First Steps Towards an Imprecise Poisson Process

Alexander Erreygers   ALEXANDER.ERREYGERS@UGENT.BE

Jasper De Bock   JASPER.DEBOCK@UGENT.BE

ELIS – FLip, Ghent University, Belgium

Abstract

The Poisson process is the most elementary continuous-time stochastic process that models a stream of repeating events. It is uniquely characterised by a single parameter called the rate. Instead of a single value for this rate, we here consider a rate interval and let it characterise two nested sets of stochastic processes. We call these two sets of stochastic process imprecise Poisson processes, explain why this is justified, and study the corresponding lower and upper (conditional) expectations. Besides a general theoretical framework, we also provide practical methods to compute lower and upper (conditional) expectations of functions that depend on the number of events at a single point in time.

Keywords: Poisson process, counting process, continuous-time Markov chain, imprecision

1. Introduction

The *Poisson process* is arguably one of the most basic stochastic processes. At the core of this model is our subject, who is interested in something specific that occurs repeatedly over time, where time is assumed to be continuous. For instance, our subject could be interested in the arrival of a customer to a queue, to give an example from queueing theory. For the sake of brevity, we will call such a specific occurrence a *Poisson-event*,[1] whence our subject is interested in a stream of Poisson-events. The time instants at which subsequent Poisson-events occur are uncertain to our subject, hence the need for a probabilistic model. This set-up is not exclusive to queueing theory; it is also used in renewal theory and reliability theory, to name but a few applications.

There is a plethora of alternative but essentially equivalent characterisations of this Poisson process. Some of the more well-known and basic characterisations are as the limit of the Bernoulli process [4, Chapter VI, Sections 5 and 6] or as a sequence of mutually independent and exponentially distributed inter-event times [5, Chapter 5, Section 3.A]. An alternative way to look at the Poisson process

is as a random dispersion of points in some general space—that need not be the real number line—see for instance [1, Sections 2.1 and 2.2] or [7, Chapter 2]. More theoretically involved characterisations that follow our set-up are (i) as a counting process—a type of continuous-time stochastic process—that has independent and stationary increments that are Poisson distributed, see for example [11, Definition 2.1.1] or [12, Section 3]; (ii) as a counting process with a condition on the "rate"—the rate of change of the probability of having a Poisson-event in a vanishingly small time period—see for instance [4, Chapter XVII, Section 2] or [11, Definition 2.1.2]; (iii) as a stationary counting process that has no after-effects—see for instance [4, Chapter XVII, Section 2] or [6, Section 1]; (iv) as a martingale through the Watanabe characterisation [14, Theorem 2.3]; or (v) as a pure-birth chain—a type of continuous-time Markov chain—with one birth rate, see for instance [9, Section 2.4]. Many of these characterisations are actually equivalent, see for instance [9, Theorem 2.4.3] or [11, Theorem 2.1.1].

Broadly speaking, these characterisations all make the same three assumptions: (i) *orderliness*, in the sense that the probability that two or more Poisson-events occur at the same time is zero; (ii) independence, more specifically the absence of after-effects or *Markovianity*; and (iii) *homogeneity*. It is essentially well-known that these three assumptions imply the existence of a parameter called the rate, and that this rate uniquely characterises the Poisson process. We here weaken the three aforementioned assumptions. First and foremost, we get rid of the implicit assumption that our subject's beliefs can be accurately modelled by a single stochastic process; instead, we assume that her beliefs only allow us to consider a *set* of stochastic processes. Specifically, we consider a rate interval instead of a precise value for the rate, and examine two distinct sets: (i) the set of all Poisson processes whose rate belongs to this rate interval; and (ii) the set of all processes that are orderly and "consistent" with the rate interval. We then define lower and upper conditional expectations as the infimum and supremum of the conditional expectations with respect to the stochastic processes in these respective sets. Aside from this general theoretical framework, we focus on computing the lower and upper expectation of functions that depend on the number of occurred Poisson-events at a single future time point. For the set of Poisson processes, we show that this requires the solution of a one-parameter

---

1. We use the term "Poisson-event" rather than just "event" to avoid confusion with the standard usage of event in probability theory, where event refers to a subset of the sample space; we are indebted to an anonymous reviewer for pointing out this potential confusion, and to Gert de Cooman for suggesting the adopted terminology.

optimisation problem; for the second set, we show that this can be computed using backwards recursion. Furthermore, we argue that both sets can be justifiably called imprecise Poisson processes: imprecise because their lower and upper expectations are not equal, and Poisson because their lower and upper expectations satisfy imprecise versions of the defining properties of the (precise) Poisson process. The interested reader can find proofs for all our results in the Appendix.

Our approach is heavily inspired by the theory of imprecise continuous-time Markov chains [8]. For instance, we define the imprecise Poisson process via consistency with a rate interval, whereas Krak et al. [8] use consistency with a set of transition rate matrices. In the bigger picture, our contribution can therefore be seen as the first steps towards generalising the theory of imprecise continuous-time Markov chains from finite to countably infinite state spaces.

2. Counting Processes in General

Recall from the Introduction that our subject is interested in the occurrences of a Poisson-event. In this setting, it makes sense to consider the number of Poisson-events that have occurred from the initial time point $t_{\text{ini}} = 0$ up to a time point t, where t is a non-negative real number.

2.1. Counting Paths and the Sample Space

The temporal evolution of the number of occurred Poisson-events is given by a counting path $\omega \colon \mathbb{R}_{\geq 0} \to \mathbb{Z}_{\geq 0}$; at any time point t in $\mathbb{R}_{\geq 0}$, $\omega(t)$ is the number of Poisson-events that have occurred from $t_{\text{ini}} = 0$ up to t.[2] Since the actual temporal evolution of the number of occurred Poisson-events is unknown to the subject, we need a probabilistic model, more specifically a continuous-time stochastic process. The sample space—the space of all possible outcomes—of this process is a set of counting paths, denoted by Ω. One popular choice for Ω is the set of all càdlàg—right-continuous with left limits—counting paths, in this set-up usually also assumed to be non-decreasing. However, our results do not require such a strong assumption. Before we state our assumptions on Ω, we first introduce some notation.

In the remainder, we frequently use increasing sequences $t_1, \ldots, t_n$ of time points, that is, sequences $t_1, \ldots, t_n$ in $\mathbb{R}_{\geq 0}$ of arbitrary length—that is, with n in $\mathbb{N}$—such that $t_i < t_{i+1}$ for all i in $\{1, \ldots, n-1\}$. For the sake of brevity, we follow [8, Section 2.1] in denoting such a sequence by u. We collect all increasing—but possibly empty—sequences of time points in $\mathscr{U}$, and let $\mathscr{U}_\emptyset := \mathscr{U} \setminus \{\emptyset\}$. Observe that as a sequence of time points u in $\mathscr{U}$ is just a finite and

ordered set of non-negative real numbers, we can perform common set-theoretic operations on them like unions. In order to lighten our notation, we identify the single time point t with a sequence; as such, we can use $u \cup t$ as a notational shorthand for $u \cup \{t\}$. Also, a statement of the form $\max u < t$ is taken to be true if $u = \emptyset$; see for instance Lemma 3. With this convention, for any t in $\mathbb{R}_{\geq 0}$, we let $\mathscr{U}_{<t} := \{u \in \mathscr{U} \colon \max u < t\}$ be the set of all sequences of time points of which the last time point precedes t. Note that if $t = 0$, then there is no such non-empty sequence and so $\mathscr{U}_{<t} = \{\emptyset\}$.

In order to better distinguish between general non-negative integers and counts, we let $\mathscr{X} := \mathbb{Z}_{\geq 0}$. For any $u = t_1, \ldots, t_n$ in $\mathscr{U}_\emptyset$, we let $\mathscr{X}_u$ be the set of all n-tuples $x_u = (x_{t_1}, \ldots, x_{t_n})$ of non-negative integers that are non-decreasing:

$$\mathscr{X}_u := \{(x_{t_1}, \ldots, x_{t_n}) \in \mathscr{X}^n \colon x_{t_1} \leq \cdots \leq x_{t_n}\}. \quad (1)$$

If u is the empty sequence $\emptyset$, then we let $\mathscr{X}_u = \mathscr{X}_\emptyset$ denote the singleton containing the empty tuple, denoted by $x_\emptyset$.

With all this notation in place, we can now formally state our requirements on Ω:

A1. $(\forall \omega \in \Omega)(\forall t, \Delta \in \mathbb{R}_{\geq 0})\ \omega(t) \leq \omega(t + \Delta)$;

A2. $(\forall u \in \mathscr{U}_\emptyset)(\forall x_u \in \mathscr{X}_u)(\exists \omega \in \Omega)(\forall t \in u)\ \omega(t) = x_t$.

Assumption (A1) ensures that all paths are non-decreasing, which is essential if we interpret $\omega(t)$ as the number of Poisson-events that have occurred up to time t. Assumption (A2) ensures that the set Ω is sufficiently large, essentially ensuring that the finitary events of Equation (2) further on are non-empty.

2.2. Coherent Conditional Probabilities

We follow Krak et al. [8] in using the framework of coherent conditional probabilities to model the beliefs of our subject. What follows is a brief introduction to coherent conditional probabilities; we refer to [10] and [8, Section 4.1] for a more detailed exposition. For any sample space—that is, a non-empty set—S, we let $\mathscr{E}(S)$ denote the set all events—that is, subsets of S—and let $\mathscr{E}_\emptyset(S) := \mathscr{E}(S) \setminus \{\emptyset\}$ denote the set of all non-empty events. Before we introduce coherent conditional probabilities, we first look at full conditional probabilities.

Definition 1 *Let S be a sample space. A full conditional probability P is a real-valued map on $\mathscr{E}(S) \times \mathscr{E}_\emptyset(S)$ such that, for all A, B in $\mathscr{E}(S)$ and C, D in $\mathscr{E}_\emptyset(S)$,*

P1. $P(A \mid C) \geq 0$;

P2. $P(A \mid C) = 1$ if $C \subseteq A$;

P3. $P(A \cup B \mid C) = P(A \mid C) + P(B \mid C)$ if $A \cap B = \emptyset$;

P4. $P(A \cap D \mid C) = P(A \mid D \cap C)P(D \mid C)$ if $D \cap C \neq \emptyset$.

---

2. We use $\mathbb{Z}_{\geq 0}$ and $\mathbb{N}$ to denote the non-negative integers and natural numbers (or positive integers), respectively. Furthermore, the real numbers, non-negative real numbers and positive real numbers are denoted by $\mathbb{R}$, $\mathbb{R}_{\geq 0}$ and $\mathbb{R}_{>0}$, respectively.

Note that (P1)–(P3) just state that $P(\cdot \,|\, C)$ is a finitely-additive probability measure, and that (P4) is a multiplicative version of Bayes' rule. We use the adjective *full* because the domain of P is $\mathscr{E}(S) \times \mathscr{E}_\emptyset(S)$. Next, we move to domains that are a subset of $\mathscr{E}(S) \times \mathscr{E}_\emptyset(S)$.

Definition 2 *Let S be a sample space. A coherent conditional probability is a real-valued map P on $\mathscr{D} \subseteq \mathscr{E}(S) \times \mathscr{E}_\emptyset(S)$ that can be extended to a full conditional probability.*

Important to emphasise here is that simply demanding that (P1)–(P4) hold on the domain $\mathscr{D}$ is in general *not* sufficient to guarantee that P can be extended to a full conditional probability. A necessary and sufficient condition for the existence of such an extension can be found in [10, Theorem 3] or [8, Corollary 4.3], but we refrain from stating it here because of its technical nature. We here only mention that this so-called *coherence* condition—hence explaining the use of the adjective coherent—has an intuitive betting interpretation, and that checking this condition is usually feasible while explicitly constructing the full conditional extension is typically not; this is extremely useful when constructing proofs. Another strong argument for using coherent conditional probabilities is that they can always be extended to a coherent conditional probability on a larger domain [10, Theorem 4]. This too is an essential tool in the proof of many of our main results, including Theorems 6, 15 and 19.

2.3. Events and Fields

For any $v = t_1,\ldots,t_n$ in $\mathscr{U}_\emptyset$ and $B \subseteq \mathscr{X}_v$, we define the *finitary event*

$$(X_v \in B) := \{\omega \in \Omega \colon (\omega(t_1),\ldots,\omega(t_n)) \in B\}. \quad (2)$$

Furthermore, we also let $(X_\emptyset = x_\emptyset) := \Omega =: (X_\emptyset \in \mathscr{X}_\emptyset)$. Then for any u in $\mathscr{U}$, we let $\mathscr{F}_u$ be the field of events—or algebra of sets—generated by the finitary events for all sequences with time points in or succeeding u:

$$\mathscr{F}_u := \langle \{(X_v \in B) \colon v \in \mathscr{U}, B \subseteq \mathscr{X}_v, \\ (\forall t \in v)\, t \in u \cup [\max u, +\infty)\} \rangle. \quad (3)$$

Lemma 3 *Consider some u in $\mathscr{U}$ and A in $\mathscr{F}_u$. Then there is some v in $\mathscr{U}$ with $\min v > \max u$ and some $B \subseteq \mathscr{X}_w$ with $w := u \cup v$ such that $A = (X_w \in B)$.*

2.4. Counting Processes as Coherent Conditional Probabilities

From here on, we focus on coherent conditional probabilities with the domain

$$\mathscr{D}_{CP} := \{(A,X_u = x_u) \colon u \in \mathscr{U}, A \in \mathscr{F}_u, x_u \in \mathscr{X}_u\},$$

which essentially consists of future events conditional on the number of occurred Poisson-events at specified past time-points. The rationale behind this domain is twofold. First and foremost, it is sufficiently large to make most inferences that one is usually interested in. For example, this domain allows us to compute—tight lower and upper bounds on—the expectation of a real-valued function on the number of occurred Poisson-events at a single future time point, as we will see in Section 6. Second, it guarantees that every rate corresponds to a unique Poisson process, as we will see in Section 3.

Definition 4 *A counting process is a coherent conditional probability P on $\mathscr{D}_{CP}$ such that*

CP1. $P(X_0 = 0) = 1$;

CP2. *for all t in $\mathbb{R}_{\geq 0}$, u in $\mathscr{U}_{<t}$ and (x_u,x) in $\mathscr{X}_{u \cup t}$,*

$$\lim_{\Delta \to 0^+} \frac{P(X_{t+\Delta} \geq x+2 \,|\, X_u = x_u, X_t = x)}{\Delta} = 0$$

and, if $t > 0$,

$$\lim_{\Delta \to 0^+} \frac{P(X_t \geq x+2 \,|\, X_u = x_u, X_{t-\Delta} = x)}{\Delta} = 0.$$

The second requirement (CP2) is—our version of—the *orderliness* property that we previously mentioned in the Introduction. In essence, it ensures that the probability that two or more Poisson-events occur at the same time is zero. We collect all counting processes in the set $\mathbb{P}$.

2.5. Conditional Expectation with Respect to a Counting Process

For any counting process P, we let E_P denote the associated (conditional) expectation, defined in the usual sense as an integral with respect to the measure P—see for instance [10, Theorem 6] or [13, Section 15.10.1].

Let $\mathscr{K}_b(\Omega)$ denote the set of all real-valued functions on Ω that are bounded below.[3] Fix some u in $\mathscr{U}$. Then f in $\mathscr{K}_b(\Omega)$ is $\mathscr{F}_u$-*measurable* if for all α in $[\inf f, +\infty)$, the level set $\{f > \alpha\} := \{\omega \in \Omega \colon f(\omega) > \alpha\}$ is an element of $\mathscr{F}_u$. We collect all such $\mathscr{F}_u$-measurable functions in $\mathscr{G}_u$.

The (conditional) expectation E_P has domain

$$\mathscr{G} := \{(f, X_u = x_u) \in \mathscr{K}_b(\Omega) \times \mathscr{E}_\emptyset(\Omega) \colon \\ u \in \mathscr{U}, x_u \in \mathscr{X}_u, f \in \mathscr{G}_u\}.$$

For any $(f, X_u = x_u)$ in $\mathscr{G}$, we have

---

3. Note that we could just as well consider arbitrary real-valued functions instead of restricting ourselves to bounded-below functions. Our main reason for doing so is that this facilitates a more elegant treatment. Furthermore, many functions of practical interest are bounded-below.

$$E_P(f \,|\, X_u = x_u)$$
$$:= \inf f + \int_{\inf f}^{\sup f} P(\{f > \alpha\} \,|\, X_u = x_u) \,d\alpha,$$

where the integral is a—possibly improper—Riemann integral. Note that this integral always exists because $P(\{f > \alpha\} \,|\, X_u = x_u)$ is a non-increasing function of α. This expression simplifies if f is an $\mathscr{F}_u$-*simple function*. To define these, we let $\mathbb{I}_A : \Omega \to \mathbb{R}$ denote the indicator of an event $A \subseteq \Omega$, defined for all ω in Ω as $\mathbb{I}_A(\omega) := 1$ if $\omega \in A$ and 0 otherwise. We then say that f is $\mathscr{F}_u$-simple if it can be written as $f = \sum_{i=1}^{n} a_i \mathbb{I}_{A_i}$, with n in $\mathbb{N}$ and, for all i in $\{1, \ldots, n\}$, a_i in $\mathbb{R}$ and A_i in $\mathscr{F}_u$. In this case, the integral expression reduces to

$$E_P(f \,|\, X_u = x_u) = \sum_{i=1}^{n} a_i P(A_i \,|\, X_u = x_u). \qquad (4)$$

For unconditional expectations, we have that

$$E(\cdot) := E_P(\cdot \,|\, \Omega) = E_P(\cdot \,|\, X_\emptyset = x_\emptyset) = E_P(\cdot \,|\, X_0 = 0),$$

where the final equality holds due to (CP1). Therefore, in the remainder, we can restrict ourselves to expectations of the form $E_P(\cdot \,|\, X_u = x_u, X_t = x)$, as $E(\cdot)$ corresponds to the case $u = \emptyset$, $t = 0$ and $x = 0$.

3. The Poisson Process in Particular

We now turn to the most well-known counting process, namely the Poisson process. As explained in the Introduction, there are plenty of alternative characterisations of the Poisson process. The following definition turns out to capture all its essential properties in our framework.

Definition 5 *A Poisson process P is a counting process such that, for all t, Δ in $\mathbb{R}_{\geq 0}$, u in $\mathscr{U}_{<t}$, (x_u, x) in $\mathscr{X}_{u \cup t}$ and y in $\mathscr{X}$ with $y \geq x$,*

PP1. $P(X_{t+\Delta} = y \,|\, X_u = x_u, X_t = x) = P(X_{t+\Delta} = y \,|\, X_t = x)$;

PP2. $P(X_{t+\Delta} = y \,|\, X_t = x) = P(X_{t+\Delta} = y - x \,|\, X_t = 0)$;

PP3. $P(X_{t+\Delta} = y \,|\, X_t = x) = P(X_\Delta = y \,|\, X_0 = x)$.

The first condition (PP1) states that the Poisson process is Markovian, while conditions (PP2) and (PP3) state that the Poisson process is homogeneous. Note that—unlike many of the characterisations mentioned in the Introduction—we do *not* impose that the transition probabilities are Poisson distributed, nor do we impose some value for the "rate". It was already observed by Feller [4, Chapter XVII, Section 2, Footnote 4] and Khintchine [6, Sections 1 and 2] that assuming—their version of—(PP1)–(PP3) is sufficient to obtain the Poisson process. Our results basically extend these characterisations to our framework for counting processes using coherent conditional probabilities.

First and foremost, we obtain that the transition probabilities are Poisson distributed, hence explaining the name of the process.

Theorem 6 *Consider a Poisson process P. Then there is a rate λ in $\mathbb{R}_{\geq 0}$ such that, for all t, Δ in $\mathbb{R}_{\geq 0}$, u in $\mathscr{U}_{<t}$, (x_u, x) in $\mathscr{X}_{u \cup t}$ and y in $\mathscr{X}$,*

$$P(X_{t+\Delta} = y \,|\, X_u = x_u, X_t = x)$$
$$= \begin{cases} \psi_{\lambda\Delta}(y - x) & \text{if } y \geq x, \\ 0 & \text{otherwise,} \end{cases} \qquad (5)$$

where $\psi_{\lambda\Delta}$ is the Poisson distribution with parameter $\lambda\Delta$, defined for all k in $\mathbb{Z}_{\geq 0}$ as

$$\psi_{\lambda\Delta}(k) := e^{-\lambda\Delta} \frac{(\lambda\Delta)^k}{k!}.$$

Conversely, for every λ in $\mathbb{R}_{\geq 0}$, there is a unique coherent conditional probability P on $\mathscr{D}_{\mathrm{CP}}$ that satisfies (CP1) *and Equation* (5)*, and this P is a Poisson process.*

Theorem 6 might seem somewhat trivial, but its proof is surprisingly lengthy. Note that it also establishes that any rate λ gives rise to a unique Poisson process, so in the remainder we can talk of *the* Poisson process with rate λ. Finally, it has the following obvious corollary.

Corollary 7 *Consider a Poisson process P. Then there is a rate λ in $\mathbb{R}_{\geq 0}$ such that, for all t in $\mathbb{R}_{\geq 0}$, u in $\mathscr{U}_{<t}$ and (x_u, x) in $\mathscr{X}_{u \cup t}$,*

$$\lim_{\Delta \to 0^+} \frac{P(X_{t+\Delta} = x + 1 \,|\, X_u = x_u, X_t = x)}{\Delta} = \lambda \qquad (6)$$

and, if $t > 0$,

$$\lim_{\Delta \to 0^+} \frac{P(X_t = x + 1 \,|\, X_u = x_u, X_{t-\Delta} = x)}{\Delta} = \lambda. \qquad (7)$$

We end our discussion of Poisson processes with the following result, which actually is a—not entirely immediate—consequence of Theorem 15 further on.

Theorem 8 *Consider a counting process P. If there is a rate λ in $\mathbb{R}_{\geq 0}$ such that P satisfies Equations* (6) *and* (7)*, then P is the Poisson process with rate λ.*

4. Sets of Counting Processes

Instead of considering a single counting process, we now study *sets* of counting processes. With any subset $\mathscr{P}$ of $\mathbb{P}$, we associate a *lower expectation*

$$\underline{E}_{\mathscr{P}}(\cdot \,|\, \cdot) := \inf\{E_P(\cdot \,|\, \cdot) : P \in \mathscr{P}\} \qquad (8)$$

and, similarly, an *upper expectation*

$$\overline{E}_{\mathscr{P}}(\cdot \,|\, \cdot) := \sup\{E_P(\cdot \,|\, \cdot) : P \in \mathscr{P}\}. \qquad (9)$$

Since the expectation E_P associated with any counting process P in $\mathscr{P}$ has domain $\mathscr{G}$, $\underline{E}_\mathscr{P}$ and $\overline{E}_\mathscr{P}$ are well-defined on the same domain $\mathscr{G}$. Observe that for any $(f, X_u = x_u)$ in $\mathscr{G}$ such that f is bounded, the lower and upper expectations are conjugate in the sense that $\overline{E}_\mathscr{P}(f \mid X_u = x_u) = -\underline{E}_\mathscr{P}(-f \mid X_u = x_u)$. Therefore, it suffices to study one of the two if only considering bounded functions; we will focus on lower expectations in the remainder.

4.1. The Obvious Imprecise Poisson Process

From here on, we consider a closed interval $\Lambda := [\underline{\lambda}, \overline{\lambda}] \subset \mathbb{R}_{\geq 0}$ of rates instead of a single value for the rate λ. In order not to unnecessarily repeat ourselves, we fix one rate interval Λ that we use throughout the remainder. Due to Theorem 6, there is one obvious set of counting processes that is entirely characterised by this rate interval Λ: the set

$$\mathbb{P}_\Lambda^\star := \{P_\lambda : \lambda \in \Lambda\}$$

that consists of all Poisson processes with rate in this interval, where P_λ denotes the Poisson process with rate λ.

The lower and upper expectation associated with this set $\mathbb{P}_\Lambda^\star$ according to Equations (8) and (9) are denoted by $\underline{E}_\Lambda^\star$ and $\overline{E}_\Lambda^\star$, respectively. It is clear that by construction, determining $\underline{E}_\Lambda^\star(f \mid X_u = x_u)$ and/or $\overline{E}_\Lambda^\star(f \mid X_u = x_u)$ boils down to solving a one-parameter optimisation problem: one has to minimise and/or maximise $E_{P_\lambda}(f \mid X_u = x_u)$—the conditional expectation of f with respect to the Poisson process with rate λ—with respect to all values of λ in the rate interval Λ. For some specific functions f, see for example Proposition 16 further on, this one-parameter optimisation problem can be solved analytically. For more involved functions, the optimisation problem has to be solved numerically, for instance by evaluating $E_{P_\lambda}(f \mid X_u = x_u)$ over a (sufficiently fine) grid of values of λ in the rate interval Λ, where $E_{P_\lambda}(f \mid X_u = x_u)$ might also have to be numerically approximated.

4.2. A More Involved Imprecise Poisson Process

A second set of counting processes characterised by the rate interval Λ is inspired by Theorem 8. This theorem suggests that the dynamics of a counting process are captured by the rate—that is, the limit expressions in Equations (6) and (7) of Corollary 7. Essential to our second characterisation is the notion of consistency.

Definition 9 *A counting process P is* consistent *with the rate interval Λ, denoted by $P \sim \Lambda$, if for all t in $\mathbb{R}_{\geq 0}$, u in $\mathscr{U}_{<t}$ and (x_u, x) in $\mathscr{X}_{u \cup t}$,*

$$\underline{\lambda} \leq \liminf_{\Delta \to 0^+} \frac{P(X_{t+\Delta} = x+1 \mid X_u = x_u, X_t = x)}{\Delta}$$
$$\leq \limsup_{\Delta \to 0^+} \frac{P(X_{t+\Delta} = x+1 \mid X_u = x_u, X_t = x)}{\Delta} \leq \overline{\lambda} \quad (10)$$

and, if $t > 0$,

$$\underline{\lambda} \leq \liminf_{\Delta \to 0^+} \frac{P(X_t = x+1 \mid X_u = x_u, X_{t-\Delta} = x)}{\Delta}$$
$$\leq \limsup_{\Delta \to 0^+} \frac{P(X_t = x+1 \mid X_u = x_u, X_{t-\Delta} = x)}{\Delta} \leq \overline{\lambda}. \quad (11)$$

We let

$$\mathbb{P}_\Lambda := \{P \in \mathbb{P} : P \sim \Lambda\}$$

denote the set of all *counting* processes that are consistent with the rate interval Λ. Observe that, as every Poisson process is a counting process,

$$\mathbb{P}_\Lambda^\star \subseteq \mathbb{P}_\Lambda. \quad (12)$$

It is essential to realise that $\mathbb{P}_\Lambda^\star$ is *not* equal to $\mathbb{P}_\Lambda$, at least not in general. Indeed, the set $\mathbb{P}_\Lambda$ will contain counting processes that have much more exotic dynamics than Poisson processes, in the sense that they need not be Markovian nor homogeneous. However, if Λ is equal to the degenerate interval $[\lambda, \lambda]$, then it follows from Theorem 8 that

$$\mathbb{P}_\Lambda^\star = \mathbb{P}_\Lambda = \{P_\lambda\}, \quad (13)$$

where P_λ is the Poisson process with rate λ, as before. Therefore, both $\mathbb{P}_\Lambda$ and $\mathbb{P}_\Lambda^\star$ are proper generalisations of the Poisson process.

We let $\underline{E}_\Lambda$ and $\overline{E}_\Lambda$ denote the lower and upper expectations associated with the set $\mathbb{P}_\Lambda$ according to Equations (8) and (9). It is an immediate consequence of Equations (8), (9) and (12) that

$$\underline{E}_\Lambda(\cdot \mid \cdot) \leq \underline{E}_\Lambda^\star(\cdot \mid \cdot) \leq \overline{E}_\Lambda^\star(\cdot \mid \cdot) \leq \overline{E}_\Lambda(\cdot \mid \cdot). \quad (14)$$

The remainder of this contribution is concerned with computing these lower and upper expectations for a specific type of functions, with a particular focus on the outer ones.

We end this section by mentioning that $\mathbb{P}_\Lambda^\star$ and $\mathbb{P}_\Lambda$ are not the only two sets of counting processes that are of potential interest, but they are—to some extent—the two most extreme sets. One set of counting process that lies in between the two is that of the time-inhomogeneous Poisson processes—see for instance [6, Section 5] or [11, Section 2.4]—that are consistent with the rate interval Λ. In order not to unnecessarily complicate our exposition, we have chosen to limit ourselves to the two extreme cases.

5. The Poisson Generator and Its Corresponding Semi-Group

Our method for computing lower expectations is based on the method used in the theory of imprecise continuous-time Markov chains [8]. Essential to this method of Krak et al. [8] is a semi-group of "lower transition operators" that is generated by a "lower transition rate operator". In

Section 5.2, we extend their method for generating this semi-group to a countably infinite state space, be it only for one specific type of lower transition rate operator. First, however, we introduce some necessary concepts and terminology.

5.1. Functions, Operators and Norms

Consider some non-empty ordered set $\mathscr{Y}$ that is at most countably infinite, and let $\mathscr{L}(\mathscr{Y})$ be the set of all bounded real-valued functions on $\mathscr{Y}$. Observe that $\mathscr{L}(\mathscr{Y})$ is clearly a vector space. Even more, it is well-known that this vector space is complete under the supremum norm

$$\|f\| := \sup\{|f(x)| : x \in \mathscr{Y}\} \quad \text{for all } f \in \mathscr{L}(\mathscr{Y}).$$

A *transformation* is any operator $A : \mathscr{L}(\mathscr{Y}) \to \mathscr{L}(\mathscr{Y})$. Such a transformation A is *non-negatively homogeneous* if, for all f in $\mathscr{L}(\mathscr{Y})$ and γ in $\mathbb{R}_{\geq 0}$, $A(\gamma f) = \gamma A f$. The supremum norm induces an operator norm for non-negatively homogeneous transformations A:

$$\|A\| := \sup\{\|Af\| : f \in \mathscr{L}(\mathscr{Y}), \|f\| = 1\};$$

see [2] for a proof that this is indeed a norm. An important non-negatively homogeneous transformation is the *identity map I* that maps any f in $\mathscr{L}(\mathscr{Y})$ to itself.

5.2. The Poisson Generator

A non-negatively homogeneous transformation that will be essential in the remainder is the *Poisson generator $\underline{Q} : \mathscr{L}(\mathscr{X}) \to \mathscr{L}(\mathscr{X})$* associated with the rate interval Λ, defined for all f in $\mathscr{L}(\mathscr{X})$ and x in $\mathscr{X}$ as

$$[\underline{Q}f](x) := \min\{\lambda f(x+1) - \lambda f(x) : \lambda \in [\underline{\lambda}, \overline{\lambda}]\}.$$

Fix any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$. If $t < s$, then we let $\mathscr{U}_{[t,s]}$ denote the set of all non-empty and increasing sequences of time points $t_0, \ldots, t_n$ that start with $t_0 = t$ and end with $t_n = s$. For any sequence u in this set $\mathscr{U}_{[t,s]}$, we let

$$\Phi_u := \prod_{i=1}^{n}(I + \Delta_i \underline{Q}), \tag{15}$$

where for any i in $\{1, \ldots, n\}$, we denote the difference between the consecutive time points t_i and t_{i-1} by $\Delta_i := t_i - t_{i-1}$. In the remainder, we let $\sigma(u) := \max\{\Delta_i : i \in \{1, \ldots, n\}\}$ be the largest of these time differences. If $t = s$, then we let $\mathscr{U}_{[t,s]} := \{t\}$, $\sigma(t) := 0$ and $\Phi_t := I$.

The Poisson generator $\underline{Q}$ generates a family of transformations, as is evident from the following result. This result is very similar to [8, Corollary 7.11], which establishes an analogous result for imprecise Markov chains with finite state spaces; it should therefore not come as a surprise that their proofs are largely similar as well.

Theorem 10 *Fix any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$. For any sequence $\{u_i\}_{i \in \mathbb{N}}$ in $\mathscr{U}_{[t,s]}$ such that $\lim_{i \to +\infty} \sigma(u_i) = 0$, the corresponding sequence $\{\Phi_{u_i}\}_{i \in \mathbb{N}}$ converges to a unique non-negatively homogeneous transformation that does not depend on the chosen sequence $\{u_i\}_{i \in \mathbb{N}}$.*

For any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, Theorem 10 allows us to define the non-negatively homogeneous transformation

$$\underline{T}_t^s := \lim_{\sigma(u) \to 0}\{\Phi_u : u \in \mathscr{U}_{[t,s]}\}, \tag{16}$$

where this unconventional notation for the limit denotes the unique limit mentioned in Theorem 10. The family of transformations thus defined has some very interesting properties: in the Appendix, we prove that for any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, f, g in $\mathscr{L}(\mathscr{X})$ and γ in $\mathbb{R}_{\geq 0}$,

SG1. $\underline{T}_t^s(\gamma f) = \gamma \underline{T}_t^s f$;

SG2. $\underline{T}_t^s(f + g) \geq \underline{T}_t^s f + \underline{T}_t^s g$;

SG3. $\underline{T}_t^s f \geq \inf f$.

We furthermore prove that this family forms a time-homogeneous semi-group, in the sense that

SG4. $\underline{T}_t^t = I$;

SG5. $\underline{T}_t^s = \underline{T}_t^r \underline{T}_r^s$ for all r in $\mathbb{R}_{\geq 0}$ with $t \leq r \leq s$;

SG6. $\underline{T}_t^s = \underline{T}_0^{s-t}$.

While the induced transformation $\underline{T}_t^s$ is interesting in its own right, we will be mainly interested in (a single component of) the image $\underline{T}_t^s f$ of some bounded function f. Therefore, for any x in $\mathscr{X}$ and t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, we define the operator $\underline{P}_t^s(\cdot \mid x) : \mathscr{L}(\mathscr{X}) \to \mathbb{R}$ as

$$\underline{P}_t^s(f \mid x) := [\underline{T}_t^s f](x) \qquad \text{for all } f \in \mathscr{L}(\mathscr{X}).$$

The following follows immediately from (SG1)–(SG3).

Corollary 11 *For any x in $\mathscr{X}$ and t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, $\underline{P}_t^s(\cdot \mid x)$ is a coherent lower prevision in the sense of [13, Definition 4.10].*

In the remainder, we let $\overline{P}_t^s(\cdot \mid x) := -\underline{P}_t^s(-\cdot \mid x)$ be the conjugate coherent upper prevision of the coherent lower prevision $\underline{P}_t^s(\cdot \mid x)$.

5.3. The Reduced Poisson Generator

Fix any $\underline{x}, \overline{x}$ in $\mathscr{X}$ such that $\underline{x} \leq \overline{x}$, and let

$$\chi := \{x \in \mathscr{X} : \underline{x} \leq x \leq \overline{x}\}.$$

We define the *reduced Poisson generator $\underline{Q}^\chi : \mathscr{L}(\chi) \to \mathscr{L}(\chi)$* for all g in $\mathscr{L}(\chi)$ and x in χ as

$$[\underline{Q}^\chi g](x) := \begin{cases} \min_{\lambda \in [\underline{\lambda}, \overline{\lambda}]} (\lambda g(x+1) - \lambda g(x)) & \text{if } \underline{x} \leq x < \overline{x}, \\ 0 & \text{if } x = \overline{x}. \end{cases}$$

In the Appendix, we verify that this reduced Poisson generator $\underline{Q}^\chi$ is a *lower transition rate operator* in the sense of [8, Definition 7.2]. As outlined in [8, Section 7], this lower transition rate operator generates a family of transformations as well. For any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$ and any u in $\mathscr{U}_{[t,s]}$, we let

$$\Phi_u^\chi := \prod_{i=1}^n (I + \Delta_i \underline{Q}^\chi).$$

Note the similarity between the equation above and Equation (15). Because $\underline{Q}^\chi$ is a lower transition rate operator, it follows from [8, Corollary 7.11]—a result similar to Theorem 10—that the transformation

$$\underline{T}_{t,s}^\chi := \lim_{\sigma(u) \to 0} \{\Phi_u^\chi : u \in \mathscr{U}_{[t,s]}\} \qquad (17)$$

is non-negatively homogeneous. The limit in this definition is to be interpreted as the limit in Equation (16): it does not depend on the actual sequence $\{u_i\}_{i \in \mathbb{N}}$ as long as $\lim_{i \to +\infty} \sigma(u_i) = 0$. Unsurprisingly, Krak et al. [8] show that this family of transformations $\underline{T}_{t,s}^\chi$ also satisfies (SG1)–(SG6). Observe that Equation (17) suggests a method to evaluate $\underline{T}_{t,s}^\chi$ for some g in $\mathscr{L}(\chi)$: choose a sufficiently fine grid u, and compute $\Phi_u^\chi g$ via backwards recursion. There is much more to this approximation method than we can cover here; the interested reader is referred to [8, Section 8.2] and [3].

5.4. The Essential Case of Eventually Constant Functions

Our reason for introducing the restricted Poisson generator $\underline{Q}^\chi$ and its induced transformation $\underline{T}_{t,s}^\chi$ is because the latter can be used to compute $\underline{P}_t^s(f \mid x)$. Essential to our method are those functions f in $\mathscr{L}(\mathscr{X})$ that are *eventually constant*, in the sense that

$$(\exists \overline{x} \in \mathscr{X})(\forall x \in \mathscr{X}, x \geq \overline{x}) \, f(x) = f(\overline{x}).$$

In this case, we say that f is *constant starting from* $\overline{x}$. We collect all real-valued bounded functions f on $\mathscr{X}$ that are eventually constant in $\mathscr{L}^c(\mathscr{X})$.

Our next result establishes a link between $\underline{P}_t^s(\cdot \mid x)$ and $\underline{T}_{t,s}^\chi$ for eventually constant functions.

Proposition 12 *Fix some t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$ and some f in $\mathscr{L}^c(\mathscr{X})$ that is constant starting from $\overline{x}$. Choose some $\underline{x}$ in $\mathscr{X}$ with $\underline{x} \leq \overline{x}$, and let $\chi := \{x \in \mathscr{X} : \underline{x} \leq x \leq \overline{x}\}$. Then for any x in $\mathscr{X}$ with $x \geq \underline{x}$,*

$$\underline{P}_t^s(f \mid x) = [\underline{T}_t^s f](x) = \begin{cases} [\underline{T}_{t,s}^\chi f^\chi](x) & \text{if } x \leq \overline{x}, \\ f(\overline{x}) & \text{if } x \geq \overline{x}, \end{cases}$$

where f^χ is the restriction of f to χ.

Note that we are free to choose $\underline{x}$. If we are interested in $\underline{P}_t^s(f \mid x)$ for a specific value of x, then choosing $\underline{x} = \min\{\overline{x}, x\}$ is the optimal choice. However, if we are interested in $\underline{P}_t^s(f \mid x)$ for a finite range $R \subset \mathscr{X}$ of different x values, the obvious choice is $\underline{x} = \min(R \cup \{\overline{x}\})$ because we then only have to determine $\underline{T}_{t,s}^\chi f^\chi$ once!

A method to compute $\underline{P}_t^s(\cdot \mid x)$ for all bounded functions f follows from combining Proposition 12 with the following result.

Proposition 13 *For any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, f in $\mathscr{L}(\mathscr{X})$ and x in $\mathscr{X}$,*

$$\underline{P}_t^s(f \mid x) = \lim_{\overline{x} \to +\infty} \underline{P}_t^s(\mathbb{I}_{\leq \overline{x}} f + f(\overline{x}) \mathbb{I}_{>\overline{x}} \mid x),$$

where $\mathbb{I}_{\leq \overline{x}}$ and $\mathbb{I}_{>\overline{x}}$ are the indicators of $\{z \in \mathscr{X} : z \leq \overline{x}\}$ and $\{z \in \mathscr{X} : z > \overline{x}\}$, respectively.

Observe that $\mathbb{I}_{\leq \overline{x}} f + f(\overline{x}) \mathbb{I}_{>\overline{x}}$—with $\mathbb{I}_{\leq \overline{x}} f$ the point-wise multiplication of $\mathbb{I}_{\leq \overline{x}}$ and f—is constant starting from $\overline{x}$. Therefore, it follows from Proposition 12 that $\underline{P}_t^s(\mathbb{I}_{\leq \overline{x}} f + f(\overline{x}) \mathbb{I}_{>\overline{x}} \mid x) = [\underline{T}_{t,s}^\chi f^\chi](x)$, where f^χ is the restriction of f to χ. We can combine this observation and Proposition 13 to obtain a method to compute $\underline{P}_t^s(f \mid x)$ for any bounded function f: (i) choose some sufficiently large $\overline{x}$ and let $\chi := \{y \in \mathscr{X} : x \leq y \leq \overline{x}\}$; (ii) compute $\underline{P}_t^s(\mathbb{I}_{\leq \overline{x}} f + f(\overline{x}) \mathbb{I}_{>\overline{x}} \mid x) = [\underline{T}_{t,s}^\chi f^\chi](x)$, using one of the existing approximation methods mentioned at the end of Section 5.3; (iii) repeat (i)–(ii) for increasingly larger $\overline{x}$ until convergence is empirically observed.

6. Computing Lower Expectations of Functions on X_s

Let $\mathscr{K}_b(\mathscr{X})$ denote the set of all real-valued bounded-below functions on $\mathscr{X}$. With any f in $\mathscr{K}_b(\mathscr{X})$ and s in $\mathbb{R}_{\geq 0}$, we associate the real-valued bounded-below function

$$f(X_s) \colon \Omega \to \mathbb{R} \colon \omega \mapsto [f(X_s)](\omega) := f(\omega(s)).$$

In other words, and as suggested by our notation, $f(X_s)$ is the functional composition of f with the projector

$$X_s \colon \Omega \to \mathscr{X} \colon \omega \mapsto X_s(\omega) := \omega(s).$$

The (conditional) expectation of $f(X_s)$ exists for any counting process P, as is established by the following rather obvious result.

Lemma 14 *Consider some s in $\mathbb{R}_{\geq 0}$ and u in $\mathscr{U}$ with $\max u \leq s$. Then for any f in $\mathscr{K}_b(\mathscr{X})$, $f(X_s)$ is an $\mathscr{F}_u$-measurable function.*

In the remainder, we provide several methods for computing lower and upper expectations; first for those with respect to the consistent Poisson processes and second for those with respect to all consistent counting processes. For the latter, we first limit ourselves to bounded functions and subsequently move on to functions that are bounded-below.

6.1. With Respect to the Consistent Poisson Processes

Fix some rate λ in $\mathbb{R}_{\geq 0}$, and let P be the Poisson process with rate λ. It is essentially well-known—and a consequence of Theorem 6—that for any t,s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, u in $\mathscr{U}_{<t}$, (x_u,x) in $\mathscr{X}_{u \cup t}$ and f in $\mathscr{K}_b(\mathscr{X})$,

$$E_P(f(X_s)\,|\,X_u = x_u, X_t = x) = \sum_{y=x}^{+\infty} f(y)\,\psi_{\lambda(s-t)}(y-x). \quad (18)$$

Because of this expression, $\underline{E}_\Lambda^\star(f(X_s)\,|\,X_u = x_u, X_t = x)$ can be computed using the straightforward method that we already discussed in Section 4.1: (i) fix a finite grid of λ's in $\Lambda = [\underline{\lambda}, \overline{\lambda}]$, (ii) (numerically) evaluate the infinite sum in Equation (18) for each λ in this grid, and (iii) compute the minimum. In some specific cases, it is even possible to know beforehand for which λ this minimum will be achieved. For example, if f is monotone and bounded, or bounded below and non-decreasing, then as we will see in Propositions 16 and 17, it suffices to consider $\lambda = \underline{\lambda}$ or $\lambda = \overline{\lambda}$.

6.2. With Respect to the Consistent Counting Processes

Computing $\underline{E}_\Lambda(f(X_s)\,|\,X_u = x_u, X_t = x)$ is less straightforward, as in general this does not reduce to a one-parameter optimisation problem. Nevertheless, as we are about to show, the semi-group of Section 5 allows us to circumvent this issue. Our first result establishes a method to compute the lower—and hence also the upper—expectation of bounded functions.

Theorem 15   *For any t,s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, u in $\mathscr{U}_{<t}$, f in $\mathscr{L}(\mathscr{X})$ and (x_u,x) in $\mathscr{X}_{u \cup t}$,*

$$\underline{E}_\Lambda(f(X_s)\,|\,X_u = x_u, X_t = x) = \underline{P}_t^s(f\,|\,x).$$

Indeed, because of this result, we can use the method that was introduced at the end of Section 5.4 to compute the lower expectation of f.

For the special case of monotone bounded functions, we obtain an even stronger result.

Proposition 16   *Fix any t,s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, u in $\mathscr{U}_{<t}$, (x_u,x) in $\mathscr{X}_{u \cup t}$ and f in $\mathscr{L}(\mathscr{X})$. If f is monotone, then*

$$\begin{aligned}
\underline{E}_\Lambda(f(X_s)\,&|\,X_u = x_u, X_t = x) \\
&= \underline{E}_\Lambda^\star(f(X_s)\,|\,X_u = x_u, X_t = x) \\
&= E_{P_\lambda}(f(X_s)\,|\,X_u = x_u, X_t = x),
\end{aligned}$$

where P_λ is the Poisson process with rate $\lambda = \underline{\lambda}$ if f is non-decreasing and rate $\lambda = \overline{\lambda}$ if f is non-increasing.

Almost everything has now been set up to consider a general real-valued bounded below function of X_s. An essential intermediary step is an extension of Proposition 16.

Proposition 17   *Fix any t,s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, u in $\mathscr{U}_{<t}$ and (x_u,x) in $\mathscr{X}_{u \cup t}$. Then for any f in $\mathscr{K}_b(\mathscr{X})$ that is non-decreasing,*

$$\begin{aligned}
\underline{E}_\Lambda(f(X_s)\,&|\,X_u = x_u, X_t = x) \\
&= \underline{E}_\Lambda^\star(f(X_s)\,|\,X_u = x_u, X_t = x) \\
&= E_{P_{\underline{\lambda}}}(f(X_s)\,|\,X_u = x_u, X_t = x)
\end{aligned}$$

and

$$\begin{aligned}
\overline{E}_\Lambda(f(X_s)\,&|\,X_u = x_u, X_t = x) \\
&= \overline{E}_\Lambda^\star(f(X_s)\,|\,X_u = x_u, X_t = x) \\
&= E_{P_{\overline{\lambda}}}(f(X_s)\,|\,X_u = x_u, X_t = x),
\end{aligned}$$

where $P_{\underline{\lambda}}$ and $P_{\overline{\lambda}}$ are the Poisson processes with rates $\underline{\lambda}$ and $\overline{\lambda}$, respectively.

As an immediate corollary of Proposition 17, we obtain an interpretation for the rate interval Λ: its bounds provide tight lower and upper bounds on the expected number of Poisson-events in any time period.

Corollary 18   *Fix any t,s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, u in $\mathscr{U}_{<t}$ and (x_u,x) in $\mathscr{X}_{u \cup t}$. Then*

$$\underline{E}_\Lambda(X_s\,|\,X_u = x_u, X_t = x) = x + \underline{\lambda}(s-t)$$

and

$$\overline{E}_\Lambda(X_s\,|\,X_u = x_u, X_t = x) = x + \overline{\lambda}(s-t),$$

and similarly for $\underline{E}_\Lambda^\star$ and $\overline{E}_\Lambda^\star$.

A more important consequence of Proposition 17 is the following result, which can be regarded as an extension of (the combination of) Proposition 13 and Theorem 15.

Theorem 19   *Fix any t,s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, u in $\mathscr{U}_{<t}$, (x_u,x) in $\mathscr{X}_{u \cup t}$ and f in $\mathscr{K}_b(\mathscr{X})$. If*

$$\sum_{y=x}^{+\infty} f_{\max}(y)\,\psi_{\overline{\lambda}(s-t)}(y-x) < +\infty,$$

where $f_{\max}$ in $\mathscr{K}_b(\mathscr{X})$ is defined for all y in $\mathscr{X}$ as

$$f_{\max}(y) := \max\{f(z)\colon z \in \mathscr{X}, z \leq y\},$$

then

$$\underline{E}_\Lambda(f(X_s)\,|\,X_u = x_u, X_t = x) = \lim_{\overline{x} \to +\infty} \underline{P}_t^s(\mathbb{I}_{\leq \overline{x}}f + f(\overline{x})\mathbb{I}_{>\overline{x}}\,|\,x),$$

$$\overline{E}_\Lambda(f(X_s)\,|\,X_u = x_u, X_t = x) = \lim_{\overline{x} \to +\infty} \overline{P}_t^s(\mathbb{I}_{\leq \overline{x}}f + f(\overline{x})\mathbb{I}_{>\overline{x}}\,|\,x),$$

where the two limits are finite.

Because of this result, we can compute the lower and upper expectation using the same method as before. Note that it makes no difference that f is no longer bounded; the method still works because $\mathbb{I}_{\leq \overline{x}}f + f(\overline{x})\mathbb{I}_{>\overline{x}}$ is bounded.
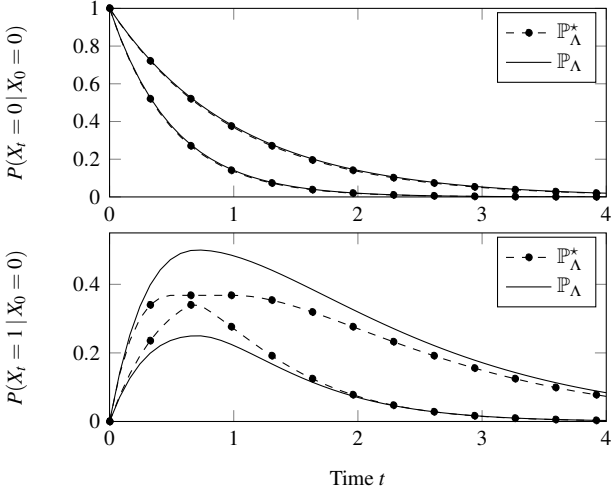
Figure 1: Bounds on transition probabilities as a function of t for the rate interval $\Lambda = [1,2]$.

6.3. A Numerical Example

We end this section with a basic numerical example. We determine tight lower and upper bounds on

$$P(X_t = x \mid X_0 = 0) = E_P(\mathbb{I}_x(X_t) \mid X_0 = 0),$$

with x equal to 0 or 1. We use the methods outlined in Sections 6.1 and 6.2 to compute lower and upper bounds with respect to the sets $\mathbb{P}_\Lambda^\star$ and $\mathbb{P}_\Lambda$ for $\Lambda = [1,2]$. The resulting bounds are depicted in Figure 1. Observe that for $x = 0$, the bounds with respect to $\mathbb{P}_\Lambda^\star$ and $\mathbb{P}_\Lambda$ are equal, as is to be expected due to Proposition 16 because $\mathbb{I}_x$ is monotone for $x = 0$. For $x = 1$, $\mathbb{I}_x$ is not monotone and the bounds with respect to $\mathbb{P}_\Lambda^\star$ are clearly *not* equal to those with respect to $\mathbb{P}_\Lambda$.

7. Justification for the Term Imprecise Poisson Process

Until now, we have provided little justification for why we call both $\mathbb{P}_\Lambda^\star$ and $\mathbb{P}_\Lambda$ imprecise Poisson processes. In Section 4.2, we already briefly mentioned that the two sets are proper generalisations of the Poisson process: if the rate interval Λ is degenerate, meaning that $\underline{\lambda} = \lambda = \overline{\lambda}$, then both sets reduce to the singleton containing the Poisson process with rate λ. Another argument for referring to $\mathbb{P}_\Lambda^\star$ and $\mathbb{P}_\Lambda$ as imprecise Poisson processes concerns the (tight lower and upper bounds on the) expected number of Poisson events in a time period of length Δ. For a Poisson process, it is well-known that this expectation is equal to $\Delta\lambda$, and we know from Corollary 18 that the corresponding lower and upper expectations are equal to $\Delta\underline{\lambda}$ and $\Delta\overline{\lambda}$, respectively.

We end this section with our strongest argument for using the term imprecise Poisson process to refer to both $\mathbb{P}_\Lambda^\star$ and $\mathbb{P}_\Lambda$. The following result establishes that the corresponding lower expectations $\underline{E}_\Lambda^\star$ and $\underline{E}_\Lambda$—and, due to conjugacy, also the corresponding upper expectations $\overline{E}_\Lambda^\star$ and $\overline{E}_\Lambda$—satisfy imprecise generalisations of (CP1), (CP2) and (PP1)–(PP3), which are the defining properties of a Poisson process.

Proposition 20 *For all t, Δ in $\mathbb{R}_{\geq 0}$, u in $\mathcal{U}_{<t}$, (x_u, x) in $\mathcal{X}_{u \cup t}$ and f in $\mathcal{L}(\mathcal{X})$,*

(i) $\underline{E}_\Lambda(f(X_0)) = f(0)$;

(ii)
$$\lim_{\Delta \to 0^+} \frac{\underline{E}_\Lambda(\mathbb{I}_{(X_{t+\Delta} \geq x+2)} \mid X_u = x_u, X_t = x)}{\Delta} = 0$$

and, if $t > 0$,

$$\lim_{\Delta \to 0^+} \frac{\underline{E}_\Lambda(\mathbb{I}_{(X_t \geq x+2)} \mid X_u = x_u, X_{t-\Delta} = x)}{\Delta} = 0;$$

(iii) $\underline{E}_\Lambda(f(X_{t+\Delta}) \mid X_u = x_u, X_t = x) = \underline{E}_\Lambda(f(X_{t+\Delta}) \mid X_t = x)$;

(iv) $\underline{E}_\Lambda(f(X_{t+\Delta}) \mid X_t = x) = \underline{E}_\Lambda(f'_x(X_{t+\Delta}) \mid X_t = 0)$;

(v) $\underline{E}_\Lambda(f(X_{t+\Delta}) \mid X_t = x) = \underline{E}_\Lambda(f(X_\Delta) \mid X_0 = x)$;

with $f'_x \colon \mathcal{X} \to \mathbb{R} \colon z \mapsto f'_x(z) := f(x+z)$. The same equalities also hold for $\underline{E}_\Lambda^\star$.

8. Conclusion

In this contribution, we proposed two generalisations of the Poisson process in the form of two sets of counting processes: the set $\mathbb{P}_\Lambda^\star$ of all Poisson processes with rate λ in the rate interval Λ, and the set $\mathbb{P}_\Lambda$ of all counting process that are consistent with the rate interval Λ. We argued why both of these sets can be seen as proper generalisations of the Poisson process. First and foremost, for a degenerate rate interval they both reduce to the singleton containing the Poisson process with this rate. Second, the lower and upper expectations with respect to both sets satisfy imprecise generalisations of (CP1), (CP2) and (PP1)–(PP3), the defining properties of a Poisson process. We also presented several methods for computing lower and upper expectations for functions that depend on the number of occurred Poisson-events at a single time point.

We end with two suggestions for future research. An obvious open question is whether we can efficiently compute lower and upper expectations for functions that depend on the number of occurred Poisson-events at multiple points in time. Based on similar results of Krak et al. [8] for imprecise continuous-time Markov chains with a finite state space, we strongly believe that this will be the case for $\mathbb{P}_\Lambda$ but not for $\mathbb{P}_\Lambda^\star$, whence providing a practical argument in favour of the former. A perhaps slightly less obvious open

question is whether Theorem 6 and Corollary 7 can be generalised to sets of counting processes, in the sense that we can infer the existence of a rate interval rather than specify one, by imposing appropriate conditions on the set of counting processes, including the imprecise generalisations of (CP1), (CP2) and (PP1)–(PP3).

Acknowledgments

References

[1] Daryl J. Daley and David Vere-Jones. *An Introduction to the Theory of Point Processes*, volume I: Elementary Theory and Methods. Springer, 2003.

[2] Jasper De Bock. The limit behaviour of imprecise continuous-time Markov chains. *Journal of Nonlinear Science*, 27(1):159–196, 2017. doi: 10.1007/s00332-016-9328-3.

[3] Alexander Erreygers and Jasper De Bock. Imprecise continuous-time Markov chains: Efficient computational methods with guaranteed error bounds. In *Proceedings of ISIPTA'17*, pages 145–156. PMLR, 2017.

[4] William Feller. *An Introduction to Probability Theory and Its Applications*, volume 1. Wiley, 1968.

[5] Samuel Karlin and Howard M. Taylor. *A first course in stochastic processes*. Academic Press, 1975.

[6] Aleksandr Yakovlevich Khintchine. *Mathematical methods in the theory of queueing*. Charles Griffin & Company, 1960.

[7] John F.C. Kingman. *Poisson Processes*. Clarendon Press, 1993.

[8] Thomas Krak, Jasper De Bock, and Arno Siebes. Imprecise continuous-time Markov chains. *International Journal of Approximate Reasoning*, 88:452–528, 2017. doi: 10.1016/j.ijar.2017.06.012.

[9] James R. Norris. *Markov chains*. Cambridge University Press, 1997.

[10] Eugenio Regazzini. Finitely additive conditional probabilities. *Rendiconti del Seminario Matematico e Fisico di Milano*, 55(1):69–89, 1985. doi: 10.1007/BF02924866.

[11] Sheldon M. Ross. *Stochastic Processes*. Wiley, 1996.

[12] Ke-iti Sato. *Lévy processes and infinitely divisible distributions*. Cambridge University Press, 1999.

[13] Matthias C.M. Troffaes and Gert de Cooman. *Lower Previsions*. Wiley, 2014.

[14] Shinzo Watanabe. On discontinuous additive functionals and Lévy measures of a Markov process. *Japanese journal of mathematics: transactions and abstracts*, 34:53–70, 1964. doi: 10.4099/jjm1924.34.0_53.

Appendix

In this appendix, we will not entirely follow the same order as we did in the main text. Our reason for doing so is that to prove the results in Section 3, we need some results that are very much related to the transformations that we introduce in Section 5. Therefore, we have chosen to start off this appendix with some general results regarding transformations.

Appendix A. Some Preliminary Results Regarding Transformations

Throughout this appendix, and as mentioned in Section 5.1, we let $\mathscr{Y}$ be any non-empty set that is at most countably finite; furthermore, we assume that $\mathscr{Y}$ is endowed with a total order "$\leq$".

A.1. General Non-Negatively Homogeneous Transformations

We start of with some essential properties of non-negatively homogeneous transformations.

Lemma 21 *Consider two non-negatively homogeneous transformations A and B on $\mathscr{L}(\mathscr{Y})$. Then*

NH1. *$A + B$ is non-negatively homogeneous;*

NH2. *μA is non-negatively homogeneous for any μ in $\mathbb{R}$;*

NH3. *AB is non-negatively homogeneous;*

NH4. *$\|Af\| \leq \|A\|\|f\|$ for any f in $\mathscr{L}(\mathscr{Y})$;*

NH5. *$\|AB\| \leq \|A\|\|B\|$.*

Proof The proof of (NH1)–(NH3) is a matter of straightforward verification. We therefore move on to proving (NH4). Observe first that if $\|f\| = 0$, then $f = 0$ and it follows from the non-negative homogeneity of A that $Af = A(0f) = 0(Af) = 0$. Therefore, $\|Af\| = 0$; hence the stated is true. Next, we assume that $\|f\| > 0$. Then

$$Af = A\left(\frac{\|f\|f}{\|f\|}\right) = \|f\|A\left(\frac{f}{\|f\|}\right) = \|f\|Af',$$

where we let $f' := f/\|f\|$. Note that $\|f'\| = \|f\|/\|f\| = 1$. Consequently,

$$
\begin{aligned}
\|Af\| &= \sup\{|[Af](x)|\colon x \in \mathscr{Y}\} \\
&= \sup\{|\|f\|[Af'](x)|\colon x \in \mathscr{Y}\} = \|f\|\sup\{|[Af'](x)|\colon x \in \mathscr{Y}\} = \|f\|\|Af'\| \\
&\leq \|A\|\|f\|,
\end{aligned}
$$

where the final inequality holds because $\|f'\| = 1$ implies that $\|Af'\| \leq \|A\|$.

Finally, we prove (NH5). To that end, we observe that

$$
\begin{aligned}
\|AB\| &= \sup\{\|ABg\|\colon g \in \mathscr{L}(\mathscr{Y}), \|g\| = 1\} \\
&\leq \sup\{\|A\|\|Bg\|\colon g \in \mathscr{L}(\mathscr{Y}), \|g\| = 1\} \\
&= \|A\|\sup\{\|Bg\|\colon g \in \mathscr{L}(\mathscr{Y}), \|g\| = 1\} = \|A\|\|B\|,
\end{aligned}
$$

where the inequality follows from (NH4). ∎

Time and time again, we will consider transformations on $\mathscr{L}(\mathscr{Y})$ that are constructed using a finite succession of the operations (NH1)–(NH3). For instance, we will often be interested in the (norm of the) "difference" $\|A_1\cdots A_k - B_1\cdots B_\ell\|$ between the two transformations $A := A_1\cdots A_k$ and $B := B_1\cdots B_\ell$, where $A_1, \ldots A_k$ and $B_1, \ldots, B_\ell$ are non-negatively homogeneous transformations. That A is a non-negatively homogeneous transformation follows from repeated application of (NH3), and similarly for B. Furthermore, $A - B$ is a non-negatively homogeneous transformation due to (NH2) with $\mu = -1$ and (NH1), so the norm of $A - B$ is indeed well-defined. In order not to needlessly repeat ourselves, we will usually refrain from explicitly mentioning that the operator constructed by a finite succession of the operations (NH1)–(NH3) is a non-negatively homogeneous transformation.

Next, we verify that our norm for non-negatively homogeneous transformations on $\mathscr{L}(\mathscr{Y})$ satisfies the three conditions of a norm: it should (i) be absolutely homogeneous, (ii) be sub-additive; and (iii) separate points. We here simply repeat the arguments of De Bock [2], who restricted himself to finite sets $\mathscr{Y}$.

(i) Let A be some non-negatively homogeneous operator, and fix some real number μ. Observe that μA is a non-negatively homogeneous transformation by (NH2). Furthermore, some straightforward manipulations yield that

$$\|\mu A\| = \sup\{\|\mu A f\| \colon f \in \mathscr{L}(\mathscr{Y}), \|f\| = 1\} = \sup\{|\mu|\|A f\| \colon f \in \mathscr{L}(\mathscr{Y}), \|f\| = 1\}$$
$$= |\mu| \sup\{\|A f\| \colon f \in \mathscr{L}(\mathscr{Y}), \|f\| = 1\} = |\mu|\|A\|,$$

where the second equality holds due to the absolute homogeneity of the supremum norm.

(ii) Let A and B be two non-negatively homogeneous transformations. Recall from (NH1) that $A + B$ is also a non-negatively homogeneous transformation. It now follows from the sub-additivity of the supremum norm that

$$\|A + B\| = \sup\{\|(A + B)f\| \colon f \in \mathscr{L}(\mathscr{Y}), \|f\| = 1\} = \sup\{\|A f + B f\| \colon f \in \mathscr{L}(\mathscr{Y}), \|f\| = 1\}$$
$$\leq \sup\{\|A f\| + \|B f\| \colon f \in \mathscr{L}(\mathscr{Y}), \|f\| = 1\} \leq \sup\{\|A\| + \|B\| \colon f \in \mathscr{L}(\mathscr{Y}), \|f\| = 1\} = \|A\| + \|B\|,$$

where for the second inequality we have used (NH4).

(iii) Let A be a non-negatively homogeneous transformation. Recall from (NH4) that, for any f in $\mathscr{L}(\mathscr{Y})$,

$$0 \leq \|A f\| \leq \|A\|\|f\|.$$

Hence, if $\|A\| = 0$, then it follows from these inequalities that $\|A f\| = 0$ for all f in $\mathscr{L}(\mathscr{Y})$. From this, we conclude that $A f = 0$ for all f in $\mathscr{L}(\mathscr{Y})$, and so $A = 0$, because the supremum norm separates points.

A.2. Lower Counting Transformations

The first two types of non-negatively homogeneous transformations that will be essential in the remainder are lower transition transformations and lower counting transformations. The following definition is a straightforward generalisation (or modification) of the existing concept of a *lower transition operator* on a finite state space, see for instance [8, Definition 7.1].

Definition 22 *A lower transition transformation $\underline{T} \colon \mathscr{L}(\mathscr{Y}) \to \mathscr{L}(\mathscr{Y})$ is a transformation such that*

LT1. $\underline{T}(\gamma f) = \gamma \underline{T} f$, *for all f in $\mathscr{L}(\mathscr{Y})$ and γ in $\mathbb{R}_{\geq 0}$;*   *[non-negative homogeneity]*

LT2. $\underline{T}(f + g) \geq \underline{T} f + \underline{T} g$, *for all f, g in $\mathscr{L}(\mathscr{Y})$;*   *[super-additivity]*

LT3. $\underline{T} f \geq \inf f$, *for all f in $\mathscr{L}(\mathscr{Y})$.*   *[bound]*

A lower counting transformation $\underline{T}$ is a lower transition transformation with

LT4. $[\underline{T} f](x) = [\underline{T}(\mathbb{I}_{\geq x} f)](x)$, *for all f in $\mathscr{L}(\mathscr{Y})$ and x in $\mathscr{Y}$.*

Lower transition transformations have many interesting properties. We start with some basic ones.

Lemma 23 *Consider a lower transition transformation $\underline{T} \colon \mathscr{L}(\mathscr{Y}) \to \mathscr{L}(\mathscr{Y})$. Then*

LT5. $\inf f \leq \underline{T} f \leq -\underline{T}(-f) \leq \sup f$ *for all f in $\mathscr{L}(\mathscr{X})$;*

LT6. $\underline{T} \mu = \mu$ *for all μ in $\mathbb{R}$;*

LT7. $\underline{T}(f + \mu) = \underline{T} f + \mu$ *for all f in $\mathscr{L}(\mathscr{Y})$ and all μ in $\mathbb{R}$;*

LT8. $\underline{T} f \leq \underline{T} g$ *for all f, g in $\mathscr{L}(\mathscr{Y})$ such that $f \leq g$;*

LT9. $|\underline{T} f - \underline{T} g| \leq -\underline{T}(-|f - g|)$ *for all f, g in $\mathscr{L}(\mathscr{Y})$.*

Proof For any x in $\mathscr{Y}$, the operator component $\underline{T}_x := [\underline{T} \cdot](x)$ is a coherent lower prevision with domain $\mathscr{L}(\mathscr{Y})$, the linear space of all bounded functions on $\mathscr{Y}$. The properties therefore follow from their respective counterparts for lower previsions, see for instance [13, Theorem 4.13]. ∎

De Bock [2] states some additional basic properties that follow from those of Lemma 23.

Lemma 24 *Consider a lower transition transformation $\underline{T} \colon \mathscr{L}(\mathscr{Y}) \to \mathscr{L}(\mathscr{Y})$. Then for all f, g in $\mathscr{L}(\mathscr{Y})$ and all non-negatively homogeneous transformations A and B,*

LT10. $\|\underline{T}\| \leq 1$*;*

LT11. $\|\underline{T}f - \underline{T}g\| \leq \|f - g\|$*;*

LT12. $\|\underline{T}A - \underline{T}B\| \leq \|A - B\|$*.*

Proof (LT10) can be verified by combining the definition of the operator norm and (LT5). Next, (LT11) follows from the definition of the (supremum) norm, (LT9) and (LT5). Finally, (LT12) follows from the definition of the operator norm and (LT11). ∎

The following is an obvious extension/adaptation of [8, Proposition 7.1] to our setting.

Lemma 25 *For any two lower transition (counting) transformations $\underline{T}_1$ and $\underline{T}_2$, their composition $\underline{T}_1\underline{T}_2$ is again a lower transition (counting) transformation.*

Proof To verify the four conditions of Definition 22, we fix some f, g in $\mathscr{L}(\mathscr{Y})$, γ in $\mathbb{R}_{\geq 0}$ and x in $\mathscr{Y}$. The first condition follows immediately from applying (LT1) twice:

$$\underline{T}_1\underline{T}_2(\gamma f) = \underline{T}_1(\gamma \underline{T}_2 f) = \gamma \underline{T}_1\underline{T}_2 f.$$

Next, we move on to the second condition. As $\underline{T}_2$ is a lower transition transformation, it follows from (LT2) that $\underline{T}_2(f + g) \geq \underline{T}_2 f + \underline{T}_2 g$. Hence, it follows from (LT8) that

$$\underline{T}_1\underline{T}_2(f + g) \geq \underline{T}_1(\underline{T}_2 f + \underline{T}_2 g) \geq \underline{T}_1\underline{T}_2 f + \underline{T}_1\underline{T}_2 g,$$

where for the final inequality we have again used (LT2). The third condition follows from (LT3), (LT8) and (LT6):

$$\underline{T}_1\underline{T}_2 f \geq \underline{T}_1 \inf f = \inf f.$$

Finally, we assume that $\underline{T}_1$ and $\underline{T}_2$ are both lower counting transformations, and verify that their composition $\underline{T}_1\underline{T}_2$ satisfies the fourth condition. To that end, we let $h := \underline{T}_2 f$ and $h' := \underline{T}_2(\mathbb{I}_{\geq x} f)$. Observe that, for all z in $\mathscr{X}$,

$$
\begin{aligned}
\mathbb{I}_{\geq x}(z)h(z) &= \mathbb{I}_{\geq x}(z)[\underline{T}_2 f](z) = \mathbb{I}_{\geq x}(z)[\underline{T}_2(\mathbb{I}_{\geq z} f)](z) \\
&= \begin{cases} [\underline{T}_2(\mathbb{I}_{\geq z} f)](z) & \text{if } z \geq x \\ 0 & \text{otherwise} \end{cases} = \begin{cases} [\underline{T}_2(\mathbb{I}_{\geq z}(\mathbb{I}_{\geq x} f))](z) & \text{if } z \geq x \\ 0 & \text{otherwise} \end{cases} = \begin{cases} [\underline{T}_2(\mathbb{I}_{\geq x} f)](z) & \text{if } z \geq x \\ 0 & \text{otherwise} \end{cases} \\
&= \mathbb{I}_{\geq x}(z)[\underline{T}_2(\mathbb{I}_{\geq x} f)](z) = \mathbb{I}_{\geq x}(z)h'(z),
\end{aligned}
$$

where the second and fifth equality hold due to (LT4) because $\underline{T}_2$ is a lower counting transformation. Hence, $\mathbb{I}_{\geq x} h = \mathbb{I}_{\geq x} h'$. As $\underline{T}_1$ is a lower counting transformation as well, it now follows that

$$[\underline{T}_1\underline{T}_2 f](x) = [\underline{T}_1 h](x) = [\underline{T}_1(\mathbb{I}_{\geq x} h)](x) = [\underline{T}_1(\mathbb{I}_{\geq x} h')](x) = [\underline{T}_1 h'](x) = [\underline{T}_1\underline{T}_2(\mathbb{I}_{\geq x} f)](x).$$

∎

The following is an extension of [8, Lemma E.4] to our—slightly—more general setting.

Lemma 26 *Consider some n in $\mathbb{N}$ and two sequences $\underline{T}_1, \dots, \underline{T}_n$ and $\underline{T}'_1, \dots, \underline{T}'_n$ of lower transition transformations on $\mathscr{L}(\mathscr{Y})$. Then*

$$\left\| \prod_{i=1}^{n} \underline{T}_i - \prod_{i=1}^{n} \underline{T}'_i \right\| \leq \sum_{i=1}^{n} \|\underline{T}_i - \underline{T}'_i\|.$$

Proof Our proof is entirely the same as that of [8, Lemma E.4], and is one using induction. Observe that the stated clearly holds for $n = 1$. Fix some m in $\mathbb{N}$ and assume that the stated holds for $n = m$. We now show that the stated then also holds for $m + 1$.

$$\left\| \prod_{i=1}^{m+1} \underline{T}_i - \prod_{i=1}^{m+1} \underline{T}'_i \right\| = \left\| \prod_{i=1}^{m+1} \underline{T}_i - \left(\prod_{i=1}^{m} \underline{T}_i \right) \underline{T}'_{m+1} + \left(\prod_{i=1}^{m} \underline{T}_i \right) \underline{T}'_{m+1} - \prod_{i=1}^{m+1} \underline{T}'_i \right\|$$

$$\leq \left\| \prod_{i=1}^{m+1} \underline{T}_i - \left(\prod_{i=1}^{m} \underline{T}_i \right) \underline{T}'_{m+1} \right\| + \left\| \left(\prod_{i=1}^{m} \underline{T}_i \right) \underline{T}'_{m+1} - \prod_{i=1}^{m+1} \underline{T}'_i \right\|$$

$$= \left\| \left(\prod_{i=1}^{m} \underline{T}_i \right) \underline{T}_{m+1} - \left(\prod_{i=1}^{m} \underline{T}_i \right) \underline{T}'_{m+1} \right\| + \left\| \left(\prod_{i=1}^{m} \underline{T}_i - \prod_{i=1}^{m} \underline{T}'_i \right) \underline{T}'_{m+1} \right\|$$

$$\leq \| \underline{T}_{m+1} - \underline{T}'_{m+1} \| + \left\| \prod_{i=1}^{m} \underline{T}_i - \prod_{i=1}^{m} \underline{T}'_i \right\| \| \underline{T}'_{m+1} \|$$

$$\leq \| \underline{T}_{m+1} - \underline{T}'_{m+1} \| + \left\| \prod_{i=1}^{m} \underline{T}_i - \prod_{i=1}^{m} \underline{T}'_i \right\|$$

$$\leq \| \underline{T}_{m+1} - \underline{T}'_{m+1} \| + \sum_{i=1}^{m} \| \underline{T}_i - \underline{T}'_i \| = \sum_{i=1}^{m+1} \| \underline{T}_i - \underline{T}'_i \|,$$

where the second inequality follows from Lemma 25, (LT12) and (NH5), the third inequality follows from (LT10) and the penultimate inequality follows from the induction hypothesis. ∎

Appendix B. Lower Transition Rate Transformations and the Corresponding Semi-Group of Lower Transition Transformations

In this short section, we introduce lower transition rate transformations, a second type of non-negatively homogeneous transformations. Additionally, we also briefly explain how these lower transition transformations generate a semi-group of lower transition transformations. Rather than working with a non-empty, ordered and possibly countably infinite set $\mathscr{Y}$, in this section we will consider a non-empty and *finite* set χ.

Definition 27 (Definition 7.2 in [8]) *A lower transition rate transformation is a transformation $\underline{R}^{\chi} \colon \mathscr{L}(\chi) \to \mathscr{L}(\chi)$ such that*

LR1. $\underline{R}^{\chi}(\gamma f) = \gamma \underline{R}^{\chi} f$, *for all f in $\mathscr{L}(\chi)$ and γ in $\mathbb{R}_{\geq 0}$;*   *[non-negative homogeneity]*

LR2. $\underline{R}^{\chi}(f + g) \geq \underline{R}^{\chi} f + \underline{R}^{\chi} g$, *for all f, g in $\mathscr{L}(\chi)$;*   *[super-additivity]*

LR3. $\underline{R}^{\chi} \mu = 0$, *for all μ in $\mathbb{R}$;*   *[zero row-sums]*

LR4. $[\underline{R}^{\chi} \mathbb{I}_y](x) \geq 0$, *for all $x, y \in \chi$ with $x \neq y$.*   *[non-negative off-diagonal elements]*

B.1. The Corresponding Semi-Group

We first repeat two intermediate results that are essential to the construction method of the semi-group. We will see in Appendix C further one that similar results hold in the setting of (generalised) Poisson generators.

Lemma 28 (Proposition 4 in [3]) *If $\underline{R}^{\chi} \colon \mathscr{L}(\chi) \to \mathscr{L}(\chi)$ is a lower transition rate transformation, then*

$$\| \underline{R}^{\chi} \| = 2 \max \{ |[\underline{R}^{\chi} \mathbb{I}_x](x)| \colon x \in \chi \}.$$

Lemma 29 (Proposition 3 in [3]) *Consider any lower transition rate transformation $\underline{R}^{\chi} \colon \mathscr{L}(\chi) \to \mathscr{L}(\chi)$ and Δ in $\mathbb{R}_{\geq 0}$. Then $(I + \Delta \underline{R}^{\chi})$ is a lower transition transformation if and only if $\Delta \| \underline{R}^{\chi} \| \leq 2$.*

Next, we repeat the two results that establish how a lower transition rate transformation generates a family of lower transition transformations; they are our direct inspiration for Theorem 10, as well as for Theorems 44 and 45 further on.

Proposition 30 (Corollary 7.11 in [8]) *Consider some lower transition rate transformation $\underline{R}^{\chi} \colon \mathscr{L}(\chi) \to \mathscr{L}(\chi)$, and fix some t, s in $\mathbb{R}_{\geq 0}$ such that $t \leq s$. For every sequence $\{u_i\}_{i \in \mathbb{N}}$ in $\mathscr{U}_{[t,s]}$ with $\lim_{i \to +\infty} \sigma(u_i) = 0$, the corresponding sequence*

$$\left\{ \prod_{k=1}^{k_i} (I + \Delta_k^i \underline{R}^{\chi}) \right\}_{i \in \mathbb{N}}$$

converges to a lower transition transformation, where for every i in $\mathbb{N}$, $k_i + 1$ is the length of the sequence $u_i = t_0^i, \ldots, t_{k_i}^i$ and, for every k in $\{1, \ldots, k_i\}, \Delta_k^i$ is the difference between the consecutive time points t_k^i and t_{k-1}^i of this sequence.

14

Proposition 31 (Theorem 7.12 in [8]) *Consider a lower transition rate transformation $\underline{R}^\chi \colon \mathscr{L}(\chi) \to \mathscr{L}(\chi)$. Then for any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, there is a unique lower transition transformation $\underline{T} \colon \mathscr{L}(\chi) \to \mathscr{L}(\chi)$ such that*

$$(\forall \varepsilon \in \mathbb{R}_{>0})(\exists \delta \in \mathbb{R}_{>0})(\forall u \in \mathscr{U}_{[t,s]}, \sigma(u) \leq \delta) \left\| \underline{T} - \prod_{i=1}^{n}(I + \Delta_i \underline{R}^\chi) \right\| \leq \varepsilon.$$

Consider some lower transition rate transformation $\underline{R}^\chi \colon \mathscr{L}(\chi) \to \mathscr{L}(\chi)$, and fix some t, s in $\mathbb{R}_{\geq 0}$ such that $t \leq s$. As explained by Krak et al. [8, Section 7.3], the two results above allow us to define the corresponding lower transition transformation

$$\underline{T}_{t,s}^\chi := \lim_{\sigma(u) \to 0} \left\{ \prod_{i=1}^{n}(I + \Delta_i \underline{R}^\chi) \colon u \in \mathscr{U}_{[t,s]} \right\}.$$

In this definition, the unconventional notation for the limit is used to indicate that the limit does not depend on the choice of sequence $\{u_i\}_{i\in\mathbb{N}}$ in $\mathscr{U}_{[t,s]}$, all that is required is that $\lim_{i\to+\infty}\sigma(u_i) = 0$. We conclude this brief section by repeating the result that establishes that the family of corresponding lower transition transformations forms a semi-group.

Proposition 32 (Propositions 7.13–14 in [8]) *Consider any lower transition rate operator $\underline{R}^\chi \colon \mathscr{L}(\chi) \to \mathscr{L}(\chi)$. Then for any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$,*

(i) $\underline{T}_{t,t}^\chi = I$;

(ii) $\underline{T}_{t,s}^\chi = \underline{T}_{t,r}^\chi \underline{T}_{r,s}^\chi$ *for any r in $\mathbb{R}_{\geq 0}$ such that $t \leq r \leq s$;*

(iii) $\underline{T}_{t,s}^\chi = \underline{T}_{0,s-t}^\chi$.

Appendix C. The Generalised Poisson Generator

In this section, we essentially generalise the results of Appendix B to the setting of a countably infinite state space; however, we limit ourselves to one specific type of (a generalisation of) lower transition rate transformations. Essential to our exposition are sequences $S := \{(\underline{\lambda}_x, \overline{\lambda}_x)\}_{x\in\mathscr{X}}$ in $\mathbb{R}_{\geq 0}^2$ such that $\underline{\lambda}_x \leq \overline{\lambda}_x$ for all $x \in \mathscr{X}$. Even more, we will usually demand that $\underline{\lambda}_x$ and $\overline{\lambda}_x$ are both contained in $\Lambda = [\underline{\lambda}, \overline{\lambda}]$. We collect all such sequences in the set

$$\mathscr{S}_\Lambda := \left\{ \{(\underline{\lambda}_x, \overline{\lambda}_x)\}_{x\in\mathscr{X}} \text{ in } [\underline{\lambda}, \overline{\lambda}]^2 \colon (\forall x \in \mathscr{X}) \, \underline{\lambda}_x \leq \overline{\lambda}_x \right\}.$$

With any $S = \{(\underline{\lambda}_x, \overline{\lambda}_x)\}_{x\in\mathscr{X}}$ in $\mathscr{S}_\Lambda$, we associate the generalised Poisson generator $\underline{Q}_S$, defined for all f in $\mathscr{L}(\mathscr{X})$ and x in $\mathscr{X}$ as

$$[\underline{Q}_S f](x) := \min\{\lambda f(x+1) - \lambda f(x) \colon \lambda \in [\underline{\lambda}_x, \overline{\lambda}_x]\}. \tag{19}$$

Observe that the generalised Poisson generator is a generalisation of the Poisson generator, because clearly

$$\underline{Q} = \underline{Q}_S \quad \text{with } S = \{(\underline{\lambda}, \overline{\lambda})\}_{x\in\mathscr{X}}. \tag{20}$$

C.1. From Generalised Poisson Generators . . .

We first establish that generalised Poisson generators can be seen as lower transition rate transformations with a countably infinite state space; more precisely, we establish that they are transformations that furthermore satisfy properties that are similar to conditions (LR1)–(LR4) of Definition 27.

Proposition 33 *Consider a sequence $S = \{(\underline{\lambda}_x, \overline{\lambda}_x)\}_{x\in\mathscr{X}}$ in $\mathscr{S}_\Lambda$. Then $\underline{Q}_S$ is a transformation on $\mathscr{L}(\mathscr{X})$. Furthermore,*

GP1. $\underline{Q}_S(\gamma f) = \gamma \underline{Q}_S f$ *for all γ in $\mathbb{R}_{\geq 0}$ and f in $\mathscr{L}(\mathscr{X})$;*   *[non-negative homogeneity]*

GP2. $\underline{Q}_S(f + g) \geq \underline{Q}_S f + \underline{Q}_S g$ *for all f, g in $\mathscr{L}(\mathscr{X})$;*   *[super-additivity]*

GP3. $\underline{Q}_s \mu = 0$ *for all μ in $\mathbb{R}$;*   *[zero row-sums]*

GP4. $[\underline{Q}_S \mathbb{I}_y](x) \geq 0$ *for all x, y in $\mathscr{X}$ with $x \neq y$;*   *[non-negative off-diagonal elements]*

GP5. $[\underline{Q}_S f](x) = [\underline{Q}_S(\mathbb{I}_{\geq x} f)](x)$ *for all x in $\mathscr{X}$ and f in $\mathscr{L}(\mathscr{X})$;*

GP6. $[\underline{Q}_S f](x) = [\underline{Q}_S(\mathbb{I}_{\leq x+1} f)](x)$ *for all x in $\mathscr{X}$ and f in $\mathscr{L}(\mathscr{X})$.*

Proof We first verify that $\underline{Q}_S$ is a transformation on $\mathscr{L}(\mathscr{X})$. To that end, we fix any f in $\mathscr{L}(\mathscr{X})$. Observe that, for any x in $\mathscr{X}$,

$$|[\underline{Q}_S f](x)| = \left|\min\{\lambda f(x+1) - \lambda f(x) \colon \lambda \in (\underline{\lambda}_x, \overline{\lambda}_x)\}\right| \leq \overline{\lambda}_x |f(x+1) - f(x)| \leq 2\overline{\lambda}_x \|f\| \leq 2\overline{\lambda}\|f\|.$$

Hence, $\underline{Q}_S f$ is clearly bounded. Since f was arbitrary, this proves that $\underline{Q}_S$ is a transformation, as required.

For the second part of the stated, we observe that properties (GP1)–(GP6) follow immediately from the definition of $\underline{Q}_S$.
∎

Next, we consider the norm of a generalised Poisson generator. Note that the following result is similar to—or an extension of—Lemma 28 because

$$2\sup\{|[\underline{Q}_S \mathbb{I}_x](x)| \colon x \in \mathscr{X}\} = 2\sup\{|-\overline{\lambda}_x| \colon x \in \mathscr{X}\} = 2\sup\{\overline{\lambda}_x \colon x \in \mathscr{X}\}.$$

Lemma 34 *For any sequence $S = \{(\underline{\lambda}_x, \overline{\lambda}_x)\}_{x \in \mathscr{X}}$ in $\mathscr{S}_\Lambda$,*

$$\|\underline{Q}_S\| = 2\sup\{\overline{\lambda}_x \colon x \in \mathscr{X}\}.$$

Proof We first show that $\|\underline{Q}_S\| \geq 2\sup\{\overline{\lambda}_x \colon x \in \mathscr{X}\}$. For any y in $\mathscr{X}$, we let $f_y := \mathbb{I}_y - \mathbb{I}_{y+1}$. Fix any y in $\mathscr{X}$. Then for any z in $\mathscr{X}$,

$$\begin{aligned}
[\underline{Q}_S f_y](z) &= \min\{\lambda f_y(z+1) - \lambda f_y(z) \colon \lambda \in [\underline{\lambda}_z, \overline{\lambda}_z]\} \\
&= \min\{\lambda \mathbb{I}_y(z+1) - \lambda \mathbb{I}_{y+1}(z+1) - \lambda \mathbb{I}_y(z) + \lambda \mathbb{I}_{y+1}(z) \colon \lambda \in [\underline{\lambda}_z, \overline{\lambda}_z]\} \\
&= \begin{cases}
\underline{\lambda}_z & \text{if } z = y-1 \text{ or } z = y+1, \\
-2\overline{\lambda}_z & \text{if } z = y, \\
0 & \text{otherwise.}
\end{cases}
\end{aligned}$$

Observe that for any y in $\mathscr{X}$, f_y is a bounded real-valued function on $\mathscr{X}$ and that

$$\|\underline{Q}_S f_y\| = \sup\{|[\underline{Q}_S f_y](z)| \colon z \in \mathscr{X}\} = \max\{\underline{\lambda}_{y-1}, 2\overline{\lambda}_y, \underline{\lambda}_{y+1}\},$$

where $\underline{\lambda}_{y-1}$ is not included in the set if $y = 0$. Therefore

$$\sup\{\|\underline{Q}_S f_x\| \colon x \in \mathscr{X}\} = \sup\{\max\{\underline{\lambda}_{x-1}, 2\overline{\lambda}_x, \underline{\lambda}_{x+1}\} \colon x \in \mathscr{X}\} = 2\sup\{\overline{\lambda}_x \colon x \in \mathscr{X}\}.$$

Since $\|f_x\| = 1$ for all x in $\mathscr{X}$, we observe that

$$\|\underline{Q}_S\| = \sup\{\|\underline{Q}_S f\| \colon f \in \mathscr{L}(\mathscr{X}), \|f\| = 1\} \geq \sup\{\|\underline{Q}_S f_x\| \colon x \in \mathscr{X}\} = 2\sup\{\overline{\lambda}_x \colon x \in \mathscr{X}\},$$

as we set out to prove.

Next, we prove that $\|\underline{Q}_S\| \leq 2\sup\{\overline{\lambda}_x \colon x \in \mathscr{X}\}$. Fix any ε in $\mathbb{R}_{>0}$. Then by the definition of $\|\underline{Q}_S\|$ and $\|\underline{Q}_S f\|$, there is an f in $\mathscr{L}(\mathscr{X})$ with $\|f\| = 1$ and an y in $\mathscr{X}$ such that

$$\|\underline{Q}_S\| - \varepsilon < \|\underline{Q}_S f\| - \frac{\varepsilon}{2} < |[\underline{Q}_S f](y)|.$$

From the definition of $\underline{Q}_S$, it now follows that

$$\begin{aligned}
|[\underline{Q}_S f](y)| &= |\min\{\lambda f(y+1) - \lambda f(y) \colon \lambda \in [\underline{\lambda}_y, \overline{\lambda}_y]\}| \\
&\leq \max\{\lambda |f(y+1) - f(y)| \colon \lambda \in [\underline{\lambda}_y, \overline{\lambda}_y]\} \\
&= \overline{\lambda}_y |f(y+1) - f(y)| \leq 2\overline{\lambda}_y \leq 2\sup\{\overline{\lambda}_x \colon x \in \mathscr{X}\},
\end{aligned}$$

where the second inequality follows from the fact that $\|f\| = 1$. Hence,

$$\|\underline{Q}_S\| - \varepsilon < |[\underline{Q}_S f](y)| \leq 2\sup\{\overline{\lambda}_x \colon x \in \mathscr{X}\}.$$

As this holds for any arbitrary positive real number ε, we conclude that $\|\underline{Q}_S\| \leq 2\sup\{\overline{\lambda}_x \colon x \in \mathscr{X}\}$, as required. The stated now follows because $\|\underline{Q}_S\| \geq 2\sup\{\overline{\lambda}_x \colon x \in \mathscr{X}\}$ and $\|\underline{Q}_S\| \leq 2\sup\{\overline{\lambda}_x \colon x \in \mathscr{X}\}$. ∎

C.2. …to Lower Counting Transformations

The generalised Poisson generator naturally defines a family of lower transition (or, more precisely, counting) transformations. Crucial to our exposition are Theorems 44 and 45 further on. In essence, these two results extend Propositions 30 and 31 to the setting of generalised Poisson generators. Even more, our reasoning that uniquely defines this family of transformations is largely analoguous to the line of reasoning followed in [8, Appendix E]. Our first step is the following observation.

Lemma 35 *Consider some S in $\mathscr{S}_\Lambda$ and some Δ in $\mathbb{R}_{\geq 0}$. Then $(I + \Delta \underline{Q}_S)$ is a lower counting transformation if and only if $\Delta \|\underline{Q}_S\| \leq 2$.*

Proof We first check the sufficiency of the condition $\Delta \|\underline{Q}_S\| \leq 2$. To that end, we fix any Δ in $\mathbb{R}_{\geq 0}$ that satisfies this condition, and let $\underline{T} := I + \Delta \underline{Q}_S$. That $\underline{T}$ satisfies (LT1) follows immediately from the non-negative homogeneity of I and that of $\underline{Q}$—that is, (GP1); similarly, (LT2) follows immediately from the super-additivity of I and that of $\underline{Q}$—that is, (GP2). Next, we verify that (LT3) holds. To that, we fix any f in $\mathscr{L}(\mathscr{X})$ and x in $\mathscr{X}$, and observe that

$$[\underline{T}f](x) = [If](x) + [\Delta \underline{Q}_S f](x) = f(x) + \Delta[\underline{Q}_S f](x) = f(x) + \Delta \min\{\lambda_x f(x+1) - \lambda_x f(x) \colon \lambda_x \in [\underline{\lambda}_x, \overline{\lambda}_x]\}$$
$$= \min\{f(x) + \Delta \lambda_x f(x+1) - \Delta \lambda_x f(x) \colon \lambda_x \in [\underline{\lambda}_x, \overline{\lambda}_x]\}$$
$$= \min\{(1 - \Delta \lambda_x)f(x) + \Delta \lambda_x f(x+1) \colon \lambda_x \in [\underline{\lambda}_x, \overline{\lambda}_x]\}, \tag{21}$$

where the fourth equality holds because $\Delta \geq 0$. Observe now that $0 \leq \Delta \lambda_x$ because Δ and λ_x are non-negative by assumption, and that furthermore $\Delta \lambda_x \leq 1$ because $\Delta \|\underline{Q}_S\| \leq 2$ by assumption and $\lambda_x \leq \sup\{\lambda_y \colon y \in \mathscr{X}\} = \|\underline{Q}_S\|/2$ due to Lemma 34. Consequently, the sum in the minimum in Equation (21) is a convex combination of $f(x)$ and $f(x+1)$. Because a convex combination of two real numbers is always greater than or equal to the minimum of these two numbers, it now follows that

$$[\underline{T}f](x) = \min\{(1 - \Delta \lambda_x)f(x) + \Delta \lambda_x f(x+1) \colon \lambda_x \in [\underline{\lambda}_x, \overline{\lambda}_x]\} \geq \min\{f(x), f(x+1)\} \geq \inf f.$$

Since this holds for all f in $\mathscr{L}(\mathscr{X})$ and x in $\mathscr{X}$, this implies (LT3). Finally, (LT4) follows immediately from (GP5) and the observation that $[If](x) = [I(\mathbb{I}_{\geq x}f)](x)$ for all f in $\mathscr{L}(\mathscr{X})$ and x in $\mathscr{X}$.

That the condition $\Delta \|\underline{Q}_S\| \leq 2$ is necessary follows from a counterexample. Fix some Δ in $\mathbb{R}_{>0}$ such that $\Delta \|\underline{Q}_S\| > 2$, and assume ex-absurdo that $(I + \Delta \underline{Q}_S)$ is a lower counting transformation. Fix any ε in $\mathbb{R}_{>0}$ such that $2\Delta \varepsilon < \Delta \|\underline{Q}_S\| - 2$. Then there is an x in $\mathscr{X}$ such that

$$[\underline{Q}_S \mathbb{I}_x](x) = -\overline{\lambda}_x \leq -\sup\{\overline{\lambda}_y \colon y \in \mathscr{X}\} + \varepsilon = -\frac{\|\underline{Q}_S\|}{2} + \varepsilon, \tag{22}$$

where the second equality follows from Lemma 34. Therefore

$$[(I + \Delta \underline{Q}_S)\mathbb{I}_x](x) = \mathbb{I}_x(x) + \Delta[\underline{Q}_S \mathbb{I}_x](x) = 1 + \Delta[\underline{Q}_S \mathbb{I}_x](x) \leq 1 - \Delta \frac{\|\underline{Q}_S\|}{2} + \Delta \varepsilon = \frac{1}{2}\Big(2 - \Delta \|\underline{Q}_S\| + 2\Delta \varepsilon\Big) < 0,$$

where the final inequality follows from our condition on ε. Since $\inf \mathbb{I}_x = 0$, this is a clear contradiction with (LT3). Hence, the condition $\Delta \|\underline{Q}_S\| \leq 2$ is indeed necessary. ∎

For our second step, we construct a lower counting transformation as the composition of lower counting transformations of the form of Lemma 35. More formally, we construct this transformation as follows. For any sequence $S = \{(\underline{\lambda}_x, \overline{\lambda}_x)\}_{x \in \mathscr{X}}$ in $\mathscr{S}_\Lambda$, any t, s in $\mathbb{R}_{\geq 0}$ such that $t \leq s$ and any u in $\mathscr{U}_{[t,s]}$, we let

$$\Phi_u := \begin{cases} \prod_{i=1}^{n}(I + \Delta_i \underline{Q}_S) & \text{if } t < s, \\ I & \text{otherwise.} \end{cases} \tag{23}$$

This notation is clearly reminiscent of the notation that was previously introduced in Section 5.2; in fact, the latter is a special case of the former because—as was previously observed in Equation (20)—the Poisson generator $\underline{Q}$ is a special case of the generalised Poisson generator. The following result establishes that the operator Φ_u thus defined is a lower counting transformation.

Corollary 36 *Fix a sequence $S = \{(\underline{\lambda}_x, \overline{\lambda}_x)\}_{x \in \mathscr{X}}$ in $\mathscr{S}_\Lambda$, some t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$ and a sequence $u = t_0, \ldots, t_n$ in $\mathscr{U}_{[t,s]}$. If $\sigma(u)\|\underline{Q}_S\| \leq 2$, then Φ_u, as defined in Equation (23), is a lower counting transformation.*

Proof Follows immediately from Lemmas 25 and 35. ∎

Next, we establish some results that will allow us to determine the difference between Φ_u and $\Phi_{u'}$, where u and u' are two sequences of time points in $\mathscr{U}_{[t,s]}$.

Lemma 37 *Consider a sequence $S = \{(\underline{\lambda}_x, \overline{\lambda}_x)\}_{x \in \mathscr{X}}$ in $\mathscr{S}_\Lambda$ and a sequence $\Delta_1, \ldots, \Delta_n$ in $\mathbb{R}_{\geq 0}$ with n in $\mathbb{N}$. If $\Delta_i \|\underline{Q}_S\| \leq 2$ for all i in $\{1, \ldots, n\}$, then*

$$\left\| \prod_{i=1}^{n}(I + \Delta_i \underline{Q}_S) - (I + \Delta \underline{Q}_S) \right\| \leq \|\underline{Q}_S\|^2 \sum_{i=1}^{n} \Delta_i \sum_{j=i+1}^{n} \Delta_j,$$

where $\Delta := \sum_{i=1}^{n} \Delta_i$.

In our proof of Lemma 37, we will make use of the following corollary.

Corollary 38 *Consider a sequence $S = \{(\underline{\lambda}_x, \overline{\lambda}_x)\}_{x \in \mathscr{X}}$ in $\mathscr{S}_\Lambda$. Then for any two non-negatively homogeneous transformations A and B on $\mathscr{L}(\mathscr{Y})$,*

$$\|\underline{Q}_S A - \underline{Q}_S B\| \leq \|\underline{Q}_S\| \|A - B\|.$$

Proof Our proof is—almost—equal to that of [2, R12]. Since the stated is clearly true for $\|\underline{Q}_S\| = 0$, we may assume that $\|\underline{Q}_S\| > 0$ without loss of generality. If we let $\Delta := 2/\|\underline{Q}_S\|$, then it follows from Lemma 35 that $\underline{T} := I + \Delta \underline{Q}$ is a lower counting transformation. Observe that

$$\|\underline{Q}_S A - \underline{Q}_S B\| = \left\| \frac{\|\underline{Q}_S\|}{2}(\underline{T}A - A) - \frac{\|\underline{Q}_S\|}{2}(\underline{T}B - B) \right\| = \frac{\|\underline{Q}_S\|}{2} \|(\underline{T}A - \underline{T}B) - (A - B)\|$$

$$\leq \frac{\|\underline{Q}_S\|}{2}\|\underline{T}A - \underline{T}B\| + \frac{\|\underline{Q}_S\|}{2}\|A - B\| \leq \frac{\|\underline{Q}_S\|}{2}\|A - B\| + \frac{\|\underline{Q}_S\|}{2}\|A - B\| = \|\underline{Q}_S\|\|A - B\|,$$

where the final inequality follows from Lemma 24 (LT12) because $\underline{T}$ is a lower counting transformation by construction. ∎

Proof of Lemma 37 Our proof is one by induction, and is almost equivalent to the one that Krak et al. provide for [8, Lemma E.5], although ours yields a (marginally) smaller upper bound. First, we observe that for $n = 1$, the stated is trivially true. Next, we fix some $n \geq 2$ and assume that the stated holds for $1 \leq n' < n$. We now show that this then implies that the stated also holds for n. Some straightforward manipulations yield

$$\left\| \prod_{i=1}^{n}(I + \Delta_i \underline{Q}_S) - (I + \Delta \underline{Q}_S) \right\| = \left\| \prod_{i=2}^{n}(I + \Delta_i \underline{Q}_S) + \Delta_1 \underline{Q}_S \prod_{i=2}^{n}(I + \Delta_i \underline{Q}_S) - I - \left(\sum_{i=2}^{n} \Delta_i\right)\underline{Q}_S - \Delta_1 \underline{Q}_S \right\|$$

$$\leq \left\| \prod_{i=2}^{n}(I + \Delta_i \underline{Q}_S) - I - \left(\sum_{i=2}^{n} \Delta_i\right)\underline{Q}_S \right\| + \left\| \Delta_1 \underline{Q}_S \prod_{i=2}^{n}(I + \Delta_i \underline{Q}_S) - \Delta_1 \underline{Q}_S \right\|$$

$$\leq \|\underline{Q}_S\|^2 \sum_{i=2}^{n} \Delta_i \sum_{j=i+1}^{n} \Delta_j + \left\| \Delta_1 \underline{Q}_S \prod_{i=2}^{n}(I + \Delta_i \underline{Q}_S) - \Delta_1 \underline{Q}_S \right\|$$

$$\leq \|\underline{Q}_S\|^2 \sum_{i=2}^{n} \Delta_i \sum_{j=i+1}^{n} \Delta_j + \Delta_1 \|\underline{Q}_S\| \left\| \prod_{i=2}^{n}(I + \Delta_i \underline{Q}_S) - I \right\|$$

$$\leq \|\underline{Q}_S\|^2 \sum_{i=2}^{n} \Delta_i \sum_{j=i+1}^{n} \Delta_j + \Delta_1 \|\underline{Q}_S\| \sum_{i=2}^{n} \|(I + \Delta_i \underline{Q}_S) - I\|$$

$$= \|\underline{Q}_S\|^2 \sum_{i=2}^{n} \Delta_i \sum_{j=i+1}^{n} \Delta_j + \Delta_1 \|\underline{Q}_S\|^2 \sum_{i=2}^{n} \Delta_i = \|\underline{Q}_S\|^2 \sum_{i=1}^{n} \Delta_i \sum_{j=i+1}^{k} \Delta_j,$$

where the first inequality is a consequence of the triangle inequality, the second inequality follows from the induction hypothesis, the third inequality follows from Corollaries 36 and 38, and the fourth inequality follows from Lemma 26. ∎

18

Lemma 39 *Fix a sequence S in $\mathscr{S}_\Lambda$ and some n in $\mathbb{N}$. Furthermore, for all i in $\{1,\dots,n\}$, we fix some sequence $\Delta_{i,1},\dots\Delta_{i,k_i}$ in $\mathbb{R}_{\geq 0}$ and let $\Delta_i := \sum_{j=1}^{k_i} \Delta_{i,j}$. Let $\Delta := \sum_{i=1}^{n} \Delta_i$ and $\Delta^\star := \max\{\Delta_i : i \in \{1,\dots,n\}\}$. If $\Delta^\star\|\underline{Q}_S\| \leq 2$, then*

$$\left\| \prod_{i=1}^{n}\left(\prod_{j=1}^{k_i}(I+\Delta_{i,j}\underline{Q}_S) \right) - \prod_{i=1}^{n}(I+\Delta_i\underline{Q}_S) \right\| \leq \|\underline{Q}_S\|^2 \sum_{i=1}^{n}\Delta_i^2 \leq \|\underline{Q}_S\|^2\Delta\Delta^\star.$$

Proof Our proof is almost the same as that of [8, Lemma E.6]. Observe that

$$\left\| \prod_{i=1}^{n}\left(\prod_{j=1}^{k_i}(I+\Delta_{i,j}\underline{Q}_S) \right) - \prod_{i=1}^{n}(I+\Delta_i\underline{Q}_S) \right\| \leq \sum_{i=1}^{n}\left\| \prod_{j=1}^{k_i}(I+\Delta_{i,j}\underline{Q}_S) - (I+\Delta_i\underline{Q}_S) \right\| \leq \sum_{i=1}^{n}\|\underline{Q}_S\|^2 \sum_{j=1}^{k_i}\Delta_{i,j} \sum_{\ell=j+1}^{k_i}\Delta_{i,\ell}$$

$$\leq \|\underline{Q}_S\|^2 \sum_{i=1}^{n}\sum_{j=1}^{k_i}\Delta_{i,j}\Delta_i \leq \|\underline{Q}_S\|^2 \sum_{i=1}^{n}\Delta_i\Delta_i$$

$$\leq \|\underline{Q}_S\|^2 \sum_{i=1}^{n}\Delta_i\Delta^\star = \|\underline{Q}_S\|^2\Delta\Delta^\star,$$

where the first inequality follows from Corollary 36 and Lemma 26, and the second inequality follows from Lemma 37. ∎

Everything is now set up to establish the following two results regarding the difference between Φ_u and $\Phi_{u'}$.

Corollary 40 *Consider a sequence S in $\mathscr{S}_\Lambda$, some t,s in $\mathbb{R}_{\geq 0}$ with $t \leq s$ and some u in $\mathscr{U}_{[t,s]}$ such that $\sigma(u)\|\underline{Q}_S\| \leq 2$. Then for any u' in $\mathscr{U}_{[t,s]}$ such that $u \subseteq u'$,*

$$\|\Phi_u - \Phi_{u'}\| \leq \sigma(u)(s-t)\|\underline{Q}_S\|^2.$$

Proof Follows almost immediately from Lemma 39 ∎

Lemma 41 *Fix a sequence S in $\mathscr{S}_\Lambda$, t,s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, δ in $\mathbb{R}_{\geq 0}$ with $\delta\|\underline{Q}_S\| \leq 2$ and u,u' in $\mathscr{U}_{[t,s]}$. If $\sigma(u) \leq \delta$ and $\sigma(u') \leq \delta$, then*

$$\|\Phi_u - \Phi_{u'}\| \leq 2\delta(s-t)\|\underline{Q}_S\|^2.$$

Proof Our proof is entirely similar to that of Krak et al. [8, Proposition 7.9]. Let $u^\star$ be the sequence of time points in $\mathscr{U}_{[t,s]}$ that contains all time points in u and u'. It then follows immediately from Corollary 40 that $\|\Phi_u - \Phi_{u^\star}\| \leq \delta(s-t)\|\underline{Q}_S\|^2$ and $\|\Phi_{u'} - \Phi_{u^\star}\| \leq \delta(s-t)\|\underline{Q}_S\|^2$, whence

$$\|\Phi_u - \Phi_{u'}\| \leq \|\Phi_u - \Phi_{u^\star}\| + \|\Phi_{u^\star} - \Phi_{u'}\| \leq 2\delta(s-t)\|\underline{Q}_S\|^2.$$

∎

Now that we have an upper bound on the measure of the distance between Φ_u and $\Phi_{u'}$, we can fix some sequence $\{u_i\}_{i\in\mathbb{N}}$ in $\mathscr{U}_{[t,s]}$ and study the behaviour of the corresponding sequence $\{\Phi_{u_i}\}_{i\in\mathbb{N}}$ in the limit for $i \to +\infty$.

Lemma 42 *Fix a sequence S in $\mathscr{S}_\Lambda$ and some t,s in $\mathbb{R}_{\geq 0}$ with $t \leq s$. Then for every sequence $\{u_i\}_{i\in\mathbb{N}}$ in $\mathscr{U}_{[t,s]}$ such that $\lim_{i\to+\infty}\sigma(u_i) = 0$, the corresponding sequence $\{\Phi_{u_i}\}_{i\in\mathbb{N}}$ is Cauchy.*

Proof In order to prove the stated, we need to show that for every ε in $\mathbb{R}_{>0}$, there exists an $i^\star$ in $\mathbb{N}$ such that $\|\Phi_{u_i} - \Phi_{u_j}\| \leq \varepsilon$ for all i,j in $\mathbb{N}$ with $i \geq i^\star$ and $j \geq i^\star$. Fix now any ε in $\mathbb{R}_{>0}$. Because $\lim_{i\to+\infty}\sigma(u_i) = 0$, there is an $i^\star$ in $\mathbb{N}$ such that (i) $\sigma(u_i)\|\underline{Q}_S\| \leq 2$ for $i \geq i^\star$, and (ii) $2\sigma(u_i)(s-t)\|\underline{Q}_S\|^2 \leq \varepsilon$ for all $i \geq i^\star$. From this and Lemma 41, it now follows that, for all i,j in $\mathbb{N}$ with $i \geq i^\star$ and $j \geq i^\star$,

$$\|\Phi_{u_i} - \Phi_{u_j}\| \leq 2\max\{\sigma(u_i),\sigma(u_j)\}(s-t)\|\underline{Q}_S\|^2 \leq \varepsilon.$$

Because ε was an arbitrary positive real number, this proves the stated. ∎

Lemma 43 *Fix a sequence S in $\mathscr{S}_\Lambda$, some t,s in $\mathbb{R}_{\geq 0}$ with $t \leq s$ and some f in $\mathscr{L}(\mathscr{X})$. For every sequence $\{u_i\}_{i \in \mathbb{N}}$ in $\mathscr{U}_{[t,s]}$ such that $\lim_{i \to +\infty} \sigma(u_i) = 0$, the corresponding sequence $\{\Phi_{u_i} f\}_{i \in \mathbb{N}}$ converges to a limit $f_{\lim}$ in $\mathscr{L}(\mathscr{X})$ that does not depend on the chosen sequence $\{u_i\}_{i \in \mathbb{N}}$.*

Proof Our proof consists of two parts. In the first part, we will prove that $\{\Phi_{u_i} f\}_{i \in \mathbb{N}}$ converges to a limit; in the second part, we will prove that this limit does not depend on the chosen sequence $\{u_i\}_{i \in \mathbb{N}}$.

Fix some sequence $\{u_i\}_{i \in \mathbb{N}}$ in $\mathscr{U}_{[t,s]}$ such that $\lim_{i \to +\infty} \sigma(u_i) = 0$. The corresponding sequence $\{\Phi_{u_i} f\}_{i \in \mathbb{N}}$ converges to a limit because (i) $\mathscr{L}(\mathscr{X})$ is a complete normed vector space, and (ii) $\{\Phi_{u_i} f\}_{i \in \mathbb{N}}$ is a Cauchy sequence in $\mathscr{L}(\mathscr{X})$. We now prove that $\{\Phi_{u_i} f\}_{i \in \mathbb{N}}$ is a Cauchy sequence. To that end, we fix some ε in $\mathbb{R}_{>0}$. If $\|f\| = 0$, then $f = 0$. Hence, it follows almost immediately from Equations (23) and (19) and (LT6) that $\Phi_{u_i} f = \Phi_{u_i} 0 = 0$ for all i in $\mathbb{N}$. Consequently, $\|\Phi_{u_i} f - \Phi_{u_j}\| = 0 \leq \varepsilon$ for all i, j in $\mathbb{N}$, and so the veracity of the claim is immediate.

Next, we consider the alternative case that $\|f\| \neq 0$. By Lemma 42, there is a $i^\star$ in $\mathbb{N}$ such that

$$(\forall i, j \in \mathbb{N}, i \geq i^\star, j \geq i^\star) \; \|\Phi_{u_i} - \Phi_{u_j}\| \leq \frac{\varepsilon}{\|f\|}.$$

Observe now that

$$(\forall i, j \in \mathbb{N}, i \geq i^\star, j \geq i^\star) \; \|\Phi_{u_i} f - \Phi_{u_j} f\| \leq \|\Phi_{u_i} - \Phi_{u_j}\| \|f\| \leq \frac{\varepsilon}{\|f\|} \|f\| = \varepsilon,$$

where the first inequality holds due to (NH4) because—for reasons explained right after Lemma 21—$A := \Phi_{u_i} - \Phi_{u_j}$ is a non-negatively homogeneous transformation. Since ε was an arbitrary positive real number, we conclude from this that $\{\Phi_{u_i} f\}_{i \in \mathbb{N}}$ is Cauchy.

Next, we prove that the limit does not depend on the chosen sequence. To that end, we fix two sequences $\{u_i\}_{i \in \mathbb{N}}$ and $\{u'_i\}_{i \in \mathbb{N}}$ such that $\lim_{i \to +\infty} \sigma(u_i) = 0$ and $\lim_{i \to +\infty} \sigma(u'_i) = 0$. Furthermore, we let $f_{\lim}$ and $f'_{\lim}$ denote the limits of $\{\Phi_{u_i} f\}_{i \in \mathbb{N}}$ and $\{\Phi_{u'_i} f\}_{i \in \mathbb{N}}$, respectively. In order to prove the stated, we need to verify that $f_{\lim} = f'_{\lim}$. To that end, we observe that, for all i in $\mathbb{N}$,

$$\|f_{\lim} - f'_{\lim}\| = \|f_{\lim} - \Phi_{u_i} f + \Phi_{u_i} f - \Phi_{u'_i} f + \Phi_{u'_i} f - f'_{\lim}\| \leq \|f_{\lim} - \Phi_{u_i} f\| + \|\Phi_{u_i} f - \Phi_{u'_i}\| f + \|f'_{\lim} - \Phi_{u'_i} f\|, \quad (24)$$

where the inequality follows from the triangle inequality. Fix now any ε in $\mathbb{R}_{>0}$, and choose any ε' in $\mathbb{R}_{>0}$ such that $3\varepsilon' \leq \varepsilon$, and additionally choose any δ in $\mathbb{R}_{>0}$ such that $2\delta(s-t)\|\underline{Q}_S\|^2 \|f\| \leq \varepsilon'$. Due to the first part of the statement, and because $\lim_{i \to +\infty} \sigma(u_i) = 0 = \lim_{i \to +\infty} \sigma(u'_i)$, there is some j in $\mathbb{N}$ such that

$$\|\Phi_{u_j} - f_{\lim}\| \leq \varepsilon' \qquad \text{and} \qquad \|\Phi_{u'_j} - f'_{\lim}\| \leq \varepsilon' \quad (25)$$

and

$$\sigma(u_j) \leq \delta \qquad \text{and} \qquad \sigma(u'_j) \leq \delta. \quad (26)$$

Fix any such j. Observe furthermore that

$$\|\Phi_{u_j} f - \Phi_{u'_j} f\| \leq \|\Phi_{u_j} - \Phi_{u'_j}\| \|f\| \leq 2\delta(s-t)\|\underline{Q}_S\|^2 \|f\| \leq \varepsilon', \quad (27)$$

where the first inequality holds due to (NH4) because—for reasons mentioned right after Lemma 21—$\Phi_{u_j} - \Phi_{u'_j}$ is a non-negatively homogeneous transformation, the second inequality follows from Equation (26) and Corollary 40 and the final inequality is precisely our condition on δ. We now use Equations (25) and (27) to bound the terms in Equation (24) for $i = j$, to yield

$$\|f_{\lim} - f'_{\lim}\| \leq \|f_{\lim} - \Phi_{u_j}\| + \|\Phi_{u_j} - \Phi_{u'_j}\| + \|f'_{\lim} - \Phi_{u'_j}\| \leq 3\varepsilon' \leq \varepsilon,$$

where the final inequality is precisely our condition on ε'. Because ε was an arbitrary positive real number, we infer from this inequality that $\|f_{\lim} - f'_{\lim}\| = 0$, which in turn implies that $f_{\lim} = f'_{\lim}$. ∎

We now have all the necessary intermediary results to establish the two main results regarding the limit behaviour of the sequences $\{\Phi_{u_i}\}_{i \in \mathbb{N}}$. Our first result establishes that the sequence always converges to a lower counting transformation. In this sense, it is similar to Proposition 30—that is, [8, Corollary 7.11].

Theorem 44 *Consider a sequence S in $\mathscr{S}_\Lambda$, and fix some t,s in $\mathbb{R}_{\geq 0}$ with $t \leq s$. For any sequence $\{u_i\}_{i \in \mathbb{N}}$ in $\mathscr{U}_{[t,s]}$ such that $\lim_{i \to +\infty} \sigma(u_i) = 0$, the corresponding sequence $\{\Phi_{u_i}\}_{i \in \mathbb{N}}$ converges to a lower counting transformation.*

Proof Recall from Lemma 43 that, for all f in $\mathscr{L}(\mathscr{X})$, the sequence $\{\Phi_{u_i}f\}_{i\in\mathbb{N}}$ converges to the bounded function $f_{\lim}$. Let $\underline{T}$ be the transformation that maps any f in $\mathscr{L}(\mathscr{X})$ to the corresponding limit $f_{\lim}$:

$$\underline{T}\colon \mathscr{L}(\mathscr{X}) \to \mathscr{L}(\mathscr{X})\colon f \mapsto \underline{T}f := \lim_{i\to+\infty} \Phi_{u_i}f = f_{\lim}. \tag{28}$$

We now first verify that $\underline{T}$ is a lower counting transformation—and therefore also a non-negatively homogeneous transformation. Because $\lim_{i\to+\infty}\sigma(u_i)=0$, there is an $i^\star$ in $\mathbb{N}$ such that $\sigma(u_i)\|\underline{Q}_S\| \leq 2$ for all $i \geq i^\star$. From this and Corollary 36, it follows that Φ_{u_i} is a lower counting transformation for all $i \geq i^\star$. This implies that $\underline{T}$, as defined in Equation (28), is a lower counting transformation as well because the (in)equalities in the conditions (LT1)–(LT4) are preserved under taking limits.

Next, we verify that $\{\Phi_{u_i}\}_{i\in\mathbb{N}}$ converges to $\underline{T}$. To that end, we fix any ε in $\mathbb{R}_{>0}$, and choose some ε' in $\mathbb{R}_{>0}$ such that $3\varepsilon' \leq \varepsilon$. Recall from Lemma 42 that $\{\Phi_{u_i}\}_{i\in\mathbb{N}}$ is a Cauchy sequence. Hence, there is an i_ε in $\mathbb{N}$ such that, for all i, j in $\mathbb{N}$ with $i \geq i_\varepsilon$ and $j \geq i_\varepsilon$,

$$\|\Phi_{u_i} - \Phi_{u_j}\| \leq \varepsilon'. \tag{29}$$

Fix now any i in $\mathbb{N}$ such that $i \geq i_\varepsilon$. From the definition of the norm for non-negatively homogeneous transformations, it follows that there is some f_1 in $\mathscr{L}(\mathscr{X})$ with $\|f_1\| = 1$ such that

$$\|\underline{T} - \Phi_{u_i}\| \leq \|\underline{T}f_1 - \Phi_{u_i}f_1\| + \varepsilon'. \tag{30}$$

Furthermore, due to Equation (28), there is a j in $\mathbb{N}$ such that $j \geq i_\varepsilon$ and $\|\underline{T}f_1 - \Phi_{u_j}f_1\| \leq \varepsilon'$. We now use this and Equation (30), to yield

$$\begin{aligned}
\|\underline{T} - \Phi_{u_i}\| &\leq \|\underline{T}f_1 - \Phi_{u_i}f_1\| + \varepsilon' = \|\underline{T}f_1 - \Phi_{u_j}f_1 + \Phi_{u_j}f_1 - \Phi_{u_i}f_1\| + \varepsilon' \\
&\leq \|\underline{T}f_1 - \Phi_{u_j}f_1\| + \|\Phi_{u_j}f_1 - \Phi_{u_i}f_1\| + \varepsilon' \\
&\leq \|\Phi_{u_j}f_1 - \Phi_{u_i}f_1\| + 2\varepsilon'.
\end{aligned}$$

Finally, we use (NH4) and the fact that $\|f_1\| = 1$, to yield

$$\|\underline{T} - \Phi_{u_i}\| \leq \|\Phi_{u_j} - \Phi_{u_i}\|\|f_1\| + 2\varepsilon' = \|\Phi_{u_j} - \Phi_{u_i}\| + 2\varepsilon' \leq 3\varepsilon' \leq \varepsilon,$$

where the penultimate inequality follows from Equation (29) because $i \geq i_\varepsilon$ and $j \geq i_\varepsilon$, and where the final inequality is precisely our condition on ε'. Because this inequality holds for any $i \geq i_\varepsilon$, and because ε was an arbitrary positive real number, we infer from this that $\lim_{i\to+\infty}\Phi_{u_i} = \underline{T}$, as required. ∎

Our second result establishes that the limit of $\{\Phi_{u_i}\}_{i\in\mathbb{N}}$ is unique, in the sense that it does not depend on the choice of $\{u_i\}_{i\in\mathbb{N}}$. Note the similarity with Proposition 31—that is, [8, Theorem 7.12].

Theorem 45 *Consider a sequence S in $\mathscr{S}_\Lambda$. For any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, there is a unique lower counting transformation $\underline{T}$ such that*

$$(\forall \varepsilon \in \mathbb{R}_{>0})(\exists \delta \in \mathbb{R}_{>0})(\forall u \in \mathscr{U}_{[t,s]}, \sigma(u) \leq \delta)\ \|\underline{T} - \Phi_u\| \leq \varepsilon.$$

Proof Consider any sequence $\{u_i\}_{i\in\mathbb{N}}$ such that $\lim_{i\to+\infty}\sigma(u_i)=0$. Let $\underline{T} := \lim_{i\to+\infty}\Phi_{u_i}$, where this limit exists and is a lower counting transformation due to Theorem 44. We now verify that this lower counting transformation $\underline{T}$ satisfies the condition of the statement. To that end, we fix any ε in $\mathbb{R}_{>0}$, and choose any ε' in $\mathbb{R}_{>0}$ such that $3\varepsilon' \leq \varepsilon$. Additionally, we choose any δ in $\mathbb{R}_{>0}$ such that $2\delta(s-t)\|\underline{Q}_S\|^2 \leq \varepsilon'$. We now proceed in a similar fashion as in the second part of the proof of Theorem 44. Fix any u in $\mathscr{U}_{[t,s]}$ such that $\sigma(u) \leq \delta$. By definition of the norm for non-negatively homogeneous transformations, there is some f_1 in $\mathscr{L}(\mathscr{X})$ with $\|f_1\| = 1$ such that

$$\|\underline{T} - \Phi_u\| \leq \|\underline{T}f_1 - \Phi_uf_1\| + \varepsilon'.$$

Because $\lim_{i\to+\infty}\Phi_{u_i} = \underline{T}$ and $\lim_{i\to+\infty}\sigma(u_i)=0$, there is an i in $\mathbb{N}$ such that $\sigma(u_i) \leq \delta$ and $\|\underline{T} - \Phi_{u_i}\| \leq \varepsilon'$. Observe now that

$$\begin{aligned}
\|\underline{T} - \Phi_u\| &\leq \|\underline{T}f_1 - \Phi_uf_1\| + \varepsilon_1 \leq \|\underline{T}f_1 - \Phi_{u_i}f_1\| + \|\Phi_{u_i}f_1 - \Phi_uf_1\| + \varepsilon' \\
&\leq \|\underline{T} - \Phi_{u_i}\| + \|\Phi_{u_i} - \Phi_u\| + \varepsilon' \leq \|\Phi_{u_i} - \Phi_u\| + 2\varepsilon' \\
&\leq 2\delta(s-t)\|\underline{Q}_S\|^2 + 2\varepsilon' \leq 3\varepsilon' \leq \varepsilon,
\end{aligned}$$

where the second inequality follows from the triangle inequality, the third inequality follows from (NH4) and $\|f_1\| = 1$, the fourth inequality holds because i was fixed in such a way that $\|\underline{T} - \Phi_{u_i}\| \leq \varepsilon'$, the fifth inequality follows from Lemma 41 because $\sigma(u) \leq \delta$ and $\sigma(u_i) \leq \delta$ and the penultimate inequality follows from our condition on δ. Because this inequality holds for any u in $\mathscr{U}_{[t,s]}$ such that $\sigma(u) \leq \delta$, and because ε was an arbitrary positive real number, this verifies the condition of the stated.

Finally, we verify that $\underline{T}$ is unique. To that end, we let $\underline{T}'$ be any lower counting transformation that (also) satisfies the condition of the stated. For any ε in $\mathbb{R}_{>0}$, we then clearly have that there is a u in $\mathscr{U}_{[t,s]}$ such that $\|\underline{T} - \Phi_u\| \leq \varepsilon/2$ and $\|\underline{T}' - \Phi_u\| \leq \varepsilon/2$. Hence, $\|\underline{T} - \underline{T}'\| \leq \|\underline{T} - \Phi_u\| + \|\underline{T}' - \Phi_u\| \leq \varepsilon$. Since ε is an arbitrary positive real number, we conclude from this that $\|\underline{T} - \underline{T}'\| = 0$, which in turn implies that $\underline{T} = \underline{T}'$, as required. ∎

We end with two useful properties of the approximation Φ_u.

Lemma 46 *Consider a sequence $S = \{(\underline{\lambda}_x, \overline{\lambda}_x)\}_{x \in \mathscr{X}}$ in $\mathscr{S}_\Lambda$. Fix an n in $\mathbb{N}$ and, for all i in $\{1, \ldots, n\}$, a Δ_i in $\mathbb{R}_{\geq 0}$ with $\Delta_i \|\underline{Q}_S\| \leq 2$. Then for any f in $\mathscr{L}(\mathscr{X})$ and x in $\mathscr{X}$,*

$$\left[\prod_{i=1}^{n}(I + \Delta_i \underline{Q}_S)f\right](y) = \left[\prod_{i=1}^{n}(I + \Delta_i \underline{Q}_S)f'_x\right](y - x) \qquad \text{for all } y \in \mathscr{X} \text{ with } y \geq x,$$

where $f'_x \colon \mathscr{X} \to \mathbb{R} \colon z \mapsto f'_x(z) := f(x + z)$.

Proof Our proof is one by induction. First, we consider the case $n = 1$. Then clearly

$$[(I + \Delta_1 \underline{Q}_S)f](y) = f(y) + \Delta_1[\underline{Q}_S f](y) = f(y) + \Delta_1 \min_{\lambda \in [\underline{\lambda}_y, \overline{\lambda}_y]} \lambda(f(y+1) - f(y))$$

$$= f'_x(y - x) + \Delta_1 \min_{\lambda \in [\underline{\lambda}_y, \overline{\lambda}_y]} \lambda(f'_x(y+1-x) - f'_x(y-x)) = f'_x(y-x) + \Delta_1[\underline{Q}_S f'_x](y-x)$$

$$= [(I + \Delta_1 \underline{Q}_S)f'_x](y-x).$$

Next, we fix some n in $\mathbb{N}$ with $n \geq 2$ and assume that the stated holds for all n' in $\mathbb{N}$ with $1 \leq n' < n$. We now show that this implies that the stated holds for n as well. Let $g := \prod_{i=2}^{n}(I + \Delta_i \underline{Q}_S)f$. Then

$$\left[\prod_{i=1}^{n}(I + \Delta_i \underline{Q}_S)f\right](y) = [(I + \Delta_1 \underline{Q}_S)g](y) = [(I + \Delta_1 \underline{Q}_S)g'_x](y-x), \tag{31}$$

where we let $g'_x \colon \mathscr{X} \to \mathbb{R} \colon z \mapsto g'_x(z) := g(z+x)$ and where the second equality follows from the induction hypothesis for $n' = 1$. Observe now that, for any z in $\mathscr{X}$,

$$g'_x(z) = g(z+x) = \left[\prod_{i=2}^{n}(I + \Delta_i \underline{Q}_S)f\right](z+x) = \left[\prod_{i=2}^{n}(I + \Delta_i \underline{Q}_S)f'_x\right](z),$$

where the third equality follows from the induction hypothesis for $n' = n - 1$. Since this holds for all z, this implies that $g'_x = \prod_{i=2}^{n}(I + \Delta_i \underline{Q}_S)f'_x$. We now substitute this equality in Equation (31) to obtain the stated:

$$\left[\prod_{i=1}^{n}(I + \Delta_i \underline{Q}_S)f\right](y) = \left[(I + \Delta_1 \underline{Q}_S)\prod_{i=2}^{n}(I + \Delta_i \underline{Q}_S)f'_x\right](y-x) = \left[\prod_{i=1}^{n}(I + \Delta_i \underline{Q}_S)f'_x\right](y-x).$$

∎

Lemma 47 *Consider a sequence $S = \{(\underline{\lambda}_x, \overline{\lambda}_x)\}_{x \in \mathscr{X}}$ in $\mathscr{S}_\Lambda$. Fix some t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, a sequence $u = t_0, \ldots, t_n$ in $\mathscr{U}_{[t,s]}$ with $\sigma(u)\|\underline{Q}_S\| \leq 2$ and an f in $\mathscr{L}(\mathscr{X})$. Then for any x, y in $\mathscr{X}$ with $y \geq x + n$,*

$$[\Phi_u f](x) = [\Phi_u(\mathbb{I}_{\leq y} f + f(y)\mathbb{I}_{>y})](x).$$

Proof In order to simplify our notation, for any y in $\mathscr{X}$ we let $f_y := \mathbb{I}_{\leq y}f + f(y)\mathbb{I}_{>y}$. Observe that f_y is eventually constant starting from y by construction, and that $f_y(x) = f(x)$ for all x in $\mathscr{X}$ such that $x \leq y$.

We first consider the case $t = s$. In this case, we have $n = 0$ and $\Phi_u = I$. Fix any $y \geq x + n = x$. We immediately see that

$$[\Phi_u f](x) = [If](x) = f(x) = \mathbb{I}_{\leq y}(x)f(x) + \mathbb{I}_{>y}(x)f(y) = f_y(x) = [If_y](x) = [\Phi_u f_y](x) = [\Phi_u(\mathbb{I}_{\leq y}f + f(y)\mathbb{I}_{>y})](x),$$

as required.

Next, we consider the case $t < s$. We will prove the stated by induction. Assume first that $n = 1$. Fix any $y \geq x + n = x + 1$. Observe that

$$[\Phi_u f](x) = [(I + \Delta_1 \underline{Q}_S)f](x) = f(x) + \Delta_1[\underline{Q}_S f](x).$$

Recall from the beginning of this proof that $f(x) = f_y(x)$ by construction. Furthermore, because $\mathbb{I}_{\leq x+1}f = \mathbb{I}_{\leq x+1}f_y$, it follows from Proposition 33 (GP6) that $[\underline{Q}_S f](x) = [\underline{Q}_S(\mathbb{I}_{\leq x+1}f)](x) = [\underline{Q}_S(\mathbb{I}_{\leq x+1}f_y)](x) = [\underline{Q}_S f_y](x)$. Hence,

$$[\Phi_u f](x) = f(x) + \Delta_1[\underline{Q}_S f](x) = f_y(x) + \Delta_1[\underline{Q}_S f_y](x) = [(I + \Delta_1 \underline{Q}_S)f_y](x) = [\Phi_u f_y](x),$$

as required.

Fix now any n in $\mathbb{N}$ with $n \geq 2$, and assume that the stated holds for all $1 \leq n' < n$. We now show that this implies the stated for n. Fix any y in $\mathscr{X}$ with $y \geq x + n$. Let $v := t_1, \ldots, t_n$. Then

$$[\Phi_u f](x) = [(I + \Delta_1 \underline{Q}_S)\Phi_v f](x) = [(I + \Delta_1 \underline{Q}_S)g](x) = [(I + \Delta_1 \underline{Q}_S)(\mathbb{I}_{\leq x+1}g + \mathbb{I}_{>x+1}g(x+1))](x), \tag{32}$$

where we let $g := \Phi_v f$ and where the final equality follows from the induction hypothesis with $n' = 1$.

It now follows from the induction hypothesis with $n' = n - 1$ that, for any z in $\mathscr{X}$ such that $y \geq z + n - 1$ (or equivalently, $z \leq y - n + 1$), $[\Phi_v f](z) = [\Phi_v f_y](z)$. As furthermore $x \leq y - n$ by assumption, we conclude from this that, for any z in $\mathscr{X}$ such that $z \leq x + 1 \leq y - n + 1$, $[\Phi_v f](z) = [\Phi_v f_y](z)$. Consequently,

$$\mathbb{I}_{\leq x+1}g + \mathbb{I}_{>x+1}g(x+1) = \mathbb{I}_{\leq x+1}(\Phi_v f) + \mathbb{I}_{>x+1}[\Phi_v f](x+1) = \mathbb{I}_{\leq x+1}(\Phi_v f_y) + \mathbb{I}_{>x+1}[\Phi_v f_y](x+1).$$

We now substitute this equality in Equation (32), to yield

$$[\Phi_u f](x) = [(I + \Delta_1 \underline{Q}_S)(\mathbb{I}_{\leq x+1}\Phi_v f_y + \mathbb{I}_{>x+1}[\Phi_v f_y](x+1))](x).$$

We now invoke the induction hypothesis with $n' = 1$ for the second time, to yield

$$[\Phi_u f](x) = [(I + \Delta_1 \underline{Q}_S)(\mathbb{I}_{\leq x+1}\Phi_v f_y + \mathbb{I}_{>x+1}[\Phi_v f_y](x+1))](x) = [(I + \Delta_1 \underline{Q}_S)\Phi_v f_y](x) = [\Phi_u f_y](x).$$

∎

C.3. The Corresponding Semi-Group of Lower Counting Transformations

Let $S = \{(\underline{\lambda}_x, \overline{\lambda}_x)\}_{x \in \mathscr{X}}$ be a sequence in $\mathscr{S}_\Lambda$. Due to Theorems 44 and 45, for any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, we can uniquely define the corresponding lower counting transformation

$$\underline{T}_t^s := \lim_{\sigma(u) \to 0} \{\Phi_u : u \in \mathscr{U}_{[t,s]}\}.$$

As explained right after Equation (16), this unconventional notation for the limit is used to emphasise that the limit does *not* depend on the chosen sequence $\{u_i\}_{i \in \mathbb{N}}$ in $\mathscr{U}_{[t,s]}$ so long as $\lim_{i \to +\infty} \sigma(u_i) = 0$.

This way, we have defined an entire family $\{\underline{T}_t^s : t, s \in \mathbb{R}_{\geq 0}, t \leq s\}$ of lower counting transformations. The following result establishes that this family is a time-homogeneous semi-group.

Proposition 48 *Consider a sequence S in $\mathscr{S}_\Lambda$. Then*

(i) $\underline{T}_t^t = I$ *for all t in $\mathbb{R}_{\geq 0}$;*

(ii) $\underline{T}_t^s = \underline{T}_t^r \underline{T}_r^s$ *for all t, r, s in $\mathbb{R}_{\geq 0}$ with $t \leq r \leq s$;*

(iii) $\underline{T}_t^s = \underline{T}_0^{s-t}$ *for all t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$.*

Proof The proofs of these properties are almost entirely the same as their counterparts in [8].

(i) This is an immediate consequence of Theorem 45 because $\mathscr{U}_{[t,t]} = \{t\}$, $\sigma(t) = 0$ and $\Phi_t = I$.

(ii) Fix any arbitrary ε in $\mathbb{R}_{>0}$. By Theorem 45, there are sequences u_1 in $\mathscr{U}_{[t,r]}$ and u_2 in $\mathscr{U}_{[r,s]}$ such that (i) $\|\underline{T}_t^r - \Phi_{u_1}\| \leq \varepsilon/3$ and $\sigma(u_1)\|\underline{Q}_S\| \leq 2$; (ii) $\|\underline{T}_r^s - \Phi_{u_2}\| \leq \varepsilon/3$ and $\sigma(u_2)\|\underline{Q}_S\| \leq 2$; and (iii) $\|\underline{T}_t^s - \Phi_u\| \leq \varepsilon/3$ and $\sigma(u)\|\underline{Q}_S\| \leq 2$, where $u := u_1 \cup u_2$ is an element of $\mathscr{U}_{[t,s]}$. Observe that

$$\|\underline{T}_t^s - \underline{T}_t^r \underline{T}_r^s\| = \|\underline{T}_t^s - \Phi_u + \Phi_{u_1}\Phi_{u_2} - \underline{T}_t^r \underline{T}_r^s\| \leq \|\underline{T}_t^s - \Phi_u\| + \|\Phi_{u_1}\Phi_{u_2} - \underline{T}_t^r \underline{T}_r^s\|$$

$$\leq \|\underline{T}_t^s - \Phi_u\| + \|\Phi_{u_1} - \underline{T}_t^r\| + \|\Phi_{u_2} - \underline{T}_r^s\| \leq 3\frac{\varepsilon}{3} = \varepsilon,$$

where the first inequality follows from the triangle inequality and the second inequality follows from Lemma 26 and Corollary 36. Because ε was any arbitrary positive real number, we conclude from this inequality that the stated holds.

(iii) Fix any sequence $\{u_i\}_{i\in\mathbb{N}}$ in $\mathscr{U}_{[t,s]}$ such that $\lim_{i\to+\infty} \sigma(u_i) = 0$. For any i in $\mathbb{N}$, we define $u_i' := t_0 - t, t_1 - t, \ldots, t_n - t$, with $u_i = t_0, \ldots, t_n$. Observe that, by construction, $\Phi_{u_i} - \Phi_{u_i'}$ for all i in $\mathbb{N}$. Because furthermore $\lim_{i\to+\infty} \Phi_{u_i} = \underline{T}_t^s$ and $\lim_{i\to+\infty} \Phi_{u_i'} = \underline{T}_0^{s-t}$ due to Theorems 44 and 45, the stated now follows immediately. ∎

We conclude this section of the Appendix with some technical results regarding the family $\{\underline{T}_t^s : t, s \in \mathbb{R}_{\geq 0}, t \leq s\}$ of lower counting transformations.

Lemma 49 *Consider a sequence S in $\mathscr{S}_\Lambda$. Then for all t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$,*

(i) $\|\underline{T}_t^s - (I + (s-t)\underline{Q}_S)\| \leq (s-t)^2\|\underline{Q}_S\|^2$;

(ii) $\|\underline{T}_t^s - I\| \leq (s-t)\|\underline{Q}_S\|$.

Proof

(i) Fix any arbitrary ε in $\mathbb{R}_{>0}$. By Theorem 45, there is a sequence $u = t_0, \ldots, t_n$ in $\mathscr{U}_{[t,s]}$ such that $\|\underline{T}_t^s - \Phi_u\| \leq \varepsilon$ and $\sigma(u)\|\underline{Q}_S\| \leq 2$. Observe that

$$\|\underline{T}_t^s - (I + (s-t)\underline{Q}_S)\| \leq \|\underline{T}_t^s - \Phi_u\| + \|\Phi_u - (I + (s-t)\underline{Q}_S)\| \leq \varepsilon + \|\underline{Q}_S\|^2 \sum_{i=1}^n \Delta_i \sum_{j=i+1}^n \Delta_j \leq \varepsilon + (s-t)^2\|\underline{Q}_S\|^2,$$

where the first inequality follows from the triangle inequality and the second inequality follows from Lemma 37. The stated now holds because this inequality holds for any arbitrary positive real number ε.

(ii) Fix any n in $\mathbb{N}$ such that $(s-t)\|\underline{Q}_S\| \leq 2n$, and let $\Delta := (s-t)/n$. Observe that

$$\|\underline{T}_0^\Delta - I\| = \|\underline{T}_0^\Delta - (I + \Delta\underline{Q}_S) + \Delta\underline{Q}_S\| \leq \|\underline{T}_0^\Delta - (I + \Delta\underline{Q}_S)\| + \Delta\|\underline{Q}_S\| \leq \Delta^2\|\underline{Q}_S\|^2 + \Delta\|\underline{Q}_S\|,$$

where the first inequality follows from the triangle inequality and the second inequality follows from (i). By repeatedly applying Proposition 48 (ii) and Proposition 48 (iii), we obtain that $\underline{T}_t^s = (\underline{T}_0^\Delta)^n$. We combine our findings, to yield

$$\|\underline{T}_t^s - I\| = \left\|(\underline{T}_0^\Delta)^n - I^n\right\| \leq n\|\underline{T}_0^\Delta - I\| \leq n\Delta^2\|\underline{Q}_S\|^2 + n\Delta\|\underline{Q}_S\| = \frac{(s-t)^2}{n}\|\underline{Q}_S\|^2 + (s-t)\|\underline{Q}_S\|,$$

where the first inequality follows from Lemma 26. The stated now follows if we take the limit for n going to $+\infty$. ∎

The second of these technical results establishes an upper bound on the error made by approximating $\underline{T}_t^s$ by Φ_u.

Lemma 50 *Fix a sequence S in $\mathscr{S}_\Lambda$, some t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$ and a sequence u in $\mathscr{U}_{[t,s]}$. If $\sigma(u)\|\underline{Q}_S\| \leq 2$, then*

$$\|\underline{T}_t^s - \Phi_u\| \leq \sigma(u)(s-t)\|\underline{Q}_S\|^2.$$

Proof Our proof is entirely similar to that of [8, Lemma E.8]. Fix any ε in $\mathbb{R}_{>0}$. By Theorem 45, there is a u_ε in $\mathscr{U}_{[t,s]}$ such that $u \subseteq u_\varepsilon$ and $\|\underline{T}_t^s - \Phi_{u_\varepsilon}\| \leq \varepsilon$. From this, the triangle inequality and Corollary 40, it follows that

$$\|\underline{T}_t^s - \Phi_u\| = \|\underline{T}_t^s - \Phi_{u_\varepsilon} + \Phi_{u_\varepsilon} - \Phi_u\| \leq \|\underline{T}_t^s - \Phi_{u_\varepsilon}\| + \|\Phi_{u_\varepsilon} - \Phi_u\| \leq \varepsilon + \sigma(u)(t-s)\|\underline{Q}_S\|^2.$$

The stated now follows because ε is an arbitrary positive real number. ∎

Our next technical result can be interpreted as dealing with the "time-derivative" of $\underline{T}_t^s$, as is explained by Krak et al. [8, Right after Proposition 7.15].

Lemma 51 *Consider a sequence S in $\mathscr{S}_\Lambda$. Then for any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$,*

$$(\forall \varepsilon \in \mathbb{R}_{>0})(\exists \delta \in \mathbb{R}_{>0})(\forall \Delta \in \mathbb{R}, 0 < |\Delta| < \delta, 0 \leq t + \Delta \leq s) \left\| \frac{T_{t+\Delta}^s - T_t^s}{\Delta} + \underline{Q}_S \underline{T}_t^s \right\| \leq \varepsilon.$$

and

$$(\forall \varepsilon \in \mathbb{R}_{>0})(\exists \delta \in \mathbb{R}_{>0})(\forall \Delta \in \mathbb{R}, 0 < |\Delta| < \delta, t \leq s + \Delta) \left\| \frac{T_t^{s+\Delta} - T_t^s}{\Delta} - \underline{Q}_S \underline{T}_t^s \right\| \leq \varepsilon.$$

Proof The proof is just the proof of Krak et al. for [8, Proposition 7.15] with some obvious modifications. Fix any ε in $\mathbb{R}_{>0}$, and fix any δ in $\mathbb{R}_{>0}$ such that $2\delta\|\underline{Q}_S\|^2 \leq \varepsilon$. Consider now any Δ in $\mathbb{R}$ such that $0 < |\Delta| < \delta$ and $0 \leq t + \Delta \leq s$. If we let $t^\star := \max\{t, t+\Delta\}$, then

$$\|\underline{T}_{t+\Delta}^s - \underline{T}_t^s + \Delta \underline{Q}_S \underline{T}_t^s\| = \|\underline{T}_{t^\star}^s - \underline{T}_{t^\star - |\Delta|}^s + |\Delta| \underline{Q}_S \underline{T}_t^s\| = \|\underline{T}_{t^\star}^s - \underline{T}_{t^\star - |\Delta|}^{t^\star} \underline{T}_{t^\star}^s + |\Delta| \underline{Q}_S \underline{T}_t^{t^\star} \underline{T}_{t^\star}^s\|,$$

where the last equality follows Proposition 48 (ii) because $t \leq t^\star \leq s$. We now use (NH5) and (LT10), to yield

$$\|\underline{T}_{t+\Delta}^s - \underline{T}_t^s + \Delta \underline{Q}_S \underline{T}_t^s\| \leq \|I - \underline{T}_{t^\star - |\Delta|}^{t^\star} + |\Delta| \underline{Q}_S \underline{T}_t^{t^\star}\| \|\underline{T}_{t^\star}^s\| \leq \|I - \underline{T}_{t^\star - |\Delta|}^{t^\star} + |\Delta| \underline{Q}_S \underline{T}_t^{t^\star}\|.$$

Further manipulations now yield

$$
\begin{aligned}
\|\underline{T}_{t+\Delta}^s - \underline{T}_t^s + \Delta \underline{Q}_S \underline{T}_t^s\| &\leq \|I - \underline{T}_{t^\star - |\Delta|}^{t^\star} + |\Delta| \underline{Q}_S - |\Delta| \underline{Q}_S + |\Delta| \underline{Q}_S \underline{T}_t^{t^\star}\| \\
&\leq \|I + |\Delta| \underline{Q}_S - \underline{T}_{t^\star - |\Delta|}^{t^\star}\| + \||\Delta| \underline{Q}_S - |\Delta| \underline{Q}_S \underline{T}_t^{t^\star}\| \\
&\leq \|I + |\Delta| \underline{Q}_S - \underline{T}_{t^\star - |\Delta|}^{t^\star}\| + |\Delta| \|\underline{Q}_S\| \|\underline{T}_t^{t^\star} - I\| \\
&\leq \Delta^2 \|\underline{Q}_S\|^2 + |\Delta| \|\underline{Q}_S\| \|\underline{T}_t^{t^\star} - I\| \leq \Delta^2 \|\underline{Q}_S\|^2 + |\Delta|(t^\star - t) \|\underline{Q}_S\| \\
&\leq 2\Delta^2 \|\underline{Q}_S\|^2,
\end{aligned}
$$

where the third inequality follows from Corollary 38, the fourth inequality follows from Lemma 49 (i), the fifth inequality follows from Lemma 49 (ii) and the final inequality follows from the fact that $0 \leq t^\star - t \leq |\Delta|$. From this inequality, it now follows that

$$\left\| \frac{T_{t+\Delta}^s - T_t^s}{\Delta} + \underline{Q}_S \underline{T}_t^s \right\| = \frac{1}{|\Delta|} \|\underline{T}_{t+\Delta}^s - \underline{T}_t^s + \Delta \underline{Q}_S \underline{T}_t^s\| \leq 2|\Delta| \|\underline{Q}_S\|^2 \leq 2\delta \|\underline{Q}_S\|^2 \leq \varepsilon,$$

which proves the first part of the stated.

The second part of the stated follows from the first part. To see this, we fix any ε, τ in $\mathbb{R}_{>0}$, and let $t' := t + \tau$ and $s' := s + \tau$. It follows from the first part of the stated that there is some δ' in $\mathbb{R}_{>0}$ such that

$$(\forall \Delta' \in \mathbb{R}, 0 < |\Delta'| < \delta', 0 \leq t' + \Delta' \leq s') \left\| \frac{T_{t'+\Delta'}^{s'} - T_{t'}^{s'}}{\Delta'} + \underline{Q}_S \underline{T}_{t'}^{s'} \right\| \leq \varepsilon. \tag{33}$$

We now let $\delta := \min\{\delta', \tau\}$, and fix any Δ in $\mathbb{R}$ such that $0 < |\Delta| < \delta$. Observe that

$$\left\| \frac{T_t^{s+\Delta} - T_t^s}{\Delta} - \underline{Q}_S \underline{T}_t^s \right\| = \left\| \frac{T_{t'}^{s'+\Delta} - T_{t'}^{s'}}{\Delta} - \underline{Q}_S \underline{T}_{t'}^{s'} \right\| = \left\| \frac{T_{t'-\Delta}^{s'} - T_{t'}^{s'}}{\Delta} - \underline{Q}_S \underline{T}_{t'}^{s'} \right\| = \left\| \frac{T_{t'-\Delta}^{s'} - T_{t'}^{s'}}{-\Delta} + \underline{Q}_S \underline{T}_{t'}^{s'} \right\|,$$

where the first and second equality follow from Proposition 48 (iii). Since furthermore $t' - \Delta = t + \tau - \Delta \geq t \geq 0$ and $t' - \Delta = t + \tau - \Delta \leq s + \tau = s'$, it follows from Equation (33) with $\Delta' = -\Delta$ that

$$\left\| \frac{\underline{T}_t^{s+\Delta} - \underline{T}_t^s}{\Delta} - \underline{Q}_S \underline{T}_t^s \right\| = \left\| \frac{\underline{T}_{t'}^{s'} - \underline{T}_{t'}^{s'}}{-\Delta} + \underline{Q}_S \underline{T}_{t'}^{s'} \right\| \leq \varepsilon,$$

as required. ∎

We now use Proposition 51 to establish the limit behaviour of $[\underline{T}_t^{t+\Delta} \mathbb{I}_{\geq x+2}](x)$ and $[\underline{T}_{t-\Delta}^t \mathbb{I}_{\geq x+2}](x)$ for $\Delta \to 0^+$.

Lemma 52 *Consider a sequence S in $\mathscr{S}_\Lambda$. Then for any t in $\mathbb{R}_{\geq 0}$ and x in $\mathscr{X}$,*

$$\lim_{\Delta \to 0^+} \frac{[\underline{T}_t^{t+\Delta} \mathbb{I}_{\geq x+2}](x)}{\Delta} = 0$$

and, if $t > 0$,

$$\lim_{\Delta \to 0^+} \frac{[\underline{T}_{t-\Delta}^t \mathbb{I}_{\geq x+2}](x)}{\Delta} = 0.$$

Proof Fix any ε in $\mathbb{R}_{>0}$. From Lemma 51 (with $s = t$), we know that there is a δ in $\mathbb{R}_{>0}$ such that

$$(\forall \Delta \in \mathbb{R}_{>0}, \Delta < \delta) \; \left\| \frac{\underline{T}_t^{t+\Delta} - \underline{T}_t^t}{\Delta} - \underline{Q}_S \underline{T}_t^t \right\| \leq \varepsilon. \tag{34}$$

Fix any Δ in $\mathbb{R}_{>0}$ such that $\Delta < \delta$, and observe that

$$\left| \frac{[\underline{T}_t^{t+\Delta} \mathbb{I}_{\geq x+2}](x) - [\underline{T}_t^t \mathbb{I}_{\geq x+2}](x)}{\Delta} - [\underline{Q}_S \underline{T}_t^t \mathbb{I}_{\geq x+2}](x) \right| \leq \left\| \frac{\underline{T}_t^{t+\Delta} \mathbb{I}_{\geq x+2} - \underline{T}_t^t \mathbb{I}_{\geq x+2}}{\Delta} - \underline{Q}_S \underline{T}_t^t \mathbb{I}_{\geq x+2} \right\|$$

$$\leq \left\| \frac{\underline{T}_t^{t+\Delta} - \underline{T}_t^t}{\Delta} - \underline{Q}_S \underline{T}_t^t \right\| \|\mathbb{I}_{\geq x+2}\| = \left\| \frac{\underline{T}_t^{t+\Delta} - \underline{T}_t^t}{\Delta} - \underline{Q}_S \underline{T}_t^t \right\|$$

$$\leq \varepsilon,$$

where the first inequality follows from the definition of the supremum norm, the second inequality follows from (NH4), the equality holds because $\|\mathbb{I}_{\geq x+2}\| = 1$ and the final inequality follows from Equation (34) because $0 < \Delta < \delta$. Next, we take a closer look at the terms in the absolute value on the left hand side of the above inequality: (a) it follows from (LT3) that $[\underline{T}_t^{t+\Delta} \mathbb{I}_{\geq x+2}](x) \geq \inf \mathbb{I}_{\geq x+2} = 0$; (b) it follows from Proposition 48 (i) that $[\underline{T}_t^t \mathbb{I}_{\geq x+2}](x) = [I \mathbb{I}_{\geq x+2}](x) = \mathbb{I}_{\geq x+2}(x) = 0$; and (c) it follows from Proposition 48 (i) and Equation (19) that $[\underline{Q}_S \underline{T}_t^t \mathbb{I}_{\geq x+2}](x) = [\underline{Q}_S \mathbb{I}_{\geq x+2}](x) = 0$. Consequently,

$$0 \leq \frac{[\underline{T}_t^{t+\Delta} \mathbb{I}_{\geq x+2}](x)}{\Delta} = \left| \frac{[\underline{T}_t^{t+\Delta} \mathbb{I}_{\geq x+2}](x) - [\underline{T}_t^t \mathbb{I}_{\geq x+2}](x)}{\Delta} - [\underline{Q}_S \underline{T}_t^t \mathbb{I}_{\geq x+2}](x) \right| \leq \varepsilon.$$

Since this holds for all Δ in $\mathbb{R}_{>0}$ such that $\Delta < \delta$, and because ε was an arbitrary positive real number, we have shown that

$$\lim_{\Delta \to 0^+} \frac{[\underline{T}_t^{t+\Delta} \mathbb{I}_{\geq x+2}](x)}{\Delta} = 0,$$

which is the first limit of the stated. The second limit of the stated follows from the the first limit and Proposition 48 (iii). ∎

Finally, we conclude this section with a useful "translation-invariance" property.

Lemma 53 *Consider a sequence S in $\mathscr{S}_\Lambda$. For any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, f in $\mathscr{L}(\mathscr{X})$ and x in $\mathscr{X}$,*

$$[\underline{T}_t^s f](y) = [\underline{T}_t^s f_x'](y - x) \quad \text{for all } y \in \mathscr{X} \text{ with } y \geq x,$$

where $f_x' \colon \mathscr{X} \to \mathbb{R} \colon z \mapsto f_x'(z) := f(x+z)$.

Proof Fix any ε in $\mathbb{R}_{>0}$, and choose some ε' in $\mathbb{R}_{>0}$ such that $2\varepsilon'\|f\| \leq \varepsilon$. By Theorem 45, there is a u in $\mathscr{U}_{[t,s]}$ such that

$$\|\underline{T}_t^s - \Phi_u\| \leq \varepsilon'.$$

Observe now that

$$|[\underline{T}_t^s f](y) - [\Phi_u f](y)| \leq \|\underline{T}_t^s f - \Phi_u f\| \leq \|\underline{T}_t^s - \Phi_u\|\|f\| \leq \varepsilon'\|f\|,$$

where the second inequality follows from (NH4). Similarly, we find that

$$\left|[\underline{T}_t^s f_x'](y-x) - [\Phi_u f_x'](y-x)\right| \leq \|\underline{T}_t^s - \Phi_u\|\|f_x'\| \leq \varepsilon'\|f_x'\| \leq \varepsilon'\|f\|,$$

where for the penultimate inequality we have used the obvious inequality $\|f_x'\| \leq \|f\|$. Using Lemma 46, we furthermore find that $[\Phi_u f](y) = [\Phi_u f_x'](y-x)$. We now combine our two previous findings, to yield

$$\begin{aligned}
\left|[\underline{T}_t^s f](y) - [\underline{T}_t^s f_x'](y-x)\right| &= \left|[\underline{T}_t^s f](y) - [\Phi_u f](y) + [\Phi_u f_x'](y-x) - [\underline{T}_t^s f_x'](y-x)\right| \\
&\leq \left|[\underline{T}_t^s f](y) - [\Phi_u f](y)\right| + \left|[\Phi_u f_x'](y-x) - [\underline{T}_t^s f_x'](y-x)\right| \\
&\leq \varepsilon'\|f\| + \varepsilon'\|f\| \leq \varepsilon,
\end{aligned}$$

where the final inequality is precisely our condition on ε'. Since ε was an arbitrary positive real number, this proves the stated. ∎

Appendix D. Linear Transformations

Another important type of transformations on $\mathscr{L}(\mathscr{Y})$ are the linear ones. A transformation A on $\mathscr{L}(\mathscr{Y})$ is *linear* if it is (i) homogeneous, in the sense that $A(\mu f) = \mu A f$ for all f in $\mathscr{L}(\mathscr{Y})$ and μ in $\mathbb{R}$; and (ii) additive, in the sense that $A(f+g) = Af + Ag$ for all f, g in $\mathscr{L}(\mathscr{Y})$. Observe that a linear transformation is always non-negatively homogeneous, and conversely, that a non-negatively homogeneous transformation is linear if and only if it is additive.

The special case that $\mathscr{Y}$ is finite deserves some additional attention. It is well-known that in this case, the linear transformation A can be identified with a $|\mathscr{Y}| \times |\mathscr{Y}|$ matrix, the (x,y)-component of which is equal to $A(x,y) := [A\mathbb{I}_y](x)$. If convenient, we will sometimes prefer the matrix interpretation over the transformation interpretation. It is also well-known that if $\mathscr{Y}$ is finite,

$$\|A\| = \max\left\{\sum_{y \in \mathscr{Y}} |A(x,y)| : x \in \mathscr{Y}\right\}. \tag{35}$$

D.1. Linear Transition Transformations

If a lower transition (counting) transformation $\underline{T}$ is linear and not just super-additive, in the sense that the inequality in (LT2) is actually an equality, then we will call it a linear transition (counting) transformation. More formally, we have the following definition.

Definition 54 *A* linear transition transformation *is any transformation $T \colon \mathscr{L}(\mathscr{Y}) \to \mathscr{L}(\mathscr{Y})$ such that*

T1. $T(\mu f) = \mu T f$ *for all f in $\mathscr{L}(\mathscr{Y})$ and μ in $\mathbb{R}$;* *[homogeneity]*

T2. $T(f+g) = Tf + Tg$ *for all f, g in $\mathscr{L}(\mathscr{Y})$;* *[additivity]*

T3. $Tf \geq \inf f$ *for all f in $\mathscr{L}(\mathscr{Y})$.* *[bound]*

A linear counting transformation *is a linear transition transformation T with*

T4. $[Tf](x) = [T(\mathbb{I}_{\geq x} f)](x)$ *for all f in $\mathscr{L}(\mathscr{Y})$ and x in $\mathscr{Y}$.*

We now state some useful properties of linear counting transformations. The first result establishes some basic properties of transition/counting transformations that follow almost immediately from (T1)–(T4).

Lemma 55 *Consider a linear transition transformation T. Then*

T5. $\inf f \leq Tf \leq \sup f$ *for all f in $\mathscr{L}(\mathscr{X})$;*

T6. $T\mu = \mu$ *for all μ in $\mathbb{R}$;*

T7. $Tf \le Tg$ *for all f, g in $\mathscr{L}(\mathscr{Y})$ such that $f \le g$.*

If T is a linear counting transformation, then

T8. $[T\mathbb{I}_y](x) = 0$ *for all x, y in $\mathscr{X}$ such that $y < x$.*

Proof Properties (T5)–(T7) follow immediately from Lemma 24 (LT5)–(LT8). Property (T8) follows from (T4) with $f = \mathbb{I}y$ and (T6), as $\mathbb{I}_{\ge x}f = \mathbb{I}_{\ge x}\mathbb{I}_y = 0$. ∎

The second result is a specialisation of Lemma 25

Lemma 56 *For any two linear transition (counting) transformations T_1 and T_2, their composition T_1T_2 is again a linear transition (counting) transformation.*

Proof The linearity—that is, (T1) and (T2)—of T_1 and T_2 implies the linearity of their composition T_1T_2, as one can easily verify. Because T_1 and T_2 are lower transition transformations (as they are linear ones by assumption), it follows from Lemma 25 that their composition T_1T_2 satisfies (LT3), which is equivalent to (T3). Similarly, the composition T_1T_2 of the two linear counting transformations satisfies (LT4), which is equivalent to (T4). ∎

The final general result will play an important role in the proof of Lemma 78 further on.

Lemma 57 *Consider two linear counting transformations T_1 and T_2. Then for all x, y in $\mathscr{X}$,*

$$[T_1T_2\mathbb{I}_y](x) = \begin{cases} \sum_{z=x}^{y}[T_1\mathbb{I}_z](x)[T_2\mathbb{I}_y](z) & \text{if } x \le y, \\ 0 & \text{otherwise}. \end{cases}$$

Proof We first consider the case $y < x$. By Lemma 56 and (T4),

$$[T_1T_2\mathbb{I}_y](x) = [T_1T_2(\mathbb{I}_{\ge x}\mathbb{I}_y)](x) = [T_1T_20](x) = 0,$$

where the last equality follows from (T1).

Next, we consider the case $y \ge x$. By (T4),

$$[T_1T_2\mathbb{I}_y](x) = [T_1(\mathbb{I}_{\ge x}(T_2\mathbb{I}_y))](x). \tag{36}$$

Fix any z in $\mathscr{X}$, and consider $\mathbb{I}_{\ge x}(z)[T_2\mathbb{I}_y](z)$. If $z < x$, then $\mathbb{I}_{\ge x}(z)[T_2\mathbb{I}_y](z) = 0$. If $z > y$, then

$$\mathbb{I}_{\ge x}(z)[T_2\mathbb{I}_y](z) = [T_2\mathbb{I}_y](z) = [T_2(\mathbb{I}_{\ge z}\mathbb{I}_y)](z) = [T_20](z) = 0,$$

where the second equality follows from (T4) and the final equality follows from (T1). From this, we conclude that

$$\mathbb{I}_{\ge x}(T_2\mathbb{I}_y) = \sum_{z=x}^{y}[T_2\mathbb{I}_y](z)\mathbb{I}_z.$$

We now substitute this equality in Equation (36), to yield

$$[T_1T_2\mathbb{I}_y](x) = [T_1(\mathbb{I}_{\ge x}(T_2\mathbb{I}_y))](x) = \left[T_1\left(\sum_{z=x}^{y}[T_2\mathbb{I}_y](z)\mathbb{I}_z\right)\right](x) = \sum_{z=x}^{y}[T_1([T_2\mathbb{I}_y](z)\mathbb{I}_z)](x) = \sum_{z=x}^{y}[T_2\mathbb{I}_y](z)[T_1\mathbb{I}_z](x),$$

using (T2) and (T1) for the final two equalities. ∎

28

D.2. Linear Generalised Poisson Generators And The Semi-Groups They Induce

We now follow the same pattern as we did in Appendix C, but the with *linear* transformations: we introduce linear generalised Poisson generators and subsequently show that these transformations generate a family of linear counting transformations. With any sequence $S := \{\lambda_x\}_{x \in \mathscr{X}}$ in $\Lambda = [\underline{\lambda}, \overline{\lambda}]$, we associate the operator Q_S defined by

$$[Q_S f](x) := \lambda_x f(x+1) - \lambda_x f(x) \quad \text{for all } x \in \mathscr{X}, f \in \mathscr{L}(\mathscr{X}).$$

Observe that Q_S is indeed a linear generalised Poisson generator. That it is linear follows immediately from its definition. That is is a generalised Poisson generator follows from the fact that

$$Q_S = \underline{Q}_{S'} \qquad \text{with } S = \{\lambda_x\}_{x \in \mathscr{X}} \text{ and } S' = \{(\lambda_x, \lambda_x)\}_{x \in \mathscr{X}}. \tag{37}$$

This relation allows us to immediately establish the following result.

Corollary 58 *Consider a sequence $S := \{\lambda_x\}_{x \in \mathscr{X}}$ in $\Lambda = [\underline{\lambda}, \overline{\lambda}]$. Then*

$$\|Q_S\| = 2 \sup\{\lambda_x \colon x \in \mathscr{X}\}.$$

Proof This is a simple corollary of Equation (37) and Lemma 34. ∎

We now establish that the linear generalised Poisson generator Q_S generates a family of linear counting transformations. In essence, we simply combine the results of Appendix C.2 with Equation (37).

Corollary 59 *Consider a sequence $S := \{\lambda_x\}_{x \in \mathscr{X}}$ in $\Lambda = [\underline{\lambda}, \overline{\lambda}]$. Then for any Δ in $\mathbb{R}_{\geq 0}$ with $\Delta \|Q_S\| \leq 2$, $(I + \Delta Q_S)$ is a linear counting transformation.*

Proof It follows from Equation (37) and Lemma 35 that $(I + \Delta Q_S)$ is a lower counting transformation. That it is furthermore linear follows directly from the linearity of I and Q_S. ∎

Because the linear generalised Poisson generator Q_S is a generalised Poisson generator, we can use here use the notation Φ_u as introduced in Equation (23) as well; we here simply replace $\underline{Q}_S$ by $Q_S = \underline{Q}_{S'}$ in the definition.

Corollary 60 *Fix a sequence $S = \{\lambda_x\}_{x \in \mathscr{X}}$ in Λ, some t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$ and a sequence $u = t_0, \ldots, t_n$ in $\mathscr{U}_{[t,s]}$. If $\sigma(u)\|Q_S\| \leq 2$, then Φ_u, as defined in Equation (23), is a linear counting transformation.*

Proof Follows immediately from Lemma 56 and Corollary 59. ∎

Corollary 61 *Consider a sequence $S := \{\lambda_x\}_{x \in \mathscr{X}}$ in $\Lambda = [\underline{\lambda}, \overline{\lambda}]$, and fix some t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$. For any sequence $\{u_i\}_{i \in \mathbb{N}}$ in $\mathscr{U}_{[t,s]}$ such that $\lim_{i \to +\infty} \sigma(u_i) = 0$, the corresponding sequence $\{\Phi_{u_i}\}_{i \in \mathbb{N}}$ converges to a linear counting transformation.*

Proof Recall from Equation (37) that Q_S is equal to the generalised Poisson generator $\underline{Q}_{S'}$ associated with the sequence $S' = \{(\lambda_x, \lambda_x)\}_{x \in \mathscr{X}}$. It therefore follows from Theorem 44 (with the sequence $S' = \{(\lambda_x, \lambda_x)\}_{x \in \mathscr{X}}$) that the corresponding sequence $\{\Phi_{u_i}\}_{i \in \mathbb{N}}$ converges to a lower counting transformation. Hence, what remains for us to verify is that this limit is a linear counting transformation. To that end, we observe that, because $\lim_{i \to +\infty} \sigma(u_i) = 0$, there is a $i^\star$ in $\mathbb{N}$ such that $\sigma(u_i)\|Q_S\| \leq 2$ for all $i \geq i^\star$. Hence, it follows from Corollary 60 that Φ_{u_i} is a linear counting transformation for all $i \geq i^\star$. As (T1)–(T4) are preserved under taking the limit for $i \to +\infty$, this implies that the limit of the corresponding sequence is a linear counting transformation, as required. ∎

Finally, we establish that the limit mentioned in Corollary 61 does not depend on the exact choice of the sequence $\{u_i\}_{i \in \mathbb{N}}$.

Corollary 62 *Consider a sequence $S := \{\lambda_x\}_{x \in \mathscr{X}}$ in $\Lambda = [\underline{\lambda}, \overline{\lambda}]$. Then for any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, there is a unique linear counting transformation T such that*

$$(\forall \varepsilon \in \mathbb{R}_{>0})(\exists \delta \in \mathbb{R}_{>0})(\forall u \in \mathscr{U}_{[t,s]}, \sigma(u) \leq \delta) \ \|T - \Phi_u\| \leq \varepsilon.$$

Proof Recall from Equation (37) that Q_S is equal to the generalised Poisson generator $\underline{Q}_{S'}$ associated with the sequence $S' = \{(\lambda_x, \lambda_x)\}_{x \in \mathscr{X}}$. It therefore follows from Theorem 45 (with the sequence $S' = \{(\lambda_x, \overline{\lambda}_x)\}_{x \in \mathscr{X}}$) that there is a unique lower counting transformation $\underline{T}$ that satisfies the condition of the stated. All that remains for us is to verify that this unique lower counting transformation $\underline{T}$ is linear. To that end, we fix any sequence $\{u_i\}_{i \in \mathbb{N}}$ such that $\lim_{i \to +\infty} \sigma(u_i) = 0$. Because $\lim_{i \to +\infty} \sigma(u_i) = 0$, it follows from the first part of the proof that

$$(\forall \varepsilon \in \mathbb{R}_{>0})(\exists i^\star \in \mathbb{N})(\forall i \in \mathbb{N}, i \geq i^\star) \; \|\underline{T} - \Phi_{u_i}\| \leq \varepsilon.$$

Hence, $\lim_{i \to +\infty} \Phi_{u_i} = \underline{T}$. That $\underline{T}$ is a linear counting transformation now follows from this equality if we recall from Corollary 61 that $\lim_{i \to +\infty} \Phi_{u_i}$ is a linear counting transformation. ∎

Using Corollaries 61 and 62, we now define the unique family of linear counting transformations that is generated by the linear generalised Poisson generator Q_S. Consider any sequence $S = \{\lambda_x\}_{x \in \mathscr{X}}$. Then for any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, we define the corresponding linear counting transformation

$$T_{t,S}^s := \lim_{\sigma(u) \to +\infty} \{\Phi_u : u \in \mathscr{U}_{[t,s]}\}.$$

We collect all these transformations in the family $\mathscr{T}_S := \{T_{t,S}^s : t, s \in \mathbb{R}_{\geq 0}, t \leq s\}$.

D.3. Counting Transformation Systems

We now provide a method to construct more intricate families of linear counting transformations. This construction method is essential in the proof of Proposition 100, where we will construct a counting process with transition probabilities that are derived from these linear counting transformations. Specifically, we are interested in families of the following type, the definition of which is based on [8, Definition 3.3].

Definition 63 *A counting transformation system is a family* $\mathscr{T} = \{T_t^s : t, s \in \mathbb{R}_{\geq 0}, t \leq s\}$ *of linear counting transformations such that*

S1. $T_t^t = I$ *for all t in* $\mathbb{R}_{\geq 0}$;

S2. $T_t^s = T_t^r T_r^s$ *for all t, r, s in* $\mathbb{R}_{\geq 0}$ *with* $t \leq r \leq s$;

S3. $\lim_{\Delta \to 0^+} \frac{[T_t^{t+\Delta} \mathbb{I}_{\geq x+2}](x)}{\Delta} = 0$ *and, if $t > 0$,* $\lim_{\Delta \to 0^+} \frac{[T_{t-\Delta}^t \mathbb{I}_{\geq x+2}](x)}{\Delta} = 0$, *for all t in* $\mathbb{R}_{\geq 0}$ *and x in* $\mathscr{X}$.

One example of a counting transformation system is the family $\mathscr{T}_S$ of linear counting transformations generated by the linear generalised Poisson generator Q_S, as is established in the next result.

Corollary 64 *Consider a sequence* $S := \{\lambda_x\}_{x \in \mathscr{X}}$ *in* $\Lambda = [\underline{\lambda}, \overline{\lambda}]$. *Then* $\mathscr{T}_S = \{T_{t,S}^s : t, s \in \mathbb{R}_{\geq 0}, t \leq s\}$, *the corresponding family of linear counting transformations, is a counting transformation system. Furthermore,*

$$T_{t,S}^s = T_{0,S}^{s-t} \qquad \text{for all } t, s \in \mathbb{R}_{\geq 0}, t \leq s. \tag{38}$$

Proof This is a corollary of Equation (37), Proposition 48 (i)–(iii) and Lemma 52. ∎

These simple systems can be used to construct more intricate systems. First, we restrict these counting transformation systems. Consider any counting transformation system $\mathscr{T} = \{T_t^s : t, s \in \mathbb{R}_{\geq 0}, t \leq s\}$ and any interval $\mathscr{I}$ in $\mathbb{R}_{\geq 0}$, here and in the remainder assumed to be of the form $[t, s]$ or $[t, +\infty)$. With this system $\mathscr{T}$ and the interval $\mathscr{I}$, we associate the *restricted counting transformation system*

$$\mathscr{T}^{\mathscr{I}} := \{T_t^s \in \mathscr{T} : t, s \in \mathscr{I}, t \leq s\}.$$

Next, we concatenate two restricted transformation systems. Consider two counting transformation systems $\mathscr{T}_1 = \{T_t^{s,1} : t, s \in \mathbb{R}_{\geq 0}, t \leq s\}$ and $\mathscr{T}_2 = \{T_t^{s,2} : t, s \in \mathbb{R}_{\geq 0}, t \leq s\}$ and two intervals $\mathscr{I}_1$ and $\mathscr{I}_2$ in $\mathbb{R}_{\geq 0}$ such that $\mathscr{I}_1$ is closed and $\max \mathscr{I}_1 = \min \mathscr{I}_2$. Then the associated *concatenated transformation system* $\mathscr{T}_1^{\mathscr{I}_1} \otimes \mathscr{T}_2^{\mathscr{I}_2}$ is defined as the family of transformations $\{T_t^s : t, s \in \mathscr{I}_1 \cup \mathscr{I}_2, t \leq s\}$ such that for all t, s in $\mathscr{I}_1 \cup \mathscr{I}_2$ with $t \leq s$,

$$T_t^s := \begin{cases} T_t^{s,1} & \text{if } s \leq r, \\ T_t^{r,1} T_r^{s,2} & \text{if } t \leq r \leq s, \\ T_t^{s,2} & \text{if } r \leq t, \end{cases} \tag{39}$$

where $r := \max \mathscr{I}_1 = \min \mathscr{I}_2$. The following result establishes that the concatenated counting transformation system is again a (restricted) counting transformation system.

Lemma 65 *Consider two counting transformation systems $\mathscr{T}_1 = \{T_t^{s,1} : t, s \in \mathbb{R}_{\geq 0}, t \leq s\}$ and $\mathscr{T}_2 = \{T_t^{s,2} : t, s \in \mathbb{R}_{\geq 0}, t \leq s\}$ and fix some r in $\mathbb{R}_{>0}$. Then the concatenated transformation system $\mathscr{T}_1^{[0,r]} \otimes \mathscr{T}_2^{[r,+\infty)}$ is a restricted counting transformation system.*

Proof It follows from Equation (39) and Lemma 56 that every operator T_t^s in the concatenation $\mathscr{T}_1^{[0,r]} \otimes \mathscr{T}_2^{[r,+\infty)}$ is a linear counting transformation. That $\mathscr{T}_1^{[0,r]} \otimes \mathscr{T}_2^{[r,+\infty)}$ furthermore satisfies (S1)–(S3) follows immediately from Equation (39) and the fact that $\mathscr{T}_1$ and $\mathscr{T}_2$ satisfy (S1)–(S3). ∎

Corollary 66 *Consider some $u = t_0, \ldots, t_n$ in $\mathcal{U}_\emptyset$ with $t_0 = 0$ and, for all i in $\{0, \ldots, n\}$, some sequence $S_i := \{\lambda_{i,x}\}_{x \in \mathcal{X}}$ in $[\underline{\lambda}, \overline{\lambda}]$. Then*

$$\mathscr{T} := \mathscr{T}_{S_0}^{[0,t_1]} \otimes \mathscr{T}_{S_1}^{[t_1,t_2]} \otimes \cdots \otimes \mathscr{T}_{S_{n-1}}^{[t_{n-1},t_n]} \otimes \mathscr{T}_{S_n}^{[t_n,+\infty)}$$

is a counting transformation system.

Proof This essentially follows from Corollary 64 and Lemma 65. Let $\mathscr{T}_n' := \mathscr{T}_{S_n}$ and $\mathscr{T}_{n-1}' := \mathscr{T}_{S_{n-1}}^{[0,t_n]} \otimes \mathscr{T}_n^{[t_n,+\infty)}$. Recall from Corollary 64 that $\mathscr{T}_{S_{n-1}}$ and $\mathscr{T}_n' = \mathscr{T}_{S_n}$ are counting transformation systems. Hence, it follows from Lemma 65 with $\mathscr{T}_1 = \mathscr{T}_{S_{n-1}}, \mathscr{T}_2 = \mathscr{T}_n'$ and $r = t_n$ that $\mathscr{T}_{n-1}' = \mathscr{T}_{S_{n-1}}^{[0,t_n]} \otimes \mathscr{T}_n^{[t_n,+\infty)}$ is a counting transformation system.

Next, we let $\mathscr{T}_{n-2}' := \mathscr{T}_{S_{n-2}}^{[0,t_{n-1}]} \otimes \mathscr{T}_{n-1}'^{[t_{n-1},+\infty)}$. We have just proven that $\mathscr{T}_{n-1}'$ is a counting transformation system, and it follows from Corollary 64 that $\mathscr{T}_{S_{n-2}}$ is a counting transformation system. Hence, it follows from Lemma 65 with $\mathscr{T}_1 = \mathscr{T}_{S_{n-2}}, \mathscr{T}_2 = \mathscr{T}_{n-1}'$ and $r = t_{n-1}$ that $\mathscr{T}_{n-2} = \mathscr{T}_{S_{n-2}}^{[0,t_{n-1}]} \otimes \mathscr{T}_{n-1}'^{[t_{n-1},+\infty)}$ is a counting transformation system.

It is now clear that if we repeat the same argument an additional $n-2$ times, we have verified the statement. ∎

D.4. From a Linear Counting Transformation System to the Poisson Distribution

We conclude this section of the Appendix with a study of the special case of constant sequences $S = \{\lambda\}_{x \in \mathcal{X}}$ in $\mathbb{R}_{\geq 0}$. For any λ in $\mathbb{R}_{\geq 0}$, we let

$$Q_\lambda := Q_S = \underline{Q}_{S'} \qquad \text{with } S := \{\lambda\}_{x \in \mathcal{X}} \text{ and } S' := \{(\lambda, \lambda)\}_{x \in \mathcal{X}}. \tag{40}$$

Similarly, we let $T_{t,\lambda}^s := T_{t,S}^s$ for all t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, and $\mathscr{T}_\lambda := \mathscr{T}_S$ to simplify our notation.

Corollary 67 *Consider any λ in $\mathbb{R}_{\geq 0}$. Then for any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, f in $\mathscr{L}(\mathcal{X})$ and x in $\mathcal{X}$,*

$$\left[T_{t,\lambda}^s f\right](y) = \left[T_{t,\lambda}^s f_x'\right](y - x) \qquad \text{for all } y \in \mathcal{X} \text{ with } y \geq x,$$

where $f_x' \colon \mathcal{X} \to \mathbb{R} \colon z \mapsto f_x'(z) := f(x + z)$.

Proof From Equation (40), we know that $T_{t,\lambda}^s = \underline{T}_t^s$, where $\underline{T}_t^s$ is the lower counting transformation generated by the (linear) generalised Poisson generator $\underline{Q}_{S'} = Q_\lambda$. Therefore, it follows from Lemma 53 that

$$[T_{t,\lambda}^s f](y) = [\underline{T}_t^s f](y) = [\underline{T}_t^s f_x'](y - x) = [T_{t,\lambda}^s f_x'](y - x),$$

as required. ∎

Corollary 68 *Consider any λ in $\mathbb{R}_{\geq 0}$. Then for any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$,*

$$(\forall \varepsilon \in \mathbb{R}_{>0})(\exists \delta \in \mathbb{R}_{>0})(\forall \Delta \in \mathbb{R}, 0 < |\Delta| < \delta, 0 \leq t + \Delta \leq s) \left\| \frac{T_{t+\Delta,\lambda}^s - T_{t,\lambda}^s}{\Delta} + Q_\lambda T_{t,\lambda}^s \right\| \leq \varepsilon.$$

and

$$(\forall \varepsilon \in \mathbb{R}_{>0})(\exists \delta \in \mathbb{R}_{>0})(\forall \Delta \in \mathbb{R}, 0 < |\Delta| < \delta, t \leq s + \Delta) \left\| \frac{T_{t,\lambda}^{s+\Delta} - T_{t,\lambda}^s}{\Delta} - Q_\lambda T_{t,\lambda}^s \right\| \leq \varepsilon.$$

Proof This is a specialisation of Lemma 51, as by Equation (40), $T_{t,\lambda}^s = \underline{T}_t^s$ where $\underline{T}_t^s$ is the lower counting transformation generated by the (linear) generalised Poisson generator $\underline{Q}_{S'} = Q_\lambda$ associated with $S' = \{(\lambda, \lambda)\}_{x \in \mathcal{X}}$. ∎

Everything is now set up to state and prove the main result of this section, namely how the Poisson distribution is obtained from a counting transformation system.

Proposition 69 *Consider any λ in $\mathbb{R}_{\geq 0}$. Then for all t, Δ in $\mathbb{R}_{\geq 0}$ and x, y in $\mathcal{X}$,*

$$[T_{t,\lambda}^{t+\Delta} \mathbb{I}_y](x) = \begin{cases} \psi_{\lambda\Delta}(y - x) & \text{if } y \geq x, \\ 0 & \text{otherwise.} \end{cases}$$

Proof Let $f := \mathbb{I}_y$, and consider the function $f_x' \colon \mathcal{X} \to \mathbb{R} \colon z \mapsto f_x'(z) := f(x + z) = \mathbb{I}_y(x + z)$. We first consider the case that $x > y$. Observe that $f_x' = 0$ because $x > y$, whence it follows from Corollary 67 with $f = \mathbb{I}_y$ that

$$[T_{t,\lambda}^{t+\Delta} \mathbb{I}_y](x) = [T_{t,\lambda}^{t+\Delta} f](x) = [T_{t,\lambda}^{t+\Delta} f_x'](x - x) = [T_{t,\lambda}^{t+\Delta} 0](0) = 0,$$

where for the final equality we have used Corollary 64 and (T6). This equality clearly agrees with the stated.

Second, we consider the case that $x \leq y$. Observe that $f_x' = \mathbb{I}_{y-x}$. Hence, it follows from Corollary 67 with $f = \mathbb{I}_y$ that

$$[T_{t,\lambda}^s \mathbb{I}_y](x) = [T_{t,\lambda}^{t+\Delta} f](x) = [T_{t,\lambda}^{t+\Delta} f_x'](x - x) = [T_{t,\lambda}^{t+\Delta} \mathbb{I}_{y-x}](0) = [T_{0,\lambda}^\Delta \mathbb{I}_{y-x}](0).$$

where for the final equality we have used Equation (40) and Equation (38) of Corollary 64. Hence, to verify the stated we need to show that

$$\phi_z(\Delta) := [T_{0,\lambda}^\Delta \mathbb{I}_z](0) = \psi_{\lambda\Delta}(z) \qquad \text{for all } \Delta \in \mathbb{R}_{\geq 0} \text{ and } z \in \mathcal{X}. \tag{41}$$

Due to Equation (40), Corollary 64 and (S1), we already know that

$$\phi_z(0) = [T_{0,\lambda}^0 \mathbb{I}_z](0) = [I\mathbb{I}_z](0) = \mathbb{I}_z(0) = \begin{cases} 1 & \text{if } z = 0, \\ 0 & \text{otherwise.} \end{cases}$$

To determine the other values, we start by fixing any Δ in $\mathbb{R}_{\geq 0}$ and z in $\mathcal{X}$. Fix an ε in $\mathbb{R}_{>0}$. By Corollary 68, there is a $\delta^\star$ in $\mathbb{R}_{>0}$ such that

$$(\forall \delta \in \mathbb{R}, 0 < |\delta| < \delta^\star, 0 \leq \Delta + \delta) \left\| \frac{T_{0,\lambda}^{\Delta+\delta} - T_{0,\lambda}^\Delta}{\delta} - Q_\lambda T_{0,\lambda}^\Delta \right\| \leq \varepsilon.$$

Fix any real number δ such that $0 < |\delta| < \delta^\star$ and $0 \leq \Delta + \delta$, and observe that

$$\left| \frac{[T_{0,\lambda}^{\Delta+\delta} \mathbb{I}_z](0) - [T_{0,\lambda}^\Delta \mathbb{I}_z](0)}{\delta} - [Q_\lambda T_{0,\lambda}^\Delta \mathbb{I}_z](0) \right| \leq \left\| \frac{T_{0,\lambda}^{\Delta+\delta} \mathbb{I}_z - T_{0,\lambda}^\Delta \mathbb{I}_z}{\delta} - Q_\lambda T_{0,\lambda}^\Delta \mathbb{I}_z \right\|$$

$$\leq \left\| \frac{T_{0,\lambda}^{\Delta+\delta} - T_{0,\lambda}^\Delta}{\delta} - Q_\lambda T_{0,\lambda}^\Delta \right\| \|\mathbb{I}_z\| = \left\| \frac{T_{0,\lambda}^{\Delta+\delta} - T_{0,\lambda}^\Delta}{\delta} - Q_\lambda T_{0,\lambda}^\Delta \right\| \leq \varepsilon,$$

where for the second inequality we have used (NH4) and the equality holds because $\|\mathbb{I}_z\| = 1$. Note that

$$[Q_\lambda T_{0,\lambda}^\Delta \mathbb{I}_z](0) = \lambda [T_{0,\lambda}^\Delta \mathbb{I}_z](1) - \lambda [T_{0,\lambda}^\Delta \mathbb{I}_z](0) = \lambda [T_{0,\lambda}^\Delta \mathbb{I}_{z-1}](0) - \lambda [T_{0,\lambda}^\Delta \mathbb{I}_z](0) = \lambda \phi_{z-1}(\Delta) - \lambda \phi_z(\Delta),$$

where for the second equality we have used Corollary 67 and where for ease of notation we let $\phi_{-1} := 0$ because if $z = 0$, then $\mathbb{I}_{z-1} = 0$ and hence it follows from Equation (40), Corollary 64 and (T1) that $[T_{0,\lambda}^\Delta \mathbb{I}_{z-1}](0) = [T_{0,\lambda}^\Delta 0](0) = 0$. We substitute this equality in our previous inequality, to obtain

$$\left| \frac{\phi_z(\Delta + \delta) - \phi_z(\Delta)}{\delta} - \lambda \phi_{z-1}(\Delta) + \lambda \phi_z(\Delta) \right| \leq \varepsilon.$$

Since this holds for any δ in $\mathbb{R}$ such that $0 < |\delta| < \delta^\star$, and because ε was an arbitrary positive real number, it follows from this inequality and the definition of the derivative that

$$\mathrm{D}\phi_z(\Delta) = \lambda \phi_{z-1}(\Delta) - \lambda \phi_z(\Delta) \qquad \text{for all } z \in \mathcal{X} \text{ and } \Delta \in \mathbb{R}_{\geq 0},$$

where $\mathrm{D}\phi_z(\Delta)$ denotes the derivative of ϕ_z evaluated in Δ. It is well-known—see for instance [6, Section 3]—that together with the initial condition $\phi_z(0) = \mathbb{I}_z(0)$, the resulting family of recursively defined initial value problems has a unique solution, namely $\phi_z(\Delta) = \psi_{\lambda\Delta}(z)$ for all Δ in $\mathbb{R}_{\geq 0}$ and z in $\mathcal{X}$. ∎

Appendix E. Supplementary Material for Section 2

E.1. Coherent Conditional Probabilities

We start this section with establishing some well-known properties of coherent conditional probabilities. Our first result establishes a necessary and sufficient condition for the real-valued map to be a coherent conditional probability. It is actually the condition that Regazzini [10] uses to define coherent conditional probabilities, but it follows from [10, Theorems 3 and 4] that our definition—that is, Definition 2—is equivalent; see for instance also [8, Appendix B].

Proposition 70 *Let S be a sample space. The real-valued map P on $\mathscr{D} \subseteq \mathscr{E}(S) \times \mathscr{E}_\emptyset(S)$ is a coherent conditional probability if and only if for all n in $\mathbb{N}$, $\alpha_1, \ldots, \alpha_n$ in $\mathbb{R}$ and $(A_1, C_1), \ldots, (A_n, C_n)$ in $\mathscr{D}$,*

$$\max\left\{ \sum_{i=1}^{n} \alpha_i \mathbb{I}_{C_i}(s)(P(A_i \,|\, C_i) - \mathbb{I}_{A_i}(s)) \colon s \in \bigcup_{i=1}^{n} C_i \right\} \geq 0. \tag{42}$$

The next result allows us to always extend coherent conditional probabilities to a larger domain.

Proposition 71 (Theorem 4 in [10]) *Consider a sample space S and a coherent conditional probability P on $\mathscr{D} \subseteq \mathscr{E}(S) \times \mathscr{E}_\emptyset(S)$. Then for any $\mathscr{D}^\star$ with $\mathscr{D} \subseteq \mathscr{D}^\star \subseteq \mathscr{E}(S) \times \mathscr{E}_\emptyset(S)$, P can be extended to a coherent conditional probability $P^\star$ on $\mathscr{D}^\star$.*

Finally, we establish some well-known properties of coherent conditional probabilities. First, we here recall from [10, Section 2] that any coherent conditional probability P satisfies (P1)–(P4) on its domain $\mathscr{D}_{\text{CP}}$. Additionally, it satisfies the following well-known properties; we refer to [8, Appendix B] for proofs.

Lemma 72 *Consider a sample space S and a coherent conditional probability P on $\mathscr{D} \subseteq \mathscr{E}(S) \times \mathscr{E}_\emptyset(S)$. Then for any (A, C) in $\mathscr{D}$,*

P5. $0 \leq P(A \,|\, C) \leq 1$;

P6. $P(A \,|\, C) = P(A \cap C \,|\, C)$ if $(A \cap C, C) \in \mathscr{D}$;

P7. $P(\emptyset \,|\, C) = 0$ if $(\emptyset, C) \in \mathscr{D}$;

P8. $P(S \,|\, C) = 1$ if $(S, C) \in \mathscr{D}$.

In the remainder, we will make frequent use of (P1)–(P8). As these are just the standard laws of probability, we will usually do this without explicitly referring to them.

E.2. Counting Processes in Particular

We first establish two obvious properties of coherent conditional probabilities on $\mathscr{D}_{\text{CP}}$ that will be useful throughout the remainder; see for example Lemma 79 or Proposition 82 further on.

Lemma 73 *Let P be a coherent conditional probability on the domain $\mathscr{D} \subseteq \mathscr{E}(\Omega) \times \mathscr{E}_\emptyset(\Omega)$ that contains $\mathscr{D}_{\text{CP}}$. Fix some t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, u in $\mathscr{U}_{<t}$ and (x_u, x) in $\mathscr{X}_{u \cup t}$. Then for all y in $\mathscr{X}$,*

$$P(X_s \leq y \,|\, X_u = x_u, X_t = x) = \begin{cases} \sum_{z=x}^{y} P(X_s = z \,|\, X_u = x_u, X_t = x) & \text{if } y \geq x, \\ 0 & \text{otherwise.} \end{cases}$$

Consequently, if $y < x$, then

$$P(X_s = y \,|\, X_u = x_u, X_t = x) = 0.$$

Proof To prove the first part of the statement, we observe that

$$P(X_s \leq y \,|\, X_u = x_u, X_t = x) = P(X_u = x_u, X_t = x, X_s \leq y \,|\, X_u = x_u, X_t = x).$$

If $y < x$, then it follows from (A1) that $(X_u = x_u, X_t = x, X_s \leq y) = \emptyset$. Hence,

$$P(X_s \leq x \,|\, X_u = x_u, X_t = x) = P(\emptyset \,|\, X_u = x_u, X_t = x) = 0,$$

which agrees with the stated. Alternatively, if $y \geq x$, then it follows from (A1) and the finite additivity of P that

$$P(X_s \leq y \,|\, X_u = x_u, X_t = x) = \sum_{z=x}^{y} P(X_u = x_u, X_t = x, X_s = z \,|\, X_u = x_u, X_t = x) = \sum_{z=x}^{y} P(X_s = z \,|\, X_u = x_u, X_t = x),$$

as required.

For the second part of the statement, we observe that $(X_s = y) \subseteq (X_s \leq y)$. Together with the first part of the statement and the monotonicity of P, this implies that

$$P(X_s = y \,|\, X_u = x_u, X_t = x) \leq P(X_s \leq y \,|\, X_u = x_u, X_t = x) = 0.$$

As furthermore $P(X_s = y \,|\, X_u = x_u, X_t = x) \geq 0$ by (P1), this clearly implies the second part of the statement. ∎

Lemma 74 *Let P be a coherent conditional probability on the domain $\mathscr{D} \subseteq \mathscr{E}(\Omega) \times \mathscr{E}_{\emptyset}(\Omega)$ that contains $\mathscr{D}_{\mathrm{CP}}$. Fix some t, s in $\mathbb{R}_{\geq 0}$ with $t < s$, u in $\mathscr{U}_{<u}$ and (x_u, x, y) in $\mathscr{X}_{u \cup \{t,s\}}$. Then for all r in $\mathbb{R}_{\geq 0}$ such that $t < r < s$,*

$$P(X_s = y \,|\, X_u = x_u, X_t = x) = \sum_{z=x}^{y} P(X_s = y \,|\, X_u = x_u, X_t = x, X_r = z) P(X_r = z \,|\, X_u = x_u, X_t = x)$$

and

$$P(X_s \geq y \,|\, X_u = x_u, X_t = x) = \sum_{z=x}^{y-1} P(X_s \geq y \,|\, X_u = x_u, X_t = x, X_r = z) P(X_r = z \,|\, X_u = x_u, X_t = x)$$
$$+ P(X_r \geq y \,|\, X_u = x_u, X_t = x).$$

Proof Observe that

$$P(X_s = y \,|\, X_u = x_u, X_t = x) = P(X_u = x_u, X_t = x, X_s = y \,|\, X_u = x_u, X_t = x)$$

due to (P6). As $(X_u = x_u, X_t = x, X_s = y) = \cup_{z=x}^{y}(X_u = x_u, X_t = x, X_r = z, X_s = y)$ due to (A1), it follows from (P3) and (P6) that

$$P(X_s = y \,|\, X_u = x_u, X_t = x) = \sum_{z=x}^{y} P(X_u = x_u, X_t = x, X_r = z, X_s = y \,|\, X_u = x_u, X_t = x)$$
$$= \sum_{z=x}^{y} P(X_r = z, X_s = y \,|\, X_u = x_u, X_t = x).$$

Finally, we use (P4), to yield

$$P(X_s = y \,|\, X_u = x_u, X_t = x) = \sum_{z=x}^{y} P(X_r = z \,|\, X_u = x_u, X_t = x) P(X_s = y \,|\, X_u = x_u, X_t = x, X_r = z),$$

which is the first equality of the statement.

For the second equality of the stated, we observe that due to (A1), $(X_u = x_u, X_t = x, X_s \geq y)$ is the union of the pairwise disjoint events $(X_u = x_u, X_t = x, X_r \geq y)$ and $(X_u = x_u, X_t = x, X_r = z, X_s \geq y)$ for all z in $\mathscr{X}$ such that $x \leq z < y$. We now again use (P3), (P6) and (P4) to yield the second equality of the statement:

$$P(X_s \geq y \,|\, X_u = x_u, X_t = x) = \sum_{z=x}^{y-1} P(X_r = z, X_s \geq y \,|\, X_u = x_u, X_t = x)$$
$$+ P(X_r \geq y \,|\, X_u = x_u, X_t = x)$$
$$= \sum_{z=x}^{y-1} P(X_r = z \,|\, X_u = x_u, X_t = x) P(X_s \geq y \,|\, X_u = x_u, X_t = x, X_r = z)$$
$$+ P(X_r \geq y \,|\, X_u = x_u, X_t = x).$$

∎

Next, we prove the first result that is given in the main text. In this proof, we will need the following—slightly stronger—lemma.

Lemma 75 *Consider some u in $\mathcal{U}$. Then the corresponding collection of finitary events*

$$\mathcal{C}_u := \{(X_v \in B) \colon v \in \mathcal{U}, B \subseteq \mathcal{X}_v, (\forall t \in v)\, t \in u \cup [\max u, +\infty)\}. \tag{43}$$

is an algebra. Therefore, $\mathcal{F}_u = \mathcal{C}_u$.

Proof The second part of the stated—that is, that $\mathcal{F}_u = \mathcal{C}_u$—is an immediate consequence of the first part—that is, that the collection of finitary events $\mathcal{C}_u$ is an algebra—because $\mathcal{F}_u$ is defined in Equation (3) as the smallest algebra of sets that contains $\mathcal{C}_u$. Hence, we only need to verify that the collection of finitary events $\mathcal{C}_u$ is an algebra of sets (sometimes also called a field of events), in the sense that

F1. $\emptyset$ belongs to $\mathcal{C}_u$;

F2. $\Omega \setminus A_1$ belongs to $\mathcal{C}_u$ for all A_1 in $\mathcal{C}_u$;

F3. $A_1 \cup A_2$ belongs to $\mathcal{C}_u$ for all A_1, A_2 in $\mathcal{C}_u$.

Observe first that (F1) holds, as the empty set $\emptyset$ belongs to $\mathcal{C}_u$. For instance, take any t in $[\max u, +\infty)$ and let $B := \emptyset$. Then clearly

$$(X_t \in B) = (X_t \in \emptyset) = \{\omega \in \Omega \colon \omega(t) \in \emptyset\} = \emptyset,$$

and $(X_t \in B) = \emptyset$ belongs to $\mathcal{C}_u$ due to Equation (43). Similarly, Ω belongs to $\mathcal{C}_u$ because $\Omega = (X_t \in \mathcal{X})$ for any t in $[\max u, +\infty)$. Observe that (F2) and (F3) are trivially satisfied for $A_1 = \emptyset$ and $A_1 = \Omega$.

Hence, we now fix any $A_1 = (X_{v_1} \in B_1)$ and $A_2 = (X_{v_2} = B_2)$ in $\mathcal{C}_u$ such that $\emptyset \neq A_1 \neq \Omega$ and $\emptyset \neq A_2 \neq \Omega$, and verify that (F2) and (F3) hold. To that end, we recall that $(X_\emptyset \in \mathcal{X}_\emptyset) = (X_\emptyset = x_\emptyset) = \Omega$, so the condition $\emptyset \neq A_1 \neq \mathcal{X}$ implies that $v_1 \neq \emptyset$, and similarly $v_2 \neq \emptyset$. Hence, we can enumerate the time points in v_1 as $t_1, \ldots, t_n$ and the time points in v_2 as $s_1, \ldots, s_m$.

In order to verify (F2), we observe that

$$\Omega \setminus A_1 = \Omega \setminus A_1 = \Omega \setminus \{\omega \in \Omega \colon (\omega(t_1), \cdots, \omega(t_n)) \in B_1\} = \{\omega \in \Omega \colon (\omega(t_1), \cdots, \omega(t_n)) \notin B_1\}$$
$$= \{\omega \in \Omega \colon (\omega(t_1), \cdots, \omega(t_n)) \in B_1^c\} = (X_{v_1} \in B_1^c),$$

with $B_1^c = \mathcal{X}_{v_1} \setminus B_1$ and where the third equality follows from (A1) and Equation (1). As $B_1^c \subseteq \mathcal{X}_{v_1}$, it follows from Equation (43) that $(X_{v_1} \in B_1^c)$ belongs to $\mathcal{C}_u$. Hence, we may conclude that $\Omega \setminus A_1$ belongs to $\mathcal{C}_u$, as required.

To verify (F3), we observe that

$$A_1 \cup A_2 = \{\omega \in \Omega \colon (\omega(t_1), \ldots, \omega(t_n)) \in B_1\} \cup \{\omega \in \Omega \colon (\omega(s_1), \ldots, \omega(s_m)) \in B_2\}$$
$$= \{\omega \in \Omega \colon (\omega(t_1), \ldots, \omega(t_n)) \in B_1 \text{ or } (\omega(s_1), \ldots, \omega(s_m)) \in B_2\}.$$

Let v be the ordered union of $v_1 = t_1, \ldots, t_n$ and $v_2 = s_1, \ldots, s_m$. We now furthermore enumerate the time-points in v as $r_1, \ldots, r_k$, and let

$$B := \{(x_{r_1}, \ldots, x_{r_k}) \in \mathcal{X}_v \colon (x_{t_1}, \ldots, x_{t_n}) \in B_1 \text{ or } (x_{s_1}, \ldots, x_{s_m}) \in B_2\}. \tag{44}$$

It then follows from (A1) and Equation (1) that

$$A_1 \cup A_2 = \{\omega \in \Omega \colon (\omega(t_1), \ldots, \omega(t_n)) \in B_1 \text{ or } (\omega(s_1), \ldots, \omega(s_m)) \in B_2\}$$
$$= \{\omega \in \Omega \colon (\omega(r_1), \ldots, \omega(r_k)) \in B\} = (X_v \in B).$$

Because v and B trivially satisfy the requirements of Equation (43), we may conclude from this equality that $A_1 \cup A_2$ belongs to $\mathcal{C}_u$, as required. ∎

Proof of Lemma 3 It follows from Lemma 43 that $A = (X_v \in B_v)$, where v is some sequence of time points in $\mathcal{U}$ such that t belongs to $u \cup [\max u, +\infty)$ for all t in v, and where B_v is a subset of $\mathcal{X}_v$.

Let $w' := v \setminus u$. If $w' = \emptyset$, then we fix any t in the open interval $(\max u, +\infty)$, and we let w be the sequence containing t. Alternatively, if $w' \neq \emptyset$, then we let w be the (ordered) sequence of time points in w.

In any case, if we enumerate the time points in $u \cup w$ as $t_1, \ldots, t_n$ and those in v as $s_1, \ldots, s_m$, then we can define the set

$$B := \{(x_{t_1}, \ldots, x_{t_n}) \in \mathcal{X}_{u \cup w} \colon (x_{s_1}, \ldots, x_{s_m}) \in B_v\}. \tag{45}$$

To obtain the stated, we observe that

$$A = (X_v \in B_v) = \{\omega \in \Omega \colon (\omega(s_1), \ldots, \omega(s_m)) \in B_v\} = \{\omega \in \Omega \colon (\omega(t_1), \ldots, \omega(t_n)) \in B\} = (X_{u \cup w} \in B),$$

where for the third equality we have used Equation (45), Equation (1) and (A1). ∎

E.3. Constructing Counting Processes

We end this section with a number of interesting technical results about the construction of counting processes. In the proof of the first of these technical results, we need the following intermediary result.

Lemma 76 *Consider some non-empty finite index set $\mathscr{I}$ and, for all i in $\mathscr{I}$, some α_i and p_i in $\mathbb{R}$. Let $\alpha^\star := \min\{\alpha_i : i \in \mathscr{I}\}$. If $p_i \geq 0$ for all i in $\mathscr{I}$ and $\sum_{i \in \mathscr{I}} p_i \leq 1$, then*

$$\sum_{i \in \mathscr{I}} \alpha_i p_i \geq \min\{0, \alpha^\star\}.$$

Proof We distinguish two cases based on the sign of $\alpha^\star$. If $\alpha^\star \geq 0$, then $\alpha_i \geq 0$ for all i in $\mathscr{I}$. Since furthermore $p_i \geq 0$ for all i in $\mathscr{I}$, we observe that $\sum_{i \in \mathscr{I}} \alpha_i p_i \geq 0 = \min\{0, \alpha^\star\}$ as this is a sum of non-negative terms.

Next, we consider the case $\alpha^\star < 0$. If $\sum_{i \in \mathscr{I}} p_i = 0$, then clearly $\sum_{i \in \mathscr{I}} \alpha_i p_i = 0 \geq \min\{0, \alpha^\star\}$. If $\sum_{i \in \mathscr{I}} p_i > 0$, then we observe that

$$\sum_{i \in \mathscr{I}} \alpha_i p_i = \left(\sum_{j \in \mathscr{I}} p_j\right) \sum_{i \in \mathscr{I}} \alpha_i \frac{p_i}{\sum_{j \in \mathscr{I}} p_j} \geq \left(\sum_{j \in \mathscr{I}} p_j\right) \alpha^\star \geq \alpha^\star = \min\{0, \alpha^\star\},$$

where the first inequality holds because a convex combination of real numbers is greater than or equal to the minimum of these real numbers, and the second inequality holds because $\alpha^\star < 0$ and $0 < \sum_{j \in \mathscr{I}} p_j \leq 1$. ∎

Everything is now set up to prove the following technical lemma, which is crucial when constructing counting processes in general and Poisson processes in particular.

Lemma 77 *Let $w = w_0, \ldots, w_\ell$ be an element of $\mathscr{U}_\emptyset$. Let P_w be a real-valued function on*

$$\mathscr{D}_w := \{(X_{w_j} = y, X_{w^j} = x_{w^j}) \in \mathscr{D}_{\mathrm{CP}} : j \in \{0, \ldots, \ell\}, w^j := \{w_0, \ldots, w_{j-1}\}, x_{w^j} \in \mathscr{X}_{w^j}, y \in \mathscr{X}\} \qquad (46)$$

such that, for any j in $\{0, \ldots, \ell\}$ and x_{w^j} in $\mathscr{X}_{w^j}$ with $w^j = \{w_0, \ldots, w_{j-1}\}$—and specifically $w^0 = \emptyset$—(i) if $j > 0$, then $P_w(X_{w_j} = y \mid X_{w^j} = x_{w^j}) = 0$ for all y in $\mathscr{X}$ with $y < x_{w_{j-1}}$; (ii) $0 \leq P_w(X_{w_j} = y \mid X_{w^j} = x_{w^j}) \leq 1$ for all y in $\mathscr{X}$ with $y \geq x_{w_{j-1}}$; and (iii) $0 \leq \sum_{y \in B} P_w(X_{w_j} = y \mid X_{w^j} = x_{w^j}) \leq 1$ for all finite subsets B of $\mathscr{X}$. Then P_w is a coherent conditional probability.

Proof Our proof by induction is inspired by that of Krak et al. [8, Lemma C.1]: we verify that P_w satisfies the necessary and sufficient condition for coherence of Proposition 70. First, we observe that this is the case if $\ell = 0$. To verify this, we fix any n in $\mathbb{N}$ and, for any i in $\{1, \ldots, n\}$, some $(A_i, C_i) = (X_{w_0} = y_i, \Omega)$ in $\mathscr{D}_w$ and α_i in $\mathbb{R}$. Observe that

$$\max\left\{\sum_{i=1}^n \alpha_i \mathbb{I}_{C_i}(\omega)(P_w(A_i \mid C_i) - \mathbb{I}_{A_i}(\omega)) : \omega \in \bigcup_{i=1}^n C_i\right\} = \max\left\{\sum_{i=1}^n \alpha_i(P_w(X_{w_0} = y_i \mid \Omega) - \mathbb{I}_{A_i}(\omega)) : \omega \in \Omega\right\}. \qquad (47)$$

Let $B := \{y \in \mathscr{X} : (\exists i \in \{1, \ldots, n\}) \, y_i = y\}$, and observe that B is non-empty and has a finite number of elements. We now partition $\{1, \ldots, n\}$ according to the events $A_i = (X_{w_\ell} = y_i)$: for any y in B, we let

$$\mathscr{I}_y := \{i \in \{1, \ldots, n\} : y_i = y\}.$$

Furthermore, we let $\alpha^\star := \min\{\sum_{i \in \mathscr{I}_y} \alpha_i : y \in B\}$, and let $y^\star$ be the element of B that reaches this minimum. Observe that

$$\sum_{i=1}^n \alpha_i P(X_{w_0} = y_i \mid \Omega) = \sum_{y \in B} \sum_{i \in \mathscr{I}_y} \alpha_i P(X_{w_0} = y \mid \Omega) = \sum_{y \in B} \left(\sum_{i \in \mathscr{I}_y} \alpha_i\right) P(X_{w_0} = y \mid \Omega)$$

$$\geq \min\{0, \alpha^\star\}, \qquad (48)$$

where the inequality follows from Lemma 76 and the conditions (ii) and (iii) on P_w of the stated. If $\alpha^\star \geq 0$, we let $\omega^\star$ be any path such that $\omega^\star(w_0) \notin B$; note that this path exists by (A2) because B is a finite subset of $\mathscr{X}$. This way,

$$\sum_{y \in B} \left(\sum_{i \in \mathscr{I}_y} \alpha_i\right) \mathbb{I}_{(X_{w_0} = y)}(\omega^\star) = 0 = \min\{0, \alpha^\star\} \qquad (49)$$

because $\mathbb{I}_{(X_{w_0}=y)}(\omega^\star) = 0$ for all y in B. Otherwise, that is if $\alpha^\star < 0$, we let $\omega^\star$ be any path such that $\omega^\star(w_0) = y^\star$; again, this path exists due to (A2). This way,

$$\sum_{y \in B} \left(\sum_{i \in \mathscr{I}_y} \alpha_i \right) \mathbb{I}_{(X_{w_0}=y)}(\omega^\star) = \alpha^\star = \min\{0, \alpha^\star\} \tag{50}$$

because, for all y in B, $\mathbb{I}_{(X_{w_0}=y)}(\omega^\star) = 0$ if $y \neq y^\star$ and $\mathbb{I}_{(X_{w_0}=y)}(\omega^\star) = 1$ if $y = y^\star$. Our choice of $\omega^\star$ guarantees that

$$\begin{aligned}
\sum_{i=1}^{n} \alpha_i(P(X_{w_0} = y_i \mid \Omega) - \mathbb{I}_{A_i}(\omega^\star)) &= \sum_{i=1}^{n} \alpha_i P(X_{w_0} = y_i \mid \Omega) - \sum_{i=1}^{n} \alpha_i \mathbb{I}_{A_i}(\omega^\star) \\
&= \sum_{i=1}^{n} \alpha_i P(X_{w_0} = y_i \mid \Omega) - \sum_{y \in B} \sum_{i \in \mathscr{I}_y} \alpha_i \mathbb{I}_{A_i}(\omega^\star) \\
&= \sum_{i=1}^{n} \alpha_i P(X_{w_0} = y \mid \Omega) - \sum_{y \in B} \left(\sum_{i \in \mathscr{I}_y} \alpha_i \right) \mathbb{I}_{(X_{w_0}=y)}(\omega^\star) \\
&= \sum_{i=1}^{n} \alpha_i P(X_{w_0} = y \mid \Omega) - \min\{0, \alpha^\star\} \\
&\geq \min\{0, \alpha^\star\} - \min\{0, \alpha^\star\} = 0, \tag{51}
\end{aligned}$$

where the fourth equality follows from Equations (49) and (50), and the inequality holds due to Equation (48). From this we infer that Equation (47) holds, which by Proposition 70 implies that P_w is a coherent conditional probability.

Next, we fix any ℓ in $\mathbb{N}$ with $\ell \geq 1$ and assume assume that the stated holds for any ℓ' in $\mathbb{N}$ with $0 \leq \ell' < \ell$. We will now show that the stated then follows for ℓ. To verify the coherence of P_w, we fix any n in $\mathbb{N}$, $(A_1, C_1), \ldots, (A_n, C_n)$ in $\mathscr{D}_w$ and $\alpha_1, \ldots, \alpha_n$ in $\mathbb{R}$. We need to show that

$$\max\left\{ \sum_{i=1}^{n} \alpha_i \mathbb{I}_{C_i}(\omega)(P_w(A_i \mid C_i) - \mathbb{I}_{A_i}(\omega)) \colon \omega \in C \right\} \geq 0, \tag{52}$$

where $C := \cup_{i=1}^{k} C_i$.

For any i in $\{1, \ldots, n\}$, there is some j_i in $\{0, \ldots, \ell\}$, some y_i in $\mathscr{X}$ and $x_{u^i}^i$ in $\mathscr{X}_{u^i}$ with $u^i := \{w_0, \ldots, w_{j_i-1}\}$, such that

$$A_i = (X_{w_{j_i}} = y_i) \quad \text{and} \quad C_i = (X_{u^i} = x_{u^i}^i).$$

Let $\mathscr{I}_{<\ell} := \{i \in \{1, \ldots, n\} \colon j_i < \ell\}$. If $\mathscr{I}_{<\ell} \neq \emptyset$, then it follows from the induction hypothesis that

$$\max\left\{ \sum_{i \in \mathscr{I}_{<\ell}} \alpha_i \mathbb{I}_{C_i}(\omega)(P_w(A_i \mid C_i) - \mathbb{I}_{A_i}(\omega)) \colon \omega \in C_{<\ell} \right\} \geq 0,$$

with $C_{<\ell} := \cup_{i \in \mathscr{I}_{<\ell}} C_i$. From this, it follows that there is some $\omega^\star$ in $C_{<\ell} \subseteq C$ such that

$$\sum_{i \in \mathscr{I}_{<\ell}} \alpha_i \mathbb{I}_{C_i}(\omega^\star)(P_w(A_i \mid C_i) - \mathbb{I}_{A_i}(\omega^\star)) \geq 0. \tag{53}$$

If $\mathscr{I}_{<\ell} = \emptyset$, then we let $\omega^\star$ be an arbitrary element of C. In any case, we have chosen a $\omega^\star$ in C that satisfies Equation (53).

Let $C^\star := \cap_{j=0}^{\ell-1}(X_{w_j} = \omega^\star(w_j))$ and $\mathscr{I}_{C^\star} := \{i \in \{1, \ldots, n\} \colon C_i = C^\star\}$. Observe that, by construction, $j_i = \ell$ for all i in $\mathscr{I}_{C^\star}$. We now execute the same trick as we did before. Since by construction $\mathscr{I}_{C^\star}$ is clearly finite, the sets $B_1^\star := \{y \in \mathscr{X} \colon y < \omega^\star(w_{\ell-1}), (\exists i \in \mathscr{I}_{C^\star}) \, y_i = y\}$ and $B_2^\star := \{y \in \mathscr{X} \colon y \geq \omega^\star(w_{\ell-1}), (\exists i \in \mathscr{I}_{C^\star}) \, y_i = y\}$ have (at most) a finite number of elements. We now partition $\mathscr{I}_{C^\star}$ according to the events $A_i = (X_{w_\ell} = y_i)$: for any y in $B_1^\star \cup B_2^\star$, we let

$$\mathscr{I}_y := \{i \in \mathscr{I}_{C^\star} \colon A_i = (X_{w_\ell} = y)\}.$$

If we furthermore let $\alpha^\star := \min\{\sum_{i \in \mathscr{I}_y} \alpha_i \colon y \in B_2^\star\}$—where we follow the convention that the minimum of the empty set is zero—then

$$\sum_{i \in \mathscr{I}_{C^\star}} \alpha_i P_w(A_i \mid C_i) = \sum_{i \in \mathscr{I}_{C^\star}} \alpha_i P_w(A_i \mid C^\star) = \sum_{y \in B_1^\star} \sum_{i \in \mathscr{I}_y} \alpha_i P_w(A_i \mid C^\star) + \sum_{y \in B_2^\star} \sum_{i \in \mathscr{I}_y} \alpha_i P_w(A_i \mid C^\star)$$

$$
\begin{aligned}
&= \sum_{y \in B_1^\star} \left(\sum_{i \in \mathscr{I}_y} \alpha_i \right) P_w(X_{w_\ell} = y \mid C^\star) + \sum_{y \in B_2^\star} \left(\sum_{i \in \mathscr{I}_y} \alpha_i \right) P_w(X_{w_\ell} = y \mid C^\star) \\
&= \sum_{y \in B_2^\star} \left(\sum_{i \in \mathscr{I}_y} \alpha_i \right) P_w(X_{w_\ell} = y \mid C^\star) \\
&= \sum_{y \in B_2^\star} \left(\sum_{i \in \mathscr{I}_y} \alpha_i \right) P_w(X_{w_\ell} = y \mid X_{w_0} = \omega^\star(w_0), \ldots, X_{w_{\ell-1}} = \omega^\star(w_{\ell-1})) \\
&\geq \min\{0, \alpha^\star\},
\end{aligned}
\tag{54}
$$

where for the fourth equality we use that $P_w(X_{w_\ell} = y \mid C^\star) = 0$ for all y in $B_1^\star$—which follows from condition (i) on P_w of the statement because, by definition of $B_1^\star$, $y < \omega^\star(w_{\ell-1})$ for all y in $B_1^\star$—and where the inequality follows from Lemma 76 due to the conditions (ii) and (iii) on P_w of the statement.

If $B_2^\star$ is non-empty and $\alpha^\star < 0$, then we let $y^\star$ be any element of $B_2^\star$ that reaches the minimum in the definition of $\alpha^\star$; if $B_2^\star$ is non-empty and $\alpha^\star \geq 0$, then we let $y^\star$ be any element of $\mathscr{X}$ such that $y^\star \geq \omega^\star(w_{\ell-1})$ and $y^\star \notin B_2^\star$; and finally, if $B_2^\star$ is empty, then we set $y^\star := \omega^\star(w_{\ell-1})$. Because $\omega^\star$ belongs to $C^\star$ by definition and $y^\star \geq \omega^\star(w_{\ell-1})$ by construction, it follows from (A2) that there is at least one path ω in $C^\star$ with $\omega(w_\ell) = y^\star$. Let $\omega^{\star\star}$ be any such path. We have chosen $\omega^{\star\star}$ such that

$$
\sum_{y \in B_1^\star} \sum_{i \in \mathscr{I}_y} \alpha_i \mathbb{I}_{A_i}(\omega^{\star\star}) = 0
\tag{55}
$$

because $\mathbb{I}_{A_i}(\omega^{\star\star}) = \mathbb{I}_{(X_{w_\ell} = y_i)}(\omega^{\star\star}) = 0$ for all i in $\cup_{y \in B_1^\star} \mathscr{I}_y$ as $\omega^{\star\star}(w_\ell) = y^\star \geq \omega^\star(w_{\ell-1}) > y_i$. Similarly,

$$
\sum_{y \in B_2^\star} \sum_{i \in \mathscr{I}_y} \alpha_i \mathbb{I}_{A_i}(\omega^{\star\star}) = \begin{cases} \sum_{i \in \mathscr{I}_{y^\star}} \alpha_i & \text{if } B_2^\star \neq \emptyset \text{ and } \alpha_\star < 0 \\ 0 & \text{otherwise} \end{cases} = \min\{0, \alpha^\star\}
\tag{56}
$$

because, for all i in $\cup_{y \in B_2^\star} \mathscr{I}_y$, we have that $\mathbb{I}_{A_i}(\omega^{\star\star}) = \mathbb{I}_{(X_{w_\ell} = y_i)}(\omega^{\star\star}) = 0$ if $y_i \neq y^\star$ and 1 if $y_i = y^\star$—where the latter only occurs if $B_2^\star$ is non-empty and $\alpha^\star < 0$. Furthermore, $\mathbb{I}_{C_i}(\omega^{\star\star}) = 1$ for all i in $\mathscr{I}_{C^\star}$ because $\omega^{\star\star}$ is an element of $C^\star$ by construction. Hence,

$$
\begin{aligned}
\sum_{i \in \mathscr{I}_{C^\star}} \alpha_i \mathbb{I}_{C_i}(\omega^{\star\star})(P_w(A_i \mid C_i) - \mathbb{I}_{A_i}(\omega^{\star\star})) &= \sum_{i \in \mathscr{I}_{C^\star}} \alpha_i (P_w(A_i \mid C_i) - \mathbb{I}_{A_i}(\omega^{\star\star})) = \sum_{i \in \mathscr{I}_{C^\star}} \alpha_i P_w(A_i \mid C_i) - \sum_{i \in \mathscr{I}_{C^\star}} \alpha_i \mathbb{I}_{A_i}(\omega^{\star\star}) \\
&= \sum_{i \in \mathscr{I}_{C^\star}} \alpha_i P_w(A_i \mid C_i) - \sum_{y \in B_1^\star} \sum_{i \in \mathscr{I}_y} \alpha_i \mathbb{I}_{A_i}(\omega^{\star\star}) - \sum_{y \in B_2^\star} \sum_{i \in \mathscr{I}_y} \alpha_i \mathbb{I}_{A_i}(\omega^{\star\star}) \\
&= \sum_{i \in \mathscr{I}_{C^\star}} \alpha_i P_w(A_i \mid C_i) - 0 - \min\{0, \alpha^\star\} \\
&\geq \min\{0, \alpha^\star\} - \min\{0, \alpha^\star\} = 0,
\end{aligned}
\tag{57}
$$

where for the fourth equality we have used Equations (55) and (56) and for the inequality we have used Equation (54).

We now summarise our findings. Recall that, by construction, $\omega^{\star\star}$ belongs to $C^\star$ and that $C^\star$ is a subset of $C_{<\ell}$ if the latter is non-empty. Therefore, it follows from Equation (53) that

$$
\sum_{i \in \mathscr{I}_{<\ell}} \alpha_i \mathbb{I}_{C_i}(\omega^{\star\star})(P_w(A_i \mid C_i) - \mathbb{I}_{A_i}(\omega^{\star\star})) \geq 0.
$$

Similarly, it follows from Equation (57) that

$$
\sum_{i \in \mathscr{I}_{C^\star}} \alpha_i \mathbb{I}_{C_i}(\omega^{\star\star})(P_w(A_i \mid C_i) - \mathbb{I}_{A_i}(\omega^{\star\star})) \geq 0.
$$

Finally, if we let $\mathscr{I} := \{1, \ldots, n\} \setminus (\mathscr{I}_{<\ell} \cup \mathscr{I}_{C^\star})$, then clearly $\mathbb{I}_{C_i}(\omega^{\star\star}) = 0$ for all i in $\mathscr{I}$. Therefore,

$$
\sum_{i \in \mathscr{I}} \alpha_i \mathbb{I}_{C_i}(\omega^{\star\star})(P_w(A_i \mid C_i) - \mathbb{I}_{A_i}(\omega^{\star\star})) = 0.
$$

From the three previous (in)equalities, we infer that

$$
\sum_{i=1}^n \alpha_i \mathbb{I}_{C_i}(\omega^{\star\star})(P_w(A_i \mid C_i) - \mathbb{I}_{A_i}(\omega^{\star\star})) = \sum_{i \in (\mathscr{I}_{<\ell} \cup \mathscr{I}_{C^\star} \cup \mathscr{I})} \alpha_i \mathbb{I}_{C_i}(\omega^{\star\star})(P_w(A_i \mid C_i) - \mathbb{I}_{A_i}(\omega^{\star\star})) \geq 0.
$$

If the path $\omega^{\star\star}$ belongs to C, then we may conclude from this that Equation (52) holds, which is what we set out to prove. Recall that this inequality holds for any path $\omega^{\star\star}$ as long as (i) it belongs to $C^\star$, in the sense that it coincides with $\omega^\star$ on the time points $w_0, \ldots, w_{\ell-1}$; and (ii) it satisfies $\omega^{\star\star}(w_\ell) = y^\star \geq \omega^{\star\star}(w_{\ell-1}) = \omega^\star(w_{\ell-1})$. Furthermore, we recall that $\omega^\star$ belongs to C, and that the paths in C are only "specified" on (a subset of) the time points $w_0, \ldots, w_{\ell-1}$, because

$$C = \bigcup_{i=1}^{k} C_i = \bigcup_{i=1}^{k} (X_{u^i} = x_{u^i}^i)$$

with $u^i = \{w_0, \ldots, w_{j_i} - 1\}$ and $j_i \leq \ell$. Consequently, it follows from (A1) and (A2) that there is a path $\omega^{\star\star}$ in C that satisfies the two requirements. ∎

We continue with a second construction lemma, this time using a counting transformation system.

Lemma 78 *Consider a non-empty and ordered sequence of time points $w = \{w_0, \ldots, w_\ell\}$ in $\mathscr{U}$ and a counting transformation system $\mathscr{T} = \{T_t^s : t, s \in \mathbb{R}_{\geq 0}, t \leq s\}$. Let $P_w^\star$ be any coherent conditional probability on $\mathscr{E}(\Omega) \times \mathscr{E}_\emptyset(\Omega)$ such that, for all j in $\{1, \ldots, \ell\}$, x_{w^j} in $\mathscr{X}_{w^j}$ with $w^j := \{w_0, \ldots, w_{j-1}\}$ and x in $\mathscr{X}$,*

$$P_w^\star(X_{w_j} = x \mid X_{w^j} = x_{w^j}) = [T_{w_{j-1}}^{w_j} \mathbb{I}_x](x_{w_{j-1}}).$$

Then for any t in w and u in $\mathscr{U}_{<t}$ with $0 \neq u \subseteq w$, x_u in $\mathscr{X}_u$ and x in $\mathscr{X}$,

$$P_w^\star(X_t = x \mid X_u = x_u) = [T_{\max u}^t \mathbb{I}_x](x_{\max u}).$$

Proof Our proof of follows that of [8, Lemma C.2] very closely. By assumption, there is some j in $\{1, \ldots, \ell\}$ such that $t = w_j$ and $u \subseteq \{w_0, \ldots, w_{j-1}\}$. We prove the stated using induction. First, we observe that if $t = w_1$, then $u = \{w_0\}$ and the stated is trivially satisfied. Next, we assume that the stated holds for $t = w_{j-1}$ with $1 < j \leq \ell$, and prove that the stated then also holds for $t = w_j$. In the remainder, we distinguish between two cases: $\max u = w_{j-1}$ and $\max u < w_{j-1}$.

Let us first consider the case that $\max u = w_{j-1}$. Observe that, due to the laws of probability,

$$P_w^\star(X_t = x \mid X_u = x_u) = P_w^\star((X_t = x) \cap (X_u = x_u) \mid X_u = x_u). \tag{58}$$

Note that, due to (A1) and Equation (1),

$$(X_u = x_u) = (X_{w^j} \in B), \tag{59}$$

with

$$B := \{y_{w^j} \in \mathscr{X}_{w^j} : (\forall s \in u)\, y_s = x_s\}.$$

Note that B is clearly a finite set, as by construction the last component $y_{w_{j-1}}$ of any y_{w^j} in B is equal to $x_{w_{j-1}}$. Substituting this equality in the previous, we now obtain that

$$\begin{aligned}
P_w^\star(X_t = x \mid X_u = x_u) &= P_w^\star((X_t = x) \cap (X_{w^j} \in B) \mid X_u = x_u) = P_w^\star((X_t = x) \cap (\cup_{y_{w^j} \in B}(X_{w^j} = y_{w^j})) \mid X_u = x_u) \\
&= P_w^\star(\cup_{y_{w^j} \in B}(X_t = x) \cap (X_{w^j} = y_{w^j}) \mid X_u = x_u) = \sum_{y_{w^j} \in B} P_w^\star((X_t = x) \cap (X_{w^j} = y_{w^j}) \mid X_u = x_u) \\
&= \sum_{y_{w^j} \in B} P_w^\star(X_t = x \mid (X_{w^j} = y_{w^j}) \cap (X_u = x_u)) P_w^\star(X_{w^j} = y_{w^j} \mid X_u = x_u) \\
&= \sum_{y_{w^j} \in B} P_w^\star(X_t = x \mid X_{w^j} = y_{w^j}) P_w^\star(X_{w^j} = y_{w^j} \mid X_u = x_u) \\
&= \sum_{y_{w^j} \in B} [T_{w_{j-1}}^{w_j} \mathbb{I}_x](y_{w_{j-1}}) P_w^\star(X_{w^j} = y_{w^j} \mid X_u = x_u) \\
&= [T_{w_{j-1}}^{w_j} \mathbb{I}_x](x_{w_{j-1}}) \sum_{y_{w^j} \in B} P_w^\star(X_{w^j} = y_{w^j} \mid X_u = x_u),
\end{aligned}$$

where for the penultimate equality we have used the equality $t = w_j$ and the condition on $P_w^\star$ of the stated, and where the last equality holds because, by construction, $y_{w_{j-1}} = x_{w_{j-1}}$ for all y_{w^j} in B. If $x < x_{w_{j-1}}$, then because $\mathscr{T}$ is a counting transformation system, it follows from (T8) that $[T_{w_{j-1}}^{w_j} \mathbb{I}_x](x_{w_{j-1}}) = 0$. This agrees with the stated, since

$$[T_{\max u}^t \mathbb{I}_x](x_{\max u}) = [T_{w_{j-1}}^t \mathbb{I}_x](x_{w_{j-1}}) = 0,$$

where the second equality follows from (T8) because $x < x_{w_{j-1}}$. In case $x \geq x_{w_{j-1}}$, we observe that

$$P_w^\star(X_t = x \mid X_u = x_u) = [T_{w_{j-1}}^{w_j} \mathbb{I}_x](x_{w_{j-1}}) \sum_{y_{w^j} \in B} P_w^\star(X_{w^j} = y_{w^j} \mid X_u = x_u)$$

$$= [T_{w_{j-1}}^{w_j} \mathbb{I}_x](x_{w_{j-1}}) P_w^\star(\cup_{y_{w^j} \in B}(X_{w^j} = y_{w^j}) \mid X_u = x_u) = [T_{w_{j-1}}^{w_j} \mathbb{I}_x](x_{w_{j-1}}) P_w^\star(X_{w^j} \in B \mid X_u = x_u)$$

$$= [T_{w_{j-1}}^{w_j} \mathbb{I}_x](x_{w_{j-1}}) P_w^\star(X_u = x_u \mid X_u = x_u) = [T_{w_{j-1}}^t \mathbb{I}_x](x_{w_{j-1}}),$$

where for the penultimate equality we have used Equation (59). This proves the induction step in case $\max u = w_{j-1}$.

Next, we consider the case $\max u < w_{j-1}$. We execute exactly the same trick, but now we consider the—clearly finite—set

$$B := \{y_{w^j} \in \mathscr{X}_{w^j} \colon y_{w_{j-1}} \leq x, (\forall s \in u)y_s = x_s\} \tag{60}$$

and, for any y in $\mathscr{X}$ such that $x_{\max u} \leq y \leq x$, the—again clearly finite—set

$$B_y := \{y_{w^{j-1}} \in \mathscr{X}_{w^{j-1}} \colon (y_{w^{j-1}}, y) \in B\} = \{y_{w^{j-1}} \in \mathscr{X}_{w^{j-1}} \colon y_{w_{j-2}} \leq y, (\forall s \in u)\ y_s = x_s\}. \tag{61}$$

Note that these two sets are connected, as

$$B = \bigcup_{y=x_{\max u}}^{x} \{(y_{w^{j-1}}, y) \colon y_{w^{j-1}} \in B_y\}. \tag{62}$$

Observe that

$$(X_u = x_u) \cap (X_t = x) = (X_{w^j} \in B) \cap (X_t = x).$$

Therefore

$$P_w^\star(X_t = x \mid X_u = x_u) = P_w^\star((X_t = x) \cap (X_u = x_u) \mid X_u = x_u) = P_w^\star((X_t = x) \cap (X_{w^j} \in B) \mid X_u = x_u).$$

If $x < x_{\max u}$, then it follows from Equations (60) and (1) that $B = \emptyset$. Therefore

$$P_w^\star(X_t = x \mid X_u = x_u) = P_w^\star(\emptyset \mid X_u = x_u) = 0 = [T_{\max u}^t \mathbb{I}_x](x_{\max u}),$$

where the final equality holds due to (T8). This case therefore agrees with the stated.

Next, we consider the case $x \geq x_{\max u}$. Then

$$P_w^\star(X_t = x \mid X_u = x_u) = \sum_{y_{w^j} \in B} P_w^\star((X_{w_j} = x) \cap (X_{w^j} = y_{w^j}) \mid X_u = x_u)$$

$$= \sum_{y_{w^j} \in B} P_w^\star(X_{w_j} = x \mid (X_{w^j} = y_{w^j}) \cap (X_u = x_u)) P_w^\star(X_{w^j} = y_{w^j} \mid X_u = x_u)$$

$$= \sum_{y_{w^j} \in B} P_w^\star(X_{w_j} = x \mid X_{w^j} = y_{w^j}) P_w^\star(X_{w^j} = y_{w^j} \mid X_u = x_u)$$

$$= \sum_{y=x_{\max u}}^{x} \sum_{y_{w^{j-1}} \in B_y} P_w^\star(X_{w_j} = x \mid X_{w^{j-1}} = y_{w^{j-1}}, X_{w_{j-1}} = y)$$

$$P_w^\star(X_{w^{j-1}} = y_{w^{j-1}}, X_{w_{j-1}} = y \mid X_u = x_u).$$

where for the third equality we have used Equation (60) and for the final equality we have used Equation (62). We now use the condition on P_w of the stated to substitute the terms of the form $P_w^\star(X_{w_j} = x \mid X_{w^{j-1}} = y_{w^{j-1}}, X_{w_{j-1}} = y)$ with $[T_{w_{j-1}}^{w_j} \mathbb{I}_x](y)$, to yield

$$P_w^\star(X_t = x \mid X_u = x_u) = \sum_{y=x_{\max u}}^{x} [T_{w_{j-1}}^{w_j} \mathbb{I}_x](y) \sum_{y_{w^{j-1}} \in B_y} P_w^\star(X_{w^{j-1}} = y_{w^{j-1}}, X_{w_{j-1}} = y \mid X_u = x_u)$$

$$= \sum_{y=x_{\max u}}^{x} [T_{w_{j-1}}^{w_j} \mathbb{I}_x](y) P_w^\star(X_{w^{j-1}} \in B_y, X_{w_{j-1}} = y \mid X_u = x_u)$$

$$= \sum_{y=x_{\max u}}^{x} [T_{w_{j-1}}^{w_j} \mathbb{I}_x](y) P_w^\star(X_u = x_u, X_{w^{j-1}} \in B_y, X_{w_{j-1}} = y \mid X_u = x_u).$$

Observe now that, for all y in $\mathscr{X}$ with $x_{\max u} \le y \le x$,

$$(X_u = x_u, X_{w_{j-1}} = y) = \{\omega \in \Omega \colon \omega(w_{j-1}) = y, (\forall s \in u)\ \omega(s) = x_s\}$$
$$\subseteq \{\omega \in \Omega \colon \omega(w_{j-2}) \le y, (\forall s \in u)\ \omega(s) = x_s\} = (X_{w^{j-1}} \in B_y),$$

where the inclusion follows from (A1) and the final equality follows from Equation (61). We infer from this that $(X_u = x_u, X_{w^{j-1}} \in B_y, X_{w_{j-1}} = y) = (X_u = x_u, X_{w_{j-1}} = y)$. We now use this equality to yield

$$\begin{aligned}
P_w^\star(X_t = x \,|\, X_u = x_u) &= \sum_{y=x_{\max u}}^{x} [T_{w_{j-1}}^{w_j}\mathbb{I}_x](y) P_w^\star(X_u = x_u, X_{w_{j-1}} = y \,|\, X_u = x_u) \\
&= \sum_{y=x_{\max u}}^{x} [T_{w_{j-1}}^{w_j}\mathbb{I}_x](y) P_w^\star(X_{w_{j-1}} = y \,|\, X_u = x_u) \\
&= \sum_{y=x_{\max u}}^{x} [T_{w_{j-1}}^{w_j}\mathbb{I}_x](y)[T_{\max u}^{w_{j-1}}\mathbb{I}_y](x_{\max u}) = [T_{\max u}^{w_j}\mathbb{I}_y](x_{\max u}) = [T_{\max u}^{t}\mathbb{I}_y](x_{\max u}),
\end{aligned}$$

where the third equality follows from the induction hypothesis and the penultimate equality follows from Lemma 57. ∎

Lemma 79 *Consider a counting transformation system $\mathscr{T}$. Let $\tilde{P}$ be the real-valued map with domain*

$$\tilde{\mathscr{D}} := \{(X_{t+\Delta} = y, (X_u = x_u, X_t = x)) \in \mathscr{D}_{\mathrm{CP}} \colon t, \Delta \in \mathbb{R}_{\ge 0}, u \in \mathscr{U}_{<t}, (x_u, x) \in \mathscr{X}_{u \cup t}, y \in \mathscr{X}\} \cup \{(X_0 = x, \Omega) \in \mathscr{D}_{\mathrm{CP}} \colon x \in \mathscr{X}\},$$

that is defined for all t, Δ in $\mathbb{R}_{\ge 0}$, u in $\mathscr{U}_{<t}$, (x_u, x) in $\mathscr{X}_{u \cup t}$ and y in $\mathscr{X}$ as

$$\tilde{P}(X_{t+\Delta} = y \,|\, X_u = x_u, X_t = x) := [T_t^{t+\Delta}\mathbb{I}_y](x)$$

and for all x in $\mathscr{X}$ as

$$\tilde{P}(X_0 = x \,|\, \Omega) := \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{otherwise}. \end{cases}$$

Then $\tilde{P}$ is coherent, and any coherent extension of $\tilde{P}$ to $\mathscr{D}_{\mathrm{CP}}$ is a counting process.

Proof Our proof follows that of Krak et al. [8, Therorem 5.2] closely. We first verify that $\tilde{P}$ is coherent using Proposition 70. To that end, we fix any arbitrary n in $\mathbb{N}$ and, for all i in $\{1, \ldots, n\}$, some $(A_i, C_i) = (X_{t_i} = x, X_{u_i} = x_{u_i})$ in $\tilde{\mathscr{D}}$ and α_i in $\mathbb{R}$. We need to show that

$$\max\left\{ \sum_{i=1}^{n} \alpha_i \mathbb{I}_{C_i}(\omega)\big(\tilde{P}(A_i \,|\, C_i) - \mathbb{I}_{A_i}(\omega)\big) \colon \omega \in \bigcup_{i=1}^{k} C_i \right\} \ge 0. \tag{63}$$

Clearly, there is some non-empty, finite and increasing sequence $w = w_0, \ldots, w_\ell$ of time points with $w_0 = 0$ such that $u_i \subseteq w$ and $t_i \in w$ for all i in $\{1, \ldots, \ell\}$. Let $\tilde{P}_w$ be the restriction of $\tilde{P}$ to $\mathscr{D}_w$, with $\mathscr{D}_w$ as defined in Equation (46) of Lemma 77. In order to verify that $\tilde{P}_w$ satisfies the three conditions of Lemma 77, we fix some j in $\{0, \ldots, \ell\}$ and x_{w^j} in $\mathscr{X}_{w^j}$, with $w^j := w_0, \ldots, w_{j-1}$.

(i) Assume that $j > 0$, and fix some y in $\mathscr{X}$ such that $y < x_{w_{j-1}}$. Observe that

$$\tilde{P}_w(X_{w_j} = y \,|\, X_{w^j} = x_{w^j}) = \tilde{P}(X_{w_j} = y \,|\, X_{w^j} = x_{w^j}) = [T_{w_{j-1}}^{w_j}\mathbb{I}_y](x_{w_{j-1}}) = 0,$$

where the last equality holds due to (T8) because $y < x_{w_{j-1}}$.

(ii) Fix some y in $\mathscr{X}$ such that $y \ge x_{w_{j-1}}$. Observe that

$$\tilde{P}_w(X_{w_j} = y \,|\, X_{w^j} = x_{w^j}) = \tilde{P}(X_{w_j} = y \,|\, X_{w^j} = x_{w^j}) = [T_{w_{j-1}}^{w_j}\mathbb{I}_y](x_{w_{j-1}}).$$

The second condition is now satisfied because $0 = \inf \mathbb{I}_y \le [T_{w_{j-1}}^{w_j}\mathbb{I}_y](x_{w_{j-1}}) \le \sup \mathbb{I}_y = 1$ due to (T5).

41

(iii) Fix some finite subset B of $\mathscr{X}$, and observe that

$$\tilde{P}_w(X_{w_j} \in B \,|\, X_{w^j} = x_{w^j}) = \tilde{P}(X_{w_j} \in B \,|\, X_{w^j} = x_{w^j}) = \sum_{y \in B} \tilde{P}(X_{w_j} = y \,|\, X_{w^j} = x_{w^j})$$

$$= \sum_{y \in B} [T^{w_j}_{w_{j-1}} \mathbb{I}_y](x_{w_{j-1}}) = \left[T^{w_j}_{w_{j-1}} \left(\sum_{y \in B} \mathbb{I}_y \right) \right](x_{w_{j-1}}) = [T^{w_j}_{w_{j-1}} \mathbb{I}_B](x_{w_{j-1}}),$$

where for the fourth equality we have used the additivity (T2) of the linear counting transformation $T^{w_j}_{w_{j-1}}$. The third condition is now satisfied because $0 = \inf \mathbb{I}_B \le [T^{w_j}_{w_{j-1}} \mathbb{I}_B](x_{w_{j-1}}) \le \sup \mathbb{I}_B = 1$ due to (T5).

Consequently, it follows from Lemma 77 that $\tilde{P}_w$ is a coherent conditional probability. By Proposition 71, we can therefore extend $\tilde{P}_w$ to a coherent conditional probability on $\mathscr{E}(\Omega) \times \mathscr{E}_\emptyset(\Omega)$. Let $\tilde{P}^\star_w$ be any such extension. It then follows from Proposition 70 that

$$\max \left\{ \sum_{i=1}^n \alpha_i \mathbb{I}_{C_i}(\omega) \big(\tilde{P}^\star_w(A_i \,|\, C_i) - \mathbb{I}_{A_i}(\omega) \big) \colon \omega \in \bigcup_{i=1}^k C_i \right\} \ge 0. \tag{64}$$

We now claim that $\tilde{P}^\star_w(A_i \,|\, C_i) = \tilde{P}(A_i \,|\, C_i)$ for all i in $\{1, \ldots, n\}$. To verify this claim, we fix any such i. If $u_i = \emptyset$, then $t_i = 0$ and so (A_i, C_i) is an element of $\mathscr{D}_w$; therefore, $\tilde{P}^\star_w(A_i \,|\, C_i) = \tilde{P}_w(A_i \,|\, C_i) = \tilde{P}(A_i \,|\, C_i)$. If $u_i \ne \emptyset$, then $u_i \subseteq w$ and $t_i \in w$. In this case, since $\tilde{P}^\star_w$ satisfies the conditions of Lemma 78, it follows from this lemma that $\tilde{P}^\star_w(A_i \,|\, C_i) = \tilde{P}(A_i \,|\, C_i)$. Since $\tilde{P}^\star_w(A_i \,|\, C_i) = \tilde{P}(A_i \,|\, C_i)$ for all i in $\{1, \ldots, n\}$, Equation (63) now follows from Equation (64).

Now that we have verified that $\tilde{P}$ is coherent, it follows from Proposition 71 that it can be extended to a coherent conditional probability $\tilde{P}^\star$ on $\mathscr{D}_{\mathrm{CP}}$. Let $\tilde{P}^\star$ be any such coherent extension.

We need to verify that $\tilde{P}^\star$ is a counting process. That (CP1) is satisfied is immediate:

$$\tilde{P}^\star(X_0 = 0) = \tilde{P}^\star(X_0 = 0 \,|\, \Omega) = \tilde{P}(X_0 = 0 \,|\, \Omega) = 1.$$

To check (CP2), we fix any t, Δ in $\mathbb{R}_{\ge 0}$, u in $\mathscr{U}_{<t}$ and (x_u, x) in $\mathscr{X}_{u \cup t}$. Observe that

$$\tilde{P}^\star(X_{t+\Delta} \ge x+2 \,|\, X_u = x_u, X_t = x) = 1 - \tilde{P}^\star(X_{t+\Delta} < x+2 \,|\, X_u = x_u, X_t = x)$$
$$= 1 - \tilde{P}^\star(X_{t+\Delta} \le x+1 \,|\, X_u = x_u, X_t = x)$$
$$= 1 - \tilde{P}^\star(X_{t+\Delta} = x \,|\, X_u = x_u, X_t = x) - \tilde{P}^\star(X_{t+\Delta} = x+1 \,|\, X_u = x_u, X_t = x)$$
$$= 1 - \tilde{P}(X_{t+\Delta} = x \,|\, X_u = x_u, X_t = x) - \tilde{P}(X_{t+\Delta} = x+1 \,|\, X_u = x_u, X_t = x)$$
$$= 1 - [T^{t+\Delta}_t \mathbb{I}_x](x) - [T^{t+\Delta}_t \mathbb{I}_{x+1}](x),$$

where we have used Lemma 73 for the third equality. Observe that $1 = [T^{t+\Delta}_t \mathbb{I}_{\ge x}](x)$ because $T^{t+\Delta}_t 1 = 1$ due to (T6) and $[T^{t+\Delta}_t 1](x) = [T^{t+\Delta}_t \mathbb{I}_{\ge x}](x)$ due to (T4). Therefore,

$$\tilde{P}^\star(X_{t+\Delta} \ge x+2 \,|\, X_u = x_u, X_t = x) = [T^{t+\Delta}_t \mathbb{I}_{\ge x}](x) - [T^{t+\Delta}_t \mathbb{I}_x](x) - [T^{t+\Delta}_t \mathbb{I}_{x+1}](x)$$
$$= [T^{t+\Delta}_t (\mathbb{I}_{\ge x} - \mathbb{I}_x - \mathbb{I}_{x+1})](x) = [T^{t+\Delta}_t \mathbb{I}_{\ge x+2}](x),$$

where for the second equality we have used the additivity (T2) of the linear counting transformation $T^{t+\Delta}_t$. Similarly, in case $\max u < t - \Delta$,

$$\tilde{P}^\star(X_t \ge x+2 \,|\, X_u = x_u, X_{t-\Delta} = x) = [T^t_{t-\Delta} \mathbb{I}_{\ge x+2}](x).$$

That (CP2) is satisfied now follows if we combine these two equalities with (S3). ∎

Appendix F. Supplementary Material for Section 3

Our proof of Theorem 6 is rather lengthy, and therefore we split it into two parts. For the first part, we first establish some convenient properties of a Poisson process.

Lemma 80 *Consider a Poisson process P. Then*

(i) $0 \le P(X_t = x \,|\, X_0 = 0) \le 1$ *for all t in $\mathbb{R}_{\ge 0}$ and x in $\mathscr{X}$;*

(ii) $P(X_0 = x \mid X_0 = 0) = 1$ *if* $x = 0$ *and* 0 *otherwise for all* x *in* $\mathscr{X}$;

(iii) $P(X_{t_1+t_2} = x \mid X_0 = 0) = \sum_{y=0}^{x} P(X_{t_1} = y \mid X_0 = 0) P(X_{t_2} = x - y \mid X_0 = 0)$ *for all* t_1, t_2 *in* $\mathbb{R}_{\geq 0}$ *and* x *in* $\mathscr{X}$;

(iv) $\lim_{t \to 0^+} P(X_t = x \mid X_0 = 0) = P(X_0 = x \mid X_0 = 0)$ *for all* x *in* $\mathscr{X}$.

Proof

(i) Follows immediately from (P1).

(ii) Follows almost immediately from (CP1) and (P2).

(iii) If t_1 or t_2 is zero, then this follows almost immediately from (ii). Hence, we now consider the case that $t_1 \neq 0 \neq t_2$. It follows from Lemma 74 with $t = 0$, $r = t_1$ and $s = t_1 + t_2$ that

$$P(X_{t_1+t_2} = x \mid X_0 = 0) = \sum_{y=0}^{x} P(X_{t_1+t_2} = x \mid X_0 = 0, X_{t_1} = y) P(X_{t_1} = y \mid X_0 = 0).$$

We now use (PP1), (PP3) and (PP2), to yield

$$P(X_{t_1+t_2} = x \mid X_0 = 0) = \sum_{y=0}^{x} P(X_{t_1+t_2} = x \mid X_{t_1} = y) P(X_{t_1} = y \mid X_0 = 0)$$

$$= \sum_{y=0}^{x} P(X_{t_2} = x \mid X_0 = y) P(X_{t_1} = y \mid X_0 = 0)$$

$$= \sum_{y=0}^{x} P(X_{t_1} = y \mid X_0 = 0) P(X_{t_2} = x - y \mid X_0 =).$$

(iv) Observe that if $x \geq 2$, then it follows from (P5) and the monotonicity of P that for all t in $\mathbb{R}_{>0}$,

$$0 \leq P(X_t = x \mid X_0 = 0) \leq P(X_t \geq 2 \mid X_0 = 0).$$

The stated now follows from this inequality because $\lim_{t \to 0^+} P(X_t \geq 2 \mid X_0 = 0) = 0$ due to (CP2).

Next, we consider the case $x = 0$. Recall from (i) that $P(X_t = x \mid X_0 = 0)$ is bounded. Furthermore, as $(X_{t+\Delta} = 0) \subseteq (X_t = 0)$ due to (A1), it follows from the monotonicity of P that $P(X_t = x \mid X_0 = 0) \geq P(X_{t+\Delta} = x \mid X_0 = 0)$ for all t, Δ in $\mathbb{R}_{\geq 0}$. In other words, $P(X_t = x \mid X_0 = 0)$ is a bounded and non-increasing function of t. It is a standard result from analysis that the left limit of a bounded and non-increasing function on $\mathbb{R}_{\geq 0}$ exists everywhere. Consequently, $\lim_{t \to 0^+} P(X_t = 0 \mid X_0 = 0)$ exists, and we denote this limit by ℓ.

We need to show that $\ell = P(X_0 = 0 \mid X_0 = 0) = 1$, where the final equality holds due to (ii). Observe that $0 \leq \ell \leq 1$ due to (i). Our proof is one by contradiction: we assume ex-absurdo that $\ell < 1$. Fix any t in $\mathbb{R}_{>0}$ and n in $\mathbb{N}$, and let $\Delta := t/n$. It then follows from (iii) that

$$P(X_t = 1 \mid X_0 = 0) = P(X_\Delta = 0 \mid X_0 = 0) P(X_{t-\Delta} = 1 \mid X_0 = 0) + P(X_\Delta = 1 \mid X_0 = 0) P(X_{t-\Delta} = 0 \mid X_0 = 0).$$

We now apply (iii) $(n-1)$ additional times, to yield

$$P(X_t = 1 \mid X_0 = 0) = n P(X_\Delta = 1 \mid X_0 = 0) P(X_\Delta = 0 \mid X_0 = 0)^{n-1}.$$

Observe that $P(X_\Delta = 1 \mid X_0 = 0) \leq 1$ due to (P5), and that $P(X_\Delta = 0 \mid X_0 = 0) \leq \ell$ as $P(X_s = 0 \mid X_0 = 0)$ is non-increasing in s and has ℓ as its right limit in $s = 0$. Hence,

$$P(X_t = 1 \mid X_0 = 0) \leq n \ell^{n-1}.$$

It is clear that if we take n sufficiently large, then this upper bound is arbitrarily close to 0. From this and (i), it follows that $P(X_t = 1 \mid X_0 = 0) = 0$ for all t in $\mathbb{R}_{>0}$. Consequently, $\lim_{t \to 0^+} P(X_t = 1 \mid X_0 = 0) = 0$. To obtain our contradiction, we observe that it follows from (P3) and (P2) that, for all t in $\mathbb{R}_{>0}$,

$$P(X_t = 0 \mid X_0 = 0) = 1 - P(X_t = 1 \mid X_0 = 0) - P(X_t \geq 2 \mid X_0 = 0). \tag{65}$$

Taking the limit for $t \to 0^+$ on both sides of the equality yields our contradiction, as

$$\lim_{t \to +0^+} P(X_t = 0 \,|\, X_0 = 0) = 1 - \lim_{t \to 0^+} P(X_t = 1 \,|\, X_0 = 0) - \lim_{t \to 0^+} P(X_t \geq 2 \,|\, X_0 = 0) = 1,$$

where for the final equality we have also used that $\lim_{t \to 0^+} P(X_t \geq 2 \,|\, X_0 = 0) = 0$, a consequence of (CP2).

For the remaining case that $x = 1$, we use Equation (65) to yield the stated:

$$\lim_{t \to 0^+} P(X_t = 1 \,|\, X_0 = 0) = 1 - \lim_{t \to 0^+} P(X_t = 0 \,|\, X_0 = 0) - \lim_{t \to 0^+} P(X_t \geq 2 \,|\, X_0 = 0) = 0 = P(X_0 = 1 \,|\, X_0 = 0),$$

where the second equality follows from the previous and the final equality follows from (ii).

■

Next, we establish the main result of the first part.

Proposition 81 *Consider a Poisson process P. Then there is a rate λ in $\mathbb{R}_{\geq 0}$ such that, for all t, Δ in $\mathbb{R}_{\geq 0}$, u in $\mathscr{U}_{<t}$, (x_u, x) in $\mathscr{X}_{u \cup t}$ and y in $\mathscr{X}$,*

$$P(X_{t+\Delta} = y \,|\, X_u = x_u, X_t = x) = \begin{cases} \psi_{\lambda\Delta}(y - x) & \text{if } y \geq x, \\ 0 & \text{otherwise,} \end{cases}$$

Proof The stated for the case $y < x$ follows immediately from Lemma 73. Hence, we immediately move on to the case that $y \geq x$. In this case,

$$P(X_{t+\Delta} = y \,|\, X_u = x_u, X_t = x) = P(X_{t+\Delta} = y \,|\, X_t = x) = P(X_{t+\Delta} = y - x \,|\, X_t = 0) = P(X_\Delta = y - x \,|\, X_0 = 0),$$

where the first equality follows from (PP1), the second from (PP2) and the third from (PP3). To verify the stated, we now need to show that there is a λ in $\mathbb{R}_{\geq 0}$ such that

$$(\forall \Delta \in \mathbb{R}_{\geq 0})(\forall z \in \mathscr{X}) \ P(X_\Delta = z \,|\, X_0 = 0) = \psi_{\lambda\Delta}(z). \tag{66}$$

What follows is a standard argument; see for instance [4, Chapter XVII, Section 6] or [6, Section 2]. We start with the case $z = 0$. For notational simplicity, we let $\theta \coloneqq P(X_1 = 0 \,|\, X_0 = 0)$. Recall from Lemma 80 (i) that $0 \leq \theta \leq 1$. Furthermore, we observe that, for any n in $\mathbb{N}$,

$$P(X_1 = 0 \,|\, X_0 = 0) = P(X_{\frac{1}{n}} = 0 \,|\, X_0 = 0)^n,$$

where the equality follows from applying Lemma 80 (iii) n times. Clearly, this implies that, for any n in $\mathbb{N}$,

$$P(X_{\frac{1}{n}} = 0 \,|\, X_0 = 0) = \theta^{\frac{1}{n}}. \tag{67}$$

Recall from before that $0 \leq \theta \leq 1$. We infer from these inequalities and Equation (67) that $0 < \theta \leq 1$. Indeed, in case $\theta = 0$, then it follows from Equation (67) that $\lim_{n \to +\infty} P(X_{\frac{1}{n}} = 0 \,|\, X_0 = 0) = 0$, which is not correct because this limit is equal to 1 due to Lemma 80 (ii) and (iv).

Next, we observe that, for any n and k in $\mathbb{N}$,

$$P(X_{\frac{k}{n}} = 0 \,|\, X_0 = 0) = P(X_{\frac{1}{n}} = 0 \,|\, X_0 = 0)^k = \theta^{\frac{k}{n}}, \tag{68}$$

where for the first equality we have applied Lemma 80 (iii) k times and for the second equality we have used Equation (67).

Next, we fix any Δ in $\mathbb{R}_{>0}$. Choose any n in $\mathbb{N}$, and let k be the non-negative integer such that

$$\frac{k-1}{n} \leq \Delta \leq \frac{k}{n}.$$

Because $P(X_s = 0 \,|\, X_0 = 0)$ is a non-increasing function of s—as we have argued in the proof of Lemma 80 (iv)—it follows from these inequalities and Equation (68) that

$$\theta^{\frac{k-1}{n}} \geq P(X_\Delta = 0 \,|\, X_0 = 0) \geq \theta^{\frac{k}{n}}.$$

It is now clear that in the limit for $n \to +\infty$, the lower and upper bound both converge to θ^Δ. We now let $\lambda := -\ln(\theta)$, which yields a non-negative real number as $0 < \theta \le 1$. Observe furthermore that $\psi_{\lambda 0}(0) = 1 = P(X_0 = 0 \,|\, X_0 = 0)$, where the final equality is precisely Lemma 80 (ii). In conclusion,

$$(\forall \Delta \in \mathbb{R}_{\ge 0}) \; P(X_\Delta = 0 \,|\, X_0 = 0) = \theta^\Delta = e^{-\lambda \Delta}, \tag{69}$$

so Equation (66) is satisfied for the case $z = 0$.

Next, we verify Equation (66) for $z > 0$. First, we introduce some additional notation. For any z in $\mathscr{X}$, we let

$$\phi_z \colon \mathbb{R}_{\ge 0} \to \mathbb{R} \colon \Delta \mapsto \phi_z(\Delta) := P(X_\Delta = z \,|\, X = 0).$$

Additionally, we also let $\phi_{-1} := 0$. Observe that $\psi_z(0) = \mathbb{I}_z(0)$ due to Lemma 80 (ii). We now claim that

$$\mathrm{D}\phi_z(\Delta) = \lambda \phi_{z-1}(\Delta) - \lambda \phi_z(\Delta) \qquad \text{for all } z \in \mathscr{X} \text{ and } \Delta \in \mathbb{R}_{\ge 0}, \tag{70}$$

where—as in the proof of Lemma 69—$\mathrm{D}\phi_z(\Delta)$ denotes the derivative of ϕ_z evaluated in Δ. It is well-known—see for instance [6, Section 3]—that together with the initial condition $\phi_z(0) = \mathbb{I}_z(0)$, the resulting family of recursively defined initial value problems has a unique solution, namely $\phi_z(\Delta) = \psi_{\lambda \Delta}(z)$ for all Δ in $\mathbb{R}_{\ge 0}$ and z in $\mathscr{X}$. Hence, our claim Equation (70) implies Equation (66).

In order to verify our claim, we first study the derivative of ϕ_z in 0. It follows immediately from Equation (69) that

$$\lim_{\delta \to 0^+} \frac{\phi_0(\delta) - \phi_0(0)}{\delta} = \lim_{\delta \to 0^+} \frac{e^{-\lambda \delta} - 1}{\delta} = \lambda. \tag{71}$$

Next, we use that $\phi_1(0) = 0$, execute some straightforward manipulations, use Equation (69) and also (CP2), to yield

$$\begin{aligned}
\lim_{\delta \to 0^+} \frac{\phi_1(\delta) - \phi_1(0)}{\delta} &= \lim_{\delta \to 0^+} \frac{\phi_1(\delta)}{\delta} = \lim_{\delta \to 0^+} \frac{P(X_\delta = 1 \,|\, X_0 = 0)}{\delta} \\
&= \lim_{\delta \to 0^+} \frac{1 - P(X_\delta = 0 \,|\, X_0 = 0) - P(X_\delta \ge 2 \,|\, X_0 = 0)}{\delta} \\
&= \lim_{\delta \to 0^+} \frac{1 - P(X_\delta = 0 \,|\, X_0 = 0)}{\delta} - \lim_{\delta \to 0^+} \frac{P(X_\delta \ge 2 \,|\, X_0 = 0)}{\delta} \\
&= \lim_{\delta \to 0^+} \frac{1 - e^{-\lambda \delta}}{\delta} - \lim_{\delta \to 0^+} \frac{P(X_\delta \ge 2 \,|\, X_0 = 0)}{\delta} \\
&= -\lambda.
\end{aligned} \tag{72}$$

Finally, we observe that, for any z in $\mathscr{X}$ such that $z \ge 2$ and any δ in $\mathbb{R}_{>0}$,

$$0 \le \phi_z(\delta) = P(X_\delta = z \,|\, X_0 = 0) \le P(X_\delta \ge 2 \,|\, X_0 = 0),$$

where the second inequality follows from the monotonicity of P. From these inequalities, from the equality $\phi_z(0) = \mathbb{I}_z(0)$ and from (CP2), we infer that, for all z in $\mathscr{X}$ such that $z \ge 2$,

$$\lim_{\delta \to 0^+} \frac{\phi_z(\delta) - \phi_z(0)}{\delta} = 0. \tag{73}$$

We are now finally ready to study the derivative of ϕ_z in a general time point—that is, to verify our claim Equation (70). First, we observe that it is an immediate consequence of Equation (69) that, for all Δ in $\mathbb{R}_{\ge 0}$,

$$\mathrm{D}\phi_0(\Delta) = \lim_{\delta \to 0} \frac{\phi_0(\Delta + \delta) - \phi_0(\Delta)}{\delta} = \lambda \phi_0(\Delta) = -\lambda \phi_{-1}(\Delta) + \lambda \phi_0(\Delta),$$

where the limit is a right limit if $\Delta = 0$. Hence, we move on to the case $z \ge 1$. We will only consider the right limit, the left limit can be verified using a similar—but slightly more involved— argument. To that end, we fix any Δ in $\mathbb{R}_{\ge 0}$ and z in $\mathscr{X}$ with $z \ge 1$. We use Lemma 80 (iii) and Equations (71)–(73), to yield

$$\lim_{\delta \to 0^+} \frac{\phi_z(\Delta + \delta) - \phi_z(\Delta)}{\delta} = \lim_{\delta \to 0^+} \frac{\sum_{z'=0}^{z} \phi_{z-z'}(\Delta) \phi_{z'}(\delta) - \phi_z(\Delta)}{\delta}$$

$$= \sum_{z'=0}^{z-1} \phi_{z'}(\Delta) \lim_{\delta \to 0^+} \frac{\phi_{z'}(\delta)}{\delta} + \phi_z(\Delta) \lim_{\delta \to 0^+} \frac{\phi_0(\delta) - 1}{\delta}$$
$$= -\lambda \phi_{z-1}(\Delta) + \lambda \phi_z(\Delta),$$

as required. ∎

For the second part of the proof of Theorem 6, we construct a Poisson process using only the transition probabilities.

Proposition 82 *Fix any λ in $\mathbb{R}_{\geq 0}$. Consider the real-valued map $\tilde{P}$ with domain*

$$\tilde{\mathscr{D}} := \{(X_{t+\Delta} = y, (X_u = x_u, X_t = x)) \in \mathscr{D}_{CP} : t, \Delta \in \mathbb{R}_{\geq 0}, u \in \mathscr{U}_{<t}, (x_u, x) \in \mathscr{X}_{u \cup t}, y \in \mathscr{X}\} \cup \{(X_0 = x, \Omega) \in \mathscr{D}_{CP} : x \in \mathscr{X}\},$$

that is defined for all t, Δ in $\mathbb{R}_{\geq 0}$, u in $\mathscr{U}_{<t}$, (x_u, x) in $\mathscr{X}_{u \cup t}$ and y in $\mathscr{X}$ as

$$\tilde{P}(X_{t+\Delta} = y \,|\, X_u = x_u, X_t = x) := \begin{cases} \psi_{\lambda\Delta}(y - x) & \text{if } y \geq x, \\ 0 & \text{otherwise}, \end{cases}$$

and for all x in $\mathscr{X}$ as

$$\tilde{P}(X_0 = 0 \,|\, \Omega) := \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{otherwise.} \end{cases} \quad 1$$

Then $\tilde{P}$ is a coherent conditional probability that has a unique extension $\tilde{P}^\star$ to $\mathscr{D}_{CP}$. Even more, this extension $\tilde{P}^\star$ is a Poisson process.

Proof Observe that for all t, Δ in $\mathbb{R}_{\geq 0}$, u in $\mathscr{U}_{<t}$, (x_u, x) in $\mathscr{X}_{u \cup t}$ and y in $\mathscr{X}$,

$$\tilde{P}(X_{t+\Delta} = y \,|\, X_u = x_u, X_t = x) = \begin{cases} \psi_{\lambda\Delta}(y - x) & \text{if } y \geq x \\ 0 & \text{otherwise} \end{cases} = [T_{t,\lambda}^{t+\Delta} \mathbb{I}_y](x), \tag{74}$$

where the final equality follows from Proposition 69. Therefore, it follows from Corollary 64 and Lemma 79 that the map $\tilde{P}$ is indeed coherent, and any coherent extension $\tilde{P}^\star$ of this map to $\mathscr{D}_{CP}$ is a counting process.

We now set out to prove that this coherent extension $\tilde{P}^\star$ is unique. To that end, we fix any $(A, X_u = x_u)$ in $\mathscr{D}_{CP}$. We distinguish two cases: $u = \emptyset$ and $u \neq \emptyset$. For the first case, we let $\tilde{P}'$ be any coherent extension of $\tilde{P}^\star$ to $\mathscr{E}(\Omega) \times \mathscr{E}_\emptyset(\Omega)$. Observe that

$$\begin{aligned} \tilde{P}^\star(A \,|\, X_\emptyset = x_\emptyset) &= \tilde{P}'(A \,|\, X_\emptyset = x_\emptyset) = \tilde{P}'(A \,|\, \Omega) = \tilde{P}'(A \cap ((X_0 = 0) \cup (X_0 > 0)) \,|\, \Omega) \\ &= \tilde{P}'(A \cap (X_0 = 0) \,|\, \Omega) + \tilde{P}'(A \cap (X_0 > 0) \,|\, \Omega) \\ &= \tilde{P}'(A \,|\, X_0 = 0)\tilde{P}'(X_0 = 0 \,|\, \Omega) + \tilde{P}'(A \,|\, X_0 > 0)\tilde{P}'(X_0 > 0 \,|\, \Omega) \\ &= \tilde{P}^\star(A \,|\, X_0 = 0)\tilde{P}^\star(X_0 = 0 \,|\, \Omega) + \tilde{P}'(A \,|\, X_0 > 0)\tilde{P}^\star(X_0 > 0 \,|\, \Omega) \\ &= \tilde{P}^\star(A \,|\, X_0 = 0)\tilde{P}^\star(X_0 = 0 \,|\, \Omega) + \tilde{P}'(A \,|\, X_0 > 0)(1 - \tilde{P}^\star(X_0 = 0 \,|\, \Omega)) \\ &= \tilde{P}^\star(A \,|\, X_0 = 0), \end{aligned} \tag{75}$$

where the third equality follows from (P6) and the obvious fact that $(X_0 = 0)$ and $(X_0 > 0)$ partition Ω, the fourth equality follows from (P3), the fifth equality follows from (P4) and the last equality holds because $\tilde{P}^\star(X_0 = 0 \,|\, \Omega) = \tilde{P}(X_0 = 0 \,|\, \Omega) = 1$ by the conditions of the statement. Hence, $\tilde{P}^\star(A \,|\, X_\emptyset = x_\emptyset)$ is uniquely defined if $\tilde{P}^\star(A \,|\, X_0 = 0)$ is.

We therefore immediately move on to the case $u \neq \emptyset$. From Lemma 3, we know that there is a v in $\mathscr{U}$ with $\min v > \max u$ and a subset B of $\mathscr{X}_w$ with $w := u \cup v$ such that $A = (X_w \in B)$. It follows from (P6) that

$$\tilde{P}^\star(A \,|\, X_u = x_u) = \tilde{P}^\star(A \cap (X_u = x_u) \,|\, X_u = x_u) = \tilde{P}^\star(X_u = x_u, X_v \in B' \,|\, X_u = x_u),$$

where we let $B' := \{y_v \in \mathscr{X}_v : (x_u, y_v) \in B\}$.

If $B' = \emptyset$, then it immediately follows from (P7) that

$$\tilde{P}^\star(A \,|\, X_u = x_u) = \tilde{P}^\star(X_u = x_u, X_v \in B' \,|\, X_u = x_u) = \tilde{P}^\star(\emptyset \,|\, X_u = x_u) = 0.$$

We therefore assume that $B' \neq \emptyset$. Fix any arbitrary ε in $\mathbb{R}_{>0}$, and let $\Delta := \max w - \max u$. As the Poisson distribution $\psi_{\lambda\Delta}$ is sigma-additive and normed, it follows that there is a z_ε in $\mathscr{X}$ such that for all z in $\mathscr{X}$ with $z \geq z_\varepsilon \geq x_{\max u}$,

$$1 - \varepsilon \leq \sum_{y=x_{\max u}}^{z} \psi_{\lambda\Delta}(y - x_{\max u}) \leq 1. \tag{76}$$

Fix now any such z. Observe that

$$\tilde{P}^\star(X_{\max w} > z \,|\, X_u = x_u) = 1 - \tilde{P}^\star(X_{\max u + \Delta} \leq z \,|\, X_u = x_u) = 1 - \sum_{y=x_{\max u}}^{z} \tilde{P}^\star(X_{\max u + \Delta} = y \,|\, X_u = x_u)$$

$$= 1 - \sum_{y=x_{\max u}}^{z} \tilde{P}(X_{\max u + \Delta} = y \,|\, X_u = x_u) = 1 - \sum_{y=x_{\max u}}^{z} \psi_{\lambda\Delta}(y - x_{\max u}),$$

where for the second equality we have used Lemma 73. We combine this equality with Equation (76), to yield

$$0 \leq \tilde{P}^\star(X_{\max w} > z \,|\, X_u = x_u) \leq \varepsilon. \tag{77}$$

Let $B'_z := \{y_v \in B' : y_{\max v} \leq z\}$, and observe that

$$(X_u = x_u, X_v \in B'_z) \subseteq (X_u = x_u, X_v \in B') \subseteq (X_u = x_u, X_v \in B'_z) \cup (X_{\max w} > z).$$

Note that $(X_u = x_u, X_v \in B'_z)$ and $(X_{\max w} > z)$ are clearly disjoint. Therefore, it follows from the monotonicity and additivity of $\tilde{P}^\star$ that

$$\tilde{P}^\star(X_u = x_u, X_v \in B'_z \,|\, X_u = x_u) \leq \tilde{P}^\star(X_u = x_u, X_v \in B' \,|\, X_u = x_u) \leq \tilde{P}^\star(X_u = x_u, X_v \in B'_z \,|\, X_u = x_u) + \varepsilon, \tag{78}$$

where for the last inequality we have also used the upper bound on $\tilde{P}^\star(X_{\max w} > z \,|\, X_u = x_u)$ of Equation (77).

We now consider the communal term in Equation (78). Since B'_z is finite by construction, it follows from (P3) that

$$\tilde{P}^\star(X_u = x_u, X_v \in B'_z \,|\, X_u = x_u) = \tilde{P}^\star\left(\bigcup_{y_v \in B'_z} (X_u = x_u, X_v = y_v) \,\middle|\, X_u = x_u\right) = \sum_{y_v \in B'_z} \tilde{P}^\star(X_u = x_u, X_v = y_v \,|\, X_u = x_u).$$

If we enumerate the time points in v as $t_1, \ldots, t_\ell$, then for any y_v in B'_z, we find that

$$\tilde{P}^\star(X_u = x_u, X_v = y_v \,|\, X_u = x_u)$$
$$= \tilde{P}^\star(X_u = x_u, X_{t_1} = y_{t_1} \,|\, X_u = x_u) \cdots \tilde{P}^\star(X_u = x_u, X_{t_\ell} = y_{t_\ell} \,|\, X_u = x_u, X_{t_1} = y_{t_1}, \ldots, X_{t_{\ell-1}} = y_{t_{\ell-1}})$$
$$= \tilde{P}^\star(X_{t_1} = y_{t_1} \,|\, X_u = x_u) \cdots \tilde{P}^\star(X_{t_\ell} = y_{t_\ell} \,|\, X_u = x_u, X_{t_1} = y_{t_1}, \ldots, X_{t_{\ell-1}} = y_{t_{\ell-1}})$$
$$= \tilde{P}(X_{t_1} = y_{t_1} \,|\, X_u = x_u) \cdots \tilde{P}(X_{t_\ell} = y_{t_\ell} \,|\, X_u = x_u, X_{t_1} = y_{t_1}, \ldots, X_{t_{\ell-1}} = y_{t_{\ell-1}}),$$

where the last equality holds because all arguments of $\tilde{P}^\star$ in the factors of this product are elements of $\tilde{\mathscr{D}}$. Hence, $\tilde{P}^\star(X_u = x_u, X_v \in B'_z \,|\, X_u = x_u)$ is uniquely defined by $\tilde{P}$. Because of this, and also because Equation (78) holds for any positive real number ε, we infer from this that $\tilde{P}^\star(A \,|\, X_u = x_u)$ is completely defined by the values of $\tilde{P}$ on its domain. As $(A, X_u = x_u)$ was an arbitrary element of $\mathscr{D}_{CP}$, we conclude that there is a unique extension of $\tilde{P}$ to $\mathscr{D}_{CP}$, and so λ indeed uniquely characterises a coherent conditional probability on $\mathscr{D}_{CP}$.

Finally, we verify that the unique extension $\tilde{P}^\star$ is a Poisson process. To that end, we first verify that the coherent conditional probability $\tilde{P}^\star$ is a counting process. That $\tilde{P}^\star$ satisfies (CP1) follows immediately from the definition of $\tilde{P}$. Furthermore, (CP2) is also satisfied due to the definition of $\tilde{P}$. In order to verify this, we fix any t in $\mathbb{R}_{\geq 0}$, u in $\mathscr{U}_{<t}$ and (x_u, x) in $\mathscr{X}_{u \cup t}$. Observe that, for any Δ in $\mathbb{R}_{>0}$,

$$\tilde{P}^\star(X_{t+\Delta} \geq x + 2 \,|\, X_u = x_u, X_t = x) = 1 - \tilde{P}^\star(X_{t+\Delta} \leq x + 1 \,|\, X_u = x_u, X_t = x)$$
$$= 1 - \tilde{P}^\star(X_{t+\Delta} = x \,|\, X_u = x_u, X_t = x) - \tilde{P}^\star(X_{t+\Delta} = x + 1 \,|\, X_u = x_u, X_t = x)$$
$$= 1 - \tilde{P}(X_{t+\Delta} = x \,|\, X_u = x_u, X_t = x) - \tilde{P}(X_{t+\Delta} = x + 1 \,|\, X_u = x_u, X_t = x)$$
$$= 1 - \psi_{\lambda\Delta}(0) - \psi_{\lambda\Delta}(1),$$

where we have used Lemma 73 for the second equality. Consequently,

$$\lim_{\Delta \to 0^+} \frac{\tilde{P}^\star(X_{t+\Delta} \geq x + 2 \,|\, X_u = x_u, X_t = x)}{\Delta} = \lim_{\Delta \to 0^+} \frac{1 - e^{-\lambda\Delta} - \lambda\Delta e^{-\lambda\Delta}}{\Delta} = 0,$$

as required. If $t > 0$, then similar reasoning can be used to verify this equality for the limit from the left. Hence, $\tilde{P}^\star$ is a counting process. That $\tilde{P}^\star$ furthermore satisfies (PP1)–(PP3) follows immediately from the definition of $\tilde{P}$. ∎

Proof of Theorem 6 The first part of the stated follows from Proposition 81. The second part essentially follows from Proposition 82. The requirement of Proposition 82 regarding $P(X_0 = x \,|\, \Omega)$ seems to be more restrictive, but it is not. To see this, we observe that any coherent conditional probability on $\mathscr{D}_{\mathrm{CP}}$ that satisfies (CP1), will also satisfy $P(X_0 = x \,|\, \Omega) = 0$ for all $x > 0$, as

$$0 \leq P(X_0 = x \,|\, \Omega) = 1 - P(X_0 \in \mathscr{X} \setminus \{x\} \,|\, \Omega) \leq 1 - P(X_0 = 0 \,|\, \Omega) = 0.$$

∎

Proof of Corollary 7 Let P be a Poisson process, and let λ be the rate mentioned in Theorem 6. Then

$$\lim_{\Delta \to 0^+} \frac{P(X_{t+\Delta} = x + 1 \,|\, X_u = x_u, X_t = x)}{\Delta} = \lim_{\Delta \to 0^+} \frac{\psi_{\lambda\Delta}(1)}{\Delta} = \lim_{\Delta \to 0^+} \frac{\lambda\Delta e^{-\lambda\Delta}}{\Delta} = \lambda$$

and, if $t > 0$,

$$\lim_{\Delta \to 0^+} \frac{P(X_t = x + 1 \,|\, X_u = x_u, X_{t-\Delta} = x)}{\Delta} = \lim_{\Delta \to 0^+} \frac{\psi_{\lambda\Delta}(1)}{\Delta} = \lim_{\Delta \to 0^+} \frac{\lambda\Delta e^{-\lambda\Delta}}{\Delta} = \lambda.$$

∎

Proof of Theorem 8 To prove this result, we make use of the terminology and results of Sections 4–6. Observe that, by assumption, P is consistent with the degenerate rate interval $\Lambda = [\lambda, \lambda]$. Fix some t, Δ in $\mathbb{R}_{\geq 0}$, u in $\mathscr{U}_{<t}$, (x_u, x) in $\mathscr{X}_{u \cup t}$ and y in $\mathscr{X}$. Observe that

$$P(X_{t+\Delta} = y \,|\, X_u = x_u, X_t = x) = E_P(\mathbb{I}_{(X_{t+\Delta} = y)} \,|\, X_u = x_u, X_t = x) = E_P(\mathbb{I}_y(X_{t+\Delta}) \,|\, X_u = x_u, X_t = x)$$
$$\geq \underline{E}_{[\lambda,\lambda]}(\mathbb{I}_y(X_{t+\Delta}) \,|\, X_u = x_u, X_t = x) = [T_{t,\lambda}^{t+\Delta} \mathbb{I}_y](x),$$

where for the first equality we have used Equation (4), for the inequality we have used Equation (8) and for the final equality we have used Theorem 15. Similarly, we also find that

$$P(X_{t+\Delta} = y \,|\, X_u = x_u, X_t = x) \leq \overline{E}_{[\lambda,\lambda]}(\mathbb{I}_y(X_{t+\Delta}) \,|\, X_u = x_u, X_t = x) = -\underline{E}_{[\lambda,\lambda]}(-\mathbb{I}_y(X_{t+\Delta}) \,|\, X_u = x_u, X_t = x)$$
$$= -[T_{t,\lambda}^{t+\Delta}(-\mathbb{I}_y)](x) = [T_{t,\lambda}^{t+\Delta} \mathbb{I}_y](x),$$

where for the first equality we have used conjugacy and for the final equality we have used (T1). Therefore,

$$P(X_{t+\Delta} = y \,|\, X_u = x_u, X_t = x) = [T_{t,\lambda}^{t+\Delta} \mathbb{I}_y](x) = \begin{cases} \psi_{\lambda\Delta}(y - x) & \text{if } y \geq x, \\ 0 & \text{otherwise,} \end{cases}$$

where the last equality follows from Proposition 69. This implies the stated due to Theorem 6. ∎

Appendix G. Supplementary Material for Section 5

Most of the results in this section are specialisations of results in Appendices A, B and C.

G.1. The Poisson Generator

All of the properties mentioned in the main text concerning the Poisson generator and the semi-group it induces, follow from results in Appendix C. Indeed, as we have previously mentioned—see Equation (20)—the Poisson generator $\underline{Q}$ as defined in Section 5.2 is precisely the generalised Poisson generator $\underline{Q}_S$ corresponding to the sequence $S = \{(\underline{\lambda}, \overline{\lambda})\}_{x \in \mathscr{X}}$. Hence, we obtain Theorem 10 as a corollary of Theorems 44 and 45.

Proof of Theorem 10 Recall from Theorem 44 that the corresponding sequence $\{\Phi_{u_i}\}_{i \in \mathbb{N}}$ converges to a lower counting transformation, which is a special type of non-negatively homogeneous transformation. Furthermore, it follows from Theorem 45 that this limit does not depend on the chosen sequence $\{u_i\}_{i \in \mathbb{N}}$. ∎

The properties (SG1)–(SG3) of the family of transformations of the form $\underline{T}_t^s$ just state that $\underline{T}_t^s$ is a lower counting transformation, as is established in Theorem 45. Furthermore, the properties (SG4)–(SG6) are stated in Proposition 48.

In the remainder, more specifically in the proof of Proposition 100 further on, we will need the following intermediary result.

Lemma 83 *Fix some t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$ and f in $\mathscr{L}(\mathscr{X})$. Then for any ε in $\mathbb{R}_{>0}$, there is a sequence $u = t_0, \ldots, t_n$ in $\mathscr{U}_{[t,s]}$ and, for all i in $\{1, \ldots, n\}$, a sequence $S_i = \{\lambda_x^i\}_{x \in \mathscr{X}}$ in $[\underline{\lambda}, \overline{\lambda}]$ such that*

$$\left\| \underline{T}_t^s f - \prod_{i=1}^n T_{t_{i-1}, S_i}^{t_i} f \right\| \leq \varepsilon.$$

.

In our proof of Lemma 83, we need the following obvious observation.

Lemma 84 *For any f in $\mathscr{L}(\mathscr{X})$, there is a sequence $S = \{\lambda_x\}_{x \in \mathscr{X}}$ in $\Lambda = [\underline{\lambda}, \overline{\lambda}]$ such that $\underline{Q}f = Q_S f$.*

Proof This is immediate from the definition of $\underline{Q}$ and Q_S. ∎

Proof of Lemma 83 Fix any ε' in $\mathbb{R}_{>0}$ such that $2\varepsilon'\|f\| \leq \varepsilon/2$ and δ in $\mathbb{R}_{>0}$ such that $\delta(s-t)\|\underline{Q}\|^2\|f\| \leq \varepsilon/2$. By Theorem 45, there is a sequence $u = t_0, \ldots, t_n$ in $\mathscr{U}_{[t,s]}$ such that $\sigma(u) \leq \delta$, $\sigma(u)\|\underline{Q}\| \leq 2$ and

$$\|\underline{T}_t^s - \Phi_u\| \leq \varepsilon'. \tag{79}$$

Let $g_i := \prod_{j=i+1}^n (I + \Delta_j \underline{Q})f$ for all i in $\{1, \ldots, n\}$, where $g_n = If = f$. It now follows from Lemma 84 that for any i in $\{1, \ldots, n\}$, there is a sequence $S_i = \{\lambda_x^i\}_{x \in \mathscr{X}}$ in $[\underline{\lambda}, \overline{\lambda}]$ such that $\underline{Q}g_i = Q_{S_i}g_i$. By construction,

$$\Phi_u f = \prod_{i=1}^n (I + \Delta_i \underline{Q})f = \prod_{i=1}^n (I + \Delta_i Q_{S_i})f. \tag{80}$$

Furthermore, we use Lemma 34, to yield

$$(\forall i \in \{1, \ldots, n\}) \ \|Q_{S_i}\| = 2\sup\{\lambda_x^i \colon x \in \mathscr{X}\} \leq 2\overline{\lambda} = \|\underline{Q}\|. \tag{81}$$

Observe now that

$$\left\| \underline{T}_t^s f - \prod_{i=1}^n T_{t_{i-1}, S_i}^{t_i} f \right\| = \left\| \underline{T}_t^s f - \prod_{i=1}^n (I + \Delta_i Q_{S_i})f + \prod_{i=1}^n (I + \Delta_i Q_{S_i})f - \prod_{i=1}^n T_{t_{i-1}, S_i}^{t_i} f \right\|$$

$$\leq \left\| \underline{T}_t^s f - \prod_{i=1}^n (I + \Delta_i Q_{S_i})f \right\| + \left\| \prod_{i=1}^n (I + \Delta_i Q_{S_i})f - \prod_{i=1}^n T_{t_{i-1}, S_i}^{t_i} f \right\|$$

$$= \|\underline{T}_t^s f - \Phi_u f\| + \left\| \prod_{i=1}^n (I + \Delta_i Q_{S_i})f - \prod_{i=1}^n T_{t_{i-1}, S_i}^{t_i} f \right\|$$

$$\leq \|\underline{T}_t^s - \Phi_u\|\|f\| + \left\| \prod_{i=1}^n (I + \Delta_i Q_{S_i}) - \prod_{i=1}^n T_{t_{i-1}, S_i}^{t_i} \right\|\|f\|,$$

where for the final equality we have used Equation (80) and for the final inequality we have used (NH4). We use Equation (79), to yield

$$\left\| \underline{T}_t^s f - \prod_{i=1}^n T_{t_{i-1},S_i}^{t_i} f \right\| \leq \varepsilon' \|f\| + \left\| \prod_{i=1}^n (I + \Delta_i Q_{S_i}) - \prod_{i=1}^n T_{t_{i-1},S_i}^{t_i} \right\| \|f\|.$$

Next, we use Lemma 26 to rewrite the second term on the right hand side of the inequality, to yield

$$\left\| \underline{T}_t^s f - \prod_{i=1}^n T_{t_{i-1},S_i}^{t_i} f \right\| \leq \varepsilon' \|f\| + \sum_{i=1}^n \left\| (I + \Delta_i Q_{S_i}) - T_{t_{i-1},S_i}^{t_i} \right\| \|f\|.$$

Finally, we use Lemma 49 (i) and Equation (81), to yield

$$\left\| \underline{T}_t^s f - \prod_{i=1}^n T_{t_{i-1},S_i}^{t_i} f \right\| \leq \varepsilon' \|f\| + \sum_{i=1}^n \Delta_i^2 \|Q_{S_i}\|^2 \|f\| \leq \varepsilon' \|f\| + \sum_{i=1}^n \Delta_i^2 \|\underline{Q}\|^2 \|f\| \leq \varepsilon' \|f\| + \sum_{i=1}^n \Delta_i \delta \|\underline{Q}\|^2 \|f\|$$

$$= \varepsilon' \|f\| + (s-t)\delta \|\underline{Q}\|^2 \|f\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

as required. $\blacksquare$

G.2. The Reduced Poisson Generator

The claims in Section 5.3 of the main text are essentially a consequence of the following simple result.

Lemma 85 *Consider some $\underline{x}, \overline{x}$ in $\mathscr{X}$ with $\underline{x} \leq \overline{x}$. If we let $\chi := \{x \in \mathscr{X} : \underline{x} \leq x \leq \overline{x}\}$, then $\underline{Q}^\chi$—as defined in Section 5.3—is a lower transition rate transformation.*

Proof The four conditions of Definition 27 are trivially satisfied. $\blacksquare$

That the limit in Equation (17) exists and is independent of the chosen sequence $\{u_i\}_{i \in \mathbb{N}}$ now immediately follows from Lemma 85 and Propositions 30 and 31. Furthermore, these three results also imply that the transformations of the form $\underline{T}_{t,s}^\chi$ are lower transition transformations, and hence satisfy (LT1)–(LT3)—or equivalently, (SG1)–(SG3). Furthermore, we additionally use Proposition 32 to yield that the family also satisfies—properties similar to—(SG4)–(SG6). Finally, we end this section on the reduced Poisson generator $\underline{Q}^\chi$ with some technical results.

Lemma 86 *Consider some $\underline{x}, \overline{x}$ in $\mathscr{X}$ with $\underline{x} \leq \overline{x}$, and let $\chi := \{x \in \mathscr{X} : \underline{x} \leq x \leq \overline{x}\}$. Then the set of dominating transition rate matrices*

$$\mathscr{Q}^\chi := \{Q^\chi \in \mathscr{R}(\chi) \colon (\forall g \in \mathscr{L}(\chi))\, \underline{Q}^\chi g \leq Q^\chi g\},$$

where $\mathscr{R}(\chi)$ denotes the set of all transition rate matrices—that is, linear lower transition rate transformations—on $\mathscr{L}(\chi)$, is non-empty, bounded, closed and convex. Furthermore, Q^χ is an element of $\mathscr{Q}^\chi$ if and only if there is a sequence $\{\lambda_x\}_{x \in \chi \setminus \{\overline{x}\}}$ in $[\underline{\lambda}, \overline{\lambda}]$ such that

$$[Q^\chi g](x) = \begin{cases} \lambda_x(g(x+1) - g(x)) & \text{if } \underline{x} \leq x < \overline{x} \\ 0 & \text{if } x = \overline{x} \end{cases} \quad \text{for all } g \in \mathscr{L}(\chi) \text{ and } x \in \chi.$$

Proof The first part of the stated follows immediately from [8, Proposition 7.8]. The second part is a matter of straightforward verification. $\blacksquare$

Corollary 87 *Consider some $\underline{x}, \overline{x}$ in $\mathscr{X}$ with $\underline{x} \leq \overline{x}$, and let $\chi := \{x \in \mathscr{X} : \underline{x} \leq x \leq \overline{x}\}$. Then for any Q^χ in $\mathscr{Q}^\chi$, with $\mathscr{Q}^\chi$ as defined in Lemma 86,*

$$\|Q^\chi\| \leq \|\underline{Q}^\chi\|.$$

Proof This is an immediate corollary of Lemmas 28 and 86:

$$\|Q^\chi\| = 2\max\{\lambda_x \colon x \in \chi, x < \overline{x}\} \cup \{0\} \leq 2\overline{\lambda} = \|\underline{Q}^\chi\|.$$

$\blacksquare$

Corollary 88 *Consider some $\underline{x}, \overline{x}$ in $\mathscr{X}$ with $\underline{x} \leq \overline{x}$, and let $\chi := \{x \in \mathscr{X} : \underline{x} \leq x \leq \overline{x}\}$. Fix some n in $\mathbb{N}$ and, for every i in $\{1, \ldots, n\}$, some Δ_i in $\mathbb{R}_{\geq 0}$ with $\Delta_i \|\underline{Q}^\chi\| \leq 2$ and some Q_i^χ in $\mathscr{Q}^\chi$. Then for all f^χ in $\mathscr{L}(\chi)$,*

$$\prod_{i=1}^{n} (I + \Delta_i \underline{Q}^\chi) f^\chi \leq \prod_{i=1}^{n} (I + \Delta_i Q_i^\chi) f^\chi.$$

Proof This is an immediate corollary of Lemma 86 and [8, Lemma F.4]. ∎

G.3. The Essential Case of Eventually Constant Functions

Before we can prove our two results concerning eventually constant functions, we first need to establish two technical lemmas.

Lemma 89 *For any Δ in $\mathbb{R}_{\geq 0}$ and any f in $\mathscr{L}^c(\mathscr{X})$ that is constant after $\overline{x}$, $(I + \Delta \underline{Q}) f$ is constant after $\overline{x}$. Furthermore, if we fix any $\underline{x}$ in $\mathscr{X}$ with $\underline{x} \leq \overline{x}$ and let $\chi := \{y \in \mathscr{X} : \underline{x} \leq x \leq \overline{x}\}$, then*

$$[(I + \Delta \underline{Q}) f](x) = \begin{cases} [(I + \Delta \underline{Q}^\chi) f^\chi](x) & \text{if } x \leq \overline{x} \\ f(\overline{x}) & \text{if } x \geq \overline{x} \end{cases} \qquad \text{for all } x \in \mathscr{X} \text{ with } x \geq \underline{x},$$

where f^χ is the restriction of f to χ.

Proof That $(I + \Delta \underline{Q}) f$ is constant after $\overline{x}$ is obvious: for any y in $\mathscr{X}$ with $y \geq \overline{x}$,

$$[(I + \Delta \underline{Q}) f](y) = f(y) + \Delta [\underline{Q} f](y) = f(y) + \Delta \min\{\lambda f(y+1) - \lambda f(y) : \lambda \in [\underline{\lambda}, \overline{\lambda}]\} = f(\overline{x}) + \Delta 0 = f(\overline{x}),$$

where for the penultimate equality we use that f is constant after $\overline{x}$.

To verify the second part of the statement, we fix any x in $\mathscr{X}$ with $x \geq \underline{x}$. Observe first that if $x \geq \overline{x}$, then

$$[(I + \Delta \underline{Q}) f](x) = f(\overline{x}) = f^\chi(\overline{x}) = f^\chi(\overline{x}) + \Delta [\underline{Q}^\chi f^\chi](\overline{x}) = [(I + \Delta \underline{Q}^\chi) f^\chi](\overline{x}),$$

where the first equality follows from the first part of our proof. If $x < \overline{x}$, then

$$\begin{aligned} [(I + \Delta \underline{Q}) f](x) &= f(x) + \Delta [\underline{Q} f](x) = f(x) + \Delta \min\{\lambda f(x+1) - \lambda f(x) : \lambda \in [\underline{\lambda}, \overline{\lambda}]\} \\ &= f^\chi(x) + \Delta \min\{\lambda f^\chi(x+1) - \lambda f^\chi(x) : \lambda \in [\underline{\lambda}, \overline{\lambda}]\} = [(I + \Delta \underline{Q}^\chi) f^\chi](x). \end{aligned}$$

∎

Lemma 90 *Fix some n in $\mathbb{N}$ and, for all i in $\{1, \ldots, n\}$, a Δ_i in $\mathbb{R}_{\geq 0}$. Then for any f in $\mathscr{L}^c$ that is constant after $\overline{x}$, $\prod_{i=1}^{n} (I + \Delta_i \underline{Q}) f$ is constant after $\overline{x}$. Furthermore, if we fix any $\underline{x}$ in $\mathscr{X}$ with $\underline{x} \leq \overline{x}$ and let $\chi := \{y \in \mathscr{X} : \underline{x} \leq x \leq \overline{x}\}$, then*

$$\left[\prod_{i=1}^{n} (I + \Delta_i \underline{Q}) f \right](x) = \begin{cases} \left[\prod_{i=1}^{n} (I + \Delta_i \underline{Q}^\chi) f^\chi \right](x) & \text{if } x \leq \overline{x} \\ f(\overline{x}) & \text{if } x \geq \overline{x} \end{cases} \qquad \text{for all } x \in \mathscr{X} \text{ with } x \geq \underline{x},$$

where f^χ is the restriction of f to χ.

Proof We provide a proof by induction. First, observe that for $n = 1$, the stated follows immediately from Lemma 89.

Second, fix some n in $\mathbb{N}$ with $n \geq 2$ and assume that the stated holds for $1 \leq n' < n$. We show that in this case, the stated then follows for n as well. Let $g := \prod_{i=2}^{n} (I + \Delta_i \underline{Q}) f$. Then by the induction hypothesis for $n' = n - 1$, g is constant after $\overline{x}$, and

$$g(x) = \left[\prod_{i=2}^{n} (I + \Delta_i \underline{Q}) f \right](x) = \begin{cases} \left[\prod_{i=2}^{n} (I + \Delta_i \underline{Q}^\chi) f^\chi \right](x) & \text{if } x \leq \overline{x}, \\ f(\overline{x}) & \text{if } x \geq \overline{x} \end{cases} \qquad \text{for all } x \in \mathscr{X} \text{ with } x \geq \underline{x}.$$

So clearly, g^{χ}, the restriction of g to χ, is equal to $\prod_{i=2}^{n}(I+\Delta_i\underline{Q}^{\chi})f^{\chi}$. Observe now that

$$g' := \prod_{i=1}^{n}(I+\Delta_i\underline{Q})f = (I+\Delta_1\underline{Q})g.$$

From the induction hypothesis for $n' = 1$, we know that g' is constant after $\bar{x}$ and that, for all x in $\mathscr{X}$ with $x \geq \underline{x}$,

$$g'(x) = \begin{cases} [(I+\Delta_1\underline{Q}^{\chi})g^{\chi}](x) & \text{if } x \leq \bar{x}, \\ g(\bar{x}) & \text{if } x \geq \bar{x} \end{cases} = \begin{cases} \left[\prod_{i=1}^{n}(I+\Delta_i\underline{Q}^{\chi})f^{\chi}\right](x) & \text{if } x \leq \bar{x}, \\ f(\bar{x}) & \text{if } x \geq \bar{x}, \end{cases}$$

where the second equality holds because $g^{\chi} = \prod_{i=2}^{n}(I+\Delta_i\underline{Q}^{\chi})f^{\chi}$ and $g(\bar{x}) = f(\bar{x})$. ∎

Proof of Proposition 12 Fix any x in $\mathscr{X}$ with $x \geq \underline{x}$. First, observe that if $x \geq \bar{x}$, then

$$[\underline{T}_t^s f](x) = [\underline{T}_t^s f_x'](0) = [\underline{T}_t^s(f(\bar{x}))](0) = f(\bar{x}),$$

where we let $f_x': \mathscr{X} \to \mathbb{R}: z \mapsto f(x+z)$, and where the first equality follows from Lemma 53, the second equality holds because clearly $f_x' = f(\bar{x})$ and the third equality follows from (LT6). The obtained equality clearly agrees with the stated.

Next, we consider the case that $x \leq \bar{x}$. Fix any ε in $\mathbb{R}_{>0}$, and choose any ε' in $\mathbb{R}_{>0}$ such that $2\|f\|\varepsilon' \leq \varepsilon$. Then by Theorem 45—in combination with Equation (20)—and Proposition 31—in combination with Lemma 85—there is a sequence u in $\mathscr{U}_{[t,s]}$ such that

$$\|\underline{T}_t^s - \Phi_u\| \leq \varepsilon' \qquad \text{and} \qquad \|\underline{T}_{t,s}^{\chi} - \Phi_u^{\chi}\| \leq \varepsilon'. \tag{82}$$

Observe now that

$$\begin{aligned}
\left|[\underline{T}_t^s f](x) - [\underline{T}_{t,s}^{\chi} f^{\chi}](x)\right| &= \left|[\underline{T}_t^s f](x) - [\Phi_u f](x) + [\Phi_u f](x) - [\underline{T}_{t,s}^{\chi} f^{\chi}](x)\right| \\
&\leq \left|[\underline{T}_t^s f](x) - [\Phi_u f](x)\right| + \left|[\Phi_u f](x) - [\underline{T}_{t,s}^{\chi} f^{\chi}](x)\right| \\
&= \left|[\underline{T}_t^s f](x) - [\Phi_u f](x)\right| + \left|[\Phi_u^{\chi} f^{\chi}](x) - [\underline{T}_{t,s}^{\chi} f^{\chi}](x)\right| \\
&\leq \|\underline{T}_t^s f - \Phi_u f\| + \|\Phi_u^{\chi} f^{\chi} - \underline{T}_{t,s}^{\chi} f^{\chi}\| \leq \|\underline{T}_t^s - \Phi_u\|\|f\| + \|\Phi_u^{\chi} - \underline{T}_{t,s}^{\chi}\|\|f^{\chi}\| \\
&\leq \varepsilon'\|f\| + \varepsilon'\|f^{\chi}\| \leq 2\varepsilon'\|f\| \leq \varepsilon,
\end{aligned}$$

where the second equality follows from Lemma 90, the third inequality follows from (NH4), the fourth inequality follows from Equation (82), the penultimate inequality holds because clearly $\|f^{\chi}\| \leq \|f\|$ and the final inequality is precisely our condition on ε'. Since ε was any arbitrary positive real number, we conclude from this inequality that $[\underline{T}_t^s f](x) = [\underline{T}_{t,s}^{\chi} f^{\chi}](x)$, as required. ∎

Proof of Proposition 13 Fix any ε in $\mathbb{R}_{>0}$. To prove the stated, we need to verify that

$$(\exists x^{\star} \in \mathscr{X})(\forall \bar{x} \in \mathscr{X}, \bar{x} \geq x^{\star})\,|\underline{P}_t^s(f\,|\,x) - \underline{P}_t^s(\mathbb{I}_{\leq\bar{x}}f + \mathbb{I}_{>\bar{x}}f(\bar{x})\,|\,x)| = |[\underline{T}_t^s f](x) - [\underline{T}_t^s(\mathbb{I}_{\leq\bar{x}}f + f(\bar{x})\mathbb{I}_{>\bar{x}})](x)| \leq \varepsilon.$$

To that end, we fix any ε' in $\mathbb{R}_{>0}$ with $2\varepsilon'\|f\| \leq \varepsilon$ and we recall from Theorem 45 that there is a sequence $u = t_0,\ldots,t_n$ in $\mathscr{U}_{[t,s]}$ such that $\sigma(u)\|\underline{Q}\| \leq 2$ and

$$\|\underline{T}_t^s - \Phi_u\| \leq \varepsilon'.$$

Let $x^{\star} := x + n$, fix any $\bar{x}$ in $\mathscr{X}$ such that $\bar{x} \geq x^{\star}$ and let $f_{\bar{x}} := \mathbb{I}_{\leq\bar{x}}f + f(\bar{x})\mathbb{I}_{>\bar{x}}$. Then

$$\begin{aligned}
|[\underline{T}_t^s f](x) - [\underline{T}_t^s(\mathbb{I}_{\leq\bar{x}}f + f(\bar{x})\mathbb{I}_{>\bar{x}})](x)| &= |[\underline{T}_t^s f](x) - [\underline{T}_t^s f_{\bar{x}}](x)| = |[\underline{T}_t^s f](x) - [\Phi_u f](x) + [\Phi_u f](x) - [\underline{T}_t^s f_{\bar{x}}](x)| \\
&\leq |[\underline{T}_t^s f](x) - [\Phi_u f](x)| + |[\underline{T}_t^s f_{\bar{x}}](x) - [\Phi_u f](x)| \\
&= |[\underline{T}_t^s f](x) - [\Phi_u f](x)| + |[\underline{T}_t^s f_{\bar{x}}](x) - [\Phi_u f_{\bar{x}}](x)| \\
&\leq \|\underline{T}_t^s f - \Phi_u f\| + \|\underline{T}_t^s f_{\bar{x}} - \Phi_u f_{\bar{x}}\| \\
&\leq \|\underline{T}_t^s - \Phi_u\|\|f\| + \|\underline{T}_t^s - \Phi_u\|\|\mathbb{I}_{\leq\bar{x}}f + f(\bar{x})\mathbb{I}_{>\bar{x}}\| \\
&\leq \varepsilon'\|f\| + \varepsilon'\|\mathbb{I}_{\leq\bar{x}}f + f(\bar{x})\mathbb{I}_{>\bar{x}}\| \leq 2\varepsilon'\|f\| \leq \varepsilon,
\end{aligned}$$

as required. In this expression, the third equality follows from Lemma 47, the third inequality follows from (NH4) and the fifth inequality follows from the obvious inequality $\|\mathbb{I}_{\leq\bar{x}}f + f(\bar{x})\mathbb{I}_{>\bar{x}}\| \leq \|f\|$. ∎

Appendix H. Supplementary Material for Section 6

In this section of the Appendix, we will focus on expectations of the form $E_P(f(X_s) \,|\, X_u = x_u, X_t = x)$. Therefore, we first establish that $(f(X_s), X_u = x_u, X_t = x)$ belongs to the domain $\mathscr{G}$ of E_P. By definition of $\mathscr{G}$, this is true if $f(X_s)$ is bounded below and $\mathscr{F}_{u \cup t}$ measurable—that is, belongs to $\mathscr{G}_{u \cup t}$.

Proof of Lemma 14 Fix any α in $[\inf f, +\infty)$. Then

$$\{f(X_s) > \alpha\} = \{\omega \in \Omega \colon f(\omega(s)) > \alpha\} = (X_s \in B_\alpha),$$

where $B_\alpha := \{y \in \mathscr{X} \colon f(y) > \alpha\}$. As $B_\alpha \subseteq \mathscr{X}$, it follows from Equation (3), the fundamental event $(X_s \in B_\alpha)$ is an element of the field $\mathscr{F}_u$. Because α was an arbitrary real number in $[\inf f, +\infty)$, we infer from this that $f(X_s)$ is $\mathscr{F}_{u \cup t}$ measurable. ∎

H.1. With Respect to the Consistent Poisson Processes

Section 6.1 of the main text contains only one (implicit) result: Equation (18). We will here formally establish this—not exactly immediate—consequence of Theorem 6 in Proposition 92. First, however, we state a helpful—and essentially well-known—technical lemma.

Lemma 91 *Consider the Poisson process P with rate λ in $\mathbb{R}_{\geq 0}$. For any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, u in $\mathscr{U}_{<t}$, (x_u, x) in $\mathscr{X}_{u \cup t}$, the measure*

$$\mu \colon \mathscr{E}(\mathscr{X}) \to \mathbb{R} \colon A \mapsto \mu(A) := P(X_s \in A \,|\, X_u = x_u, X_t = x). \tag{83}$$

is the σ-additive probability measure corresponding to the probability mass function

$$\pi \colon \mathscr{X} \to \mathbb{R} \colon y \mapsto \pi(y) := \begin{cases} \psi_{\lambda(s-t)}(y-x) & \text{if } y \geq x, \\ 0 & \text{otherwise.} \end{cases}$$

Proof That μ is a probability measure on $\mathscr{E}(\mathscr{X})$ follows immediately from the fact that P is a coherent conditional probability. Therefore, we only need to verify that μ is σ-additive and corresponds to the probability mass function π. To that end, we observe that, for any y in $\mathscr{X}$,

$$\mu(\{y\}) = P(X_s = y \,|\, X_u = x_u, X_t = x) = \begin{cases} \psi_{\lambda(s-t)}(y-x) & \text{if } y \geq x \\ 0 & \text{otherwise} \end{cases} = \pi(y), \tag{84}$$

where the second equality follows from Theorem 6. Therefore, the stated holds if we can prove that, for any subset A of $\mathscr{X}$,

$$\mu(A) = \sum_{y \in A} \pi(y). \tag{85}$$

To verify this, we fix any subset A of $\mathscr{X}$, and distinguish two cases: A is finite and A is infinite. In the first case, it follows from the finite additivity of P and Equation (84) that

$$\mu(A) = P(X_s \in A \,|\, X_u = x_u, X_t = x) = \sum_{y \in A} P(X_s = A \,|\, X_u = x_u, X_t = x) = \sum_{y \in A} \mu(\{y\}) = \sum_{y \in A} \pi(y), \tag{86}$$

as required.

The case that A is infinite is slightly more involved. Observe first that, for any $\bar{z}$ in $\mathscr{X}$ with $\bar{z} \geq x$,

$$P(X_s \geq \bar{z} + 1 \,|\, X_u = x_u, X_t = x) = 1 - P(X_s \leq \bar{z} \,|\, X_u = x_u, X_t = x)$$

$$= 1 - \sum_{y=x}^{\bar{z}} P(X_s = y \,|\, X_u = x_u, X_t = x) = 1 - \sum_{y=x}^{\bar{z}} \psi_{\lambda(s-t)}(y-x),$$

where for the penultimate equality we have used Lemma 73 and for the last equality we have used Theorem 6. Fix any ε in $\mathbb{R}_{>0}$. As the Poisson distribution $\psi_{\lambda(s-t)}$ is normed, there is some $\bar{z}$ in $\mathscr{X}$ such that $\bar{z} \geq x$ and

$$1 - \varepsilon \leq \sum_{y=x}^{\bar{z}} \psi_{\lambda(s-t)}(y-x) \leq 1.$$

Fix any such $\bar{z}$ such that there is at least one y in A with $y \leq \bar{z}$. Then from the left inequality and the previous equality, it now follows that

$$P(X_s \geq \bar{z} + 1 \mid X_u = x_u, X_t = x) \leq \varepsilon. \tag{87}$$

Let $A' := \{y \in A : y \leq \bar{z}\}$. As $A' \subseteq A \subseteq A' \cup \{y \in \mathscr{X} : y \geq \bar{z} + 1\}$ and $A' \cap \{y \in \mathscr{X} : y \geq \bar{z} + 1\} = \emptyset$, it follows from the monotonicity and finite additivity of P that

$$P(X_s \in A' \mid X_u = x_u, X_t = x) \leq P(X_s \in A \mid X_u = x_u, X_t = x) \leq P(X_s \in A' \mid X_u = x_u, X_t = x) + \varepsilon,$$

where we have also used Equation (87) for the right inequality. As A' is finite, it follows from these two inequalities and Equation (86) that

$$\sum_{y \in A'} \pi(y) \leq \mu(A) = P(X_s \in A \mid X_u = x_u, X_t = x) \leq \sum_{y \in A'} \pi(y) + \varepsilon.$$

Because ε was an arbitrary positive real number, we conclude from these inequalities that $\mu(A) = \sum_{y \in A} \pi(y)$, as required. ∎

Proposition 92 *Consider the Poisson process P with rate λ in $\mathbb{R}_{\geq 0}$. Then for any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, u in $\mathscr{U}_{<t}$, (x_u, x) in $\mathscr{X}_{u \cup t}$ and f in $\mathscr{K}_b(\mathscr{X})$,*

$$E_P(f(X_s) \mid X_u = x_u, X_t = x) = \sum_{y=x}^{+\infty} f(y) \psi_{\lambda(s-t)}(y - x).$$

Proof Recall from Section 2.5 that

$$E_P(f(X_s) \mid X_u = x_u, X_t = x) = \inf f + \int_{\inf f}^{\sup f} P(\{f(X_s) > \alpha\} \mid X_u = x_u, X_t = x) \, d\alpha.$$

Observe now that $\{f(X_s) > \alpha\} = (X_s \in A_\alpha)$ with $A_\alpha := \{y \in \mathscr{X} : f(y > \alpha)\} \subseteq \mathscr{X}$. Consequently,

$$E_P(f(X_s) \mid X_u = x_u, X_t = x) = \inf f + \int_{\inf f}^{\sup f} \mu(\{f > \alpha\}) \, d\alpha,$$

with $\{f > \alpha\} := \{y \in \mathscr{X} : f(y) > \alpha\}$ and where μ is the σ-additive measure defined as in Lemma 91:

$$\mu : \mathscr{E}(\mathscr{X}) \to \mathbb{R} : A \mapsto \mu(A) := P(X_s \in A \mid X_u = x_u, X_t = x). \tag{88}$$

Let $f' := f - \inf f$. It then follows from Levi's Monotone Convergence Theorem—with the sequence $\{f' \mathbb{I}_{\leq i}\}_{i \in \mathbb{N}}$—that

$$\inf f + \int_{\inf f}^{\sup f} \mu(\{f > \alpha\}) \, d\alpha = \inf f + \int_0^{\sup f'} \mu(\{f' > \alpha\}) \, d\alpha = \inf f + \sum_{y=0}^{+\infty} f'(y) \mu(\{y\}) = \sum_{y=0}^{+\infty} f(y) \mu(\{y\}),$$

where for the final equality we have also used that $1 = \sum_{y=0}^{+\infty} \mu(\{y\})$. The stated now follows from this equality and Lemma 91. ∎

H.2. With Respect to the Consistent Counting Processes

H.2.1. SOME NOTATION AND INTERMEDIARY TECHNICAL RESULTS

Before we can prove our main results, we introduce some additional notation and establish some useful technical results. Consider a counting process P. Fix some t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, some u in $\mathscr{U}_{<t}$, some x_u in $\mathscr{X}_u$ and some $\underline{x}, \bar{x}$ in $\mathscr{X}$ such that $x_{\max u} \leq \underline{x} \leq \bar{x}$, and let $\chi := \{x \in \mathscr{X} : \underline{x} \leq x \leq \bar{x}\}$. We now consider the operator $T_{x_u, t, s}^\chi : \mathscr{L}(\chi) \to \mathscr{L}(\chi)$, defined for all f^χ in $\mathscr{L}(\chi)$ by

$$[T_{x_u, t, s}^\chi f^\chi](x) := \sum_{y=x}^{\bar{x}-1} f^\chi(y) P(X_s = y \mid X_u = x_u, X_t = x) + f^\chi(\bar{x}) P(X_s \geq \bar{x} \mid X_u = x_u, X_t = x) \quad \text{for all } x \in \chi, \tag{89}$$

where for notational simplicity we let the empty sum equal zero.

Lemma 93 *Consider a counting process P. Fix some t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, some u in $\mathscr{U}_{<t}$, some x_u in $\mathscr{X}_u$ and some $\underline{x}, \overline{x}$ in $\mathscr{X}$ with $x_{\max u} \leq \underline{x} \leq \overline{x}$. If we let $\chi := \{x \in \mathscr{X} : \underline{x} \leq x \leq \overline{x}\}$, then $T^{\chi}_{x_u,t,s} \colon \mathscr{L}(\chi) \to \mathscr{L}(\chi)$ is a linear counting transformation and*

$$[T^{\chi}_{x_u,t,s} f^{\chi}](\overline{x}) = f^{\chi}(\overline{x}) \quad \text{for all } f^{\chi} \in \mathscr{L}(\chi).$$

Proof We first verify that $T^{\chi}_{x_u,t,s}$ is a linear counting transformation. To that end, we just check the four conditions of Definition 54. That $T^{\chi}_{x_u,t,s}$ is linear transformation—that is, (T1) and (T2)—follows immediately from Equation (89). To verify (T3), we fix some f^{χ} in $\mathscr{L}(\chi)$ and some x in χ, and observe that

$$[T^{\chi}_{x_u,t,s} f^{\chi}](x) = \sum_{y=x}^{\overline{x}-1} f^{\chi}(y) P(X_s = y \mid X_u = x_u, X_t = x) + f^{\chi}(\overline{x}) P(X_s \geq \overline{x} \mid X_u = x_u, X_t = x)$$

$$\geq \sum_{y=x}^{\overline{x}-1} (\inf f^{\chi}) P(X_s = y \mid X_u = x_u, X_t = x) + (\inf f^{\chi}) P(X_s \geq \overline{x} \mid X_u = x_u, X_t = x),$$

where the inequality holds because we replace each term by a lower term. From this, it now follows that

$$[T^{\chi}_{x_u,t,s} f^{\chi}](x) = (\inf f^{\chi}) P(X_s \geq x \mid X_u = x_u, X_t = x) = (\inf f^{\chi}) P(X_u = x_u, X_t = x, X_s \geq x \mid X_u = x_u, X_t = x)$$

$$= \inf f^{\chi} P(X_u = x_u, X_t = x \mid X_u = x_u, X_t = x) = \inf f^{\chi},$$

where for the first equality we have used the additivity of P and for the second equality we have used (A1). The final condition (T4) again follows immediately from Equation (89).

To verify the second part of the statement, we fix some f^{χ} in $\mathscr{L}(\chi)$ and observe that

$$[T^{\chi}_{x_u,t,s} f^{\chi}](\overline{x}) = f^{\chi}(\overline{x}) P(X_s \geq \overline{x} \mid X_u = x_u, X_t = \overline{x}) = f^{\chi}(\overline{x}) P(X_u = x_u, X_t = \overline{x}, X_s \geq \overline{x} \mid X_u = x_u, X_t = \overline{x})$$

$$= f^{\chi}(\overline{x}) P(X_u = x_u, X_t = \overline{x} \mid X_u = x_u, X_t = \overline{x}) = f^{\chi}(\overline{x}),$$

where for the third equality we have again used (A1). ∎

The following result is heavily inspired by [8, Proposition 4.7], and is essential to the proof of Lemma 95.

Lemma 94 *Consider a counting process P that is consistent with the rate interval Λ. Fix some t in $\mathbb{R}_{\geq 0}$, u in $\mathscr{U}_{<t}$, x_u in $\mathscr{X}_u$ and $\underline{x}, \overline{x}$ in $\mathscr{X}$ with $x_{\max u} \leq \underline{x} \leq \overline{x}$, and let $\chi := \{x \in \mathscr{X} : \underline{x} \leq x \leq \overline{x}\}$. Then*

$$(\forall \varepsilon \in \mathbb{R}_{>0})(\exists \delta \in \mathbb{R}_{>0})(\forall \Delta \in \mathbb{R}_{>0}, \Delta < \delta)(\exists Q^{\chi} \in \mathscr{Q}^{\chi}) \left\| \frac{T^{\chi}_{x_u,t,t+\Delta} - I}{\Delta} - Q^{\chi} \right\| \leq \varepsilon$$

and, if $t > 0$,

$$(\forall \varepsilon \in \mathbb{R}_{>0})(\exists \delta \in \mathbb{R}_{>0})(\forall \Delta \in \mathbb{R}_{>0}, \Delta < \delta)(\exists Q^{\chi} \in \mathscr{Q}^{\chi}) \left\| \frac{T^{\chi}_{x_u,t-\Delta,t} - I}{\Delta} - Q^{\chi} \right\| \leq \varepsilon.$$

Proof Observe that if $\underline{x} = \overline{x}$, then χ is the singleton containing $\underline{x} = \overline{x}$. In this case, $\mathscr{Q}^{\chi}(\{\underline{x}\})$ because 0 is the only linear transformation on $\mathscr{L}(\chi) = \mathscr{L}(\{\underline{x}\})$ that can satisfy all four conditions of Definition 27—and specifically (LR3). Observe now that for any Δ in $\mathbb{R}_{>0}$,

$$\left\| \frac{T^{\chi}_{x_u,t,t+\Delta} - I}{\Delta} - 0 \right\| = \frac{T^{\chi}_{x_u,t,t+\Delta}(\underline{x},\underline{x}) - I(\underline{x},\underline{x})}{\Delta} - 0(\underline{x},\underline{x}) = \frac{T^{\chi}_{x_u,t,t+\Delta}(\underline{x},\underline{x}) - 1}{\Delta} = \frac{[T^{\chi}_{x_u,t,t+\Delta} \mathbb{I}_{\underline{x}}](\underline{x}) - 1}{\Delta} = \frac{\mathbb{I}_{\underline{x}}(\underline{x}) - 1}{\Delta} = 0,$$

where for the second equality we have used Equation (35) and for the penultimate equality we have used Lemma 93. Hence, the first part of the stated is trivially verified. Similar reasoning yields that the first part is trivially satisfies for all Δ in $\mathbb{R}_{>0}$ such that $\Delta < t - \max u$.

Next, we consider the alternative case that $\underline{x} < \overline{x}$. We here only prove the first inequality of the stated, the proof of the second inequality is entirely analogous. Fix any arbitrary ε in $\mathbb{R}_{>0}$, and choose some ε_1 and ε_2 in $\mathbb{R}_{>0}$ such that $2\varepsilon_1 + |\chi|\varepsilon_2 \leq \varepsilon$. Because P is consistent with Λ, it follows from Equation (10) that for any x in $\chi' := \chi \setminus \{\overline{x}\}$, there is a $\delta_{1,x}$ in $\mathbb{R}_{>0}$ such that for all Δ in $\mathbb{R}_{>0}$, there is a $\lambda_{x,\Delta}$ in Λ such that

$$\left| \frac{P(X_{t+\Delta} = x+1 \mid X_u = x_u, X_t = x)}{\Delta} - \lambda_{x,\Delta} \right| \leq \varepsilon_1. \tag{90}$$

Additionally, as P is a counting process it follows from (CP2) that for all x in χ' there is a $\delta_{2,x}$ in $\mathbb{R}_{>0}$ such that

$$(\forall \Delta \in \mathbb{R}_{>0}, \Delta < \delta_{2,x}) \; 0 \leq \frac{P(X_{t+\Delta} \geq x+2 \,|\, X_u = x_u, X_t = x)}{\Delta} \leq \varepsilon_2. \tag{91}$$

Observe that clearly $(X_{t+\Delta} = y) \subseteq (X_{t+\Delta} \geq x+2)$ if $x+2 \leq y$, whence it follows from Equation (91) and the monotonicity of P that, for all x and y in χ with $x+2 \leq y$,

$$(\forall \Delta \in \mathbb{R}_{>0}, \Delta < \delta_{2,x}) \; 0 \leq \frac{P(X_{t+\Delta} = y \,|\, X_u = x_u, X_t = x)}{\Delta} \leq \varepsilon_2. \tag{92}$$

Similarly, it is clear that $(X_{t+\Delta} \geq \bar{x}) \subseteq (X_{t+\Delta} \geq x+2)$ if $x+2 \leq \bar{x}$, whence it follows that for all x in χ with $x+2 \leq \bar{x}$,

$$(\forall \Delta \in \mathbb{R}_{>0}, \Delta < \delta_{2,x}) \; 0 \leq \frac{P(X_{t+\Delta} \geq \bar{x} \,|\, X_u = x_u, X_t = x)}{\Delta} \leq \varepsilon_2. \tag{93}$$

Additionally, we use the finite additivity of P and Equations (90) and (91) with $x = \bar{x} - 1$, to yield

$$(\forall \Delta \in \mathbb{R}_{>0}, \Delta < \delta_{1,\bar{x}-1}, \Delta < \delta_{2,\bar{x}-1}) \; \left| \frac{P(X_{t+\Delta} \geq \bar{x} \,|\, X_u = x_u, X_t = \bar{x}-1)}{\Delta} + \lambda_{\bar{x}-1,\Delta} \right| \leq \varepsilon_1 + \varepsilon_2. \tag{94}$$

To bound $P(X_{t+\Delta} = x \,|\, X_u = x_u, X_t = x)$, we recall from Lemma 73 that

$$P(X_{t+\Delta} = x \,|\, X_u = x_u, X_t = x) = 1 - P(X_{t+\Delta} = x+1 \,|\, X_u = x_u, X_t = x) - P(X_{t+\Delta} \geq x+2 \,|\, X_u = x_u, X_t = x).$$

We now combine this equality and Equations (90) and (91), to yield that for all x in χ',

$$(\forall \Delta \in \mathbb{R}_{>0}, \Delta < \delta_{1,x}, \Delta < \delta_{2,x}) \; \left| \frac{P(X_{t+\Delta} = x \,|\, X_u = x_u, X_t = x) - 1}{\Delta} + \lambda_{x,\Delta} \right| \leq \varepsilon_1 + \varepsilon_2. \tag{95}$$

Let $\delta := \min \cup_{x \in \chi'} \{\delta_{1,x}, \delta_{2,x}\}$, and fix any Δ in $\mathbb{R}_{>0}$ with $\Delta < \delta$. Let Q_Δ^χ be the element of $\mathscr{Q}^\chi$ that is characterised by the sequence $\{\lambda_{x,\Delta}\}_{x \in \chi'}$ in Λ, as explained in Lemma 86. To verify the stated, we now set out to bound

$$\left\| \frac{T_{x_u,t,t+\Delta}^\chi - I}{\Delta} - Q_\Delta^\chi \right\| = \max_{x \in \chi} \sum_{y \in \chi} \left| \frac{T_{x_u,t,t+\Delta}^\chi(x,y) - I(x,y)}{\Delta} - Q_\Delta^\chi(x,y) \right|,$$

where the equality follows from Equation (35). To that end, we take a closer look at the expression on the right for all x in χ. Observe first that

$$\sum_{y \in \chi} \left| \frac{T_{x_u,t,t+\Delta}^\chi(x,y) - I(x,y)}{\Delta} - Q_\Delta^\chi(x,y) \right| = \sum_{y=x}^{\bar{x}} \left| \frac{T_{x_u,t,t+\Delta}^\chi(x,y) - I(x,y)}{\Delta} - Q_\Delta^\chi(x,y) \right|,$$

because, by construction, $T_{x_u,t,t+\Delta}^\chi(x,y) = 0 = Q^\chi(x,y)$ and, by definition, $I(x,y) = 0$ for all y in χ with $y < x$. We now use the definitions of $T_{x_u,t,t+\Delta}^\chi$ and Q_Δ^χ to rewrite the expression on the right hand side. If $x+2 \leq \bar{x}$, then this yields

$$\begin{aligned}
\sum_{y \in \chi} \left| \frac{T_{x_u,t,t+\Delta}^\chi(x,y) - I(x,y)}{\Delta} - Q_\Delta^\chi(x,y) \right| &= \left| \frac{P(X_{t+\Delta} = x \,|\, X_u = x_u, X_t = x) - 1}{\Delta} + \lambda_{x,\Delta} \right| \\
&+ \left| \frac{P(X_{t+\Delta} = x+1 \,|\, X_u = x_u, X_t = x)}{\Delta} - \lambda_{x,\Delta} \right| \\
&+ \sum_{y=x+2}^{\bar{x}} \left| \frac{P(X_{t+\Delta} = y \,|\, X_u = x_u, X_t = x)}{\Delta} \right| \\
&+ \left| \frac{P(X_{t+\Delta} \geq \bar{x} \,|\, X_u = x_u, X_t = x)}{\Delta} \right| \\
&\leq \varepsilon_1 + \varepsilon_2 + \varepsilon_1 + (\bar{x} - x - 2)\varepsilon_2 + \varepsilon_2 \leq 2\varepsilon_1 + |\chi|\varepsilon_2 \leq \varepsilon,
\end{aligned}$$

where for the first inequality we have used Equations (90), (92), (93) and (95). If $x+1 = \bar{x}$, then this yields

$$\sum_{y \in \chi} \left| \frac{T^{\chi}_{x_u,t,t+\Delta}(x,y) - I(x,y)}{\Delta} - Q^{\chi}_{\Delta}(x,y) \right| = \left| \frac{P(X_{t+\Delta} = x \mid X_u = x_u, X_t = x) - 1}{\Delta} + \lambda_{x,\Delta} \right|$$
$$+ \left| \frac{P(X_{t+\Delta} \geq \bar{x} \mid X_u = x_u, X_t = x)}{\Delta} - \lambda_{x,\Delta} \right|$$
$$\leq \varepsilon_1 + \varepsilon_2 + \varepsilon_1 + \varepsilon_2 \leq 2\varepsilon_1 + |\chi|\varepsilon_2 \leq \varepsilon,$$

where for the first inequality we have used Equations (94) and (95). Finally, if $x = \bar{x}$, then it follows from Lemma 93 that

$$\sum_{y \in \chi} \left| \frac{T^{\chi}_{x_u,t,t+\Delta}(x,y) - I(x,y)}{\Delta} - Q^{\chi}_{\Delta}(x,y) \right| = \left| \frac{[T^{\chi}_{x_u,t,t+\Delta}\mathbb{I}_{\bar{x}}](\bar{x}) - 1}{\Delta} \right| = \left| \frac{\mathbb{I}_{\bar{x}}(\bar{x}) - 1}{\Delta} \right| = 0 \leq \varepsilon.$$

From these three cases, we infer that

$$\left\| \frac{T^{\chi}_{x_u,t,t+\Delta} - I}{\Delta} - Q^{\chi}_{\Delta} \right\| = \max_{x \in \chi} \sum_{y \in \chi} \left| \frac{T^{\chi}_{x_u,t,t+\Delta}(x,y) - I(x,y)}{\Delta} - Q^{\chi}_{\Delta}(x,y) \right| \leq \varepsilon,$$

as required. ∎

We now use Lemma 94 to establish a result—heavily inspired by [8, Lemma F.1]—that will be essential in the proof of Proposition 98.

Lemma 95 *Consider a counting process P that is consistent with $\Lambda = [\underline{\lambda}, \overline{\lambda}]$, some t, s in $\mathbb{R}_{\geq 0}$ with $t < s$, some u in $\mathscr{U}_{<t}$ and some x_u in $\mathscr{X}_u$. Fix some $\underline{x}, \bar{x}$ in $\mathscr{X}$ with $x_{\max u} \leq \underline{x} \leq \bar{x}$, and let $\chi := \{x \in \mathscr{X} : \underline{x} \leq x \leq \bar{x}\}$. Then for all ε, δ in $\mathbb{R}_{>0}$, there is a $v = t_0, \ldots, t_n$ in $\mathscr{U}_{[t,s]}$ such that $\sigma(v) < \delta$ and*

$$(\forall i \in \{1, \ldots, n\})(\exists Q^{\chi}_i \in \mathscr{Q}^{\chi}) \, \|T^{\chi}_{x_u,t_{i-1},t_i} - (I + \Delta_i Q^{\chi}_i)\| \leq \Delta_i \varepsilon.$$

Proof Our proof is almost entirely equivalent to the proof of [8, Lemma F.1]; the only difference is that we invoke Lemma 94 instead of [8, Proposition 4.7]. Therefore, and also because it is rather lengthy, we have chosen to omit the proof. ∎

H.2.2. EVENTUALLY CONSTANT FUNCTIONS

Before we consider general bounded functions, we first limit ourselves to eventually constant functions. We first establish the following useful intermediary result.

Lemma 96 *Consider an f in $\mathscr{L}^c(\mathscr{X})$ that is constant from $\bar{x}$, some s in $\mathbb{R}_{\geq 0}$ and some u in $\mathscr{U}$ such that $\max u \leq s$. Then*

$$f(X_s) = \sum_{x=0}^{\bar{x}-1} f(x)\mathbb{I}_{(X_s = x)} + f(\bar{x})\mathbb{I}_{(X_s \geq \bar{x})}$$

such that $f(X_s)$ is an $\mathscr{F}_u$-simple function.

Proof It is easy to see that

$$f(X_s) = \sum_{y=0}^{\bar{x}-1} f(y)\mathbb{I}_y(X_s) + f(\bar{x})\mathbb{I}_{\geq \bar{x}}(X_s) = \sum_{y=0}^{\bar{x}-1} f(y)\mathbb{I}_{(X_s = y)} + f(\bar{x})\mathbb{I}_{(X_s \geq \bar{x})}.$$

As all events in the indicators are clearly contained in $\mathscr{F}_u$, it furthermore follows that $f(X_s)$ is an $\mathscr{F}_u$-simple function. ∎

The following result is inspired by [8, Lemma F.2], and one of our main reasons for introducing the notation $T^{\chi}_{x_u,t,s}$.

Lemma 97 *Consider a counting process P. Fix some f in $\mathscr{L}^{c}(\mathscr{X})$ that is constant from $\bar{x}$, some t,s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, some u in $\mathscr{U}_{<t}$ and some (x_u, x_t) in $\mathscr{X}_{u \cup t}$. If $x_t < \bar{x}$, then for any $v = t_0, t_1, \ldots, t_n$ in $\mathscr{U}_{[t,s]}$,*

$$E_P(f(X_s) \mid X_u = x_u, X_t = x_t) = \left[T^{\chi}_{x_u,t,t_1} \prod_{i=2}^{n} T^{\chi}_{x_u \cup t, t_{i-1}, t_i} f^{\chi} \right](x_t),$$

where $\chi := \{x \in \mathscr{X} : x_t \leq x \leq \bar{x}\}$, and f^{χ} is the restriction of f to χ.

Proof Fix some $\bar{x}$ in $\mathscr{X}$ such that f is constant starting from $\bar{x}$. Our proof is one by induction. First, it is an immediate consequence of Lemma 96 and Equation (4) that

$$E_P(f(X_s) \mid X_u = x_u, X_t = x_t) = \sum_{y=0}^{\bar{x}-1} f(y) P(X_s = y \mid X_u = x_u, X_t = x_t) + f(\bar{x}) P(X_s \geq \bar{x} \mid X_u = x_u, X_t = x_t).$$

We use Lemma 73, that f^{χ} is the restriction of f to χ and Equation (89), to yield

$$
\begin{aligned}
E_P(f(X_s) \mid X_u = x_u, X_t = x_t) &= \sum_{y=x_t}^{\bar{x}-1} f(y) P(X_s = y \mid X_u = x_u, X_t = x_t) + f(\bar{x}) P(X_s \geq \bar{x} \mid X_u = x_u, X_t = x_t) \\
&= \sum_{y=x_t}^{\bar{x}-1} f^{\chi}(y) P(X_s = y \mid X_u = x_u, X_t = x_t) + f^{\chi}(\bar{x}) P(X_s \geq \bar{x} \mid X_u = x_u, X_t = x_t) \\
&= [T^{\chi}_{x_u,t,s} f^{\chi}](x_t),
\end{aligned}
$$

as required.

For the induction step, we fix some n in $\mathbb{N}$ with $n \geq 2$ and assume that the stated holds for any sequence v of length $n' + 1$, with n' in $\mathbb{N}$ such that $1 \leq n' < n$. The stated then follows for any sequence v of length $n + 1$, as we will now prove. We start with applying the induction hypothesis to the sequence $t_0, t_2, \ldots, t_n$, to yield

$$
\begin{aligned}
E_P(f(X_s) \mid X_u = x_u, X_t = x_t) &= \left[T^{\chi}_{x_u,t,t_2} \prod_{i=3}^{n} T^{\chi}_{x_u \cup t, t_{i-1}, t_i} f^{\chi} \right](x_t) \\
&= \sum_{y_2=x_t}^{\bar{x}-1} P(X_{t_2} = y_2 \mid X_u = x_u, X_t = x_t) \left[\prod_{i=3}^{n} T^{\chi}_{x_u \cup t, t_{i-1}, t_i} f^{\chi} \right](y_2) \\
&\quad + P(X_{t_2} \geq \bar{x} \mid X_u = x_u, X_t = x_t) \left[\prod_{i=3}^{n} T^{\chi}_{x_u \cup t, t_{i-1}, t_i} f^{\chi} \right](\bar{x}). \tag{96}
\end{aligned}
$$

We now substitute the probabilities in this sum with an expanded expression From Lemma 74, it follows that, for any y_2 in $\chi \setminus \{\bar{x}\}$,

$$P(X_{t_2} = y_2 \mid X_u = x_u, X_t = x_t) = \sum_{y_1=x_t}^{y_2} P(X_{t_1} = y_1 \mid X_u = x_u, X_t = x_t) P(X_{t_2} = y_2 \mid X_u = x_u, X_t = x_t, X_{t_1} = y_1), \tag{97}$$

and

$$
\begin{aligned}
P(X_{t_2} \geq \bar{x} \mid X_u = x_u, X_t = x_t) &= \sum_{y_1=x_t}^{\bar{x}-1} P(X_{t_1} = y_1 \mid X_u = x_u, X_t = x_t) P(X_{t_2} \geq \bar{x} \mid X_u = x_u, X_t = x_t, X_{t_1} = y_1) \\
&\quad + P(X_{t_1} \geq \bar{x} \mid X_u = x_u, X_t = x_t). \tag{98}
\end{aligned}
$$

We substitute Equations (97) and (98) in Equation (96), to yield

$$
\begin{aligned}
&E_P(f(X_s) \mid X_u = x_u, X_t = x_t) \\
&= \sum_{y_2=x_t}^{\bar{x}-1} \sum_{y_1=x_t}^{y_2} P(X_{t_1} = y_1 \mid X_u = x_u, X_t = x_t) P(X_{t_2} = y_2 \mid X_u = x_u, X_t = x_t, X_{t_1} = y_1) \left[\prod_{i=3}^{n} T^{\chi}_{x_u \cup t, t_{i-1}, t_i} f^{\chi} \right](y_2)
\end{aligned}
$$

$$+ \sum_{y_1=x_t}^{\bar{x}-1} P(X_{t_1}=y_1 \,|\, X_u=x_u, X_t=x_t) P(X_{t_2} \geq \bar{x} \,|\, X_u=x_u, X_t=x_t, X_{t_1}=y_1) \left[\prod_{i=3}^{n} T^{\chi}_{x_{u \cup t}, t_{i-1}, t_i} f^{\chi} \right] (\bar{x})$$

$$+ P(X_{t_1} \geq \bar{x} \,|\, X_u=x_u, X_t=x_t) \left[\prod_{i=3}^{n} T^{\chi}_{x_{u \cup t}, t_{i-1}, t_i} f^{\chi} \right] (\bar{x}).$$

Recall from Lemma 93 that $\left[\prod_{i=2}^{n} T^{\chi}_{x_{u \cup t}, t_{i-1}, t_i} f^{\chi} \right] (\bar{x}) = \left[\prod_{i=3}^{n} T^{\chi}_{x_{u \cup t}, t_{i-1}, t_i} f^{\chi} \right] (\bar{x})$. We substitute this in the previous equality and change the summation order, to yield

$$E_P(f(X_s) \,|\, X_u=x_u, X_t=x_t)$$
$$= \sum_{y_1=x_t}^{\bar{x}-1} \sum_{y_2=y_1}^{\bar{x}-1} P(X_{t_1}=y_1 \,|\, X_u=x_u, X_t=x_t) P(X_{t_2}=y_2 \,|\, X_u=x_u, X_t=x_t, X_{t_1}=y_1) \left[\prod_{i=3}^{n} T^{\chi}_{x_{u \cup t}, t_{i-1}, t_i} f^{\chi} \right] (y_2)$$

$$+ \sum_{y_1=x_t}^{\bar{x}-1} P(X_{t_1}=y_1 \,|\, X_u=x_u, X_t=x_t) P(X_{t_2} \geq \bar{x} \,|\, X_u=x_u, X_t=x_t, X_{t_1}=y_1) \left[\prod_{i=3}^{n} T^{\chi}_{x_{u \cup t}, t_{i-1}, t_i} f^{\chi} \right] (\bar{x})$$

$$+ P(X_{t_1} \geq \bar{x} \,|\, X_u=x_u, X_t=x_t) \left[\prod_{i=2}^{n} T^{\chi}_{x_{u \cup t}, t_{i-1}, t_i} f^{\chi} \right] (\bar{x}).$$

Finally, we use the definition of $T^{\chi}_{x_{u \cup t}, t_1, t_2}$ and $T^{\chi}_{x_u, t, t_1}$ to yield the stated:

$$E_P(f(X_s) \,|\, X_u=x_u, X_t=x_t) = \sum_{y_1=x_t}^{\bar{x}-1} P(X_{t_1}=y_1 \,|\, X_u=x_u, X_t=x_t) \left[\prod_{i=2}^{n} T^{\chi}_{x_{u \cup t}, t_{i-1}, t_i} f^{\chi} \right] (y_2)$$

$$+ P(X_{t_1} \geq \bar{x} \,|\, X_u=x_u, X_t=x_t) \left[\prod_{i=2}^{n} T^{\chi}_{x_{u \cup t}, t_{i-1}, t_i} f^{\chi} \right] (\bar{x})$$

$$= \left[T^{\chi}_{x_u, t, t_1} \prod_{i=2}^{n} T^{\chi}_{x_{u \cup t}, t_{i-1}, t_i} f^{\chi} \right] (x_t).$$

∎

The following two propositions are essential to the proof of Theorem 101.

Proposition 98 *Consider any counting process P that is consistent with the rate interval Λ. Fix any f in $\mathscr{L}^c(\mathscr{X})$, any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$ and any u in $\mathscr{U}_{<t}$. Then for any (x_u, x) in $\mathscr{X}_{u \cup t}$,*

$$\underline{P}^s_t(f \,|\, x) \leq E_P(f(X_s) \,|\, X_u=x_u, X_t=x).$$

Proof Our proof is for a large part similar to that of [8, Proposition 8.1]. Let $\bar{x}$ be in $\mathscr{X}$ such that f is constant after $\bar{x}$. From Lemma 96 and Equation (4), it then follows that

$$E_P(f(X_s) \,|\, X_u=x_u, X_t=x) = \sum_{y=0}^{\bar{x}-1} f(y) P(X_s=y \,|\, X_u=x_u, X_t=x) + f(\bar{x}) P(X_s \geq \bar{x} \,|\, X_u=x_u, X_t=x),$$

where we let the empty sum equal zero. Recall from Lemma 73 that $P(X_s=y \,|\, X_u=x_u, X_t=x) = 0$ for all y in $\mathscr{X}$ with $y < x$. Therefore,

$$E_P(f(X_s) \,|\, X_u=x_u, X_t=x) = \sum_{y=x}^{\bar{x}-1} f(y) P(X_s=y \,|\, X_u=x_u, X_t=x) + f(\bar{x}) P(X_s \geq \bar{x} \,|\, X_u=x_u, X_t=x). \quad (99)$$

We distinguish two cases. First, we consider the case $x > \bar{x}$. In this case, $(X_u=x_u, X_t=x) \subseteq (X_s \geq \bar{x})$ due to (A1), such that $P(X_s \geq \bar{x} \,|\, X_u=x_u, X_t=x) = 1$ due to (P2). We substitute this equality in Equation (99), to yield

$$E_P(f(X_s) \,|\, X_u=x_u, X_t=x) = f(\bar{x}).$$

As furthermore $\underline{P}^s_t(f\,|\,x) = f(\bar{x})$ due to Proposition 12, this verifies the inequality of the statement in this case.

Next, we consider the case $x \le \bar{x}$, and distinguish two additional cases. If $s = t$, then the equality of the statement follows immediately from Equation (99), some obvious properties of counting processes and Proposition 48 (i):

$$E_P(f(X_s)\,|\,X_u = x_u, X_t = x) = f(x) = [If](x) = [\underline{T}^t_t f](x) = \underline{P}^s_t(f\,|\,x).$$

Hence, from now on we furthermore assume that $t < s$. Let $\chi := \{y \in \mathscr{X} : x \le y \le \bar{x}\}$, and recall from Proposition 12 that $\underline{P}^s_t(f\,|\,x) = [\underline{T}^{\chi}_{t,s} f^{\chi}](x)$. Thus, the equality of the statement is verified if we can show that

$$[\underline{T}^{\chi}_{t,s} f^{\chi}](x) \le E_P(f(X_s)\,|\,X_u = x_u, X_t = x). \tag{100}$$

To that end, we fix any ε in $\mathbb{R}_{>0}$, and choose any $\varepsilon_1, \varepsilon_2$ in $\mathbb{R}_{>0}$ such that $\varepsilon_1\|f^{\chi}\| \le \varepsilon/2$ and $\varepsilon_2(s-t)\|f^{\chi}\| \le \varepsilon/2$. It follows from Lemma 85 and Proposition 31 that there is some δ in $\mathbb{R}_{>0}$ such that $\delta\|\underline{Q}^{\chi}\| \le 2$ and

$$(\forall v \in \mathscr{U}_{[t,s]}, \sigma(v) \le \delta)\ \|\underline{T}^{\chi}_{t,s} - \Phi^{\chi}_v\| \le \varepsilon_1. \tag{101}$$

As P is consistent with $[\underline{\lambda}, \overline{\lambda}]$, it follows from Lemma 94 that there is some Δ_1 in $\mathbb{R}_{>0}$ with $\Delta_1 < \max\{\delta, s-t\}$ and some Q^{χ}_1 in $\mathscr{Q}^{\chi}$ such that

$$\|T^{\chi}_{x_u, t, t+\Delta} - (I + \Delta_1 Q^{\chi}_1)\| \le \Delta_1 \varepsilon_2. \tag{102}$$

Furthermore, since $t_1 := t + \Delta_1 < s$ by construction, it follows from Lemma 95 that there is a sequence $v' = t_1, t_2, \ldots, t_n$ in $\mathscr{U}_{[t_1, s]}$ with $\sigma(v') < \delta$ and some $Q^{\chi}_2, \ldots, Q^{\chi}_n$ in $\mathscr{Q}^{\chi}$ such that

$$(\forall i \in \{2, \ldots, n\})\ \|T^{\chi}_{x_{u\cup t}, t_{i-1}, t_i} - (I + \Delta_i Q^{\chi}_i)\| \le \Delta_i \varepsilon_2. \tag{103}$$

Let $t_0 := t$ and $v^\star := t_0, t_1, \ldots, t_n$. Recall from Lemma 97 that

$$E_P(f(X_s)\,|\,X_u = x_u, X_t = x) = \left[T^{\chi}_{x_u, t, t_1} \prod_{i=2}^{n} T^{\chi}_{x_{u\cup t}, t_{i-1}, t_i} f^{\chi}\right](x),$$

such that

$$\left|E_P(f(X_s)\,|\,X_u = x_u, X_t = x) - \left[\prod_{i=1}^{n}(I + \Delta_i Q^{\chi}_i) f^{\chi}\right](x)\right| = \left|\left[T^{\chi}_{x_u, t, t_1} \prod_{i=1}^{n} T^{\chi}_{x_{u\cup t}, t_i, t_{i+1}} f^{\chi}\right](x) - \left[\prod_{i=1}^{n}(I + \Delta_i Q^{\chi}_i) f^{\chi}\right](x)\right|.$$

We now use (NH4), Lemma 26 and Equations (102) and (103), to yield

$$\left|E_P(f(X_s)\,|\,X_u = x_u, X_t = x) - \left[\prod_{i=1}^{n}(I + \Delta_i Q^{\chi}_i) f^{\chi}\right](x)\right|$$

$$\le \left\|T^{\chi}_{x_u, t, t_1} \prod_{i=2}^{n} T^{\chi}_{x_{u\cup t}, t_{i-i}, t_i} f^{\chi} - \prod_{i=1}^{n}(I + \Delta_i Q^{\chi}_i) f^{\chi}\right\| \le \left\|T^{\chi}_{x_u, t, t_1} \prod_{i=2}^{n} T^{\chi}_{x_{u\cup t}, t_{i-1}, t_i} - \prod_{i=1}^{n}(I + \Delta_i Q^{\chi}_i)\right\| \|f^{\chi}\|$$

$$\le \|T^{\chi}_{x_u, t, t_1} - (I + \Delta_1 Q^{\chi}_1)\| \|f^{\chi}\| + \sum_{i=2}^{n} \|T^{\chi}_{x_{u\cup t}, t_{i-1}, t_i} - (I + \Delta_i Q^{\chi}_i)\| \|f^{\chi}\|$$

$$\le \sum_{i=1}^{n} \Delta_i \varepsilon_2 \|f^{\chi}\| = (s-t)\varepsilon_2 \|f^{\chi}\| \le \frac{\varepsilon}{2}.$$

Furthermore, it follows from (NH4) and Equation (101), which holds because $\sigma(v^\star) < \delta$ by construction, that

$$\left|[\underline{T}^{\chi}_{t,s} f^{\chi}](x) - [\Phi^{\chi}_{v^\star} f^{\chi}](x)\right| \le \|\underline{T}^{\chi}_{t,s} f^{\chi} - \Phi^{\chi}_{v^\star} f^{\chi}\| \le \|\underline{T}^{\chi}_{t,s} - \Phi^{\chi}_{v^\star}\| \|f^{\chi}\| \le \varepsilon_1 \|f^{\chi}\| \le \frac{\varepsilon}{2}.$$

Combining the two previous inequalities, we find that

$$[\underline{T}^{\chi}_{t,s} f^{\chi}](x) \le \left[\prod_{i=1}^{n}(I + \Delta_i \underline{Q}^{\chi}) f^{\chi}\right](x) + \frac{\varepsilon}{2} \le \left[\prod_{i=1}^{n}(I + \Delta_i Q^{\chi}_i) f^{\chi}\right](x) + \frac{\varepsilon}{2} \le E_P(f(X_s)\,|\,X_u = x_u, X_t = x) + \varepsilon,$$

where the second inequality holds due to Corollary 88. Since ε was an arbitrary positive real number, this implies Equation (100), as required. $\blacksquare$

Lemma 99 *Consider a counting transformation system $\mathcal{T}$ of the form of Corollary 66. Let $\tilde{P}$ be the real-valued map with domain*

$$\tilde{\mathcal{D}} := \{(X_{t+\Delta} = y, (X_u = x_u, X_t = x)) \in \mathcal{D}_{\mathrm{CP}} : t, \Delta \in \mathbb{R}_{\geq 0}, u \in \mathcal{U}_{<t}, (x_u, x) \in \mathcal{X}_{u \cup t}, y \in \mathcal{X}\} \cup \{(X_0 = x, \Omega) \in \mathcal{D}_{\mathrm{CP}} : x \in \mathcal{X}\},$$

that is defined for all t, Δ in $\mathbb{R}_{\geq 0}$, u in $\mathcal{U}_{<t}$, (x_u, x) in $\mathcal{X}_{u \cup t}$ and y in $\mathcal{X}$ as

$$\tilde{P}(X_{t+\Delta} = y \mid X_u = x_u, X_t = x) := [T_t^{t+\Delta} \mathbb{I}_y](x)$$

and for all x in $\mathcal{X}$ as

$$\tilde{P}(X_0 = x \mid \Omega) := \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\tilde{P}$ is coherent, and any extension of $\tilde{P}$ to $\mathcal{D}_{\mathrm{CP}}$ is a counting process that is consistent with Λ.

Proof Fix some $u = t_0, \ldots, t_n$ in $\mathcal{U}_{\emptyset}$ with $t_0 = 0$ and, for all i in $\{0, \ldots, n\}$, some sequence $S_i := \{\lambda_{i,x}\}_{x \in \mathcal{X}}$ in $[\underline{\lambda}, \overline{\lambda}]$. Due to Corollary 66,

$$\mathcal{T} := \mathcal{T}_{S_0}^{[0, t_1]} \otimes \mathcal{T}_{S_1}^{[t_1, t_2]} \otimes \cdots \otimes \mathcal{T}_{S_{n-1}}^{[t_{n-1}, t_n]} \otimes \mathcal{T}_{S_n}^{[t_n, +\infty)}$$

is a counting transformation system. Therefore, it follows from Lemma 79 that $\tilde{P}$ is coherent, and that any coherent extension $\tilde{P}^{\star}$ of $\tilde{P}$ to $\mathcal{D}_{\mathrm{CP}}$ is a counting process. Hence, all that remains for us is to prove that any such coherent extension $\tilde{P}^{\star}$ is consistent with $[\underline{\lambda}, \overline{\lambda}]$. We will only verify Equation (10), Equation (11) can be verified in a similar fashion. To that end, we fix any t in $\mathbb{R}_{\geq 0}$, u in $\mathcal{U}_{<t}$, Δ in $\mathbb{R}_{>0}$ and (x_u, x) in $\mathcal{X}_{u \cup t}$. Observe that

$$\frac{\tilde{P}(X_{t+\Delta} = x+1 \mid X_u = x_u, X_t = x)}{\Delta} = \frac{[T_t^{t+\Delta} \mathbb{I}_{x+1}](x)}{\Delta}.$$

We let i be the greatest element of $\{0, \ldots, n\}$ such that $t_i \leq t$. We now claim that

$$\lim_{\Delta \to 0^+} \frac{[T_t^{t+\Delta} \mathbb{I}_{x+1}](x)}{\Delta} = \lambda_{i,x}.$$

If this were true, then

$$\lim_{\Delta \to 0^+} \frac{\tilde{P}^{\star}(X_{t+\Delta} = x+1 \mid X_u = x_u, X_t = x)}{\Delta} = \lambda_{i,x}$$

which implies Equation (10) because $\underline{\lambda} \leq \lambda_{i,x} \leq \overline{\lambda}$.

We now set out to verify our claim. By construction, $T_t^{t+\Delta} = T_{t,S_i}^{t+\Delta}$ for any Δ in $\mathbb{R}_{\geq 0}$ with $t_i + \Delta < t_{i+1}$ if $i < n$, such that

$$\left| \frac{[T_t^{t+\Delta} \mathbb{I}_{x+1}](x)}{\Delta} - \lambda_{i,x} \right| = \left| \frac{[T_{t,S_i}^{t+\Delta} \mathbb{I}_{x+1}](x)}{\Delta} - \lambda_{i,x} \right|.$$

We use the two obvious equalities $[I \mathbb{I}_{x+1}](x)$ and $[Q_{S_i} \mathbb{I}_{x+1}](x) = \lambda_{i,x}$ and (NH4), to yield

$$\left| \frac{[T_t^{t+\Delta} \mathbb{I}_{x+1}](x)}{\Delta} - \lambda_{i,x} \right| = \left| \frac{[T_{t,S_i}^{t+\Delta} \mathbb{I}_{x+1}](x) - [I \mathbb{I}_{x+1}](x)}{\Delta} - [Q_{S_i} \mathbb{I}_{x+1}](x) \right|$$

$$\leq \left\| \frac{T_{t,S_i}^{t+\Delta} \mathbb{I}_{x+1} - I \mathbb{I}_{x+1}}{\Delta} - [Q_{S_i} \mathbb{I}_{x+1}] \right\| \leq \left\| \frac{T_{t,S_i}^{t+\Delta} - I}{\Delta} - Q_{S_i} \right\| \|\mathbb{I}_{x+1}\| = \left\| \frac{T_{t,S_i}^{t+\Delta} - I}{\Delta} - Q_{S_i} \right\|.$$

Fix any ε in $\mathbb{R}_{>0}$. Because, by definition, $\mathcal{T}_{S_i}$ is the family of (linear) lower counting transformations induced by the generalised Poisson generator $Q_{S_i} = \underline{Q}_{S_i'}$ characterised by the sequence $S_i' = \{(\lambda_{i,x}, \lambda_{i,x})\}_{x \in \mathcal{X}}$, it follows from Lemma 51 that there is a δ in $\mathbb{R}_{>0}$ such that for any Δ in $\mathbb{R}_{>0}$ with $\Delta < \delta$,

$$\left\| \frac{T_{t,S_i}^{t+\Delta} - I}{\Delta} - Q_{S_i} \right\| = \left\| \frac{T_{t,S_i}^{t+\Delta} - T_{t,S_i}^t}{\Delta} - Q_{S_i} T_{t,S_i}^t \right\| \leq \varepsilon, \tag{104}$$

where the equality follows from Proposition 48 (i). For any Δ in $\mathbb{R}_{>0}$ such that $\Delta < \delta$ and $t_i + \Delta < t_{i+i}$ if $i < n$, it follows from the previous two inequalities that

$$\left| \frac{[T_t^{t+\Delta} \mathbb{I}_{x+1}](x)}{\Delta} - \lambda_{i,x} \right| \leq \varepsilon.$$

Since ε was any positive real number, we infer from this inequality that

$$\lim_{\Delta \to 0^+} \frac{[T_t^{t+\Delta} \mathbb{I}_{x+1}](x)}{\Delta} = \lambda_{i,x},$$

as required. ∎

Proposition 100 *Consider any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, any u in $\mathcal{U}_{<t}$, any (x_u, x) in $\mathcal{X}_{u \cup t}$ and any f in $\mathcal{L}^c(\mathcal{X})$. Then for any ε in $\mathbb{R}_{>0}$, there is a counting process P that is consistent with the rate interval Λ such that*

$$|\underline{P}_t^s(f \mid x) - E_P(f(X_s) \mid X_u = x_u, X_t = x)| \leq \varepsilon.$$

Proof By Lemma 83, there is a sequence $v = t_0, \ldots, t_n$ in $\mathcal{U}_{[t,s]}$ and, for all i in $\{1, \ldots, n\}$, a sequence $S_i = \{\lambda_x^i\}_{x \in \mathcal{X}}$ in $[\underline{\lambda}, \overline{\lambda}]$ such that

$$\left| [\underline{T}_t^s f](x) - \left[\prod_{i=1}^n T_{t_{i-1}, S_i}^{t_i} f \right](x) \right| \leq \left\| \underline{T}_t^s f - \prod_{i=1}^n T_{t_{i-1}, S_i}^{t_i} f \right\| \leq \varepsilon. \tag{105}$$

Furthermore, we fix any arbitrary sequence $S = \{\lambda_x\}_{x \in \mathcal{X}}$ in $[\underline{\lambda}, \overline{\lambda}]$. It now follows from Corollary 66 that

$$\mathcal{T} = \{T_r^q : r, q \in \mathbb{R}_{\geq 0}, r \leq q\} := \mathcal{T}_S^{[0, t_1]} \otimes \mathcal{T}_{S_1}^{[t_1, t_2]} \otimes \mathcal{T}_{S_3}^{[t_2, t_3]} \otimes \cdots \otimes \mathcal{T}_{S_n}^{[t_{n-1}, t_n]} \otimes \mathcal{T}_S^{[t_n, +\infty)}$$

is a counting transformation system. Furthermore, it follows from Lemma 99 that there is a counting process P that is consistent with $[\underline{\lambda}, \overline{\lambda}]$ and that satisfies

$$P(X_s = y \mid X_u = x_u, X_t = x) = [T_t^s \mathbb{I}_y](x) = \left[\prod_{i=1}^n T_{t_{i-1}, S_i}^{t_i} \mathbb{I}_y \right](x) \qquad \text{for all } y \in \mathcal{X}. \tag{106}$$

Observe furthermore that for any y in $\mathcal{X}$ with $y \geq 1$,

$$P(X_s \geq y \mid X_u = x_u, X_t = x) = 1 - P(X_s \leq y - 1 \mid X_u = x_u, X_t = x) = 1 - \sum_{z=0}^{y-1} P(X_s = z \mid X_u = x_u, X_t = x)$$

$$= 1 - \sum_{z=0}^{y-1} [T_t^s \mathbb{I}_z](x) = \left[T_t^s \left(1 - \sum_{z=0}^{y-1} \mathbb{I}_z \right) \right](x) = [T_t^s \mathbb{I}_{\geq y}](x), \tag{107}$$

where for the third equality we have used Equation (106) and for the fourth equality we have used the linearity of the linear counting transformation T_t^s.

Fix some $\bar{x}$ in $\mathcal{X}$ such that f is constant starting from $\bar{x}$. By Lemma 96 and Equation (4),

$$E_P(f(X_s) \mid X_u = x_u, X_t = x) = \sum_{y=0}^{\bar{x}-1} f(y) P(X_s = y \mid X_u = x_u, X_t = x) + f(\bar{x}) P(X_s \geq \bar{x} \mid X_u = x_u, X_t = x).$$

We now substitute Equations (106) and (107) and again use the linearity of the linear counting transformation T_t^s, to yield

$$E_P(f(X_s) \mid X_u = x_u, X_t = x) = \sum_{y=0}^{\bar{x}-1} f(y) [T_t^s \mathbb{I}_y](x) + f(\bar{x}) [T_t^s \mathbb{I}_{\geq \bar{x}}](x)$$

$$= \left[T_t^s \left(\sum_{y=0}^{\bar{x}-1} f(y) \mathbb{I}_y + f(\bar{x}) \mathbb{I} \geq \bar{x} \right) \right](x) = [T_t^s f](x) = \left[\prod_{i=1}^n T_{t_{i-1}, S_i}^{t_i} f \right](x),$$

where the final equality holds due to the construction of $\mathcal{T}$. The stated now follows if we substitute this equality in Equation (105). ∎

Everything is now set up to prove our main result regarding the expectation of eventually constant functions.

Theorem 101 *For any t,s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, u in $\mathscr{U}_{<t}$, f in $\mathscr{L}^c(\mathscr{X})$ and (x_u,x) in $\mathscr{X}_{u\cup t}$,*

$$\underline{E}_\Lambda(f(X_s) \,|\, X_u = x_u, X_t = x) = \underline{P}_t^s(f \,|\, x).$$

Proof On the one hand, it follows from Proposition 98 and Equation (8) that

$$\underline{P}_t^s(f(X_s) \,|\, x) \leq \underline{E}_\Lambda(f(X_s) \,|\, X_u = x_u, X_t = x) = \inf\{E_P(f(X_s) \,|\, X_u = x_u, X_t = x) : P \in \mathbb{P}_\Lambda\}.$$

On the other hand, it follows from Proposition 100 that

$$(\forall \varepsilon \in \mathbb{R}_{>0})(\exists P^\star \in \mathbb{P}_\Lambda) \, |\underline{P}_t^s(f \,|\, x) - E_{P^\star}(f(X_s) \,|\, X_u = x_u, X_t = x)| \leq \varepsilon.$$

From these two inequalities, it follows that for any ε in $\mathbb{R}_{>0}$,

$$\underline{E}_\Lambda(f(X_s) \,|\, X_u = x_u, X_t = x) \geq \underline{P}_t^s(f(X_s) \,|\, x) \geq E_{P^\star}(f(X_s) \,|\, X_u = x_u, X_t = x) - \varepsilon \geq \underline{E}_\Lambda(f(X_s) \,|\, X_u = x_u, X_t = x) - \varepsilon.$$

The equality of the statement follows from these inequalities because ε is an arbitrary positive real number. ∎

H.2.3. BOUNDED FUNCTIONS

Next, we move from eventually constant functions to general bounded functions. Essential to our proof of Theorem 15 are the following two observations.

Lemma 102 *For any s,t in $\mathbb{R}_{\geq 0}$ with $t \leq s$,*

$$\lim_{\bar{x} \to +\infty} \overline{P}_t^s(\mathbb{I}_{\geq \bar{x}} \,|\, x) = 0,$$

where $\mathbb{I}_{\geq \bar{x}}$ is the indicator of the set $\{z \in \mathscr{X} : z \geq \bar{x}\}$.

Proof Fix any ε in $\mathbb{R}_{>0}$. To prove the stated, we need to verify that

$$(\exists x^\star \in \mathscr{X})(\forall \bar{x} \in \mathscr{X}, \bar{x} \geq x^\star) \, 0 \leq \overline{P}_t^s(\mathbb{I}_{\geq \bar{x}} \,|\, x) = -\underline{P}_t^s(-\mathbb{I}_{\geq \bar{x}} \,|\, x) = -[\underline{T}_t^s(-\mathbb{I}_{\geq \bar{x}})](x) \leq \varepsilon.$$

To that end, we recall from Theorem 45 that there is a sequence u in $\mathscr{U}_{[t,s]}$ such that $\sigma(u)\|\underline{Q}\| \leq 2$ and

$$\|\underline{T}_t^s - \Phi_u\| \leq \varepsilon.$$

Let n be the number of time points in u. Furthermore, we let $x^\star := x + n + 1$ and fix any $\bar{x}$ in $\mathscr{X}$ such that $\bar{x} \geq x^\star$. It then follows from Lemma 47 that

$$[\Phi_u(-\mathbb{I}_{\geq \bar{x}})](x) = [\Phi_u(-\mathbb{I}_{\leq x+n}\mathbb{I}_{\geq \bar{x}} - \mathbb{I}_{\geq \bar{x}}(x+n)\mathbb{I}_{>x+n})](x) = [\Phi_u 0](x) = 0,$$

where for the final equality we have used (LT6), which holds because Φ_u is a lower counting transformation due to Corollary 36. We combine these two observations, to yield

$$0 \leq \overline{P}_t^s(\mathbb{I}_{\geq \bar{x}} \,|\, x) = -[\underline{T}_t^s(-\mathbb{I}_{\geq \bar{x}})](x) + [\Phi_u(-\mathbb{I}_{\geq \bar{x}})](x) \leq |[\underline{T}_t^s(-\mathbb{I}_{\geq \bar{x}})](x) - [\Phi_u(-\mathbb{I}_{\geq \bar{x}})](x)|$$
$$\leq \|\underline{T}_t^s(-\mathbb{I}_{\geq \bar{x}}) - \Phi_u(-\mathbb{I}_{\geq \bar{x}})\| \leq \|\underline{T}_t^s - \Phi_u\|\|-\mathbb{I}_{\geq \bar{x}}\| = \|\underline{T}_t^s - \Phi_u\| \leq \varepsilon,$$

where for the first equality we have used that $\overline{P}_t^s(\cdot \,|\, x)$ is a coherent upper prevision, for the third inequality we have used (NH4) and for the final equality we have used that $\|-\mathbb{I}_{>\bar{x}}\| = 1$. ∎

Lemma 103 *Let P be any counting process that is consistent with the rate interval Λ. Then for any t,s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, u in $\mathscr{U}_{<t}$, (x_u,x) in $\mathscr{X}_{u\cup t}$, f in $\mathscr{L}(\mathscr{X})$ and ε in $\mathbb{R}_{>0}$,*

$$(\exists x^\star \in \mathscr{X})(\forall \bar{x} \in \mathscr{X}, \bar{x} \geq x^\star)(\forall P \in \mathbb{P}_\Lambda) \, |E_P(f(X_s) \,|\, X_u = x_u, X_t = x) - E_P([\mathbb{I}_{\leq \bar{x}}f + f(\bar{x})\mathbb{I}_{>\bar{x}}](X_s) \,|\, X_u = x_u, X_t = x)| \leq \varepsilon.$$

Proof Fix any $\bar{x}$ in $\mathscr{X}$, and let $f_{\bar{x}} := f\mathbb{I}_{\leq\bar{x}} + f(\bar{x})\mathbb{I}_{>\bar{x}}$. Observe that

$$f_{\bar{x}} - 2\|f\|\mathbb{I}_{>\bar{x}} \leq f \leq f_{\bar{x}} + 2\|f\|\mathbb{I}_{>\bar{x}}.$$

Let P be any counting process that is consistent with Λ. Due to the previous inequalities and the monotonicity of E_P,

$$E_P([f_{\bar{x}} - 2\|f\|\mathbb{I}_{>\bar{x}}](X_s) \,|\, X_u = x_u, X_t = x) \leq E_P(f(X_s) \,|\, X_u = x_u, X_t = x)$$
$$\leq E_P([f_{\bar{x}} + \|f\|\mathbb{I}_{>\bar{x}}](X_s) \,|\, X_u = x_u, X_t = x). \quad (108)$$

Because $f_{\bar{x}} - 2\|f\|\mathbb{I}_{>\bar{x}}$ and $f_{\bar{x}}$ are both constant starting from $\bar{x}+1$, it follows from Lemma 96 and Equation (4) that

$$E_P([f_{\bar{x}} - 2\|f\|\mathbb{I}_{>\bar{x}}](X_s) \,|\, X_u = x_u, X_t = x) = E_P(f_{\bar{x}}(X_s) \,|\, X_u = x_u, X_t = x) - 2\|f\|E_P(\mathbb{I}_{>\bar{x}}(X_s) \,|\, X_u = x_u, X_t = x). \quad (109)$$

Similarly,

$$E_P([f_{\bar{x}} + 2\|f\|\mathbb{I}_{>\bar{x}}](X_s) \,|\, X_u = x_u, X_t = x) = E_P(f_{\bar{x}}(X_s) \,|\, X_u = x_u, X_t = x) + 2\|f\|E_P(\mathbb{I}_{>\bar{x}}(X_s) \,|\, X_u = x_u, X_t = x). \quad (110)$$

We now combine Equations (108)–(110), to yield

$$|E_P(f(X_s) \,|\, X_u = x_u, X_t = x) - E_P(f_{\bar{x}}(X_s) \,|\, X_u = x_u, X_t = x)| \leq 2\|f\|E_P(\mathbb{I}_{>\bar{x}}(X_s) \,|\, X_u = x_u, X_t = x). \quad (111)$$

It now follows from Proposition 98, the conjugacy of $\underline{P}_t^s(\cdot\,|\,x)$ and $\overline{P}_t^s(\cdot\,|\,x)$ and the obvious equality $E_P(\mathbb{I}_{>\bar{x}}(X_s) \,|\, X_u = x_u, X_t = x) = -E_P(-\mathbb{I}_{>\bar{x}}(X_s) \,|\, X_u = x_u, X_t = x)$ that

$$E_P(\mathbb{I}_{>\bar{x}}(X_s) \,|\, X_u = x_u, X_t = x) \leq \overline{P}_t^s(\mathbb{I}_{>\bar{x}} \,|\, x). \quad (112)$$

Fix any ε in $\mathbb{R}_{>0}$, and choose any ε' in $\mathbb{R}_{>0}$ such that $2\|f\|\varepsilon' \leq \varepsilon$. It now follows from and Lemma 102 that there is an $x^\star$ in $\mathscr{X}$ such that if $\bar{x} \geq x^\star$, then

$$\overline{P}(\mathbb{I}_{>\bar{x}} \,|\, x) \leq \varepsilon'. \quad (113)$$

Finally, we now combine Equations (111)–(113) and recall that $f_{\bar{x}} = f\mathbb{I}_{\leq\bar{x}} + f(\bar{x})\mathbb{I}_{>\bar{x}}$, to yield

$$(\forall \bar{x} \in \mathscr{X}, \bar{x} \geq x^\star)(\forall P \in \mathbb{P}_\Lambda) \; |E_P(f(X_s) \,|\, X_u = x, X_t = x) - E_P([f\mathbb{I}_{\leq\bar{x}} + f(\bar{x})\mathbb{I}_{>\bar{x}}](X_s) \,|\, X_u = x, X_t = x)| \leq 2\|f\|\varepsilon' \leq \varepsilon.$$

∎

Proof of Theorem 15 Our proof is similar to the proof of Theorem 101 In the first part, we will show that

$$\underline{P}_t^s(f \,|\, x) \leq \underline{E}_\Lambda(f(X_s) \,|\, X_u = x_u, X_t = x) = \inf\{E_P(f(X_s) \,|\, X_u = x_u, X_t = x) : P \in \mathbb{P}_\Lambda\}. \quad (114)$$

In the second part, we will subsequently show that

$$(\forall \varepsilon \in \mathbb{R}_{>0})(\exists P^\star \in \mathbb{P}_\Lambda) \; |\underline{P}_t^s(f \,|\, x) - E_{P^\star}(f(X_s) \,|\, X_u = x_u, X_t = x)| \leq \varepsilon. \quad (115)$$

The stated follows from Equations (114) and (115). Indeed, from these equations it follows that, for all ε in $\mathbb{R}_{>0}$,

$$\underline{E}_\Lambda(f(X_s) \,|\, X_u = x_u, X_t = x) \geq \underline{P}_t^s(f \,|\, x) \geq E_{P^\star}(f(X_s) \,|\, X_u = x_u, X_t = x) - \varepsilon \geq \underline{E}_\Lambda(f(X_s) \,|\, X_u = x_u, X_t = x) - \varepsilon.$$

The equality of the statement now follows from these inequalities because ε is an arbitrary positive real number.

We now set out to prove Equations (114) and (115). To that end, we fix any ε in $\mathbb{R}_{>0}$, and choose some $\varepsilon_1, \varepsilon_2, \varepsilon_3$ in $\mathbb{R}_{>0}$ such that $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 \leq \varepsilon$. Recall from Proposition 13 that there is an $x_1^\star$ such that

$$(\forall \bar{x} \in \mathscr{X}, \bar{x} \geq x_1^\star) \; |\underline{P}_t^s(f \,|\, x) - \underline{P}_t^s(\mathbb{I}_{\leq\bar{x}}f + f(\bar{x})\mathbb{I}_{>\bar{x}} \,|\, x)| \leq \varepsilon_1. \quad (116)$$

Due to Lemma 103, there is an $x_2^\star$ such that

$$(\forall \bar{x} \in \mathscr{X}, \bar{x} \geq x_2^\star)(\forall P \in \mathbb{P}_\Lambda) \; |E_P(f(X_s) \,|\, X_u = x_u, X_t = x) - E_P([\mathbb{I}_{\leq\bar{x}}f + f(\bar{x})\mathbb{I}_{>\bar{x}}](X_s) \,|\, X_u = x_u, X_t = x)| \leq \varepsilon_2. \quad (117)$$

Let $x^\star := \max\{x_1^\star, x_2^\star\}$, and fix any $\bar{x}$ in $\mathscr{X}$ such that $\bar{x} \geq x^\star$. It now follows from Equations (116) and (117) that, for any P in $\mathbb{P}_\Lambda$,

$$\underline{P}_t^s(f \,|\, x) - \varepsilon \le \underline{P}_t^s(f \,|\, x) - \varepsilon_1 - \varepsilon_2 \le \underline{P}_t^s(\mathbb{I}_{\le \overline{x}} f + f(\overline{x})\mathbb{I}_{> \overline{x}} \,|\, x) - \varepsilon_1$$
$$\le E_P([\mathbb{I}_{\le \overline{x}} f + f(\overline{x})\mathbb{I}_{> \overline{x}}](X_s) \,|\, X_u = x_u, X_t = x) - \varepsilon_1 \le E_P(f(X_s) \,|\, X_u = x_u, X_t = x),$$

where the third inequality follows from Proposition 98. Since ε was an arbitrary positive real number, we infer from this inequality that

$$(\forall P \in \mathbb{P}_\Lambda) \, \underline{P}_t^s(f \,|\, x) \le E_P(f(X_s) \,|\, X_u = x_u, X_t = x).$$

We combine this with Equation (8) and use that non-strict inequalities are preserved when taking infima, to yield Equation (114).

Next, we prove Equation (115). Due to Proposition 100, there is a $P^\star$ in $\mathbb{P}_\Lambda$ such that

$$|\underline{P}_t^s(\mathbb{I}_{\le \overline{x}} f + f(\overline{x})\mathbb{I}_{> \overline{x}} f \,|\, x) - E_{P^\star}([\mathbb{I}_{\le \overline{x}} f + f(\overline{x})\mathbb{I}_{> \overline{x}}](X_s) \,|\, X_u = x_u, X_t = x)| \le \varepsilon_3.$$

We now use this inequality and Equations (116) and (117), to yield

$$\begin{aligned}
|\underline{P}_t^s&(f \,|\, x) - E_{P^\star}(f(X_s) \,|\, X_u = x_u, X_t = x)| \\
&\le |\underline{P}_t^s(f \,|\, x) - \underline{P}_t^s(\mathbb{I}_{\le x^\star} f + f(x^\star)\mathbb{I}_{> x^\star} \,|\, x)| \\
&\quad + |\underline{P}_t^s(\mathbb{I}_{\le x^\star} f + f(x^\star)\mathbb{I}_{> x^\star} \,|\, x) - E_{P^\star}([\mathbb{I}_{\le x^\star} f + f(x^\star)\mathbb{I}_{> x^\star}](X_s) \,|\, X_u = x_u, X_t = x)| \\
&\quad\quad + |E_{P^\star}([\mathbb{I}_{\le x^\star} f + f(x^\star)\mathbb{I}_{> x^\star}](X_s) \,|\, X_u = x_u, X_t = x) - E_{P^\star}(f(X_s) \,|\, X_u = x_u, X_t = x)| \\
&\le \varepsilon_1 + \varepsilon_2 + \varepsilon_3 \le \varepsilon,
\end{aligned}$$

as required by Equation (115). ∎

Next, we consider the special case of monotone bounded functions.

Lemma 104 *Fix any f in $\mathscr{L}(\mathscr{X})$, n in $\mathbb{N}$ and, for all i in $\{1, \dots, n\}$, some Δ_i in $\mathbb{R}_{\ge 0}$ with $\Delta_i \|\underline{Q}\| \le 2$. If f is non-decreasing, then*

$$\prod_{i=1}^n (I + \Delta_i \underline{Q}) f = \prod_{i=1}^n (I + \Delta_i Q_{\underline{\lambda}}) f$$

is non-decreasing. Similarly, if f is non-increasing, then

$$\prod_{i=1}^n (I + \Delta_i \underline{Q}) f = \prod_{i=1}^n (I + \Delta_i Q_{\overline{\lambda}}) f$$

is non-increasing.

Proof We only prove the stated for a non-decreasing f, the proof for non-increasing f is entirely similar. Our proof is one by induction. In case $n = 1$, the equality of the stated follows trivially from the definition of $\underline{Q}$: for any x in $\mathscr{X}$,

$$\begin{aligned}
[(I + \Delta_1 \underline{Q}) f](x) &= f(x) + \Delta_1 [\underline{Q} f](x) = f(x) + \Delta_1 \min\{\lambda f(x+1) - \lambda f(x) : \lambda \in [\underline{\lambda}, \overline{\lambda}]\} \\
&= f(x) + \Delta_1 (\underline{\lambda} f(x+1) - \underline{\lambda} f(x)) = f(x) + \Delta_1 [Q_{\underline{\lambda}} f](x) \\
&= [(I + \Delta_1 Q_{\underline{\lambda}}) f](x).
\end{aligned}$$

That $(I + \Delta_1 \underline{Q}) f$ is non-decreasing as well is easily verified in a similar fashion:

$$\begin{aligned}
[(I + \Delta_1 \underline{Q}) f](x+1) - [(I + \Delta_1 \underline{Q}) f](x) &= [(I + \Delta_1 Q_{\underline{\lambda}}) f](x+1) - [(I + \Delta_1 Q_{\underline{\lambda}}) f](x) \\
&= f(x+1) + \Delta_1 \underline{\lambda}(f(x+2) - f(x+1)) - f(x) - \Delta_1 \underline{\lambda}(f(x+1) - f(x)) \\
&= (1 - \Delta_1 \underline{\lambda})(f(x+1) - f(x)) + \Delta_1 \underline{\lambda}(f(x+2) - f(x+1)) \\
&\ge 0,
\end{aligned}$$

where the inequality follows from the inequality $1 - \Delta_1 \underline{\lambda} \ge 0$, which holds because $\underline{\lambda} \le \overline{\lambda}$ and $\Delta_1 2\overline{\lambda} = \Delta_1 \|\underline{Q}\| \le 2$, where the first inequality follows from Equation (20) and Lemma 34.

For the induction step, we fix any m in $\mathbb{N}$ with $m \ge 2$ and assume that the stated then holds for all n in $\mathbb{N}$ with $n < m$. We now show that the stated then follows for m as well. Let

$$g' := \prod_{i=2}^m (I + \Delta_i \underline{Q}) f \qquad \text{and} \qquad g := \prod_{i=1}^m (I + \Delta_i \underline{Q}) f = (I + \Delta_1 \underline{Q}) g'.$$

By the induction hypothesis, g' is non-decreasing and equal to

$$\prod_{i=2}^{m}(I+\Delta_i Q_{\underline{\lambda}})f$$

HIt follows from this and the induction hypothesis for $n=1$ that

$$g=(I+\Delta_1\underline{Q})g'=(I+\Delta_1 Q_{\underline{\lambda}})g'=\prod_{i=1}^{m}(I+\Delta_i Q_{\underline{\lambda}})f$$

is non-decreasing, as required. ∎

Proof of Proposition 16 We only prove the stated for a non-decreasing function f, the proof for a non-increasing f is entirely similar. We first set out to prove that

$$[\underline{T}_t^s f](x)=[T_{t,\underline{\lambda}}^s f](x). \tag{118}$$

To that end, we fix any ε in $\mathbb{R}_{>0}$, and choose any ε' in $\mathbb{R}_{>0}$ such that $2\varepsilon'\|f\|\le\varepsilon$. By Theorem 45 and Corollary 62, there is a u in $\mathscr{U}_{[t,s]}$ with $\sigma(u)\|\underline{Q}\|\le 2$—and, due to Lemma 34 and Corollary 58, therefore also $\sigma(u)\|Q_{\underline{\lambda}}\|\le 2$—such that

$$\left\|\underline{T}_t^s-\prod_{i=1}^{n}(I+\Delta_i\underline{Q})\right\|\le\varepsilon' \qquad\text{and}\qquad \left\|T_{t,\underline{\lambda}}^s-\prod_{i=1}^{n}(I+\Delta_i Q_{\underline{\lambda}})\right\|\le\varepsilon'. \tag{119}$$

Observe now that

$$\left|[\underline{T}_t^s f](x)-[T_{t,\underline{\lambda}}^s f](x)\right|\le\|\underline{T}_t^s f-T_{t,\underline{\lambda}}^s f\|=\left\|\underline{T}_t^s f-\prod_{i=1}^{n}(I+\Delta_i\underline{Q})f+\prod_{i=1}^{n}(I+\Delta_i\underline{Q})f-T_{t,\underline{\lambda}}^s f\right\|$$

$$=\left\|\underline{T}_t^s f-\prod_{i=1}^{n}(I+\Delta_i\underline{Q})f+\prod_{i=1}^{n}(I+\Delta_i Q_{\underline{\lambda}})f-T_{t,\underline{\lambda}}^s f\right\|$$

$$\le\left\|\underline{T}_t^s f-\prod_{i=1}^{n}(I+\Delta_i\underline{Q})f\right\|+\left\|T_{t,\underline{\lambda}}^s f-\prod_{i=1}^{n}(I+\Delta_i Q_{\underline{\lambda}})f\right\|$$

$$\le\left\|\underline{T}_t^s-\prod_{i=1}^{n}(I+\Delta_i\underline{Q})\right\|\|f\|+\left\|T_{t,\underline{\lambda}}^s-\prod_{i=1}^{n}(I+\Delta_i Q_{\underline{\lambda}})\right\|\|f\|$$

$$\le 2\varepsilon'\|f\|\le\varepsilon,$$

where the second equality follows from Lemma 104, the second inequality follows from (NH4) and the penultimate inequality follows from Equation (119). Since ε was an arbitrary positive real number, these inequalities imply Equation (118).

The stated basically follows from Equation (118). To see this, we let P_λ be a Poisson process with rate λ in the rate interval Λ. Then clearly

$$\underline{E}_\Lambda(f(X_s)\,|\,X_u=x_u,X_t=x)\le\underline{E}_\Lambda^\star(f(X_s)\,|\,X_u=x_u,X_t=x)\le E_{P_\lambda}(f(X_s)\,|\,X_u=x_u,X_t=x). \tag{120}$$

Recall from Theorem 15 that

$$\underline{E}_\Lambda(f(X_s)\,|\,X_u=x_u,X_t=x)=[\underline{T}_t^s f](x). \tag{121}$$

Similarly, it follows from Equation (13) and Theorem 15 that

$$E_{P_{\underline{\lambda}}}(f(X_s)\,|\,X_u=x_u,X_t=x)=\underline{E}_{[\underline{\lambda},\underline{\lambda}]}(f(X_s)\,|\,X_u=x_u,X_t=x)=[T_{t,\underline{\lambda}}^s f](x). \tag{122}$$

It now follows from Equation (118), Equation (120) with $\lambda=\underline{\lambda}$ and Equations (121) and (122) that

$$E_{P_{\underline{\lambda}}}(f(X_s)\,|\,X_u=x_u,X_t=x)=[T_{t,\underline{\lambda}}^s f](x)=[\underline{T}_t^s f](x)=\underline{E}_\Lambda(f(X_s)\,|\,X_u=x_u,X_t=x)\le E_{P_{\underline{\lambda}}}(f(X_s)\,|\,X_u=x_u,X_t=x),$$

which clearly implies the equality of the statement for non-decreasing functions f. ∎

66

H.2.4. NON-BOUNDED FUNCTIONS

Finally, we are ready to make the transition towards bounded-below functions.

Proof of Proposition 17 Observe that if f is non-decreasing and bounded, then the stated follows immediately from Proposition 16. We therefore only have to prove the stated for a non-decreasing f that is not bounded but bounded below. Observe that in this case, $\inf f = f(0)$.

We now first set out to prove that

$$(\forall P \in \mathbb{P}_\Lambda) E_{P_{\underline{\lambda}}}(f(X_S) \mid X_u = x, X_t = x) \leq E_P(f(X_s) \mid X_u = x_u, X_t = x). \tag{123}$$

To that end, we fix any P in $\mathbb{P}_\Lambda$. Recall that

$$E_P(f(X_s) \mid X_u = X_u, X_t = x) = \int_{\inf f}^{\sup f} P(\{f(X_s) > \alpha\} \mid X_u = x_u, X_t = x) \, d\alpha = \int_{f(0)}^{+\infty} P(\{f(X_s) > \alpha\} \mid X_u = x_u, X_t = x) \, d\alpha.$$

We now fix any β in $\mathbb{R}$ with $\beta \geq f(x)$, and let y_β be an element of $\mathscr{X}$ such that $f(y_\beta) \leq \beta \leq f(y_\beta + 1)$—this is possible because f is non-decreasing and unbounded. Observe that

$$\int_{f(0)}^{\beta} P(\{f(X_s) > \alpha\} \mid X_u = x_u, X_t = x) \, d\alpha$$

$$= \int_{f(0)}^{f(1)} P(\{f(X_s) > \alpha\} \mid X_u = x_u, X_t = x) \, d\alpha + \int_{f(1)}^{f(2)} P(\{f(X_s) > \alpha\} \mid X_u = x_u, X_t = x) \, d\alpha$$

$$+ \cdots + \int_{f(y_\beta)}^{\beta} P(\{f(X_s) > \alpha\} \mid X_u = x_u, X_t = x) \, d\alpha.$$

As a consequence of our assumptions on f, it follows that

$$(\forall y \in \mathscr{X})(\forall \alpha \in [f(y), f(y+1))) \; \{f(X_s) > \alpha\} = (X_s > y). \tag{124}$$

Hence,

$$\int_{f(0)}^{\beta} P(\{f(X_s) > \alpha\} \mid X_u = x_u, X_t = x) \, d\alpha = \sum_{y=0}^{y_\beta - 1} (f(y+1) - f(y)) P(X_s > y \mid X_u = x_u, X_t = x)$$

$$+ (\beta - f(y_\beta)) P(X_s > y_\beta \mid X_u = x_u, X_t = x)$$

For any y in $\mathscr{X}$, it follows from Equation (4)—with the $\mathscr{F}_{u \cup t}$-simple function $\mathbb{I}_{>y}(X_s)$—and Proposition 16—with the non-decreasing function $\mathbb{I}_{>y}$—that

$$P(X_s > y \mid X_u = x_u, X_t = x) = E_P(\mathbb{I}_{>y}(X_s) \mid X_u = x_u, X_t = x)$$

$$\geq E_{P_{\underline{\lambda}}}(\mathbb{I}_{>y}(X_s) \mid X_u = x_u, X_t = x) = P_{\underline{\lambda}}(X_s > y \mid X_u = x_u, X_t = x),$$

where $P_{\underline{\lambda}}$ is the Poisson process with rate $\underline{\lambda}$. We combine this inequality with the previous equality, to yield

$$\int_{f(0)}^{\beta} P(\{f(X_s) > \alpha\} \mid X_u = x_u, X_t = x) \, d\alpha \geq \sum_{y=0}^{y_\beta - 1} (f(y+1) - f(y)) P_{\underline{\lambda}}(X_s > y \mid X_u = x_u, X_t = x)$$

$$+ (\beta - f(y_\beta)) P_{\underline{\lambda}}(X_s > y_\beta \mid X_u = x_u, X_t = x)$$

$$= \int_{f(0)}^{\beta} P_{\underline{\lambda}}(\{f(X_s) > \alpha\} \mid X_u = x_u, X_t = x) \, d\alpha.$$

We take the limit for β going to $+\infty$ on both sides of the inequality, to yield Equation (123):

$$E_P(f(X_s) \mid X_u = x_u, X_t = x) = \lim_{\beta \to +\infty} \int_{f(0)}^{\beta} P(\{f(X_s) > \alpha\} \mid X_u = x_u, X_t = x) \, d\alpha$$

$$\geq \lim_{\beta \to +\infty} \int_{f(0)}^{\beta} P_{\underline{\lambda}}(\{f(X_s) > \alpha\} \,|\, X_u = x_u, X_t = x) \, d\alpha = E_{P_{\underline{\lambda}}}(f(X_s) \,|\, X_u = x_u, X_t = x).$$

Finally, we verify that the stated. On the one hand, we recall from Equation (14) that

$$\underline{E}_{\Lambda}(f(X_s) \,|\, X_u = x_u, X_t = x) \leq \underline{E}_{\Lambda}^{\star}(f(X_s) \,|\, X_u = x_u, X_t = x) \leq E_{P_{\underline{\lambda}}}(f(X_s) \,|\, X_u = x_u, X_t = x),$$

where the final equality holds because $P_{\underline{\lambda}}$ belongs to $\mathbb{P}_{\Lambda}^{\star}$. On the other hand, because non-strict inequalities are preserved when taking the infimum, it follows from Equations (8) and (123) that

$$\underline{E}_{\Lambda}(f(X_s) \,|\, X_u = x_u, X_t = x) \geq E_{P_{\underline{\lambda}}}(f(X_s) \,|\, X_u = x_u, X_t = x).$$

The stated equality now follows immediately from these two observations ∎

Proof of Corollary 18 As $f(X_s) = X_s$ is non-decreasing, it follows from Proposition 17 that

$$\underline{E}_{\Lambda}(X_s \,|\, X_u = x_u, X_t = x) = \underline{E}_{\Lambda}^{\star}(X_s \,|\, X_u = x_u, X_t = x) = E_{P_{\underline{\lambda}}}(X_s \,|\, X_u = x_u, X_t = x)$$

and

$$\overline{E}_{\Lambda}(X_s \,|\, X_u = x_u, X_t = x) = \overline{E}_{\Lambda}^{\star}(X_s \,|\, X_u = x_u, X_t = x) = E_{P_{\overline{\lambda}}}(X_s \,|\, X_u = x_u, X_t = x)$$

The stated now immediately follows if we recall that

$$E_{P_{\underline{\lambda}}}(X_s \,|\, X_u = x_u, X_t = x) = \sum_{y=x}^{+\infty} y \, \psi_{\underline{\lambda}(s-t)}(y - x) = x + \underline{\lambda}(s - t),$$

and

$$E_{P_{\overline{\lambda}}}(X_s \,|\, X_u = x_u, X_t = x) = \sum_{y=x}^{+\infty} y \, \psi_{\overline{\lambda}(s-t)}(y - x) = x + \overline{\lambda}(s - t),$$

where both times the first equality follows from Proposition 92. ∎

Lemma 105 *Consider any counting process P. Fix any t, s in $\mathbb{R}_{\geq 0}$ with $t \leq s$, u in $\mathscr{U}_{<t}$ and (x_u, x) in $\mathscr{X}_{u \cup t}$. Then for any f in $\mathscr{K}_b(\mathscr{X})$,*

$$E_P(f(X_s) \,|\, X_u = x_u, X_t = x) = \inf f + E_P(f'(X_s) \,|\, X_u = x_u, X_t = x),$$

with $f' := f - \inf f$.

Proof Follows immediately from the definition of E_P. ∎

Proof of Theorem 19 First, we observe that

$$\inf f \leq f \leq f_{\max}.$$

Therefore, for any P in $\mathbb{P}_{\Lambda}$,

$$\inf f \leq E_P(f(X_s) \,|\, X_u = x_u, X_t = x) \leq E_P(f_{\max}(X_s) \,|\, X_u = x_u, X_t = x) \tag{125}$$

due to the monotonicity of E_P. Since $f_{\max}$ is clearly a non-decreasing bounded-below function, it follows from Proposition 17 and Proposition 92 that, for any P in $\mathbb{P}_{\Lambda}$,

$$E_P(f_{\max}(X_s) \,|\, X_u = x_u, X_t = x) \leq E_{P_{\overline{\lambda}}}(f_{\max}(X_s) \,|\, X_u = x_u, X_t = x) = \sum_{y=x}^{+\infty} f_{\max}(y) \, \psi_{\overline{\lambda}(s-t)}(y - x) < +\infty,$$

where the final inequality is precisely the condition on f of the statement. Because non-strict inequalities are preserved when taking infima and suprema, we infer from this that

$$\inf f \leq \underline{E}_{\Lambda}(f(X_s) \,|\, X_u = x_u, X_t = x) \leq \overline{E}_{\Lambda}(f(X_s) \,|\, X_u = x_u, X_t = x) \leq \sum_{y=x}^{+\infty} f_{\max}(y) \, \psi_{\overline{\lambda}(s-t)}(y - x) < +\infty.$$

This already settles the second part of the stated, namely that the lower and upper expectations are finite.

Next, we set out to prove the equalities of the statement. For any $\bar{x}$ in $\mathscr{X}$, we let $f_{\bar{x}} := f\mathbb{I}_{\leq\bar{x}} + f(\bar{x})\mathbb{I}_{>\bar{x}}$. By definition of the limit, we need to prove that

$$(\forall\varepsilon\in\mathbb{R}_{>0})(\exists x^\star\in\mathscr{X})(\forall\bar{x}\in\mathscr{X},\bar{x}\geq x^\star)\,|\underline{E}_\Lambda(f(X_s)\,|\,X_u=x_u,X_t=x)-\underline{P}^s_t(f_{\bar{x}}\,|\,x)|\leq\varepsilon \tag{126}$$

and

$$(\forall\varepsilon\in\mathbb{R}_{>0})(\exists x^\star\in\mathscr{X})(\forall\bar{x}\in\mathscr{X},\bar{x}\geq x^\star)\,|\overline{E}_\Lambda(f(X_s)\,|\,X_u=x_u,X_t=x)-\overline{P}^s_t(f_{\bar{x}}\,|\,x)|\leq\varepsilon. \tag{127}$$

To that end, we fix any ε in $\mathbb{R}_{>0}$, and choose any ε_1 and ε_2 in $\mathbb{R}_{>0}$ such that $\varepsilon_1+\varepsilon_2\leq\varepsilon$.

Our first step is to obtain a bound on

$$|E_P(f(X_s)\,|\,X_u=x_u,X_t=x)-E_P(f_{\bar{x}}(X_s)\,|\,X_u=x_u,X_t=x)|,$$

with P in $\mathbb{P}_\Lambda$ and $\bar{x}$ in $\mathscr{X}$. To that end, we let $f':=f-\inf f$, $f'_{\max}:=f_{\max}-\inf f$ and $f'_{\bar{x}}:=f'\mathbb{I}_{\leq\bar{x}}+f'(\bar{x})\mathbb{I}_{>\bar{x}}$ for any $\bar{x}$ in $\mathscr{X}$. Due to the condition on f of the statement and the properties of the Poisson distribution,

$$\sum_{y=x}^{+\infty}f'_{\max}(y)\psi_{\overline{\lambda}(s-t)}(y-x)=\sum_{y=x}^{+\infty}(f_{\max}(y)-\inf f)\psi_{\overline{\lambda}(s-t)}(y-x)=\sum_{y=x}^{+\infty}f_{\max}(y)\psi_{\overline{\lambda}(s-t)}(y-x)-\inf f<+\infty.$$

Hence, there is an $x^\star$ in $\mathscr{X}$ with $x^\star\geq x$ such that

$$(\forall\bar{x}\in\mathscr{X},\bar{x}\geq x^\star)\left|\sum_{y=\bar{x}+1}^{+\infty}f'_{\max}(y)\psi_{\overline{\lambda}(s-t)}(y-x)\right|=\left|\sum_{y=x}^{+\infty}f'_{\max}(y)\psi_{\overline{\lambda}(s-t)}(y-x)-\sum_{y=x}^{\bar{x}}f'_{\max}(y)\psi_{\overline{\lambda}(s-t)}(y-x)\right|\leq\varepsilon_1. \tag{128}$$

Fix any $\bar{x}$ in $\mathscr{X}$ with $\bar{x}\geq x^\star$, and observe that

$$\mathbb{I}_{\leq\bar{x}}f'\leq f'\leq\mathbb{I}_{\leq x}f'+f'_{\max}\mathbb{I}_{>\bar{x}}. \tag{129}$$

and

$$\mathbb{I}_{\leq\bar{x}}f'\leq f'_{\bar{x}}\leq\mathbb{I}_{\leq x}f'+f'_{\max}\mathbb{I}_{>\bar{x}}. \tag{130}$$

Fix any P in $\mathbb{P}_\Lambda$. Then due to Equation (129) and the monotonicity of E_P,

$$E_P([\mathbb{I}_{\leq\bar{x}}f'](X_s)\,|\,X_u=x_u,X_t=x)\leq E_P(f'(X_s)\,|\,X_u=x_u,X_t=x)\leq E_P([\mathbb{I}_{\leq\bar{x}}f'+\mathbb{I}_{>\bar{x}}f'_{\max}](X_s)\,|\,X_u=x_u,X_t=x). \tag{131}$$

Similarly,

$$E_P([\mathbb{I}_{\leq\bar{x}}f'](X_s)\,|\,X_u=x_u,X_t=x)\leq E_P(f'_{\bar{x}}(X_s)\,|\,X_u=x_u,X_t=x)\leq E_P([\mathbb{I}_{\leq\bar{x}}f'+\mathbb{I}_{>\bar{x}}f'_{\max}](X_s)\,|\,X_u=x_u,X_t=x). \tag{132}$$

Furthermore, it follows from the linearity of E_P that

$$E_P([\mathbb{I}_{\leq\bar{x}}f'+\mathbb{I}_{>\bar{x}}f'_{\max}](X_s)\,|\,X_u=x_u,X_t=x)=E_P([\mathbb{I}_{\leq\bar{x}}f'](X_s)\,|\,X_u=x_u,X_t=x)+E_P([\mathbb{I}_{>\bar{x}}f'_{\max}](X_s)\,|\,X_u=x_u,X_t=x). \tag{133}$$

Alternatively, we obtain Equation (133) as follows. Recall that

$$E_P([\mathbb{I}_{\leq\bar{x}}f'+\mathbb{I}_{>\bar{x}}f'_{\max}](X_s)\,|\,X_u=x_u,X_t=x)=\int_0^{\sup f'}P(\{[\mathbb{I}_{\leq\bar{x}}f'+\mathbb{I}_{>\bar{x}}f'_{\max}](X_s)>\alpha\}\,|\,X_u=x_u,X_t=x)\,d\alpha.$$

Observe now that for all α in $\mathbb{R}_{\geq 0}$,

$$\begin{aligned}\{[\mathbb{I}_{\leq\bar{x}}f'+\mathbb{I}_{>\bar{x}}f'_{\max}](X_s)>\alpha\}&=\{\omega\in\Omega\colon\mathbb{I}_{\leq\bar{x}}(\omega(s))f'(\omega(s))+\mathbb{I}_{>\bar{x}}(\omega(s))f'_{\max}(\omega(s))>\alpha\}\\&=\{\omega\in\Omega\colon\big(\omega(s)\leq\bar{x},f'(\omega(s))>\alpha\big)\text{ or }\big(\omega(s)>\bar{x},f'_{\max}(\omega(s))>\alpha\big)\}\\&=\{\omega\in\Omega\colon\omega(s)\leq\bar{x},f'(\omega(s))>\alpha\}\cup\{\omega\in\Omega\colon\omega(s)>\bar{x},f'_{\max}(\omega(s))>\alpha\}\\&=\{[\mathbb{I}_{\leq\bar{x}}f'](X_s)>\alpha\}\cup\{[\mathbb{I}_{>\bar{x}}f'_{\max}](X_s)>\alpha\},\end{aligned}$$

where the union is one of two disjoint sets. Furthermore, the first set of this union is clearly empty if $\alpha\geq f_{\max}(\bar{x})$. We use our decomposition of the level sets, to yield

$$E_P([\mathbb{I}_{\leq\bar{x}}f'+\mathbb{I}_{>\bar{x}}f'_{\max}](X_s)\,|\,X_u=x_u,X_t=x)$$

$$= \int_0^{\sup f'} P(\{[\mathbb{I}_{\leq \bar{x}} f'](X_s) > \alpha\} \mid X_u = x_u, X_t = x) + P(\{[\mathbb{I}_{>\bar{x}} f'_{\max}](X_s) > \alpha\} \mid X_u = x_u, X_t = x) \, d\alpha$$

$$= \int_0^{\sup \mathbb{I}_{\leq \bar{x}} f'} P(\{[\mathbb{I}_{\leq \bar{x}} f'](X_s) > \alpha\} \mid X_u = x_u, X_t = x) + \int_0^{\sup f'} P(\{[\mathbb{I}_{>\bar{x}} f'_{\max}](X_s) > \alpha\} \mid X_u = x_u, X_t = x) \, d\alpha$$

$$= E_P([\mathbb{I}_{\leq \bar{x}} f'](X_s) \mid X_u = x_u, X_t = x) + E_P([\mathbb{I}_{>\bar{x}} f'_{\max}](X_s) \mid X_u = x_u, X_t = x),$$

where the for the second equality we have furthermore used the linearity of the (improper) Riemann integral.

In any case, it now follows from Equations (131)–(133) that

$$\left| E_P(f'(X_s) \mid X_u = x_u, X_t = x) - E_P(f'_{\bar{x}}(X_s) \mid X_u = x_u, X_t = x) \right| \leq E_P([\mathbb{I}_{>\bar{x}} f'_{\max}](X_s) \mid X_u = x_u, X_t = x).$$

Observe now that $\mathbb{I}_{>\bar{x}} f'_{\max}$ is a non-decreasing and bounded below function. Therefore, it follows from Proposition 17 and Proposition 92 that

$$\left| E_P(f'(X_s) \mid X_u = x_u, X_t = x) - E_P(f'_{\bar{x}}(X_s) \mid X_u = x_u, X_t = x) \right| \leq \sum_{y=\bar{x}+1}^{+\infty} f'_{\max} \psi_{\bar{\lambda}(s-t)}(y - x).$$

From this, Lemma 105 and Equation (128), we now infer that

$$(\forall \bar{x} \in \mathscr{X}, \bar{x} \geq x^\star)(\forall P \in \mathbb{P}_\Lambda) \, |E_P(f(X_s) \mid X_u = x_u, X_t = x) - E_P(f_{\bar{x}}(X_s) \mid X_u = x_u, X_t = x)| \leq \varepsilon_1 \tag{134}$$

Next, we recall from Proposition 100 that

$$(\forall \bar{x} \in \mathscr{X})(\exists P_\mathrm{l} \in \mathbb{P}_\Lambda) \, \left| \underline{P}_t^s(f_{\bar{x}} \mid x) - E_{P_\mathrm{l}}(f_{\bar{x}}(X_s) \mid X_u = x_u, X_t = x) \right| \leq \varepsilon_2 \tag{135}$$

and, due to conjugacy, that

$$(\forall \bar{x} \in \mathscr{X})(\exists P_\mathrm{u} \in \mathbb{P}_\Lambda) \, \left| \overline{P}_t^s(f_{\bar{x}} \mid x) - E_{P_\mathrm{u}}(f_{\bar{x}}(X_s) \mid X_u = x_u, X_t = x) \right| \leq \varepsilon_2. \tag{136}$$

Everything is now set up for us to verify Equations (126) and (127). We here only verify the former, the latter follows from entirely similar reasoning. Fix any $\bar{x}$ in $\mathscr{X}$ such that $\bar{x} \geq x^\star$. On the one hand, we observe that

$$\underline{E}_\Lambda(f(X_s) \mid X_u = x_u, X_t = x) \leq E_{P_\mathrm{l}}(f(X_s) \mid X_u = x_u, X_t = x) \leq E_{P_\mathrm{l}}(f_{\bar{x}}(X_s) \mid X_u = x_u, X_t = x) + \varepsilon_1$$
$$\leq \underline{P}_t^s(f_{\bar{x}} \mid x) + \varepsilon_1 + \varepsilon_2 \leq \underline{P}_t^s(f_{\bar{x}} \mid x) + \varepsilon,$$

where the first equality holds because P_l belongs to $\mathbb{P}_\Lambda$, and where for the subsequent inequalities we have used Equation (134), Equation (135) and our condition on ε_1 and ε_2. On the other hand, we observe that

$$\underline{E}_\Lambda(f(X_s) \mid X_u = x_u, X_t = x) = \inf\{E_P(f(X_s) \mid X_u = x_u, X_t = x) \colon P \in \mathbb{P}_\Lambda\}$$
$$\geq \inf\{E_P(f_{\bar{x}}(X_s) \mid X_u = x_u, X_t = x) - \varepsilon_1 \colon P \in \mathbb{P}_\Lambda\}$$
$$= \inf\{E_P(f_{\bar{x}}(X_s) \mid X_u = x_u, X_t = x) \colon P \in \mathbb{P}_\Lambda\} - \varepsilon_1$$
$$= \underline{E}_\Lambda(f_{\bar{x}}(X_s) \mid X_u = x_u, X_t = x) - \varepsilon_1$$
$$= \underline{P}_t^s(f_{\bar{x}} \mid x) - \varepsilon_1 \geq \underline{P}_t^s(f_{\bar{x}} \mid x) - \varepsilon,$$

where the first two equalities follow from Equation (8), the first inequality follows from Equation (134) and the final equality follows from Theorem 15 because $f_{\bar{x}}$ is clearly bounded. It is now clear that these two observations imply Equation (126), as required. ∎

Appendix I. Supplementary Material for Section 7

Proof of Proposition 20 We first consider the five properties for $\underline{E}_\Lambda^\star$. Properties (i) and (iii)–(v) follow almost immediately from Equation (8) and Proposition 92. To verify (ii), we observe that, for all Δ in $\mathbb{R}_{>0}$,

$$\underline{E}_\Lambda^\star(\mathbb{I}_{(X_{t+\Delta} \geq x+2)} \mid X_u = x_u, X_t = x) = \inf\{E_{P_\lambda}(\mathbb{I}_{(X_{t+\Delta} \geq x+2)} \mid X_u = x_u, X_t = x) \colon \lambda \in \Lambda\}$$
$$= \inf\{P_\lambda(X_{x+2} \geq x + 2 \mid X_u = x_u, X_t = x) \colon \lambda \in \Lambda\}$$

$$= \inf\{1 - \psi_{\lambda\Delta}(0) - \psi_{\lambda\Delta}(1) \colon \lambda \in \Lambda\}$$
$$= 1 - \psi_{\overline{\lambda}\Delta}(0) - \psi_{\overline{\lambda}\Delta}(1).$$

where we have used Equation (8) for the first equality, Equation (4) for the second equality and Proposition 81 for the third equality. Similarly, if $t > 0$, then

$$\underline{E}_{\Lambda}^{\star}(\mathbb{I}_{(X_t \geq x+2)} \,|\, X_u = x_u, X_{t-\Delta} = x)1 - \psi_{\overline{\lambda}\Delta}(0) - \psi_{\overline{\lambda}\Delta}(1).$$

Therefore,

$$\lim_{\Delta \to 0^+} \frac{\underline{E}_{\Lambda}^{\star}(\mathbb{I}_{(X_{t+\Delta} \geq x+2)} \,|\, X_u = x_u, X_t = x)}{\Delta} = \lim_{\Delta \to 0^+} \frac{1 - \psi_{\overline{\lambda}\Delta}(0) - \psi_{\overline{\lambda}\Delta}(1)}{\Delta} = \lim_{\Delta \to 0^+} \frac{1 - e^{-\overline{\lambda}\Delta} - \overline{\lambda}\Delta e^{-\overline{\lambda}\Delta}}{\Delta} = 0,$$

and similarly for the limit from the left if $t > 0$.

Next, we consider the five properties for $\underline{E}_{\Lambda}$. Properties (i), (iii) and (v) follow almost immediately from Theorem 15. Property (ii) follows immediately from Theorem 15, Equation (20) and Lemma 52. Finally, property (iv) follows immediately from Theorem 15, Equation (20) and Lemma 53. ∎