Ann Oper Res (2008) 160: 31–50
DOI 10.1007/s10479-007-0295-9

Markovian trees: properties and algorithms

Nigel G. Bean · Nectarios Kontoleon · Peter G. Taylor

Published online: 6 December 2007

Abstract In this paper we introduce a structure called the Markovian tree (MT). We define the MT and explore its alternative representation as a continuous-time Markovian Multitype Branching Process. We then develop two algorithms, the Depth and Order algorithms to determine the probability of eventual extinction of the MT process. We show that both of these algorithms have very natural physically intuitive interpretations and are analogues of the Neuts and U algorithms in Matrix-analytic Methods. Furthermore, we show that a special case of the Depth algorithm sheds new light on the interpretation of the sample paths of the Neuts algorithm.

Keywords Branching processes · Matrix analytic methods

1 Introduction

The continuous-time Markovian Multitype Branching Process (ctMMTBP) (Athreya and Ney 1971; Harris 1963) can be used for modeling a variety of phenomena. A special case of a ctMMTBP occurs when branch points are restricted to generate at most two new offspring in which case it is known as the binary branch point ctMMTBP, a process that was used in Kontoleon (2005) and Pinelis (2003) to model biological phenomena such as phylogenetic processes and macroevolution.

In Kontoleon (2005) the evolution of branches in tree histories was given an interpretation different from that which underlies the ctMMTBP. Under the ctMMBTP interpretation, a branch point can occur in which a branch of one type dies and a single branch of another type is born. In the interpretation of Kontoleon (2005), this is thought of as a continuation of the same branch with a change of *phase* of some underlying modulating process. Such a

N.G. Bean (✉) · N. Kontoleon
Applied Mathematics, University of Adelaide, South Australia 5005, Australia
e-mail: nigel.bean@adelaide.edu.au

P.G. Taylor
Department of Mathematics and Statistics, The University of Melbourne, Victoria 3010, Australia

transition is hidden from the point of view of the tree topology. There are two other types of branch points: (i) a binary branch point generating one new daughter and the continuation of the parent, and (ii) a terminating branch point. These are both observable from the point of view of the tree topology.

This interpretation led to an alternative representation for this type of branching process called the Markovian Binary Tree (MBT) (Kontoleon 2005). The evolution of each branch is governed by a (Transient) Markovian Arrival Process (MAP) (Latouche et al. 2003; Lucantoni 1991; Neuts 1979). MAPs have two types of transitions: hidden and observable, which correspond to the hidden and observable events described above. In essence the binary branch point ctMMTBP was represented as a level-dependent quasi-birth and death process (Bright and Taylor 1995). This representation was shown to be much more amenable to algorithmic analysis than the binary branch point ctMMTBP and thus more suitable for use as a macroevolutionary model (Kontoleon 2005).

Using a similar argument, in this paper we re-interpret the general ctMMTBP. This interpretation allows us to introduce the Markovian Tree. Treating singular branch points as hidden transitions and all others as observable ones, confers to the ctMMTBP a structure where the branch length distribution and the offspring distributions are in fact correlated.

The Markovian tree representation of the general ctMMTBP enables a field that is almost devoid of algorithmic approaches (Dorman et al. 2004) to become subject to the powerful techniques of matrix-analytic theory (Latouche and Ramaswami 1999). We show in this paper that this representation leads to some simple, physically intuitive algorithms for calculating the probability of eventual extinction for Markovian trees. These algorithms are called the Depth and Order algorithms, which are the generalisations of the Depth and Order algorithms from Kontoleon (2005). These algorithms have natural analogues in the matrix-analytic literature (Latouche and Ramaswami 1999; Neuts 1976, 1981) as the Neuts and U algorithms for QBDs. In fact, we show that there is a one-to-one mapping from the sample paths of the Depth algorithm to the sample paths of the Neuts algorithm (Neuts 1976, 1981) if we restrict the Markovian tree process to be binary.

In Sect. 2 we define the MT representation. In Sect. 3 the equivalent ctMMTBP representation is stated and explored. In Sect. 4 a general Markovian tree labelling system is introduced. In Sect. 5 the Depth algorithm is discussed. The Depth algorithm is equivalent to the Harris algorithm for the discrete-time multi-type branching process (Harris 1963), the difference being however, that we have given this algorithm a novel and interesting physical interpretation. Section 6 proves that there exists a one-to-one mapping from the sample paths of the Depth algorithm to the sample paths of the Neuts algorithm. Finally, we define the order of a Markovian tree in Sect. 7 and then in Sect. 8 we discuss the Order algorithm for the MT.

2 The Markovian tree: definition

The Markovian tree is a continuous-time Markov chain with states,

$$X(t) = \big(N(t), \phi_1(t), \ldots, \phi_{N(t)}(t)\big)$$

defined on the state space $\bigcup_{k=0}^{\infty}\{\{k\} \times \{1, \ldots, n\}^k\}$. The random variable, $N(t)$, denotes the number of living branches at time t and the random variables, $\phi_k(t) \in \{1, 2, \ldots, n\}$, for all $k \in \{1, \ldots, N(t)\}$ denote the phase of the k-th branch at time t. Level m of a Markovian

Tree, denoted $\mathcal{L}(m)$, is the set of states with m branches alive and so contains the n^m possible states

$$\mathcal{L}(m) = \{(m, \phi_1, \ldots, \phi_m) \big| (\phi_1, \ldots, \phi_m) \in \{1, 2, \ldots, n\}^m\}.$$

Level $\mathcal{L}(0)$ contains the single state with zero branches, that is, $\mathcal{L}(0) = \{(0)\}$.

Now, suppose that at time t the process is in a state with M branches and let branch $k \leq M$ be in phase r. The current state of the process therefore has the form

$$\begin{array}{cccccccc} (M, & a, & \ldots, & b, & r, & c, & \ldots, & d) \\ & 1 & \cdots & k-1 & k & k+1 & \cdots & M \end{array},$$

where each branch is labelled by the number below that branch. The following transitions involving branch k are then possible:

- A hidden transition to phase $j \neq r$, occurs with rate $(D_0)_{rj}$. This transition causes the state of the MT to become

$$\begin{array}{cccccccc} (M, & a, & \ldots, & b, & j, & c, & \ldots, & d) \\ & 1 & \cdots & k-1 & k & k+1 & \cdots & M \end{array}.$$

- A catastrophe occurs on branch k with rate d_r. This causes branch k to cease to exist and the new state is

$$\begin{array}{ccccccc} (M-1, & a, & \ldots, & b, & c, & \ldots, & d) \\ & 1 & \cdots & k-1 & k & \cdots & M-1 \end{array}.$$

  The branches that were previously labelled $k+1, \ldots, M$ have been re-labelled to $k, \ldots, M-1$.
- An observable transition that generates m branches for some $m \geq 1$ occurs. Such an observable transition occurs with rate $(B_m)_{r,i_0 i_1 \ldots i_{m-1} i_m}$. For reasons to be made clear later, we designate the right most branch of a branch point, in this case the branch in phase i_m, to be the continuation of the parental branch and all the other branches to be the daughter branches of that parental branch. The new state of the MT is

$$\begin{array}{cccccccccccc} (M+m, & a, & \ldots, & b, & i_0, & \ldots, & i_{m-1}, & i_m, & c, & \ldots, & d) \\ & 1 & \cdots & k-1 & k & \cdots & k+m-1 & k+m & k+m+1 & \cdots & M+m \end{array},$$

  where the daughter branches in phases $i_0, i_1, \ldots, i_{m-1}$ are designated $k, \ldots, k+m-1$, the parental branch is now the $k+m$-th branch and the branches that were previously labelled $k+1, \ldots, M$ have been re-labelled to $k+m+1, \ldots, M+m$.

The above discussion illustrates that the evolution of each branch is governed by a (transient) BMAP, see Latouche et al. (2003), Lucantoni (1991), with parameters $(\boldsymbol{d}, D_0, B_m, m \geq 1)$. At a branch point, all of the branches have their own realisation of that BMAP. So each branch is free to evolve independently of the rest of the tree. The fact that their evolution is governed by a BMAP also implies that each branch length is not exponentially distributed but rather has a phase type distribution (Latouche et al. 2003; Lucantoni 1991; Neuts 1979) which is a generalisation of the exponential distribution. Furthermore, the branch length distribution and the offspring distribution are correlated (Kontoleon 2005). It is these correlations that are of some importance in phylogenetic tree topology modelling (Kontoleon 2005; Mooers and Heard 1997; Pinelis 2003).

The transition rate matrix for the process has the block structure

$$Q = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \cdots \\ A_{-1}^{(1)} & A_0^{(1)} & A_1^{(1)} & A_2^{(1)} & A_3^{(1)} & \ddots \\ \mathbf{0} & A_{-1}^{(2)} & A_0^{(2)} & A_1^{(2)} & A_2^{(2)} & \ddots \\ \mathbf{0} & 0 & A_{-1}^{(3)} & A_0^{(3)} & A_1^{(3)} & \ddots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{1}$$

The $n^k \times n^k$ matrices $A_0^{(k)}$ are given by,

$$A_0^{(k)} = A_0^{(k-1)} \oplus D_0, \quad \text{for } k \geq 1, \tag{2}$$

with $A_0^{(0)} = 0$. Note the matrix D_0 has the property that $(D_0)_{ii} < 0$ and $(D_0)_{ij} \geq 0$ for $1 \leq i \neq j \leq n$. The interpretation of $A_0^{(k)}$ is as follows: $A_0^{(k)}$ gives the rate at which one of the k branches undergoes a hidden transition that does not generate any new branches. The matrices $A_{-1}^{(k)}$ are of dimension $n^k \times n^{k-1}$ and are given by

$$A_{-1}^{(k)} = \sum_{j=0}^{k-1} I^{(j)} \otimes \boldsymbol{d} \otimes I^{(k-1-j)}, \quad k \geq 1 \tag{3}$$

where $I^{(\ell)}$ is the $n^\ell \times n^\ell$ identity matrix. The $n \times 1$ vector $\boldsymbol{d}$ has components $d_i \geq 0$ for $1 \leq i \leq n$ with at least one component being strictly greater than zero. The matrix $A_{-1}^{(k)}$ is the rate matrix for transitions in which one of the k branches becomes extinct, leaving $k-1$ remaining branches. For $m \geq 1$, the $A_m^{(k)}$ matrices are of dimension $n^k \times n^{k+m}$ and can be written in the form

$$A_m^{(k)} = \sum_{j=0}^{k-1} I^{(j)} \otimes B_m \otimes I^{(k-1-j)}, \quad k \geq 1. \tag{4}$$

The nonnegative element, $(B_m)_{i,i_0 i_1 \ldots i_{m-1} i_m}$, gives the rate at which a branch point occurs at which immediately after birth the parental branch is in phase i_m and the m new daughter branches are in phases, $i_0, i_1, \ldots, i_{m-1}$, given that the parent branch was in phase i immediately before the birth. Therefore, $A_m^{(k)}$ gives the rate at which a single branch from the k possible branches will give rise to m daughter branches in the one transition.

Finally it should also be noted that

$$\boldsymbol{d} = -\left(D_0 \boldsymbol{e}_0 + \sum_{m=1}^{\infty} B_m \boldsymbol{e}_m \right),$$

where $\boldsymbol{e}_m$, for all $m = 0, 1, \ldots$, are vectors of the appropriate dimensions whose elements are all ones.

Having now discussed the definition of the MT we write the MT in its alternative representation as a continuous-time Markovian Multitype Branching Process.

3 The Markovian tree: ctMMTBP representation

3.1 Definition

Let $\boldsymbol{f}(\boldsymbol{s}) = (f^{(1)}(\boldsymbol{s}), f^{(2)}(\boldsymbol{s}), \dots, f^{(n)}(\boldsymbol{s}))$ be the generating function of the offspring probability distribution for the Markovian tree. Let $\boldsymbol{x}$ be an $n \times 1$ vector, and let $\boldsymbol{x}^{(m)}$ denote the m-fold Kronecker product of the vector $\boldsymbol{x}$. In other words, $\boldsymbol{x}^{(m)}$ is defined by,

$$\boldsymbol{x}^{(m)} = \boldsymbol{x}^{(m-1)} \otimes \boldsymbol{x}, \tag{5}$$

with $\boldsymbol{x}^{(1)} = \boldsymbol{x}$. We define the vector, $\widehat{\boldsymbol{d}}$, by,

$$\hat{d}_i = \frac{d_i}{-(D_0)_{ii}}, \tag{6}$$

for all $i = 1, \dots, n$, the matrix, $\widehat{D}_0$, to be

$$(\widehat{D}_0)_{ij} = \frac{(D_0)_{ij}}{-(D_0)_{ii}}, \tag{7}$$

for $1 \le i \ne j \le n$ with $(\widehat{D}_0)_{ii} = 0$, and for $m \ge 1$ the matrix $\widehat{B}_m$ to be

$$(\widehat{B}_m)_{i,i_0\dots i_{m-1}i_m} = \frac{(B_m)_{i,i_0\dots i_{m-1}i_m}}{-(D_0)_{ii}}, \tag{8}$$

for all $i_0, i_1, \dots, i_{m-1}, i_m \in \{1, \dots, n\}$.

The generating function of the offspring probability distribution can then be written as

$$\boldsymbol{f}(\boldsymbol{s}) = \widehat{\boldsymbol{d}} + \widehat{D}_0 \boldsymbol{s} + \sum_{m=1}^{\infty} \widehat{B}_m \boldsymbol{s}^{(m+1)}. \tag{9}$$

The MT representation of the ctMMTBP directly focuses on the space of tree topologies. This is so, because the MT representation distinguishes transitions that result in new branches being formed, the observable transitions, and those that do not result in new branches, the hidden transitions. A hidden transition in a realisation of a ctMMTBP can be considered not to generate a new branch. Thus for the MT we consider the branches of a realisation and the phase process that acts on these branches. Furthermore, the phase process on a branch generates correlations between branch lifetimes, which now may be non-exponential, and the phases of the daughter branches at their birth (Kontoleon 2005); correlations that are not obvious from the ctMMTBP representation.

3.2 Regularity and mean number of branches

In order for the process to be regular (that is, non-explosive) (Athreya and Ney 1971, p. 201) we require that

$$\frac{\partial f^{(i)}(\boldsymbol{s})}{\partial s_j}\Big|_{\boldsymbol{s}=\boldsymbol{e}} < \infty, \text{ for all } i, j = 1, \dots, n, \tag{10}$$

where $\boldsymbol{e} = (1, 1, \dots, 1)$.

The expected number of branches in phase j given that the process began in phase i can be calculated using (Athreya and Ney 1971, p. 184)

$$M(t) = \exp(At), \tag{11}$$

where $A_{ij} = -(D_0)_{ii} b_{ij}$ and

$$b_{ij} = \left.\frac{\partial f^{(i)}(\boldsymbol{s})}{\partial s_j}\right|_{s=e} - \delta_{ij}. \tag{12}$$

In the case of the MT, the matrix A can be expressed in the form

$$A = D_0 + \sum_{m=1}^{\infty} B_m C_m, \tag{13}$$

where $(C_m)_{i_0 \dots i_{m-1} i_m, j}$ is the matrix that counts how many of the $m+1$ branches emanating from a node are in phase j immediately after the creation of that node, that is,

$$(C_m)_{i_0 \dots i_{m-1} i_m, j} = \sum_{k=0}^{m} I\{i_k = j\}. \tag{14}$$

Let λ_A be the dominant eigenvalue of A. Then the process is

- Subcritical if $\lambda_A < 0$,
- Critical if $\lambda_A = 0$, and
- Supercritical if $\lambda_A > 0$.

3.3 Probability of eventual extinction

The final property we wish to discuss in this section is the probability of eventual extinction, denoted by the vector $\boldsymbol{q}$. It is well known from branching process theory (Athreya and Ney 1971, p. 186) that if $\lambda_A \le 0$ the process will eventually become extinct almost surely and if $\lambda_A > 0$ then $\boldsymbol{q} < \boldsymbol{e}$ component-wise. It is this final case that interests us the most. The probability of eventual extinction of a continuous-time Markovian multi-type branching process, (Athreya and Ney 1971, p. 108), is the minimal non-negative solution to

$$\boldsymbol{u}(\boldsymbol{s}) = \boldsymbol{0}, \tag{15}$$

where $u_i(\boldsymbol{s}) = (-D_0)_{ii}(f_i(\boldsymbol{s}) - s_i)$. For the MT

$$\boldsymbol{u}(\boldsymbol{s}) = \boldsymbol{d} + D_0 \boldsymbol{s} + \sum_{m=1}^{\infty} B_m \boldsymbol{s}^{(m+1)}, \tag{16}$$

using (9). We therefore have that $\boldsymbol{q}$ is the minimal non-negative solution to

$$\boldsymbol{d} + D_0 \boldsymbol{s} + \sum_{m=1}^{\infty} B_m \boldsymbol{s}^{(m+1)} = \boldsymbol{0}. \tag{17}$$

We multiply this equation by $(-D_0)^{-1}$ and re-arrange to obtain,

$$\boldsymbol{s} = (-D_0)^{-1} \boldsymbol{d} + \sum_{m=1}^{\infty} (-D_0)^{-1} B_m \boldsymbol{s}^{(m+1)}, \tag{18}$$

which is the form that is most useful for the discussion of the Depth and Order algorithms in Sects. 5 and 8.

4 Labelling the nodes of an MT

In the sections that follow it is important to have a labelling system for the branch points, or nodes of a tree. We begin by labelling the first non-root node of a tree by [0]. Now, consider the node that has label,

$$[0, i_1, i_2, \dots, i_l]$$

where $i_1, \dots, i_l$ are non-negative integers, and suppose that $m+1$ branches emanate from this node; m of these being the daughter branches and one of these being the parental branch. We label the nodes at the tips of the daughter branches as

$$[0, i_1, i_2, \dots, i_l, 0]\ [0, i_1, i_2, \dots, i_l, 1]\ \dots\ [0, i_1, i_2, \dots, i_l, m-1]$$

and we label the tip of the parental branch, called the parental subnode, by

$$[0, i_1, i_2, \dots, i_l, m].$$

For a node labelled $[\psi]$, the number $|\psi|$ of indices gives the *depth* of the node. Let $\alpha(\psi)$ be the mapping that moves up the tree from $[\psi] = [0, i_1, \dots, i_{l-1}, i_l]$ to its parent node, $[0, i_1, \dots, i_{l-1}]$. With this notation, the root node of the tree is $[\alpha(0)]$. This labelling system is depicted in Fig. 1 for a Markovian tree.

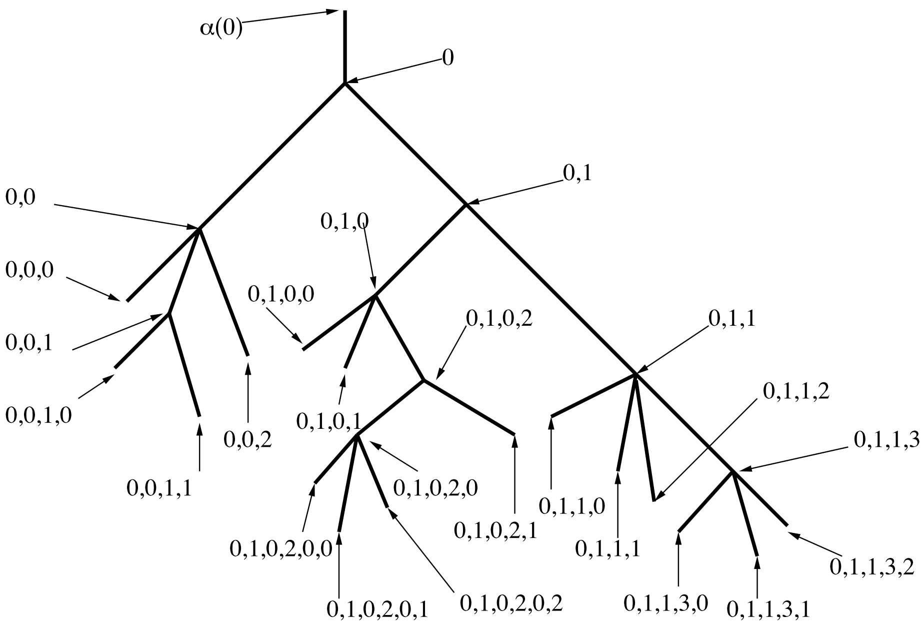

Fig. 1 Labelling the nodes in a tree and tree topology

The segment of a branch between the nodes $[\alpha(\psi)]$ and $[\psi]$ is the ordered pair, $([\alpha(\psi)],[\psi])$. We write $([\alpha(\psi)],[\psi])^{(i)}$ if this segment is an internal segment. We write $([\alpha(\psi)],[\psi])^{(e)}$ if the branch is extinct, and finally we write $([\alpha(\psi)],[\psi])^{(u)}$ if this branch has not yet become extinct or internal (that is, it is a non-extinct leaf). We call these branches *unevolved*. If a superscript is not specified then we just refer to the segment generically; its type is unimportant.

Definition 1 (Tree Topology) The topology of a tree is the branching pattern of that tree when the lengths of the branches are ignored. Note that this is still a function of time, as the tree (and its topology) continue to evolve until its extinction.

Now, suppose there are $m+1$ subtrees that emanate from node $[\psi]$; m of these being the daughter branches and one of these being the parental branch. We represent the topology of the tree based around $[\psi]$ by the ordered set,

$$\{([\alpha(\psi)],[\psi])^{(i)},\mathcal{T}_{[\psi,0]},\mathcal{T}_{[\psi,1]},\ldots,\mathcal{T}_{[\psi,m-1]},\mathcal{T}_{[\psi,m]}\},$$

where $\mathcal{T}_{[\psi,j]}$ is the topology of the j-th daughter subtree that is based around $[\psi,j]$, for all $j=0,1,2,\ldots,m-1$ and $\mathcal{T}_{[\psi,m]}$ is the topology of the parental subtree based around the parental subnode $[\psi,m]$.

Let the set of branch points of a tree of topology $\mathcal{T}$ be denoted by $\mathbb{B}_{\mathcal{T}}$ and let the set of leaf nodes of a tree of that same topology be denoted by $\mathbb{L}_{\mathcal{T}}$. We then have

$$\mathbb{N}_{\mathcal{T}}=\mathbb{B}_{\mathcal{T}}\cup\mathbb{L}_{\mathcal{T}},$$

as the set of nodes of a tree of topology $\mathcal{T}$.

Since the number of daughter branches that are generated at each internal node, $[\psi]$, is finite but possibly unbounded, we let,

$$\sigma(\psi)=\max\{j:[\psi,j]\in\mathbb{N}_{\mathcal{T}}\},\tag{19}$$

be the total number of branches that emanate from $[\psi]$. The parental branch is always the branch that is created from nodes, $[\psi]$ and $[\psi,\sigma(\psi)]$, that is

$$([\psi],[\psi,\sigma(\psi)]).$$

Suppose that $[\psi]$ is either the root node or an internal node, then let the function, θ be defined by,

$$\theta(\alpha(0))=[0],$$

and for $[\psi]\neq[\alpha(0)]$,

$$\theta(\psi)=[\psi,\sigma(\psi)].$$

The function θ is well defined, and maps a node, $[\psi]$ of depth $|\psi|$ to the parental subnode, $[\psi,\sigma(\psi)]$ that emanates from ψ and which is at a depth of $|\psi|+1$. Therefore $\theta^k(\psi)$ traces the pathway of the parental branch that commences from node $[\psi]$, provided that the parental branch is of at least length k from node $[\psi]$. Clearly, $\theta^0(\psi)=[\psi]$. Finally, if $[\psi]$ is a node, then, $\phi(\alpha(\psi))$ denotes the phase of the branch $([\alpha(\psi)],[\psi])$ immediately after $[\alpha(\psi)]$. The phase of the parental branch, $([\psi],[\theta(\psi)])$ immediately after the node $[\psi]$ is denoted by $\phi_p(\psi)$.

Before we commence the discussion of the algorithms to determine the probability of eventual extinction we define topological isomorphism which is important below.

Definition 2 (Topological Isomorphism) Two trees are topologically isomorphic if by a suitable interchange of the branch labels at some or all of the nodes of a tree, the two trees can be made identical.

Note that although two trees may be topologically isomorphic we do not consider them to have identical topologies.

5 The depth algorithm

The Depth algorithm for an MT is the continuous-time analogue of the algorithm of Harris (Harris 1963). We show that the Depth algorithm has a very interesting physical interpretation, but first we define the Depth of a Markovian tree.

Definition 3 (Markovian Tree Depth) The depth, $\delta(\mathcal{T})$, of an MT of topology $\mathcal{T}$ is defined to be,

$$\delta(\mathcal{T}) = \max_{\psi \in \mathbb{B}_{\mathcal{T}}} \{|\psi|\}.$$

The depth of the MT depicted in Fig. 2 can be easily calculated as shown below,

$$\begin{aligned}
\delta(\mathcal{T}) &= \max_{\psi \in \mathbb{B}_{\mathcal{T}a}} \{|\psi|\} \\
&= \{|0|, |0, 1|, |0, 2|, |0, 3|, |0, 1, 0|, |0, 1, 3|, |0, 3, 1|, |0, 3, 1, 1|, |0, 3, 1, 1, 0|\} \\
&= 5.
\end{aligned}$$

Lemma 1 $\lim_{t\to\infty} |\mathcal{T}(t)| < \infty$, *almost surely, if and only if* $\lim_{t\to\infty} \delta(\mathcal{T}(t)) < \infty$, *almost surely.*

Proof Since all states with a non-zero number of branches are transient (Athreya and Ney 1971, p. 186), (Harris 1963, Sects. II.6, II.7), $\lim_{t\to\infty} |\mathcal{T}(t)| < \infty$, almost surely, if and

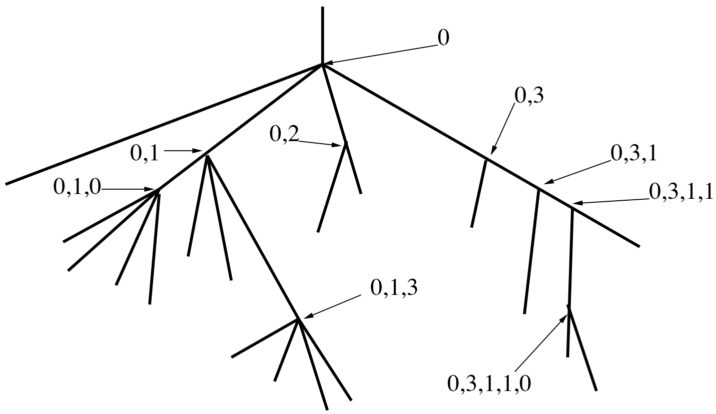

Fig. 2 An MT of depth 5

only if the tree becomes extinct, almost surely. Further, the tree is extinct, in the limit as $t \to \infty$, if and only if on every branch there are a finite number of nodes and so $\lim_{t\to\infty} \delta(\mathcal{T}(t)) < \infty$. □

Since trees that eventually become extinct are of finite depth, almost surely, the probability of eventual extinction of an MT, conditional on starting in phase i is

$$\begin{aligned} q_i &= P\big[|\mathcal{T}_{[\psi]}| < \infty \big| \phi\big(\alpha(\psi)\big) = i\big] \\ &= P\big[\delta(\mathcal{T}_{[\psi]}) < \infty \big| \phi\big(\alpha(\psi)\big) = i\big], \end{aligned} \tag{20}$$

and we write $\boldsymbol{q} = (q_1, \dots q_n)$. Let $\boldsymbol{q}(l)$ be the probability that a tree, $\mathcal{T}_{[\psi]}$ commencing with one branch will eventually become extinct and $\delta(\mathcal{T}_{[\psi]}) < l + 1$, for all $l \geq 0$. That is,

$$\boldsymbol{q}(l) = P\big[\delta(\mathcal{T}_{[\psi]}) < l + 1 \big| \phi\big(\alpha(\psi)\big)\big]. \tag{21}$$

Notice that

$$\boldsymbol{q}(0) = (-D_0)^{-1}\boldsymbol{d}, \tag{22}$$

because an extinct tree of zero depth cannot undergo any observable transitions.

Theorem 2 *The sequence* $\{\boldsymbol{q}(l)\}$, *for* $l \geq 0$, *defined by* (21) *and* (22) *is monotonically increasing and converges to the vector* $\boldsymbol{q}$. *The sequence* $\{\boldsymbol{q}(l)\}$ *also satisfies*,

$$\boldsymbol{s}(0) = (-D_0)^{-1}\boldsymbol{d}, \tag{23}$$

$$\boldsymbol{s}(l) = (-D_0)^{-1}\boldsymbol{d} + \sum_{m=1}^{\infty} (-D_0)^{-1} B_m \boldsymbol{s}^{m+1}(l-1), \quad l \geq 1. \tag{24}$$

Proof The fact that $\{\boldsymbol{q}(l)\}$ is monotonically increasing is obvious. That it converges to $\boldsymbol{q}$ is also obvious since $\lim_{l\to\infty} \boldsymbol{q}(l) = \lim_{l\to\infty} P[\delta(\mathcal{T}_{[\psi]}) < l + 1|\phi(\alpha(\psi))] = P[\delta(\mathcal{T}_{[\psi]}) < \infty|\phi(\alpha(\psi))] = \boldsymbol{q}$.

To show that $\boldsymbol{q}(l)$ satisfies (24) we describe the physical evolution of the process. There are only two pathways with which a tree can eventually become extinct with depth $\delta(\mathcal{T}_{[\psi]}) < l + 1$. The first is a direct extinction, also called a catastrophic event (Latouche et al. 2003), where the parent branch becomes extinct before any non-singular branch points. The probability of this scenario is just $(-D_0)^{-1}\boldsymbol{d}$. In the second pathway, the parent undergoes an observable transition at node $[\psi]$ spawning a finite but unbounded number of daughters with probability $\sum_{m=1}^{\infty} (-D_0)^{-1} B_m$. Clearly, in order for the tree to eventually become extinct with depth $\delta < l + 1$, all the daughter subtrees and the parental subtree must each independently become extinct with depths $\delta < l$. The probability of this second pathway is given by $\sum_{m=1}^{\infty} (-D_0)^{-1} B_m \boldsymbol{q}^{m+1}(l-1)$. Hence we have that,

$$\boldsymbol{q}(l) = (-D_0)^{-1}\boldsymbol{d} + \sum_{m=1}^{\infty} (-D_0)^{-1} B_m \boldsymbol{q}^{m+1}(l-1). \tag{25}$$

and the proof is complete. □

At the l-th step of the algorithm the space of extinct topologies that are measured includes all those topologies from step $l - 1$, that is, those topologies that have depths, $\delta < l$, plus all those topologies of depth $\delta = l$. Consequently, at each step the number of new topologies included is finite.

6 A new interpretation of the sample paths of the Neuts algorithm

The Depth algorithm sheds a new light on the sample paths that are included at each iteration of the Neuts algorithm (1979, 1981) which was the first algorithm proposed for determining the minimal non-negative solution G to the matrix quadratic equation

$$A_2 + A_1 X + A_0 X^2 = 0, \tag{26}$$

that arises in the study of Quasi-Birth-and-Death (QBD) processes. The matrices A_0, A_1 and A_2 in (26) define the transition rates for a level-independent QBD.

The algorithm of Neuts (1976, 1981) was developed by considering the sequence of matrices defined by,

$$G(0) = (-A_1)^{-1} A_2, \tag{27}$$

$$G(l) = (-A_1)^{-1} A_2 + (-A_1)^{-1} A_0 G^2(l-1), \tag{28}$$

for $l \geq 1$. Neuts showed that the sequence $G(\ell)$ is non-decreasing and converges to G.

Consider the set $\mathbb{S}^l$, which contains all sample paths of the QBD that begin in $\mathcal{L}(m)$ and which eventually visit, $\mathcal{L}(m-1)$, such that the maximum level reached is $\mathcal{L}(m+l)$ with the added restriction that each of the sample paths has at most 2^l left transitions (that is, transitions that decrease the level by one).

The l-th iteration of the Neuts algorithm takes into account the probability mass of sample paths in the set $\mathbb{S}^l$. However, the space of sample paths included at the l-th iteration is a strict subset of $\mathbb{S}^l$. To illustrate this, Fig. 3 depicts two sample paths from $\mathbb{S}^2$. Sample path 1 is included at the step $l = 2$ of the Neuts algorithm but sample path 2 is not; sample path 2 is considered at the next step of the algorithm. Up to now, a description of the subset has not been given. Below, we give such a description in terms of the Depth algorithm.

If we re-write (24) for the case when the MT is a binary tree we obtain

$$\boldsymbol{s}(l) = (-D_0)^{-1} \boldsymbol{d} + (-D_0)^{-1} B_1 \boldsymbol{s}^2(l-1), \quad l \geq 1. \tag{29}$$

This equation looks remarkably similar to

$$G(l) = (-A_1)^{-1} A_2 + (-A_1)^{-1} A_0 G^2(l-1), \tag{30}$$

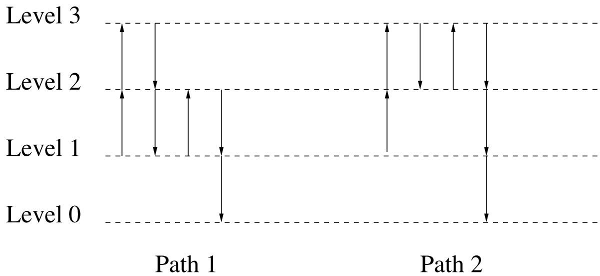

Fig. 3 Two sample paths in $\mathbb{S}^2$

for the level-independent QBD process. This suggests that there exists some relationship between the sample paths of the Neuts algorithm and the tree topologies of the Depth algorithm. In fact, there is a very intimate relationship: the set of sample paths measured at each step of the Neuts algorithm can be transformed to the set of tree topologies that are measured at the equivalent step of the Depth algorithm; this transformation is one-to-one.

There are two types of transitions in the Neuts algorithm, left and right transitions (that is, transitions that decrease the level by one and increase the level by one, respectively). There are also two types of transitions in the Depth algorithm, branch extinctions and internal branch point generation. It then seems natural to apply the following transformation,

- Left transition $\rightarrow$ branch extinctions, and
- Right transition $\rightarrow$ internal node.

More formally, if the left most unevolved branch is $([\alpha(\psi)], [\psi])^{(u)}$, then the next right transition in the Neuts algorithm generates an internal node at $[\psi]$ by creating a new daughter branch. Immediately after this transition then, we have

$$[\psi] \rightarrow \left(([\alpha(\psi)], [\psi])^{(i)}, ([\psi], [\psi, 0])^{(u)}, ([\psi], [\psi, 1])^{(u)}\right),$$

where $([\psi], [\psi, 0])^{(u)}$ and $([\psi], [\psi, 1])^{(u)}$ are the unevolved daughter and parental branches, respectively. On the other hand the next left transition of the Neuts algorithm makes the branch $([\alpha(\psi)], [\psi])^{(u)}$ extinct. In other words,

$$([\alpha(\psi)], [\psi])^{(u)} \rightarrow ([\alpha(\psi)], [\psi])^{(e)}.$$

Consider a sequence of Right, Right and then Left transitions in the Neuts algorithm, applied as the transformation to the tree $[0]$. Applying the first Right transition we obtain,

$$\left(([\alpha(0)], [0])^{(i)}, ([0], [0, 0])^{(u)}, ([0], [0, 1])^{(u)}\right).$$

After the completion of the first transition we are left with a tree that has two unevolved branches, $([0], [0, 0])^{(u)}$ and $([0], [0, 1])^{(u)}$. The left most unevolved branch is $([0], [0, 0])^{(u)}$. The next transition therefore acts on the branch $([0], [0, 0])^{(u)}$. This is again a Right transition and so generates a branch point at $[0, 0]$, leaving,

$$\left(([\alpha(0)], [0])^{(i)}, ([0], [0, 0])^{(i)}, ([0, 0], [0, 0, 0])^{(u)}, ([0, 0], [0, 0, 1])^{(u)}, ([0], [0, 1])^{(u)}\right).$$

The next transition is a Left transition, and we apply this to the left-most unevolved branch which in this case is $([0, 0], [0, 0, 0])^{(u)}$, and so we get,

$$\left(([\alpha(0)], [0])^{(i)}, ([0], [0, 0])^{(i)}, ([0, 0], [0, 0, 0])^{(e)}, ([0, 0], [0, 0, 1])^{(u)}, ([0], [0, 1])^{(u)}\right).$$

It is obvious that this transformation is indeed well defined and one-to-one. This is due to the simple mapping procedure that we apply; each transition has its unique position. At each step, k, we can now easily map each sample path of that step to a binary tree of depth less than or equal to k and therefore give an easily identifiable description to the sample paths of the Neuts algorithm in terms of binary tree topology.

The formal details of this correspondence including all the proofs can be found in Kontoleon (2005).

7 The order of an MT: definition

Let $N(\mathcal{T}_{[\psi]})$ denote the set of internal nodes along the parental branch of $\mathcal{T}_{[\psi]}$, that is $N(\mathcal{T}_{[\psi]}) = \{\theta^k([\psi])\}_{k\in\{0,1,2,\dots\}}$.

Definition 4 (Order)

1. A single branch MT is of order zero.
2. An MT of topology $\mathcal{T}$ is order 1 if at each internal parental node $n \in N(\mathcal{T})$ all the daughter MTs are single branches and hence of order zero.
3. An MT of topology $\mathcal{T}$ is order l if at each internal parental node $n \in N(\mathcal{T})$ all the daughter MTs are trees of order at most $l-1$ and where for some n there exists at least one daughter MT of order $l-1$.

We denote the order of a tree with topology $\mathcal{T}$ to be $\mathbb{O}(\mathcal{T})$.

It is clear from the above definition that we need to calculate the order of an MT recursively. To illustrate the procedure we calculate the order of the topology depicted in Fig. 4. It is not difficult to see that the order of the tree, $\mathcal{T}$, depicted in the figure is $\mathbb{O}(\mathcal{T}) = 3$. This can be seen because the most complex subtree, $\mathcal{T}_{[0,0]}$, has order 2 since $\mathcal{T}_{[0,0,1]}$ has order 1.

Trees that are topologically isomorphic can have different orders. Rotating the nodes changes the parental branch and the daughter subtrees at each node along the parental branch and this changes the calculation of the order.

At each and every internal node of the parental branch if the order of the node is at most $l-1$ then the order of the tree is at most l. Parental branches with any number of internal nodes can be created and so the number of topologies with order l is infinite. This fact is important in understanding the convergence of the Order algorithm, which we discuss in the next section.

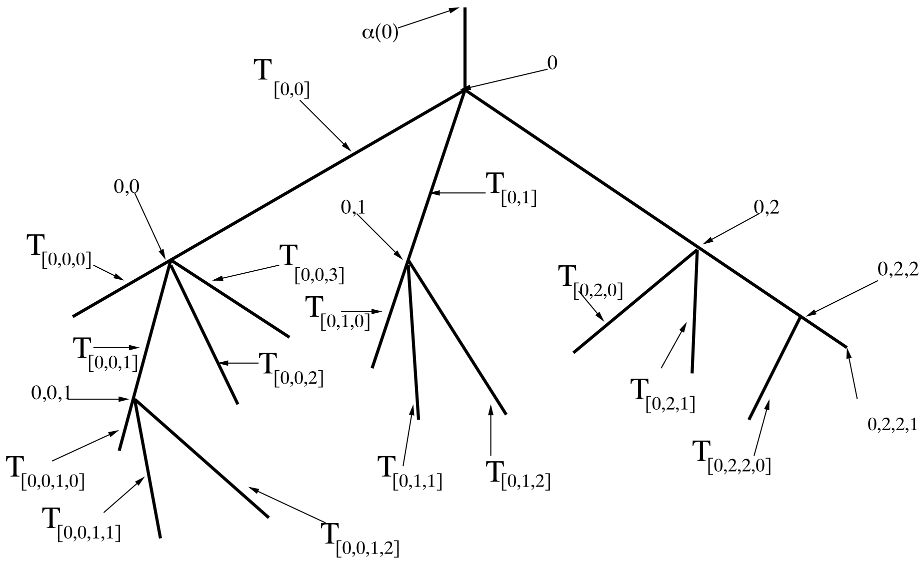

Fig. 4 An example of an order calculation

Lemma 3 $\lim_{t\to\infty}|\mathcal{T}|<\infty$ *if and only if* $\mathbb{O}(\mathcal{T})<\infty$, *almost surely.*

Proof Since all the states with a non-zero number of branches are transient (Athreya and Ney 1971, p. 186; Harris 1963, Sects. II.6, II.7), we have that $\lim_{t\to\infty}|\mathcal{T}(t)|<\infty$, almost surely, if and only if the tree becomes extinct, almost surely. Further, the tree is extinct if and only if on every branch there are a finite number of nodes and so, $\mathbb{O}(\mathcal{T}(t))<\infty$ as $t\to\infty$, almost surely. □

8 The order algorithm

The Order algorithm, as we shall see, is a significant improvement on the Depth algorithm. Let us re-write (18) as

$$\boldsymbol{s}=(-D_0)^{-1}\boldsymbol{d}+\sum_{m=1}^{\infty}(-D_0)^{-1}B_m\big(\boldsymbol{s}^m\otimes I^{(1)}\big)\boldsymbol{s}, \tag{31}$$

where $I^{(1)}$ is the $n\times n$ identity matrix. If we substitute this equation into the right hand side we obtain,

$$\begin{aligned}\boldsymbol{s}&=(-D_0)^{-1}\boldsymbol{d}+\sum_{m=1}^{\infty}(-D_0)^{-1}B_m\big(\boldsymbol{s}^m\otimes I^{(1)}\big)(-D_0)^{-1}\boldsymbol{d}\\&\quad+\left(\sum_{m=1}^{\infty}(-D_0)^{-1}B_m\big(\boldsymbol{s}^m\otimes I^{(1)}\big)\right)\left(\sum_{m=1}^{\infty}(-D_0)^{-1}B_m\big(\boldsymbol{s}^m\otimes I^{(1)}\big)\right)\boldsymbol{s}\\&=(-D_0)^{-1}\boldsymbol{d}+U(\boldsymbol{s})(-D_0)^{-1}\boldsymbol{d}+U^2(\boldsymbol{s})\boldsymbol{s},\end{aligned}$$

where $U(\boldsymbol{s})=\sum_{m=1}^{\infty}(-D_0)^{-1}B_m(\boldsymbol{s}^m\otimes I^{(1)})$. Now if we do this l times, we obtain,

$$\boldsymbol{s}=\sum_{k=0}^{l}U^k(\boldsymbol{s})(-D_0)^{-1}\boldsymbol{d}+R_l(\boldsymbol{s}), \tag{32}$$

where $R_l(\boldsymbol{s})=U^{l+1}(\boldsymbol{s})\boldsymbol{s}$ is the remainder term. Now if we take the limit as $l\to\infty$ we obtain,

$$\boldsymbol{s}=\sum_{k=0}^{\infty}U^k(\boldsymbol{s})(-D_0)^{-1}\boldsymbol{d}+R(\boldsymbol{s}), \tag{33}$$

where $R(\boldsymbol{s})=\lim_{l\to\infty}R_l(\boldsymbol{s})$. In general the remainder term does not need to be zero, but we show below that for $\boldsymbol{q}$, the minimal non-negative solution of equation (18), the remainder term $R(\boldsymbol{q})$ is zero on physical grounds.

Let

$$U=U(\boldsymbol{q})=\sum_{m=1}^{\infty}(-D_0)^{-1}B_m\big(\boldsymbol{q}^m\otimes I^{(1)}\big). \tag{34}$$

Consider the m-th term in the summation, that is,

$$U_m=(-D_0)^{-1}B_m\big(\boldsymbol{q}^m\otimes I^{(1)}\big). \tag{35}$$

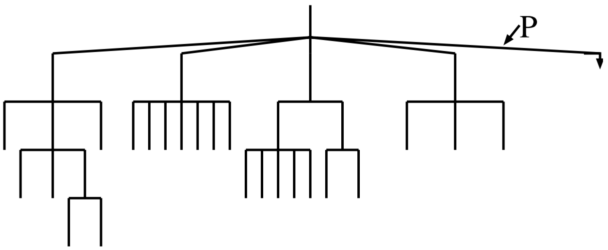

Fig. 5 An example of a U_4-unit

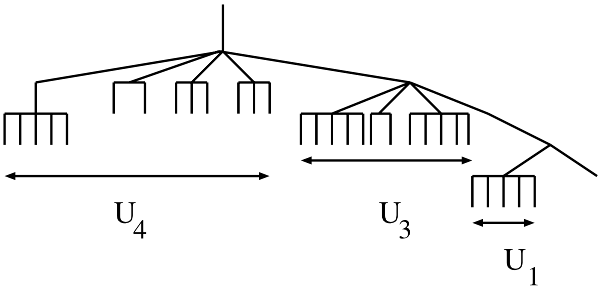

Fig. 6 An example of a tree with three U_k-units

This term gives the probability that there are m daughters spawned at a branch point each of which generates a subtree that eventually becomes extinct. For the purposes of this expression the parental branch remains alive. We call such a structure a U_m-unit; if we do not specify how many daughter branches there are we call the structure a U-unit. Figure 5 represents a U_4-unit. The parent branch gives birth to four daughter branches. These four daughter branches generate subtrees that eventually become extinct, whilst the arrow on the parent branch indicates that its evolution has been suspended, that is, it is an unevolved branch. In general, the suspension of the evolution of the parental branch is made manifest in (35) by the identity matrix, $I^{(1)}$, while the m daughter branches are made extinct by the $\boldsymbol{q}^m$ term. Allowing the parental branch to remain idle while its daughters all become extinct is possible because of the independent evolution of each branch subsequent to its birth. To evaluate the matrix U we sum over all m, since there is no restriction on the number of births.

To construct an extinct tree, we connect the parental subnode of the previous U-unit to the root node of a new U-unit. An extinct tree must be of finite size, almost surely, so only a finite number of U-units can be connected. Following the connection of the final U-unit, the parent branch must undergo a catastrophe before any other observable transition. As an example, Fig. 6 depicts an extinct tree that is constructed from a U_4-unit a U_3-unit and a U_1-unit before final extinction. The probability of obtaining a tree with this description is easily deduced to be, $U_4U_3U_1(-D_0)^{-1}\boldsymbol{d}$. More generally,

$$U^k(-D_0)^{-1}\boldsymbol{d}, \tag{36}$$

gives the probability of generating an extinct tree from k U-units. In other words, the above expression is the probability measure of the space of extinct trees that are constructed from any combination of k U-units, with the parent branch of the k-th unit undergoing a catastrophe to render the tree extinct. Therefore,

$$\sum_{k=0}^{\infty} U^k(-D_0)^{-1}\boldsymbol{d},$$

is the probability measure of the space of all possible extinct trees. Thus

$$\boldsymbol{q} = \sum_{k=0}^{\infty} U^k(-D_0)^{-1}\boldsymbol{d}, \tag{37}$$

and therefore, by (33), $R(\boldsymbol{q}) = 0$.

To recap, the structural subunits that can be used to generate extinct trees are the U-units. These units are connected using their parental branches as we have shown above. Consider a node, $[\psi] \in \mathbb{B}_{\mathcal{T}}$, we can represent a U-unit whose daughter subtrees are spawned at $[\psi]$, as the following ordered set of branches and subtrees,

$$\{([\alpha(\psi)], [\psi])^{(i)}, \mathcal{T}_{[\psi,0]}, \mathcal{T}_{[\psi,1]}, \dots, \mathcal{T}_{[\psi,\sigma(\psi)-1]}, ([\psi], [\theta(\psi)])^{(u)}\},$$

where the parental branch is denoted by $([\psi], [\theta(\psi)])^{(u)}$ to emphasize the fact that this branch is unevolved.

Now, let $\mathcal{B}_m[\theta^k(\psi)]$ be the event that a parental branch which commences from node $[\psi]$ has undergone k observable transitions since $[\psi]$ such that at $\theta^k(\psi)$ m daughter branches are spawned. If we do not specify the number of daughter branches we write, $\mathcal{B}[\theta^k(\psi)]$ for the event that a parental branch which commences from node $[\psi]$ has undergone k observable transitions since $[\psi]$. Let $\mathcal{A}[\theta^k(\psi)]$ be the event that the parental branch has undergone $k-1$ observable transitions since node $[\psi]$ followed by a catastrophic transition at node $[\theta^k(\psi)]$. As before, let $\phi_p(\theta^k(\psi))$ be the phase that the parental branch was in immediately after the $[\theta^k(\psi)]$-th branch point. The initial phase of the parent branch is $\phi_p(\psi)$. The initial phase of a tree of topology $\mathcal{T}_{[\psi]}$ is denoted by $\phi(\alpha(\psi))$.

It is not difficult to see that the order of a U-unit is equal to the order of the daughter tree with the highest order. The above comments tell us that the matrix U has the physical interpretation

$$U = P\big[\mathcal{B}[\theta(\psi)],\ \max_{j<\sigma(\psi)} \mathbb{O}\big(\mathcal{T}_{[\psi,j]}\big) < \infty \;\&\; \phi_p\big(\theta(\psi)\big)\big|\phi_p(\psi)\big], \tag{38}$$

for all $[\psi]$ and the vector $\boldsymbol{q}$ has the physical interpretation

$$\boldsymbol{q} = P\big[\mathbb{O}(\mathcal{T}_{[\psi]}) < \infty\big|\phi\big(\alpha(\psi)\big)\big], \tag{39}$$

for all $[\psi] \neq \alpha(0)$.

Expressed in this way, we see U as being the probability that beginning at some node, $[\psi]$ in phase $\phi_p(\psi)$, a branch point eventually occurs at node $[\theta(\psi)]$, the orders of each of the daughter subtrees are finite, and the parent branch is suspended at node $[\theta(\psi)]$ in phase $\phi_p(\theta(\psi))$. The matrix U is independent of the position of the initial node because the subsequent evolution of any branch that is spawned from that node is independent of the rest of the tree immediately after its birth. The probability $\boldsymbol{q}$ of eventual extinction of the tree $\mathcal{T}_{[\psi]}$ is the probability that $\mathcal{T}_{[\psi]}$ has finite order as $t \to \infty$, given that it commenced from node $[\alpha(\psi)]$ in phase, $\phi(\alpha(\psi))$.

The Order algorithm to determine the minimal non-negative solution to (31) is

$$\boldsymbol{s}(0) = (-D_0)^{-1}\boldsymbol{d}, \tag{40}$$

$$X(l) = \sum_{m=1}^{\infty}(-D_0)^{-1}B_m\big(\boldsymbol{s}^m(l-1) \otimes I^{(1)}\big), \tag{41}$$

$$\boldsymbol{s}(l) = \sum_{k=0}^{\infty} X^k(l)(-D_0)^{-1}\boldsymbol{d}, \tag{42}$$

for $l \geq 1$.

Theorem 4 *The sequences* $\{U(l)|l \geq 1\}$ *and* $\{\boldsymbol{q}(l)|l \geq 0\}$ *defined by*

$$U(l) = P\big[\mathcal{B}[\theta(\psi)],\ \max_{j<\sigma(\psi)} \mathbb{O}\big(\mathcal{T}_{[\psi,j]}\big) < l \;\&\; \phi_p\big(\theta(\psi)\big)|\phi_p(\psi)\big]. \tag{43}$$

and

$$\boldsymbol{q}(l) = P\big[\mathbb{O}(\mathcal{T}_{[\psi]}) < l+1\big|\phi\big(\alpha(\psi)\big)\big], \tag{44}$$

satisfy (40)–(42). *The two sequences are monotonically increasing and respectively converge to the matrix* U *and the vector* $\boldsymbol{q}$.

Proof Trivially the sequences $\{U(l)\}$ and $\{\boldsymbol{q}(l)\}$ defined by (43) and (44) are monotonically increasing and converge to U and $\boldsymbol{q}$ respectively.

The matrix, $U_m(l)$ is the probability that beginning at some node $[\psi]$ and phase $\phi_p(\psi)$, a branch point eventually occurs at node $[\theta(\psi)]$, generating m daughter subtrees such that the orders of each of the daughter subtrees is less than l, so, $\mathbb{O}(\mathcal{T}_{[\psi]}))) \leq l$, and the parental branch is suspended at node $[\theta(\psi)]$ in phase $\phi_p(\theta(\psi))\}$. Therefore we have,

$$\begin{aligned}
U_m(l) &= P\big[\mathcal{B}_m[\theta(\psi)],\ \max_{k<m}\{\mathbb{O}\big(\mathcal{T}_{[\psi,k]}\big)\} < l,\ \&\ \phi_p\big(\theta(\psi)\big)|\phi_p(\psi)\big] \\
&= P\big[\mathcal{B}_m[\theta(\psi)],\ \mathbb{O}\big(\mathcal{T}_{[\psi,0]}\big) < l, \dots, \mathbb{O}\big(\mathcal{T}_{[\psi,m-1]}\big) < l,\ \&\ \phi_p\big(\theta(\psi)\big)|\phi_p(\psi)\big],
\end{aligned}$$

and so each and every daughter subtree must be extinct with order less than l. Each subtree evolves independently conditional on its initial phase and so we can write,

$$\begin{aligned}
U_m(l) = \sum_{\phi(\alpha(\psi,0))} \cdots \sum_{\phi(\alpha(\psi,m-1))} &P\big[([\psi],[\theta(\psi)])^{(i)}, \phi(\alpha(\psi,0)), \dots, \phi(\alpha(\psi,m-1)), \\
&\& \ \phi_p(\theta(\psi))\big|\phi_p(\psi)\big] \\
&\times P\big[\mathbb{O}\big(\mathcal{T}_{[\psi,0]}\big) < l|\phi(\alpha(\psi,0))\big] \dots P\big[\mathbb{O}\big(\mathcal{T}_{[\psi,m-1]}\big) < l|\phi(\alpha(\psi,m-1))\big].
\end{aligned} \tag{45}$$

This follows easily because $\mathcal{B}_m[\theta(\psi)]$ is the event that a branch point occurs to make $([\psi],[\theta(\psi)])$ an internal branch with m daughter branches, and the probability of this is just $(-D_0)^{-1}B_m$; the terms $P[\mathbb{O}(\mathcal{T}_{[\psi,k]}) < l|\phi(\psi,k)]$ are just the probability that the subtree $\mathcal{T}_{[\psi,k]}$ becomes extinct with order at most $l-1$, given the initial phase, in other words, $\boldsymbol{q}(l-1)$, for all $k = 0, 1, \dots, m-1$. We therefore have that (45) is

$$U_m(l) = (-D_0)^{-1}B_m\big(\boldsymbol{q}^m(l-1) \otimes I^{(1)}\big), \tag{46}$$

where the Kronecker product with $I^{(1)}$ represents the fact that the parental branch is suspended with probability 1. Now since a U-unit may be constructed by any finite number of daughter branches we sum over all possibilities and therefore we have that,

$$U(l) = \sum_{m=1}^{\infty} (-D_0)^{-1} B_m \big(\boldsymbol{q}^m(l-1) \otimes I^{(1)}\big). \tag{47}$$

Let $\mathcal{T}_{[\psi]}$ be a topology that is based around node $[\psi]$ whose parental branch has undergone k branch points before undergoing a catastrophic transition. The first of these internal branch points is $[\psi]$. The probability that a tree has k branch points before a catastrophe and has order $\mathbb{O}(\mathcal{T}_{[\psi]}) < l+1$ is given by

$$P\big[\mathcal{A}[\theta^k(\psi)], \mathbb{O}(\mathcal{T}_{[\psi]}) < l+1 \big| \phi(\alpha(\psi))\big], \tag{48}$$

because at the $(k+1)$-st parental subnode, $[\theta^k(\psi)]$, from $[\alpha(\psi)]$, a catastrophic transition occurs rendering the parental branch extinct.

For notational convenience we let

$$\mathcal{O}(\theta^k(\psi)) = \max_{j<\sigma(\theta^k(\psi))} \{\mathbb{O}(\mathcal{T}_{[\theta^k(\psi),j]})\},$$

for $k = 0, 1, 2, \ldots, |N(\mathcal{T}_{[\psi]})|$. Now, following a similar argument to that just used to determine the expression for $U(l)$, we have

$$\begin{aligned}
&P\big[\mathcal{A}\big[\theta^k(\psi)\big], \mathbb{O}(\mathcal{T}_{[\psi]}) < l+1 \big| \phi\big(\alpha(\psi)\big)\big] \\
&\quad = P\big[\mathcal{A}\big[\theta^k(\psi)\big], \max_{i=0,\ldots,k-1} \mathcal{O}\big(\theta^i(\psi)\big) < l \big| \phi\big(\alpha(\psi)\big)\big] \\
&\quad = P\big[\mathcal{A}\big[\theta^k(\psi)\big],\ \mathcal{O}(\psi) < l,\ \mathcal{O}\big(\theta(\psi)\big) < l, \ldots,\ \mathcal{O}\big(\theta^{k-1}(\psi)\big) < l \big| \phi\big(\alpha(\psi)\big)\big].
\end{aligned}$$

But since each of the subtrees evolve independently, conditionally on the initial phases, we can instead write,

$$\begin{aligned}
&P\big[\mathcal{A}[\theta^k(\psi)], \mathbb{O}(\mathcal{T}_{[\psi]}) < l+1 \big| \phi(\alpha(\psi))\big] \\
&\quad = \sum_{\phi_p(\psi)} \cdots \sum_{\phi_p(\theta^{k-1}(\psi))} P\big[\mathcal{B}_{\sigma(\psi)-1}[\psi],\ \mathcal{O}(\psi) < l,\ \Phi(\psi)\ \&\ \phi_p(\psi) \big| \phi\big(\alpha(\psi)\big)\big] \\
&\qquad \times P\big[\mathcal{B}_{\sigma(\theta(\psi))-1}\big[\theta\big(\psi\big)\big],\ \mathcal{O}\big(\theta(\psi)\big) < l,\ \Phi(\theta(\psi))\ \&\ \phi_p\big(\theta(\psi)\big) \big| \phi_p(\psi))\big] \cdots \\
&\qquad \times P\big[\mathcal{B}_{\sigma(\theta^{k-1}(\psi))-1}[\theta^{k-1}(\psi)],\ \mathcal{O}\big(\theta^{k-1}(\psi)\big) < l,\ \Phi(\theta^{k-1}(\psi)) \\
&\qquad\qquad \&\ \phi_p\big(\theta^{k-1}(\psi)\big) \big| \phi_p\big(\theta^{k-2}(\psi)\big)\big] \\
&\qquad \times P\big[\mathcal{A}[\theta^k(\psi)] \big| \phi_p\big(\theta^{k-1}(\psi)\big)\big],
\end{aligned} \tag{49}$$

where $\Phi(\psi)$ is just shorthand for $\phi(\psi, 0), \ldots, \phi(\psi, \sigma(\psi) - 1)$, the phases of each daughter branch emanating from an internal node $[\psi]$. The first k terms of (49) are each just the definition of $U(l)$, and the last term is just equal to $(-D_0)^{-1}\boldsymbol{d}$, because after the k-th branch point of the parental branch, a catastrophe must occur. Hence we can write,

$$P\big[\mathcal{A}[\theta^k(\psi)], \mathbb{O}(\mathcal{T}_{[\psi]}) < l+1 \big| \phi\big(\alpha(\psi)\big)\big] = U^k(l)(-D_0)^{-1}\boldsymbol{d}. \tag{50}$$

However, to obtain $\boldsymbol{q}(l)$ we must sum over all the possible number of branch points of the parental branch, so,

$$\begin{aligned}\boldsymbol{q}(l) &= \sum_{k=0}^{\infty} P\big[\mathcal{A}[\theta^k(\psi)], \mathbb{O}(\mathcal{T}_{[\psi]}) < l+1 \big| \phi\big(\alpha(\psi)\big)\big] \\ &= \sum_{k=0}^{\infty} U^k(l)(-D_0)^{-1}\boldsymbol{d}. \end{aligned} \tag{51}$$

□

A direct consequence of the definition of the order of an MT is that the Order algorithm converges at a faster rate than the Depth algorithm. To show this to be true all we need to do is prove that the set of trees included in each iteration of the Order algorithm contains at least all the trees included in the equivalent iteration of the Depth algorithm. We do this using an inductive argument. It is clear that the set of trees of order 1 contain all the trees of depth 1. Consider now the set of order l trees and suppose that this set also contains all the trees of depth l. Now take any $n \in \mathbb{N}$ trees from the set of order l trees and connect all these trees by a branch point. By definition this gives us all the trees of order $l+1$. Now consider a subset of the set of all order l trees that are restricted to be of depth at most l, since we connect these trees by an initial branch point the resulting tree will necessarily be of depth $l+1$ also. This concludes the proof.

Consider now the Order algorithm restricted to binary branching processes. In this case (40), (41) and (42) become,

$$\boldsymbol{s}(0) = (-D_0)^{-1}\boldsymbol{d}, \tag{52}$$

$$X(l) = (-D_0)^{-1}B_1\big(\boldsymbol{s}^1(l-1)\otimes I^{(1)}\big), \tag{53}$$

$$\boldsymbol{s}(l) = \sum_{k=0}^{\infty} X^k(l)(-D_0)^{-1}\boldsymbol{d}, \tag{54}$$

for $l \geq 1$. It is not difficult to see that (54) can also be written as,

$$\boldsymbol{s}(l) = \big(I^{(1)} - (-D_0)^{-1}B_1\big(\boldsymbol{s}^1(l-1)\otimes I^{(1)}\big)\big)^{-1}(-D_0)^{-1}\boldsymbol{d}. \tag{55}$$

This equation is remarkably similar to the U algorithm as described in (Latouche and Ramaswami 1999),

$$G(l) = \big(I - (-A_1)^{-1}A_0 G(l-1)\big)^{-1}(-A_1)^{-1}A_2. \tag{56}$$

Thus for the binary branching process, the Order algorithm is analogous to the U algorithm just as the binary branching process Depth algorithm is analogous to the Neuts algorithm.

Finally, some elegant observations can be made:

- The set of sample paths included at the ℓ-th step of the Neuts algorithm requires a recursive calculation, whereas the definition of the depth of a tree is a direct definition. The set of sample paths included at the ℓ-th step of the U algorithm can be calculated directly, whereas the definition of the order of a tree involves a recursive definition.
- The U-algorithm converges at a faster rate than the Neuts algorithm, the set of additional sample paths is infinite as opposed to finite. Similarly the Order algorithm converges at a faster rate than the Depth algorithm for the same reason.

9 Conclusion and further work

In this paper we gave the ctMMTBP an alternative interpretation: singular transitions are no longer observable whereas non-singular transitions are. This interpretation led us to an alternative representation of the ctMMTBP, which we have called the Markovian Tree. The Markovian tree, which can be represented using matrix analytic methods, is far more amenable to algorithmic analysis than the usual ctMMTBP representation (Dorman et al. 2004). Furthermore, the MT serves as a bridge between matrix-analytic methodology and branching process theory.

In this paper we presented two algorithms to determine the probability of eventual extinction of Markovian trees: the Depth and Order algorithms. Both these algorithms have physically intuitive interpretations based on the depth of tree topologies, or on the order of tree topologies. The Depth algorithm is essentially the algorithm of Harris (1963) and is linearly convergent with respect to depth. The Order algorithm, which is linearly convergent with respect to order, is a substantial improvement. This follows from the fact that there are an infinite number of topologies of each order whereas there are only a finite number of topologies of each depth.

Since the MT is more amenable to algorithmic analysis, this representation serves as a starting point to develop other algorithms for the ctMMTBP. The development of such algorithms will enable the ctMMTBP to play a more prominent role in modelling physical phenomena.

Acknowledgements The second author would like to thank the Australian Government for providing the Australian Postgraduate Award without which this work would not have been possible. The first and third authors would like to acknowledge the financial support of the Australian Research Council through Discovery-Project DP0209921. Finally, the authors would also like to thank the referees for their invaluable advice and comments.

References

Athreya, K., & Ney, P. (1971). *Branching processes*. Berlin: Springer.
Bright, L., & Taylor, P. (1995). Calculating the equilibrium distribution in level-dependent quasi-birth-and-death processes. *Communications in Statistics. Stochastic Models*, *11*, 497–525.
Dorman, K., Sinsheimer, J., & Lange, K. (2004). In the garden of branching processes. *SIAM Review*, *46*, 202–229.
Harris, T. (1963). *The theory of branching processes*. Berlin: Springer.
Kontoleon, N. (2005). The Markovian binary tree: a model of macroevolution. PhD thesis, The University of Adelaide, Adelaide, S.A. 5005.
Latouche, G., & Ramaswami, V. (1999). *Introduction to matrix analytic methods in stochastic modelling*. Philadelphia: ASA SIAM.
Latouche, G., Remiche, M.-A., & Taylor, P. G. (2003). Transient Markov arrival processes. *Annals of Applied Probability*, *13*, 628–640.
Lucantoni, D. (1991). New results on the single server queue with a batch Markovian arrival process. *Communications in Statistics. Stochastic Models*, *7*, 1–46.
Mooers, A. O., & Heard, S. B. (1997). Inferring evolutionary processes from phylogenetic tree shape. *The Quarterly Review of Biology*, *71*, 31–54.
Neuts, M. (1976). Moment formulas for the Markov renewal branching process. *Advances in Applied Probability*, *8*, 690–711.
Neuts, M. (1979). A versatile Markovian point process. *Journal of Applied Probability*, *16*, 764–779.
Neuts, M. F. (1981). *Matrix-geometric solutions in stochastic models*. Baltimore: John Hopkins Press.
Pinelis, I. (2003). Evolutionary models of phylogenetic trees. *Proceedings of the Royal Society Statistical Society of London*, *203*, 1425–1431. Appendix A available online at http://journals.royalsoc.ac.uk/app/home/contribution.asp?wasp=a8dec894974947a5b93cf7579dc72f2e&referrer-parent&backto=issue,15,15;journal,60,206;homemainpublications,6,7;.