}_{\varphi_0}[\boldsymbol{X}(t) = (k_1 + 1, k_2)] \\ &+ (k_2 + 1)\,d_2\,\mathrm{P}_{\varphi_0}[\boldsymbol{X}(t) = (k_1, k_2 + 1)] \\ &+ (k_1 + 1)\,(D_0)_{12}\,\mathrm{P}_{\varphi_0}[\boldsymbol{X}(t) = (k_1 + 1, k_2 - 1)] \\ &+ (k_2 + 1)\,(D_0)_{21}\,\mathrm{P}_{\varphi_0}[\boldsymbol{X}(t) = (k_1 - 1, k_2 + 1)] \\ &+ [k_1\,(D_0)_{11} + k_2\,(D_0)_{22}]\,\mathrm{P}_{\varphi_0}[\boldsymbol{X}(t) = (k_1, k_2)] \end{aligned}$$

for $k_1, k_2 \geq 1$, where the probability that $X_1(t)$ or $X_2(t)$ takes a strictly negative value is equal to zero, by convention. Finally, $\mathrm{P}[\boldsymbol{X}(0) = (1,0)|\varphi_0 = 1] = 1$ and $\mathrm{P}[\boldsymbol{X}(0) = (0,1)|\varphi_0 = 2] = 1$.

We multiply (6) by $s_1^{k_1}\, s_2^{k_2}$, and sum over all values of k_1, k_2, and obtain (3) after some algebraic manipulations.

The equation (4) may also be obtained from the Kolmogorov equations but a less cumbersome argument follows from conditioning on the time of the first event: either the initial individual has not undergone any observable event yet at time t, which occurs with probability given by $e^{D_0\,t}$, or it dies at some time $u \leq t$, which occurs with probability given by $e^{D_0\,u}\,\boldsymbol{d}$, or it has a child at some time $u \leq t$, which occurs with probability given by $e^{D_0\,u}\,B$, and the two subprocesses evolve independently of each others afterwards. In matrix notation, this gives

$$\begin{aligned} \boldsymbol{F}(\boldsymbol{s}, t) &= e^{D_0\,t}\,\boldsymbol{s} + \int_0^t e^{D_0\,u}\,\boldsymbol{d}\,du \\ &\quad + \int_0^t e^{D_0\,u}\,B\,(\boldsymbol{F}(\boldsymbol{s}, t - u) \otimes \boldsymbol{F}(\boldsymbol{s}, t - u))\,du. \end{aligned}$$

Now, taking the derivative with respect to t on both sides, we have

$$\frac{\partial}{\partial t}\boldsymbol{F}(\boldsymbol{s}, t) = \boldsymbol{d} + D_0\,\boldsymbol{F}(\boldsymbol{s}, t) + B\,(\boldsymbol{F}(\boldsymbol{s}, t) \otimes \boldsymbol{F}(\boldsymbol{s}, t)),$$

which is (4).

We do not have analytical solutions for the (partial) differential equations but there exist powerful numerical tools such as the solvers in MATLAB. This, combined with numerical techniques for inverting probability generating functions [1], allows in principle to obtain approximations for the distribution of the population size at some given time. Further study is necessary, however, since the whole procedure has shown signs of numerical instability.

Nevertheless, as we show in the next section, we may obtain information about the time to extinction, and we may extract factorial moments of $\boldsymbol{X}(t)$: the kth factorial moment $M^{(k)}(t)$ is an $n \times n^k$ matrix given by

$$\begin{aligned} M^{(k)}(t) &= \frac{\partial^k}{(\partial\boldsymbol{s}^T)^k}\,\boldsymbol{F}(\boldsymbol{s}, t)\,|_{\boldsymbol{s}=\boldsymbol{1}} \\ &= \boldsymbol{F}(\boldsymbol{s}, t) \otimes \underbrace{\frac{\partial}{\partial\boldsymbol{s}^T} \otimes \cdots \frac{\partial}{\partial\boldsymbol{s}^T}}_{k \text{ times}}\,|_{\boldsymbol{s}=\boldsymbol{1}}\;. \end{aligned}$$

THEOREM 5. *The matrices $M^{(k)}(t)$, $k \geq 1$, are solutions of the two following recursive differential equations, with initial conditions $M^{(1)}(0) = I_n$ and $M^{(k)}(0) = 0$ for $k \geq 2$. The first system is*

$$\begin{aligned} \frac{\partial}{\partial t}M^{(k)}(t) &= M^{(k)}(t)\,(\Omega \otimes I_{n^{k-1}})\,A(k) \quad (7) \\ &\quad + M^{(k-1)}(t)\,[B\,(I_{n^2} + I_{(n,n)}) \otimes I_{n^{k-2}}]\,C(k) \end{aligned}$$

where Ω is defined in (2) and the $n^k \times n^k$ coefficients matrices $A(k)$ and $C(k)$ are recursively defined as

$$\begin{aligned} A(1) &= I_n \\ C(1) &= 0_n \\ A(k) &= I_{(n^{k-1},n)} + (A(k-1) \otimes I_n), \quad (8) \\ C(k) &= I_{(n^{k-2},n^2)}\,(I_{(n,n^{k-2})} \otimes I_n)\,(A(k-1) \otimes I_n) \\ &\quad + (C(k-1) \otimes I_n), \quad (9) \end{aligned}$$

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7428
http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7428

for $k \geq 2$.

One also has

$$\frac{\partial}{\partial t} M^{(k)}(t) = \Omega\, M^{(k)}(t) + \sum_{i=1}^{k-1} B\,(M^{(i)}(t) \otimes M^{(k-i)}(t))\, C(i, k-i), \qquad (10)$$

where the matrix coefficients $C(i, k-i)$ are recursively defined as

$$\begin{aligned} C(0, k-i) &= I_{n^{k-i}} \\ C(i, 0) &= I_{n^i} \\ C(i, k-i) &= I_{(n^{k-i}, n^i)}\, [I_{(n^{i-1}, n^{k-i})}\, C(i-1, k-i) \otimes I_n] \\ &\quad +[C(i, k-i-1) \otimes I_n]. \end{aligned}$$

Proof. The first recurrence is proved by taking successive derivatives of (3) and the second recurrence is based on (4). We only give details for (3).

First, note that

$$\begin{aligned} \boldsymbol{a}(\boldsymbol{1}) &= \boldsymbol{d} + D_0 \boldsymbol{s} + B(\boldsymbol{s} \otimes \boldsymbol{s})|_{\boldsymbol{s}=\boldsymbol{1}} = \boldsymbol{0} \\ \frac{d}{d\boldsymbol{s}^T} \boldsymbol{a}(\boldsymbol{s})|_{\boldsymbol{s}=\boldsymbol{1}} &= D_0 + B\,(I_n \otimes \boldsymbol{s} + \boldsymbol{s} \otimes I_n)|_{\boldsymbol{s}=\boldsymbol{1}} = \Omega \\ \frac{d^2}{(d\boldsymbol{s}^T)^2} \boldsymbol{a}(\boldsymbol{s})|_{\boldsymbol{s}=\boldsymbol{1}} &= B(I_n \otimes I_n) + B(I_n \otimes I_n) I_{(n,n)} \\ &= B(I_{n^2} + I_{(n,n)}), \end{aligned}$$

by the derivative rule of Kronecker products.

Taking derivatives of both sides of (3), we obtain

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{s}^T} \frac{\partial}{\partial t} \boldsymbol{F}(\boldsymbol{s}, t) &= \frac{\partial^2}{(\partial \boldsymbol{s}^T)^2} \boldsymbol{F}(\boldsymbol{s}, t) \cdot (\boldsymbol{a}(\boldsymbol{s}) \otimes I_n) \\ &\quad + \frac{\partial}{\partial \boldsymbol{s}^T} \boldsymbol{F}(\boldsymbol{s}, t) \cdot \frac{d}{d\boldsymbol{s}^T}\, \boldsymbol{a}(\boldsymbol{s}) \end{aligned}$$

and by setting $\boldsymbol{s} = \boldsymbol{1}$ we obtain (7) for $k = 1$.

Now, we make the induction assumption that

$$\begin{aligned} &\frac{\partial^k}{(\partial \boldsymbol{s}^T)^k} \frac{\partial}{\partial t} \boldsymbol{F}(\boldsymbol{s}, t) = \qquad (11) \\ &\quad \frac{\partial^{k+1}}{(\partial \boldsymbol{s}^T)^{k+1}} \boldsymbol{F}(\boldsymbol{s}, t) \cdot (\boldsymbol{a}(\boldsymbol{s}) \otimes I_{n^k}) \\ &\quad + \frac{\partial^k}{(\partial \boldsymbol{s}^T)^k} \boldsymbol{F}(\boldsymbol{s}, t) \cdot (\frac{d}{d\boldsymbol{s}^T} \boldsymbol{a}(\boldsymbol{s}) \otimes I_{n^{k-1}})\, A(k) \\ &\quad + \frac{\partial^{k-1}}{(\partial \boldsymbol{s}^T)^{k-1}} \boldsymbol{F}(\boldsymbol{s}, t) \cdot (\frac{d^2}{(d\boldsymbol{s}^T)^2}\, \boldsymbol{a}(\boldsymbol{s}) \otimes I_{n^{k-2}})\, C(k) \end{aligned}$$

for some $k \geq 1$ and, differentiating again with respect to $\boldsymbol{s}^T$, we obtain

$$\begin{aligned} &\frac{\partial^{k+1}}{(\partial \boldsymbol{s}^T)^{k+1}} \frac{\partial}{\partial t} \boldsymbol{F}(\boldsymbol{s}, t) = \\ &\quad \frac{\partial^{k+2}}{(\partial \boldsymbol{s}^T)^{k+2}} \boldsymbol{F}(\boldsymbol{s}, t) \cdot (\boldsymbol{a}(\boldsymbol{s}) \otimes I_{n^k} \otimes I_n) \\ &\quad + \frac{\partial^{k+1}}{(\partial \boldsymbol{s}^T)^{k+1}} \boldsymbol{F}(\boldsymbol{s}, t) \cdot (\frac{d}{d\boldsymbol{s}^T} \boldsymbol{a}(\boldsymbol{s}) \otimes I_{n^k})\, I_{(n^k, n)} \\ &\quad + \frac{\partial^{k+1}}{(\partial \boldsymbol{s}^T)^{k+1}} \boldsymbol{F}(\boldsymbol{s}, t) \cdot (\frac{d}{d\boldsymbol{s}^T} \boldsymbol{a}(\boldsymbol{s}) \otimes I_{n^{k-1}} \otimes I_n)\, (A(k) \otimes I_n) \\ &\quad + \frac{\partial^k}{(\partial \boldsymbol{s}^T)^k} \boldsymbol{F}(\boldsymbol{s}, t) \cdot (\frac{d^2}{(d\boldsymbol{s}^T)^2}\, \boldsymbol{a}(\boldsymbol{s}) \otimes I_{n^{k-1}})\, I_{(n^{k-1}, n^2)} \\ &\quad \cdot (I_{(n, n^{k-1})} \otimes I_n)\, (A(k) \otimes I_n) \\ &\quad + \frac{\partial^k}{(\partial \boldsymbol{s}^T)^k} \boldsymbol{F}(\boldsymbol{s}, t) \cdot (\frac{d^2}{(d\boldsymbol{s}^T)^2}\, \boldsymbol{a}(\boldsymbol{s}) \otimes I_{n^{k-1}})\, (C(k) \otimes I_n) \end{aligned}$$

Using (8, 9), we obtain

$$\begin{aligned} &\frac{\partial^{k+1}}{(\partial \boldsymbol{s}^T)^{k+1}} \frac{\partial}{\partial t} \boldsymbol{F}(\boldsymbol{s}, t) = \\ &\quad \frac{\partial^{k+2}}{(\partial \boldsymbol{s}^T)^{k+2}} \boldsymbol{F}(\boldsymbol{s}, t) \cdot (\boldsymbol{a}(\boldsymbol{s}) \otimes I_{n^{k+1}}) \\ &\quad + \frac{\partial^{k+1}}{(\partial \boldsymbol{s}^T)^{k+1}} \boldsymbol{F}(\boldsymbol{s}, t) \cdot (\frac{d}{d\boldsymbol{s}^T} \boldsymbol{a}(\boldsymbol{s}) \otimes I_{n^k})\, A(k+1) \\ &\quad + \frac{\partial^k}{(\partial \boldsymbol{s}^T)^k} \boldsymbol{F}(\boldsymbol{s}, t) \cdot (\frac{d^2}{(d\boldsymbol{s}^T)^2}\, \boldsymbol{a}(\boldsymbol{s}) \otimes I_{n^{k-1}})\, C(k+1) \end{aligned}$$

which shows that (11) holds for all k. The recursion formula for the factorial moments follows when we set $\boldsymbol{s} = \boldsymbol{1}$ in (11). □

The size of the matrices $M^{(k)}(t)$ very rapidly increases, which somewhat limits the usefulness of Theorem 5 beyond the first few moments, but the mean and variance are easily obtained, as we show now.

Corollary 1. *The first two moments are given by*

$$\begin{aligned} M^{(1)}(t) &= e^{\Omega\, t} \qquad (12) \\ M^{(2)}(t) &= X(t)\, (I_{n^2} + I_{(n,n)}) \qquad (13) \end{aligned}$$

where $X(t)$ is the solution of the Lyapunov equation

$$X(t)\, (\Omega \oplus \Omega) - \Omega\, X(t) + e^{\Omega\, t}\, B - B\, (e^{\Omega\, t} \otimes e^{\Omega\, t}) = 0. \qquad (14)$$

Proof. Taking $k = 1$ in (10), we see that $M^{(1)}(t)$ is the solution of $\partial/\partial t M^{(1)}(t) = \Omega\, M^{(1)}(t)$, with $M^{(1)}(0) = I$. This proves (12).

For the second moment, (10, 12) yield

$$M^{(2)}(t) = e^{\Omega\, t}\, Y(t)\ (I_{n^2} + I_{(n,n)}),$$

where

$$Y(t) = \int_0^t e^{-\Omega\, u}\, B\, (e^{\Omega\, u} \otimes e^{\Omega\, u})\, du.$$

We premultiply by Ω and integrate by parts, to find that $Y(t)$ satisfies the Lyapunov equation

$$Y(t)\, (\Omega \oplus \Omega) - \Omega\, Y(t) + B - e^{-\Omega\, t}\, B\, (e^{\Omega\, t} \otimes e^{\Omega\, t}) = 0.$$

Writing $X(t) = e^{\Omega\, t}\, Y(t)$, we obtain (13). □

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7428
http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7428

The Lyapunov equation is solved using the `lyap` solver in MATLAB. This uses a triangular decomposition approach and the complexity is $\mathcal{O}(n^3)$.

The standard deviation of the total population size at time t, given the initial phase, is thus given as

$$\boldsymbol{\sigma}_M(t) = \left[M^{(2)}(t)\,\mathbf{1} + \mathrm{diag}(M^{(1)}(t)\,\mathbf{1})\,(\mathbf{1} - M^{(1)}(t)\,\mathbf{1})\right]^{1/2}.$$

Remark 1. We can give a physical interpretation to the inverse $(-\Omega)^{-1}$ when it is well defined, that is in the sub-critical case. Indeed,

$$\begin{aligned}
[(-\Omega)^{-1}]_{ij} &= \int_0^\infty \left[e^{\Omega\, t}\right]_{ij} dt \\
&= \int_0^\infty \mathrm{E}[X_j(t)\,|\,\varphi_0 = i]\,dt \\
&= \int_0^\infty \mathrm{E}\left[\sum_{n\geq 0} n\,\mathbb{1}_{\{X_j(t)=n\}}\,|\,\varphi_0 = i\right] dt \\
&= \mathrm{E}\left[\sum_{n\geq 0} \int_0^\infty n\,\mathbb{1}_{\{X_j(t)=n\}}\,dt\,|\,\varphi_0 = i\right]
\end{aligned}$$

If we define the total *cumulated* amount of time the process is in phase j as the sum of the lengths of intervals where $X_j(\cdot) > 0$, weighted by X_j, then the entry (i,j) of $(-\Omega)^{-1}$ may be interpreted as the expectation of that sum, given that the process starts with a first individual in phase i.

4. TIME UNTIL EXTINCTION

Our next measure of interest is the distribution of the time T_e until the MBT becomes extinct. For reasons that will appear below, we assume that the MBT is either super-critical or sub-critical.

We denote by $\boldsymbol{F}(t) = \mathrm{P}[T_e \leq t\,|\varphi_0]$ the distribution function of T_e and we observe that the extinction probability $\boldsymbol{q}$ defined in Section 2.2 actually is the limit of $\boldsymbol{F}(t)$ as $t\to\infty$. It is clear that $T_e \leq t$ if and only if $\boldsymbol{X}(t) = 0$, so that $\boldsymbol{F}(t) = \boldsymbol{F}(\boldsymbol{s},t)|_{\boldsymbol{s}=\mathbf{0}}$.

Thus, taking $\boldsymbol{s} = \mathbf{0}$ in (4), we find that the distribution function $\boldsymbol{F}(t)$ satisfies the differential equation

$$\frac{d}{dt}\boldsymbol{F}(t) = \boldsymbol{d} + D_0\,\boldsymbol{F}(t) + B\,(\boldsymbol{F}(t)\otimes\boldsymbol{F}(t)), \qquad (15)$$

with $\boldsymbol{F}(0) = \mathbf{0}$.

This is a quadratic matrix differential equation, and numerical tools allow us to approximate the distribution, as we show in Section 6. We have used the solver `ode45` in MATLAB, which is based on an explicit Runge-Kutta (4,5)-formula. This is a one-step solver, and its complexity is linear in the number n of phases. Further details are to be found in Dormand and Prince [4].

If $\boldsymbol{q} < 1$, then T_e is infinite with a positive probability and its expectation is infinite. In order to investigate a meaningful quantity, we define $\boldsymbol{M}_e$ as the conditional expectation of T_e, given the initial phase and given that extinction occurs:

$$\boldsymbol{M}_e = \Phi^{-1}\int_0^\infty [\boldsymbol{q} - \boldsymbol{F}(u)]\,du = \int_0^\infty [\mathbf{1} - \Phi^{-1}\,\boldsymbol{F}(u)]\,du$$

where

$$\Phi = \mathrm{diag}(\boldsymbol{q}) \qquad (16)$$

is nonsingular since we assumed at the end of Section 2.2 that $q_i > 0$ for all i.

To compute $\boldsymbol{M}_e$, we proceed as follows: we fix some time T^* such that

$$||\mathbf{1} - \Phi^{-1}\,\boldsymbol{F}(T^*)|| < \varepsilon \qquad (17)$$

where ϵ is arbitrarily small, and we write that

$$\boldsymbol{M}_e = \int_0^{T^*}[\mathbf{1} - \Phi^{-1}\boldsymbol{F}(u)]\,du + \int_{T^*}^\infty [\mathbf{1} - \Phi^{-1}\boldsymbol{F}(u)]\,du. \qquad (18)$$

The first integral may be approximated using the trapezoid rule: we choose a number k^* of intervals and write

$$\begin{aligned}
\int_0^{T^*}[\mathbf{1} - \Phi^{-1}\boldsymbol{F}(u)]\,du &= h\Big\{\mathbf{1} - \Phi^{-1}[\boldsymbol{F}(0) + \boldsymbol{F}(T^*)]/2 \\
&\quad + \sum_{i=1}^{k^*-1}[\mathbf{1} - \Phi^{-1}\boldsymbol{F}(ih)]\Big\} + E
\end{aligned}$$

where $h = T^*/k^*$ and $E = T^* h^2 \Phi^{-1}\boldsymbol{F}''(c)/12$ for some c in $0 < c < T^*$ is the approximation error. In practice, we choose ϵ and h and compute $F(ih)$ for successive values of i until $F(k^*h)$ satisfies the inequality (17).

The second integral may be computed using a function $\tilde{\boldsymbol{F}}(t)$ which approximates $\boldsymbol{F}(t)$ for large values of t, that is, for values of t such that $\boldsymbol{F}(t) \approx \boldsymbol{q}$. This approximation is given by

$$\tilde{\boldsymbol{F}}(t) = \left[I - e^{\Theta\, t}\right]\boldsymbol{q}, \qquad (19)$$

where

$$\Theta = D_0 + B(\boldsymbol{q}\oplus\boldsymbol{q}). \qquad (20)$$

It is obtained as follows. Define $\boldsymbol{\epsilon}(t) = \boldsymbol{q} - \boldsymbol{F}(t)$. The vector $\boldsymbol{q}$ is a solution of

$$0 = \boldsymbol{d} + D_0\,\boldsymbol{q} + B\,(\boldsymbol{q}\otimes\boldsymbol{q}),$$

which is obtained by pre-multiplying both sides of (1) by D_0. Subtracting this from (15), we obtain

$$\begin{aligned}
\boldsymbol{\epsilon}'(t) &= D_0\boldsymbol{\epsilon}(t) + B[(\boldsymbol{q}\otimes\boldsymbol{q}) - (\boldsymbol{F}(t)\otimes\boldsymbol{F}(t))] \\
&= D_0\boldsymbol{\epsilon}(t) + B[(\boldsymbol{q} - \boldsymbol{F}(t)\otimes\boldsymbol{q}) + (\boldsymbol{F}(t)\otimes\boldsymbol{q} - \boldsymbol{F}(t))] \\
&= D_0\boldsymbol{\epsilon}(t) + B[(I\otimes\boldsymbol{q}) + (\boldsymbol{F}(t)\otimes I)]\boldsymbol{\epsilon}(t). \qquad (21)
\end{aligned}$$

Now, if t is large enough, then $\boldsymbol{F}(t)$ may be replaced by $\boldsymbol{q}$ in the equation above and we obtain the approximate system $\tilde{\boldsymbol{\epsilon}}'(t) = \Theta\,\tilde{\boldsymbol{\epsilon}}(t)$, which leads to

$$\tilde{\boldsymbol{\epsilon}}(t) = e^{\Theta\, t}\,\boldsymbol{q}.$$

Since $\boldsymbol{\epsilon}(0) = \boldsymbol{q}$ by definition, this yields (19).

Thus, the second integral in (18) may be approximated by

$$\begin{aligned}
\int_{T^*}^\infty [\mathbf{1} - \Phi^{-1}\tilde{\boldsymbol{F}}(u)]\,du &= \int_{k^*h}^\infty \Phi^{-1}e^{\Theta u}\boldsymbol{q}\,du \\
&= \Phi^{-1}[-\Theta]^{-1}e^{\Theta k^* h}\boldsymbol{q} \qquad (22)
\end{aligned}$$

provided that the maximal real part of the eigenvalue of Θ is strictly negative; we show below that this is the case when, as we assumed at the beginning of this section, the MBT is not critical.

Theorem 6. *If the MBT is super- or sub-critical, then the eigenvalues of $\Theta = D_0 + B\,(\boldsymbol{q}\oplus\boldsymbol{q})$ all have a strictly negative real part.*

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7428
http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7428

PROOF. If the MBT is sub-critical, then $\boldsymbol{q} = \mathbf{1}$, $\Theta = \Omega$ and the real part of all eigenvalue is strictly negative, as we recalled in Section 2.2, so that we only need to consider the super-critical case.

There are several ways to interpret the MBT as a multi-type branching process. We apply the standard uniformization procedure to each controlling MAP: choose $c \geq (-D_0)_{ii}$ for all i and assume that after exponential intervals of time with parameter c, individuals die, or are replaced by one new individual, or are replaced by two new individuals, with probabilities given by the following vector and matrices:

$$\boldsymbol{d^*} = 1/c\, \boldsymbol{d} \qquad D_0^* = 1/c\, D_0 + I \qquad B^* = 1/c\, B.$$

The generating function of the progeny of a particle in phase i is thus given by the ith entry of

$$\boldsymbol{G}(\boldsymbol{s}) = \boldsymbol{d^*} + D_0^*\, \boldsymbol{s} + B^*(\boldsymbol{s} \otimes \boldsymbol{s}),$$

and the matrix of first derivatives $M(\boldsymbol{s}) = \partial/\partial \boldsymbol{s}^T \boldsymbol{G}(\boldsymbol{s})$ is given by $M(\boldsymbol{s}) = D_0^* + B^*(\boldsymbol{s} \oplus \boldsymbol{s})$. Observe that $M(\boldsymbol{q}) = I + 1/c\,\Theta$, so that the theorem is proved once we show that the spectral radius of $M(\boldsymbol{q})$ is strictly less than one.

The generating function at the kth generation is $\boldsymbol{G}_k(\boldsymbol{s}) = \boldsymbol{G}(\boldsymbol{G}_{k-1}(\boldsymbol{s}))$, with $\boldsymbol{G}_1(\boldsymbol{s}) = \boldsymbol{G}(\boldsymbol{s})$, and $\lim_{k\to\infty} \boldsymbol{G}_k(\boldsymbol{s}) = \boldsymbol{q}$ for all $\boldsymbol{s} \leq \boldsymbol{q}$ (see Athreya and Ney [2, Section 5.3]), so that $\lim_{k\to\infty} \partial/\partial \boldsymbol{s}^T \boldsymbol{G}_k(\boldsymbol{s}) = 0$, in particular for $\boldsymbol{s} = \boldsymbol{q}$.

One shows by induction that

$$\partial/\partial \boldsymbol{s}^T \boldsymbol{G}_k(\boldsymbol{s}) = M(\boldsymbol{G}_{k-1}(\boldsymbol{s}))\partial/\partial \boldsymbol{s}^T \boldsymbol{G}_{k-1}(\boldsymbol{s}).$$

The vector $\boldsymbol{q}$ being the minimal solution of $\boldsymbol{q} = \boldsymbol{G}(\boldsymbol{q})$, we have $\boldsymbol{q} = \boldsymbol{G}_k(\boldsymbol{q})$ for all k, so that

$$\partial/\partial \boldsymbol{s}^T \boldsymbol{G}_k(\boldsymbol{s})|_{\boldsymbol{s}=\boldsymbol{q}} = (M(\boldsymbol{q}))^k$$

and therefore $\lim_{k\to\infty} (M(\boldsymbol{q}))^k = 0$, and so the spectral radius of $M(\boldsymbol{q})$ is strictly less than one, which concludes the proof. □

In summary, the mean time until extinction, given that extinction occurs, may be approximated by

$$\begin{aligned} \tilde{\boldsymbol{M}}_e \;&=\; h\Big\{1 - \Phi^{-1}[\boldsymbol{F}(0) + \boldsymbol{F}(T^*)]/2 \qquad (23) \\ &\quad + \sum_{i=1}^{k^*-1}[\mathbf{1} - \Phi^{-1}\boldsymbol{F}(ih)]\Big\} + \Phi^{-1}[-\Theta]^{-1}\, e^{\Theta T^*}\boldsymbol{q}. \end{aligned}$$

5. TOTAL PROGENY SIZE

Recall that the lifetime of an individual is controlled by a transient MAP which controls when children are born and when death occurs. We analyze here the distribution of $N(t)$, the total number of individuals born until time t, irrespective of their status, and its limit as t goes to infinity.

THEOREM 7. *Let $\boldsymbol{g}_i(k,t) = P[N(t) = k|\varphi_0 = i]$ denote the probability that a total of k individuals are born before time t, given the initial phase of the first individual is i. These probabilities are given by*

$$\boldsymbol{g}(0,t) = \mathbf{0} \qquad (24)$$

$$\boldsymbol{g}(1,t) = e^{D_0 t}\mathbf{1} + \int_0^t e^{D_0 u}\boldsymbol{d}\, du, \qquad (25)$$

$$\boldsymbol{g}(k,t) = \int_0^t e^{D_0 u} B \cdot \sum_{i=1}^{k-1}[\boldsymbol{g}(i,t-u) \otimes \boldsymbol{g}(k-i,t-u)]\, du. \qquad (26)$$

for $k \geq 2$, $t \geq 0$.

PROOF. The first equation is justified by the fact that there is always at least one individual, the one at the origin of the process, in the total progeny. In order to justify (25), we note that at time t there is only one individual in total if either the one at the origin has not yet given birth or has died at some time $u \leq t$. Finally, if at time t the total progeny is $k \geq 2$, it means that the first individual gives birth to a child at some time $u \leq t$ and that the sum of the total progenies in the families generated by the two individuals after the birth event equals k, this proves (26). □

We may express the probability generating function

$$\boldsymbol{G}(z,t) = \sum_{k\geq 0} \boldsymbol{g}(k,t) z^k$$

of $N(t)$ as the solution of a differential equation, by using the recursive expressions in Theorem 7 and taking derivatives with respect to t:

$$\frac{\partial}{\partial t}\boldsymbol{G}(z,t) = \boldsymbol{d}z + D_0\boldsymbol{G}(z,t) + B(\boldsymbol{G}(z,t) \otimes \boldsymbol{G}(z,t)), \quad (27)$$

with $\boldsymbol{G}(z,0) = z$. Note that $\boldsymbol{G}(1,t) = \mathbf{1}$ for all finite t.

We might numerically solve this differential equation and take the inverse transform, as suggested for the generating function $\boldsymbol{F}(\boldsymbol{s},t)$ of the population size at time t; here, however, the distribution of the total progeny size is directly computable by (24)–(26).

Let $\boldsymbol{D}^{(k)}(t) = \partial^k \boldsymbol{G}(z,t)/(\partial z)^k|_{z=1}$ denote the $n \times 1$ kth factorial moment vector of $N(t)$. These moments are recursively characterized as follows, the proof is by induction using standard scalar differentiation rules.

THEOREM 8. *The vectors $\boldsymbol{D}^{(k)}(t)$ satisfy the following recurrence*

$$\begin{aligned} \frac{\partial}{\partial t}\boldsymbol{D}^{(1)}(t) \;&=\; \Omega\, \boldsymbol{D}^{(1)}(t) + \boldsymbol{d}, \quad k \geq 2: \qquad (28) \\ \frac{\partial}{\partial t}\boldsymbol{D}^{(k)}(t) \;&=\; \Omega\, \boldsymbol{D}^{(k)}(t) + \sum_{i=1}^{k-1}\binom{k}{i} B\, [\boldsymbol{D}^{(i)}(t) \otimes \boldsymbol{D}^{(k-i)}(t)], \end{aligned}$$

with $\boldsymbol{D}^{(1)}(0) = \mathbf{1}$ and $\boldsymbol{D}^{(k)}(0) = \mathbf{0}$ for $k \geq 2$.

One easily obtains an explicit expression for the first moment: the solution of (28), is given by

$$\boldsymbol{D}^{(1)}(t) = e^{\Omega\, t}\left(\int_0^t e^{-\Omega\, u}\, \boldsymbol{d}\, du + \mathbf{1}\right).$$

If the inverse of Ω exists, then this is equivalent to

$$\boldsymbol{D}^{(1)}(t) = [I - e^{\Omega t}]\,(-\Omega)^{-1}\boldsymbol{d} + e^{\Omega t}\mathbf{1}.$$

Now let t tend to infinity, and define $\boldsymbol{g}(k) = \lim_{t\to\infty} \boldsymbol{g}(k,t)$. We easily obtain from (24–26) that

$$\begin{aligned} \boldsymbol{g}(0) &= \mathbf{0} \\ \boldsymbol{g}(1) &= (-D_0)^{-1}\boldsymbol{d} \\ \boldsymbol{g}(k) &= (-D_0)^{-1}B\sum_{i=1}^{k-1}[\boldsymbol{g}(i) \otimes \boldsymbol{g}(k-i)], \end{aligned}$$

for $k \geq 2$. The generating function $\boldsymbol{G}(z) = \sum_{k\geq 0} \boldsymbol{g}(k)\, z^k$ is such that $\boldsymbol{G}(1) = \boldsymbol{q}$ and it is, for every z, the minimal non negative solution of the fixed point equation

$$\boldsymbol{G}(z) = (-D_0)^{-1}\boldsymbol{d}z + (-D_0)^{-1}B(\boldsymbol{G}(z) \otimes \boldsymbol{G}(z)) \qquad (29)$$

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7428
http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7428

with the same structure as the extinction equation (1) and the same numerical techniques may be used to evaluate $\boldsymbol{G}(z)$ for any given z.

The partial derivatives $\boldsymbol{D}^{(k)} = \partial^k \boldsymbol{G}(z)/(\partial z)^k|_{z=1}$ are here equal to the factorial moment vector of the total progeny $N(\infty)$ *on the paths where extinction occurs.* It is easy to show by induction that

$$\begin{aligned} \boldsymbol{D}^{(1)} &= (-\Theta)^{-1}\boldsymbol{d}, \qquad (30) \\ \boldsymbol{D}^{(k)} &= (-\Theta)^{-1}\sum_{i=1}^{k-1}\binom{k}{i}B(\boldsymbol{D}^{(i)}\otimes\boldsymbol{D}^{(k-i)}) \end{aligned}$$

for $k \geq 2$, where Θ is defined in (20), provided that the MBT is not critical.

Here, the moments may be expressed explicitly. For instance, the second factorial moment is given by

$$\boldsymbol{D}^{(2)} = 2(-\Theta)^{-1}B[(-\Theta)^{-1}\boldsymbol{d}\otimes(-\Theta)^{-1}\boldsymbol{d}]$$

but such expressions rapidly become very cumbersome.

It is more natural to consider the conditional distribution of $N(\infty)$ *given that* extinction occurs. It suffices to multiply by Φ^{-1} the expressions obtained so far, where Φ is defined in (16). We obtain, for instance, that the conditional mean total progeny size, given extinction occurs, is

$$\bar{\boldsymbol{D}}^{(1)} = \Phi^{-1}(-\Theta)^{-1}\boldsymbol{d} = -(\Theta\Phi)^{-1}\boldsymbol{d}$$

and the conditional standard deviation as

$$\bar{\sigma}_D = \left[\Phi^{-1}\boldsymbol{D}^{(2)} + \mathrm{diag}(\bar{\boldsymbol{D}}^{(1)})(\boldsymbol{1}-\bar{\boldsymbol{D}}^{(1)})\right]^{1/2}.$$

6. NUMERICAL EXAMPLE

The examples in this section are borrowed from a detailed analysis of women's demography, presented in a forthcoming paper [6]. We model the life of a woman as progressing in stages through 22 age classes: 0 year, from 1 to 4 years, and then five-year intervals: from 5 to 9 years, from 10 to 14 years, ..., until the interval from 95 to 99 years, and finally a class for women beyond 100 years. These age classes correspond to the transient phases in the MAP. During the ith interval, a woman may give birth to a daughter at the rate b_i and she may die at the rate d_i.

The time unit is the year and the matrix D_0 is given by

$$D_0 = \begin{pmatrix} * & 1 & & & & \\ & * & 1/4 & & & \\ & & * & 1/5 & & \\ & & & \ddots & & \\ & & & & * & 1/5 \\ & & & & & * \end{pmatrix},$$

where a $*$ on the diagonal indicates that the corresponding value is such that $D_0\boldsymbol{1} + D_1\boldsymbol{1} + \boldsymbol{d} = \boldsymbol{0}$. The birth rates matrix B is given by $B = (\boldsymbol{\alpha}\otimes D_1)$ where $\boldsymbol{\alpha} = [1, 0, \ldots, 0]$ is a 1×22 row vector of initial probabilities, reflecting the fact that the lifetime of a newborn always starts in the 0-year age class.

We make the assumption that the mortality and fertility rates stay constant over time and that the population does not undergo any migration phenomena. We also suppose that there is no dependence between the individuals who form the populations.

The parameter values are chosen so as to represent the women population in three countries: Japan, Turkey and the Democratic Republic of Congo. The vector $\boldsymbol{d}$ of death rates is obtained by adapting the life table of each country for 2003, while the matrix $D_1 = \mathrm{diag}(\boldsymbol{b})$ is obtained by adapting the age-specific fertility rates. The data are obtained from the World Health Organization web-site [14] for the life tables, and the United Nations web-site [13] for the age-specific fertility rates, and are adapted as follows:

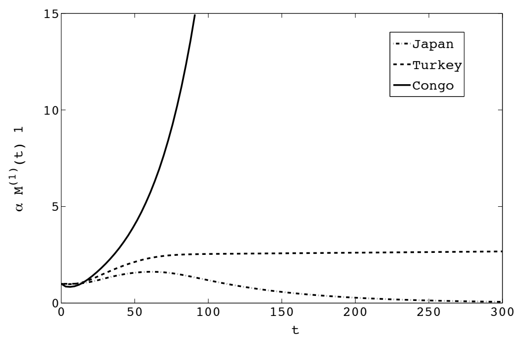

Figure 1: Mean family size

- The death rates d_x in the age class x in the life tables is computed as the ratio between the number of deaths during the year of women in age class x and the female population in the same class at mid-year;
- The age-specific fertility rates give the number of living births during the calendar year, per 1000 female resident population of a given age class x at 30 June. Thus, to obtain the birth rate b_x, we divide the fertility rate by 1000 and then again by $sr + 1$, where sr is the sex ratio, to obtain only the mean number of daughters per woman.

We chose these three countries to cover the range of possible behavior: Japan is sub-critical with $\rho(\Omega) = -0.015$, Turkey is nearly critical with $\rho(\Omega) = 2.48\cdot 10^{-4}$, and Congo is supercritical with $\rho(\Omega) = 3.15\cdot 10^{-2}$.

For the three countries, we compare different characteristic features.

- The mean total size $E[X_1(t) + \cdots + X_{22}(t)|\varphi_0 = 0]$ of the family generated by a first woman born at year 0 is shown on Figure 1 as a function of t; the explosive nature of the parameters for Congo are quite obvious, we also see that the mean value for Turkey is steadily increasing, albeit at a much slower pace.
- We show on Figure 2 the distribution of the time until extinction of the family generated by a first woman born at year 0. We observe that for Congo, the limit $\boldsymbol{q}$ is approximately equal to 0.4, and that for Turkey, the convergence is very slow. The conditional mean time until extinction of the family generated by a woman born at time 0, given extinction occurs, is given in Table 1 below; the small value for Congo reflects the rapid increase of the distribution function to its limit, on the contrary, the huge value for Turkey reflects the long tail of the distribution.
- We show on Figure 3 the mean progeny size in the family after $t = 100$ years, as a function of the age of

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7428
http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7428

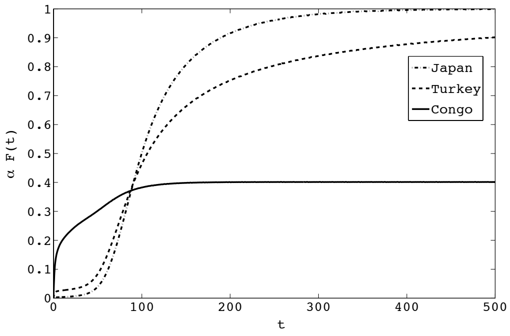

Figure 2: Distribution of the time until extinction

Japan	123.9	Turkey	8428.8	Congo	28.9

Table 1: Conditional mean time until extinction

the first woman at time 0. We observe the effect of infantile mortality in the curve for Congo: if a woman survives to age 15, her expected progeny 100 years later is about double what it is for a new-born woman.

- Finally, we give in Table 2 the conditional mean total progeny size until extinction, given extinction occurs, for a woman born at time 0 and its conditional standard deviation.

Generally, when projections are made in demography, the first woman is supposed to be born at time zero. Here, the model of the MBT allows us to consider also the cases where at time zero, the first woman may be in any age class.

Let us also emphasize that we only look at the female line of descendants of a woman, meaning that if this woman has only sons during her life, in our model, her total progeny is restricted to herself only. This probably explains why the mean family size at a given time in Japan is so low.

Nevertheless, the model of the Markovian binary tree applied to demography is useful in helping to understand the dynamics of the population growth of these countries, and to compare their behavior.

For instance, the infantile mortality, quite important in some countries, appears very clearly on Figure 3 as we already mentioned, but also on Figure 2. It is also the reason why the conditional mean time until extinction is the lowest for Congo, despite the fact that its population is exploding on average. Finally, Turkey has very large conditioned mean time until extinction, mean total progeny size and standard deviation compared to the other countries. This is somewhat unexpected, but the reason actually is quite simple: the MBT modeling the population in Turkey is almost critical, the matrix Θ is well conditioned, but it has an eigenvalue equal to $-2.48 \cdot 10^{-4}$, close to zero, and the inverse of Θ in (23, 30) amplifies the values of these moments.

Remark 2. The numerical results here have been produced without taking into account the special structure of B resulting from the fact that all new born start in phase 1 and the parent does not change its phase after giving birth.

Indeed, the special structure does not lead to new approaches for the resolution of the various equations, although

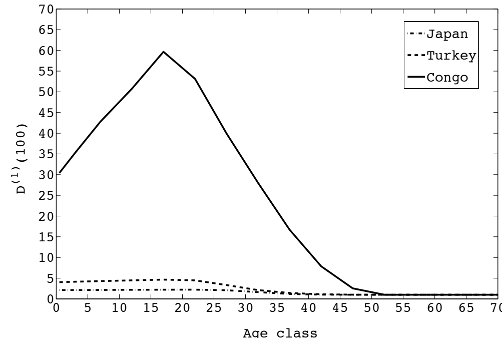

Figure 3: Expectation of $N(t)$ for $t = 100$ years

Country	$(\bar{\boldsymbol{D}}^{(1)})_1$	$(\bar{\boldsymbol{\sigma}}_D)_1$
Japan	2.57	3.5
Turkey	136.88	1806.4
Congo	1.58	1.56

Table 2: Mean and standard deviation of the total progeny size in the family generated by a woman born at time 0, given that extinction occurs

it would reduce the complexity of some computations, if taken into account. For instance, the numerous products of the form $B\,(\boldsymbol{v} \otimes \boldsymbol{v})$, where $\boldsymbol{v}$ is an $n \times 1$ vector, would have a complexity $\mathcal{O}(n)$ instead of $\mathcal{O}(n^3)$.

Acknowledgment

We thank Sophie Alexander for her help in finding the data on life tables and age-specific fertility rates, as well as Jean-Pierre Grimmeau for his useful comments about our numerical demographic illustrations.

The first author is a research fellow of the Fonds National de la Recherche Scientifique (F.N.R.S.), part of her research has been supported by a F.R.I.A. grant.

7. REFERENCES

[1] J. Abate and W. Whitt. Numerical inversion of probability generating functions. *Operations Research Letters*, 12(4):245–251, 1992.

[2] K. Athreya and P. Ney. *Branching Processes.* Springer-Verlag, New York, 1972.

[3] N. Bean, N. Kontoleon, and P. Taylor. Markovian trees: Properties and algorithms. *Ann. Oper. Res.*, 160(1):31–50, 2008.

[4] J. R. Dormand and P. J. Prince. A family of embedded Runge-Kutta formulae. *Journal of Computational and Applied Mathematics*, 6:19–26, 1980.

[5] T. Harris. *The Theory of Branching Processes.* Dover, New York, 1963.

[6] S. Hautphenne and G. Latouche. The Markovian binary tree applied to demography. Technical Report 593, Université Libre de Bruxelles, 2009.

[7] S. Hautphenne, G. Latouche, and M.-A. Remiche. Newton's iteration for the extinction probability of a Markovian Binary Tree. *Linear Algebra Appl.*, 428:2791–2804, 2008.

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7428
http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7428

[8] S. Hautphenne, G. Latouche, and M.-A. Remiche. Algorithmic approach to the extinction probability of branching processes. *Methodol Comput Appl Probab*, 2009, to appear.

[9] S. Hautphenne and B. Vanhoudt. On a link between markovian trees and tree-structured markov chains. *European J. Oper. Res.*, 2009. In press.

[10] G. Latouche and V. Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modeling.* ASA-SIAM Series on Statistics and Applied Probability, SIAM, Philadelphia, PA, 1999.

[11] G. Latouche, M.-A. Remiche, and P. Taylor. Transient Markov arrival processes. *The Annals of Applied Probability*, 13(2):628–640, 2003.

[12] E. C. MacRae. Matrix derivatives with an application to an adaptive linear decision problem. *The Annals of Statistics*, 2(2):337–346, 1974.

[13] World population prospects: The 2006 revision. Web page, United Nations Population Division, http://esa.un.org/unpp/, 20th Sept, 2007.

[14] WHO statistical information system (WHOSIS). Web page, World Health Organization, http://www.who.int/whosis/en/, 4th Dec, 2008.

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7428
http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7428