International Journal of Approximate Reasoning 80 (2017) 137–173

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

# Computing lower and upper expected first-passage and return times in imprecise birth–death chains

Stavros Lopatatzidis *, Jasper De Bock, Gert de Cooman

*Ghent University, Data Science Lab, Belgium*

ARTICLE INFO

*Article history:*
Received 13 April 2016
Received in revised form 26 August 2016
Accepted 26 August 2016
Available online 5 September 2016

*Keywords:*
Birth–death chain
First-passage time
Return time
Imprecise probability
Lower expectation
Submartingale

ABSTRACT

We provide simple methods for computing exact bounds on expected first-passage and return times in finite-state birth–death chains, when the transition probabilities are imprecise, in the sense that they are only known to belong to convex closed sets of probability mass functions. In order to do that, we model these so-called imprecise birth–death chains as a special type of time-homogeneous imprecise Markov chain, and use the theory of sub- and supermartingales to define global lower and upper expectation operators for them. By exploiting the properties of these operators, we construct a simple system of non-linear equations that can be used to efficiently compute exact lower and upper bounds for any expected first-passage or return time. We also discuss two special cases: a precise birth–death chain, and an imprecise birth–death chain for which the transition probabilities belong to linear-vacuous mixtures. In both cases, our methods simplify even more. We end the paper with some numerical examples.



## 1. Introduction

Birth–death chains [27, Section 9.4] are a special type of time-homogeneous Markov chains, where transitions from a given state are only possible to that state or to adjacent ones. They are used in various scientific fields, including evolutionary biology [1, Chapter 3] and queueing theory [16]. We consider the generalised case of an imprecise birth–death chain, which, basically, is a birth–death chain whose transition probabilities are not specified exactly, but are only known to belong to some given closed convex set of probability mass functions. This may be the case because the transition probabilities are based on partial expert knowledge or limited data, or for the purposes of conducting a sensitivity analysis. Similar models have already been considered in Reference [6], which presented results on limiting conditional distributions for imprecise birth–death chains with one absorbing state. Imprecise birth–death chains are themselves a special case of so-called (time-homogeneous) imprecise Markov chains, which were studied in—amongst others—References [11,18,25].

This paper focuses on—upward and downward—first-passage and return times.[1] For precise birth–death chains, these have for example been studied in Reference [23]. For the more general case of imprecise birth–death chains, we are not aware of any previous discussion in the literature. Our most important contributions are simple methods for computing exact lower and upper bounds for any expected first-passage and return time in finite-state imprecise birth–death chains. We also consider two special cases: precise birth–death chains and imprecise birth–death chains whose local models are linear-vacuous mixtures. In those cases, our methods lead to closed-form expressions.

* Corresponding author.
*E-mail addresses:* stavros.lopatatzidis@UGent.be (S. Lopatatzidis), jasper.debock@UGent.be (J. De Bock), gert.decooman@UGent.be (G. de Cooman).

[1] These are often called recurrence times as well.

http://dx.doi.org/10.1016/j.ijar.2016.08.008

We start in Section 2 with a brief introduction to the framework of imprecise probabilities, restricting our attention to the simple case of finite state spaces. We introduce credal sets as closed and convex sets of probability mass functions, provide a number of examples, and establish a connection with lower and upper probabilities. This section also introduces the concept of a lower and upper expectation operator, and explains that these operators are mathematically equivalent to credal sets.

Next, we explain what a birth–death chain is in Section 3, and then introduce an imprecise version in Section 4. As we will see, an imprecise birth–death chain is just a special type of a time-homogeneous imprecise Markov chain [11], which, basically, can be regarded as birth–death chain whose local models are credal sets. However, as we will explain, this should not be taken to mean that an imprecise birth–death chain is a collection of birth–death chains. Instead, an imprecise birth–death chain can be regarded as a set of probability trees, only some of which are birth–death chains.

Section 5 defines the global lower and upper expectations that correspond to these imprecise birth–death chains, using the notions of sub- and supermartingales. For real-valued functions that only depend on a finite number of variables, these lower and upper expectations are just the minimum and maximum expectations of this function, with respect to the probability trees in the imprecise birth–death chain. For more general—possibly extended real-valued—functions on the infinite sequence of all variables, the expressions become more intricate. We also recall some convenient properties of the definitions that we adopt, including a global Markov property and a generalised version of the law of iterated expectations.

With all this machinery in place, Section 6 then finally introduces our main topic of interest: return and—upward and downward—first-passage times, and in particular, their lower and upper expectations. As we will see, these lower and upper expected first-passage and return times satisfy a relatively simple system of non-linear equations. The next three sections of the paper are concerned with solving this system, and by doing so, we obtain a simple method for computing the lower and upper expected first-passage and return times that we are interested in. In Sections 7 and 8, we develop recursive methods for computing lower and upper expected upward and downward first-passage times, respectively, and Section 9 explains how these results can be used to compute lower and upper expected return times.

The next two sections are concerned with special cases. Section 10 considers the special case of precise birth–death chains and establishes closed-form expressions for their expected first-passage and return times. It also proves that even though an imprecise birth–death chain is more than just a collection of precise birth–death chains, nevertheless, the lower and upper value of an expected first-passage or return time can always be obtained by a precise birth–death chain. Section 11 discusses the special case where the local models are linear-vacuous mixtures, that is, when they are $\epsilon$-contaminated. Here too, we are able to obtain closed-form expressions.

Section 12 presents some numerical results. We apply our methods for lower and upper first-passage times to a general example, and illustrate our methods for lower and upper return times on an example with local models that are linear-vacuous mixtures.

Finally, Section 13 briefly concludes the paper and mentions some possible avenues for future research. The proofs of our main results are gathered in Appendix A.

## 2. A brief introduction to imprecise probabilities

We start by presenting some basic concepts from the theory of imprecise probabilities. For more information, we refer the reader to Walley's book [26], and to more recent textbooks [2,24].

Consider a variable $X$ that takes values in some non-empty finite set $\mathscr{X}$. A common approach to describe a subject's uncertainty about the actual value of $X$ is then to consider a probability mass function $p$ on $\mathscr{X}$, that is, an element of the set

$$\Sigma_{\mathscr{X}} := \left\{ p \in \mathbb{R}^{\mathscr{X}} : \sum_{x \in \mathscr{X}} p(x) = 1 \text{ and } (\forall x \in \mathscr{X})\, p(x) \geq 0 \right\}.$$

For any real-valued function $f$ on $\mathscr{X}$—also called a *gamble*—the corresponding expectation of $f$ is then given by

$$E_p(f) := \sum_{x \in \mathscr{X}} p(x) f(x).$$

If we now let $\mathscr{G}(\mathscr{X})$ be the set of all gambles, then the expectation operator $E_p : \mathscr{G}(\mathscr{X}) \to \mathbb{R}$ can be regarded as an alternative, equivalent representation for $p$. Indeed, $E_p$ can clearly be inferred from $p$ and, conversely, if we know $E_p$, then for all $x \in \mathscr{X}$, $p(x)$ is equal to the expectation $E_p(\mathbb{I}_x)$ of the indicator $\mathbb{I}_x \in \mathscr{G}(\mathscr{X})$ of $x$, as defined by

$$\mathbb{I}_x(y) := \begin{cases} 1 & \text{if } y = x \\ 0 & \text{if } y \neq x \end{cases} \text{ for all } y \in \mathscr{X}.$$

The success of this approach depends crucially on the assumption that our uncertainty about $X$ can be described by a probability mass function $p$, and furthermore requires that $p$ is specified precisely. However, in practice, eliciting such a probability function can be difficult, especially if it is based on—possibly disagreeing—expert opinions, or when it has to be learned from small amounts of data. Whenever this is the case, the theory of imprecise probabilities [2,24,26] does not

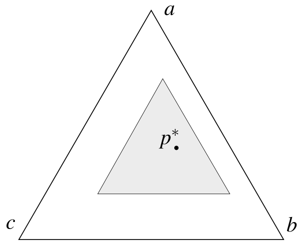


**Fig. 1.** The grey area depicts the linear-vacuous credal set $\Phi_1$ of Example 1.

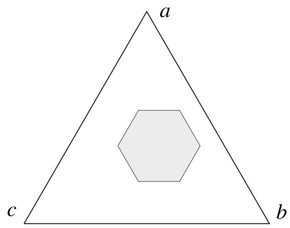


**Fig. 2.** The grey area depicts the credal set $\Phi_2$ of Example 2.

insist on using a single probability mass function $p$, but instead allows for using sets of probabilities. The central object of study then becomes a credal set $\Phi$: a non-empty, convex and closed subset of $\Sigma_{\mathscr{X}}$.

For any such credal set $\Phi$, the corresponding lower and upper probability mass of $x \in \mathscr{X}$ are defined by

$$\underline{p}(x) := \min\{p(x) : p \in \Phi\} \text{ and } \overline{p}(x) := \max\{p(x) : p \in \Phi\}, \tag{1}$$

respectively. The following example illustrates these concepts.

**Example 1.** Let $\mathscr{X} := \{a, b, c\}$, consider a probability mass function

$$p^* = (p^*(a), p^*(b), p^*(c)) = (2/5, 2/5, 1/5)$$

and let $\varepsilon := 1/2$. The corresponding so-called linear-vacuous [26, Section 2.9.2] (or $\epsilon$-contaminated) credal set is then defined by

$$\Phi_1 := \Phi^{\varepsilon}_{p^*} := \{(1-\varepsilon)p^* + \varepsilon p : p \in \Sigma_{\mathscr{X}}\}, \tag{2}$$

which can be regarded as neighbourhood model for the probability mass function $p^*$.

Since the state space $\mathscr{X}$ is ternary, this credal set can be depicted easily. We first represent $\Sigma_{\mathscr{X}}$ by an equilateral triangle of height one. The elements $p = (p(a), p(b), p(c))$ of $\Sigma_{\mathscr{X}}$ then correspond to points in this triangle. For every such $p$, the value of $p(a)$ is equal to the perpendicular distance from $p$ to the edge that opposes the corner that corresponds to $a$, and similarly for $p(b)$ and $p(c)$. In this way, the credal set $\Phi_1$ corresponds to the grey area in Fig. 1. As can be seen from this figure, $\Phi_1$ is the convex hull of the three extreme points $(\overline{p}(a), \underline{p}(b), \underline{p}(c))$, $(\underline{p}(a), \overline{p}(b), \underline{p}(c))$, and $(\underline{p}(a), \underline{p}(b), \overline{p}(c))$. The numerical values of the lower and upper masses in these expressions are

$$\underline{p}(a) = \underline{p}(b) = 1/5,\ \overline{p}(a) = \overline{p}(b) = 7/10,\ \underline{p}(c) = 1/10, \text{ and } \overline{p}(c) = 3/5.$$

They are easily obtained by combining Equations (1) and (2). $\Diamond$

It is not necessary for a credal set to be defined directly, as was the case in Example 1. It can also be specified indirectly, by means of partial constraints on probabilities. A particularly appealing way of doing so is to specify a probability interval for every $x$ in $\mathscr{X}$, and to let $\Phi$ be the largest subset of $\Sigma_{\mathscr{X}}$ that satisfies these constraints. We illustrate this in our next example.

**Example 2.** Let $\mathscr{X} := \{a, b, c\}$ and consider the following probability constraints:

$$p(a) \in [1/5, 8/15],\ p(b) \in [1/5, 8/15] \text{ and } p(c) \in [1/10, 13/30].$$

The largest set of probability mass functions $p \in \Sigma_{\mathscr{X}}$ that satisfies these constraints is then a credal set, which we denote by $\Phi_2$; see Fig. 2. This credal set $\Phi_2$ has six extreme points, the elements of which consist of a lower probability, an upper probability and their complement. These extreme points

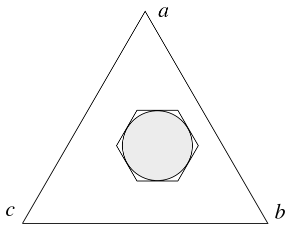


**Fig. 3.** The grey zone depicts the credal set $\Phi_3$ of Example 2. The credal set $\Phi_3$ tangents to each side of the credal set $\Phi_2$ that was described earlier in the same example.

$$(8/15, 1/5, 8/30), (8/15, 11/30, 1/10), (11/30, 8/15, 1/10),$$
$$(1/5, 8/15, 8/30), (1/5, 11/30, 13/30), (11/30, 1/5, 13/30)$$

can be identified in the above order by starting from the upper left corner of the polygon in Fig. 2 and moving clockwise. $\Diamond$

However, as our next example should clarify, a credal set is not always completely characterised by such lower and upper probability masses.

**Example 3.** Let $\Phi_3$ be the circular credal set that is depicted in Fig. 3, which has an infinite number of extreme points. In order to allow for an easy comparison, Fig. 3 also depicts the credal set $\Phi_2$ of Example 2. These two credal sets clearly have the same lower and upper probability masses. $\Diamond$

Therefore, we will not restrict our attention to lower and upper probability masses, but will instead focus on lower and upper expectations.

With any credal set $\Phi$, we can associate a lower expectation operator $\underline{E}\colon \mathscr{G}(\mathscr{X}) \to \mathbb{R}$ and an upper expectation operator $\overline{E}\colon \mathscr{G}(\mathscr{X}) \to \mathbb{R}$, defined by

$$\underline{E}(f) := \min\{E_p(f)\colon p \in \Phi\} \text{ and } \overline{E}(f) := \max\{E_p(f)\colon p \in \Phi\} \text{ for all } f \in \mathscr{G}(\mathscr{X}). \tag{3}$$

Lower and upper expectations are related by conjugacy: for any $f \in \mathscr{G}(\mathscr{X})$, we have that $\overline{E}(f) = -\underline{E}(-f)$. Therefore, it suffices to consider only one of them. We will focus on lower expectations.

For any credal set $\Phi$, the corresponding lower expectation operator $\underline{E}$ can easily be shown to satisfy the following so-called *coherence axioms*:

C1. $\underline{E}(f) \geq \min f$ for all $f \in \mathscr{G}(\mathscr{X})$; [bounds]
C2. $\underline{E}(f+g) \geq \underline{E}(f) + \underline{E}(g)$ for all $f, g \in \mathscr{G}(\mathscr{X})$; [superadditivity]
C3. $\underline{E}(\lambda f) = \lambda \underline{E}(f)$ for all $f \in \mathscr{G}(\mathscr{X})$ and real $\lambda \geq 0$, [non-negative homogeneity]

and, if we let $\overline{E}$ be the corresponding conjugate upper expectation operator, then as a consequence of C1–C3, we also find that

C4. $\underline{E}(f) \leq \underline{E}(g)$ and $\overline{E}(f) \leq \overline{E}(g)$ for all $f, g \in \mathscr{G}(\mathscr{X})$ with $f \leq g$;
C5. $\min f \leq \underline{E}(f) \leq \overline{E}(f) \leq \max f$ for all $f \in \mathscr{G}(\mathscr{X})$;
C6. $\underline{E}(f+\mu) = \underline{E}(f) + \mu$ and $\overline{E}(f+\mu) = \overline{E}(f) + \mu$ for all $f \in \mathscr{G}(\mathscr{X})$ and $\mu \in \mathbb{R}$.

The reason we refer to C1–C3 as axioms is because they capture the essence of what it means to be the lower expectation operator of a credal set. Indeed, as proved in Reference [17], for any operator $\underline{E}\colon \mathscr{G}(\mathscr{X}) \to \mathbb{R}$ that satisfies properties C1–C3,[2] the corresponding credal set

$$\Phi := \{p \in \Sigma_{\mathscr{X}}\colon E_p(f) \geq \underline{E}(f) \text{ for all } f \text{ in } \mathscr{G}(\mathscr{X})\}$$

completely characterises $\underline{E}$, in the sense that $\underline{E}$ can be derived from $\Phi$ by means of Equation (3), and this credal set is furthermore the only credal set for which this is the case. Hence, a credal set $\Phi$ and its lower expectation operator $\underline{E}$ are mathematically equivalent, and therefore, we can—and will—use them interchangeably.

[2] Such an operator is sometimes also called a *coherent lower prevision*, and is then usually given a different interpretation and notation: $\underline{E}(f)$ is then denoted by $\underline{P}(f)$, is called the lower prevision of $f$, and is interpreted as a subject's supremum buying price for the uncertain payoff $f(X)$; see References [20,24,26] for more information.

All of the above concepts and results apply when $\mathscr{G}(\mathscr{X})$ consists of real-valued functions on finite spaces, and when the lower expectation operators that we consider are unconditional. However, although the basic ideas remain similar, technicalities arise when we deal with—possibly unbounded—real-valued functions on infinite spaces, with extended real-valued functions, or with conditional operators; see for example Reference [24, Part II] for a very general treatment. Fortunately, this will not be an issue here, because in the context of stochastic processes, these technicalities can be dealt with in a specific way. We will discuss this in Section 5. For now, a basic understanding of credal sets and lower expectation operators is sufficient.

## 3. Birth–death chains

The main type of model that we will consider in this paper is an imprecise birth–death chain, which is a special stochastic process whose local models are credal sets. Before we discuss this model, we introduce the special case of 'precise' birth–death chains, which are themselves a special type of time-homogeneous Markov chains.

### 3.1. Notation

A birth–death chain models the time-evolution of a system that takes values in a state space $\mathscr{X}$ that is linearly ordered. We will restrict attention to finite state spaces that—in order to avoid trivialities—consist of at least two elements. Therefore, without loss of generality, we have that $\mathscr{X} = \{0, \ldots, L\}$ with $L \in \mathbb{N}$.[3] At any time $n \in \mathbb{N}$, the state of the chain is represented by a random variable, denoted by $X_n$, which takes values in $\mathscr{X}$. For every $n \in \mathbb{N}$, the sequence of variables $X_1, \ldots, X_n$ is denoted by $X_{1:n}$ and also, for every $k \in \mathbb{N}$ such that $k \leq n$, we denote by $X_{k:n}$ the sequence of variables from time $k$ to $n$. A sequence $X_{1:n}$ takes values $x_{1:n} := (x_1, \ldots, x_n)$ in $\mathscr{X}^n$, and similarly for $X_{k:n}$. We call any finite sequence of state values $x_{1:n} \in \mathscr{X}^n$ a *situation* and denote the set of all situations by $\Omega^\lozenge$. For the special case of $n = 0$, we have the so-called *initial situation*, denoted by $\square$. Therefore, $\mathscr{X}^0 := \{\square\}$ and $x_{1:0}$ is the empty sequence. We also allow concatenation of situations with state values or variables. For all $n \in \mathbb{N}_0$, given a situation $x_{1:n} \in \mathscr{X}^n$ and a state $i \in \mathscr{X}$, the concatenation of $x_{1:n}$ and $i$, denoted by $(x_{1:n}, i)$, is a situation in $\mathscr{X}^{n+1}$. Similarly, for any $m > n$, we write $(x_{1:n}, X_{n+1:m})$ to denote the concatenation of the situation $x_{1:n}$ and the sequence of variables $X_{n+1:m}$.

### 3.2. Stochastic processes

A birth–death chain is a special type of (time-homogeneous) Markov chain and therefore a stochastic process, so it is completely determined by its local conditional probability mass functions. For every $n \in \mathbb{N}$ and every situation $x_{1:n}$ in $\mathscr{X}^n$, we need a local probability mass function $p(X_{n+1}|x_{1:n})$ which, for every $x_{n+1} \in \mathscr{X}$, provides us with the probability $p(x_{n+1}|x_{1:n})$ that the state $X_{n+1}$ takes the value $x_{n+1}$, conditional on the information that the current situation of the process is $x_{1:n}$. Similarly, for $n = 0$, we need a local (unconditional) initial model $p(X_1|\square)$. If these local models are available, then for any $n \in \mathbb{N}$ and $m \in \mathbb{N}_0$ such that $m < n$, the probability $p(x_{m+1:n}|x_{1:m})$ that $X_{m+1:n}$ assumes the value $x_{m+1:n}$, conditional on the information that the current situation of the process is $x_{1:m}$, is easily seen to be given by

$$p(x_{m+1:n}|x_{1:m}) = \prod_{k=m}^{n-1} p(x_{k+1}|x_{1:k}). \tag{4}$$

### 3.3. Time-homogeneous Markov chains

A stochastic process is called a Markov chain if its local models satisfy the *Markov condition*:

$$p(X_{n+1}|x_{1:n}) = p(X_{n+1}|x_n) \text{ for all } n \in \mathbb{N} \text{ and } x_{1:n} \in \mathscr{X}^n, \tag{5}$$

where $p(X_{n+1}|x_n)$ is the probability of $X_{n+1}$, conditional on $X_n = x_n$. For the case of *homogeneous* Markov chains, $p(X_{n+1}|x_n)$ furthermore does not depend on $n$, but only on $x_n$. In that case, for every $x \in \mathscr{X}$, there is some probability mass function $\varphi_x$ such that

$$p(X_{n+1}|x_n) = \varphi_{x_n}(X_{n+1}) \text{ for all } n \in \mathbb{N} \text{ and } x_n \in \mathscr{X}. \tag{6}$$

If we then let $\varphi_\square$ be the unique probability mass function such that $\varphi_\square(X_1) = p(X_1|\square)$, Equation (4) reduces to

$$p(x_{m+1:n}|x_{1:m}) = \begin{cases} \varphi_\square(x_1) \displaystyle\prod_{k=1}^{n-1} \varphi_{x_k}(x_{k+1}) & \text{if } m = 0; \\ \displaystyle\prod_{k=m}^{n-1} \varphi_{x_k}(x_{k+1}) & \text{if } m > 0. \end{cases} \tag{7}$$

[3] Zero is excluded from $\mathbb{N}$ and we let $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

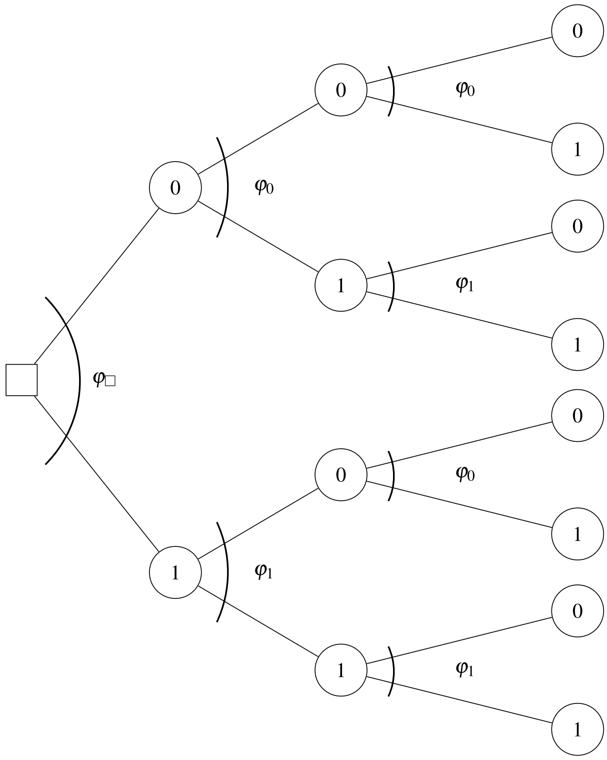


**Fig. 4.** The initial part of the probability tree of a time-homogeneous Markov chain with state space $\mathscr{X} := \{0, 1\}$.

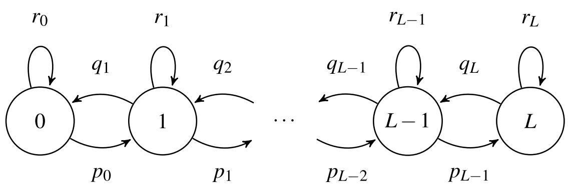


**Fig. 5.** A birth–death chain with state space $\mathscr{X} = \{0, \dots, L\}$.

By depicting the situations $x_{1:n}$ of such a time-homogeneous Markov chain in a tree, and attaching local models to these situations, we obtain a so-called *probability tree* [21]. At each time $n \in \mathbb{N}_0$ we then have a set of possible situations $x_{1:n}$ in $\mathscr{X}^n$, each of which is associated with a local model $p(X_{n+1}|x_{1:n}) = \varphi_{x_n}(X_{n+1})$. Fig. 4 depicts the initial part of an example of such a probability tree, for a time-homogeneous Markov chain with state space $\mathscr{X} = \{0, 1\}$.

### 3.4. Birth–death chains

Since the transition probabilities $p(X_{n+1}|x_n) = \varphi_{x_n}(X_{n+1})$ of a Markov chain do not depend on $n$, they can be conveniently summarised by means of a single stochastic matrix $P$ of dimension $L + 1$, the elements of which are defined by letting $P_{ij} := \varphi_i(j)$ for all $i, j \in \mathscr{X}$. In the special case of a birth–death chain, this stochastic matrix is tridiagonal, which expresses that transitions are only possible between adjacent states. Hence, in that case, $P$ is of the form

$$P = \begin{pmatrix} r_0 & p_0 & 0 & \cdots & \cdots & 0 \\ q_1 & r_1 & p_1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & q_{L-1} & r_{L-1} & p_{L-1} \\ 0 & \cdots & \cdots & 0 & q_L & r_L \end{pmatrix} \tag{8}$$

where the elements of each row sum to 1. For all $i \in \mathscr{X}/\{0, L\}$, we will assume that the probabilities $p_i$, $q_i$ and $r_i$ are strictly positive, and similarly for $r_0, p_0, q_L, r_L$. Such a birth–death chain can be depicted as a probability tree, as in Fig. 4, but it also has a chainlike representation, as depicted in Fig. 5.

## 4. Imprecise birth–death chains

Imprecise birth–death chains are similar to precise birth death chains. The main difference is that the probability mass functions that make up the matrix $P$ do not need to be specified exactly. They are only known to belong to convex closed sets of probability mass functions—credal sets.

### 4.1. Local credal sets

For every $i \in \mathscr{X} \setminus \{0, L\}$, we consider a credal set $\mathscr{Q}_i$ on $\mathscr{X}_m := \{\ell, e, u\}$, where—for reasons that should become clear soon—$m$ stands for middle and $\ell$, $e$ and $u$ stand for lower, equal and upper, respectively. For the individual probability mass functions $\pi_i \in \mathscr{Q}_i$, we will make frequent use of the notational convention that

$$(q_i, r_i, p_i) = \big(\pi_i(\ell), \pi_i(e), \pi_i(u)\big),$$

thereby establishing an intuitive connection with the matrix $P$ that characterises a precise birth–death chain. Similarly, $\mathscr{Q}_0$ and $\mathscr{Q}_L$ are taken to be credal sets on $\mathscr{X}_0 := \{e, u\}$ and $\mathscr{X}_L := \{\ell, e\}$, respectively. For their elements $\pi_0 \in \mathscr{Q}_0$ and $\pi_L \in \mathscr{Q}_L$, we adopt the following analogous notational conventions:

$$(r_0, p_0) = \big(\pi_0(e), \pi_0(u)\big) \text{ and } (q_L, r_L) = \big(\pi_L(\ell), \pi_L(e)\big).$$

Similar notational conventions are also adopted for the lower and upper probabilities that correspond to the credal sets $\mathscr{Q}_i$: in accordance with Equation (1), we define

$$(\forall i \in \mathscr{X}) \quad \underline{r}_i := \min\{r_i : \pi_i \in \mathscr{Q}_i\} \text{ and } \overline{r}_i := \max\{r_i : \pi_i \in \mathscr{Q}_i\};$$

$$(\forall i \in \mathscr{X} \setminus \{0\}) \quad \underline{q}_i := \min\{q_i : \pi_i \in \mathscr{Q}_i\} \text{ and } \overline{q}_i := \max\{q_i : \pi_i \in \mathscr{Q}_i\}; \tag{9}$$

$$(\forall i \in \mathscr{X} \setminus \{L\}) \quad \underline{p}_i := \min\{p_i : \pi_i \in \mathscr{Q}_i\} \text{ and } \overline{p}_i := \max\{p_i : \pi_i \in \mathscr{Q}_i\}. \tag{10}$$

For reasons of mathematical convenience, we will restrict ourselves to credal sets $\mathscr{Q}_i$ that satisfy the following positivity assumption.

**Assumption 1** *(Positivity assumption).* For every $i \in \mathscr{X}$, the local credal set $\mathscr{Q}_i$ consists of strictly positive probability mass functions.

This positivity assumption implies—amongst other useful consequences such as Theorem 4—that $0 < \underline{p}_i \leq \overline{p}_i < 1$ for all $i \in \mathscr{X} \setminus \{L\}$ and that $0 < \underline{q}_i \leq \overline{q}_i < 1$ for all $i \in \mathscr{X} \setminus \{0\}$.

We now use the credal sets $\mathscr{Q}_i$ to define corresponding local credal sets $\Phi_i$ on $\mathscr{X}$. For all $i \in \mathscr{X} \setminus \{0, L\}$, a probability mass function $\phi_i \in \Sigma_{\mathscr{X}}$ belongs to $\Phi_i$ if and only if there is some $\pi_i \in \mathscr{Q}_i$ such that

$$\phi_i(j) = \begin{cases} q_i & \text{if } j = i-1 \\ r_i & \text{if } j = i \\ p_i & \text{if } j = i+1 \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } j \in \mathscr{X}. \tag{11}$$

Similarly, $\phi_0$ belongs to $\Phi_0$ if and only if there is some $\pi_0 \in \mathscr{Q}_0$ such that

$$\phi_0(j) = \begin{cases} r_0 & \text{if } j = 0 \\ p_0 & \text{if } j = 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } j \in \mathscr{X} \tag{12}$$

and $\phi_L$ belongs to $\Phi_L$ if and only if there is some $\pi_L \in \mathscr{Q}_L$ such that

$$\phi_L(j) = \begin{cases} q_L & \text{if } j = L-1 \\ r_L & \text{if } j = L \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } j \in \mathscr{X}. \tag{13}$$

Finally, $\Phi_\square$ is taken to be an arbitrary credal set on $\mathscr{X}$.

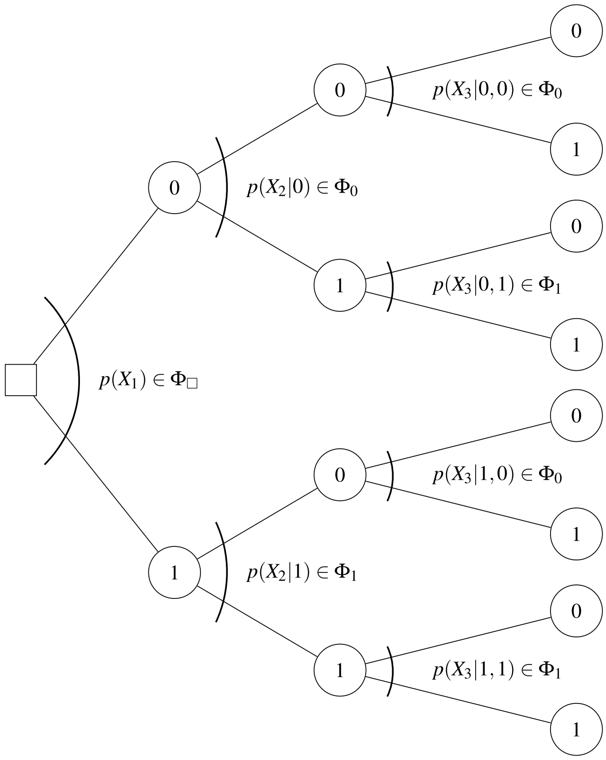


**Fig. 6.** The initial part of one of the probability trees that is compatible with an imprecise Markov chain with state space $\mathscr{X} := \{0, 1\}$.

### 4.2. Imprecise birth–death chains

The local credal sets that we have just introduced can now be used to define an imprecise birth–death chain. Simply put, we just replace the initial probability mass function $\varphi_\Box$ by the initial credal set $\Phi_\Box$ and, for all $i \in \mathscr{X}$, we replace the transition probability mass functions $\varphi_i$ by the local credal sets $\Phi_i$.

More specifically, we consider a probability tree that is partially specified, in the sense that all that we know about it is that it is compatible with these local credal sets, in the sense that

$$p(X_1|\Box) \in \Phi_\Box \tag{14}$$

and, for all $n \in \mathbb{N}$ and $x_{1:n} \in \mathscr{X}^n$,

$$p(X_{n+1}|x_{1:n}) \in \Phi_{x_n}. \tag{15}$$

We will use $\mathscr{T}^\Phi$ to denote the set of all probability trees that are compatible with these constraints. Fig. 6 depicts the initial part of a generic element of this set.

Note that the probability trees in $\mathscr{T}^\Phi$ may not correspond to a (time-homogeneous) Markov chain: we do not require $p(X_{n+1}|x_{1:n})$ and $p(X_{n+1}|x_n)$ to be equal, nor do we enforce that $p(X_{n+1}|x_n)$ should not depend on $n$. This is easily understood by comparing Fig. 4 with Fig. 6. If $\varphi_\Box \in \Phi_\Box$, $\varphi_0 \in \Phi_0$ and $\varphi_1 \in \Phi_1$, then clearly, the time-homogeneous Markov chain that is depicted in Fig. 4 is a special case of the generic probability tree that is depicted in Fig. 6. However, the probability tree that is depicted in Fig. 6 need not be of the type that is depicted in Fig. 4. Consequently, an imprecise Markov chain is not just a set of imprecisely specified birth–death chains. Instead, as we are about to show, $\mathscr{T}^\Phi$ is a set of probability trees that satisfies an imprecise Markov condition.

### 4.3. In terms of lower expectation operators

As explained in Section 2, every credal set has a corresponding lower (and upper) expectation operator, which serves as an alternative—mathematically equivalent—representation for this credal set. Therefore, the local credal sets $\Phi_i$, $i \in \mathscr{X}$, and $\Phi_\Box$ can be equivalently represented by means of lower or upper expectation operators.

For any $i \in \mathscr{X}$, we denote the lower and upper expectation operator that corresponds to $\Phi_i$ by $\underline{Q}_i$ and $\overline{Q}_i$, respectively. For any $f \in \mathscr{G}(\mathscr{X})$, they are given by

$$\underline{Q}_i(f) := \min_{\phi_i \in \Phi_i} E_{\phi_i}(f) = \min_{\phi_i \in \Phi_i} \sum_{j \in \mathscr{X}} \phi_i(j) f(j) \tag{16}$$

and

$$\overline{Q}_i(f) := \max_{\phi_i \in \Phi_i} E_{\phi_i}(f) = \max_{\phi_i \in \Phi_i} \sum_{j \in \mathscr{X}} \phi_i(j) f(j). \tag{17}$$

The lower and upper expectation operators $\underline{Q}_\square$ and $\overline{Q}_\square$ that correspond to $\Phi_\square$ are defined similarly.

In more or less the same way, we can also associate conditional lower expectation operators with the imprecise birth–death chain $\mathscr{T}^\Phi$. For example, for every $n \in \mathbb{N}_0$ and $x_{1:n} \in \mathscr{X}^n$, we can consider the conditional lower expectation operator $\underline{E}_{n+1}(\cdot|x_{1:n})$, as defined by

$$\underline{E}_{n+1}(f|x_{1:n}) := \underline{E}(f(X_{n+1})|x_{1:n}) := \min_{p \in \mathscr{T}^\Phi} \sum_{j \in \mathscr{X}} p(j|x_{1:n}) f(j) \text{ for all } f \in \mathscr{G}(\mathscr{X}), \tag{18}$$

and the conditional upper expectation operator $\overline{E}_{n+1}(\cdot|x_{1:n})$, which is defined similarly. For $n = 0$, we also write $\underline{E}_1(\cdot) := \underline{E}_1(\cdot|\square)$ and $\overline{E}_1(\cdot) := \overline{E}_1(\cdot|\square)$. As a rather straightforward consequence of Equation (15), we then find that $\underline{E}_1 = \underline{Q}_\square$ and $\overline{E}_1 = \overline{Q}_\square$ and, for all $n \in \mathbb{N}$ and $x_{1:n} \in \mathscr{X}^n$, that

$$\underline{E}_{n+1}(\cdot|x_{1:n}) = \underline{Q}_{x_n}(\cdot) \text{ and } \overline{E}_{n+1}(\cdot|x_{1:n}) = \overline{Q}_{x_n}(\cdot). \tag{19}$$

As this equation illustrates, an imprecise birth–death chain $\mathscr{T}^\Phi$ satisfies an *imprecise time-homogeneous Markov condition*, in the sense that lower and upper expectations of functions of the state at time $n+1$, conditional on the value $x_{1:n}$ of the state at all previous times, only depend on the value $x_n$ of the last state; they do not depend on the previous states, nor do they depend on the time $n$. However, as explained before, this is not the case for the individual elements of $\mathscr{T}^\Phi$, which are not required to be Markov chains.

Because they satisfy this imprecise Markov condition, the imprecise birth–death chains that we consider in this paper are a special case of the imprecise Markov chains that were introduced in Reference [11]. Hence, an imprecise birth–death chain is simply an imprecise Markov chain that has the lower and upper expectation operators $\underline{Q}_i$ and $\overline{Q}_i$—see Equations (16) and (17)—as its local models, or equivalently, the credal sets $\Phi_i$.

Imposing Equation (19) rather than Equation (5) can also be regarded as adopting a weaker notion of independence: one could say that we are adopting a notion of 'almost'-independence. To use a specific imprecise-probabilistic terminology: we impose *epistemic irrelevance* rather than strong independence. From this perspective, and for a finite time horizon, an imprecise birth–death chain can be regarded as a very special case of a credal network (an imprecise version of a Bayesian network) under epistemic irrelevance [3,7,8,12]. More information on the various independence concepts that are used in imprecise probabilities can be found in References [4,5,9,13].

## 5. Global lower and upper expectations

So far, the only lower and upper expectations that we have associated with an imprecise birth–death chain $\mathscr{T}^\Phi$ are local ones, for real-valued functions on finite spaces. However, in order to be able to study first-passage and return times, we also need to consider global lower and upper expectations of extended real-valued functions on the infinite set of all paths. We introduce these global lower and upper expectations in this section. Since they are connected through conjugacy, we focus on lower expectations.

### *5.1. Global models for real-valued functions on finite spaces*

As a first intermediate step, we consider lower expectations of real-valued functions that depend on a finite number of variables. For such functions, their lower expectation is simply the lower envelope of all the expectations that correspond to a probability tree in $\mathscr{T}^\Phi$: for all $n \in \mathbb{N}$, we have that

$$\underline{E}(h(X_{1:n})) := \min_{p \in \mathscr{T}^\Phi} \sum_{x_{1:n} \in \mathscr{X}^n} p(x_{1:n}) h(x_{1:n}) \text{ for all } h \in \mathscr{G}(\mathscr{X}^n)$$

and, for all $m \in \mathbb{N}_0$ such that $m < n$ and all $x_{1:m} \in \mathscr{X}^m$, we have that

$$\underline{E}(h(X_{1:n})|x_{1:m}) := \min_{p \in \mathscr{T}^\Phi} \sum_{x_{m+1:n} \in \mathscr{X}^n} p(x_{m+1:n}|x_{1:m}) h(x_{1:n}) \text{ for all } h \in \mathscr{G}(\mathscr{X}^n).$$

In these expressions, $p(x_{1:n}) := p(x_{1:n}|\square)$ and $p(x_{m+1:n}|x_{1:m})$ can be computed as in Equation (4). The upper expectations $\overline{E}(h(X_{1:n}))$ and $\overline{E}(h(X_{1:n})|x_{1:m})$, are defined similarly, and are related to the lower expectations through conjugacy.

In principle, similar definitions could also be used for functions that depend on an infinite number of variables, by making an assumption of countable additivity and using the resulting probability measures to define expectation operators. However, unfortunately, not much is known about the global lower expectation operators that would result from such an approach. For that reason, we will instead consider an alternative approach that is based on submartingales [14], and which is inspired by the game-theoretic probability framework of Shafer and Vovk [22]. The properties of this alternative approach are (at least for now) better understood, and it has the additional advantage that it does not require any measurability conditions, and that it can be applied easily to extended-real valued functions. In the following three subsections, we provide a brief summary of this approach, applied to our specific context of imprecise birth–death chains.

### 5.2. Sub- and supermartingales

A *real process* $\mathscr{F}$ is a real-valued map defined on $\Omega^\Diamond$, which associates a real number $\mathscr{F}(x_{1:n}) \in \mathbb{R}$ with any situation $x_{1:n} \in \mathscr{X}^n$, for all $n \in \mathbb{N}_0$. A *gamble process* is a map from $\Omega^\Diamond$ to $\mathscr{G}(\mathscr{X})$, which associates with any situation $x_{1:n} \in \mathscr{X}^n$ a gamble in $\mathscr{G}(\mathscr{X})$, for all $n \in \mathbb{N}_0$.

With any real process $\mathscr{F}$, we can always associate a corresponding gamble process $\Delta\mathscr{F}$, called the *process difference*. For every situation $x_{1:n} \in \mathscr{X}^n$, the corresponding gamble $\Delta\mathscr{F}(x_{1:n}) \in \mathscr{G}(\mathscr{X})$ is defined by

$$\Delta\mathscr{F}(x_{1:n})(x_{n+1}) := \mathscr{F}(x_{1:n+1}) - \mathscr{F}(x_{1:n}) \text{ for all } x_{n+1} \in \mathscr{X}.$$

In the specific case of an imprecise birth–death chain, a *submartingale* $\mathscr{M}$ is then a real process such that $\underline{Q}_\Box(\Delta\mathscr{M}(\Box)) \geq 0$ and

$$\underline{Q}_{x_n}(\Delta\mathscr{M}(x_{1:n})) \geq 0 \text{ for all } n \in \mathbb{N} \text{ and } x_{1:n} \in \mathscr{X}^n. \tag{20}$$

A *supermartingale* is a real process $\mathscr{M}$ such that $-\mathscr{M}$ is a submartingale, or equivalently, because of conjugacy, such that $\overline{Q}_\Box(\Delta\mathscr{M}(\Box)) \leq 0$ and

$$\overline{Q}_{x_n}(\Delta\mathscr{M}(x_{1:n})) \leq 0 \text{ for all } n \in \mathbb{N} \text{ and } x_{1:n} \in \mathscr{X}^n. \tag{21}$$

A submartingale is *uniformly bounded above* if there is some $B \in \mathbb{R}$, such that $\mathscr{M}(x_{1:n}) \leq B$ for all $n \in \mathbb{N}_0$ and $x_{1:n} \in \mathscr{X}^n$. A supermartingale $\mathscr{M}$ is *uniformly bounded below* if $-\mathscr{M}$ is uniformly bounded above, or equivalently, if there is some $B \in \mathbb{R}$, such that $\mathscr{M}(x_{1:n}) \geq B$ for all $n \in \mathbb{N}_0$ and $x_{1:n} \in \mathscr{X}^n$. The set of all uniformly bounded above submartingales is denoted by $\underline{\mathbb{M}}$, and the set of all uniformly bounded below supermartingales is denoted by $\overline{\mathbb{M}}$; clearly, we have that $\overline{\mathbb{M}} = -\underline{\mathbb{M}}$.

### 5.3. General global models

For every $n$ in $\mathbb{N}$, we use $X_{n:\infty}$ to denote the infinite sequence of variables $(X_n, X_{n+1}, \ldots)$. An instantiation $x_{1:\infty} = (x_1, x_2, \ldots)$ of $X_{1:\infty}$ is called a *path* and is also denoted by $\omega$. We call *sample space* the set of all paths and we denote it by $\Omega := \mathscr{X}^{\mathbb{N}}$. For any path $\omega \in \Omega$, the initial sequence that consists of its first $n$ elements is a situation in $\mathscr{X}^n$, and we denote it by $\omega^n$. Its $n$-th element belongs to $\mathscr{X}$ and is denoted by $\omega_n$. We let its 0-th element be the initial situation $\omega^0 = \omega_0 = \Box$. Furthermore, with any situation $x_{1:n} \in \mathscr{X}^n$, we associate a set

$$\Gamma(x_{1:n}) := \{\omega \in \Omega \colon \omega^n = x_{1:n}\}$$

that consists of all the paths $\omega \in \Omega$ whose initial part is equal to $x_{1:n}$. For $n = 0$, we have that $\Gamma(x_{1:0}) = \Gamma(\Box) = \Omega$.

The global models that we are about to construct will provide lower and upper expectations of *extended real-valued* functions on $\Omega$, where by extended real-valued, we mean that these functions take values in the set $\mathbb{R}^* := \mathbb{R} \cup \{-\infty, +\infty\}$. For any such extended real-valued function $g$ on $\Omega$, we will often explicitly indicate that its value depends on the variables $\{X_i\}_{i\in\mathbb{N}}$, by writing $g(X_{1:\infty})$. The advantage of this notational convention is that it also allows us to create new functions. For example, for any $n \in \mathbb{N}$, we can write $g(X_{n+1:\infty})$ to denote a shifted version of $g(X_{1:\infty})$. Similarly, for any $f \in \mathscr{G}(\mathscr{X})$ and $n \in \mathbb{N}$, we can write $f(X_n)$ to denote a real-valued function on $\Omega$ whose value in $\omega \in \Omega$ is equal to $f(\omega_n)$.

The link between the lower and upper expectations of these extended real-valued functions and the local models of our imprecise birth death-chain is now established by means of the sub- and supermartingales of the previous subsection. First, for any real process $\mathscr{F}$ (and therefore, in particular, for any sub- or supermartingale), we consider the extended real-valued functions $\liminf \mathscr{F}$ and $\limsup \mathscr{F}$, defined for all $\omega \in \Omega$ by

$$\liminf \mathscr{F}(\omega) := \liminf_{n\to\infty} \mathscr{F}(\omega^n) \text{ and } \limsup \mathscr{F}(\omega) := \limsup_{n\to\infty} \mathscr{F}(\omega^n).$$

Next, we use these functions to define global conditional lower and upper expectations. For any $n \in \mathbb{N}_0$, any $x_{1:n} \in \mathscr{X}^n$ and any extended real-valued function $g$ on $\mathscr{X}^{\mathbb{N}}$, the conditional lower expectation of $g$ is defined as

$$\underline{E}(g|x_{1:n}) := \sup\{\mathscr{M}(x_{1:n}) \colon \mathscr{M} \in \underline{\mathbb{M}} \text{ and } \limsup \mathscr{M}(\omega) \leq g(\omega) \text{ for all } \omega \in \Gamma(x_{1:n})\}. \tag{22}$$

Similarly, the conjugate conditional upper expectation of $g$ is defined as

$$\overline{E}(g|x_{1:n}) := \inf\{\mathscr{M}(x_{1:n}) : \mathscr{M} \in \overline{\mathbb{M}} \text{ and } \liminf \mathscr{M}(\omega) \geq g(\omega) \text{ for all } \omega \in \Gamma(x_{1:n})\}. \tag{23}$$

Detailed technical and philosophical discussions about these and other closely related so-called game-theoretic definitions of lower and upper expectations can be found in References [10,14,22]. Basically, the starting point is the observation that a supermartingale $\mathscr{M}$ can be interpreted as a capital process (it represents the evolution of your monetary capital) of which the local changes $\Delta\mathscr{M}(x_{1:n})$ are expected (on 'average') to either increase your capital or keep it steady (because their local lower expectation $\underline{Q}_{x_n}(\Delta\mathscr{M}(x_{1:n}))$ is non-negative). The assumption is then that, since all these local increases are expected to be at least non-negative, the value of $\limsup \mathscr{M}(x_{1:n}, X_{n+1:\infty})$ should be expected to be at least $\mathscr{M}(x_{1:n})$. Since this assumption applies to all $\mathscr{M}$ in $\underline{\mathbb{M}}$, we arrive at Equation (22). Equation (23) follows from a similar argument or from conjugacy. A more detailed discussion—including the reason why we only consider submartingales that are bounded above—can be found in Reference [14].

For our present purposes, it suffices to know that the lower and upper expectations in Equations (22) and (23) are a proper generalisation of the lower and upper expectations that we considered earlier on in this paper. For example, it can be shown that the lower expectations $\underline{E}(h(X_{1:n}))$ and $\underline{E}(h(X_{1:n})|x_{1:m})$ that we defined in Section 5.1 correspond to special cases of Equation (22), with $g = h(X_{1:n})$. Similarly, the lower expectation $\underline{E}(f(X_{n+1})|x_{1:n})$ that was defined in Equation (18) is also a special case of Equation (22), with $g = f(X_{n+1})$, and therefore, it follows from the imprecise Markov condition (19) that

$$\underline{E}(f(X_{n+1})|x_{1:n}) = \underline{Q}_{x_n}(f) \text{ for all } n \in \mathbb{N}, x_{1:n} \in \mathscr{X}^n \text{ and } f \in \mathscr{G}(\mathscr{X}). \tag{24}$$

Proofs of these results, as well as further discussions on the connection between Equations (22) and (23) and the probability trees in $\mathscr{T}^\Phi$ can be found in Reference [14].

### 5.4. Properties of the general global models

We end this section by presenting a number of technical properties of the global expectations that we have just defined. Proofs for these—as well as other and more general—properties can be found in Reference [14].[4]

First of all, our global models satisfy generalised versions of the coherence properties C1–C6 that were discussed in Section 2. The generalised versions of C5 and C6 look as follows:

**C5**. $\inf\{f(\omega) : \omega \in \Gamma(x_{1:n})\} \leq \underline{E}(f|x_{1:n}) \leq \overline{E}(f|x_{1:n}) \leq \sup\{f(\omega) : \omega \in \Gamma(x_{1:n})\}$;
**C6**. $\underline{E}(g + \mu|x_{1:n}) = \mu + \underline{E}(g|x_{1:n})$ and $\overline{E}(g + \mu|x_{1:n}) = \mu + \overline{E}(g|x_{1:n})$,

for all extended real-valued functions $g$ on $\Omega$, all $n \in \mathbb{N}_0$, all $x_{1:n} \in \mathscr{X}^n$ and all $\mu \in \mathbb{R}$. Generalised versions of C1–C4 can also be stated, but since we will not need them in this paper, we restrict attention to the generalised versions of C5 and C6.

Secondly, as a rather immediate consequence of Equation (22), we obtain the following basic proposition.

**Proposition 1.** *Consider any $n \in \mathbb{N}$. Then for any extended real-valued function $g$ on $\Omega$:*

$$\underline{E}(g(X_{1:\infty})|x_{1:n}) = \underline{E}(g(x_{1:n}, X_{n+1:\infty})|x_{1:n}).$$

Thirdly, as the following result establishes, the conditional lower expectations in Equation (22) satisfy a global Markov property: the lower expectation of a (possibly extended) real-valued function of the future variables $X_{n+1:\infty}$, conditional on the value of the past and present variables $X_{1:n}$, depends only on the value of the present variable $X_n$, and not on the value of the past variables, nor on the specific time $n$.

**Proposition 2.** *Consider any $n \in \mathbb{N}$, $x_{1:n-1} \in \mathscr{X}^{n-1}$ and $i \in \mathscr{X}$. Then for any extended real-valued function $g$ on $\Omega$:*

$$\underline{E}(g(X_{n+1:\infty})|x_{1:n-1}, i) = \underline{E}(g(X_{2:\infty})|i).$$

Finally, if we regard the conditional lower expectation $\underline{E}(g|x_{1:m})$ as a function of $x_{1:m}$, and we interpret this function $\underline{E}(g|X_{1:m})$ as a (possibly extended) real-valued function on $\Omega$, then we obtain the following result, which can be regarded as a generalised version of the law of iterated expectations.

**Proposition 3.** *Consider any $n, m \in \mathbb{N}_0$ such that $m \geq n$ and any $x_{1:n} \in \mathscr{X}^n$. Then for any extended real-valued function $g$ on $\Omega$:*

$$\underline{E}(g|x_{1:n}) = \underline{E}(\underline{E}(g|X_{1:m})|x_{1:n}).$$

[4] **C5** and **C6** follow from [14, Proposition 14]; Proposition 1 is identical to [14, Proposition 15]; Proposition 2 is a special case of [14, Proposition 19]; Proposition 3 is identical to [14, Theorem 16].

The conditional upper expectations of Equation (23) satisfy suitably adapted versions of Propositions 1–3; they follow immediately from conjugacy and the respective version for lower expectations.

Hence, in summary, we have found that the conditional lower and upper expectation operators that we defined in Equations (22) and (23) not only extend the finite models in Section 5.1, but also satisfy suitably extended versions of the coherence properties in Section 2 as well as a number of other powerful properties—see Propositions 1–3. The results for first-passage and return times that we are about to present are derived directly from these properties; the only result that explicitly uses Equations (22) and (23) directly, is Theorem 4. Any other (for example measure-theoretic) global lower and upper expectation operator that satisfies Theorem 4 and the properties in this section, will therefore lead to completely identical results.

## 6. First passage and return times

Consider a time $n \in \mathbb{N}$ and two—possibly identical—states $i$ and $j$ in $\mathscr{X}$. Suppose that the value of the current state $X_n$ is equal to $i$, that the values of the previous states $X_{1:n-1}$ are given by $x_{1:n-1}$, and consider the number of time-steps required to reach $j$, or equivalently, the value of $\tau_{\to j}(X_{n+1:\infty})$, where the extended real-valued function $\tau_{\to j}$ is defined by[5]

$$\tau_{\to j}(\omega) := \inf\{m \in \mathbb{N} \colon \omega_m = j\}. \tag{25}$$

We call this number of time-steps the *first-passage time* of $j$ conditional on $X_{1:n} = (x_{1:n-1}, i)$, and when $i = j$, we call it the *return time* of $i$. The so-called upward and downward first-passage times correspond to the cases $i < j$ and $i > j$, respectively. The goal of this paper is to compute the lower and upper expectations $\underline{E}(\tau_{\to j}(X_{n+1:\infty})|x_{1:n-1}, i)$ and $\overline{E}(\tau_{\to j}(X_{n+1:\infty})|x_{1:n-1}, i)$ of these first-passage and return times.

For the lower expected first-passage time of $j$ conditional on $X_{1:n} = (x_{1:n-1}, i)$, it follows from Proposition 2—with $g = \tau_{\to j}$—that

$$\underline{E}(\tau_{\to j}(X_{n+1:\infty})|x_{1:n-1}, i) = \underline{E}(\tau_{\to j}(X_{2:\infty})|i).$$

Hence, for all $i, j$ in $\mathscr{X}$, the lower expected first-passage time $\underline{E}(\tau_{\to j}(X_{n+1:\infty})|x_{1:n-1}, i)$ neither depends on $n$ nor on $x_{1:n-1}$, and we can therefore simply refer to it as the lower expected first-passage time from $i$ to $j$. Similarly, by combining Proposition 2 with conjugacy, it can be shown that $\overline{E}(\tau_{\to j}(X_{n+1:\infty})|x_{1:n-1}, i)$ only depends on $i$ and $j$, and therefore, we can refer to it as the upper expected first-passage time from $i$ to $j$.

In order to reflect these findings in our notation, we will from now on denote the lower and upper expected first-passage time from $i$ to $j$ by $\underline{\tau}_{i\to j}$ and $\overline{\tau}_{i\to j}$, respectively; they are defined by

$$\underline{\tau}_{i\to j} := \underline{E}(\tau_{\to j}(X_{2:\infty})|i) = \underline{E}(\tau_{\to j}(X_{n+1:\infty})|x_{1:n-1}, i) \tag{26}$$

and

$$\overline{\tau}_{i\to j} := \overline{E}(\tau_{\to j}(X_{2:\infty})|i) = \overline{E}(\tau_{\to j}(X_{n+1:\infty})|x_{1:n-1}, i). \tag{27}$$

The following theorem establishes a first convenient property of $\underline{\tau}_{i\to j}$ and $\overline{\tau}_{i\to j}$ that follows from Assumption 1; see Appendix A for a proof.

**Theorem 4.** *For all $i, j \in \mathscr{X}$, the lower and upper first-passage times $\underline{\tau}_{i\to j}$ and $\overline{\tau}_{i\to j}$ are real-valued and strictly positive.*

In the rest of this section, we will derive a system of non-linear equations for these lower and upper expected first-passage times. The starting point for this derivation is the fact that

$$\underline{\tau}_{i\to j} = \underline{E}(\tau_{\to j}(X_{2:\infty})|i) = \underline{E}(\underline{E}(\tau_{\to j}(X_{2:\infty})|X_{1:2})|i), \tag{28}$$

which is a direct consequence of Proposition 3. Next, in order to express $\underline{E}(\tau_{\to j}(X_{2:\infty})|X_{1:2})$ in terms of lower expected first-passage times, we start by observing that

$$\tau_{\to j}(X_{2:\infty}) = \begin{cases} 1 & \text{if } X_2 = j \\ 1 + \tau_{\to j}(X_{3:\infty}) & \text{if } X_2 \neq j \end{cases} = 1 + \mathbb{I}_{\neg j}(X_2)\tau_{\to j}(X_{3:\infty}), \tag{29}$$

where we let $\mathbb{I}_{\neg j} := 1 - \mathbb{I}_j$ and we adopt the convention that $0 \cdot +\infty = 0$. If we now consider any $x_{1:2} \in \mathscr{X}^2$, then it follows from Equation (29) that

[5] The reason why this function is extended real-valued is because $\tau_{\to j}(\omega) = +\infty$ if $\omega_m \neq j$ for all $m \in \mathbb{N}$.

$$\begin{aligned}\underline{E}(\tau_{\to j}(X_{2:\infty})|x_{1:2}) &= \underline{E}(1+\mathbb{I}_{\neg j}(X_2)\tau_{\to j}(X_{3:\infty})|x_{1:2})\\ &= 1+\underline{E}(\mathbb{I}_{\neg j}(X_2)\tau_{\to j}(X_{3:\infty})|x_{1:2}) = 1+\underline{E}(\mathbb{I}_{\neg j}(x_2)\tau_{\to j}(X_{3:\infty})|x_{1:2}),\end{aligned}$$

where the second equality is a consequence of **C6** and the last equality follows from Proposition 1. Furthermore, since **C5** implies that $\underline{E}(0|x_{1:2}) = 0$, it follows from Equation (26)—with $n = 2$—that

$$\begin{aligned}&\underline{E}(\mathbb{I}_{\neg j}(x_2)\tau_{\to j}(X_{3:\infty})|x_{1:2})\\ &= \begin{cases}\underline{E}(0|x_{1:2}) & \text{if } x_2 = j\\ \underline{E}(\tau_{\to j}(X_{3:\infty})|x_{1:2}) & \text{if } x_2 \neq j\end{cases} = \begin{cases}0 & \text{if } x_2 = j\\ \underline{\tau}_{x_2\to j} & \text{if } x_2 \neq j\end{cases} = \mathbb{I}_{\neg j}(x_2)\underline{\tau}_{x_2\to j}.\end{aligned}$$

Hence, we find that $\underline{E}(\tau_{\to j}(X_{2:\infty})|x_{1:2}) = 1+\mathbb{I}_{\neg j}(x_2)\underline{\tau}_{x_2\to j}$. Since $x_{1:2}\in\mathscr{X}^2$ was arbitrary, this implies that

$$\underline{E}(\tau_{\to j}(X_{2:\infty})|X_{1:2}) = 1+\mathbb{I}_{\neg j}(X_2)\underline{\tau}_{X_2\to j} = 1+\mathbb{I}_{\neg j}(X_2)\underline{\tau}_{\bullet\to j}(X_2),$$

with $\underline{\tau}_{\bullet\to j}\in\mathscr{G}(\mathscr{X})$ defined by $\underline{\tau}_{\bullet\to j}(x) := \underline{\tau}_{x\to j}$ for all $x\in\mathscr{X}$; the fact that $\underline{\tau}_{\bullet\to j}$ is indeed an element of $\mathscr{G}(\mathscr{X})$—is real-valued rather than extended real-valued—is a consequence of Theorem 4. In combination with Equation (28), this implies that

$$\underline{\tau}_{i\to j} = \underline{E}(1+\mathbb{I}_{\neg j}(X_2)\underline{\tau}_{\bullet\to j}(X_2)|i) = \underline{Q}_i(1+\mathbb{I}_{\neg j}\underline{\tau}_{\bullet\to j}) = 1+\underline{Q}_i(\mathbb{I}_{\neg j}\underline{\tau}_{\bullet\to j}),$$

where we use Equation (24) to establish the second equality, and where the last equality follows from C6. Because of Equations (11)–(13), we now finally obtain the following system of non-linear equations: for all $j\in\mathscr{X}$, we have that

$$\underline{\tau}_{0\to j} = 1+\min_{\pi_0\in\mathscr{Q}_0}\left\{r_0\mathbb{I}_{\neg j}(0)\underline{\tau}_{0\to j}+p_0\mathbb{I}_{\neg j}(1)\underline{\tau}_{1\to j}\right\} \tag{30}$$

and

$$\underline{\tau}_{L\to j} = 1+\min_{\pi_L\in\mathscr{Q}_L}\left\{q_L\mathbb{I}_{\neg j}(L-1)\underline{\tau}_{L-1\to j}+r_L\mathbb{I}_{\neg j}(L)\underline{\tau}_{L\to j}\right\} \tag{31}$$

and, for all $i\in\mathscr{X}/\{0,L\}$, we have that

$$\underline{\tau}_{i\to j} = 1+\min_{\pi_i\in\mathscr{Q}_i}\left\{q_i\mathbb{I}_{\neg j}(i-1)\underline{\tau}_{i-1\to j}+r_i\mathbb{I}_{\neg j}(i)\underline{\tau}_{i\to j}+p_i\mathbb{I}_{\neg j}(i+1)\underline{\tau}_{i+1\to j}\right\}. \tag{32}$$

Using a completely analogous derivation, we also find that

$$\overline{\tau}_{i\to j} = 1+\overline{Q}_i(\mathbb{I}_{\neg j}\overline{\tau}_{\bullet\to j}), \tag{33}$$

which gives rise to a similar system of non-linear equations.

The main goal of the rest of this paper is to solve these systems of equations in order to compute $\underline{\tau}_{i\to j}$ and $\overline{\tau}_{i\to j}$ for all $i,j\in\mathscr{X}$. However, since the equations in these systems are non-linear, it is not feasible to solve them directly. Fortunately, as we will show in the following three sections, it is possible to transform them into simple recursive expressions, which can then be used to compute $\underline{\tau}_{i\to j}$ and $\overline{\tau}_{i\to j}$ for all $i,j\in\mathscr{X}$.

## 7. Lower and upper expected upward first-passage times

We start by computing lower expectations of upward first-passage times, that is, for all $i,j\in\mathscr{X}$ such that $i<j$, we will compute $\underline{\tau}_{i\to j}$. We initially focus on computing $\underline{\tau}_{i\to i+1}$, for $i\in\mathscr{X}\setminus\{L\}$, and then show that any lower expected upward first-passage time can be obtained as a sum of such terms. Similar results are obtained for upper expected upward first-passage times.

Finding $\underline{\tau}_{0\to 1}$ is easy. It follows from Equation (30), with $j=1$, that

$$\begin{aligned}\underline{\tau}_{0\to 1} &= 1+\min_{\pi_0\in\mathscr{Q}_0} r_0\underline{\tau}_{0\to 1} = 1+\min_{\pi_0\in\mathscr{Q}_0}(1-p_0)\underline{\tau}_{0\to 1}\\ &= 1+\underline{\tau}_{0\to 1}-\max_{\pi_0\in\mathscr{Q}_0} p_0\underline{\tau}_{0\to 1} = 1+\underline{\tau}_{0\to 1}-\overline{p}_0\underline{\tau}_{0\to 1},\end{aligned}$$

where the second equality holds because $\pi_0$ is a probability mass function on a binary set and the last equality holds because we know from Theorem 4 that $\underline{\tau}_{0\to 1}$ is real-valued and therefore finite. Hence, since we know from Theorem 4 that $\underline{\tau}_{0\to 1}$ is strictly positive and real-valued, it follows that

$$\underline{\tau}_{0\to 1} = \frac{1}{\overline{p}_0}. \tag{34}$$

Finding $\underline{\tau}_{0\to j}$, for $j \in \{2, \dots, L\}$, is more involved. We start by establishing a connection with $\underline{\tau}_{1\to j}$. By applying Equation (30), we find that

$$\begin{aligned}\underline{\tau}_{0\to j} = 1 + \min_{\pi_0 \in \mathscr{Q}_0} \{r_0 \underline{\tau}_{0\to j} + p_0 \underline{\tau}_{1\to j}\} &= 1 + \min_{\pi_0 \in \mathscr{Q}_0} \{(1-p_0)\underline{\tau}_{0\to j} + p_0 \underline{\tau}_{1\to j}\} \\ &= 1 + \underline{\tau}_{0\to j} + \min_{\pi_0 \in \mathscr{Q}_0} p_0(\underline{\tau}_{1\to j} - \underline{\tau}_{0\to j}),\end{aligned}$$

which implies, due to Theorem 4, that

$$\min_{\pi_0 \in \mathscr{Q}_0} p_0(\underline{\tau}_{1\to j} - \underline{\tau}_{0\to j}) = -1. \tag{35}$$

Since the minimum in Equation (35) is negative and $p_0$ is a probability and therefore non-negative, it must be that $\underline{\tau}_{1\to j} - \underline{\tau}_{0\to j} < 0$. Therefore, Equation (35) is minimised for $p_0 = \overline{p}_0$ and we find that

$$\underline{\tau}_{0\to j} = \frac{1}{\overline{p}_0} + \underline{\tau}_{1\to j}. \tag{36}$$

By combining Equations (34) and (36), we see that

$$\underline{\tau}_{0\to j} = \underline{\tau}_{0\to 1} + \underline{\tau}_{1\to j} \text{ for all } j \in \{2, \dots, L\}. \tag{37}$$

Since we already know $\underline{\tau}_{0\to 1}$—see Equation (34)—we are now left to find $\underline{\tau}_{1\to j}$.

We first consider the case $j = 2$. In that case, it follows from Equation (32), with $i = 1$ and $j = 2$, that

$$\begin{aligned}\underline{\tau}_{1\to 2} = 1 + \min_{\pi_1 \in \mathscr{Q}_1} \{q_1 \underline{\tau}_{0\to 2} + r_1 \underline{\tau}_{1\to 2}\} &= 1 + \min_{\pi_1 \in \mathscr{Q}_1} \{q_1 \underline{\tau}_{0\to 2} + (1 - q_1 - p_1)\underline{\tau}_{1\to 2}\} \\ &= 1 + \underline{\tau}_{1\to 2} + \min_{\pi_1 \in \mathscr{Q}_1} \{q_1(\underline{\tau}_{0\to 2} - \underline{\tau}_{1\to 2}) - p_1 \underline{\tau}_{1\to 2}\},\end{aligned}$$

which implies, due to Theorem 4, that

$$\min_{\pi_1 \in \mathscr{Q}_1} \{q_1(\underline{\tau}_{0\to 2} - \underline{\tau}_{1\to 2}) - p_1 \underline{\tau}_{1\to 2}\} = -1.$$

By applying Equation (37) for $j = 2$ we then find that

$$\min_{\pi_1 \in \mathscr{Q}_1} \{q_1 \underline{\tau}_{0\to 1} - p_1 \underline{\tau}_{1\to 2}\} = -1. \tag{38}$$

Therefore, and because we already know $\underline{\tau}_{0\to 1}$, it follows from Assumption 1 and the following lemma that $\underline{\tau}_{1\to 2}$ is the unique solution to Equation (38).

**Proposition 5.** *Consider a credal set $\mathscr{Q}$ on $\mathscr{X}_m$ that consists of strictly positive probability mass functions and let $c$ be a real constant. Then*

$$\min_{\pi \in \mathscr{Q}} \{qc - p\mu\}$$

*is a strictly decreasing function of $\mu$.*

This unique solution $\underline{\tau}_{1\to 2}$ is furthermore easy to compute. It follows from Proposition 5 that a simple bisection method suffices.

Next, we consider the case $j \in \{3, \dots, L\}$. By applying Equation (32), for such a $j$ and with $i = 1$, we find that

$$\begin{aligned}\underline{\tau}_{1\to j} &= 1 + \min_{\pi_1 \in \mathscr{Q}_1} \{q_1 \underline{\tau}_{0\to j} + r_1 \underline{\tau}_{1\to j} + p_1 \underline{\tau}_{2\to j}\} \\ &= 1 + \min_{\pi_1 \in \mathscr{Q}_1} \{q_1 \underline{\tau}_{0\to j} + (1 - q_1 - p_1)\underline{\tau}_{1\to j} + p_1 \underline{\tau}_{2\to j}\} \\ &= 1 + \underline{\tau}_{1\to j} + \min_{\pi_1 \in \mathscr{Q}_1} \{q_1(\underline{\tau}_{0\to j} - \underline{\tau}_{1\to j}) + p_1(\underline{\tau}_{2\to j} - \underline{\tau}_{1\to j})\},\end{aligned}$$

which implies, due to Theorem 4, that

$$\min_{\pi_1 \in \mathscr{Q}_1} \{q_1(\underline{\tau}_{0\to j} - \underline{\tau}_{1\to j}) + p_1(\underline{\tau}_{2\to j} - \underline{\tau}_{1\to j})\} = -1.$$

In combination with Equation (37), this results in

$$\min_{\pi_1 \in \mathscr{Q}_1} \{q_1 \underline{\tau}_{0\to 1} + p_1(\underline{\tau}_{2\to j} - \underline{\tau}_{1\to j})\} = -1. \tag{39}$$

Since we know from Assumption 1 and Proposition 5 that the equation

$$\min_{\pi_1 \in \mathscr{Q}_1} \{q_1 \underline{\tau}_{0\to 1} + p_1 \mu\} = -1$$

has a unique solution $\mu$, it follows directly from Equations (38) and (39) that

$$\underline{\tau}_{1\to j} = \underline{\tau}_{1\to 2} + \underline{\tau}_{2\to j} \text{ for all } j \in \{3, \dots, L\}. \tag{40}$$

At this point, we already know how to compute $\underline{\tau}_{0\to 1}$ and $\underline{\tau}_{1\to 2}$ and we have also established the following additivity property:

$$\underline{\tau}_{i\to j} = \underline{\tau}_{i\to i+1} + \underline{\tau}_{i+1\to j}$$

for all $i \in \{0, 1\}$ and $j \in \{i+2, \dots, L\}$. By continuing in this way, we obtain the following two results; see Appendix A for a proof.

**Proposition 6.** *For any $i \in \mathscr{X} \setminus \{0, L\}$, we have that*

$$\min_{\pi_i \in \mathscr{Q}_i} \{q_i \underline{\tau}_{i-1\to i} - p_i \underline{\tau}_{i\to i+1}\} = -1. \tag{41}$$

**Proposition 7.** *For all $i, j \in \mathscr{X}$ such that $i + 1 < j$, we have that*

$$\underline{\tau}_{i\to j} = \underline{\tau}_{i\to i+1} + \underline{\tau}_{i+1\to j}.$$

For all $i \in \mathscr{X} \setminus \{L\}$, the value of $\underline{\tau}_{i\to i+1}$ can therefore be computed recursively. For $i = 0$, we simply apply Equation (34). For any other $i \in \mathscr{X} \setminus \{0, L\}$, it follows from Assumption 1 and Propositions 5 and 6 that $\underline{\tau}_{i\to i+1}$ is the unique solution to Equation (41), which can be obtained by means of a bisection method. In this equation, the value of $\underline{\tau}_{i-1\to i}$ has already been computed earlier on in this recursive procedure.

The following additivity result is a direct consequence of Proposition 7.

**Corollary 8.** *For all $i, j \in \mathscr{X}$ such that $i < j$, we have that*

$$\underline{\tau}_{i\to j} = \sum_{k=i}^{j-1} \underline{\tau}_{k\to k+1}.$$

It implies that the recursive techniques that we developed in this section can be used to compute any lower expected upward first-passage time.

Similar results can be proved for upper expectations of upward first-passage times. We only provide the final expressions; the derivation is completely analogous. In this case, the starting point is that

$$\overline{\tau}_{0\to 1} = \frac{1}{\underline{p}_0}. \tag{42}$$

For all $i \in \mathscr{X} \setminus \{0, L\}$, the value of $\overline{\tau}_{i\to i+1}$ can then be computed recursively, due to Assumption 1 and the following two results.

**Proposition 9.** *For all $i \in \mathscr{X} \setminus \{0, L\}$, we have that*

$$\max_{\pi_i \in \mathscr{Q}_i} \{q_i \overline{\tau}_{i-1\to i} - p_i \overline{\tau}_{i\to i+1}\} = -1. \tag{43}$$

**Proposition 10.** *Consider a credal set $\mathscr{Q}$ on $\mathscr{X}_m$ that consists of strictly positive probability mass functions and let $c$ be a real constant. Then*

$$\max_{\pi \in \mathscr{Q}} \{qc - p\mu\}$$

*is a strictly decreasing function of $\mu$.*

Due to our next result, this recursive technique allows us to compute arbitrary upper expected upward first-passage times.

**Proposition 11.** *For all $i, j \in \mathscr{X}$ such that $i < j$, we have that*

$$\overline{\tau}_{i \to j} = \sum_{k=i}^{j-1} \overline{\tau}_{k \to k+1}.$$

## 8. Lower and upper expected downward first-passage times

Lower and upper expectations of downward first-passage times can be computed in more or less the same way. The main difference is that the recursive expressions now start from the other side, that is, from $i = L$.[6] We find that

$$\underline{\tau}_{L \to L-1} = \frac{1}{\overline{q}_L} \text{ and } \overline{\tau}_{L \to L-1} = \frac{1}{\underline{q}_L}. \tag{44}$$

For all $i \in \mathscr{X} \setminus \{0, L\}$, due to Assumption 1, the values of $\underline{\tau}_{i \to i-1}$ and $\overline{\tau}_{i \to i-1}$ can now be computed recursively, using the following two results.

**Proposition 12.** *For all $i \in \mathscr{X} \setminus \{0, L\}$, we have that*

$$\min_{\pi_i \in \mathscr{Q}_i} \{-q_i \underline{\tau}_{i \to i-1} + p_i \underline{\tau}_{i+1 \to i}\} = -1 \text{ and } \max_{\pi_i \in \mathscr{Q}_i} \{-q_i \overline{\tau}_{i \to i-1} + p_i \overline{\tau}_{i+1 \to i}\} = -1.$$

**Proposition 13.** *Consider a credal set $\mathscr{Q}$ on $\mathscr{X}_m$ that consists of strictly positive probability mass functions and let $c$ be a real constant. Then*

$$\min_{\pi \in \mathscr{Q}} \{-q\mu + pc\} \text{ and } \max_{\pi \in \mathscr{Q}} \{-q\mu + pc\}$$

*are strictly decreasing functions of $\mu$.*

Once we have computed $\underline{\tau}_{i \to i-1}$ and $\overline{\tau}_{i \to i-1}$ for all $i \in \mathscr{X} \setminus \{L\}$, the following result enables us to easily obtain all other lower and upper expected downward first-passage times.

**Proposition 14.** *For all $i, j \in \mathscr{X}$ such that $i > j$, we have that*

$$\underline{\tau}_{i \to j} = \sum_{k=j}^{i-1} \underline{\tau}_{k+1 \to k} \text{ and } \overline{\tau}_{i \to j} = \sum_{k=j}^{i-1} \overline{\tau}_{k+1 \to k}.$$

## 9. Lower and upper expected return times

Given the results in the previous two sections, lower and upper expected return times can now be computed very easily. By applying Equations (30)–(32), with $j$ equal to 0, $L$ and $i$, respectively, we find that

$$\underline{\tau}_{0 \to 0} = 1 + \min_{\pi_0 \in \mathscr{Q}_0} p_0 \underline{\tau}_{1 \to 0} = 1 + \underline{p}_0 \underline{\tau}_{1 \to 0} \tag{45}$$

and

$$\underline{\tau}_{L \to L} = 1 + \min_{\pi_L \in \mathscr{Q}_L} q_L \underline{\tau}_{L-1 \to L} = 1 + \underline{q}_L \underline{\tau}_{L-1 \to L} \tag{46}$$

and, for all $i \in \mathscr{X} \setminus \{0, L\}$, that

$$\underline{\tau}_{i \to i} = 1 + \min_{\pi_i \in \mathscr{Q}_i} \{q_i \underline{\tau}_{i-1 \to i} + p_i \underline{\tau}_{i+1 \to i}\}. \tag{47}$$

In these expressions, the lower expected first-passage times $\underline{\tau}_{1 \to 0}$, $\underline{\tau}_{L-1 \to L}$, $\underline{\tau}_{i-1 \to i}$ and $\underline{\tau}_{i+1 \to i}$ can be computed using the recursive techniques that we developed in the previous two sections. Similarly, for the upper case, we find that

[6] Our presentation of—and proofs for—the results in this section are adapted versions of the ones in Section 7. An alternative method would be to observe that a downward first-passage time from $i$ to $j$ is the same as an upward first-passage time from $L - i$ to $L - j$ in a new imprecise birth–death chain, obtained by reversing the order of the states, and by switching the role of $p$ and $q$ accordingly.

$$\overline{\tau}_{0\to 0} = 1 + \max_{\pi_0\in\mathscr{Q}_0} p_0\overline{\tau}_{1\to 0} = 1 + \overline{p}_0\overline{\tau}_{1\to 0} \tag{48}$$

and

$$\overline{\tau}_{L\to L} = 1 + \max_{\pi_L\in\mathscr{Q}_L} q_L\overline{\tau}_{L-1\to L} = 1 + \overline{q}_L\overline{\tau}_{L-1\to L} \tag{49}$$

and, for all $i \in \mathscr{X} \setminus \{0, L\}$, that

$$\overline{\tau}_{i\to i} = 1 + \max_{\pi_i\in\mathscr{Q}_i} \{q_i\overline{\tau}_{i-1\to i} + p_i\overline{\tau}_{i+1\to i}\}. \tag{50}$$

Again, the upper expected first-passage times $\overline{\tau}_{1\to 0}$, $\overline{\tau}_{L-1\to L}$, $\overline{\tau}_{i-1\to i}$ and $\overline{\tau}_{i+1\to i}$ that appear in these expressions can be computed with the recursive techniques that were introduced above.

## 10. The precise case and its connection with the imprecise one

Since birth–death chains are a special case of imprecise birth–death chains, our results for imprecise birth–death chains can also be applied to birth–death chains. As we will now show, in that case, lower and upper expected first-passage and return times coincide, and our recursive methods then lead to closed-form expressions for them. Furthermore, although an imprecise birth–death chain is not the same as a set of birth death chains—see Section 4.2—we will see that for the purposes of computing lower and upper expected first-passage and return times, it so happens that this makes no difference.

### *10.1. Expected first-passage and return times in precise birth–deaths chains*

Clearly, a birth–death chain can be regarded as a special type of imprecise birth–death chain. It corresponds to the case where all the local credal sets are singletons, that is, $\Phi_\square = \{\varphi_\square\}$ and, for all $i \in \mathscr{X}$, $\mathscr{Q}_i = \{\pi_i\}$. We refer to this special type of imprecise birth–death chain as a precise birth–death chain. For these precise birth–death chains, as the following result implies, lower and upper expected first-passage and return times coincide.

**Proposition 15.** *Consider any imprecise birth–death chain such that, for all $i \in \mathscr{X}$, $\mathscr{Q}_i = \{\pi_i\}$. Then*

$$\underline{\tau}_{i\to j} = \overline{\tau}_{i\to j} \text{ for all } i, j \in \mathscr{X}.$$

Notice that this result does not require that the initial credal set $\Phi_\square$ should be a singleton. This is not surprising: since none of the methods in Sections 7–9 require the use of $\Phi_\square$, it follows that $\Phi_\square$ does not have any effect on first-passage or return times. Therefore, for our present purposes, all the relevant parameters of a birth–death chain can be represented by a single stochastic matrix $P$, the form of which is given by Equation (8). Due to Proposition 15, with any such matrix $P$, we can associate a unique expected first-passage time from $i \in \mathscr{X}$ to $j \in \mathscr{X}$, defined by

$$\tau^P_{i\to j} := \underline{\tau}_{i\to j} = \overline{\tau}_{i\to j} \quad \text{for all } i, j \in \mathscr{X}. \tag{51}$$

If $i = j$, then $\tau^P_{i\to i}$ is called an expected return time. As it turns out, we can derive closed-form expressions for these expected first-passage and return times. The following lemma presents such an expression for a specific type of expected upward first-passage times.

**Proposition 16.** *Consider a precise birth–death chain of which the stochastic matrix $P$ is given by Equation (8). Then for all $i \in \mathscr{X} \setminus \{L\}$, we have that*

$$\tau^P_{i\to i+1} = \sum_{k=0}^{i} \frac{\prod_{\ell=k+1}^{i} q_\ell}{\prod_{m=k}^{i} p_m}. \tag{52}$$

Based on this result, it is now easy to obtain expressions for all the other expected upward first-passage times, because it follows from Corollary 8 and Equation (51) that

$$\tau^P_{i\to j} = \sum_{k=i}^{j-1} \tau^P_{k\to k+1} \text{ for all } i, j \in \mathscr{X} \text{ such that } i < j. \tag{53}$$

Similarly, for expected downward first-passage times, it follows from Proposition 14 and Equation (51) that

$$\tau^P_{i\to j} = \sum_{k=j}^{i-1} \tau^P_{k+1\to k} \text{ for all } i, j \in \mathscr{X} \text{ such that } i > j, \tag{54}$$

where the individual terms in the summation are given by Proposition 17.

**Proposition 17.** *Consider a precise birth–death chain of which the stochastic matrix $P$ is given by Equation (8). Then for all $i \in \mathscr{X} \setminus \{0\}$, we have that*

$$\tau^P_{i\to i-1} = \sum_{k=i}^{L} \frac{\prod_{\ell=i}^{k-1} p_\ell}{\prod_{m=i}^{k} q_m}.$$

Closed-form expressions for expected return times can now be derived from Equations (45)–(50), which, for precise birth–death chains, reduce to the following simple expressions:

$$\tau^P_{0\to 0} = 1 + p_0 \tau^P_{1\to 0} \text{ and } \tau^P_{L\to L} = 1 + q_L \tau^P_{L-1\to L} \tag{55}$$

and

$$\tau^P_{i\to i} = 1 + q_i \tau^P_{i-1\to i} + p_i \tau^P_{i+1\to i} \quad \text{for all } i \in \mathscr{X} \setminus \{0, L\}. \tag{56}$$

We conclude from this section that the expected first-passage and return times of a precise birth–death chain are easy to compute. However, this is not surprising. In fact, the closed-form expressions that we have obtained are—although we did not find a reference that states them explicitly—essentially well known from the traditional theory of precise birth–death chains. The only difference is that the traditional theory of precise birth–death chains adopts a measure-theoretic definition for its global expectation operator, whereas we define it by means of the game-theoretic framework that was discussed in Section 5. The main result of this section is therefore that for the purpose of computing expected first-passage and return times in precise birth–death chains, these two definitions are identical.

### *10.2. Connecting imprecise birth–death chains with precise ones*

As explained in Section 4.2, an imprecise birth–death chain is not just a set of precise birth–death chains. Of course, clearly, any precise birth–death chain for which the initial model $\varphi_\square$ belongs to $\Phi_\square$ and each of the transition probabilities $\varphi_i$, $i \in \mathscr{X}$, belongs to $\Phi_i$ is an element of the imprecise birth–death chain $\mathscr{T}^\Phi$. However, the imprecise birth–death chain $\mathscr{T}^\Phi$ also contains other probability trees, which do not correspond to a (time-homogeneous and Markovian) precise birth–death chain.

However, for the purposes of computing lower and upper expected first-passage or return times, these extra probability trees are not essential, because as we will now show, for first-passage and return times, the lower and upper expectations are achieved by (time-homogeneous and Markovian) precise birth–death chains in $\mathscr{T}^\Phi$.

Since we already know from Section 10.1 that the initial model $\varphi_\square \in \Phi_\square$ of these precise birth–death chains does not influence their expected first-passage and return times, we can conveniently represent them by means of their stochastic matrix $P$. Depending on the type of bound that we are considering, a different type of stochastic matrix $P$ will be needed to achieve the bound. We will specify the essential features of these different types by means of selection methods. For any given imprecise birth–death chain, such a selection method describes a specific way of choosing a stochastic matrix $P$.

For lower expected upward first-passage times, we use the following selection method.

**Selection Method LU$_k$**
Let $P$ be any stochastic matrix of the form in Equation (8) and such that

(1) if $k \neq 0$, then $p_0 = \overline{p}_0$;
(2) for all $\ell \in \{1, \dots, k-1\}$, $(q_\ell, r_\ell, p_\ell) \in \arg\min_{\pi_\ell \in \mathscr{Q}_\ell} \{q_\ell \underline{\tau}_{\ell-1\to\ell} - p_\ell \underline{\tau}_{\ell\to\ell+1}\}$.

Indeed, as our next result establishes, for any given imprecise birth–death chain, its lower expected upward first-passage time can be obtained by a precise birth–death chain whose stochastic matrix $P$ is selected according to the method above.

**Theorem 18.** *Consider an imprecise birth–death chain, some $k \in \mathscr{X}$, and a precise birth death chain whose stochastic matrix $P$ is obtained from this imprecise birth–death chain by means of Selection Method LU$_k$. Then for all $i, j \in \mathscr{X}$ such that $i < j \leq k$, $\underline{\tau}_{i\to j} = \tau^P_{i\to j}$.*

This result is at its most powerful if we choose $k = L$. In that case, it follows that all the lower expected upward first-passage times $\underline{\tau}_{i\to j}$, with $i, j \in \mathscr{X}$ such that $i < j$, can be obtained by the same precise birth–death chain.

Similar results also hold for upper expected upward first-passage times.

**Selection Method UU$_k$**
Let $P$ be any stochastic matrix of the form in Equation (8) and such that

(1) if $k \neq 0$, then $p_0 = \underline{p}_0$;
(2) for all $\ell \in \{1, \dots, k-1\}$, $(q_\ell, r_\ell, p_\ell) \in \arg\max_{\pi_\ell \in \mathcal{Q}_\ell} \{q_\ell \overline{\tau}_{\ell-1\to\ell} - p_\ell \overline{\tau}_{\ell\to\ell+1}\}$.

**Theorem 19.** *Consider an imprecise birth–death chain, some $k \in \mathscr{X}$, and a precise birth death chain whose stochastic matrix $P$ is obtained from this imprecise birth–death chain by means of Selection Method UU$_k$. Then for all $i, j \in \mathscr{X}$ such that $i < j \leq k$, $\overline{\tau}_{i\to j} = \tau^P_{i\to j}$.*

As before, this result is most powerful if we choose $k = L$, because it then implies that every upper expected upward first-passage time can be obtained by the same precise birth–death chain.

At first sight, it seems as though Theorems 18 and 19 could provide us with a simple method for computing lower and upper expected upward first-passage times, thereby providing an alternative for the recursive equations in Section 7. All we have to do is (a) construct a stochastic matrix $P$ according to an appropriate selection method and then (b) use this matrix $P$ to apply the closed-form expressions in Section 10.1. However, this method is not practical. The issue here is step (a). For example, executing Selection Method LU$_k$ (2) is not just a matter of choosing $p_\ell = \overline{p}_\ell$ and $q_\ell = \underline{q}_\ell$, because, depending on the shape of $\mathcal{Q}_\ell$, it may not be possible to attain these extrema simultaneously. Therefore, finding the optimal tuples $(p_\ell, r_\ell, q_\ell)$ requires us to know the value of $\underline{\tau}_{\ell\to\ell+1}$ for all $\ell \in \{0, \dots, k-1\}$. However, in practice, we don't know these values yet. In fact, the whole point of computing lower expected upward first-passage times is to obtain these values. Therefore, Theorems 18 and 19 should not be regarded as the basis of a computational method. Instead, their main importance is the theoretical insight that the lower and upper expected upward first-passage times that correspond to an imprecise birth death chain are achieved by (time-homogeneous and Markovian) precise birth–death chains in $\mathscr{T}^\Phi$.

Completely analogous conclusions can be drawn for lower and upper downward first-passage times, using the following selection methods and results.

**Selection Method LD$_k$**
Let $P$ be any stochastic matrix of the form in Equation (8) and such that

(1) if $k \neq L$, then $q_L = \overline{q}_L$;
(2) for all $\ell \in \{k+1, \dots, L-1\}$, $(q_\ell, r_\ell, p_\ell) \in \arg\min_{\pi_\ell \in \mathcal{Q}_\ell} \{-q_\ell \underline{\tau}_{\ell\to\ell-1} + p_\ell \underline{\tau}_{\ell+1\to\ell}\}$.

**Theorem 20.** *Consider an imprecise birth–death chain, some $k \in \mathscr{X}$, and a precise birth death chain whose stochastic matrix $P$ is obtained from this imprecise birth–death chain by means of Selection Method LD$_k$. Then for all $i, j \in \mathscr{X}$ such that $k \leq j < i$, $\underline{\tau}_{i\to j} = \tau^P_{i\to j}$.*

**Selection Method UD$_k$**
Let $P$ be any stochastic matrix of the form in Equation (8) and such that

(1) if $k \neq L$, then $q_L = \underline{q}_L$;
(2) for all $\ell \in \{k+1, \dots, L-1\}$, $(q_\ell, r_\ell, p_\ell) \in \arg\max_{\pi_\ell \in \mathcal{Q}_\ell} \{-q_\ell \overline{\tau}_{\ell\to\ell-1} + p_\ell \overline{\tau}_{\ell+1\to\ell}\}$.

**Theorem 21.** *Consider an imprecise birth–death chain, some $k \in \mathscr{X}$, and a precise birth death chain whose stochastic matrix $P$ is obtained from this imprecise birth–death chain by means of Selection Method UD$_k$. Then for all $i, j \in \mathscr{X}$ such that $k \leq j < i$, $\overline{\tau}_{i\to j} = \tau^P_{i\to j}$.*

These results are at their most powerful if we choose $k = 0$. In that case, all the lower expected upward first-passage times $\underline{\tau}_{i\to j}$, with $i, j \in \mathscr{X}$ such that $j < i$, can be obtained by the same precise birth–death chain.

This is not the case for lower and upper expected return times: there may not be a single precise birth–death chain for which all lower expected return times are obtained, nor is there guaranteed to be a single precise birth–death chain for which all the upper expected return times are obtained. Nevertheless, as we show in Theorems 22 and 23 below, it is always possible to select one precise birth–death chain for each specific lower expected return time and one for each specific upper expected return time, using the following selection methods.

**Selection Method LR$_k$**
Let $P$ be any stochastic matrix of the form in Equation (8) and such that

(1) if $k \neq 0$, then $p_0 = \overline{p}_0$, and if $k = 0$, then $p_0 = \underline{p}_0$;
(2) for all $\ell \in \{1, \dots, k-1\}$, $(q_\ell, r_\ell, p_\ell) \in \arg\min_{\pi_\ell \in \mathcal{Q}_\ell} \{q_\ell \underline{\tau}_{\ell-1\to\ell} - p_\ell \underline{\tau}_{\ell\to\ell+1}\}$.
(3) if $k \neq 0$ and $k \neq L$, then $(q_k, r_k, p_k) \in \arg\min_{\pi_k \in \mathcal{Q}_k} \{q_k \underline{\tau}_{k-1\to k} + p_k \underline{\tau}_{k+1\to k}\}$;
(4) for all $\ell \in \{k+1, \dots, L-1\}$, $(q_\ell, r_\ell, p_\ell) \in \arg\min_{\pi_\ell \in \mathcal{Q}_\ell} \{-q_\ell \underline{\tau}_{\ell\to\ell-1} + p_\ell \underline{\tau}_{\ell+1\to\ell}\}$.
(5) if $k \neq L$, then $q_L = \overline{q}_L$, and if $k = L$, then $q_L = \underline{q}_L$.

**Theorem 22.** *Consider an imprecise birth–death chain, some $k \in \mathscr{X}$, and a precise birth death chain whose stochastic matrix $P$ is obtained from this imprecise birth–death chain by means of Selection Method LR$_k$. Then $\underline{\tau}_{k\to k} = \tau^P_{k\to k}$.*

**Selection Method UR$_k$**
Let $P$ be any stochastic matrix of the form in Equation (8) and such that

(1) if $k \neq 0$, then $p_0 = \underline{p}_0$, and if $k = 0$, then $p_0 = \overline{p}_0$;
(2) for all $\ell \in \{1, \dots, k-1\}$, $(q_\ell, r_\ell, p_\ell) \in \arg\max_{\pi_\ell \in \mathcal{Q}_\ell} \{q_\ell \overline{\tau}_{\ell-1\to\ell} - p_\ell \overline{\tau}_{\ell\to\ell+1}\}$.
(3) if $k \neq 0$ and $k \neq L$, then $(q_k, r_k, p_k) \in \arg\max_{\pi_k \in \mathcal{Q}_k} \{q_k \overline{\tau}_{k-1\to k} + p_k \overline{\tau}_{k+1\to k}\}$;
(4) for all $\ell \in \{k+1, \dots, L-1\}$, $(q_\ell, r_\ell, p_\ell) \in \arg\max_{\pi_\ell \in \mathcal{Q}_\ell} \{-q_\ell \overline{\tau}_{\ell\to\ell-1} + p_\ell \overline{\tau}_{\ell+1\to\ell}\}$.
(5) if $k \neq L$, then $q_L = \underline{q}_L$, and if $k = L$, then $q_L = \overline{q}_L$.

**Theorem 23.** *Consider an imprecise birth–death chain, some $k \in \mathscr{X}$, and a precise birth death chain whose stochastic matrix $P$ is obtained from this imprecise birth–death chain by means of Selection Method UR$_k$. Then $\overline{\tau}_{k\to k} = \tau^P_{k\to k}$.*

## 11. Linear-vacuous mixtures

Precise birth–death chains are not the only special case for which it is possible to obtain closed-form expressions. As we are about to show, such expressions can also be obtained for the special case that all the local models are linear-vacuous mixtures.

In order to define this type of model, we start from given strictly positive probability mass functions $\pi^*_0 = (r^*_0, p^*_0) \in \Sigma_{\mathscr{X}_0}$, $\pi^*_L = (q^*_L, r^*_L) \in \Sigma_{\mathscr{X}_L}$ and, for all $i \in \mathscr{X} \setminus \{0, L\}$, $\pi^*_i = (q^*_i, r^*_i, p^*_i) \in \Sigma_{\mathscr{X}_m}$. Furthermore, for all $i \in \mathscr{X}$, we consider some real-valued $\varepsilon_i \in [0, 1)$. We use these parameters to define linear-vacuous local credal sets. Similarly to Equation (2) in Section 2, our local credal sets are defined as follows:

$$\mathcal{Q}_0 = \mathcal{Q}^{\varepsilon_0}_{\pi^*_0} := \{(1-\varepsilon_0)\pi^*_0 + \varepsilon_0 \pi'_0 \colon \pi'_0 \in \Sigma_{\mathscr{X}_0}\}$$

and

$$\mathcal{Q}_L = \mathcal{Q}^{\varepsilon_L}_{\pi^*_L} := \{(1-\varepsilon_L)\pi^*_L + \varepsilon_L \pi'_L \colon \pi'_L \in \Sigma_{\mathscr{X}_L}\}$$

and, for all $i \in \mathscr{X} \setminus \{0, L\}$,

$$\mathcal{Q}_i = \mathcal{Q}^{\varepsilon_i}_{\pi^*_i} := \{(1-\varepsilon_i)\pi^*_i + \varepsilon_i \pi'_i \colon \pi'_i \in \Sigma_{\mathscr{X}_m}\}.$$

Furthermore, due to (9), for all $i \in \mathscr{X} \setminus \{0\}$, we have that

$$\underline{q}_i := (1-\varepsilon_i) q^*_i \text{ and } \overline{q}_i := (1-\varepsilon_i) q^*_i + \varepsilon_i$$

and, due to (10), for all $i \in \mathscr{X} \setminus \{L\}$, we have that

$$\underline{p}_i := (1-\varepsilon_i) p^*_i \text{ and } \overline{p}_i := (1-\varepsilon_i) p^*_i + \varepsilon_i.$$

In this special case, Equation (41) can be solved analytically. For all $i \in \mathscr{X} \setminus \{0, L\}$, we find that

$$\min_{\pi_i \in \mathcal{Q}_i} \{q_i \underline{\tau}_{i-1\to i} - p_i \underline{\tau}_{i\to i+1}\} \tag{57}$$

$$= \min_{\pi' \in \Sigma_{\mathscr{X}_m}} \left\{[(1-\varepsilon_i)q^*_i + \varepsilon_i q'_i]\underline{\tau}_{i-1\to i} - [(1-\varepsilon_i)p^*_i + \varepsilon_i p'_i]\underline{\tau}_{i\to i+1}\right\} \tag{58}$$

$$= (1-\varepsilon_i)(q^*_i \underline{\tau}_{i-1\to i} - p^*_i \underline{\tau}_{i\to i+1}) + \varepsilon_i \min_{\pi'_i \in \Sigma_{\mathscr{X}_m}} \left\{q'_i \underline{\tau}_{i-1\to i} - p'_i \underline{\tau}_{i\to i+1}\right\}$$

$$= (1-\varepsilon_i)(q^*_i \underline{\tau}_{i-1\to i} - p^*_i \underline{\tau}_{i\to i+1}) - \varepsilon_i \underline{\tau}_{i\to i+1} = \underline{q}_i \underline{\tau}_{i-1\to i} - \overline{p}_i \underline{\tau}_{i\to i+1},$$

where the third equation holds because we know from Theorem 4 that $\underline{\tau}_{i-1\to i}$ and $\underline{\tau}_{i\to i+1}$ are real-valued and positive. Therefore, for all $i \in \mathscr{X} \setminus \{0, L\}$, it follows directly from Equation (41) that

$$\underline{\tau}_{i\to i+1} = \frac{1}{\overline{p}_i} + \frac{\underline{q}_i}{\overline{p}_i}\underline{\tau}_{i-1\to i}.$$

By combining this recursive expression with Equation (34), we can derive explicit expressions. For all $i \in \mathscr{X} \setminus \{L\}$, we find that:

$$\underline{\tau}_{i\to i+1} = \sum_{k=0}^{i} \frac{\prod_{\ell=k+1}^{i} \underline{q}_\ell}{\prod_{m=k}^{i} \overline{p}_m}. \tag{59}$$

In combination with Corollary 8, this equation allows us to easily compute all lower expected upward first-passage times for the linear-vacuous case.

Similar results can be obtained for upper expected upward first-passage times and for lower and upper expected downward first-passage times. For all $i \in \mathscr{X} \setminus \{0, L\}$, we find that

$$\overline{\tau}_{i\to i+1} = \frac{1}{\underline{p}_i} + \frac{\overline{q}_i}{\underline{p}_i}\overline{\tau}_{i-1\to i},\ \underline{\tau}_{i\to i-1} = \frac{1}{\overline{q}_i} + \frac{\underline{p}_i}{\overline{q}_i}\underline{\tau}_{i+1\to i} \text{ and } \overline{\tau}_{i\to i-1} = \frac{1}{\underline{q}_i} + \frac{\overline{p}_i}{\underline{q}_i}\overline{\tau}_{i+1\to i}.$$

By combining these recursive equations with Equations (42) and (44), we can obtain explicit expressions. For all $i \in \mathscr{X} \setminus \{L\}$, we find that

$$\overline{\tau}_{i\to i+1} = \sum_{k=0}^{i} \frac{\prod_{\ell=k+1}^{i} \overline{q}_\ell}{\prod_{m=k}^{i} \underline{p}_m}$$

and, for all $i \in \mathscr{X} \setminus \{0\}$, we find that

$$\underline{\tau}_{i\to i-1} = \sum_{k=i}^{L} \frac{\prod_{\ell=i}^{k-1} \underline{p}_\ell}{\prod_{m=i}^{k} \overline{q}_m} \text{ and } \overline{\tau}_{i\to i-1} = \sum_{k=i}^{L} \frac{\prod_{\ell=i}^{k-1} \overline{p}_\ell}{\prod_{m=i}^{k} \underline{q}_m}. \tag{60}$$

In combination with Proposition 11 and 14, these equations allow us to easily compute all upper expected upward first-passage times and all lower and upper expected downward first-passage times for the linear-vacuous case.

For the lower and upper return times, we still use Equations (45) and (46) if $i = 0$ and Equations (48) and (49) if $i = L$. If $i \in \mathscr{X} \setminus \{0, L\}$, then, for this linear-vacuous case, Equations (47) and (50) can be simplified. We find that

$$\begin{aligned}
\underline{\tau}_{i\to i} &= 1 + \min_{\pi_i \in \mathscr{Q}_i} \left\{q_i \underline{\tau}_{i-1\to i} + p_i \underline{\tau}_{i+1\to i}\right\} \\
&= 1 + \min_{\pi_i' \in \Sigma_{\mathscr{X}_m}} \left\{[(1-\varepsilon_i)q_i^* + \varepsilon_i q_i']\underline{\tau}_{i-1\to i} + [(1-\varepsilon_i)p_i^* + \varepsilon_i p_i']\underline{\tau}_{i+1\to i}\right\} \\
&= 1 + (1-\varepsilon_i)(q_i^*\underline{\tau}_{i-1\to i} + p_i^*\underline{\tau}_{i+1\to i}) = 1 + \underline{q}_i\underline{\tau}_{i-1\to i} + \underline{p}_i\underline{\tau}_{i+1\to i}
\end{aligned} \tag{61}$$

and that

$$\begin{aligned}
\overline{\tau}_{i\to i} &= 1 + (1-\varepsilon_i)(q_i^*\overline{\tau}_{i-1\to i} + p_i^*\overline{\tau}_{i+1\to i}) + \varepsilon_i \max\{\overline{\tau}_{i-1\to i}, \overline{\tau}_{i+1\to i}\} \\
&= 1 + \max\{\overline{q}_i\overline{\tau}_{i-1\to i} + \underline{p}_i\overline{\tau}_{i+1\to i}, \underline{q}_i\overline{\tau}_{i-1\to i} + \overline{p}_i\overline{\tau}_{i+1\to i}\}.
\end{aligned}$$

## 12. Numerical results

In order to illustrate our computational methods, we will now calculate lower and upper expected first-passage and return times for two examples of imprecise birth–death chains. The first is a general example of an imprecise birth–death chain and the second one is an imprecise birth–death chain with linear-vacuous local models. In both examples, we take $\mathscr{Q}_i$ to be identical for all $i \in \mathscr{X} \setminus \{0, L\}$, and simply denote it by $\mathscr{Q}$, which is a credal set on $\mathscr{X}_m$. Some of the lower and upper expectations we compute have many digits after the decimal points; we round them off to the third decimal.

### 12.1. A general example

Consider an imprecise birth–death chain with state space $\mathscr{X} = \{0, \ldots, 4\}$, that is, with $L = 4$. Let $\mathscr{Q}_0$ be the unique credal set on $\mathscr{X}_0$ such that $\underline{p}_0 = 0.15$ and $\overline{p}_0 = 0.4$ and let $\mathscr{Q}_L$ be the unique credal set on $\mathscr{X}_L$ such that $\underline{q}_L = 0.2$ and $\overline{q}_L = 0.6$. The credal set $\mathscr{Q}$ is taken to be the convex hull of the following 7 extreme points, which are of the form $\pi = (q, r, p)$.

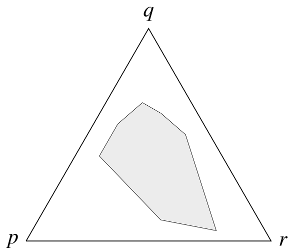


**Fig. 7.** The grey zone depicts the credal set $\mathcal{Q}$ from the birth–death chain in the general example.

**Table 1**
Final results for the general example.

| | |
|---|---|
| $\underline{\tau}_{0\to4}$ | 12.227 |
| $\overline{\tau}_{0\to4}$ | 1420 |
| $\underline{\tau}_{4\to0}$ | 8.094 |
| $\overline{\tau}_{4\to0}$ | 965 |

**Table 2**
Intermediate results for the general example.

| | | | |
|---|---|---|---|
| $\underline{\tau}_{0\to1}$ | 2.5 | $\underline{\tau}_{4\to3}$ | 1.667 |
| $\underline{\tau}_{1\to2}$ | 3.125 | $\underline{\tau}_{3\to2}$ | 2.051 |
| $\underline{\tau}_{2\to3}$ | 3.281 | $\underline{\tau}_{2\to1}$ | 2.170 |
| $\underline{\tau}_{3\to4}$ | 3.320 | $\underline{\tau}_{1\to0}$ | 2.206 |
| $\overline{\tau}_{0\to1}$ | 6.667 | $\overline{\tau}_{4\to3}$ | 5 |
| $\overline{\tau}_{1\to2}$ | 43.333 | $\overline{\tau}_{3\to2}$ | 40 |
| $\overline{\tau}_{2\to3}$ | 226.667 | $\overline{\tau}_{2\to1}$ | 180 |
| $\overline{\tau}_{3\to4}$ | 1143.333 | $\overline{\tau}_{1\to0}$ | 740 |

$$(0.65, 0.15, 0.2), (0.6, 0.25, 0.15), (0.5, 0.4, 0.1), (0.05, 0.75, 0.2),$$
$$(0.1, 0.5, 0.4), (0.4, 0.1, 0.5), (0.55, 0.1, 0.35).$$

Fig. 7 provides a graphical representation of this credal set $\mathcal{Q}$.

For this particular example, we now compute $\underline{\tau}_{0\to4}$, $\overline{\tau}_{0\to4}$, $\underline{\tau}_{4\to0}$ and $\overline{\tau}_{4\to0}$. Due to Corollary 8, we know that

$$\underline{\tau}_{0\to4} = \underline{\tau}_{0\to1} + \underline{\tau}_{1\to2} + \underline{\tau}_{2\to3} + \underline{\tau}_{3\to4}, \tag{62}$$

where, using Equation (34),

$$\underline{\tau}_{0\to1} = 1/\overline{p}_0 = 2.5.$$

By plugging this value for $\underline{\tau}_{0\to1}$ into Equation (41), for $i = 1$, we find that

$$\min_{\pi_1\in\mathcal{Q}} \{2.5q_1 - p_1\underline{\tau}_{1\to2}\} = -1. \tag{63}$$

As we know from Proposition 5, this equality has a unique solution that can for example be obtained by means of a bisection method. We find that $\underline{\tau}_{1\to2} = 3.125$. Similarly, in a recursive fashion, we find that $\underline{\tau}_{2\to3} = 3.281$ and $\underline{\tau}_{3\to4} = 3.320$. A final application of Equation (62) tells us that $\underline{\tau}_{0\to4} = 12.227$. $\overline{\tau}_{0\to4}$, $\underline{\tau}_{4\to0}$ and $\overline{\tau}_{4\to0}$ can be computed analogously; the results are given in Table 1. Intermediate results can be found in Table 2.

This example also illustrates that optimising expected first passage or return times is not just a matter of assigning extreme values to the parameters $q_i$, $r_i$ and $p_i$. For example, in order to obtain the lower expected upward first passage time $\underline{\tau}_{1\to2}$, one might think that it suffices to maximise $p_1$ and minimise $q_1$, as this leads us to move to higher states more quickly, and hence should result in a lower upward first passage time. However, despite the fact that this intuition is correct, it is not always possible to apply it. The problem is that it will sometimes be impossible to maximise $p_1$ and minimise $q_1$ simultaneously, and a tradeoff between these two optimisation criteria is then required. For example, in this case, the minimum in Equation (63) is obtained by the probability mass function $(0.1, 0.5, 0.4)$, whereas the maximal value for $p_1$ is 0.5 and the minimal value for $q_1$ is 0.05. Similar examples can be constructed for upper expected upward first-passage times, for lower and upper downward first-passage times, and for lower and upper return times.

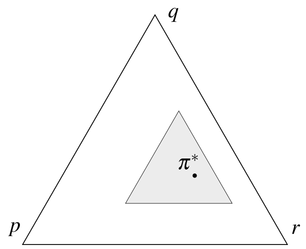


**Fig. 8.** The grey zone depicts the credal set $\mathcal{Q}^{\varepsilon}_{\pi^*}$ from the birth–death chain in the linear-vacuous example.

**Table 3**
Lower and upper expected return times for the birth–death chain in the linear-vacuous mixture example.

| $i$ | $\underline{\tau}_{i\to i}$ | $\overline{\tau}_{i\to i}$ |
|---|---|---|
| 0 | 1.585 | 91.411 |
| 1 | 1.525 | 24.956 |
| 2 | 1.679 | 17.846 |
| 3 | 1.656 | 79.711 |
| 4 | 2.037 | 503.725 |

*12.2. Linear-vacuous example*

Consider a precise birth–death chain with state space $\mathscr{X} = \{0, 1, 2, 3, 4\}$—$L = 4$—and the following probability matrix:

$$P^* = \begin{pmatrix} 0.55 & 0.45 & 0 & 0 & 0 \\ 0.3 & 0.5 & 0.2 & 0 & 0 \\ 0 & 0.3 & 0.5 & 0.2 & 0 \\ 0 & 0 & 0.3 & 0.5 & 0.2 \\ 0 & 0 & 0 & 0.6 & 0.4 \end{pmatrix},$$

which is completely characterised by the probability mass functions $\pi^*_0 = (0.55, 0.45)$, $\pi^*_L = (0.6, 0.4)$ and $\pi^*_i = \pi^* = (0.3, 0.5, 0.2)$ for all $i \in \mathscr{X} \setminus \{0, L\}$.

We now let $\varepsilon_i = \varepsilon = 0.4$ for all $i \in \mathscr{X}$ and consider the imprecise birth–death chain that has the corresponding linear-vacuous credal sets as its local models. In this way, we obtain the following lower and upper probabilities:

$$\underline{p}_0 = 0.27,\ \overline{p}_0 = 0.67,\ \underline{q}_L = 0.36 \text{ and } \overline{q}_L = 0.76$$

and, for all $i \in \mathscr{X} \setminus \{0, L\}$:

$$\underline{q}_i = 0.18,\ \overline{q}_i = 0.58,\ \underline{p}_i = 0.12 \text{ and } \overline{p}_i = 0.52.$$

For all $i \in \mathscr{X} \setminus \{0, L\}$, the credal set $\mathcal{Q}_i$ is equal to $\mathcal{Q}^{\varepsilon}_{\pi^*}$, which is the convex hull of the following three extreme points:

$$(0.58, 0.3, 0.12),\ (0.18, 0.7, 0.12) \text{ and } (0.18, 0.3, 0.52).$$

Fig. 8 provides a graphical representation of this credal set $\mathcal{Q}^{\varepsilon}_{\pi^*}$.

The lower and upper expected return times that correspond to this imprecise birth–death chain can be found in Table 3. For the sake of this example, we compute $\underline{\tau}_{1\to 1}$ explicitly. We start by applying Equation (61) for $i = 1$, which tells us that

$$\underline{\tau}_{1\to 1} = 1 + \underline{q}_1 \underline{\tau}_{0\to 1} + \underline{p}_1 \underline{\tau}_{2\to 1} = 1 + 0.18\underline{\tau}_{0\to 1} + 0.12\underline{\tau}_{2\to 1}.$$

Therefore, and because we know from Equations (59) and (60) that

$$\underline{\tau}_{0\to 1} = \frac{1}{\overline{p}_0} = 1.493 \text{ and } \underline{\tau}_{2\to 1} = \frac{1}{\overline{q}_2} + \frac{\underline{p}_2}{\overline{q}_2\overline{q}_3} + \frac{\underline{p}_2\underline{p}_3}{\overline{q}_2\overline{q}_3\overline{q}_4} = 2.137,$$

we find that $\underline{\tau}_{1\to 1} = 1.525$.

## 13. Conclusion and future work

The main conclusion of this paper is that the lower and upper expected—upward and downward—first-passage times and return times of an imprecise birth–death chain can be computed easily. In particular, by exploiting the properties of the global lower expectation operator of such an imprecise birth–death chain, it is possible to derive a simple system of non-linear equations, and by solving this system, we can then compute any lower or upper expected first-passage or return time through a simple recursive scheme. If the local models of the imprecise birth–death chain are precise or linear vacuous, a simple closed-form expression even suffices. The feasibility of these methods was confirmed by numerical examples. Furthermore, even though an imprecise birth–death chain is not just a set of precise birth–death chains, we have shown that a lower or upper expected first-passage or return time is always achieved by a precise birth–death chain.

For now, our methods impose a strict positivity assumption on the local models of the imprecise birth–death chain. In future work, we would like to drop this local positivity assumption, thereby allowing us to consider some other important types of discrete time imprecise Markov chains, such as pure birth processes, pure death processes and birth–death chains with absorbing states. Finally, we would like to try and apply—suitably adapted versions of—our methods to the Bonus–Malus systems that are described in Reference [15], and to continuous-time imprecise birth–death chains.

## Acknowledgements

The research of Stavros Lopatatzidis and Gert de Cooman was funded through project number 3G012512 of the Research Foundation Flanders (FWO). Jasper De Bock is a Postdoctoral Fellow of the FWO and wishes to acknowledge its financial support. The authors would also like to express their gratitude to the anonymous referees who commented on a previous version of this paper and on a reduced conference version that preceded it [19], for their effort, and for their many useful suggestions and corrections.

## Appendix A. Proofs of results

**Theorem 4.** *For all $i, j \in \mathscr{X}$, the lower and upper first-passage times $\underline{\tau}_{i\to j}$ and $\overline{\tau}_{i\to j}$ are real-valued and strictly positive.*

**Proof.** Since it follows from Equation (25) that $\inf\{\tau_{\to j}(\omega)\colon \omega\in\Gamma(i)\}\geq 1$ and from Equations (26) and (27) that

$$\underline{\tau}_{i\to j}=\underline{E}(\tau_{\to j}(X_{2:\infty})|i) \text{ and } \overline{\tau}_{i\to j}=\overline{E}(\tau_{\to j}(X_{2:\infty})|i), \tag{64}$$

**C5** implies that $1\leq\underline{\tau}_{i\to j}\leq\overline{\tau}_{i\to j}$, and therefore, the only thing that we still need to prove is that $\overline{\tau}_{i\to j}<+\infty$. We will do this by showing that

$$(\exists\mathscr{M}\in\overline{\mathbb{M}})\liminf\mathscr{M}(\omega)\geq\tau^1_{\to j}(\omega) \text{ for all } \omega\in\Gamma(i), \tag{65}$$

where $\tau^1_{\to j}$ is an extended real-valued function on $\Omega$ that is defined by[7]

$$\tau^1_{\to j}(\omega)=\tau^1_{\to j}(x_{1:\infty}):=\tau_{\to j}(x_{2:\infty})=\inf\{m\in\mathbb{N}\colon \omega_{m+1}=j\} \text{ for all } \omega=x_{1:\infty}\in\Omega. \tag{66}$$

Indeed, since it follows from Equations (64), (66) and (23)—in that order—that

$$\overline{\tau}_{i\to j}=\overline{E}(\tau^1_{\to j}|i):=\inf\{\mathscr{M}(i)\colon \mathscr{M}\in\overline{\mathbb{M}} \text{ and } \liminf\mathscr{M}(\omega)\geq\tau^1_{\to j}(\omega) \text{ for all } \omega\in\Gamma(i)\},$$

Equation (65) clearly implies that $\overline{\tau}_{i\to j}<+\infty$.

Consider the values $\varepsilon^u_0:=1/\underline{p}_0$ and $\varepsilon^d_L:=1/\underline{q}_L$. Using $\varepsilon^u_0$ and $\varepsilon^d_L$, we now define recursively, for all $x\in\mathscr{X}\setminus\{0,L\}$:

$$\varepsilon^u_x:=\frac{1}{\underline{p}_x}+\frac{\overline{q}_x}{\underline{p}_x}\varepsilon^u_{x-1} \text{ and } \varepsilon^d_x:=\frac{1}{\underline{q}_x}+\frac{\overline{p}_x}{\underline{q}_x}\varepsilon^d_{x+1}. \tag{67}$$

Due to Assumption 1, we have that $\varepsilon^u_0$ and $\varepsilon^d_L$, as well as $\varepsilon^u_x$ and $\varepsilon^d_x$, for all $x\in\mathscr{X}\setminus\{0,L\}$, are strictly positive and real-valued. Now let $\Delta_j\in\mathscr{G}(\mathscr{X})$ be defined by

$$\Delta_j(i'):=\begin{cases}0 & \text{if } i'=j\\ \sum_{\ell=i'}^{j-1}\varepsilon^u_\ell & \text{if } i'<j\\ \sum_{\ell=j+1}^{i'}\varepsilon^d_\ell & \text{if } i'>j\end{cases} \quad \text{for all } i'\in\mathscr{X}, \tag{68}$$

[7] Basically, $\tau^1_{\to j}$ is just a shifted version of $\tau_{\to j}$. If $x_1\neq j$, $\tau^1_{\to j}(x_{1:\infty})$ provides us with the time immediately before the first occurrence of $j$ (and is equal to $+\infty$ if it never occurs). If $x_1=j$, $\tau^1_{\to j}(x_{1:\infty})$ provides us with the time immediately before the second occurrence of $j$ (and is equal to $+\infty$ if it does not occur a second time).

and consider a real process $\mathscr{M}$, defined for all $m \in \mathbb{N}_0$ and $x_{1:m} \in \mathscr{X}^m$ by

$$\mathscr{M}(x_{1:m}) := \begin{cases} 1 + \overline{Q}_i(\Delta_j) & \text{if } m \in \{0,1\} \text{ or } x_1 \neq i; \\ m - 1 + \Delta_j(x_m) & \text{if } x_1 = i, m \geq 2 \text{ and } (\forall k \in \{2,\dots,m-1\})\, x_k \neq j; \\ \mathscr{M}(x_{1:m-1}) & \text{if } x_1 = i, m \geq 2 \text{ and } (\exists k \in \{2,\dots,m-1\})\, x_k = j. \end{cases} \tag{69}$$

In the remainder of this proof, we show that (65) holds, by proving that $\mathscr{M} \in \overline{\mathbb{M}}$ and that $\liminf \mathscr{M}(\omega) \geq \tau^1_{\to j}(\omega)$ for all $\omega \in \Gamma(i)$.

We start by proving that $\liminf \mathscr{M}(\omega) \geq \tau^1_{\to j}(\omega)$ for all $\omega \in \Gamma(i)$. We consider two cases: $\tau^1_{\to j}(\omega) < +\infty$ and $\tau^1_{\to j}(\omega) = +\infty$. If $\tau^1_{\to j}(\omega) < +\infty$, then with $m := \tau^1_{\to j}(\omega) + 1$, Equation (66) implies that

$$\omega_m = j \text{ and } (\forall k \in \{2,\dots,m-1\})\omega_k \neq j$$

and therefore, because of Equations (69) and (68), for all $n \geq m$, it follows that

$$\mathscr{M}(\omega^n) = \mathscr{M}(\omega^m) = m - 1 + \Delta_j(j) = \tau^1_{\to j}(\omega),$$

which implies that $\liminf_{n\to\infty} \mathscr{M}(\omega^n) = \tau^1_{\to j}(\omega)$. If $\tau^1_{i\to j}(\omega) = +\infty$, Equation (66) implies that $\omega_k \neq j$ for all $k \geq 2$, and therefore, it follows from Equation (69) that

$$\liminf_{n\to\infty} \mathscr{M}(\omega^n) = \liminf_{n\to\infty}(n - 1 + \Delta_j(\omega_n)) \geq \liminf_{n\to\infty}(n-1) = +\infty = \tau^1_{\to j}(\omega),$$

where the inequality holds because it follows from Equation (68) that $\Delta_j(\omega_n) \geq 0$.

We now prove that $\mathscr{M}$ belongs to $\overline{\mathbb{M}}$. From Equation (68), we infer that $\Delta_j \geq 0$ and therefore, it follows from C5 that $\overline{Q}_i(\Delta_j) \geq 0$. Hence, due to Equation (69), it follows that $\mathscr{M}$ is bounded below by 0. Therefore, in order to prove that $\mathscr{M} \in \overline{\mathbb{M}}$, it remains now to prove that $\mathscr{M}$ is a supermartingale, or equivalently, that $\overline{Q}_\square(\Delta\mathscr{M}(\square)) \leq 0$ and $\overline{Q}_{x_m}(\Delta\mathscr{M}(x_{1:m})) \leq 0$ for all $m \in \mathbb{N}$ and $x_{1:m} \in \mathscr{X}^m$.

The first inequality is easily proved: since Equation (69) implies that $\Delta\mathscr{M}(\square) = 0$, it follows from C5 that $\overline{Q}_\square(\Delta\mathscr{M}(\square)) = 0$. So consider any $m \in \mathbb{N}$ and $x_{1:m} \in \mathscr{X}^m$. We need to prove that $\overline{Q}_{x_m}(\Delta\mathscr{M}(x_{1:m})) \leq 0$. We distinguish among three types of situations $x_{1:m}$.

If $x_1 \neq i$ or $x_k = j$ for at least one $k$ in $\{2,\dots,m\}$, then as before, Equation (69) implies that $\Delta\mathscr{M}(x_{1:m}) = 0$, and therefore, it follows from C5 that $\overline{Q}_{x_m}(\Delta\mathscr{M}(x_{1:m})) = 0$.

If $m = 1$ and $x_1 = i$, then

$$\begin{aligned}\overline{Q}_i(\Delta\mathscr{M}(i)) &= \overline{Q}_i(1 + \Delta_j - [1 + \overline{Q}_i(\Delta_j)]) \\ &= \overline{Q}_i(\Delta_j - \overline{Q}_i(\Delta_j)) = \overline{Q}_i(\Delta_j) - \overline{Q}_i(\Delta_j) = 0,\end{aligned}$$

where the first equality follows from Equation (69) and the third equality from C6.

The remaining type of situations $x_{1:m}$ are those for which $m \geq 2$, $x_1 = i$ and $x_k \neq j$ for all $k$ in $\{2,\dots,m\}$. Before tackling this type of situation, we first present some useful equations. For all $x \in \mathscr{X}$, it follows from Equation (69) that

$$\Delta\mathscr{M}(x_{1:m})(x) = (m+1) - 1 + \Delta_j(x) - (m - 1 + \Delta_j(x_m)) = 1 + \Delta_j(x) - \Delta_j(x_m). \tag{70}$$

Combining Equation (70) with Equation (68) results in

$$\Delta\mathscr{M}(x_{1:m})(x_m) = 1. \tag{71}$$

Also, if $x_m \neq L$, then since $x_m \neq j$, we find that

$$\Delta\mathscr{M}(x_{1:m})(x_m + 1) = 1 + \Delta_j(x_m + 1) - \Delta_j(x_m) = \begin{cases} 1 - \varepsilon^u_{x_m} & \text{if } x_m < j; \\ 1 + \varepsilon^d_{x_m+1} & \text{if } x_m > j; \end{cases} \tag{72}$$

similarly, if $x_m \neq 0$, we have that

$$\Delta\mathscr{M}(x_{1:m})(x_m - 1) = 1 + \Delta_j(x_m - 1) - \Delta_j(x_m) = \begin{cases} 1 + \varepsilon^u_{x_m-1} & \text{if } x_m < j; \\ 1 - \varepsilon^d_{x_m} & \text{if } x_m > j. \end{cases} \tag{73}$$

We now consider three cases: $x_m = 0$, $x_m = L$ and $x_m \notin \{0, L\}$.

If $x_m = 0$, it follows from Equations (12) and (17) that

$$\begin{aligned}\overline{Q}_{x_m}(\Delta\mathscr{M}(x_{1:m})) &= \max_{\pi_0 \in \mathscr{Q}_0} \{(1 - p_0)\Delta\mathscr{M}(x_{1:m})(0) + p_0\Delta\mathscr{M}(x_{1:m})(1)\} \\ &= \max_{\pi_0 \in \mathscr{Q}_0} \{(1 - p_0) + p_0(1 - \varepsilon^u_0)\} = \max_{\pi_0 \in \mathscr{Q}_0} \{-p_0\varepsilon^u_0\} + 1 = -\underline{p}_0\varepsilon^u_0 + 1 = 0,\end{aligned}$$

where the second equality follows from Equations (71) and (72) and the fourth holds because $\varepsilon^u_0$ is strictly positive.

If $x_m = L$, it follows from Equations (13) and (17) that

$$\begin{aligned}\overline{Q}_{x_m}(\Delta\mathscr{M}(x_{1:m})) &= \max_{\pi_L\in\mathscr{Q}_L}\{q_L\Delta\mathscr{M}(x_{1:m})(L-1)+(1-q_L)\Delta\mathscr{M}(x_{1:m})(L)\}\\ &= \max_{\pi_L\in\mathscr{Q}_L}\{q_L(1-\varepsilon_L^d)+(1-q_L)\} = \max_{\pi_L\in\mathscr{Q}_L}\{-q_L\varepsilon_L^d\}+1 = -\underline{q}_L\varepsilon_L^d+1=0,\end{aligned}$$

where the second equality follows from Equations (71) and (73) and the fourth holds because $\varepsilon_L^d$ is strictly positive.

If $x_m \notin \{0, L\}$, it follows from Equations (11) and (17) that

$$\begin{aligned}\overline{Q}_{x_m}(\Delta\mathscr{M}(x_{1:m})) &= \max_{\pi_{x_m}\in\mathscr{Q}_{x_m}}\{q_{x_m}\Delta\mathscr{M}(x_{1:m})(x_m-1)\\ &\qquad +(1-q_{x_m}-p_{x_m})\Delta\mathscr{M}(x_{1:m})(x_m)+p_{x_m}\Delta\mathscr{M}(x_{1:m})(x_m+1)\}\\ &= \begin{cases}\max\limits_{\pi_{x_m}\in\mathscr{Q}_{x_m}}\{q_{x_m}\varepsilon_{x_m-1}^u - p_{x_m}\varepsilon_{x_m}^u\}+1 & \text{if } x_m<j;\\ \max\limits_{\pi_{x_m}\in\mathscr{Q}_{x_m}}\{-q_{x_m}\varepsilon_{x_m}^d + p_{x_m}\varepsilon_{x_m+1}^d\}+1 & \text{if } x_m>j,\end{cases}\end{aligned}$$

where the last equality holds because of Equations (71)–(73). Hence, if $x_m < j$, we find that

$$\begin{aligned}\overline{Q}_{x_m}(\Delta\mathscr{M}(x_{1:m})) &= \max_{\pi_{x_m}\in\mathscr{Q}_{x_m}}\{q_{x_m}\varepsilon_{x_m-1}^u - p_{x_m}\varepsilon_{x_m}^u\}+1\\ &\le \max_{\pi_{x_m}\in\mathscr{Q}_{x_m}}\{q_{x_m}\varepsilon_{x_m-1}^u\} + \max_{\pi_{x_m}\in\mathscr{Q}_{x_m}}\{-p_{x_m}\varepsilon_{x_m}^u\}+1\\ &= \overline{q}_{x_m}\varepsilon_{x_m-1}^u - \underline{p}_{x_m}\varepsilon_{x_m}^u+1=0,\end{aligned}$$

where the second equality holds because $\varepsilon_{x_m-1}^u$ and $\varepsilon_{x_m}^u$ are strictly positive and the third equality follows from (67). Similarly, if $x_m > j$, we find that

$$\begin{aligned}\overline{Q}_{x_m}(\Delta\mathscr{M}(x_{1:m})) &= \max_{\pi_{x_m}\in\mathscr{Q}_{x_m}}\{-q_{x_m}\varepsilon_{x_m}^d + p_{x_m}\varepsilon_{x_m+1}^d\}+1\\ &\le \max_{\pi_{x_m}\in\mathscr{Q}_{x_m}}\{-q_{x_m}\varepsilon_{x_m}^d\} + \max_{\pi_{x_m}\in\mathscr{Q}_{x_m}}\{p_{x_m}\varepsilon_{x_m+1}^d\}+1\\ &= -\underline{q}_{x_m}\varepsilon_{x_m}^d + \underline{p}_{x_m}\varepsilon_{x_m+1}^d+1=0. \qquad \square\end{aligned}$$

**Proposition 5.** *Consider a credal set $\mathscr{Q}$ on $\mathscr{X}_m$ that consists of strictly positive probability mass functions and let $c$ be a real constant. Then*

$$\min_{\pi\in\mathscr{Q}}\{qc - p\mu\}$$

*is a strictly decreasing function of $\mu$.*

**Proof.** Consider any $\mu_1, \mu_2 \in \mathbb{R}$, such that $\mu_2 > \mu_1$. Then,

$$\begin{aligned}\min_{\pi\in\mathscr{Q}}\{qc-p\mu_1\} &= \min_{\pi\in\mathscr{Q}}\{qc-p\mu_2+p(\mu_2-\mu_1)\}\\ &\ge \min_{\pi\in\mathscr{Q}}\{qc-p\mu_2\}+\min_{\pi\in\mathscr{Q}}\{p(\mu_2-\mu_1)\} > \min_{\pi\in\mathscr{Q}}\{qc-p\mu_2\}\end{aligned}$$

where the last inequality holds because, since $\mu_2 - \mu_1 > 0$,

$$\min_{\pi\in\mathscr{Q}}\{p(\mu_2-\mu_1)\} = (\mu_2-\mu_1)\min_{\pi\in\mathscr{Q}}\{p\},$$

where $\min_{\pi\in\mathscr{Q}}\{p\} > 0$ because $\mathscr{Q}$ is a credal set that consists of strictly positive probability mass functions. $\square$

**Lemma 24.** *For all $k \in \mathscr{X} \setminus \{0, L\}$, we have that*

$$\min_{\pi_k\in\mathscr{Q}_k}\{q_k\underline{\tau}_{k-1\to k} - p_k\underline{\tau}_{k\to k+1}\} = -1,$$

*and that, for all $\ell$ in $\mathscr{X}$ such that $\ell > k$:*

$$\underline{\tau}_{k-1\to\ell} = \underline{\tau}_{k-1\to k} + \underline{\tau}_{k\to\ell}.$$

**Proof.** For $k = 1$, we have proved in the main text that the lemma holds; see Equations (37) and (38). We will now generalise it using induction. Assuming that the lemma is true for $k-1$, with $k \in \mathscr{X} \setminus \{0, 1, L\}$, we prove that it is also true for $k$.

Consider any $\ell > k$. By taking Equation (32), for $i = k-1$ and $j = \ell$, we find that

$$\begin{aligned}\underline{\tau}_{k-1\to\ell} &= 1 + \min_{\pi_{k-1}\in\mathscr{Q}_{k-1}} \{q_{k-1}\underline{\tau}_{k-2\to\ell} + r_{k-1}\underline{\tau}_{k-1\to\ell} + p_{k-1}\underline{\tau}_{k\to\ell}\}\\ &= 1 + \min_{\pi_{k-1}\in\mathscr{Q}_{k-1}} \{q_{k-1}\underline{\tau}_{k-2\to\ell} + (1-q_{k-1}-p_{k-1})\underline{\tau}_{k-1\to\ell} + p_{k-1}\underline{\tau}_{k\to\ell}\}\\ &= 1 + \underline{\tau}_{k-1\to\ell} + \min_{\pi_{k-1}\in\mathscr{Q}_{k-1}} \{q_{k-1}(\underline{\tau}_{k-2\to\ell} - \underline{\tau}_{k-1\to\ell}) - p_{k-1}(\underline{\tau}_{k-1\to\ell} - \underline{\tau}_{k\to\ell})\},\end{aligned}$$

which, due to Theorem 4, implies that

$$\min_{\pi_{k-1}\in\mathscr{Q}_{k-1}} \{q_{k-1}(\underline{\tau}_{k-2\to\ell} - \underline{\tau}_{k-1\to\ell}) - p_{k-1}(\underline{\tau}_{k-1\to\ell} - \underline{\tau}_{k\to\ell})\} = -1.$$

In combination with the induction hypothesis, which implies that $\underline{\tau}_{k-2\to\ell} = \underline{\tau}_{k-2\to k-1} + \underline{\tau}_{k-1\to\ell}$, the equation above results in

$$\min_{\pi_{k-1}\in\mathscr{Q}_{k-1}} \{q_{k-1}\underline{\tau}_{k-2\to k-1} - p_{k-1}(\underline{\tau}_{k-1\to\ell} - \underline{\tau}_{k\to\ell})\} = -1. \tag{74}$$

Due to Proposition 5 and the induction hypothesis, which implies that

$$\min_{\pi_{k-1}\in\mathscr{Q}_{k-1}} \{q_{k-1}\underline{\tau}_{k-2\to k-1} - p_{k-1}\underline{\tau}_{k-1\to k}\} = -1,$$

we infer from Equation (74) that $\underline{\tau}_{k-1\to\ell} - \underline{\tau}_{k\to\ell} = \underline{\tau}_{k-1\to k}$, and therefore that

$$\underline{\tau}_{k-1\to\ell} = \underline{\tau}_{k-1\to k} + \underline{\tau}_{k\to\ell}. \tag{75}$$

By taking now Equation (32), for $i = k$ and $j = k+1$, we find that

$$\begin{aligned}\underline{\tau}_{k\to k+1} &= 1 + \min_{\pi_k\in\mathscr{Q}_k} \{q_k\underline{\tau}_{k-1\to k+1} + r_k\underline{\tau}_{k\to k+1}\}\\ &= 1 + \min_{\pi_k\in\mathscr{Q}_k} \{q_k\underline{\tau}_{k-1\to k+1} + (1-q_k-p_k)\underline{\tau}_{k\to k+1}\}\\ &= 1 + \underline{\tau}_{k\to k+1} + \min_{\pi_k\in\mathscr{Q}_k} \{q_k(\underline{\tau}_{k-1\to k+1} - \underline{\tau}_{k\to k+1}) - p_k\underline{\tau}_{k\to k+1}\},\end{aligned}$$

which, due to Theorem 4, implies that

$$\min_{\pi_k\in\mathscr{Q}_k} \{q_k(\underline{\tau}_{k-1\to k+1} - \underline{\tau}_{k\to k+1}) - p_k\underline{\tau}_{k\to k+1}\} = -1.$$

By combining this with Equation (75), for $\ell = k+1$, we find that

$$\min_{\pi_k\in\mathscr{Q}_k} \{q_k\underline{\tau}_{k-1\to k} - p_k\underline{\tau}_{k\to k+1}\} = -1. \qquad \square$$

**Proposition 6.** *For any $i \in \mathscr{X} \setminus \{0, L\}$, we have that*

$$\min_{\pi_i\in\mathscr{Q}_i} \{q_i\underline{\tau}_{i-1\to i} - p_i\underline{\tau}_{i\to i+1}\} = -1.$$

**Proof.** This result follows trivially from Lemma 24. $\square$

**Proposition 7.** *For all $i, j \in \mathscr{X}$ such that $i+1 < j$, we have that*

$$\underline{\tau}_{i\to j} = \underline{\tau}_{i\to i+1} + \underline{\tau}_{i+1\to j}.$$

**Proof.** This result follows trivially from Lemma 24. $\square$

**Corollary 8.** *For all $i, j \in \mathscr{X}$ such that $i < j$, we have that*

$$\underline{\tau}_{i\to j} = \sum_{k=i}^{j-1} \underline{\tau}_{k\to k+1}.$$

**Proof.** For $j = i+1$, this result is trivial. For $j = i+2$, it follows from Proposition 7 that

$$\underline{\tau}_{i\to i+2} = \underline{\tau}_{i\to i+1} + \underline{\tau}_{i+1\to i+2}.$$

Similarly, for $j > i+2$, by applying Proposition 7 multiple times, we find that

$$\begin{aligned}\underline{\tau}_{i\to j} &= \underline{\tau}_{i\to i+1} + \underline{\tau}_{i+1\to j} = \underline{\tau}_{i\to i+1} + \underline{\tau}_{i+1\to i+2} + \underline{\tau}_{i+2\to j}\\ &= \underline{\tau}_{i\to i+1} + \underline{\tau}_{i+1\to i+2} + \dots + \underline{\tau}_{j-1\to j} = \sum_{k=0}^{j-1} \underline{\tau}_{k\to k+1}. \qquad \square\end{aligned}$$

**Lemma 25.** *For all $k \in \mathscr{X} \setminus \{0, L\}$, we have that*

$$\max_{\pi_k \in \mathscr{Q}_k} \{q_k \overline{\tau}_{k-1\to k} - p_k \overline{\tau}_{k\to k+1}\} = -1,$$

*and that, for all $\ell$ in $\mathscr{X}$ such that $\ell > k$:*

$$\overline{\tau}_{k-1\to \ell} = \overline{\tau}_{k-1\to k} + \overline{\tau}_{k\to \ell}.$$

**Proof.** For any $k, \ell \in \mathscr{X} \setminus \{0, L\}$, applying Equation (33) for $i = k$ and $j = \ell$ yields

$$\begin{aligned}\overline{\tau}_{k\to\ell} &= 1 + \overline{Q}_k(\mathbb{I}_{\neg\ell}\overline{\tau}_{\bullet\to\ell})\\ &= 1 + \max_{\pi_k \in \mathscr{Q}_k} \{q_k \mathbb{I}_{\neg\ell}(k-1)\overline{\tau}_{k-1\to\ell} + r_k \mathbb{I}_{\neg\ell}(k)\overline{\tau}_{k\to\ell} + p_k \mathbb{I}_{\neg\ell}(k+1)\overline{\tau}_{k+1\to\ell}\}.\end{aligned} \tag{76}$$

We now first prove the case $k = 1$. By applying Equation (33) for $i = 0$, we find that

$$\overline{\tau}_{0\to j} = 1 + \overline{Q}_0(\mathbb{I}_{\neg j}\overline{\tau}_{\bullet\to j}) = 1 + \max_{\pi_0 \in \mathscr{Q}_0} \big\{r_0 \mathbb{I}_{\neg j}(0)\overline{\tau}_{0\to j} + p_0 \mathbb{I}_{\neg j}(1)\overline{\tau}_{1\to j}\big\}. \tag{77}$$

Consider any $\ell \in \mathscr{X}$ such that $\ell > 1$. By applying Equation (77) for $j = \ell$, we then find that

$$\begin{aligned}\overline{\tau}_{0\to\ell} = 1 + \max_{\pi_0 \in \mathscr{Q}_0} \big\{r_0 \overline{\tau}_{0\to\ell} + p_0 \overline{\tau}_{1\to\ell}\big\} &= 1 + \max_{\pi_0 \in \mathscr{Q}_0} \big\{(1-p_0)\overline{\tau}_{0\to\ell} + p_0 \overline{\tau}_{1\to\ell}\big\}\\ &= 1 + \overline{\tau}_{0\to\ell} + \max_{\pi_0 \in \mathscr{Q}_0} \big\{-p_0(\overline{\tau}_{0\to\ell} - \overline{\tau}_{1\to\ell})\big\},\end{aligned}$$

which, due to Theorem 4, implies that

$$\max_{\pi_0 \in \mathscr{Q}_0} \big\{-p_0(\overline{\tau}_{0\to\ell} - \overline{\tau}_{1\to\ell})\big\} = -1 \Rightarrow \overline{\tau}_{0\to\ell} = \frac{1}{\underline{p}_0} + \overline{\tau}_{1\to\ell}.$$

By combining this with Equation (42), we find that

$$\overline{\tau}_{0\to\ell} = \overline{\tau}_{0\to 1} + \overline{\tau}_{1\to\ell}. \tag{78}$$

By taking Equation (76), for $k = 1$ and $\ell = 2$, we find that

$$\begin{aligned}\overline{\tau}_{1\to 2} = 1 + \max_{\pi_1 \in \mathscr{Q}_1} \{q_1 \overline{\tau}_{0\to 2} + r_1 \overline{\tau}_{1\to 2}\} &= 1 + \max_{\pi_1 \in \mathscr{Q}_1} \{q_1 \overline{\tau}_{0\to 2} + (1 - q_1 - p_1)\overline{\tau}_{1\to 2}\}\\ &= 1 + \overline{\tau}_{1\to 2} + \max_{\pi_1 \in \mathscr{Q}_1} \{q_1(\overline{\tau}_{0\to 2} - \overline{\tau}_{1\to 2}) - p_1 \overline{\tau}_{1\to 2}\}\end{aligned}$$

which, due to Theorem 4, implies that

$$\max_{\pi_1 \in \mathscr{Q}_1} \{q_1(\overline{\tau}_{0\to 2} - \overline{\tau}_{1\to 2}) - p_1 \overline{\tau}_{1\to 2}\} = -1.$$

By combining this with Equation (78), for $\ell = 2$, we find that

$$\max_{\pi_1 \in \mathscr{Q}_1} \{q_1 \overline{\tau}_{0\to 1} - p_1 \overline{\tau}_{1\to 2}\} = -1.$$

We will now generalise our proof using induction. Assuming that the lemma is true for $k-1$, with $k \in \mathscr{X} \setminus \{0, 1, L\}$, we prove that it is also true for $k$. Consider any $\ell > k$. By taking Equation (76), for $i = k-1$ and $j = \ell$, we find that

$$\begin{aligned}\overline{\tau}_{k-1\to\ell} &= 1 + \max_{\pi_{k-1}\in\mathscr{Q}_{k-1}} \{q_{k-1}\overline{\tau}_{k-2\to\ell} + r_{k-1}\overline{\tau}_{k-1\to\ell} + p_{k-1}\overline{\tau}_{k\to\ell}\} \\ &= 1 + \max_{\pi_{k-1}\in\mathscr{Q}_{k-1}} \{q_{k-1}\overline{\tau}_{k-2\to\ell} + (1-q_{k-1}-p_{k-1})\overline{\tau}_{k-1\to j} + p_{k-1}\overline{\tau}_{k\to\ell}\} \\ &= 1 + \overline{\tau}_{k-1\to\ell} + \max_{\pi_{k-1}\in\mathscr{Q}_{k-1}} \{q_{k-1}(\overline{\tau}_{k-2\to\ell} - \overline{\tau}_{k-1\to\ell}) - p_{k-1}(\overline{\tau}_{k-1\to\ell} - \overline{\tau}_{k\to\ell})\},\end{aligned}$$

which, due to Theorem 4, implies that

$$\max_{\pi_{k-1}\in\mathscr{Q}_{k-1}} \{q_{k-1}(\overline{\tau}_{k-2\to\ell} - \overline{\tau}_{k-1\to\ell}) - p_{k-1}(\overline{\tau}_{k-1\to\ell} - \overline{\tau}_{k\to\ell})\} = -1.$$

In combination with the induction hypothesis, which implies that $\overline{\tau}_{k-2\to\ell} = \overline{\tau}_{k-2\to k-1} + \overline{\tau}_{k-1\to\ell}$, the equation above results in

$$\max_{\pi_{k-1}\in\mathscr{Q}_{k-1}} \{q_{k-1}\overline{\tau}_{k-2\to k-1} - p_{k-1}(\overline{\tau}_{k-1\to\ell} - \overline{\tau}_{k\to\ell})\} = -1. \tag{79}$$

Due to Proposition 10 (proved further on) and the induction hypothesis, which implies that

$$\max_{\pi_{k-1}\in\mathscr{Q}_{k-1}} \{q_{k-1}\overline{\tau}_{k-2\to k-1} - p_{k-1}\overline{\tau}_{k-1\to k}\} = -1,$$

we infer from Equation (79) that $\overline{\tau}_{k-1\to\ell} - \overline{\tau}_{k\to\ell} = \overline{\tau}_{k-1\to k}$, and therefore that

$$\overline{\tau}_{k-1\to\ell} = \overline{\tau}_{k-1\to k} + \overline{\tau}_{k\to\ell}. \tag{80}$$

By taking now Equation (76), for $\ell = k+1$, we find that

$$\begin{aligned}\overline{\tau}_{k\to k+1} &= 1 + \max_{\pi_k\in\mathscr{Q}_k} \{q_k\overline{\tau}_{k-1\to k+1} + r_k\overline{\tau}_{k\to k+1}\} \\ &= 1 + \max_{\pi_k\in\mathscr{Q}_k} \{q_k\overline{\tau}_{k-1\to k+1} + (1-q_k-p_k)\overline{\tau}_{k\to k+1}\} \\ &= 1 + \overline{\tau}_{k\to k+1} + \max_{\pi_k\in\mathscr{Q}_k} \{q_k(\overline{\tau}_{k-1\to k+1} - \overline{\tau}_{k\to k+1}) - p_k\overline{\tau}_{k\to k+1}\},\end{aligned}$$

which, due to Theorem 4, implies that

$$\max_{\pi_k\in\mathscr{Q}_k} \{q_k(\overline{\tau}_{k-1\to k+1} - \overline{\tau}_{k\to k+1}) - p_k\overline{\tau}_{k\to k+1}\} = -1.$$

By combining this with Equation (80), for $\ell = k+1$, we find that

$$\max_{\pi_k\in\mathscr{Q}_k} \{q_k\overline{\tau}_{k-1\to k} - p_k\overline{\tau}_{k\to k+1}\} = -1. \qquad \square$$

**Proposition 9.** *For all $i \in \mathscr{X} \setminus \{0, L\}$, we have that*

$$\max_{\pi_i\in\mathscr{Q}_i} \{q_i\overline{\tau}_{i-1\to i} - p_i\overline{\tau}_{i\to i+1}\} = -1.$$

**Proof.** This result follows trivially from Lemma 25. $\square$

**Proposition 10.** *Consider a credal set $\mathscr{Q}$ on $\mathscr{X}_m$ that consists of strictly positive probability mass functions and let $c$ be a real constant. Then*

$$\max_{\pi\in\mathscr{Q}} \{qc - p\mu\}$$

*is a strictly decreasing function of $\mu$.*

**Proof.** Consider any $\mu_1, \mu_2 \in \mathbb{R}$, such that $\mu_2 > \mu_1$. Then,

$$\begin{aligned}\max_{\pi\in\mathscr{Q}} \{qc - p\mu_2\} &= \max_{\pi\in\mathscr{Q}} \{qc - p\mu_1 + p(\mu_1 - \mu_2)\} \\ &\le \max_{\pi\in\mathscr{Q}} \{qc - p\mu_1\} + \max_{\pi\in\mathscr{Q}} \{p(\mu_1 - \mu_2)\} < \max_{\pi\in\mathscr{Q}} \{qc - p\mu_1\}\end{aligned}$$

where the last inequality holds because

$$\max_{\pi\in\mathscr{Q}} \{p(\mu_1 - \mu_2)\} = (\mu_1 - \mu_2) \max_{\pi\in\mathscr{Q}} \{p\}$$

where $\max_{\pi\in\mathscr{Q}}\{p\} > 0$ because $\mathscr{Q}$ is a credal set that consists of strictly positive probability mass functions. $\square$

**Proposition 11.** *For all $i, j \in \mathscr{X}$ such that $i < j$, we have that*

$$\overline{\tau}_{i\to j} = \sum_{k=i}^{j-1} \overline{\tau}_{k\to k+1}.$$

**Proof.** For $j = i+1$, this result is trivial. For $j = i+2$, it follows from Lemma 25 that

$$\overline{\tau}_{i\to i+2} = \overline{\tau}_{i\to i+1} + \overline{\tau}_{i+1\to i+2}.$$

Similarly, for $j > i+2$, by applying Lemma 25 multiple times, we find that

$$\begin{aligned}\overline{\tau}_{i\to j} &= \overline{\tau}_{i\to i+1} + \overline{\tau}_{i+1\to j} = \overline{\tau}_{i\to i+1} + \overline{\tau}_{i+1\to i+2} + \overline{\tau}_{i+2\to j}\\ &= \overline{\tau}_{i\to i+1} + \overline{\tau}_{i+1\to i+2} + \ldots + \overline{\tau}_{j-1\to j} = \sum_{k=0}^{j-1} \overline{\tau}_{k\to k+1}. \qquad \square\end{aligned}$$

**Lemma 26.** *For all $k \in \mathscr{X} \setminus \{0, L\}$, we have that*

$$\min_{\pi_k\in\mathscr{Q}_k} \{-q_k\underline{\tau}_{k\to k-1} + p_k\underline{\tau}_{k+1\to k}\} = -1,$$

*and that, for all $\ell$ in $\mathscr{X}$ such that $k > \ell$:*

$$\underline{\tau}_{k+1\to\ell} = \underline{\tau}_{k+1\to k} + \underline{\tau}_{k\to\ell}.$$

**Proof.** We first prove the case $k = L-1$. Consider any $\ell \in \mathscr{X}$ such that $\ell < L-1$. By taking Equation (31), for $j = \ell$, we find that

$$\begin{aligned}\underline{\tau}_{L\to\ell} &= 1 + \min_{\pi_L\in\mathscr{Q}_L} \{q_L\underline{\tau}_{L-1\to\ell} + r_L\underline{\tau}_{L\to\ell}\} = 1 + \min_{\pi_L\in\mathscr{Q}_L} \{q_L\underline{\tau}_{L-1\to\ell} + (1-q_L)\underline{\tau}_{L\to\ell}\}\\ &= 1 + \underline{\tau}_{L\to\ell} + \min_{\pi_L\in\mathscr{Q}_L} \{-q_L(\underline{\tau}_{L\to\ell} - \underline{\tau}_{L-1\to\ell})\},\end{aligned}$$

which, due to Theorem 4, implies that

$$\min_{\pi_L\in\mathscr{Q}_L} \{-q_L(\underline{\tau}_{L\to\ell} - \underline{\tau}_{L-1\to\ell})\} = -1 \Rightarrow \underline{\tau}_{L\to\ell} = \frac{1}{\overline{q}_L} + \underline{\tau}_{L-1\to\ell}.$$

By combining this with Equation (44), we find that

$$\underline{\tau}_{L\to\ell} = \underline{\tau}_{L\to L-1} + \underline{\tau}_{L-1\to\ell}. \tag{81}$$

By applying Equation (32) for $i = L-1$ and $j = L-2$, we find that

$$\begin{aligned}\underline{\tau}_{L-1\to L-2} &= 1 + \min_{\pi_{L-1}\in\mathscr{Q}_{L-1}} \{r_{L-1}\underline{\tau}_{L-1\to L-2} + p_{L-1}\underline{\tau}_{L\to L-2}\}\\ &= 1 + \min_{\pi_{L-1}\in\mathscr{Q}_{L-1}} \{(1-q_{L-1}-p_{L-1})\underline{\tau}_{L-1\to L-2} + p_{L-1}\underline{\tau}_{L\to L-2}\}\\ &= 1 + \underline{\tau}_{L-1\to L-2} + \min_{\pi_{L-1}\in\mathscr{Q}_{L-1}} \{-q_{L-1}\underline{\tau}_{L-1\to L-2} + p_{L-1}(\underline{\tau}_{L\to L-2} - \underline{\tau}_{L-1\to L-2})\}\end{aligned}$$

which, due to Theorem 4, implies that

$$\min_{\pi_{L-1}\in\mathscr{Q}_{L-1}} \{-q_{L-1}\underline{\tau}_{L-1\to L-2} + p_{L-1}(\underline{\tau}_{L\to L-2} - \underline{\tau}_{L-1\to L-2})\} = -1.$$

Combining the equation above with Equation (81), for $\ell = L-2$, we find that

$$\min_{\pi_{L-1}\in\mathscr{Q}_{L-1}} \{-q_{L-1}\underline{\tau}_{L-1\to L-2} + p_{L-1}\underline{\tau}_{L\to L-1}\} = -1.$$

We will now generalise the proof using induction. Assuming that the lemma is true for $k+1$, with $k \in \mathscr{X} \setminus \{0, L-1, L\}$, we prove that it is also true for $k$.

Consider any $\ell \in \mathscr{X}$ such that $\ell < k$. By applying Equation (32) for $i = k+1$ and $j = \ell$, we find that

$$\begin{aligned}\underline{\tau}_{k+1\to\ell} &= 1 + \min_{\pi_{k+1}\in\mathscr{Q}_{k+1}} \{q_{k+1}\underline{\tau}_{k\to\ell} + r_{k+1}\underline{\tau}_{k+1\to\ell} + p_{k+1}\underline{\tau}_{k+2\to\ell}\}\\ &= 1 + \min_{\pi_{k+1}\in\mathscr{Q}_{k+1}} \{q_{k+1}\underline{\tau}_{k\to\ell} + (1-q_{k+1}-p_{k+1})\underline{\tau}_{k+1\to\ell} + p_{k+1}\underline{\tau}_{k+2\to\ell}\}\\ &= 1 + \underline{\tau}_{k+1\to\ell} + \min_{\pi_{k+1}\in\mathscr{Q}_{k+1}} \{-q_{k+1}(\underline{\tau}_{k+1\to\ell} - \underline{\tau}_{k\to\ell}) + p_{k+1}(\underline{\tau}_{k+2\to\ell} - \underline{\tau}_{k+1\to\ell})\},\end{aligned}$$

which, due to Theorem 4, implies that

$$\min_{\pi_{k+1}\in\mathscr{Q}_{k+1}} \{-q_{k+1}(\underline{\tau}_{k+1\to\ell} - \underline{\tau}_{k\to\ell}) + p_{k+1}(\underline{\tau}_{k+2\to\ell} - \underline{\tau}_{k+1\to\ell})\} = -1.$$

In combination with the induction hypothesis, which implies that $\underline{\tau}_{k+2\to\ell} = \underline{\tau}_{k+2\to k+1} + \underline{\tau}_{k+1\to\ell}$, the equation above results in

$$\min_{\pi_{k+1}\in\mathscr{Q}_{k+1}} \{-q_{k+1}(\underline{\tau}_{k+1\to\ell} - \underline{\tau}_{k\to\ell}) + p_{k+1}\underline{\tau}_{k+2\to k+1}\} = -1. \tag{82}$$

Due to Proposition 13 (proved further on) and the induction hypothesis, which implies that

$$\min_{\pi_{k+1}\in\mathscr{Q}_{k+1}} \{-q_{k+1}\underline{\tau}_{k+1\to k} + p_{k+1}\underline{\tau}_{k+2\to k+1}\} = -1,$$

we infer from Equation (82) that $\underline{\tau}_{k+1\to\ell} - \underline{\tau}_{k\to\ell} = \underline{\tau}_{k+1\to k}$, and therefore that

$$\underline{\tau}_{k+1\to\ell} = \underline{\tau}_{k+1\to k} + \underline{\tau}_{k\to\ell}. \tag{83}$$

By taking now Equation (32), for $i = k$ and $j = k-1$, we find that

$$\begin{aligned}\underline{\tau}_{k\to k-1} &= 1 + \min_{\pi_k\in\mathscr{Q}_k} \{r_k\underline{\tau}_{k\to k-1} + p_k\underline{\tau}_{k+1\to k-1}\}\\ &= 1 + \min_{\pi_k\in\mathscr{Q}_k} \{(1-q_k-p_k)\underline{\tau}_{k\to k-1} + p_k\underline{\tau}_{k+1\to k-1}\}\\ &= 1 + \underline{\tau}_{k\to k-1} + \min_{\pi_k\in\mathscr{Q}_k} \{-q_k\underline{\tau}_{k\to k-1} + p_k(\underline{\tau}_{k+1\to k-1} - \underline{\tau}_{k\to k-1})\},\end{aligned}$$

which, due to Theorem 4, implies that

$$\min_{\pi_k\in\mathscr{Q}_k} \{-q_k\underline{\tau}_{k\to k-1} + p_k(\underline{\tau}_{k+1\to k-1} - \underline{\tau}_{k\to k-1})\} = -1.$$

By combining this with Equation (83), for $\ell = k-1$, we find that

$$\min_{\pi_k\in\mathscr{Q}_k} \{-q_k\underline{\tau}_{k-1\to k} + p_k\underline{\tau}_{k+1\to k}\} = -1. \qquad \square$$

**Lemma 27.** *For all $k \in \mathscr{X} \setminus \{0, L\}$, we have that*

$$\max_{\pi_k\in\mathscr{Q}_k} \{-q_k\overline{\tau}_{k\to k-1} + p_k\overline{\tau}_{k+1\to k}\} = -1,$$

*and that, for all $\ell$ in $\mathscr{X}$ such that $k > \ell$:*

$$\overline{\tau}_{k+1\to\ell} = \overline{\tau}_{k+1\to k} + \overline{\tau}_{k\to\ell}.$$

**Proof.** We first prove the case $k = L-1$. By taking Equation (33), for $i = L$, we find that

$$\overline{\tau}_{L\to j} = 1 + \overline{Q}_L(\mathbb{I}_{\neg j}\overline{\tau}_{\bullet\to j}) = 1 + \max_{\pi_L\in\mathscr{Q}_L} \left\{q_L\mathbb{I}_{\neg j}(L-1)\overline{\tau}_{L-1\to j} + r_L\mathbb{I}_{\neg j}(L)\overline{\tau}_{L\to j}q_L\right\}. \tag{84}$$

Consider any $\ell \in \mathscr{X}$ such that $\ell < L-1$. By applying Equation (84) for $j = \ell$, we find that

$$\begin{aligned}\overline{\tau}_{L\to\ell} = 1 + \max_{\pi_L\in\mathscr{Q}_L} \{q_L\overline{\tau}_{L-1\to\ell} + r_L\overline{\tau}_{L\to\ell}\} &= 1 + \max_{\pi_L\in\mathscr{Q}_L} \{q_L\overline{\tau}_{L-1\to\ell} + (1-q_L)\overline{\tau}_{L\to\ell}\}\\ &= 1 + \overline{\tau}_{L\to\ell} + \max_{\pi_L\in\mathscr{Q}_L} \left\{-q_L(\overline{\tau}_{L\to\ell} - \overline{\tau}_{L-1\to\ell})\right\},\end{aligned}$$

which, due to Theorem 4, implies that

$$\max_{\pi_L\in\mathscr{Q}_L} \left\{-q_L(\overline{\tau}_{L\to\ell} - \overline{\tau}_{L-1\to\ell})\right\} = -1 \Rightarrow \overline{\tau}_{L\to\ell} = \frac{1}{\underline{q}_L} + \overline{\tau}_{L-1\to\ell}.$$

By combining this with Equation (44), we find that

$$\overline{\tau}_{L\to\ell} = \overline{\tau}_{L\to L-1} + \overline{\tau}_{L-1\to\ell}. \tag{85}$$

By taking Equation (76), for $k = L-1$ and $\ell = L-2$, we find that

$$\begin{aligned}\overline{\tau}_{L-1\to L-2} &= 1 + \max_{\pi_{L-1}\in\mathscr{Q}_{L-1}} \{r_{L-1}\overline{\tau}_{L-1\to L-2} + p_{L-1}\overline{\tau}_{L\to L-2}\}\\ &= 1 + \max_{\pi_{L-1}\in\mathscr{Q}_{L-1}} \{(1-q_{L-1}-p_{L-1})\overline{\tau}_{L-1\to L-2} + p_{L-1}\overline{\tau}_{L\to L-2}\}\\ &= 1 + \overline{\tau}_{L-1\to L-2} + \max_{\pi_{L-1}\in\mathscr{Q}_{L-1}} \{-q_{L-1}\overline{\tau}_{L-1\to L-2} + p_{L-1}(\overline{\tau}_{L\to L-2} - \overline{\tau}_{L-1\to L-2})\}\end{aligned}$$

which, due to Theorem 4, implies that

$$\max_{\pi_{L-1}\in\mathscr{Q}_{L-1}} \{-q_{L-1}\overline{\tau}_{L-1\to L-2} + p_{L-1}(\overline{\tau}_{L\to L-2} - \overline{\tau}_{L-1\to L-2})\} = -1.$$

Combining the equation above with Equation (85), for $\ell = L-2$, we find that

$$\max_{\pi_{L-1}\in\mathscr{Q}_{L-1}} \{-q_{L-1}\overline{\tau}_{L-1\to L-2} + p_{L-1}\overline{\tau}_{L\to L-1}\} = -1.$$

We will now generalise the proof using induction. Assuming that the lemma is true for $k+1$, with $k \in \mathscr{X} \setminus \{0, L-1, L\}$, we prove that it is also true for $k$.

Consider any $\ell < k$. By taking Equation (33), for $i = k+1$ and $j = \ell$, we find that

$$\begin{aligned}\overline{\tau}_{k+1\to\ell} &= 1 + \max_{\pi_{k+1}\in\mathscr{Q}_{k+1}} \{q_{k+1}\overline{\tau}_{k\to\ell} + r_{k+1}\overline{\tau}_{k+1\to\ell} + p_{k+1}\overline{\tau}_{k+2\to\ell}\}\\ &= 1 + \max_{\pi_{k+1}\in\mathscr{Q}_{k+1}} \{q_{k+1}\overline{\tau}_{k\to j} + (1-q_{k+1}-p_{k+1})\overline{\tau}_{k+1\to j} + p_{k+1}\overline{\tau}_{k+2\to\ell}\}\\ &= 1 + \overline{\tau}_{k+1\to\ell} + \max_{\pi_{k+1}\in\mathscr{Q}_{k+1}} \{-q_{k+1}(\overline{\tau}_{k+1\to\ell} - \overline{\tau}_{k\to\ell}) + p_{k+1}(\overline{\tau}_{k+2\to\ell} - \overline{\tau}_{k+1\to\ell})\},\end{aligned}$$

which, due to Theorem 4, implies that

$$\max_{\pi_{k+1}\in\mathscr{Q}_{k+1}} \{-q_{k+1}(\overline{\tau}_{k+1\to\ell} - \overline{\tau}_{k\to\ell}) + p_{k+1}(\overline{\tau}_{k+2\to\ell} - \overline{\tau}_{k+1\to\ell})\} = -1.$$

In combination with the induction hypothesis, which implies that $\overline{\tau}_{k+2\to\ell} = \overline{\tau}_{k+2\to k+1} + \overline{\tau}_{k+1\to\ell}$, the equation above results in

$$\max_{\pi_{k+1}\in\mathscr{Q}_{k+1}} \{-q_{k+1}(\overline{\tau}_{k+1\to\ell} - \overline{\tau}_{k\to\ell}) + p_{k+1}\overline{\tau}_{k+2\to k+1}\} = -1. \tag{86}$$

Due to Proposition 13 (proved further on) and the induction hypothesis, which implies that

$$\max_{\pi_{k+1}\in\mathscr{Q}_{k+1}} \{-q_{k+1}\overline{\tau}_{k+1\to k} + p_{k+1}\overline{\tau}_{k+2\to k+1}\} = -1,$$

we infer from Equation (86) that $\overline{\tau}_{k+1\to\ell} - \overline{\tau}_{k\to\ell} = \overline{\tau}_{k+1\to k}$, and therefore that

$$\overline{\tau}_{k+1\to\ell} = \overline{\tau}_{k+1\to k} + \overline{\tau}_{k\to\ell}. \tag{87}$$

By taking Equation (33), for $i = k$ and $j = k-1$, we find that

$$\begin{aligned}\overline{\tau}_{k\to k-1} &= 1 + \min_{\pi_k\in\mathscr{Q}_k} \{r_k\overline{\tau}_{k\to k-1} + p_k\overline{\tau}_{k+1\to k-1}\}\\ &= 1 + \max_{\pi_k\in\mathscr{Q}_k} \{(1-q_k-p_k)\overline{\tau}_{k\to k-1} + p_k\overline{\tau}_{k+1\to k-1}\}\\ &= 1 + \overline{\tau}_{k\to k-1} + \max_{\pi_k\in\mathscr{Q}_k} \{-q_k\overline{\tau}_{k\to k-1} + p_k(\overline{\tau}_{k+1\to k-1} - \overline{\tau}_{k\to k-1})\},\end{aligned}$$

which, due to Theorem 4, implies that

$$\max_{\pi_k\in\mathscr{Q}_k} \{-q_k\overline{\tau}_{k\to k-1} + p_k(\overline{\tau}_{k+1\to k-1} - \overline{\tau}_{k\to k-1})\} = -1.$$

By combining this with Equation (87), for $\ell = k-1$, we find that

$$\min_{\pi_k\in\mathscr{Q}_k} \{-q_k\overline{\tau}_{k-1\to k} + p_k\overline{\tau}_{k+1\to k}\} = -1. \qquad \square$$

**Proposition 12.** *For all $i \in \mathscr{X} \setminus \{0, L\}$, we have that*

$$\min_{\pi_i \in \mathscr{Q}_i} \{-q_i \underline{\tau}_{i\to i-1} + p_i \underline{\tau}_{i+1\to i}\} = -1 \text{ and } \max_{\pi_i \in \mathscr{Q}_i} \{-q_i \overline{\tau}_{i\to i-1} + p_i \overline{\tau}_{i+1\to i}\} = -1.$$

**Proof.** This result follows trivially from Lemmas 26 and 27. □

**Proposition 13.** *Consider a credal set $\mathscr{Q}$ on $\mathscr{X}_m$ that consists of strictly positive probability mass functions and let $c$ be a real constant. Then*

$$\min_{\pi \in \mathscr{Q}} \{-q\mu + pc\} \text{ and } \max_{\pi \in \mathscr{Q}} \{-q\mu + pc\}$$

*are strictly decreasing functions of $\mu$.*

**Proof.** Consider any $\mu_1, \mu_2 \in \mathbb{R}$, such that $\mu_2 > \mu_1$. Then

$$\begin{aligned}\min_{\pi \in \mathscr{Q}} \{-q\mu_1 + pc\} &= \min_{\pi \in \mathscr{Q}} \{q(\mu_2 - \mu_1) - q\mu_2 + pc\} \\ &\geq \min_{\pi \in \mathscr{Q}} \{q(\mu_2 - \mu_1)\} + \min_{\pi \in \mathscr{Q}} \{-q\mu_2 + pc\} > \min_{\pi \in \mathscr{Q}} \{-q\mu_2 + pc\},\end{aligned}$$

where the last inequality holds because

$$\min_{\pi \in \mathscr{Q}} \{q(\mu_2 - \mu_1)\} = (\mu_2 - \mu_1) \min_{\pi \in \mathscr{Q}} \{q\}$$

where $\min_{\pi \in \mathscr{Q}}\{q\} > 0$ because $\mathscr{Q}$ is a credal set that consists of strictly positive probability mass functions.

Similarly,

$$\begin{aligned}\max_{\pi \in \mathscr{Q}} \{-q\mu_2 + pc\} &= \max_{\pi \in \mathscr{Q}} \{q(\mu_1 - \mu_2) - q\mu_1 + pc\} \\ &\leq \max_{\pi \in \mathscr{Q}} \{q(\mu_1 - \mu_2)\} + \max_{\pi \in \mathscr{Q}} \{-q\mu_1 + pc\} < \max_{\pi \in \mathscr{Q}} \{-q\mu_1 + pc\},\end{aligned}$$

where the last inequality holds because

$$\max_{\pi \in \mathscr{Q}} \{q(\mu_1 - \mu_2)\} = (\mu_1 - \mu_2) \max_{\pi \in \mathscr{Q}} \{q\},$$

where $\max_{\pi \in \mathscr{Q}}\{q\} > 0$ because $\mathscr{Q}$ is a credal set that consists of strictly positive probability mass functions. □

**Proposition 14.** *For all $i, j \in \mathscr{X}$ such that $i > j$, we have that*

$$\underline{\tau}_{i\to j} = \sum_{k=j}^{i-1} \underline{\tau}_{k+1\to k} \text{ and } \overline{\tau}_{i\to j} = \sum_{k=j}^{i-1} \overline{\tau}_{k+1\to k}.$$

**Proof.** We prove first the lower case. For $j = i - 1$, this result is trivial. For $j = i - 2$, it follows from Lemma 26 that

$$\underline{\tau}_{i\to i-2} = \underline{\tau}_{i\to i-1} + \underline{\tau}_{i-1\to i-2}.$$

Similarly, for $j < i - 2$, by applying Lemma 26 multiple times, we find that

$$\begin{aligned}\underline{\tau}_{i\to j} &= \underline{\tau}_{i\to i-1} + \underline{\tau}_{i-1\to j} = \underline{\tau}_{i\to i-1} + \underline{\tau}_{i-1\to i-2} + \underline{\tau}_{i-2\to j} \\ &= \underline{\tau}_{i\to i-1} + \underline{\tau}_{i-1\to i-2} + \ldots + \underline{\tau}_{j+1\to j} = \sum_{k=j}^{i-1} \underline{\tau}_{k+1\to k}.\end{aligned}$$

Now we prove the upper case. For $j = i - 1$, this result is trivial. For $j = i - 2$, it follows from Lemma 27 that

$$\overline{\tau}_{i\to i-2} = \overline{\tau}_{i\to i-1} + \overline{\tau}_{i-1\to i-2}.$$

Similarly, for $j < i - 2$, by applying Lemma 27 multiple times, we find that

$$\begin{aligned}\overline{\tau}_{i\to j} &= \overline{\tau}_{i\to i-1} + \overline{\tau}_{i-1\to j} = \overline{\tau}_{i\to i-1} + \overline{\tau}_{i-1\to i-2} + \overline{\tau}_{i-2\to j} \\ &= \overline{\tau}_{i\to i-1} + \overline{\tau}_{i-1\to i-2} + \ldots + \overline{\tau}_{j+1\to j} = \sum_{k=j}^{i-1} \overline{\tau}_{k+1\to k}.\end{aligned}$$

□

**Proposition 15.** *Consider any imprecise birth–death chain such that, for all $i \in \mathscr{X}$, $\mathscr{Q}_i = \{\pi_i\}$. Then*

$$\underline{\tau}_{i\to j} = \overline{\tau}_{i\to j} \text{ for all } i, j \in \mathscr{X}.$$

**Proof.** This result is an immediate consequence of the fact that, in this case, our recursive equations for computing $\underline{\tau}_{i\to j}$ become identical to the equations that we use to compute $\overline{\tau}_{i\to j}$. For example, for upward first-passage times, Equation (34) is now identical to Equation (42) because $\underline{p}_0 = p_0 = \overline{p}_0$, and Equation (41) is now identical to Equation (43) because, since $\mathscr{Q}_i = \{\pi_i\}$, the minimum and maximum disappear. Similar observations can also be made for all the other recursive equations in Sections 7–9. □

**Proposition 16.** *Consider a precise birth–death chain of which the stochastic matrix $P$ is given by Equation (8). Then for all $i \in \mathscr{X} \setminus \{L\}$, we have that*

$$\tau^P_{i\to i+1} = \sum_{k=0}^{i} \frac{\prod_{\ell=k+1}^{i} q_\ell}{\prod_{m=k}^{i} p_m}.$$

**Proof.** We provide a proof by induction. Since $\overline{p}_0 = p_0$, it follows from Equations (51) and (34) that $\tau^P_{0\to 1} = \underline{\tau}_{0\to 1} = 1/\overline{p}_0 = 1/p_0$, proving Equation (52) for $i = 0$.

Consider now any $i \in \mathscr{X} \setminus \{0, L\}$ and let us assume, as our induction hypothesis, that the result is true for $i-1$. Since $\mathscr{Q}_i = \{\pi_i\}$, it follows from Proposition 6 and Equation (51) that $q_i\tau^P_{i-1\to i} - p_i\tau^P_{i\to i+1} = -1$, and therefore, Assumption (1) implies that

$$\tau^P_{i\to i+1} = \frac{1}{p_i} + \frac{q_i}{p_i}\tau^P_{i-1\to i} = \frac{1}{p_i} + \frac{q_i}{p_i}\sum_{k=0}^{i-1} \frac{\prod_{\ell=k+1}^{i-1} q_\ell}{\prod_{m=k}^{i-1} p_m} = \frac{1}{p_i} + \sum_{k=0}^{i-1} \frac{\prod_{\ell=k+1}^{i} q_\ell}{\prod_{m=k}^{i} p_m} = \sum_{k=0}^{i} \frac{\prod_{\ell=k+1}^{i} q_\ell}{\prod_{m=k}^{i} p_m},$$

where the second equality follows from the induction hypothesis. □

**Proposition 17.** *Consider a precise birth–death chain of which the stochastic matrix $P$ is given by Equation (8). Then for all $i \in \mathscr{X} \setminus \{0\}$, we have that*

$$\tau^P_{i\to i-1} = \sum_{k=i}^{L} \frac{\prod_{\ell=i}^{k-1} p_\ell}{\prod_{m=i}^{k} q_m}.$$

**Proof.** This proof is completely analogous to the proof of Proposition 17. Again, we provide a proof by induction. Since $\overline{q}_L = q_L$, it follows from Equations (51) and (44) that $\tau^P_{L\to L-1} = \underline{\tau}_{L\to L-1} = 1/\overline{q}_L = 1/q_L$, proving the result for $i = L$.

Consider now any $i \in \mathscr{X} \setminus \{0, L\}$ and let us assume, as our induction hypothesis, that the result is true for $i+1$. Since $\mathscr{Q}_i = \{\pi_i\}$, it follows from Proposition 12 and Equation (51) that $-q_i\tau^P_{i\to i-1} + p_i\tau^P_{i+1\to i} = -1$, and therefore, Assumption (1) implies that

$$\tau^P_{i\to i-1} = \frac{1}{q_i} + \frac{p_i}{q_i}\tau^P_{i+1\to i} = \frac{1}{q_i} + \frac{p_i}{q_i}\sum_{k=i+1}^{L} \frac{\prod_{\ell=i+1}^{k-1} p_\ell}{\prod_{m=i+1}^{k} q_m} = \frac{1}{q_i} + \sum_{k=i+1}^{L} \frac{\prod_{\ell=i}^{k-1} p_\ell}{\prod_{m=i}^{k} q_m} = \sum_{k=i}^{L} \frac{\prod_{\ell=i}^{k-1} p_\ell}{\prod_{m=i}^{k} q_m},$$

where the second equality follows from the induction hypothesis. □

**Theorem 18.** *Consider an imprecise birth–death chain, some $k \in \mathscr{X}$, and a precise birth death chain whose stochastic matrix $P$ is obtained from this imprecise birth–death chain by means of Selection Method $\mathrm{LU}_k$. Then for all $i, j \in \mathscr{X}$ such that $i < j \le k$, $\underline{\tau}_{i\to j} = \tau^P_{i\to j}$.*

**Proof.** Due to Corollary 8 and Equation (53), it clearly suffices to prove that

$$\underline{\tau}_{i\to i+1} = \tau^P_{i\to i+1} \text{ for all } i < k. \tag{88}$$

We provide a proof by induction. Since we know from Selection Method $\mathrm{LU}_k$ (1) that $\overline{p}_0 = p_0$, it follows from Equation (34) and Proposition 16 that $\underline{\tau}_{0\to 1} = 1/\overline{p}_0 = 1/p_0 = \tau^P_{0\to 1}$, which proves Equation (88) for $i = 0$.

Consider now any $i \in \{1, \dots, k-1\}$ and let us assume, as our induction hypothesis, that Equation (88) is true for $i-1$, that is, $\underline{\tau}_{i-1\to i} = \tau^P_{i-1\to i}$. Then on the one hand, if we apply Proposition 6 to the imprecise birth–death chain, it follows from Selection Method $\mathrm{LU}_k$ (2) that $q_i\underline{\tau}_{i-1\to i} - p_i\underline{\tau}_{i\to i+1} = -1$. On the other hand, if we apply Proposition 6 to the precise birth–death chain with stochastic matrix $P$, we find that $q_i\tau^P_{i-1\to i} - p_i\tau^P_{i\to i+1} = -1$. By combining these two statements with the induction hypothesis, it follows that $p_i\underline{\tau}_{i\to i+1} = p_i\tau^P_{i\to i+1}$, which, because of Assumption (1), implies that $\underline{\tau}_{i\to i+1} = \tau^P_{i\to i+1}$, as required. □

**Theorem 19.** *Consider an imprecise birth–death chain, some $k \in \mathscr{X}$, and a precise birth death chain whose stochastic matrix $P$ is obtained from this imprecise birth–death chain by means of Selection Method $UU_k$. Then for all $i, j \in \mathscr{X}$ such that $i < j \leq k$, $\overline{\tau}_{i\to j} = \tau^P_{i\to j}$.*

**Proof.** Due to Proposition 11 and Equation (53), it suffices to prove that

$$\overline{\tau}_{i\to i+1} = \tau^P_{i\to i+1} \text{ for all } i < k. \tag{89}$$

We provide a proof by induction. Since we know from Selection Method UU$_k$ (1) that $\underline{p}_0 = p_0$, it follows from Equation (42) and Proposition 16 that $\overline{\tau}_{0\to 1} = 1/\underline{p}_0 = 1/p_0 = \tau^P_{0\to 1}$, which proves Equation (89) for $i = 0$.

Consider now any $i \in \{1, \dots, k-1\}$ and let us assume, as our induction hypothesis, that Equation (89) is true for $i-1$, that is, $\overline{\tau}_{i-1\to i} = \tau^P_{i-1\to i}$. Then on the one hand, if we apply Proposition 9 to the imprecise birth–death chain, it follows from Selection Method UU$_k$ (2) that $q_i\overline{\tau}_{i-1\to i} - p_i\overline{\tau}_{i\to i+1} = -1$. On the other hand, if we apply Proposition 9 to the precise birth–death chain with stochastic matrix $P$, we find that $q_i\tau^P_{i-1\to i} - p_i\tau^P_{i\to i+1} = -1$. By combining these two statements with the induction hypothesis, it follows that $p_i\overline{\tau}_{i\to i+1} = p_i\tau^P_{i\to i+1}$, which, because of Assumption (1), implies that $\overline{\tau}_{i\to i+1} = \tau^P_{i\to i+1}$, as required. □

**Theorem 20.** *Consider an imprecise birth–death chain, some $k \in \mathscr{X}$, and a precise birth death chain whose stochastic matrix $P$ is obtained from this imprecise birth–death chain by means of Selection Method $LD_k$. Then for all $i, j \in \mathscr{X}$ such that $k \leq j < i$, $\underline{\tau}_{i\to j} = \tau^P_{i\to j}$.*

**Proof.** Due to Proposition 14 and Equation (54), it suffices to prove that

$$\underline{\tau}_{i\to i-1} = \tau^P_{i\to i-1} \text{ for all } i > k. \tag{90}$$

We provide a proof by induction. Since we know from Selection Method LD$_k$ (1) that $\overline{q}_L = q_L$, it follows from Equation (44) and Proposition 17 that $\underline{\tau}_{L\to L-1} = 1/\overline{q}_L = 1/q_L = \tau^P_{L\to L-1}$, which proves Equation (90) for $i = L$.

Consider now any $i \in \{k+1, \dots, L-1\}$ and let us assume, as our induction hypothesis, that Equation (90) is true for $i+1$, that is, $\underline{\tau}_{i+1\to i} = \tau^P_{i+1\to i}$. Then on the one hand, if we apply Proposition 12 to the imprecise birth–death chain, it follows from Selection Method LD$_k$ (2) that $-q_i\underline{\tau}_{i\to i-1} + p_i\underline{\tau}_{i+1\to i} = -1$. On the other hand, if we apply Proposition 12 to the precise birth–death chain with stochastic matrix $P$, we find that $-q_i\tau^P_{i\to i-1} + p_i\tau^P_{i+1\to i} = -1$. By combining these two statements with the induction hypothesis, it follows that $q_i\underline{\tau}_{i\to i-1} = q_i\tau^P_{i\to i-1}$, which, because of Assumption (1), implies that $\underline{\tau}_{i\to i-1} = \tau^P_{i\to i-1}$, as required. □

**Theorem 21.** *Consider an imprecise birth–death chain, some $k \in \mathscr{X}$, and a precise birth death chain whose stochastic matrix $P$ is obtained from this imprecise birth–death chain by means of Selection Method $UD_k$. Then for all $i, j \in \mathscr{X}$ such that $k \leq j < i$, $\overline{\tau}_{i\to j} = \tau^P_{i\to j}$.*

**Proof.** Due to Proposition 14 and Equation (54), it suffices to prove that

$$\overline{\tau}_{i\to i-1} = \tau^P_{i\to i-1} \text{ for all } i > k. \tag{91}$$

We provide a proof by induction. Since we know from Selection Method UD$_k$ (1) that $\underline{q}_L = q_L$, it follows from Equation (44) and Proposition 17 that $\overline{\tau}_{L\to L-1} = 1/\underline{q}_L = 1/q_L = \tau^P_{L\to L-1}$, which proves Equation (91) for $i = L$.

Consider now any $i \in \{k+1, \dots, L-1\}$ and let us assume, as our induction hypothesis, that Equation (90) is true for $i+1$, that is, $\overline{\tau}_{i+1\to i} = \tau^P_{i+1\to i}$. Then on the one hand, if we apply Proposition 12 to the imprecise birth–death chain, it follows from Selection Method LD$_k$ (2) that $-q_i\overline{\tau}_{i\to i-1} + p_i\overline{\tau}_{i+1\to i} = -1$. On the other hand, if we apply Proposition 12 to the precise birth–death chain with stochastic matrix $P$, we find that $-q_i\tau^P_{i\to i-1} + p_i\tau^P_{i+1\to i} = -1$. By combining these two statements with the induction hypothesis, it follows that $q_i\overline{\tau}_{i\to i-1} = q_i\tau^P_{i\to i-1}$, which, because of Assumption (1), implies that $\overline{\tau}_{i\to i-1} = \tau^P_{i\to i-1}$, as required. □

**Theorem 22.** *Consider an imprecise birth–death chain, some $k \in \mathscr{X}$, and a precise birth death chain whose stochastic matrix $P$ is obtained from this imprecise birth–death chain by means of Selection Method $LR_k$. Then $\underline{\tau}_{k\to k} = \tau^P_{k\to k}$.*

**Proof.** First observe that, since Selection Method LR$_k$ implies Selection Method LU$_k$ and LD$_k$, we can use Theorem 18 and 20 to find that

$$\left(k \neq 0 \Rightarrow \underline{\tau}_{k-1\to k} = \tau^P_{k-1\to k}\right) \text{ and } \left(k \neq L \Rightarrow \underline{\tau}_{k+1\to k} = \tau^P_{k+1\to k}\right). \tag{92}$$

We now prove the theorem for $k = 0$. Since we know from Selection Method LR$_k$ (1) that $\underline{p}_0 = p_0$, it follows from Equations (45) and (92) that

$$\underline{\tau}_{0\to 0} = 1 + \underline{p}_0 \underline{\tau}_{1\to 0} = 1 + p_0 \underline{\tau}_{1\to 0} = 1 + p_0 \tau^P_{1\to 0}$$

and therefore, due to Equation (55), we infer that $\underline{\tau}_{0\to 0} = \tau^P_{0\to 0}$.

The case $k = L$ is proved similarly. Since we know from Selection Method LR$_k$ (5) that $\underline{q}_L = q_L$, it follows from Equations (46) and (92) that

$$\underline{\tau}_{L\to L} = 1 + \underline{q}_L \underline{\tau}_{L-1\to L} = 1 + q_L \underline{\tau}_{L-1\to L} = 1 + q_L \tau^P_{L-1\to L}$$

and therefore, due to Equation (55), we find that $\underline{\tau}_{L\to L} = \tau^P_{L\to L}$.

It remains now to prove the theorem for the case $k \in \{1, \dots, L-1\}$. Since we know from Selection Method LR$_k$ (3) that $(q_k, r_k, p_k) \in \arg\min_{\pi_k \in \mathscr{Q}_k} \{q_k \underline{\tau}_{k-1\to k} + p_k \underline{\tau}_{k+1\to k}\}$, it follows from Equations (47) and (92) that

$$\underline{\tau}_{k\to k} = 1 + q_k \underline{\tau}_{k-1\to k} + p_k \underline{\tau}_{k+1\to k} = 1 + q_k \tau^P_{k-1\to k} + p_k \tau^P_{k+1\to k}$$

and therefore, Equation (56) implies that $\underline{\tau}_{k\to k} = \tau^P_{k\to k}$, as required. □

**Theorem 23.** *Consider an imprecise birth–death chain, some $k \in \mathscr{X}$, and a precise birth death chain whose stochastic matrix $P$ is obtained from this imprecise birth–death chain by means of Selection Method UR$_k$. Then $\overline{\tau}_{k\to k} = \tau^P_{k\to k}$.*

**Proof.** First observe that, since Selection Method UR$_k$ implies Selection Method UU$_k$ and UD$_k$, we can use Theorem 19 and 21 to find that

$$\left(k \neq 0 \Rightarrow \overline{\tau}_{k-1\to k} = \tau^P_{k-1\to k}\right) \text{ and } \left(k \neq L \Rightarrow \overline{\tau}_{k+1\to k} = \tau^P_{k+1\to k}\right). \tag{93}$$

We now prove the theorem for $k = 0$. Since we know from Selection Method UR$_k$ (1) that $\overline{p}_0 = p_0$, it follows from Equations (48) and (93) that

$$\overline{\tau}_{0\to 0} = 1 + \overline{p}_0 \overline{\tau}_{1\to 0} = 1 + p_0 \overline{\tau}_{1\to 0} = 1 + p_0 \tau^P_{1\to 0}$$

and therefore, due to Equation (55), we infer that $\overline{\tau}_{0\to 0} = \tau^P_{0\to 0}$.

The case $k = L$ is proved similarly. Since we know from Selection Method UR$_k$ (5) that $\overline{q}_L = q_L$, it follows from Equations (49) and (93) that

$$\overline{\tau}_{L\to L} = 1 + \overline{q}_L \overline{\tau}_{L-1\to L} = 1 + q_L \overline{\tau}_{L-1\to L} = 1 + q_L \tau^P_{L-1\to L}$$

and therefore, due to Equation (55), we find that $\overline{\tau}_{L\to L} = \tau^P_{L\to L}$.

It remains now to prove the theorem for the case $k \in \{1, \dots, L-1\}$. Since we know from Selection Method LR$_k$ (3) that $(q_k, r_k, p_k) \in \arg\max_{\pi_k \in \mathscr{Q}_k} \{q_k \overline{\tau}_{k-1\to k} + p_k \overline{\tau}_{k+1\to k}\}$, it follows from Equations (50) and (93) that

$$\overline{\tau}_{k\to k} = 1 + q_k \overline{\tau}_{k-1\to k} + p_k \overline{\tau}_{k+1\to k} = 1 + q_k \tau^P_{k-1\to k} + p_k \tau^P_{k+1\to k}$$

and therefore, Equation (56) implies that $\overline{\tau}_{k\to k} = \tau^P_{k\to k}$, as required. □

## References

[1] Linda J.S. Allen, An Introduction to Stochastic Processes with Applications to Biology, Prentice Hall, 2003.
[2] Thomas Augustin, Frank P.A. Coolen, Gert de Cooman, Matthias C.M. Troffaes (Eds.), Introduction to Imprecise Probabilities, John Wiley & Sons, 2014.
[3] Jasper De Bock, Gert de Cooman, Credal networks under epistemic irrelevance: the sets of desirable gambles approach, Int. J. Approx. Reason. 56 (Part B) (2015) 178–207.
[4] Inés Couso, Serafín Moral, Peter Walley, A survey of concepts of independence for imprecise probabilities, Risk Decis. Policy 5 (02) (2000) 165–181.
[5] Fabio G. Cozman, Sets of probability distributions, independence, and convexity, Synthese 186 (2012) 177–200.
[6] Richard J. Crossman, Pauline Coolen-Schrijner, Frank P.A. Coolen, Time-homogeneous birth–death processes with probability intervals and absorbing state, J. Stat. Theory Pract. 3 (1) (2009) 103–118.
[7] Jasper De Bock, Credal networks under epistemic irrelevance: theory and algorithms, PhD thesis, Ghent University, Faculty of Engineering and Architecture, 2015.
[8] Jasper De Bock, Gert de Cooman, Allowing for probability zero in credal networks under epistemic irrelevance, in: ISIPTA'13: Proceedings of the Eighth International Symposium on Imprecise Probability: Theories and Applications, SIPTA, 2013, pp. 109–118.
[9] Luis M. de Campos, Serafín Moral, Independence concepts for convex sets of probabilities, in: Eleventh Conference on Uncertainty in Artificial Intelligence, 1995, pp. 108–115.
[10] Gert de Cooman, Filip Hermans, Imprecise probability trees: bridging two theories of imprecise probability, Artif. Intell. 172 (11) (2008) 1400–1427.
[11] Gert de Cooman, Filip Hermans, Erik Quaeghebeur, Imprecise Markov chains and their limit behaviour, Probab. Eng. Inf. Sci. 23 (4) (2009) 597–635.

[12] Gert de Cooman, Filip Hermans, Alessandro Antonucci, Marco Zaffalon, Epistemic irrelevance in credal nets: the case of imprecise Markov trees, Int. J. Approx. Reason. 51 (2010) 1029–1052.
[13] Gert de Cooman, Enrique Miranda, Marco Zaffalon, Independent natural extension, Artif. Intell. 175 (2011) 1911–1950.
[14] Gert De Cooman, Jasper De Bock, Stavros Lopatatzidis, Imprecise stochastic processes in discrete time: global models, imprecise Markov chains, and ergodic theorems, Int. J. Approx. Reason. 76 (2016) 18–46.
[15] Michel Denuit, Xavier Maréchal, Sandra Pitrebois, Jean-François Walhin, Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus–Malus Systems, John Wiley & Sons, 2007.
[16] Donald Gross, John F. Shortle, James M. Thompson, Carl M. Harris, Fundamentals of Queueing Theory, Wiley–Interscience, 2008.
[17] Peter J. Huber, Robust Statistics, John Wiley & Sons, 1981.
[18] Igor O. Kozine, Lev V. Utkin, Interval-valued finite Markov chains, Reliab. Comput. 8 (2) (2002) 97–113.
[19] Stavros Lopatatzidis, Jasper De Bock, Gert de Cooman, Calculating bounds on expected return and first-passage times in finite-state imprecise birth–death chains, in: ISIPTA '15 – Proceedings of the Ninth International Symposium on Imprecise Probability: Theories and Applications, SIPTA, 2015, pp. 177–186.
[20] Enrique Miranda, Gert de Cooman, Lower previsions, in: Introduction to Imprecise Probabilities, John Wiley & Sons, 2014.
[21] Glenn Shafer, The Art of Causal Conjecture, The MIT Press, 1996.
[22] Glenn Shafer, Volodya Vovk, Probability and Finance: It's Only a Game!, Wiley, 2001.
[23] Ushio Sumita, Yasushi Masuda, On first-passage time structure of random walks, Stoch. Process. Appl. 20 (1) (1985) 133–147.
[24] Matthias C.M. Troffaes, Gert de Cooman, Lower Previsions, Wiley, 2014.
[25] Škulj Damjan, Regular finite Markov chains with interval probabilities, in: ISIPTA '07 – Proceedings of the Fifth International Symposium on Imprecise Probability: Theories and Applications, SIPTA, 2007, pp. 405–413.
[26] Peter Walley, Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, 1991.
[27] Peter Whittle, Probability via Expectation, Springer, 2000.