COUNTABLE-STATE STOCHASTIC PROCESSES WITH CÀDLÀG SAMPLE PATHS

ALEXANDER ERREYGERS AND JASPER DE BOCK

ABSTRACT. The Daniell–Kolmogorov Extension Theorem is a fundamental result in the theory of stochastic processes, as it allows one to construct a stochastic process with prescribed finite-dimensional distributions. However, it is well-known that the domain of the constructed probability measure – the product sigma-algebra in the set of all paths – is not sufficiently rich. This problem is usually dealt with through a modification of the stochastic process, essentially changing the sample paths so that they become càdlàg. Assuming a countable state space, we provide an alternative version of the Daniell–Kolmogorov Extension Theorem that does not suffer from this problem, in that the domain is sufficiently rich and we do not need a subsequent modification step: we assume a rather weak regularity condition on the finite-dimensional distributions, and directly obtain a probability measure on the product sigma-algebra in the set of all càdlàg paths.

1. INTRODUCTION

A stochastic process is a model for a system whose state changes over time in an uncertain manner. More formally, a stochastic process is a joint uncertainty model for a sequence $(X_t)_{t \in \mathbb{T}}$ of $\mathcal{X}$-valued maps on some sample space $\mathfrak{S}$ indexed by $\mathbb{T}$, where $\mathcal{X}$ is the state space of the system and $\mathbb{T}$ is an infinite subset of the real numbers that is interpreted as the time axis. Stochastic processes have been and still are an active field of research, as is clear from the large number of monographs that have appeared on the subject [18, 10, 3, 6, 9, 15, 24]. One usually derives the joint uncertainty model for $(X_t)_{t \in \mathbb{T}}$ from the finite-dimensional distributions: the joint uncertainty models for $(X_t)_{t \in T}$, where T ranges over all finite subsets of $\mathbb{T}$. For example, Kolmogorov [18, 19] already considered a stochastic processes with $\mathcal{X} = \mathbb{R}$ in his seminal 'Grundbegriffe', formulating what is now known as the Daniell–Kolmogorov Extension Theorem [19, Chapter III, § 4]. However, it is well-known [26, Section 5.1.2] that if $\mathbb{T}$ is not countable, then the domain of the resulting stochastic process – the product σ-algebra generated by the cylinder events – is not rich enough.

The first one to get around this problem was Doob [10], who in his influential work discusses stochastic processes in a very general setting. He explains why the product σ-algebra of events is not sufficiently rich [10, Chapter II, Section 2], but also comes up with a solution: the notion of 'separability of a stochastic process' [10, pp. 51 and 52], which ensures that the product σ-algebra of events *is* sufficiently rich. In the particular case of 'Markov processes with infinitely many states and continuous parameter', his Theorem 2.4 [10, p. 266] implies that the sample paths of any 'separable Markov process' are almost surely 'step functions' [10, p. 245 and 246]. Billingsley [3, Sections 36 and 37] does something similar, but only

for stochastic processes with the real numbers as state space: he also argues that the Daniell–Kolmogorov Extension Theorem does not suffice [3, pp. 492–494], and solves this through his notion of 'separability' [3, pp. 526–527].

Like Doob [10] and Billingsley [3], Breiman [6] treats stochastic processes in the particular setting where the state space is the set of real numbers, and explains that the standard construction is insufficient. He solves this problem a bit differently, though: he imposes a form of 'continuity in probability' [6, Definition 12.15] on the stochastic process, and shows that in combination with 'almost-sure absolute continuity' this allows the construction of a modification with continuous – often also called regular – sample paths [6, Theorem 12.6]. More recently, Borovkov [5, Section 18.2] – following Cramér & Leadbetter [9] – does not rely on 'almost-sure absolute continuity', but argues that in order to get a modification with regular sample paths, one needs a quantitative bound on the continuity. See also [8, 17].

Fewer authors consider stochastic processes with state spaces other than the real numbers. For example, Gikhman and Skorokhod [15] consider the very general setting of metric spaces for time domain and state space, but for most of their results they need compactness [15, Chapter 4, Section 4]. Rogers and Williams [24, Chapter II] also give a very broad account, but they only solve the issue of the product σ-algebra in some particular cases, for example that of Markov (or Feller-Dynkyn) processes [24, Chapter III]. Fristedt and Gray [14, Chapter 31] define what they call a '(pure-jump) Markov process' on the set of càdlàg paths (for a general Polish state space), but they never really move past the (finitary) cylinder events. Finally, König [21] gives a solution to this problem that does not involve separability or modifications: his Theorem 3.2 extends the probability measure on the product σ-algebra to a 'maximal non-sequential inner extension' on a larger domain, but never really shows that this domain contains 'all' the interesting events.

This brings us to our main contribution: we give a version of the Daniell–Kolmogorov Extension Theorem with a σ-algebra that is sufficiently rich. We will do so for a non-empty and countable state space $\mathcal{X}$, and an arbitrary time domain $\mathbb{T} \subseteq \mathbb{R}$. As is customary, we interpret the elements of $\mathbb{T}$ as time points; two important cases are $\mathbb{T} = \mathbb{N}$ and $\mathbb{T} = \mathbb{R}_{\geq 0} = [0, +\infty[$. We intend to construct a joint uncertainty model for the state of the system in a (not necessarily proper) subset of the index set $\mathbb{T}$; henceforth, we will denote this subset by $\mathcal{T}$. The usual choice is $\mathcal{T} = \mathbb{T}$, but other choices are sometimes useful as well. For example, in the setting of 'Markovian imprecise jump processes' (also known as imprecise continuous-time Markov chains) [20, 13], we encounter the case $\mathbb{T} = \mathbb{R}_{\geq 0}$ and $\mathcal{T} = \{t_1, \ldots t_{n-1}\} \cup [t_n, +\infty[$, with $t_1, \ldots, t_n \in \mathbb{R}_{\geq 0}$ such that $t_1 < \cdots < t_n$.

Our main result is Theorem 11, which is similar to the Daniell–Kolmogorov Extension Theorem in that it starts from a consistent family of finite-dimensional distributions, but also different in the following ways. A first difference is that we specify the finite-dimensional distributions for finite subsets of $\mathcal{T} \subseteq \mathbb{T}$ instead of $\mathbb{T}$. A second and more important difference is that we not only require consistency of these finite-dimensional distributions, but also regularity (Definition 4), and it is this that allows us to obtain a probability measure on the product σ-algebra for the regular (in this case càdlàg) paths instead of, as is customary, on the product σ-algebra for all paths. This way, we get rid of the additional step of constructing a regular modification.

1.1. **Some notation regarding tuples.** For any non-empty subset $\mathcal{S}$ of $\mathbb{T}$, we let $\mathcal{U}_{\mathcal{S}}$ be the set of all (non-empty) tuples $(t_1, \ldots, t_n) \in \bigcup_{k \in \mathbb{N}} \mathcal{S}^k$ that are increasing, so with $t_1 < \cdots < t_n$; if $\mathcal{S} = \mathcal{T}$, we simply write $\mathcal{U}$. We will usually denote a generic tuple in the set $\mathcal{U}_{\mathbb{T}}$ by u, but sometimes also by v or w. For any two tuples of time

points $u = (r_1, \ldots, r_n)$ and $v = (s_1, \ldots, s_m)$ in $\mathcal{U}_{\mathbb{T}}$, we write $u \sqsubseteq v$ if all time points in u are included in v, in the sense that $\{r_1, \ldots, r_n\} \subseteq \{s_1, \ldots, s_m\}$.

For any $u = (t_1, \ldots, t_n) \in \mathcal{U}_{\mathbb{T}}$, we let $\mathcal{X}_u$ denote the set of all n-tuples of states $(x_{t_1}, \ldots, x_{t_n})$ indexed by the time points in u, and we usually denote a generic n-tuple of states in $\mathcal{X}_u$ by x_u, y_u or z_u. Furthermore, given $u = (t_1, \ldots, t_n) \in \mathcal{U}_{\mathbb{T}}$ and $x_u = (x_{t_1}, \ldots, x_{t_n}) \in \mathcal{X}_u$ and for all $v = (s_1, \ldots, s_m) \in \mathcal{U}_{\mathbb{T}}$ such that $v \sqsubseteq u$, we let $x_v := (x_{s_1}, \ldots, x_{s_m})$ be the m-tuple in $\mathcal{X}_v$ that consists of those components of x_u with index in v. Rogers & Williams [24, Chapter II, Section 25 onwards] use slightly different notation: they consider finite subsets U of $\mathcal{T}$ instead of increasing finite tuples u in $\mathcal{U} \subseteq \bigcup_{k \in \mathbb{N}} \mathcal{T}^k$, and focus on the set $\mathcal{X}^U$ of $\mathcal{X}$-valued maps from U to $\mathcal{X}$ instead of on $\mathcal{X}_u$. It should be clear that these two approaches are essentially equivalent; nevertheless, we choose to use increasing tuples because these will come in handy in several places, for example in Sections 4.2 and 4.3 and Appendix C.1.

2. Constructing a stochastic process

We want to model a system whose state, which takes values in $\mathcal{X}$, changes along the time axis $\mathbb{T}$ in an uncertain manner. For this, we turn to (measure-theoretical) probability theory, as for example outlined in [24, 3, 14, 6, 17]. This means that we set out to determine a suitable probability space $(\mathfrak{S}, \mathfrak{A}, P)$, where the non-empty set $\mathfrak{S}$ is called the *sample space*, $\mathfrak{A}$ is a σ-algebra of *events* – subsets of $\mathfrak{S}$ – that are of interest to us, and P is a probability measure on $\mathfrak{A}$. As we will presently see, we will only consider couples $(\mathfrak{S}, \mathfrak{A})$ of a specific form; for example, the sample space will be a set of paths.

2.1. **Paths and cylinder events.** Since our system changes state over $\mathbb{T}$, it makes sense to think of an element $\mathfrak{s}$ of the sample space $\mathfrak{S}$ as a map from $\mathbb{T}$ to $\mathcal{X}$; we call such an $\mathcal{X}$-valued map on $\mathbb{T}$ a *path*, and we denote the set of all paths by $\mathcal{X}^{\mathbb{T}}$. Hence, the obvious choice for the sample space $\mathfrak{S}$ is some (non-empty) set of paths $\Omega \subseteq \mathcal{X}^{\mathbb{T}}$. However, in many cases – for example when using coherent conditional probabilities to deal with conditioning, see [20, Eqn. (12)] – it makes sense to require that

$$(1) \qquad (\forall u \in \mathcal{U})(\forall x_u \in \mathcal{X}_u)(\exists \omega \in \Omega) \; \omega(u) = x_u,$$

where here and in the remainder, $\omega(u) := \big(\omega(t_1), \ldots, \omega(t_n)\big)$ for all $\omega \in \mathcal{X}^{\mathbb{T}}$ and $u = (t_1, \ldots, t_n) \in \mathcal{U}$. We think this is a reasonable (and harmless) requirement on Ω, so we henceforth assume it.

Now that we have established that $\Omega \subseteq \mathcal{X}^{\mathbb{T}}$ is a sensible sample space, the question remains which events we are interested in. At the very least, we are interested in events regarding the state of the system in a finite number of time points in $\mathcal{T}$. Let us formalise these events.

For all $t \in \mathbb{T}$, we define the projector variable or evaluation map [24, Chapter II, Eqn. (25.1)]

$$X_t \colon \Omega \to \mathcal{X} \colon \omega \mapsto \omega(t);$$

we extend this notation in the obvious way to tuples of time points $u \in \mathcal{U}$:

$$X_u \colon \Omega \to \mathcal{X}_u \colon \omega \mapsto \omega(u).$$

Then an event regarding the state of the system in a finite number of time points in $\mathcal{T}$ is a *cylinder event* [24, Chapter II, Definition 25.4]: an event of the form

$$\{X_u \in A\} := \{\omega \in \Omega \colon X_u(\omega) \in A\} = \{\omega \in \Omega \colon \omega(u) \in A\}$$

with $u \in \mathcal{U}$ and $A \in \wp(\mathcal{X}_u)$.[1] The condition in Eqn. (1) ensures that for all $u \in \mathcal{U}$ and $A \in \wp(\mathcal{X}_u)$, the corresponding cylinder event $\{X_u \in A\}$ is the empty set if and

---

[1]$\wp(\mathcal{X}_u)$ denotes the powerset of $\mathcal{X}_u$.

only if $A = \varnothing$. Furthermore, it is not difficult to verify that the collection

$$\mathscr{C} \coloneqq \big\{\{X_u \in A\} \colon u \in \mathcal{U}, A \in \wp(\mathcal{X}_u)\big\}$$

of cylinder events with time points in $\mathcal{T}$ is an algebra of events.

In order to obtain the probability triple that we are after, it remains for us to determine (i) the subset Ω of $\mathcal{X}^{\mathbb{T}}$ that we want to use as possibility space, and (ii) a probability measure P on a σ-algebra $\mathscr{A}$ in Ω that includes $\mathscr{C}$. Such a probability triple is what we will call a stochastic process.

Definition 1. A *stochastic process* is a probability triple $(\Omega, \mathscr{A}, P)$ such that $\Omega \subseteq \mathcal{X}^{\mathbb{T}}$ and $\mathscr{C} \subseteq \mathscr{A}$.

Our definition of a stochastic process differs somewhat from the 'usual' one, which considers a generic probability space $(\mathfrak{S}, \mathfrak{A}, P)$ and a family $(Y_t)_{t \in \mathbb{T}}$ of $\mathfrak{A}/\wp(\mathcal{X})$-measurable maps from $\mathfrak{S}$ to $\mathcal{X}$ – see for example [10, Chapter II], [15, Chapter IV], [6, Definition 12.1], [24, Chapter II, Definition 27.1], [3, Section 36], [14, Section 1.2] or [5, Definition 18.1.2]. The first difference is that we choose to restrict the sample space $\mathfrak{S}$ to a set of paths a priori. The second difference is that we are only interested in modelling our uncertainty about the state in the time points in $\mathcal{T} \subseteq \mathbb{T}$; in that sense, our definition corresponds to the usual one with $\mathfrak{S} = \Omega$ and $\mathfrak{A} = \mathscr{A} \supseteq \mathscr{C}$, at least if we only consider the family $(X_t)_{t \in \mathcal{T}}$ instead of the one indexed by $\mathbb{T}$.

To construct a stochastic process, one usually turns to the Daniell–Kolmogorov Extension Theorem – see [24, Chapter II, Theorem 31.1] or, if $\mathcal{X}$ is finite, [16, Section 1.5.2]. The idea behind this theorem is simple: given a model for our uncertainty about X_u for all $u \in \mathcal{U}$, it constructs a (joint) uncertainty model for $(X_t)_{t \in \mathcal{T}}$. Since we are dealing with a countable state space, the uncertainty models for X_u are finite-dimensional charges or distributions.

2.2. **Finite-dimensional charges and distributions.** A *collection of finite-dimensional charges* is a collection $\mu_\bullet \coloneqq (\mu_u)_{u \in \mathcal{U}}$ such that for all $u \in \mathcal{U}$, μ_u is a probability charge[2] on $\wp(\mathcal{X}_u)$. Usually, we will assume that for all $u \in \mathcal{U}$, the corresponding charge μ_u is countably additive, making it a *distribution*[3]; whenever this is the case, we speak of a *collection of finite-dimensional distributions* instead of a collection of finite-dimensional charges.

A collection of finite-dimensional charges – or distributions – $\mu_\bullet$ is said to be *consistent* – or alternatively, 'satisfies the compatibility condition' or 'has the projective property', see [24, Chapter II, Eqn. (29.6)] or [16, Eqn. (1.9)] – if for all $u, v \in \mathcal{U}$ such that $u \sqsubseteq v$,

$$\mu_u(A) = \mu_v\big(\{x_v \in \mathcal{X}_v \colon x_u \in A\}\big) \quad \text{for all } A \in \wp(\mathcal{X}_u).$$

For the Poisson process, with $\mathcal{X} = \mathbb{Z}_{\geq 0}$ and $\mathbb{T} = \mathbb{R}_{\geq 0} = \mathcal{T}$, the finite-dimensional distributions are derived from the Poisson distribution. As explained in, for example, [24, Chapter II, Section 33], [3, Section 23] or [17, Chapter 12], one fixes a

---

[2] We use the terminology introduced in [2, Definition 2.1.1]: given an algebra $\mathfrak{A}$ of events in $\mathfrak{S}$, a probability charge P is a real-valued map on $\mathfrak{A}$ that is non-negative and finitely additive, with $P(\mathfrak{S}) = 1$. See also [6, Definition A.5], [24, Chapter II, Section 4], [14, Section 7.2], [27, Definition 1.15] or [8, Definition A.1.2].

[3] We call a probability measure on the powerset of a countable possibility space a distribution. Of course, a distribution μ_u on $\wp(\mathcal{X}_u)$ is completely determined by the values it assumes on the atoms, so by the unique corresponding probability mass function

$$\mathcal{X}_u \to \mathbb{R} \colon x \mapsto \mu_u(\{x\}).$$

rate $\lambda \in \mathbb{R}_{\geq 0}$, and for all $u = (t_1, \ldots, t_n) \in \mathcal{U}$ and $x_u \in \mathcal{X}_u$, lets

$$\mu_u(x_u) := \begin{cases} \displaystyle\prod_{k=2}^n e^{-\lambda(t_k - t_{k-1})} \frac{\left(\lambda(t_k - t_{k-1})\right)^{(x_{t_k} - x_{t_{k-1}})}}{(x_{t_k} - x_{t_{k-1}})!} & \text{if } x_{t_1} \leq \cdots \leq x_{t_n}, \\ 0 & \text{otherwise.} \end{cases}$$

In finite-state and countable-state Markov processes, where one usually also takes $\mathbb{T} = \mathbb{R}_{\geq 0} = \mathcal{T}$, the collection of finite-dimensional distributions is constructed with the help of an initial distribution and a semi-group of transition matrices – corresponding to the matrix exponential of a (transition) rate matrix – as explained in [7, Part II], [16, Chapter 8], [24, Chapter III] or [23, Chapters 2 and 3].

2.3. From finite-dimensional charges to charges on the cylinder events.
The consistency condition allows us to construct a probability charge on the cylinder events $\mathscr{C}$ from a collection of finite-dimensional charges, which is the first step in establishing the Daniell-Kolomogorov Extension Theorem.

Proposition 1. *Consider a collection of finite-dimensional charges $\mu_\bullet$. Then there is a unique probability charge P on the algebra of cylinder events $\mathscr{C}$ such that*

$$P(\{X_u \in A\}) = \mu_u(A) \quad \text{for all } u \in \mathcal{U}, A \in \wp(\mathcal{X}_u)$$

if and only if $\mu_\bullet$ is consistent.

Proof. To see that the consistency of $\mu_\bullet$ is necessary, assume *ex absurdo* that there is a unique probability charge P on $\mathscr{C}$ with the property in the statement and that $\mu_\bullet$ is not consistent. Then there are $u, v \in \mathcal{U}$ such that $u \sqsubseteq v$ and $A \in \wp(\mathcal{X}_u)$ such that

$$\mu_u(A) \neq \mu_v(A') \quad \text{with } A' := \{x_v \in \mathcal{X}_v \colon x_u \in A\}.$$

Now $\{X_u \in A\} = \{X_v \in A'\}$, and therefore

$$\mu_u(A) = P(\{X_u \in A\}) = P(\{X_v \in A'\}) = \mu_v(A'),$$

which is a clear contradiction.

Next, we assume that $\mu_\bullet$ is consistent, and show that there is a unique probability charge P with the required property. We repeat the argument in the 'Start of Proof' of Theorem 30.1 in [24, Chapter II] in our slightly different setting.

First, fix some $u, v \in \mathcal{U}$, $A \in \wp(\mathcal{X}_u)$ and $B \in \wp(\mathcal{X}_v)$. Let $w \in \mathcal{U}$ be the unique tuple of time points that consists of all time points in u and v. Then by construction, $u \sqsubseteq w$ and $v \sqsubseteq w$. Hence, we can let

$$A^\star := \{x_w \in \mathcal{X}_w \colon x_u \in A\} \quad \text{and} \quad B^\star := \{x_w \in \mathcal{X}_w \colon x_u \in B\}.$$

If $A^\star = B^\star$, then it follows from the consistency of $\mu_\bullet$ that

$$\mu_u(A) = \mu_w(A^\star) = \mu_w(B^\star) = \mu_v(B).$$

Alternatively, if $A^\star \cap B^\star = \varnothing$, then it follows from the consistency condition that

$$(2) \qquad \mu_w(A^\star \cup B^\star) = \mu_w(A^\star) + \mu_w(B^\star) = \mu_u(A) + \mu_v(B).$$

Now consider two events $\tilde{A}, \tilde{B}$ in $\mathscr{C}$. Then there are some $u, v \in \mathcal{U}$, $A \in \wp(\mathcal{X}_u)$ and $B \in \wp(\mathcal{X}_v)$ such that $\tilde{A} = \{X_u \in A\}$ and $\tilde{B} = \{X_v \in B\}$. Let w, $A^\star$ and $B^\star$ be as before. Then $\{X_u \in A\} = \{X_w \in A^\star\}$ and $\{X_v \in B\} = \{X_w \in B^\star\}$ by construction. If $\tilde{A} = \tilde{B}$, then $\{X_w \in A^\star\} = \{X_w \in B^\star\}$, and it then follows from this and Eqn. (1) that $A^\star = B^\star$. If $\tilde{A} \cap \tilde{B} = \varnothing$, then $\tilde{A} \cup \tilde{B} = \{X_w \in A^\star \cup B^\star\}$ and $\tilde{A} \cap \tilde{B} = \{X_w \in A^\star \cap B^\star\} = \varnothing$, and it follows from the latter and Eqn. (1) that $A^\star \cap B^\star = \varnothing$.

Due to the preceding two observations and because $\mu_\bullet$ is a consistent collection of finite-dimensional charges, we can define the real-valued map P on $\mathscr{C}$ for all $\tilde{A} = \{X_u \in A\} \in \mathscr{C}$ by

$$P(\tilde{A}) = P(\{X_u \in A\}) \coloneqq \mu_u(A).$$

Indeed, this map is well-defined because it does not depend on the particular choice of u and A. To see this, it suffices to consider the case $\tilde{A} = \tilde{B}$ above, for which we had found that $A^\star = B^\star$ and therefore, that $\mu_u(A) = \mu_v(B)$. Furthermore, we trivially have that P is non-negative and that $P(\Omega) = P(\{X_u \in \mathcal{X}_u\}) = \mu_u(\mathcal{X}_u) = 1$. To see that P is finitely additive, we consider the case above with $\tilde{A} \cap \tilde{B} = \varnothing$, for which we found that $\tilde{A} \cup \tilde{B} = \{X_w \in A^\star \cup B^\star\}$ and $A^\star \cap B^\star = \varnothing$. It therefore follows from Eqn. (2) that

$$P(\tilde{A} \cup \tilde{B}) = P(\{X_w \in A^\star \cup B^\star\}) = \mu_w(A^\star \cup B^\star) = \mu_u(A) + \mu_v(B) = P(\tilde{A}) + P(\tilde{B}).$$

Hence, P is a probability charge. The uniqueness of P is obvious. $\square$

Proposition 1 allows us to go from a consistent collection of finite-dimensional charges to a probability charge, but we can also go the other way around: any (consistent) triple $(\Omega, \mathscr{A}, P)$ induces a collection of finite-dimensional charges that is consistent.

Definition 2. Consider a subset Ω of $\mathcal{X}^\mathbb{T}$, an algebra $\mathscr{A} \supseteq \mathscr{C}$ of events in Ω and a probability charge P on $\mathscr{A}$. Then for all $u \in \mathcal{U}$, the map

$$\mu_u \colon \wp(\mathcal{X}_u) \to [0,1] \colon A \mapsto \mu_u(A) \coloneqq P(\{X_u \in A\})$$

is a probability charge. Moreover, the collection $(\mu_u)_{u \in \mathcal{U}}$ is consistent, which is why we call it the *(collection of) finite-dimensional charges of P*. If P is countably additive, then μ_u is a distribution for all $u \in \mathcal{U}$; in this case, we call $(\mu_u)_{u \in \mathcal{U}}$ the *(collection of) finite-dimensional distributions of P*.

Proof. Follows immediately from the properties of (countably additive) probability charges. $\square$

2.4. **From the finite-dimensional distributions to a stochastic process.** Proposition 1 and Definition 2 show that for a fixed set of paths Ω that satisfies Eqn. (1), there is a one to one correspondence between consistent collections of finite-dimensional charges and probability charges on $\mathscr{C}$. It also follows from Definition 2 that any stochastic process $(\Omega, \mathscr{A}, P)$ induces a consistent collection of finite-dimensional distributions.

This raises the following question. Given a set of paths $\Omega \subseteq \mathcal{X}^\mathbb{T}$ and a consistent collection of finite-dimensional distributions $\mu_\bullet$, is there a probability measure P on some σ-algebra $\mathscr{A} \supseteq \mathscr{C}$ in Ω such that the finite-dimensional distributions of P are $\mu_\bullet$? The key to answering this question lies in Caratheodory's Extension Theorem – see for example [6, Theorem A.9], [24, Chapter II, Theorem 5.1], [14, Chapter 7, Theorem 14] or [8, Theorem A.1.17] – which says that any probability charge P on some algebra $\mathscr{A}$ can be extended to a probability measure P_σ on the generated σ-algebra $\sigma(\mathscr{A})$ if and only if P is countably additive, and that this extension is then unique. Indeed, since the consistent collection of finite-dimensional distributions $\mu_\bullet$ corresponds to a unique probability charge P on $\mathscr{C}$, it then follows that there is a probability measure on some σ-algebra $\mathscr{A} \supseteq \mathscr{C}$ if and only if P is countably additive.

In the remainder, we investigate the countable additivity of the induced probability charge P for two important sets of paths: the set of all paths in Section 3 and the set of all 'càdlàg' paths in Section 4.

3. The set of all paths

We denote the set of all paths $\mathcal{X}^{\mathbb{T}}$ by $\dot{\Omega}$; clearly, $\dot{\Omega}$ satisfies Eqn. (1). For all $t \in \mathbb{T}$ and $u \in \mathcal{U}$, we denote the corresponding projector variables for $\dot{\Omega}$ by $\dot{X}_t$ and $\dot{X}_u$. We do something similar for the cylinder events: for any tuple of time points $u \in \mathcal{U}$ and any subset A of $\mathcal{X}_u$, we denote the corresponding cylinder event by

$$\{\dot{X}_u \in A\} := \{\dot{\omega} \in \dot{\Omega} \colon \dot{\omega}(u) \in A\}.$$

Hence, the algebra of cylinder events for the set of all paths $\dot{\Omega}$ is

$$\dot{\mathscr{C}} := \big\{\{\dot{X}_u \in A\} \colon u \in \mathcal{U}, A \in \wp(\mathcal{X}_u)\big\}.$$

3.1. **Establishing countable additivity.** Crucially, the set of all paths $\dot{\Omega}$ allows us to establish the countable additivity of any probability charge P induced by a consistent collection of finite-dimensional distributions. Since this result is essentially well-known, at least in case $\mathcal{T} = \mathbb{T}$, we have relegated our proof to Appendix A further on.

Theorem 2. *For any consistent collection $\mu_\bullet$ of finite-dimensional distributions, the corresponding probability charge P on $\dot{\mathscr{C}}$ of Proposition 1 is countably additive.*

As we explained at the end of Section 2, Theorem 2 admits us to invoke Caratheodory's Extension Theorem to obtain a stochastic process $(\dot{\Omega}, \sigma(\dot{\mathscr{C}}), P)$ such that the finite-dimensional distributions of P are the prescribed ones $\mu_\bullet$. This result is known as the Daniell–Kolmogorov Extension Theorem, and is similar to – and essentially implied by – Theorem 31.1 in [24, Chapter II], or in case $\mathcal{X}$ is finite, the result outlined in [16, Section 1.5.2].

Theorem 3. *For any consistent collection of finite-dimensional distributions $\mu_\bullet$, there is a unique probability measure P on $\sigma(\dot{\mathscr{C}})$ such that*

$$P(\{\dot{X}_u \in A\}) = \mu_u(A) \quad \text{for all } u \in \mathcal{U}, A \in \wp(\mathcal{X}_u).$$

Proof. Follows immediately from Proposition 1, Theorem 2 and Caratheodory's Extension Theorem. □

3.2. **The insufficiency of the product sigma-algebra.** There is one major issue with Theorem 3: the generated σ-algebra $\sigma(\dot{\mathscr{C}})$ is not sufficiently rich, in the sense that many practically-relevant events do not belong to it. Simply put it comes down to this: the events in $\sigma(\dot{\mathscr{C}})$ only depend on the state of the paths in the time points in a countable subset of $\mathcal{T}$. We are by no means the first to signal this issue: Breiman [6, Proposition 12.8], Billingsley [3, Theorem 36.3] and Cohen & Elliot [8, Lemma A.2.2] all mention it, to give but three examples. We give a (slightly modified version of) Lemma 25.9 in [24, Chapter II], applicable to the general case in Section 2. In it, we use the following hitherto undefined notation: for all subsets $\mathcal{S}, \mathcal{R}$ of $\mathbb{T}$ such that $\mathcal{S} \subseteq \mathcal{R}$, we denote the restriction of $\varpi \colon \mathcal{R} \to \mathcal{X}$ to $\mathcal{S}$ by $\varpi|_{\mathcal{S}}$, and we let

$$\tilde{\mathscr{A}}_{\mathcal{S}} := \sigma\big(\big\{\{\varpi \in \mathcal{X}^{\mathcal{S}} \colon \varpi(s) \in A\} \colon s \in \mathcal{S}, A \in \wp(\mathcal{X})\big\}\big).$$

Lemma 4. *Consider a non-empty subset Ω of $\mathcal{X}^{\mathbb{T}}$ and let $\mathfrak{C}$ be the set of all non-empty countable subsets of $\mathcal{T}$. Then*

$$\sigma(\mathscr{C}) = \bigcup_{\mathcal{C} \in \mathfrak{C}} \big\{\{\omega \in \Omega \colon \omega|_{\mathcal{C}} \in \tilde{A}_{\mathcal{C}}\} \colon \tilde{A}_{\mathcal{C}} \in \tilde{\mathscr{A}}_{\mathcal{C}}\big\}.$$

In our proof, we make use of the following claim (without formal proof) of Rogers & Williams [24, Chapter II, Eqn. (25.3)].

Lemma 5. *For any two subsets $\mathcal{S}, \mathcal{R}$ of $\mathbb{T}$ such that $\mathcal{S} \subseteq \mathcal{R}$, the restriction operator $\bullet|_{\mathcal{S}} \colon \mathcal{X}^{\mathcal{R}} \to \mathcal{X}^{\mathcal{S}}$ is $\tilde{\mathscr{A}}_{\mathcal{R}}/\tilde{\mathscr{A}}_{\mathcal{S}}$-measurable.*

Proof. By a standard result in measure theory – see for example Proposition 2.3 in [24, Chapter II] – we need to look only at the sets $\{\psi \in \mathcal{X}^{\mathcal{S}} \colon \psi(s) \in A\}$ that generate $\tilde{\mathscr{A}}_{\mathcal{S}}$; more formally, this is the case if (and only if) for all $s \in \mathcal{S}$ and $A \in \wp(\mathcal{X})$,

$$\{\varpi \in \mathcal{X}^{\mathcal{R}} \colon \varpi|_{\mathcal{S}} \in \{\psi \in \mathcal{X}^{\mathcal{S}} \colon \psi(s) \in A\}\} \in \tilde{\mathscr{A}}_{\mathcal{R}},$$

and this condition is trivially satisfied because

$$\{\varpi \in \mathcal{X}^{\mathcal{R}} \colon \varpi|_{\mathcal{S}} \in \{\psi \in \mathcal{X}^{\mathcal{S}} \colon \psi(s) \in A\}\} = \{\varpi \in \mathcal{X}^{\mathcal{R}} \colon \varpi(s) \in A\} \in \tilde{\mathscr{A}}_{\mathcal{R}}.$$

$\square$

Proof of Lemma 4. We adapt the argument in [24, Proof of Lemma 25.9] to our slightly different setting. Since $\mathscr{C} = \{\dot{A} \cap \Omega \colon \dot{A} \in \dot{\mathscr{C}}\}$ by definition of $\mathscr{C}$ and $\dot{\mathscr{C}}$, it follows from a standard result in measure theory – see for example Theorem 10.1 in [3] – that

$$\sigma(\mathscr{C}) = \{\dot{A} \cap \Omega \colon \dot{A} \in \sigma(\dot{\mathscr{C}})\}.$$

On the other hand, it is clear that for all $\mathcal{C} \in \mathfrak{C}$,

$$\big\{\{\omega \in \Omega \colon \omega|_{\mathcal{C}} \in \tilde{A}_{\mathcal{C}}\} \colon \tilde{A}_{\mathcal{C}} \in \tilde{\mathscr{A}}_{\mathcal{C}}\big\} = \big\{\{\dot{\omega} \in \dot{\Omega} \colon \dot{\omega}|_{\mathcal{C}} \in \tilde{A}_{\mathcal{C}}\} \cap \Omega \colon \tilde{A}_{\mathcal{C}} \in \tilde{\mathscr{A}}_{\mathcal{C}}\big\}.$$

Hence, it suffices to prove that

$$\sigma(\dot{\mathscr{C}}) = \bigcup_{\mathcal{C} \in \mathfrak{C}} \big\{\{\dot{\omega} \in \dot{\Omega} \colon \dot{\omega}|_{\mathcal{C}} \in \tilde{A}_{\mathcal{C}}\} \colon \tilde{A}_{\mathcal{C}} \in \tilde{\mathscr{A}}_{\mathcal{C}}\big\},$$

which is precisely the equality in the statement in case $\Omega = \dot{\Omega}$. To simplify our notation, we let

$$\tilde{\mathscr{A}}_{\mathcal{C}}^{\uparrow \dot{\Omega}} \coloneqq \big\{\{\dot{\omega} \in \dot{\Omega} \colon \dot{\omega}|_{\mathcal{C}} \in \tilde{A}_{\mathcal{C}}\} \colon \tilde{A}_{\mathcal{C}} \in \tilde{\mathscr{A}}_{\mathcal{C}}\big\} \quad \text{for all } \mathcal{C} \in \mathfrak{C}.$$

First, we verify that $\mathscr{A} \coloneqq \bigcup_{\mathcal{C} \in \mathfrak{C}} \tilde{\mathscr{A}}_{\mathcal{C}}^{\uparrow \dot{\Omega}}$ is a σ-algebra. It is clear that $\mathscr{A}$ includes the empty set and is closed under taking complements, so we really only need to show that $\mathscr{A}$ is closed under countable unions. To this end, we fix any sequence $(\dot{A}_n)_{n \in \mathbb{N}}$ in $\mathscr{A}$. Then by definition, for all $n \in \mathbb{N}$ there are some $\mathcal{C}_n \in \mathfrak{C}$ and $\tilde{A}_{\mathcal{C}_n} \in \tilde{\mathscr{A}}_{\mathcal{C}_n}$ such that

$$\dot{A}_n = \{\dot{\omega} \in \dot{\Omega} \colon \dot{\omega}|_{\mathcal{C}_n} \in \tilde{A}_{\mathcal{C}_n}\}.$$

Then clearly, the set $\mathcal{C} \coloneqq \bigcup_{n \in \mathbb{N}} \mathcal{C}_n$ is a countable subset of $\mathcal{T}$, and therefore $\mathcal{C} \in \mathfrak{C}$. Furthermore, for all $n \in \mathbb{N}$,

$$\tilde{A}_{\mathcal{C}}^n \coloneqq \{\varpi \in \mathcal{X}^{\mathcal{C}} \colon \varpi|_{\mathcal{C}_n} \in \tilde{A}_{\mathcal{C}_n}\} \in \tilde{\mathscr{A}}_{\mathcal{C}}$$

because the restriction operator $\bullet|_{\mathcal{C}_n} \colon \mathcal{X}^{\mathcal{C}} \to \mathcal{X}^{\mathcal{C}_n}$ is $\tilde{\mathscr{A}}_{\mathcal{C}}/\tilde{\mathscr{A}}_{\mathcal{C}_n}$-measurable due to Lemma 5, and then

$$\dot{A}_n = \{\dot{\omega} \in \dot{\Omega} \colon \dot{\omega}|_{\mathcal{C}_n} \in \tilde{A}_{\mathcal{C}_n}\} = \{\dot{\omega} \in \dot{\Omega} \colon \dot{\omega}|_{\mathcal{C}} \in \tilde{A}_{\mathcal{C}}^n\}.$$

Since $\bigcup_{n \in \mathbb{N}} \tilde{A}_{\mathcal{C}}^n \in \tilde{\mathscr{A}}_{\mathcal{C}}$ because $\tilde{\mathscr{A}}_{\mathcal{C}}$ is a σ-algebra, it follows from this that

$$\bigcup_{n \in \mathbb{N}} \dot{A}_n = \bigcup_{n \in \mathbb{N}} \{\dot{\omega} \in \dot{\Omega} \colon \dot{\omega}|_{\mathcal{C}} \in \tilde{A}_{\mathcal{C}}^n\} = \left\{\dot{\omega} \in \dot{\Omega} \colon \dot{\omega}|_{\mathcal{C}} \in \bigcup_{n \in \mathbb{N}} \tilde{A}_{\mathcal{C}}^n\right\} \in \mathscr{A},$$

as required.

Next, we show that $\sigma(\dot{\mathscr{C}}) \subseteq \mathscr{A}$. Since $\mathscr{A}$ is a σ-algebra, it suffices to show that $\dot{\mathscr{C}} \subseteq \mathscr{A}$. To this end, we fix any $\dot{A} = \{\dot{X}_u \in A\} \in \dot{\mathscr{C}}$. If we enumerate the time points in u as $(t_1, \ldots, t_n)$ and let $\mathcal{C} \coloneqq \{t_1, \ldots, t_n\}$, then clearly

$$\tilde{A}_{\mathcal{C}} \coloneqq \{\varpi \in \mathcal{X}^{\mathcal{C}} : (\varpi(t_1), \ldots, \varpi(t_n)) \in A\} = \bigcup_{x_u \in A} \bigcap_{k=1}^{n} \{\varpi \in \mathcal{X}^{\mathcal{C}} : \varpi(t_k) = x_{t_k}\} \in \tilde{\mathscr{A}}_{\mathcal{C}}$$

because $A \subseteq \mathcal{X}_u$ is countable. Hence,

$$\dot{A} = \{\dot{X}_u \in A\} = \{\dot{\omega} \in \dot{\Omega} : \dot{\omega}|_{\mathcal{C}} \in \tilde{A}_{\mathcal{C}}\} \in \mathscr{A},$$

as required.

To obtain the equality that we are after, it remains for us to verify that $\mathscr{A} \subseteq \sigma(\dot{\mathscr{C}})$, or equivalently, that $\{\dot{\omega} \in \dot{\Omega} : \dot{\omega}|_{\mathcal{C}} \in \tilde{A}_{\mathcal{C}}\} \in \sigma(\dot{\mathscr{C}})$ for all $\mathcal{C} \in \mathfrak{C}$ and $\tilde{A}_{\mathcal{C}} \in \tilde{\mathscr{A}}_{\mathcal{C}}$. So let us fix any $\mathcal{C} \in \mathfrak{C}$ and $\tilde{A}_{\mathcal{C}} \in \tilde{\mathscr{A}}_{\mathcal{C}}$. Then

$$\{\dot{\omega} \in \dot{\Omega} : \dot{\omega}|_{\mathcal{C}} \in \tilde{A}_{\mathcal{C}}\} = \{\dot{\omega} \in \mathcal{X}^{\mathbb{T}} : \dot{\omega}|_{\mathcal{C}} \in \tilde{A}_{\mathcal{C}}\} \in \tilde{\mathscr{A}}_{\mathbb{T}}$$

because the restriction operator $\bullet|_{\mathcal{C}} \colon \mathcal{X}^{\mathbb{T}} \to \mathcal{X}^{\mathcal{C}}$ is $\tilde{\mathscr{A}}_{\mathbb{T}}/\tilde{\mathscr{A}}_{\mathcal{C}}$-measurable due to Lemma 5. Since

$$\tilde{\mathscr{A}}_{\mathbb{T}} = \sigma\big(\{\{\dot{\omega} \in \mathcal{X}^{\mathbb{T}} : \dot{\omega}(s) \in A\} \colon s \in \mathbb{T}, A \in \wp(\mathcal{X})\}\big)$$

and, for all $s \in \mathbb{T}$ and $A \in \wp(\mathcal{X})$,

$$\{\dot{\omega} \in \mathcal{X}^{\mathbb{T}} : \dot{\omega}(s) \in A\} = \{\dot{\omega} \in \dot{\Omega} : \dot{\omega}(s) \in A\} = \{\dot{X}_{(s)} \in A\},$$

we furthermore have that

$$\tilde{\mathscr{A}}_{\mathbb{T}} \subseteq \sigma\big(\{\{\dot{X}_u \in A\} \colon u \in \mathcal{U}, A \in \wp(\mathcal{X}_u)\}\big) = \sigma(\dot{\mathscr{C}}).$$

It therefore follows that $\{\dot{\omega} \in \dot{\Omega} : \dot{\omega}|_{\mathcal{C}} \in \tilde{A}_{\mathcal{C}}\} \in \tilde{\mathscr{A}}_{\mathbb{T}} \subseteq \sigma(\dot{\mathscr{C}})$, as required. $\square$

In the particular case where Ω is the set $\dot{\Omega}$ of all paths, and hence $\mathscr{C} = \dot{\mathscr{C}}$, the main take-away from Lemma 4 is that if $\mathcal{T}$ is not countable, then $\sigma(\dot{\mathscr{C}})$ may *not* contain 'all' events that are of interested to us. The following example illustrates this.

Example 1. Let $\mathbb{T} \coloneqq \mathbb{R}_{\geq 0} \eqqcolon \mathcal{T}$. Suppose we are interested in the event that our system is in some given state $x \in \mathcal{X}$ before some time $T \in \mathbb{R}_{>0}$:

$$\dot{A} \coloneqq \{\dot{\omega} \in \dot{\Omega} : (\exists t \in [0, T]) \, \dot{\omega}(t) = x\}.$$

Events like this are common in applications, for example in model checking [1]. Obviously, we can write this event as an uncountable union of events in $\dot{\mathscr{C}}$:

$$\dot{A} = \bigcup_{t \in [0,T]} \{\dot{X}_t = x\}.$$

However, the generated σ-algebra $\sigma(\dot{\mathscr{C}})$ is only guaranteed to be closed under countable unions. In fact, we can use Lemma 4 to prove that $\dot{A}$ does not belong to $\sigma(\dot{\mathscr{C}})$. Assume *ex absurdo* that $\dot{A} \in \sigma(\dot{\mathscr{C}})$. Then by Lemma 4, there are some non-empty countable subset $\mathcal{C}$ of $\mathcal{T}$ and subset $\tilde{A}$ of $\mathcal{X}^{\mathcal{C}}$ such that

$$\dot{A} = \{\dot{\omega} \in \dot{\Omega} : \dot{\omega}|_{\mathcal{C}} \in \tilde{A}\}.$$

Take any path $\dot{\omega} \in \dot{A}^c$. Then clearly, $\dot{\omega}(t) \neq x$ for all $t \in [0, T]$. Now take any $t^\star \in [0, T] \setminus \mathcal{C}$, and let $\tilde{\omega} \in \dot{\Omega}$ be the path defined for all $t \in \mathcal{T}$ by $\tilde{\omega}(t) \coloneqq x$ if $t = t^\star$ and $\tilde{\omega}(t) = \dot{\omega}(t)$ otherwise. Then on the one hand, $\tilde{\omega} \in \dot{A}$ by construction. On the other hand, $\tilde{\omega}|_{\mathcal{C}} = \dot{\omega}|_{\mathcal{C}}$ by construction, and since $\dot{\omega} \notin \dot{A}$ this implies that $\tilde{\omega} \notin \dot{A}$, a clear contradiction.

The reason why this problem occurs is because, whenever $\mathcal{T}$ is uncountable, (the restriction to $\mathcal{T}$ of) a path ω in $\dot{\Omega}$ is not fully defined by the states it assumes on a countable subset of $\mathcal{T}$. The obvious solution to this problem is therefore to focus on a subset Ω of the set of all paths $\dot{\Omega}$ that does not contain such paths, and it is standard to choose the set of càdlàg paths for this; see Section 4 further on.

The usual manner to proceed is then to construct a 'modification' $\tilde{X}_\bullet$ of $\dot{X}_\bullet$ that has 'càdlàg sample paths', which can only be done under some conditions on the finite-dimensional distributions; we are only aware of work where this is done for Markov processes – that is, only for finite-dimensional distributions of a specific form. Most authors then proceed with this modified process $\tilde{X}_\bullet$ as if it was the original process $\dot{X}_\bullet$, but this modification is a drastically changed version of $\dot{X}_\bullet$. Why one would be allowed to even change the outcomes (or sample paths) of a stochastic process is often not really given much thought or motivation, expect perhaps for the (implicit) pragmatic justification that it 'just works'. That said, one could use the modified process $\tilde{X}_\bullet$ to obtain a probability measure on a (sufficiently rich) σ-algebra of events in the set of càdlàg paths, but most – if not almost all – works that we are aware of skip this crucial step. We have found two notable exceptions: Cramér & Leadbetter [9, Sections 3.2, 3.3 and 3.6] more or less explain how this can be done, as does Borovkov [5, Section 18.1] in less detail.

In contrast, we now set out to provide a necessary and sufficient regularity condition on any (so not necessarily Markov) collection of finite-dimensional distributions for there to be a corresponding probability measure on the σ-algebra of events generated by the cylinder events in the set of *càdlàg* paths – instead of the usual product σ-algebra $\sigma(\dot{\mathscr{C}})$ of events in the set of *all* paths $\dot{\Omega}$. The crucial benefit of our approach is that we do *not* need to modify our stochastic process to avoid the issue illustrated in Example 1 – that is, the lack of richness of the generated σ-algebra

4. The set of càdlàg paths

Before we can introduce the càdlàg paths, we need to deal with some topological considerations. Consider a subset $\mathcal{S}$ of $\mathbb{R}$. A real number $s \in \mathbb{R}$ is an *isolated point* of $\mathcal{S}$ if there is some $\delta \in \mathbb{R}_{>0}$ such that $]s - \delta, s + \delta[\cap \mathcal{S} = \{s\}$, a *right-sided limit point* of $\mathcal{S}$ if $]s, s + \delta[\cap \mathcal{S} \neq \varnothing$ for all $\delta \in \mathbb{R}_{>0}$ and a *left-sided limit point* of $\mathcal{S}$ if $]s - \delta, s[\cap \mathcal{S} \neq \varnothing$ for all $\delta \in \mathbb{R}_{>0}$. We collect all isolated points, right-sided limit points and left-sided limit points of $\mathcal{S}$ in $\mathrm{is}(\mathcal{S})$, $\mathrm{rlims}(\mathcal{S})$ and $\mathrm{llims}(\mathcal{S})$, respectively, and let $\mathrm{lims}(\mathcal{S}) \coloneqq \mathrm{rlims}(\mathcal{S}) \cup \mathrm{llims}(\mathcal{S})$; obviously, $\mathrm{is}(\mathcal{S}) \cap \mathrm{lims}(\mathcal{S}) = \varnothing$. The union of the isolated points and the right-sided limits points of $\mathcal{S}$ is its closure with respect to the *lower limit topology* – also known as the 'right half-open topology' – on $\mathbb{R}$ [25, Section 16.33.e], the union of the isolated points and the left-sided limits points of $\mathcal{S}$ is its closure with respect to the *upper limit topology* – also know as the 'left half-open interval topology' – on $\mathbb{R}$ [11, Chapter III, Section 3, Ex. 4], while the union of the isolated points and the limit points of $\mathcal{S}$ is the closure with respect to the usual topology on $\mathbb{R}$.

Next, we take directional limits along $\mathcal{S}$. The following is a specialization of the general (topological) definition of a limit [14, Appendix C, Definition 8 and Promblem 18] to the lower and upper limit topologies, see also [25, Section 15.21]. Consider a map ϕ from $\mathcal{S}$ to a metric space $(\mathbb{M}, \rho)$; we will need two particular cases: (i) the set of real numbers $\mathbb{R}$ with the usual metric induced by the absolute value (in Definition 4 further on), and (ii) the set of states $\mathcal{X}$ with the Kronecker metric (in Definition 3 further on). Then for any right-sided limit point t of $\mathcal{S}$, we say that *the right-sided limit of ϕ in t along $\mathcal{S}$ exists* if there is some – necessarily

unique – $\ell \in \mathbb{M}$ such that

$$\big(\forall \epsilon \in \mathbb{R}_{>0}\big)\big(\exists \delta \in \mathbb{R}_{>0}\big)\big(\forall r \in \,]t, t+\delta[\,\cap\, \mathcal{S}\big)\ \rho(\phi(r), \ell) < \epsilon;$$

whenever this is the case, we denote this limit ℓ by $\lim_{\mathcal{S}\ni r \searrow t} \phi(r)$. Similarly, for any left-sided limit point t of $\mathcal{S}$, we say that *the left-sided limit of ϕ in t along $\mathcal{S}$ exists* if there is some – necessarily unique – $\ell \in \mathbb{M}$ such that

$$\big(\forall \epsilon \in \mathbb{R}_{>0}\big)\big(\exists \delta \in \mathbb{R}_{>0}\big)\big(\forall s \in \,]t-\delta, t[\,\cap\, \mathcal{S}\big)\ \rho(\phi(s), \ell) < \epsilon,$$

and then we denote this limit ℓ by $\lim_{\mathcal{S}\ni s \nearrow t} \phi(s)$.

With this terminology and notation, we can define càdlàg paths as those paths that have left-sided and right-sided limits and are continuous from the right in all time points t in the closure of $\mathbb{T}$ where these concepts make sense.

Definition 3. Consider a non-empty subset $\mathcal{S}$ of $\mathbb{R}$. A map $\varpi \colon \mathcal{S} \to \mathcal{X}$ is *càdlàg* if it has a left-sided limit (along $\mathcal{S}$) in all $t \in \mathrm{llims}(\mathcal{S})$, meaning that

$$\big(\exists \delta \in \mathbb{R}_{>0}\big)\big(\exists x \in \mathcal{X}\big)\big(\forall s \in \mathcal{S} \cap \,]t-\delta, t[\big)\ \varpi(s) = x,$$

has a right-sided limit (along $\mathcal{S}$) in all $t \in \mathrm{rlims}(\mathcal{S})$, meaning that

$$\big(\exists \delta \in \mathbb{R}_{>0}\big)\big(\exists x \in \mathcal{X}\big)\big(\forall r \in \mathcal{S} \cap \,]t, t+\delta[\big)\ \varpi(r) = x,$$

and is right-continuous (along $\mathcal{S}$) in all $t \in \mathcal{S}\cap\mathrm{rlims}(\mathcal{S})$, meaning that $\lim_{\mathcal{S}\ni r \searrow t} \varpi(r) = \varpi(t)$. Càdlàg paths correspond to the case $\mathcal{S} = \mathbb{T}$; that is, càdlàg paths are càdlàg maps on $\mathbb{T}$. We collect all such càdlàg paths in $\mathring{\Omega}$.

It is straightforward to verify that $\mathring{\Omega}$ satisfies Eqn. (1). Another important property of càdlàg paths is that they can always be extended to a càdlàg map on $\mathbb{R}$. Because this is an intermediary technical result, we have relegated its proof to Appendix B.

Lemma 6. *Any càdlàg map $\varpi \colon \mathcal{S} \to \mathcal{X}$, and therefore also any càdlàg path $\mathring{\omega} \in \mathring{\Omega}$, can be extended to a map $\psi \colon \mathbb{R} \to \mathcal{X}$ that is right-continuous and has left-sided limits everywhere.*

For all $t \in \mathbb{T}$ and $u \in \mathcal{U}$, we denote the corresponding projector variable for the set of all càdlàg paths $\mathring{\Omega}$ by $\mathring{X}_t$ and $\mathring{X}_u$, and we use similar notation for the cylinder events: for any tuple of time points $u \in \mathcal{U}$ and any subset A of $\mathcal{X}_u$, we denote the corresponding cylinder event by

$$\{\mathring{X}_u \in A\} \coloneqq \big\{\mathring{\omega} \in \mathring{\Omega} \colon \mathring{\omega}(u) \in A\big\} = \{\dot{X}_u \in A\} \cap \mathring{\Omega}.$$

Hence, the algebra of cylinder events for the set of all càdlàg paths $\mathring{\Omega}$ is

$$\mathring{\mathscr{C}} \coloneqq \big\{\{\mathring{X}_u \in A\} \colon u \in \mathcal{U}, A \in \wp(\mathcal{X}_u)\big\} = \big\{\dot{A} \cap \mathring{\Omega} \colon \dot{A} \in \dot{\mathscr{C}}\big\}.$$

4.1. **The sufficiency of the generated sigma-algebra.** So are càdlàg paths fully defined by the states they assume on a countable subset of $\mathbb{T}$? The following result makes clear that this is the case, at least if $\mathbb{T}$ has a countable subset $\mathcal{D}$ that is dense in $\mathbb{T}$ for the lower limit topology, or equivalently, such that $\mathbb{T} \subseteq \mathcal{D}\cup\mathrm{rlims}(\mathcal{D})$; its simple proof can be found in Appendix B.

Lemma 7. *Consider a subset $\mathcal{D}$ of $\mathbb{T}$ such that $\mathbb{T} \subseteq \mathcal{D} \cup \mathrm{rlims}(\mathcal{D})$. Then for all $\mathring{\omega}_1, \mathring{\omega}_2 \in \mathring{\Omega}$,*

$$\mathring{\omega}_1 = \mathring{\omega}_2 \Leftrightarrow (\forall d \in \mathcal{D})\ \mathring{\omega}_1(d) = \mathring{\omega}_2(d).$$

Fortunately, the requirement that $\mathbb{T}$ has such a countable dense subset is always met.

Lemma 8. *Any subset $\mathcal{S}$ of $\mathbb{R}$ has a countable subset $\mathcal{D}$ that is dense with respect to the lower limit topology, meaning that*

$$\mathcal{S} = \left\{ s \in \mathcal{S} \colon (\forall \delta \in \mathbb{R}_{>0}) \ [s, s+\delta[\cap \mathcal{D} \neq \varnothing \right\} \subseteq \mathcal{D} \cup \mathrm{rlims}(\mathcal{D}).$$

Proof. According to [11, Chapter VIII, Section 7, Ex. 5], any subspace $\mathcal{S}'$ of the set of real numbers $\mathbb{R}$ equipped with the upper limit topology is separable, in the sense that it has a countable dense subset $\mathcal{D}'$ [11, Chapter VIII, Section 7, Definition 7.1]. Since the intervals $]a, b]$ form a basis for the upper limit topology [11, Chapter III, Section 3, Ex. 4], it follows from Theorem 4.13 in [11, Chapter III] that

$$(]a, b] \cap \mathcal{S}') \cap \mathcal{D}' \neq \varnothing \text{ for all } a, b \in \mathbb{R} \text{ such that } a < b,$$

and therefore

$$]s - \delta, s] \cap \mathcal{D}' \neq \varnothing \text{ for all } s \in \mathcal{S}', \delta \in \mathbb{R}_{>0}.$$

Due to symmetry, this implies that any subspace $\mathcal{S}$ of the set of real numbers equipped with the lower limit topology has a countable dense subset $\mathcal{D}$ as well, with

$$[s, s + \delta[\cap \mathcal{D} \neq \varnothing \text{ for all } s \in \mathcal{S}, \delta \in \mathbb{R}_{>0}.$$

This also implies that for any $s \in \mathcal{S} \setminus \mathcal{D}$, $]s, s + \delta[\cap \mathcal{D} \neq \varnothing$ for all $\delta \in \mathbb{R}_{>0}$, and therefore $s \in \mathrm{rlims}(\mathcal{D})$. $\qquad\square$

This is already promising, but we have left one big question unanswered: is the σ-algebra $\sigma(\mathring{\mathscr{C}})$ generated by the cylinder events sufficiently rich, meaning that it contains 'all' events that we are interested in? Due to Lemmas 4, 7 and 8, we would expect that the answer to this question is yes, but since it is hard to formalise what precisely 'all events of interest' are, we cannot answer this question definitively. We can, however, easily show that the issue we observed in Example 1 is now resolved.

Example 2. Let us return to the setting of Example 1. Note that the set of positive rationals $\mathbb{Q}_{\geq 0}$ is a countable subset of $\mathbb{T} = \mathbb{R}_{\geq 0} = \mathcal{T}$ that is dense with respect to the lower limit topology. For the set of càdlàg paths, the event of interest is

$$\mathring{A} \coloneqq \bigcup_{t \in [0, T]} \{\mathring{X}_t = x\}.$$

Fix any càdlàg path $\mathring{\omega} \in \mathring{\Omega}$. Then $\mathring{\omega} \in \mathring{A}$ if and only if there is some $t \in [0, T]$ such that $\mathring{\omega}(t) = x$; since $\mathring{\omega}$ is right-continuous in $\mathbb{T} \cap \mathrm{rlims}(\mathbb{T}) = \mathbb{R}_{\geq 0}$, either this is the case for $t = T$ or for some rational time point $t \in [0, T[\cap \mathbb{Q}_{\geq 0}$. Hence,

$$\mathring{A} = \bigcup_{t \in \mathcal{C}} \{\mathring{X}_t = x\} \quad \text{with } \mathcal{C} \coloneqq \{T\} \cup ([0, T[\cap \mathbb{Q}_{\geq 0}).$$

Since $\mathcal{C}$ is clearly countable and the σ-algebra generated by $\mathring{\mathscr{C}}$ is closed under countable unions, we infer from this equality that $\mathring{A} \in \sigma(\mathring{\mathscr{C}})$.

So it seems like $\sigma(\mathring{\mathscr{C}})$ would be a suitable domain, but can we also obtain a probability measure on $\sigma(\mathring{\mathscr{C}})$? To that end, we need to show that the probability charge on $\mathring{\mathscr{C}}$ induced by a consistent collection of finite-dimensional distributions is countably additive. We will see that this is the case if and only if an extra regularity condition is satisfied. Before we can introduce this condition in Section 4.3 further on, we need to introduce the notion of (the expected number of) jumps.

4.2. **The number of jumps.** A crucial property of càdlàg paths is that for all $s, r \in \mathbb{T}$ such that $s \leq r$, the number of 'discontinuities' or 'jumps' in $[s, r] \cap \mathbb{T}$ is finite. For specific choices of $\mathbb{T}$ – such as $\mathbb{R}$ or $\mathbb{R}_{\geq 0}$ – this property is well known – see for example Lemma 5.20 in [12] or Lemma 1 in [4, Section 12] for related results – and a rather straightforward consequence of the definition of càdlàg paths and the Heine-Borel Theorem. We will establish this result for general $\mathbb{T}$ though. But let us begin by explaining what we mean with 'jumps'.

Fix $s, r \in \mathbb{R}$ such that $s \leq r$ and a càdlàg path $\mathring{\omega} \in \mathring{\Omega}$. First, we assume that $[s, r] \subseteq \mathbb{T}$. Then the number of jumps in $[s, r] \cap \mathbb{T}$ of $\mathring{\omega}$ is

$$\left| \left\{ t \in \,]s, r] \colon \lim_{\Delta \searrow 0} \mathring{\omega}(t - \Delta) \neq \mathring{\omega}(t) \right\} \right|,$$

and it follows fairly easily from the properties of càdlàg paths and the Heine-Borel Theorem – or Billingsley's [4, Section 12, Lemma 1] argument – that this number is finite. Therefore, it is not difficult to verify that

$$(3) \qquad \left| \left\{ t \in \,]s, r] \colon \lim_{\Delta \searrow 0} \mathring{\omega}(t - \Delta) \neq \mathring{\omega}(t) \right\} \right| = \sup \{ \hat{\eta}_u(\mathring{\omega}(u)) \colon u \in \mathcal{U}_{[s,r] \cap \mathbb{T}} \},$$

where for any tuple of time points $u = (t_1, \dots, t_n) \in \mathcal{U}_{\mathbb{T}}$, we let

$$\hat{\eta}_u \colon \mathcal{X}_u \to \{0, \dots, n-1\} \colon x_u \mapsto |k \in \{2, \dots, n\} \colon x_{t_{k-1}} \neq x_{t_k}|.$$

If $[s, r]$ is not included in $\mathbb{T}$, we cannot define the number of jumps in $[s, r]$ in the same way. Instead, we generalise the right-hand side of Eqn. (3), but with the sufficient generality we need in the remainder: for all subsets $\mathcal{S}$ of $\mathbb{T}$,

$$\eta_{\mathcal{S}}(\mathring{\omega}) := \sup \{ \hat{\eta}_u(\mathring{\omega}(u)) \colon u \in \mathcal{U}_{\mathcal{S}} \} \text{ for all } \mathring{\omega} \in \mathring{\Omega}.$$

In any case, the number of jumps of a *càdlàg* path in $[s, r]$ is always finite; because the proof of this (essentially well-known) result is not very instructive, we have relegated it to Appendix B.

Proposition 9. *For all* $\mathring{\omega} \in \mathring{\Omega}$ *and* $s, r \in \mathbb{R}$ *such that* $s < r$ *and* $[s, r] \cap \mathbb{T} \neq \varnothing$, $\eta_{[s,r] \cap \mathbb{T}}(\mathring{\omega}) < +\infty$.

Finally, we will also consider the expected number of jumps. To this end, we observe that for all $u = (t_1, \dots, t_m) \in \mathcal{U}$, $\hat{\eta}_u$ is a $\wp(\mathcal{X}_u)$-simple variable[4] because

$$\hat{\eta}_u = \sum_{k=1}^{m-1} k \mathbb{I}_{\{x_u \in \mathcal{X}_u \colon \hat{\eta}_u(x_u) = k\}} = \sum_{k=2}^{m} \mathbb{I}_{\{x_u \in \mathcal{X}_u \colon x_{t_k} \neq x_{t_{k-1}}\}};$$

hence, for any probability charge μ on $\wp(\mathcal{X}_u)$, its expectation with respect to μ is well-defined through the Dunford integral:[5]
(4)
$$E_\mu(\hat{\eta}_u) = \sum_{k=1}^{m-1} k\mu(\{x_u \in \mathcal{X}_u \colon \hat{\eta}_u(x_u) = k\}) = \sum_{k=2}^{m} \mu_u\big(\{x_u \in \mathcal{X}_u \colon x_{t_k} \neq x_{t_{k-1}}\}\big).$$

If $\mu_\bullet$ is a consistent collection of finite-dimensional charges, then for all $u = (t_1, \dots, t_m) \in \mathcal{U}$, the final expression in Eqn. (4) simplifies as follows:

$$(5) \qquad E_{\mu_u}(\hat{\eta}_u) = \sum_{k=2}^{m} \mu_u\big(\{x_u \in \mathcal{X}_u \colon x_{t_k} \neq x_{t_{k-1}}\}\big) = \sum_{k=2}^{m} \mu_{(t_{k-1}, t_k)}\big(\mathcal{X}_{\neq}^2\big),$$

where here and in the remainder, we let $\mathcal{X}_{\neq}^2 := \big\{(x, y) \in \mathcal{X}^2 \colon x \neq y\big\}$ and $\mathcal{X}_{=}^2 := \big\{(x, y) \in \mathcal{X}^2 \colon x = y\big\}$, and for all $s, r \in \mathcal{T}$ such that $s < r$, with some slight

---

[4]We adhere to the definition given by Troffaes & De Cooman [27, Definition 1.16], but see also Definition 4.2.12 in [2].

[5]We adhere to the definition in [27, Definition 8.13], see also Definition 4.4.1 in [2].

abuse of notation, we also write $\mathcal{X}_{\neq}^2$ and $\mathcal{X}_{=}^2$ when we actually mean $\{(x_s, x_r) \in \mathcal{X}_{(x_s, x_r)} : x_s \neq x_r\}$ and $\{(x_s, x_r) \in \mathcal{X}_{(x_s, x_r)} : x_s = x_r\}$, respectively.

4.3. **Regularity.** Finally, we can get down to proving that the following two regularity conditions on a consistent collection of finite-dimensional distributions are necessary and sufficient for the corresponding probability charge on $\mathring{\mathscr{C}}$ to be countably additive.

Definition 4. A collection of finite-dimensional charges $\mu_\bullet$ is *regular* if

R1. for all $t \in \mathrm{rlims}(\mathcal{T}) \cap \mathcal{T}$,

$$\lim_{\mathcal{T} \ni r \searrow t} \mu_{(t,r)}\big(\mathcal{X}_{=}^2\big) = 1;$$

R2. for all $n \in \mathbb{N}$ such that $[-n, n] \cap \mathcal{T} \neq \varnothing$,

$$\lim_{k \to +\infty} \sup\big\{\mu_u\big(\{x_u \in \mathcal{X}_u : \hat{\eta}_u(x_u) \geq k\}\big) : u \in \mathcal{U}_{[-n,n] \cap \mathcal{T}}\big\} = 0.$$

The two conditions for regularity basically ensure that the collection of finite-dimensional charges is compatible with the properties of càdlàg paths. The first condition (R1) mirrors the continuity from the right: it demands that the finite-dimensional charges are (stochastically) continuous from the right – similar to the notions in, for example, [6, Definition 12.15] or [5, Definition 18.2.2]. It may be less obvious on first inspection, but the second condition (R2) mirrors the fact that càdlàg paths have a finite number of jumps in any closed and bounded interval. While Proposition 9 says that this should be the case for all intervals $[s, r]$, it clearly suffices to limit ourselves to intervals of the form $[s, r] = [-n, n]$ with $n \in \mathbb{N}$ because the number of jumps is monotone.

That regularity is indeed necessary and sufficient for a consistent collection of finite-dimensional distributions to have a corresponding probability charge P on $\mathring{\mathscr{C}}$ that is countably additive, is established by our next result. Our proof is a bit long and needs quite a bit more technical machinery; for this reason, we have relegated it to Appendix C.

Theorem 10. *Consider a consistent collection $\mu_\bullet$ of finite-dimensional distributions, and let P be the corresponding probability charge on $\mathring{\mathscr{C}}$ of Proposition 1. Then P is countably additive if an only if $\mu_\bullet$ is regular.*

We now use Theorem 10 to prove our main result: a version of the Daniell–Kolmogorov Extension Theorem where the set of all paths is replaced by the set of all càdlàg paths.

Theorem 11. *Consider a consistent collection of finitary distributions $\mu_\bullet$. Then there is a probability measure P on $\sigma(\mathring{\mathscr{C}})$ such that*

$$P(\{\mathring{X}_u \in A\}) = \mu_u(A) \quad \text{for all } u \in \mathcal{U}, A \in \wp(\mathcal{X}_u)$$

if and only if $\mu_\bullet$ is regular; if this is the case, then this probability measure is unique.

Proof. Follows immediately from Proposition 1, Theorem 10 and Caratheodory's Extension Theorem. $\square$

That we need the two conditions (R1) and (R2) for Theorem 11 is in line with Borovkov's [5] findings for $\mathbb{R}$-valued processes: right after his Definition 18.2.3, he explains that stochastic continuity (so his version of (R1)) alone does not suffice, and that 'in order to characterise the properties of trajectories, one needs quantitative bounds for [the magnitude of the jumps]'. Note that in our setting of countable state spaces, we do *not* need a quantitative bound on the 'magnitude' of the jumps but on the *number* of jumps.

The two conditions (R1) and (R2) for regularity are perhaps not the most easy ones to check. Fortunately, there are plenty of sufficient conditions for regularity that can be more easily verified. For example, the one in the following result comes in handy in the setting of Poisson processes and Markovian (imprecise) jump processes.

Proposition 12. *Consider a consistent collection $\mu_\bullet$ of finite-dimensional distributions. If for all $n \in \mathbb{N}$ with $[-n, n] \cap \mathcal{T} \neq \varnothing$ there is some $\lambda_n \in \mathbb{N}$ such that*

$$E_{\mu_u}(\hat{\eta}_u) \leq \lambda_n(t_m - t_1) \quad \text{for all } u = (t_1, \ldots, t_m) \in \mathcal{U}_{[-n,n] \cap \mathcal{T}},$$

then $\mu_\bullet$ is regular.

Proof. To verify (R1), we fix any $t \in \mathcal{T} \cap \mathrm{rlims}(\mathcal{T})$. Let n be any natural number such that $|t| < n$. Then $[-n, n] \cap \mathcal{T} \neq \varnothing$ by construction and $]t, n] \cap \mathcal{T} \neq \varnothing$ because t is a right-sided limit point of $\mathcal{T}$. For all $r \in]t, n] \cap \mathcal{T}$ and with $u := (t, r)$, $\{x_u \in \mathcal{X}_u \colon \hat{\eta}_u(x_u) = 1\} = \mathcal{X}_{\neq}^2$, and it follows from this and the condition in the statement that

$$\mu_u(\mathcal{X}_{\neq}^2) = \mu_u(\{x_u \in \mathcal{X}_u \colon \hat{\eta}_u(x_u) = 1\}) = E_{\mu_u}(\hat{\eta}_u) \leq \lambda_n(r - t).$$

From this inequality, we infer that

$$\lim_{\mathcal{T} \ni r \searrow t} \mu_{(t,r)}(\mathcal{X}_{=}^2) = \lim_{\mathcal{T} \ni r \searrow t} 1 - \mu_{(t,r)}(\mathcal{X}_{\neq}^2) = 1,$$

as required for (R1).

Checking (R2) is straightforward. Fix some $n \in \mathbb{N}$ with $[-n, n] \cap \mathcal{T} \neq \varnothing$. For all $k \in \mathbb{N}$ and $u = (t_1, \ldots, t_m) \in \mathcal{U}_{[-n,n] \cap \mathcal{T}}$, it follows from Markov's inequality and the condition in the statement that

$$\mu_u(\{x_u \in \mathcal{X}_u \colon \hat{\eta}_u(x_u) \geq k\}) \leq \frac{E_{\mu_u}(\hat{\eta}_u)}{k} \leq \frac{\lambda_n(t_m - t_1)}{k} \leq \frac{2n\lambda_n}{k}.$$

From this, we infer that

$$\lim_{k \to +\infty} \sup\{\mu_u(\{x_u \in \mathcal{X}_u \colon \hat{\eta}_u(x_u) \geq k\}) \colon u \in \mathcal{U}_{[-n,n] \cap \mathcal{T}}\} = 0,$$

as required. □

The consistent collection of finite-dimensional distributions $\mu_\bullet$ for the Poisson process with rate $\lambda \in \mathbb{R}_{\geq 0}$, as introduced at the end of Section 2.2, satisfies the sufficient condition in Proposition 12: it is not difficult to verify that for all $u = (t_1, \ldots, t_m) \in \mathcal{U}$,

$$E_{\mu_u}(\hat{\eta}_u) = \sum_{k=2}^{m} \lambda(t_k - t_{k-1}) = \lambda(t_m - t_1).$$

The same is true for the collections of finite-dimensional distributions for finite-state Markov processes whose semi-group of transition matrices is generated by a (transition) rate matrix. Even more, this sufficient condition also suffices in the more general setting of Markovian imprecise jump processes with a finite state space [20, 13]. In that setting, one considers sets of consistent collections of finite-dimensional distributions that are 'consistent' – in the sense of [20, Definition 6.1] or [12, Definition 3.50] – with a non-empty and bounded set $\mathcal{Q}$ of (transition) rate matrices, and Erreygers [12, Corollary 5.18 and Theorem 5.27] proves that for each of these collections of finite-dimensional distributions,

$$E_{\mu_u}(\hat{\eta}_u) \leq \frac{1}{2}\|\mathcal{Q}\|(t_m - t_1) < +\infty \quad \text{for all } u = (t_1, \ldots, t_m) \in \mathcal{U},$$

where $\|\mathcal{Q}\|$ is the supremum over the norms of the rate matrices in the bounded set $\mathcal{Q}$.

Our work on Markovian imprecise jump processes has also inspired us to establish two additional sufficient conditions for regularity. The main difference with the condition in Proposition 12 is that these conditions only concern the finite-dimensional distributions for two time points; they can be thought of as placing bounds on the dynamics of the process. Because these results are of a more technical nature, we have relegated them to Appendix D. Here, we only give a simplified version of Proposition 24.

Corollary 13. *Suppose that $\mathcal{T} = \mathbb{R}_{\geq 0}$, and consider a consistent collection $\mu_\bullet$ of finite-dimensional distributions. If for all $n \in \mathbb{N}$ there is some $\lambda_n \in \mathbb{R}_{\geq 0}$ such that*

$$\limsup_{r \searrow t} \frac{\mu_{(t,r)}(\mathcal{X}_{\neq}^2)}{r - t} \leq \lambda_n \quad \text{for all } t \in [0, n[$$

and

$$\limsup_{s \nearrow t} \frac{\mu_{(s,t)}(\mathcal{X}_{\neq}^2)}{t - s} \leq \lambda_n \quad \text{for all } t \in]0, n],$$

then $\mu_\bullet$ is regular.

Note that Corollary 13 is also relevant when dealing with the Poisson process: with $\mu_\bullet$ the consistent collection of finite-dimensional distributions for the Poisson process with rate λ,

$$\lim_{r \searrow t} \frac{\mu_{(t,r)}(\mathcal{X}_{\neq}^2)}{r - t} = \lim_{r \searrow t} \frac{1 - e^{-\lambda(r - t)}}{r - t} = \lambda \quad \text{for all } t \in \mathbb{R}_{\geq 0}$$

and

$$\lim_{s \nearrow t} \frac{\mu_{(s,t)}(\mathcal{X}_{\neq}^2)}{t - s} = \lim_{s \nearrow t} \frac{1 - e^{-\lambda(t - s)}}{t - s} = \lambda \quad \text{for all } t \in \mathbb{R}_{>0}$$

5. Conclusion

Our version of the Daniell–Kolmogorov Extension Theorem (Theorem 11) does not suffer from the issue with the original result (Theorem 3): by using the càdlàg paths as sample space, we obtain a probability measure whose domain is indeed sufficiently rich, in that it contains most – if not all – of the events that one can be interested in. Moreover, our results considers stochastic processes indexed by any subset $\mathbb{T}$ of the real numbers and looked at events that depend only on the time points in some subset $\mathcal{T}$ of $\mathbb{T}$, while it is standard to assume $\mathcal{T} = \mathbb{T}$; this distinction between $\mathbb{T}$ and $\mathcal{T}$ may seem somewhat strange at first, but it was motivated by our work on Markovian imprecise jump processes, where this proved to be useful.

In contrast to the original version of the theorem, our new version is limited to the specific setting of countable-state stochastic processes due to the form of our regularity condition. For this reason, a first line of follow-up research we envision is to generalise our work to uncountable state spaces. Crucial is that one should find a meaningful alternative to the number of jumps; for example, with the set of real numbers, it is customary – in the setting of modifications, at least – to look at upcrossing numbers instead of the number of jumps. A second line of follow-up research could be to apply our results. We have already briefly touched on the relevance to the Poisson process, and we are convinced that our theorem can be useful for more general counting processes as well. Furthermore, we know from our work in the setting of Markovian imprecise jump processes that our version of the Daniell–Kolmogorov Extension Theorem can be of use in the setting of jump processes that need not be time-homogeneous nor even Markovian.

References

1. Christel Baier and Joost-Pieter Katoen, *Principles of model checking*, MIT Press, Cambridge, Massachusetts, 2008.
2. K. P. S. Bhaskara Rao and M. Bhaskara Rao, *Theory of charges*, Academic Press, London, 1983.
3. Patrick Billingsley, *Probability and measure*, 3rd ed., Wiley, New York, 1995.
4. _____, *Convergence of probability measures*, 2nd ed., Wiley, New York, 1999.
5. Alexandr A. Borovkov, *Probability theory*, Springer, London, 2013.
6. Leo Breiman, *Probability*, SIAM, Philadelphia, 1992.
7. Kai Lang Chung, *Markov chains with stationary transition probabilities*, Springer, Berlin, Heidelberg, 1960.
8. Samuel N. Cohen and Robert J. Elliott, *Stochastic calculus and applications*, Birkhäuser, New York, 2015.
9. Harald Cramér and M.R. Leadbetter, *Stationary and related stochastic processes*, Wiley, New York, 1967.
10. Joseph L. Doob, *Stochastic processes*, Wiley, New York, 1953.
11. James Dugundji, *Topology*, Allyn and Bacon, Boston, 1966.
12. Alexander Erreygers, *Markovian imprecise jump processes: Foundations, algorithms and applications*, Ph.D. thesis, Ghent University, 2021, Available at https://users.ugent.be/~aerreyge/.
13. Alexander Erreygers and Jasper De Bock, *Markovian imprecise jump processes: Extension to measurable variables, convergence theorems and algorithms*, International Journal of Approximate Reasoning **147** (2022), 78–124.
14. Bert E. Fristedt and Lawrence F. Gray, *A modern approach to probability theory*, Birkhäuser Boston, Massachusetts, 1997.
15. Iosif I. Gikhman and Anatoly V. Skorokhod, *Introduction to the theory of random processes*, W.B. Saunders Company, Philadelphia, 1969.
16. Marius Iosifescu, *Finite Markov processes and their applications*, Wiley, Chichester, 1980.
17. Olav Kallenberg, *Foundations of modern probability*, 2nd ed., Springer, New York, 2002.
18. Andrei N. Kolmogorov, *Grundbegriffe der wahrscheinlichkeitsrechnung*, Springer, Berlin, 1933, Translated to English as [19].
19. _____, *Foundations of the theory of probability*, Chelsea, New York, 1950, English translation of [18].
20. Thomas Krak, Jasper De Bock, and Arno Siebes, *Imprecise continuous-time Markov chains*, International Journal of Approximate Reasoning **88** (2017), 452–528.
21. Heinz König, *Stochastic processes on the basis of new measure theory*, Proceedings of the Conference Positivity IV – Theory and Applications (Desden) (Martin R. Weber and Jürgen Voigt, eds.), Technische Universität Dresden, Institut für Analysis, 2006, Reprinted as Chapter 18 in [22], pp. 79–92.
22. _____, *Measure and integration*, Birkhäuser, Basel, 2012.
23. James R. Norris, *Markov chains*, Cambridge University Press, Cambridge, 1997.
24. L. Chris G. Rogers and David Williams, *Diffusions, Markov processes and martingales*, 2nd ed., vol. 1, Wiley, Chichester, 1994.
25. Eric Schechter, *Handbook of analysis and its foundations*, Academic Press, San Diego, 1997.
26. Glenn Shafer and Vladimir Vovk, *The Sources of Kolmogorov's Grundbegriffe*, Statistical Science **21** (2006), no. 1, 70–98.
27. Matthias C.M. Troffaes and Gert de Cooman, *Lower previsions*, Wiley, Chichester, 2014.

Appendix A. Proof for Theorem 2

In case $\mathcal{T} = \mathbb{T}$, Theorem 2 is implied by Theorem 31.1 in [24, Chapter II] – we leave it to the reader to check that $(\mathcal{X}, \wp(\mathcal{X}))$ is a Lusin space. We could also show that this in turn implies Theorem 2 in case $\mathcal{T} \subset \mathbb{T}$, but we choose to take a more direct route towards proving this result; our reasons for this are twofold: (i) it does not take a lot more work, and (ii) it would be a shame not to elucidate how the argument simplifies in our particular case where $\mathcal{X}$ is countable.

Our proof for Theorem 2 is essentially the argument used by Rogers & Williams [24, Chapter II] (who assume that $\mathcal{X}$ is a complete metric space) to prove their Lemma 30.7, but translated to the setting of a countable state space $\mathcal{X}$ – note that if $\mathcal{X}$ is countably infinite, then $(\mathcal{X}, \wp(\mathcal{X}))$ is not a complete metric space. The crucial

result that allows us to avoid the assumption that $\mathcal{X}$ is a commplete metric space, is the following intermediate lemma.

Lemma 14. *For all $n \in \mathbb{N}$, consider a finite subset B_n of $\mathcal{X}_{u_n}$, with $u_n \in \mathcal{U}$, and let $\dot{B}_n := \{\dot{X}_{u_n} \in B_n\}$. If $\bigcap_{n=1}^{m} \dot{B}_n \neq \varnothing$ for all $m \in \mathbb{N}$, then also $\bigcap_{n \in \mathbb{N}} \dot{B}_n \neq \varnothing$.*

Proof. We endow $\mathcal{X}$ with the cofinite topology [25, Section 5.15.c], the open sets of which are $\{\varnothing\} \cup \{A^{\mathrm{c}} \colon A \subseteq \mathcal{X}, |A| < +\infty\}$, and we endow $\dot{\Omega} = \mathcal{X}^{\mathbb{T}}$ with the corresponding product topology. Since any product of topological spaces, each equiped with the cofinite topology, is compact [25, Section 17.20 (AC25)], this turns $\dot{\Omega}$ into a compact space.

For all $x \in \mathcal{X}$ and $t \in \mathbb{T}$, since $\{x\}^{\mathrm{c}}$ belongs to the cofinite topology on $\mathcal{X}$, we know that $\{\dot{X}_t = x\}^{\mathrm{c}} = \{\dot{X}_t \in \{x\}^{\mathrm{c}}\} = \{\dot{\omega} \in \dot{\Omega} \colon \dot{\omega}(t) \in \{x\}^{\mathrm{c}}\}$ belongs to the product topology on $\dot{\Omega}$. For all $u = (t_1, \ldots, t_k) \in \mathcal{U}$ and finite $B \subseteq \mathcal{X}_u$, since

$$\{\dot{X}_u \in B\} = \bigcup_{x_u \in B} \{\dot{X}_u = x_u\} = \bigcup_{x_u \in B} \bigcap_{\ell=1}^{k} \{\dot{X}_{t_\ell} = x_{t_\ell}\},$$

this implies that

$$\{\dot{X}_u \in B\}^{\mathrm{c}} = \bigcap_{x_u \in B} \bigcup_{\ell=1}^{k} \{\dot{X}_{t_\ell} = x_{t_\ell}\}^{\mathrm{c}}$$

belongs to the product topology as well. Hence, for all $n \in \mathbb{N}$, $\dot{B}_n = \{\dot{X}_{u_n} \in B_n\}$ is a closed subset of $\dot{\Omega}$.

Let us now assume that $\bigcap_{n=1}^{m} \dot{B}_n \neq \varnothing$ for all $m \in \mathbb{N}$. Then clearly, for any finite $N \subseteq \mathbb{N}$, we have that $\bigcap_{n \in N} \dot{B}_n \neq \varnothing$. The compactness of $\dot{\Omega}$ therefore implies that $\bigcap_{n \in \mathbb{N}} \dot{B}_n \neq \varnothing$ – see [25, Definition 17.2.(B)]. $\qquad\square$

Our proof for Theorem 2 now follows more or less the same argument as the one for Lemma 30.7 in [24, Chapter II], but without the use of their Lemma 29.7.

Proof of Theorem 2. According to a classic result from measure theory – see for example [14, Chapter 7, Proposition 9] or [24, Lemma 4.3] – P is countably additive if and only if, for any decreasing sequence $(\dot{A}_n)_{n \in \mathbb{N}}$ in $\dot{\mathscr{C}}$ such that $\bigcap_{n \in \mathbb{N}} \dot{A}_n = \varnothing$, also $\lim_{n \to +\infty} P(\dot{A}_n) = 0$. We will prove this alternative condition instead.

By definition of $\dot{\mathscr{C}}$, there are, for all $n \in \mathbb{N}$, some $u_n \in \mathcal{U}$ and $A_n \in \mathcal{X}_{u_n}$ such that $\dot{A}_n = \{\dot{X}_{u_n} \in A_n\}$. Fix any $\epsilon \in \mathbb{R}_{>0}$. For all $n \in \mathbb{N}$, since μ_{u_n} is countably additive and $\mathcal{X}_{u_n}$ – and hence also A_n – is a countable set, we have that $\mu_{u_n}(A_n) = \sum_{x_{u_n} \in A_n} \mu_{u_n}(\{x_{u_n}\})$. Since every $\mu_{u_n}(\{x_{u_n}\})$ is positive, the convergence of this countable sum to $\mu_{u_n}(A_n)$ implies that there is a finite subset B_n of A_n such that $\mu_{u_n}(B_n) = \sum_{x_{u_n} \in B_n} \mu_{u_n}(\{x_{u_n}\}) > \mu_{u_n}(A_n) - 2^{-n}\epsilon$ and hence also $\mu_{u_n}(A_n \setminus B_n) < 2^{-n}\epsilon$. Consider any such B_n and let $\dot{B}_n := \{\dot{X}_{u_n} \in B_n\} \subseteq \dot{A}_n$.

Since $\bigcap_{n \in \mathbb{N}} \dot{A}_n = \varnothing$, we clearly also have that $\bigcap_{n \in \mathbb{N}} \dot{B}_n = \varnothing$. It therefore follows from Lemma 14 that there is some $m \in \mathbb{N}$ such that $\bigcap_{n=1}^{m} \dot{B}_n = \varnothing$. Then for any $\dot{\omega} \in \dot{A}_m$, since $\dot{\omega} \notin \bigcap_{n=1}^{m} \dot{B}_n$, there is some $1 \leq n_{\dot{\omega}} \leq m$ such that $\dot{\omega} \notin \dot{B}_{n_{\dot{\omega}}}$ and therefore $\dot{\omega} \in \dot{A}_m \setminus \dot{B}_{n_{\dot{\omega}}} \subseteq \dot{A}_{n_{\dot{\omega}}} \setminus \dot{B}_{n_{\dot{\omega}}} \subseteq \bigcup_{n=1}^{m} \dot{A}_n \setminus \dot{B}_n$. This implies that $\dot{A}_m \subseteq \bigcup_{n=1}^{m} \dot{A}_n \setminus \dot{B}_n$ and therefore, that

$$P(\dot{A}_m) \leq P\left(\bigcup_{n=1}^{m} \dot{A}_n \setminus \dot{B}_n\right) \leq \sum_{n=1}^{m} P(\dot{A}_n \setminus \dot{B}_n) = \sum_{n=1}^{m} P(\{\dot{X}_{u_n} \in A_n \setminus B_n\})$$

$$= \sum_{n=1}^{m} \mu_{u_n}(A_n \setminus B_n) < \sum_{n=1}^{m} 2^{-n}\epsilon < \epsilon.$$

Since $(\dot{A}_n)_{n\in\mathbb{N}}$ is a decreasing sequence, it follows that $\lim_{n\to+\infty} P(\dot{A}_n) < \epsilon$. As the choice of $\epsilon > 0$ was arbitrary, we conclude that $\lim_{n\to+\infty} P(\dot{A}_n) = 0$.   □

APPENDIX B. PROPERTIES OF CÀDLÀG PAHTS

In this appendix, we give the proof for the three important properties of càdlàg paths. The first property is that a càdlàg path always has a càdlàg extension to $\mathbb{R}$.

Proof of Lemma 6. Fix some state $x \in \mathcal{X}$, and for all $t \in \mathbb{R}$, let $\mathcal{S}_{\leq t} :=]-\infty, t] \cap \mathcal{S}$ and $s_t^\star := \sup \mathcal{S}_{\leq t}$. Then

$$\psi \colon \mathbb{R} \to \mathcal{X} \colon t \mapsto \begin{cases} \varpi(t) & \text{if } t \in \mathcal{S} \\ \lim_{\mathcal{S} \ni r \searrow t} \varpi(r) & \text{if } t \notin \mathcal{S}, t \in \mathrm{rlims}(\mathcal{S}) \\ \varpi(s_t^\star) & \text{if } t \notin \mathcal{S}, t \notin \mathrm{rlims}(\mathcal{S}), \mathcal{S}_{\leq t} \neq \varnothing, s_t^\star \in \mathcal{S} \\ \lim_{\mathcal{S} \ni s \nearrow s_t^\star} \varpi(s) & \text{if } t \notin \mathcal{S}, t \notin \mathrm{rlims}(\mathcal{S}), \mathcal{S}_{\leq t} \neq \varnothing, s_t^\star \notin \mathcal{S} \\ x & \text{if } t \notin \mathcal{S}, t \notin \mathrm{rlims}(\mathcal{S}), \mathcal{S}_{\leq t} = \varnothing \end{cases}$$

clearly coincides with ϖ on $\mathcal{S}$. Less obvious is that ψ (i) has a left-sided limit everywhere and (ii) is continuous from the right.

For the first condition, we fix any $t \in \mathbb{R}$, and distinguish two cases. The first case is that t is not a left-sided limit point of $\mathcal{S}$. Then there is some $\delta \in \mathbb{R}_{>0}$ such that $]t-\delta, t[\cap \mathcal{S} = \varnothing$. Then for all $s \in]t-\delta, t[$, $s \notin \mathcal{S} \cup \mathrm{rlims}(\mathcal{S})$ – because with $\delta' := t - s > 0$, $[s, s+\delta'[\cap \mathcal{S} \subseteq]t-\delta, t[\cap \mathcal{S} = \varnothing$ – and $\mathcal{S}_{\leq s} = \mathcal{S}_{\leq t-\delta}$ – because $]t-\delta, s] \cap \mathcal{S} \subseteq]t-\delta, t[\cap \mathcal{S} = \varnothing$. It follows from this and the definition of ψ that for all $s \in]t-\delta, t[$,

$$\psi(s) = \begin{cases} \varpi(s_{t-\delta}^\star) & \text{if } \mathcal{S}_{\leq t-\delta} \neq \varnothing, s_{t-\delta}^\star \in \mathcal{S} \\ \lim_{\mathcal{S} \ni r \nearrow s_{t-\delta}^\star} \varpi(r) & \text{if } \mathcal{S}_{\leq t-\delta} \neq \varnothing, s_{t-\delta}^\star \notin \mathcal{S}, \\ x & \text{otherwise.} \end{cases}$$

From this, we conclude that ψ is constant on $]t-\delta, t[$.

The second case is that t is a left-sided limit point of $\mathcal{S}$. Because ϖ is càdlàg, there are some $\delta' \in \mathbb{R}_{>0}$ and $x_t \in \mathcal{X}$ such that $\varpi(s) = x_t$ for all $s \in]t-\delta', t[\cap \mathcal{S} \neq \varnothing$. Take any $s' \in]t-\delta', t[\cap \mathcal{S}$. Then for all $s \in]s', t[\cap \mathcal{S}$, $\varpi(s) = \varpi(s') = x_t$ and $s' \leq s_s^\star < t$. From this, it follows that for all $s \in]s', t[$, (i) $\varpi(s) = x_t$ if $s \in \mathcal{S}$; (ii) $\lim_{\mathcal{S} \ni r \searrow s} \varpi(r) = x_t$ if $s \in \mathrm{rlims}(\mathcal{S}) \setminus \mathcal{S}$; (iii) $\varpi(s_s^\star) = x_t$ if $s \notin \mathcal{S} \cup \mathrm{rlims}(\mathcal{S})$ and $s_s^\star \in \mathcal{S}$; and (iv) $\lim_{\mathcal{S} \ni r \nearrow s_s^\star} \varpi(r) = x_t$ if $s \notin \mathcal{S} \cup \mathrm{rlims}(\mathcal{S})$ and $s_s^\star \notin \mathcal{S}$ – the case $\mathcal{S}_{\leq s} = \varnothing$ is clearly impossible because $s' \in \mathcal{S}_{\leq s}$. We conclude from this and the definition of ψ that $\psi(s) = x_t$ for all $s \in]s', t[$. In both cases, this shows that ψ has a left-sided limit in t.

For the second condition, we fix any $t \in \mathbb{R}$. Here too, we distinguish two cases. The first case is that t is not a right-sided limit point of $\mathcal{S}$, meaning that there is some $\delta \in \mathbb{R}_{>0}$ such that $]t, t+\delta[\cap \mathcal{S} = \varnothing$. Then for all $r \in]t, t+\delta[$, $r \notin \mathcal{S} \cup \mathrm{rlims}(\mathcal{S})$ – because with $\delta' := t+\delta-r > 0$, $[r, r+\delta'[\cap \mathcal{S} \subseteq]t, t+\delta[\cap \mathcal{S} = \varnothing$ – and $\mathcal{S}_{\leq r} = \mathcal{S}_{\leq t}$ – because $]t, r] \cap \mathcal{S} \subseteq]t, t+\delta[\cap \mathcal{S} = \varnothing$. We distinguish two subcases. On the one hand, if $t \in \mathcal{S}$ then it follows from this and our definition of ψ that (i) $\psi(t) = \varpi(t)$ and (ii) for all $r \in]t, t+\delta[$, $s_r^\star = \sup \mathcal{S}_{\leq r} = \sup \mathcal{S}_{\leq t} = t \in \mathcal{S}$ and therefore $\psi(r) = \varpi(s_r^\star) = \varpi(t)$. If on the other hand $t \notin \mathcal{S}$ and therefore $t \notin \mathcal{S} \cup \mathrm{rlims}(\mathcal{S})$, then it follows from this and the definition of ψ that for all $r \in]t, t+\delta[$, $\psi(r) = \psi(t)$.

This leaves the case that t is a right-sided limit point of $\mathcal{S}$. Because ϖ is càdlàg, there are some $\delta \in \mathbb{R}_{>0}$ and x_t such that $\varpi(r) = x_t$ for all $r \in [t, t+\delta[\cap \mathcal{S} \neq \varnothing$. Here too, we distinguish two subcases. If $t \in \mathcal{S}$, then we infer from this that (i) $\psi(t) = \varpi(t) = x_t$; (ii) for all $r \in]t, t+\delta[\cap (\mathcal{S} \cup \mathrm{rlims}(\mathcal{S}))$, $\psi(r) = x_t$; and (iii) for all $r \in]t, t+\delta[\setminus (\mathcal{S} \cup \mathrm{rlims}(\mathcal{S}))$, $\mathcal{S}_{\leq r} \neq \varnothing$ and $s_r^\star \in]t, t+\delta[$, and therefore $\psi(r) = x_t$. If on the other hand $t \notin \mathcal{S}$, then we infer from this that (i) $\psi(t) =$

$\lim_{\mathcal{S}\ni s\searrow t}\varpi(s) = x_t$; (ii) for all $r \in \,]t,t+\delta[\,\cap (\mathcal{S} \cup \mathrm{rlims}(\mathcal{S}))$, $\psi(r) = x_t$; and (iii) for all $r \in \,]t,t+\delta[\,\setminus (\mathcal{S} \cup \mathrm{rlims}(\mathcal{S}))$, $\,]t,r[\,\cap \mathcal{S} \neq \varnothing$ – as t is a right-sided limit point of $\mathcal{S}$ – and therefore $\mathcal{S}_{\leq r} \neq \varnothing$ and $s_r^\star \in \,]t,t+\delta[$, whence

$$\psi(r) = \begin{cases} \varpi(s_r^\star) = x_t & \text{if } s_r^\star \in \mathcal{S}, \\ \lim_{\mathcal{S}\ni s\nearrow s_r^\star}\varpi(s) = x_t & \text{if } s_r^\star \notin \mathcal{S}. \end{cases}$$

It is clear in both cases that ψ is continuous from the right in t, as required. $\square$

The second property we need to prove is Lemma 7.

Proof of Lemma 7. The direct implication is trivial because $\mathcal{D} \subseteq \mathbb{T}$, so it remains for us to show the converse implication. To this end, we assume that $\mathring{\omega}_1(d) = \mathring{\omega}_2(d)$ for all $d \in \mathcal{D}$, and set out to show that then $\mathring{\omega}_1(t) = \mathring{\omega}_2(t)$ for all $t \in \mathbb{T}$. Fix any time point t in $\mathbb{T}$. If $t \in \mathcal{D}$, then $\mathring{\omega}_1(t) = \mathring{\omega}_2(t)$ by assumption, as required. If $t \notin \mathcal{D}$, then $t \in \mathrm{rlims}(\mathcal{D})$ because $\mathbb{T} \subseteq \mathcal{D} \cup \mathrm{rlims}(\mathcal{D})$ by assumption. Because $\mathring{\omega}_1$ and $\mathring{\omega}_2$ are both càdlàg, there are some $\delta_1, \delta_2 \in \mathbb{R}_{>0}$ such that

$$\big(\forall r \in \mathbb{T} \cap \,]t,t+\delta_1[\big)\ \mathring{\omega}_1(r) = \mathring{\omega}_1(t) \quad \text{and} \quad \big(\forall r \in \mathbb{T} \cap \,]t,t+\delta_2[\big)\ \mathring{\omega}_2(r) = \mathring{\omega}_2(t).$$

Let $\delta := \min\{\delta_1, \delta_2\}$. Fix any $d \in \,]t,t+\delta[\,\cap \mathcal{D}$ – this is always possible because t is a right-sided limit point of $\mathcal{D}$, so $\,]t,t+\delta[\,\cap \mathcal{D} \neq \varnothing$. On the one hand, it follows from the preceding that $\mathring{\omega}_1(t) = \mathring{\omega}_1(d)$ and $\mathring{\omega}_2(t) = \mathring{\omega}_2(d)$. On the other hand, we know that $\mathring{\omega}_1(d) = \mathring{\omega}_2(d)$ because $d \in \mathcal{D}$. From this, we infer that $\mathring{\omega}_1(t) = \mathring{\omega}_2(t)$. Since this is the case for all $t \in \mathbb{T}$, we have shown that $\mathring{\omega}_1 = \mathring{\omega}_2$, as required. $\square$

Finally, we prove that a càdlàg path can only jump a finite number of times in any closed and bounded interval.

Proof of Proposition 9. By Lemma 6, there is some càdlàg map $\psi\colon \mathbb{R} \to \mathcal{X}$ that coincides with $\mathring{\omega}$ on $\mathbb{T}$. As we have explained in Section 4.2, it is not difficult to show that

$$\sup\big\{\hat{\eta}_u(\psi(u))\colon u \in \mathcal{U}_{[s,r]}\big\} = \Big|\Big\{t \in \,]s,r]\colon \lim_{\Delta\searrow 0}\psi(t-\Delta) \neq \psi(t)\Big\}\Big| < +\infty,$$

where for all $u = (t_1,\ldots,t_n) \in \mathcal{U}_{[s,r]}$, we let $\psi(u) := (\psi(t_1),\ldots,\psi(t_n))$. Since clearly

$$\eta_{[s,r]\cap\mathbb{T}}(\mathring{\omega}) = \sup\big\{\hat{\eta}_u(\psi(u))\colon u \in \mathcal{U}_{[s,r]\cap\mathbb{T}}\big\} \leq \sup\big\{\hat{\eta}_u(\psi(u))\colon u \in \mathcal{U}_{[s,r]}\big\},$$

this proves the statement. $\square$

Appendix C. Proof for Theorem 10

In this appendix, we prove Theorem 10. We will get to this in Appendix C.4, but first we need to introduce some additional technical machinery. In Appendix C.1, we take a second look at the expected number of jumps, but this time in the setting of probability measures and stochastic processes. This theme continues in Appendix C.2, where we prove that whenever a (not necessarily càdlàg) path ω has a finite number of jumps (for some countable subset of $\mathcal{T}$), we can always 'modify' this path in such a way that it becomes càdlàg. Appendix C.3 introduces the two remaining intermediary results that we will need to prove Theorem 10.

C.1. **The expected number of jumps, again.** Consider again a subset Ω of $\dot{\Omega}$ that satisfies Eqn. (1). Then for any tuple of time points $u = (t_1, \ldots, t_n) \in \mathcal{U}$, the 'number of jumps of $X_\bullet$ along u' is the functional composition of $\hat{\eta}_u$ after the projection X_u:

$$\hat{\eta}_u(X_u) \colon \Omega \to \{0, \ldots, n-1\} \colon \omega \mapsto \hat{\eta}_u(\omega(u)).$$

It is easy to see that

$$(6) \qquad\qquad \hat{\eta}_u(X_u) = \sum_{k=2}^{n} \mathbb{I}_{\{X_{t_{k-1}} \neq X_{t_k}\}},$$

where for all $s, r \in \mathcal{T}$ such that $s < r$, we let

$$\{X_s \neq X_r\} \coloneqq \{X_{(s,r)} \in \mathcal{X}_{\neq}^2\} \in \mathscr{C} \quad \text{and} \quad \{X_s = X_r\} \coloneqq \{X_{(s,r)} \in \mathcal{X}_{=}^2\} \in \mathscr{C}.$$

This means that $\hat{\eta}_u(X_u)$ is a $\mathscr{C}$-simple and therefore trivially $\sigma(\mathscr{C})/\mathscr{B}(\overline{\mathbb{R}})$-measurable variable. Consequently, we also have that

$$\{\hat{\eta}_u(X_u) \leq \alpha\} \coloneqq \{\omega \in \Omega \colon \hat{\eta}_u(\omega(u)) \leq \alpha\} \in \mathscr{C} \quad \text{for all } \alpha \in \mathbb{R}.$$

Crucial to our proof of Theorem 10 is the number of jumps of $X_\bullet$ in $[s, r] \cap \mathcal{C}$, with $\mathcal{C}$ a countable subset of $\mathcal{T}$. For any countable subset $\mathcal{C}$ of $\mathcal{T}$ and $s, r \in \mathbb{R}$ such that $s < r$ and $[s, r] \cap \mathcal{D} \neq \varnothing$, the number of jumps of $X_\bullet$ in $[s, r] \cap \mathcal{D}$ is the variable

$$\eta_{[s,r] \cap \mathcal{D}}(X_\bullet) \colon \Omega \to \mathbb{Z}_{\geq 0} \cup \{+\infty\} \colon \omega \mapsto \eta_{[s,r] \cap \mathcal{D}}(\omega);$$

in other words, $\eta_{[s,r] \cap \mathcal{D}}(X_\bullet)$ is the functional composition of $\eta_{[s,r] \cap \mathcal{D}}$ and the 'projection' $X_\bullet$ which maps any $\omega \in \Omega$ to itself. This variable is $\sigma(\mathscr{C})/\mathscr{B}(\overline{\mathbb{R}})$-measurable.

Lemma 15. *Consider a countable subset $\mathcal{C}$ of $\mathcal{T}$ and some $s, r \in \mathbb{R}$ such that $s < r$ and $[s, r] \cap \mathcal{C} \neq \varnothing$. Then there is a sequence $(u_n)_{n \in \mathbb{N}} \in \mathcal{U}_{[s,r] \cap \mathcal{C}}$ such that $(\hat{\eta}_{u_n}(X_{u_n}))_{n \in \mathbb{N}}$ is an increasing sequence of positive $\mathscr{C}$-simple – and therefore $\sigma(\mathscr{C})/\mathscr{B}(\overline{\mathbb{R}})$-measurable – variables that converges point-wise to $\eta_{[s,r] \cap \mathcal{C}}(X_\bullet)$. Hence, $\eta_{[s,r] \cap \mathcal{C}}(X_\bullet)$ is $\sigma(\mathscr{C})/\mathscr{B}(\overline{\mathbb{R}})$-measurable.*

In our proof of Lemma 15, we make use of the straightforward observation that if $u \sqsubseteq v$, then the number of jumps of $X_\bullet$ along v is greater than or equal to the number of jumps of $X_\bullet$ along u; clearly, it suffices to state this property for paths.

Lemma 16. *For all $\dot{\omega} \in \dot{\Omega}$ and $u, v \in \mathcal{U}$ such that $u \sqsubseteq v$, $\hat{\eta}_u(\dot{\omega}(u)) \leq \hat{\eta}_v(\dot{\omega}(v))$.*

Proof. Let us write $u = (r_1, \ldots, r_n)$ and $v = (s_1, \ldots, s_m)$. If $n = 1$, then the inequality in the statement holds trivially because $\hat{\eta}_u = 0$ and $\hat{\eta}_v \geq 0$. Hence, we assume that $n > 1$. Because $u \sqsubseteq v$ by the conditions of the statement, for all $k \in \{1, \ldots, n\}$ there is a unique natural number $\ell_k \in \{1, \ldots, m\}$ such that $r_k = s_{\ell_k}$. Then

$$\hat{\eta}_u(\dot{\omega}(u)) = \big| \big\{ k \in \{2, \ldots, n\} \colon \dot{\omega}(s_{\ell_{k-1}}) \neq \dot{\omega}(s_{\ell_k}) \big\} \big|.$$

Because the tuples of time points u and v are increasing, $\ell_{k-1} < \ell_k$ for all $k \in \{2, \ldots, n\}$. Furthermore, it is easy to verify that for all $k \in \{2, \ldots, n\}$, $\dot{\omega}(s_{\ell_{k-1}}) \neq \dot{\omega}(s_{\ell_k})$ implies that $\dot{\omega}(s_{\ell-1}) \neq \dot{\omega}(s_\ell)$ for at least one $\ell \in \{\ell_{k-1} + 1, \ldots, \ell_k\}$ – if this is not the case, then $\dot{\omega}(s_{\ell_{k-1}}) = \dot{\omega}(s_{\ell_{k-1}+1}) = \cdots = \dot{\omega}(s_{\ell_k-1}) = \dot{\omega}(s_{\ell_k})$, which is a clear contradiction. Hence,

$$\big| \big\{ k \in \{2, \ldots, n\} \colon \dot{\omega}(s_{\ell_{k-1}}) \neq \dot{\omega}(s_{\ell_k}) \big\} \big| \leq \big| \big\{ \ell \in \{2, \ldots, m\} \colon \dot{\omega}(s_{\ell-1}) \neq \dot{\omega}(s_\ell) \big\} \big|,$$

and therefore

$$\hat{\eta}_u(\dot{\omega}(u)) \leq \hat{\eta}_v(\dot{\omega}(v)).$$

$\square$

Proof of Lemma 15. Because $\mathcal{C}$ is countable by assumption, so is $[s, r] \cap \mathcal{C}$. Consequently, there is a sequence $(u_n)_{n \in \mathbb{N}} \in \mathcal{U}_{[s,r]\cap\mathcal{C}}$ such that (i) $u_n \sqsubseteq u_{n+1}$ for all $n \in \mathbb{N}$ and (ii) for all $d \in [s, r] \cap \mathcal{C}$, there is some $n \in \mathbb{N}$ such that d belongs to u_n; for example, simply add all the time points in $[s, r] \cap \mathcal{C}$ one by one.

For all n in $\mathbb{N}$, we know from before that $\hat{\eta}_{u_n}(X_{u_n})$ is a positive $\mathscr{C}$-simple and therefore trivially $\sigma(\mathscr{C})/\mathscr{B}(\overline{\mathbb{R}})$-measurable variable, and because $u_n \sqsubseteq u_{n+1}$ by construction, $\hat{\eta}_{u_n}(X_{u_n}) \leq \hat{\eta}_{u_{n+1}}(X_{u_{n+1}})$ due to Lemma 16. Hence, we have verified that $(\hat{\eta}_{u_n}(X_{u_n}))_{n \in \mathbb{N}}$ is a sequence of positive $\sigma(\mathscr{C})/\mathscr{B}(\overline{\mathbb{R}})$-measurable variables that is monotonously increasing. Any monotonously increasing sequence converges point-wise, so it remains for us to show that this point-wise limit is $\eta_{[s,r]\cap\mathcal{C}}(X_\bullet)$. On the one hand, for all $n \in \mathbb{N}$, u_n belongs to $\mathcal{U}_{[s,r]\cap\mathcal{C}}$ so $\hat{\eta}_{u_n}(X_{u_n}) \leq \eta_{[s,r]\cap\mathcal{C}}(X_\bullet)$; hence,

$$\lim_{n \to +\infty} \hat{\eta}_{u_n}(X_{u_n}) \leq \eta_{[s,r]\cap\mathcal{C}}(X_\bullet).$$

On the other hand, for any tuple of time points $u \in \mathcal{U}_{\mathcal{C}}^{[s,r]}$ there is some $n \in \mathbb{N}$ such that $u \sqsubseteq u_n$, and therefore

$$\eta_{[s,r]\cap\mathcal{C}}(X_\bullet) \leq \lim_{n \to +\infty} \hat{\eta}_{u_n}(X_{u_n}).$$

From these two inequalities, we infer that $(\hat{\eta}_{u_n}(X_{u_n}))_{n \in \mathbb{N}}$ converges point-wise to $\eta_{[s,r]\cap\mathcal{C}}(X_\bullet)$. Since $(\hat{\eta}_{u_n}(X_{u_n}))_{n \in \mathbb{N}}$ is an increasing sequence of positive $\sigma(\mathscr{C})/\mathscr{B}(\overline{\mathbb{R}})$-measurable variables, this implies that $\eta_{[s,r]\cap\mathcal{C}}(X_\bullet)$ is $\sigma(\mathscr{C})/\mathscr{B}(\overline{\mathbb{R}})$-measurable [14, Chapter 2, Proposition 11 or Lemma 13]. □

We end this intermezzo with an alternative expression for the probability that the number of jumps in $[s, r] \cap \mathcal{C}$, with $\mathcal{C}$ a countable subset of $\mathcal{T}$, is infinite.

Lemma 17. *Consider a stochastic process $(\Omega, \sigma(\mathscr{C}), P)$. Fix a countable subset $\mathcal{C}$ of $\mathcal{T}$. Then for all $s, r \in \mathbb{R}$ such that $s < r$ and $[s, r] \cap \mathcal{C} \neq \emptyset$,*

$$P(\{\eta_{[s,r]\cap\mathcal{C}}(X_\bullet) = +\infty\}) = \lim_{k \to +\infty} \sup\{P(\{\hat{\eta}_u(X_u) \geq k\}) \colon u \in \mathcal{U}_{[s,r]\cap\mathcal{C}}\}.$$

Proof. Clearly, the sequence of level sets $(\{\eta_{[s,r]\cap\mathcal{C}}(X_\bullet) \geq k\})_{k \in \mathbb{N}}$ decreases to the level set $\{\eta_{[s,r]\cap\mathcal{C}}(X_\bullet) = +\infty\}$. All these sets furthermore belong to $\sigma(\mathscr{C})$ because $\eta_{[s,r]\cap\mathcal{C}}(X_\bullet)$ is $\sigma(\mathscr{C})/\mathscr{B}(\overline{\mathbb{R}})$-measurable due to Lemma 15. Since the probability measure P is continuous with respect to monotone sequences, it follows that

$$P(\{\eta_{[s,r]\cap\mathcal{C}}(X_\bullet) = +\infty\}) = \lim_{k \to +\infty} P(\{\eta_{[s,r]\cap\mathcal{C}}(X_\bullet) \geq k\}).$$

The equality in the statement follows if we can prove that for all $k \in \mathbb{N}$,

(7) $$P(\{\eta_{[s,r]\cap\mathcal{C}}(X_\bullet) \geq k\}) = \sup\{P(\{\hat{\eta}_u(X_u) \geq k\}) \colon u \in \mathcal{U}_{[s,r]\cap\mathcal{C}}\}.$$

To this end, we fix any $k \in \mathbb{N}$. By definition of $\eta_{[s,r]\cap\mathcal{C}}(X_\bullet)$, $\hat{\eta}_u(X_u) \leq \eta_{[s,r]\cap\mathcal{C}}(X_\bullet)$ for all $u \in \mathcal{U}_{[s,r]\cap\mathcal{C}}$. Hence, it follows from the monotonicity of the probability measure P that

$$P(\{\eta_{[s,r]\cap\mathcal{C}}(X_\bullet) \geq k\}) \geq P(\{\hat{\eta}_u(X_u) \geq k\}) \quad \text{for all } u \in \mathcal{U}_{[s,r]\cap\mathcal{C}},$$

and therefore

$$P(\{\eta_{[s,r]\cap\mathcal{C}}(X_\bullet) \geq k\}) \geq \sup\{P(\{\hat{\eta}_u(X_u) \geq k\}) \colon u \in \mathcal{U}_{[s,r]\cap\mathcal{C}}\}.$$

Recall from Lemma 15 that there is a sequence $(u_n)_{n \in \mathbb{N}}$ in $\mathcal{U}_{[s,r]\cap\mathcal{C}}$ such that $(\hat{\eta}_{u_n}(X_{u_n}))_{n \in \mathbb{N}}$ is an increasing sequence sequence of positive $\sigma(\mathscr{C})/\mathscr{B}(\overline{\mathbb{R}})$-measurable variables that converges point-wise to $\eta_{[s,r]\cap\mathcal{C}}(X_\bullet)$. Since all of the involved variables take values in $\mathbb{Z}_{\geq 0} \cup \{+\infty\}$, this implies that the corresponding sequence of level sets $(\{\hat{\eta}_{u_n}(X_{u_n}) \geq k\})_{n \in \mathbb{N}}$ increases to the level set $\{\eta_{[s,r]\cap\mathcal{C}}(X_\bullet) \geq k\}$. Again,

it therefore follows from the continuity of the probability measure P with respect to monotone sequences that

$$P(\{\eta_{[s,r]\cap\mathcal{C}}(X_\bullet) \geq k\}) = \lim_{n\to+\infty} P(\{\hat{\eta}_{u_n}(X_{u_n}) \geq k\})$$
$$\leq \sup\{P(\{\hat{\eta}_u(X_u) \geq k\}) : u \in \mathcal{U}_{[s,r]\cap\mathcal{C}}\}.$$

Eqn. (7) now follows immediately from the preceding two inequalities. $\square$

C.2. **Modifying a path with a finite number of jumps along a countable dense subset.** We are interested in the number of jumps of $\dot{X}_\bullet$ in $[s,r]\cap\mathcal{D}$ because if this is finite for some $\dot{\omega} \in \dot{\Omega}$, then the right-sided limit of $\dot{\omega}$ along $\mathcal{D}$ exists in any right-sided limit point t of $\mathcal{D}\cap[s,r[$, and similarly for the left-sided limit in the left-sided limit points. The following intermediary result – which can be seen as a version of Theorem 62.7 in [24, Chapter II] in our setting – essentially shows this, and also establishes a sufficient condition under which we can 'modify' the path $\dot{\omega}$ in such a way that it becomes càdlàg.

Lemma 18. *Consider some $\dot{\omega} \in \dot{\Omega}$, and fix a subset $\mathcal{D}$ of $\mathcal{T}$ such that $\mathcal{T} \subseteq \mathcal{D} \cup$ rlims$(\mathcal{D})$. Suppose that for all $n \in \mathbb{N}$ such that $[-n,n]\cap\mathcal{D} \neq \varnothing$,*

$$\eta_{[-n,n]\cap\mathcal{D}}(\dot{\omega}) < +\infty.$$

Then for all $t \in$ rlims$(\mathcal{D})$, $\lim_{\mathcal{D}\ni r\searrow t}\dot{\omega}(r)$ exists and for all $t \in$ llims$(\mathcal{D})$, $\lim_{\mathcal{D}\ni s\nearrow t}\dot{\omega}(s)$ exists. Furthermore, there is a càdlàg path $\mathring{\omega} \in \mathring{\Omega}$ such that

$$\mathring{\omega}(t) = \begin{cases} \lim_{\mathcal{D}\ni r\searrow t}\dot{\omega}(r) & \text{if } t \in \text{rlims}(\mathcal{D}) \\ \dot{\omega}(t) & \text{if } t \in \mathcal{D} \setminus \text{rlims}(\mathcal{D}) \end{cases} \quad \text{for all } t \in \mathcal{T}.$$

Proof. In the first part of this proof, we prove a convenient intermediary result. Fix any time point $t \in \mathbb{R}$, and let n be a natural number such that $-n < t < n$ and $[-n,n]\cap\mathcal{D} \neq \varnothing$ – this is always possible because $\mathcal{D}$ is dense in the non-empty subset $\mathcal{T}$ of $\mathbb{T}$. By the condition in the statement, there is some tuple of time points $u = (t_1,\ldots,t_m) \in \mathcal{U}_{[-n,n]\cap\mathcal{D}}$ such that

$$\hat{\eta}_u(\dot{\omega}(u)) = \eta_{[-n,n]\cap\mathcal{D}}(\dot{\omega}) = \sup\{\hat{\eta}_v(\dot{\omega}(v)) : v \in \mathcal{U}_{[-n,n]\cap\mathcal{D}}\} < +\infty$$

and, for any tuple of time points $v \in \mathcal{U}_{[-n,n]\cap\mathcal{D}}$ such that $u \sqsubseteq v$,

$$\hat{\eta}_u(\dot{\omega}(u)) \leq \hat{\eta}_v(\dot{\omega}(v)) \leq \eta_{[-n,n]\cap\mathcal{D}}(\dot{\omega}) = \sup\{\hat{\eta}_w(\dot{\omega}(w)) : w \in \mathcal{U}_{[-n,n]\cap\mathcal{D}}\} = \hat{\eta}_u(\dot{\omega}(u)),$$

where for the first inequality we used Lemma 16. In other words, no matter how many time points from $[-n,n]\cap\mathcal{D}$ we add to u, the number of jumps of $\dot{\omega}$ along the sequence will remain the same. From this, we infer that $\dot{\omega}(s) = \dot{\omega}(t_1)$ for all $s \in [-n,t_1]\cap\mathcal{D}$, $\dot{\omega}(r) = \dot{\omega}(t_m)$ for all $r \in [t_m,n]\cap\mathcal{D}$ and that for all $k \in \{1,\ldots,m-1\}$, there is some $t_k^\star \in [t_k,t_{k+1}]$ such that $\dot{\omega}(s) = \dot{\omega}(t_k)$ for all $s \in [t_k,t_k^\star[\cap\mathcal{D}$ and $\dot{\omega}(r) = \dot{\omega}(t_{k+1})$ for all $r \in]t_k^\star,t_{k+1}]\cap\mathcal{D}$. Because $-n < t < n$, we infer from this that there are strictly positive real numbers $\delta_{t,+}, \delta_{t,-}$ and states $x_{t,+}, x_{t,-} \in \mathcal{X}$ such that

$$(8) \quad (\forall r \in]t, t+\delta_{t,+}[\cap\mathcal{D})\ \dot{\omega}(r) = x_{t,+} \quad \text{and} \quad (\forall s \in]t-\delta_{t,-},t[\cap\mathcal{D})\ \dot{\omega}(s) = x_{t,-}.$$

In the second part of this proof, we use Eqn. (8) to prove the first part of the statement. Fix a right-sided limit point t of $\mathcal{D}$. From the first part of this proof – that is, from right before Eqn. (8) – we know that there are some positive real number $\delta_{t,+}$ and state $x_{t,+}$ such that for all $r \in]t, t+\delta_{t,+}[\cap\mathcal{D}$, $\dot{\omega}(r) = x_{t,+}$. Then $\lim_{\mathcal{D}\ni r\searrow t}\dot{\omega}(r)$ exists, and is equal to $x_{t,+}$. Similarly, for any left-sided limit point t of $\mathcal{D}$, there are some positive real number $\delta_{t,-}$ and state $x_{t,-}$ such that for all $s \in]t-\delta_{t,-},t[\cap\mathcal{D}$, $\dot{\omega}(r) = x_{t,-}$; so $\lim_{\mathcal{D}\ni s\nearrow t}\dot{\omega}(s)$ exists and is equal to $x_{t,-}$.

In the third part of this proof, we show the existence of the càdlàg path $\mathring{\omega} \in \mathring{\Omega}$. We start by constructing a map $\varpi \colon \mathcal{T} \to \mathcal{X}$, defined by

$$\varpi(t) \coloneqq \begin{cases} \lim_{\mathcal{D} \ni r \searrow t} \mathring{\omega}(r) & \text{if } t \in \mathrm{rlims}(\mathcal{D}) \\ \mathring{\omega}(t) & \text{if } t \in \mathcal{D} \setminus \mathrm{rlims}(\mathcal{D}) \end{cases} \quad \text{for all } t \in \mathcal{T}.$$

In the remainder of this proof, we will show that ϖ is càdlàg. It then follows from Lemma 6 that ϖ can be extended to a map $\psi \colon \mathbb{R} \to \mathcal{X}$ that is càdlàg. Since restrictions of càdlàg maps are càdlàg too, this implies that $\mathring{\omega} \coloneqq \psi|_{\mathbb{T}}$ is a càdlàg path in $\mathring{\Omega}$ that, since it extends ϖ, clearly satisfies the condition in the statement.

To show that ϖ is càdlàg, we consider any $t \in \mathbb{R}$. We've shown earlier in this proof that there are strictly positive real numbers $\delta_{t,+}, \delta_{t,-}$ and states $x_{t,+}, x_{t,-} \in \mathcal{X}$ that satisfy Eqn. (8). Taking into account the definition of ϖ, this implies that

$$\big(\forall r \in \,]t, t+\delta_{t,+}[\, \cap \mathcal{T}\big) \; \varpi(r) = x_{t,+} \quad \text{and} \quad \big(\forall s \in \,]t-\delta_{t,-}, t[\, \cap \mathcal{T}\big) \; \varpi(s) = x_{t,-}.$$

Since this is true for any $t \in \mathbb{R}$, and therefore definitely for any $t \in \mathrm{llims}(\mathcal{T})$ or $t \in \mathrm{rlims}(\mathcal{T})$ this already establishes the first two conditions in Definition 3. To establish the third condition, we need to show that if $t \in \mathcal{T} \cap \mathrm{rlims}(\mathcal{T})$, then also $\varpi(t) = x_{t,+}$. So consider the case $t \in \mathcal{T} \cap \mathrm{rlims}(\mathcal{T})$. Since $t \in \mathrm{rlims}(\mathcal{T})$, we also have that $t \in \mathrm{rlims}(\mathcal{D})$. To see why that is the case, consider any $\delta \in \mathbb{R}_{>0}$. Since $t \in \mathrm{rlims}(\mathcal{T})$, we know that $]t, t+\delta[\, \cap \mathcal{T} \neq \varnothing$. Choose any $t_\delta \in \,]t, t+\delta[\, \cap \mathcal{T}$. Since $t_\delta \in \mathcal{T} \subseteq \mathcal{D} \cup \mathrm{rlims}(\mathcal{D})$, we then have that $[t_\delta, t+\delta[\, \cap \mathcal{D} \neq \varnothing$, which implies that also $]t, t+\delta[\, \cap \mathcal{D} \neq \varnothing$. Since $\delta \in \mathbb{R}_{>0}$ was arbitrary, we indeed have that $t \in \mathrm{rlims}(\mathcal{D})$. It therefore follows from Eqn. (8) that $\varpi(t) = x_{t,+}$. $\qquad\square$

C.3. **Two additional intermediary results.** In our proof for Theorem 10, we need the following continuity property of probability measures – for a proof, see for example Theorem 2 in [14, Section 6.1].

Lemma 19. *Consider a probability space $(\mathfrak{S}, \mathfrak{A}, P)$ and a sequence $(A_n)_{n \in \mathbb{N}}$ in $\mathfrak{A}$. If $\lim_{n \to +\infty} \mathbb{I}_{A_n}(\mathfrak{s})$ exists for all $\mathfrak{s} \in \mathfrak{S}$, then*

$$\lim_{n \to +\infty} A_n \coloneqq \Big\{ \mathfrak{s} \in \mathfrak{S} \colon \lim_{n \to +\infty} \mathbb{I}_{A_n}(\mathfrak{s}) = 1 \Big\} \in \mathfrak{A}$$

and

$$P\Big(\lim_{n \to +\infty} A_n \Big) = \lim_{n \to +\infty} P(A_n).$$

Another example where Lemma 19 comes in handy is the following intermediary result, which is the final lemma we need to prove Theorem 10. To state it, we observe that for all $n \in \mathbb{N}$ and $v, w \in \mathcal{U}$ such that $v = (v_1, \ldots, v_m)$ and $w = (w_1, \ldots, w_m)$,

$$\{X_v = X_w\} \coloneqq \bigcap_{j=1}^{m} \{X_{v_j} = X_{w_j}\} \in \mathscr{C}.$$

Lemma 20. *Consider a probability charge P on a domain $\mathscr{A}$ that includes $\mathscr{C}$. Fix some $v = (t_1, \ldots, t_m) \in \mathcal{U}$ and some countable subset $\mathcal{D}$ of $\mathcal{T}$ such that $\{t_1, \ldots, t_m\} \subseteq \mathcal{D} \cup \mathrm{rlims}(\mathcal{D})$. Suppose the collection $\mu_\bullet$ of finite-dimensional charges of P satisfies (R1). Then for all $\epsilon \in \mathbb{R}_{>0}$, there is some $w = (t'_1, \ldots, t'_m) \in \mathcal{U}_\mathcal{D}$ such that*

$$|P(A) - P(A \cap \{X_v = X_w\})| < \epsilon \quad \text{for all } A \in \mathscr{A}.$$

Furthermore, for any $\delta \in \mathbb{R}_{>0}$, we can guarantee that $t'_j \in \,]t_j, t_j + \delta[$ if $t_j \in \mathrm{rlims}(\mathcal{D})$ and $t'_j = t_j$ otherwise. Hence, there is some sequence $(w_\ell)_{\ell \in \mathbb{N}}$ in $\mathcal{U}_\mathcal{D}$ such that (i) for all $\ell \in \mathbb{N}$, $w_\ell = (t_1^\ell, \ldots, t_m^\ell)$, (ii) for all $j \in \{1, \ldots, m\}$, $(t_j^\ell)_{\ell \in \mathbb{N}}$ is a

strictly decreasing sequence in $\mathcal{D}$ *that converges to* t_j *if* $t_j \in \mathrm{rlims}(\mathcal{D})$ *and a constant sequence that is equal to* t_j *otherwise, and (iii)*

$$P(A) = \lim_{\ell \to +\infty} P(A \cap \{X_v = X_{w_\ell}\})$$

Proof. Fix some $j \in \{1, \ldots, m\}$. If $t_j \in \mathcal{D} \setminus \mathrm{rlims}(\mathcal{D})$, then we let $t'_j := t_j$, and then trivially

$$P(\{X_{t_j} = X_{t'_j}\}) = 1 > 1 - \frac{\epsilon}{m}.$$

If on the other hand $t_j \in \mathrm{rlims}(\mathcal{D})$, then it follows from the assumptions on P and $\mathcal{D}$ in the statement that there is some t'_j in $]t_j, t_{j+1}[\cap]t_j, t_j + \delta[\cap \mathcal{D}$ – where we let $t_{m+1} := +\infty$ – such that

$$P(\{X_{t_j} = X_{t'_j}\}) = \mu_{(t_j, t'_j)}(\mathcal{X}^2_=) > 1 - \frac{\epsilon}{m}.$$

Let $w := (t'_1, \ldots, t'_m) \in \mathcal{U}_{\mathcal{D}}$, and note that

$$\{X_v = X_w\} = \bigcap_{j=1}^{m} \{X_{t_j} = X_{t'_j}\} \quad \text{and} \quad \{X_v = X_w\}^{\mathrm{c}} = \bigcup_{j=1}^{m} \{X_{t_j} = X_{t'_j}\}^{\mathrm{c}}.$$

It follows from this, the properties of the probability charge P and our construction of w that

$$P(\{X_v = X_w\}^{\mathrm{c}}) = P\left(\bigcup_{j=1}^{m} \{X_{t_j} = X_{t'_j}\}^{\mathrm{c}} \right) \leq \sum_{j=1}^{m} P(\{X_{t_j} = X_{t'_j}\}^{\mathrm{c}})$$

$$= \sum_{j=1}^{m} \left(1 - P(\{X_{t_j} = X_{t'_j}\}) \right)$$

$$< \sum_{j=1}^{m} \frac{\epsilon}{m} = \epsilon.$$

Now for all $A \in \mathscr{A}$,

$$P(A) = P(A \cap \{X_v = X_w\}) + P(A \cap \{X_v = X_w\}^{\mathrm{c}}),$$

and therefore

$$|P(A) - P(A \cap \{X_v = X_w\})| = P(A \cap \{X_v = X_w\}^{\mathrm{c}}) \leq P(\{X_v = X_w\}^{\mathrm{c}}) < \epsilon,$$

as required.

The second part of the statement follows immediately from the first part, since it is easy to ensure that the sequences $(t^\ell_j)_{\ell \in \mathbb{N}}$ are decreasing if $t_j \in \mathrm{rlims}(\mathcal{D})$: when constructing $w_{\ell+1}$, simply choose δ small enough such that $t_j + \delta < t^\ell_j$ for all $j \in \{1, \ldots, m\}$ such that $t_j \in \mathrm{rlims}(\mathcal{D})$. □

C.4. **Proof for Theorem 10.** Finally, we can get around to proving the precursor to our main result.

Proof of Theorem 10. First, we prove the necessity; that is, we assume that the probability charge P is countably additive, and show that this implies that $\mu_\bullet$ is regular. Since P is countably additive, we know from Caratheodory's Extension Theorem that there is a unique probability measure P_σ on $\sigma(\mathscr{C})$ that extends P. Note that the finite-dimensional distributions of P_σ are $\mu_\bullet$. We will use this to show that $\mu_\bullet$ is regular.

For (R1), we fix some $t \in \mathrm{rlims}(\mathcal{T}) \cap \mathcal{T}$. For any decreasing sequence $(r_n)_{n \in \mathbb{N}}$ in $]t, +\infty[\cap \mathcal{T}$ with $\lim_{n \to +\infty} r_n = t$, $\lim_{n \to +\infty} \omega(r_n) = \omega(t)$ for all $\omega \in \mathring{\Omega}$ due

to the right-continuity of càdlàg paths, and therefore $\lim_{n \to +\infty} \{\mathring{X}_t = \mathring{X}_{r_n}\} = \mathring{\Omega}$; hence, it follows from Lemma 19 that

$$\lim_{n \to +\infty} P_\sigma(\{\mathring{X}_t = \mathring{X}_{r_n}\}) = P_\sigma(\mathring{\Omega}) = 1.$$

Since this is true for any such sequence, it follows that

$$\lim_{\mathcal{T} \ni r \searrow t} P_\sigma(\{\mathring{X}_t = \mathring{X}_r\}) = 1.$$

Furthermore, for all $r \in \mathcal{T}$ such that $r > t$, we have that

$$P_\sigma(\{\mathring{X}_t = \mathring{X}_r\}) = P(\{\mathring{X}_t = \mathring{X}_r\}) = P(\{\mathring{X}_{(t,r)} \in \mathcal{X}_=^2\}) = \mu_{(t,r)}(\mathcal{X}_=^2).$$

From these two observations, we infer that $\lim_{\mathcal{T} \ni r \searrow t} \mu_{(t,r)}(\mathcal{X}_=^2) = 1$, as required for (R1).

For (R2), we fix some $n \in \mathbb{N}$ such that $[-n,n] \cap \mathcal{T} \neq \varnothing$, and some countable subset $\mathcal{C}$ of $[-n,n] \cap \mathcal{T}$ such that $[-n,n] \cap \mathcal{T} \subseteq \mathcal{C} \cup \mathrm{rlims}(\mathcal{C})$; this is always possible due to Lemma 8. Since $\mathcal{C} \subseteq [-n,n] \cap \mathbb{T}$, we know from Proposition 9 that

$$\eta_\mathcal{C}(\mathring{\omega}) \leq \eta_{[-n,n] \cap \mathbb{T}}(\mathring{\omega}) < +\infty \quad \text{for all } \mathring{\omega} \in \mathring{\Omega}.$$

Hence, $\{\eta_\mathcal{C}(\mathring{X}_\bullet) = +\infty\} = \varnothing$, and therefore it must be that

$$P_\sigma(\{\eta_\mathcal{C}(\mathring{X}_\bullet) = +\infty\}) = P_\sigma(\varnothing) = 0.$$

From this and Lemma 17, it follows that

$$(9) \qquad \lim_{k \to +\infty} \sup\{P_\sigma(\{\hat{\eta}_u(\mathring{X}_u) \geq k\}) \colon u \in \mathcal{U}_\mathcal{C}\} = 0.$$

Now fix any $k \in \mathbb{N}$ and $\epsilon \in \mathbb{R}_{>0}$. Then for all $v \in \mathcal{U}_{[-n,n] \cap \mathcal{T}}$, we know from Lemma 20 – with $\mathcal{C}$ and $\{\hat{\eta}_v(\mathring{X}_v) \geq k\}$ here in the role of $\mathcal{D}$ and A there – that there is some $w \in \mathcal{U}_\mathcal{C}$ such that

$$P_\sigma(\{\hat{\eta}_v(\mathring{X}_v) \geq k\}) < P_\sigma(\{\hat{\eta}_v(\mathring{X}_v) \geq k\} \cap \{\mathring{X}_v = \mathring{X}_w\}) + \epsilon \leq P_\sigma(\{\hat{\eta}_w(\mathring{X}_w) \geq k\}) + \epsilon,$$

where for the non-strict inequality we used the monotonicity of P_σ and that

$$\{\hat{\eta}_v(\mathring{X}_v) \geq k\} \cap \{\mathring{X}_v = \mathring{X}_w\} \subseteq \{\hat{\eta}_w(\mathring{X}_w) \geq k\}.$$

From this inequality, we infer that

$$\sup\{P_\sigma(\{\hat{\eta}_u(\mathring{X}_u) \geq k\}) \colon u \in \mathcal{U}_{[-n,n] \cap \mathcal{T}}\} \leq \sup\{P_\sigma(\{\hat{\eta}_u(\mathring{X}_u) \geq k\}) \colon u \in \mathcal{U}_\mathcal{C}\} + \epsilon.$$

Since $\epsilon \in \mathbb{R}_{>0}$ is arbitrary, and $\mathcal{C} \subseteq [-n,n] \cap \mathcal{T}$, it follows that

$$\sup\{P_\sigma(\{\hat{\eta}_u(\mathring{X}_u) \geq k\}) \colon u \in \mathcal{U}_{[-n,n] \cap \mathcal{T}}\} = \sup\{P_\sigma(\{\hat{\eta}_u(\mathring{X}_u) \geq k\}) \colon u \in \mathcal{U}_\mathcal{C}\}.$$

This is true for all $k \in \mathbb{N}$, so we infer from Equation (9) that

$$\lim_{k \to +\infty} \sup\{P_\sigma(\{\hat{\eta}_u(\mathring{X}_u) \geq k\}) \colon u \in \mathcal{U}_{[-n,n] \cap \mathcal{T}}\} = 0.$$

Because the finite-dimensional distributions of P_σ are $\mu_\bullet$, we conclude from this that

$$\lim_{k \to +\infty} \sup\{\mu_u(\{x_u \in \mathcal{X}_u \colon \hat{\eta}_u(x_u) \geq k\}) \colon u \in \mathcal{U}_{[-n,n] \cap \mathcal{T}}\} = 0,$$

which is exactly what (R2) demands.

Second, we prove the sufficiency: we assume that $\mu_\bullet$ is regular, and show that the corresponding probability charge P is countably additive, or equivalently, that for any sequence $(\mathring{A}_i)_{i \in \mathbb{N}}$ in $\mathring{\mathscr{C}}$ such that $\mathring{A}_i \supseteq \mathring{A}_{i+1}$ for all $i \in \mathbb{N}$ and $\bigcap_{i \in \mathbb{N}} \mathring{A}_i = \varnothing$, $\lim_{i \to +\infty} P(\mathring{A}_i) = 0$.

To prove this, we fix any such sequence $(\mathring{A}_i)_{i \in \mathbb{N}}$. By definition of $\mathring{\mathscr{C}}$, for all $i \in \mathbb{N}$ there are some $u_i \in \mathcal{U}$ and $A_i \in \wp(\mathcal{X}_{u_i})$ such that $\mathring{A}_i = \{\mathring{X}_{u_i} \in A_i\}$. Without loss of generality, we may assume that $u_i \sqsubseteq u_{i+1}$ – if this is not the case, add the missing points to u_{i+1} to obtain u'_{i+1}, and replace A_{i+1} by $A'_{i+1} := \{x_{u'_{i+1}} \in$

$\mathcal{X}_{u'_{i+1}}\colon x_{u_{i+1}} \in A_{i+1}\}$. Then for all $i \in \mathbb{N}$, $\{x_{u_{i+1}} \in \mathcal{X}_{u_{i+1}}\colon x_{u_i} \in A_i\} \supseteq A_{i+1}$ since $(\mathring{A}_i)_{i\in\mathbb{N}}$ is decreasing.

To prove that the probability of these events converges to 0, we will rely on Theorem 3. Since $\mu_\bullet$ is consistent, this result says that there is a (unique) probability measure $\dot{P}$ on $\sigma(\mathscr{C})$ such that

$$(10) \qquad \dot{P}(\{\dot{X}_u \in A\}) = \mu_u(A) = P(\{\mathring{X}_u \in A\}) \quad \text{for all } u \in \mathcal{U}, A \in \wp(\mathcal{X}_u).$$

It follows from these equalities that

$$(11) \qquad \lim_{i\to+\infty} P(\mathring{A}_i) = \lim_{i\to+\infty} P(\{\mathring{X}_{u_i} \in A_i\}) = \lim_{i\to+\infty} \dot{P}(\{\dot{X}_{u_i} \in A_i\}).$$

Note that since $(\{\mathring{X}_{u_i} \in A_i\})_{i\in\mathbb{N}}$ is decreasing, the same must hold for $(\{\dot{X}_{u_i} \in A_i\})_{i\in\mathbb{N}}$.

Fix some countable subset $\mathcal{D}$ of $\mathcal{T}$ such that $\mathcal{T} \subseteq \mathcal{D} \cup \mathrm{rlims}(\mathcal{D})$; this is always possible due to Lemma 8. Then it follows from Eqn. (10) and (R1) that for all $t \in \mathcal{T} \cap \mathrm{rlims}(\mathcal{T})$

$$\lim_{\mathcal{T}\ni r \searrow t} \dot{P}(\{\dot{X}_t = \dot{X}_r\}) = \lim_{\mathcal{T}\ni r \searrow t} \mu_{(t,r)}(\mathcal{X}^2_=) = 1.$$

For the implication of Eqn. (10) and (R2), we fix some $n \in \mathbb{N}$ such that $[-n,n]\cap\mathcal{D} \neq \varnothing$. Recall from Lemma 17 that

$$\dot{P}(\{\eta_{[-n,n]\cap\mathcal{D}}(\dot{X}_\bullet) = +\infty\}) = \lim_{k\to+\infty} \sup\{\dot{P}(\{\hat{\eta}_u(\dot{X}_u) \geq k\})\colon u \in \mathcal{U}_{[-n,n]\cap\mathcal{D}}\}.$$

It therefore follows from Eqns. (R2) and (10) that

$$\begin{aligned}
0 &= \lim_{k\to+\infty} \sup\{\mu_u(\{x_u \in \mathcal{X}_u\colon \hat{\eta}_u(x_u) \geq k\})\colon u \in \mathcal{U}_{[-n,n]\cap\mathcal{T}}\} \\
&\geq \lim_{k\to+\infty} \sup\{\mu_u(\{x_u \in \mathcal{X}_u\colon \hat{\eta}_u(x_u) \geq k\})\colon u \in \mathcal{U}_{[-n,n]\cap\mathcal{D}}\} \\
&= \lim_{k\to+\infty} \sup\{\dot{P}(\{\hat{\eta}_u(\dot{X}_u) \geq k\})\colon u \in \mathcal{U}_{[-n,n]\cap\mathcal{D}}\} \\
&= \dot{P}(\{\eta_{[-n,n]\cap\mathcal{D}}(\dot{X}_\bullet) = +\infty\}) \geq 0,
\end{aligned}$$

which implies that $\dot{P}(\{\eta_{[-n,n]\cap\mathcal{D}}(\dot{X}_\bullet) = +\infty\}) = 0$ and therefore, that

$$\dot{P}(\{\eta_{[-n,n]\cap\mathcal{D}}(\dot{X}_\bullet) < +\infty\}) = 1 - \dot{P}(\{\eta_{[-n,n]\cap\mathcal{D}}(\dot{X}_\bullet) = +\infty\}) = 1.$$

Let $\dot{A} := \bigcap_{n=n^\star}^{+\infty}\{\eta_{[-n,n]\cap\mathcal{D}}(\dot{X}_\bullet) < +\infty\}$, where $n^\star$ is the smallest natural number such that $[-n,n]\cap\mathcal{D} \neq \varnothing$. Since each of the events $\{\eta_{[-n,n]\cap\mathcal{D}}(\dot{X}_\bullet) < +\infty\}$ has probability 1, so does their countable intersection $\dot{A}$. Hence,

$$(12) \qquad \lim_{i\to+\infty} \dot{P}(\{\dot{X}_{u_i} \in A_i\}) = \lim_{i\to+\infty} \dot{P}(\dot{A} \cap \{\dot{X}_{u_i} \in A_i\}).$$

From Lemma 20, we know that for all $i \in \mathbb{N}$ there is a sequence $(u_{i,\ell})_{\ell\in\mathbb{N}}$ in $\mathcal{U}_\mathcal{D}$ such that (i) each of the $u_{i,\ell}$'s has the same number of time points as u_i, (ii) if the j-th component of u_i is a right-sided limit point of $\mathcal{D}$, then the j-th component of $u_{i,\ell}$ strictly decreases to it as ℓ recedes to $+\infty$, and if the j-th component of u_i is not a right-sided limit point of $\mathcal{D}$, the j-th component of $u_{i,\ell}$ is equal to it, and (iii)

$$\dot{P}(\dot{A} \cap \{\dot{X}_{u_i} \in A_i\}) = \lim_{\ell\to+\infty} \dot{P}(\dot{A} \cap \{\dot{X}_{u_i} \in A_i\} \cap \{\dot{X}_{u_i} = \dot{X}_{u_{i,\ell}}\}).$$

Since for all $i, \ell \in \mathbb{N}$,

$$\begin{aligned}
\dot{A} \cap \{\dot{X}_{u_i} \in A_i\} \cap \{\dot{X}_{u_i} = \dot{X}_{u_{i,\ell}}\} &= \dot{A} \cap \{\dot{X}_{u_{i,\ell}} \in A_i\} \cap \{\dot{X}_{u_i} = \dot{X}_{u_{i,\ell}}\} \\
&\subseteq \dot{A} \cap \{\dot{X}_{u_{i,\ell}} \in A_i\},
\end{aligned}$$

we infer from this that for all $i \in \mathbb{N}$,

$$(13) \qquad \dot{P}(\dot{A} \cap \{\dot{X}_{u_i} \in A_i\}) \leq \limsup_{\ell \to +\infty} \dot{P}(\dot{A} \cap \{\dot{X}_{u_{i,\ell}} \in A_i\}).$$

For all $i \in \mathbb{N}$ and $\dot{\omega} \in \dot{A}$, it follows from the construction of $(u_{i,\ell})_{\ell \in \mathbb{N}}$ and – for the components of u_i that are a right-sided limit point of $\mathcal{D}$ – Lemma 18 that $\lim_{\ell \to +\infty} \dot{\omega}(u_{i,\ell})$ exists (with the limit taken component-wise); from this and Lemma 19, it follows that for all $i \in \mathbb{N}$

$$\dot{A}_i := \{\dot{\omega} \in \dot{A}: \lim_{\ell \to +\infty} \dot{\omega}(u_{i,\ell}) \in A_i\} = \lim_{\ell \to +\infty} \dot{A} \cap \{\dot{X}_{u_{i,\ell}} \in A_i\} \in \sigma(\mathscr{C})$$

and

$$(14) \qquad \dot{P}(\dot{A}_i) = \lim_{\ell \to +\infty} \dot{P}(\dot{A} \cap \{\dot{X}_{u_{i,\ell}} \in A_i\}) = \limsup_{\ell \to +\infty} \dot{P}(\dot{A} \cap \{\dot{X}_{u_{i,\ell}} \in A_i\}).$$

We now establish two properties of these sets A_i.

The first property is that $(\dot{A}_i)_{i \in \mathbb{N}}$ is a decreasing sequence. To prove this, we consider any $i \in \mathbb{N}$ and set out to show that $\dot{A}_{i+1} \subseteq \dot{A}_i$, or equivalently, that any $\dot{\omega}$ in $\dot{A}_{i+1}$ belongs to $\dot{A}_i$ as well. So consider any $\dot{\omega} \in \dot{A}_{i+1}$. Then $\dot{\omega} \in \dot{A}$ and $\lim_{\ell \to +\infty} \dot{\omega}(u_{i+1,\ell}) \in A_{i+1} \subseteq \{x_{u_{i+1}} \in \mathcal{X}_{u_{i+1}}: x_{u_i} \in A_i\}$. So there is some $x_{u_i} \in A_i$ such that, for each $t \in u_i$, the corresponding component of $\lim_{\ell \to +\infty} \dot{\omega}(u_{i+1,\ell})$ is equal to x_t. Furthermore, if t is a right-sided limit point of $\mathcal{D}$, then the corresponding components of $\lim_{\ell \to +\infty} \dot{\omega}(u_{i+1,\ell})$ and $\lim_{\ell \to +\infty} \dot{\omega}(u_{i,\ell})$ are equal because $\lim_{\mathcal{D} \ni r \searrow t} \dot{\omega}(r)$ exists due to Lemma 18, and if t is not a right-sided limit point of $\mathcal{D}$, then the corresponding components of $\lim_{\ell \to +\infty} \dot{\omega}(u_{i+1,\ell})$ and $\lim_{\ell \to +\infty} \dot{\omega}(u_{i,\ell})$ are equal because they are both equal to $\dot{\omega}(t)$. In all cases, we conclude that for each $t \in u_i$, the corresponding component of $\lim_{\ell \to +\infty} \dot{\omega}(u_{i,\ell})$ is equal to x_t. So $\lim_{\ell \to +\infty} \dot{\omega}(u_{i,\ell}) = x_{u_i} \in A_i$ and therefore, since $\dot{\omega} \in \dot{A}$, we see that $\dot{\omega} \in \dot{A}_i$, as required.

The second property is that $\bigcap_{i \in \mathbb{N}} \dot{A}_i = \varnothing$. To see why this is true, assume *ex absurdo* that $\bigcap_{i \in \mathbb{N}} \dot{A}_i \neq \varnothing$. Then there is some $\dot{\omega} \in \dot{A}$ such that for all $i \in \mathbb{N}$, $\lim_{\ell \to +\infty} \dot{\omega}(u_{i,\ell}) \in A_i$. By Lemma 18, there is some $\mathring{\omega} \in \mathring{\Omega}$ such that for all $i \in \mathbb{N}$, $\mathring{\omega}(u_i) = \lim_{\ell \to +\infty} \dot{\omega}(u_{i,\ell}) \in A_i$; for the components t of u_i that are right-sided limit points of $\mathcal{D}$, the equality in this statement follows because $\mathring{\omega}(t) = \lim_{\mathcal{D} \ni r \searrow t} \dot{\omega}(r)$, and for the components t of u_i that are not right-sided limit points of $\mathcal{D}$, this follows because $\mathring{\omega}(t) = \dot{\omega}(t)$ and because the corresponding components of $u_{i,\ell}$ are all equal to t. But then $\mathring{\omega} \in \bigcap_{i \in \mathbb{N}} \{\mathring{X}_{u_i} \in A_i\} = \bigcap_{i \in \mathbb{N}} \mathring{A}_i$, which is a contradiction because $\bigcap_{i \in \mathbb{N}} \mathring{A}_i = \varnothing$.

Since $\dot{P}$ is a probability measure – and therefore definitely continuous for decreasing sequences – it follows from these two properties that

$$(15) \qquad \lim_{i \to +\infty} \dot{P}(\dot{A}_i) = \dot{P}(\varnothing) = 0.$$

Finally, it follows from Eqns. (11) to (15) – and the non-negativity of P – that

$$\begin{aligned}
0 \leq \lim_{i \to +\infty} P(\mathring{A}_i) &= \lim_{i \to +\infty} \dot{P}(\{\dot{X}_{u_i} \in A_i\}) \\
&= \lim_{i \to +\infty} \dot{P}(\dot{A} \cap \{\dot{X}_{u_i} \in A_i\}) \\
&\leq \limsup_{\ell \to +\infty} \dot{P}(\dot{A} \cap \{\dot{X}_{u_{i,\ell}} \in A_i\}) = \lim_{i \to +\infty} \dot{P}(\dot{A}_i) = 0.
\end{aligned}$$

So we find that $\lim_{i \to +\infty} P(\mathring{A}_i)$, which finalises our proof of the sufficiency. $\qquad\square$

APPENDIX D. CONDITIONS FOR REGULARITY

The two regularity conditions (R1) and (R2) are perhaps not the most easy to check. We have already given a simpler sufficient condition in Proposition 12 in the main text. In this appendix, we give two alternative sufficient conditions in the special case that $\mathcal{T} \cap [-n, n]$ is closed for all $n \in \mathbb{N}$. These only involve the finite-dimensional distributions for two time points, and can be interpreted as bounding the dynamics of the stochastic process.

Proposition 21. *Suppose that $\mathcal{T} \cap [-n, n]$ is closed for all $n \in \mathbb{N}$. Consider a consistent collection $\mu_\bullet$ of finite-dimensional distributions. If for all $n \in \mathbb{N}$ there is some $\lambda_n \in \mathbb{R}_{\geq 0}$ such that*

$$(\forall t \in \mathcal{T} \cap [-n, n])(\forall \epsilon \in \mathbb{R}_{>0})(\exists \delta \in \mathbb{R}_{>0})(\forall s, r \in \mathcal{T} \cap]t - \delta, t + \delta[\colon s < r)$$

$$\frac{\mu_{(s,r)}(\mathcal{X}_{\neq}^2)}{r - s} < \lambda_n + \epsilon,$$

then $\mu_\bullet$ is regular.

In our proof, we will lean on the following two intermediary results, which will come in handy in our proof for Proposition 24 further on as well. The first one is an immediate consequence of the Heine-Borel Theorem.

Lemma 22. *Consider a closed and bounded subset $\mathcal{S}$ of $\mathcal{T}$. For all $s \in \mathcal{S}$, fix some positive real number δ_s. Then there is a tuple of time points $(s_1, \ldots, s_m) \in \mathcal{U}_{\mathcal{S}}$ such that, with $\Delta_\ell := \delta_{s_\ell}$ for all $\ell \in \{1, \ldots, m\}$,*
- (i) $\mathcal{S} \subseteq \bigcup_{\ell=1}^m]s_\ell - \Delta_\ell, s_\ell + \Delta_\ell[$; *and*
- (ii) $s_i - \Delta_i < s_j - \Delta_j$ *and* $s_i + \Delta_i < s_j + \Delta_j$ *for all* $i, j \in \{1, \ldots, m\}$ *such that $i < j$.*

Hence, for any sequence $u = (t_1, \ldots, t_n) \in \mathcal{U}_{\mathcal{S}}$ with $n \geq 2$,

$$(16) \quad \left| \left\{ k \in \{2, \ldots, n\} \colon (\nexists \ell \in \{1, \ldots, m\}) \, \{t_{k-1}, t_k\} \subseteq]s_\ell - \Delta_\ell, s_\ell + \Delta_\ell[\right\} \right| < m.$$

If furthermore $\mathcal{S}$ is convex, then

$$(17) \qquad\qquad s_{\ell-1} + \Delta_{\ell-1} > s_\ell - \Delta_\ell \text{ for all } \ell \in \{2, \ldots, m\}.$$

Proof. Statements (i) and (ii) and Eqn. (17) essentially follow from an argument in [20, Proof of Lemma F.1], which we modify to our slightly different context. For all $s \in \mathcal{S}$, we let $C_s :=]s - \delta_s, s + \delta_s[$. Note that $(C_s)_{s \in \mathcal{S}}$ is an open cover of the bounded and closed set $\mathcal{S}$. By the Heine-Borel Theorem, this open cover has a finite subcover: $C_{s_1}, \ldots, C_{s_m}$. Without loss of generality, we may assume that $s_1 < \cdots < s_m$ and that this subcover is minimal, in the sense that removing one of its elements has as a consequence that we no longer have a subcover of $\mathcal{S}$ any more. This proves (i).

Our proof for (ii) is one by contradiction. Assume *ex absurdo* that there are indices $i, j \in \{1, \ldots, m\}$ such that $i < j$ and either $s_i - \Delta_i \geq s_j - \Delta_j$ or $s_i + \Delta_i \geq s_j + \Delta_j$. Let us start with the first case. Since $i < j$ by assumption, it follows that $s_i < s_j$, and therefore, since $s_i - \Delta_i \geq s_j - \Delta_j$ by assumption, that $\Delta_j \geq s_j - s_i + \Delta_i > \Delta_i$. From this, it follows that $s_i + \Delta_i < s_i + \Delta_j < s_j + \Delta_j$, where for the second inequality we again used that $s_i < s_j$. Hence, $C_{s_i} =]s_i - \Delta_i, s_i + \Delta_i[\subseteq]s_j - \Delta_j, s_j + \Delta_j[= C_{s_j}$, but this is a contradiction because $C_{s_1}, \ldots C_{s_m}$ is minimal. For the second case that $i < j$ and $s_i + \Delta_i \geq s_j + \Delta_j$, a similar argument leads to a contradiction as well.

Next, we fix any $u = (t_1, \ldots, t_n) \in \mathcal{U}_{\mathcal{S}}$ with $n \geq 2$, and set out to prove Eqn. (16). For all $k \in \{1, \ldots, n\}$, the time point t_k belongs to $\mathcal{S}$, so it follows from (i) that the index set

$$\mathcal{L}_k := \left\{ \ell \in \{1, \ldots, m\} \colon t_k \in C_{s_\ell} \right\}$$

is non-empty and finite. Furthermore, for all $k \in \{1, \ldots, n\}$,

$$(18) \qquad \mathcal{L}_k = \{\ell \in \mathbb{N} \colon \min \mathcal{L}_k \leq \ell \leq \max \mathcal{L}_k\}.$$

Indeed, if this were not the case, there would be indices $i, \ell, j \in \{1, \ldots, m\}$ such that $i < \ell < j$, $t_k \in C_{s_i} \cap C_{s_j}$ and $t_k \notin C_{s_\ell}$. Then either $t_k \leq s_\ell - \Delta_\ell$ or $t_k \geq s_\ell + \Delta_\ell$, but by (ii) this implies that $t_k < s_j - \Delta_j$ – so $t_k \notin C_{s_j}$ – or $t_k > s_i + \Delta_i$ – so $t_k \notin C_{s_i}$ – which is a contradiction because $t_k \in C_{s_i} \cap C_{s_j}$ by assumption. Finally, for all indices $k_1, k_2 \in \{1, \ldots, n\}$ such that $k_1 < k_2$,

$$(19) \qquad \min \mathcal{L}_{k_1} \leq \min \mathcal{L}_{k_2} \text{ and } \max \mathcal{L}_{k_1} \leq \max \mathcal{L}_{k_2};$$

again, this follows more or less immediately from (ii). Indeed, assume *ex absurdo* that $j := \min \mathcal{L}_{k_1} > \min \mathcal{L}_{k_2} =: i$. Then $t_{k_1} \in {]s_j - \Delta_j, s_j + \Delta_j[}$ and $t_{k_1} \notin {]s_i - \Delta_i, s_i + \Delta_i[} \ni t_{k_2}$, and by (ii), $s_i - \Delta_i < s_j - \Delta_j$. Hence, either (i) $t_{k_1} \leq s_i - \Delta_i$ and therefore $t_{k_1} < s_j - \Delta_j < t_{k_1}$, or (ii) $t_{k_1} \geq s_i + \Delta_i$ and therefore $t_{k_1} > t_{k_2}$; in both cases, we end up with a clear contradiction. A similar argument shows that $\max \mathcal{L}_{k_1} \leq \max \mathcal{L}_{k_2}$.

To prove Eqn. (16), we observe that, by Eqns. (18) and (19),

$$\big\{k \in \{2, \ldots, n\} \colon (\nexists \ell \in \{1, \ldots, m\}) \; \{t_{k-1}, t_k\} \subseteq {]s_\ell - \Delta_\ell, s_\ell + \Delta_\ell[}\big\}$$
$$= \big\{k \in \{2, \ldots, n\} \colon \max \mathcal{L}_{k-1} < \min \mathcal{L}_k\big\}$$
$$\subseteq \big\{k \in \{2, \ldots, n\} \colon \max \mathcal{L}_{k-1} < \max \mathcal{L}_k\big\}.$$

Since $1 \leq \max \mathcal{L}_1 \leq \cdots \leq \max \mathcal{L}_n \leq m$ by Eqn. (19), we infer from this that

$$\big|\big\{k \in \{2, \ldots, n\} \colon (\nexists \ell \in \{1, \ldots, m\}) \; \{t_{k-1}, t_k\} \subseteq {]s_\ell - \Delta_\ell, s_\ell + \Delta_\ell[}\big\}\big| \leq m - 1,$$

as required.

Finally, we prove the final part of the statement, so we assume that $\mathcal{S}$ is convex. Assume *ex absurdo* that there is an index $\ell^\star \in \{2, \ldots, m\}$ such that

$$s_{\ell^\star - 1} + \Delta_{\ell^\star - 1} < s_{\ell^\star} - \Delta_{\ell^\star}.$$

Fix any $t \in {]s_{\ell^\star - 1} + \Delta_{\ell^\star - 1}, s_{\ell^\star} - \Delta_{\ell^\star}[}$. Note that $\mathcal{S}$ includes $[s_{\ell^\star - 1}, s_{\ell^\star}]$ since $\mathcal{S}$ is closed and convex, so t belongs to $\mathcal{S}$. However, it follows from (ii) that $s_i + \Delta_i \leq s_{\ell^\star - 1} - \Delta_{\ell^\star - 1}$ for all $i \in \{1, \ldots, \ell^\star - 1\}$ and $s_{\ell^\star} - \Delta_{\ell^\star} \leq s_j - \Delta_j$ for all $j \in \{\ell^\star, \ldots, m\}$. From this, we infer that t does not belong to $\bigcup_{\ell=1}^m C_{s_\ell}$, which is a contradiction because $C_{s_1}, \ldots, C_{s_m}$ is a cover of $\mathcal{S}$. $\qquad\square$

The second one is an obvious observation about the expected number of jumps.

Lemma 23. *Consider a consistent collection $\mu_\bullet$ of finite-dimensional charges. Then for all $u, v \in \mathcal{U}$ such that $u \sqsubseteq v$,*

$$E_{\mu_u}(\hat{\eta}_u) \leq E_{\mu_v}(\hat{\eta}_v).$$

Proof. Let us enumerate the time points in u as $(r_1, \ldots, r_n)$ and in v as $(s_1, \ldots, s_m)$. It follows from Eqn. (4) and the consistency of $\mu_\bullet$ that

$$E_{\mu_u}(\hat{\eta}_u) = \sum_{k=2}^n \mu_u(\{x_u \in \mathcal{X}_u \colon x_{r_{k-1}} \neq x_{r_k}\}) = \sum_{k=2}^n \mu_v(\{y_v \in \mathcal{X}_v \colon y_{r_{k-1}} \neq y_{r_k}\}).$$

Let $\ell_1, \ldots, \ell_n$ be as defined in the proof of Lemma 16. Then as explained there, for all $k \in \{2, \ldots, n\}$,

$$\{y_v \in \mathcal{X}_v \colon y_{r_{k-1}} \neq y_{r_k}\} \subseteq \bigcup_{\ell = \ell_{k-1}+1}^{\ell_k} \{y_v \in \mathcal{X}_v \colon y_{s_{\ell-1}} \neq y_{s_\ell}\}.$$

Together with the (sub-)additivity of μ_v, this implies that

$$E_{\mu_u}(\hat{\eta}_u) \le \sum_{k=2}^{n} \sum_{\ell=\ell_{k-1}+1}^{\ell_k} \mu_v(\{y_v \in \mathcal{X}_v \colon y_{s_{\ell-1}} \neq y_{s_\ell}\})$$

$$\le \sum_{\ell=2}^{m} \mu_v(\{y_v \in \mathcal{X}_v \colon y_{s_{\ell-1}} \neq y_{s_\ell}\}).$$

The inequality in the statement follows immediately due to Eqn. (4). □

Proof of Proposition 21. To verify (R1), we fix any $t \in \mathcal{T} \cap \mathrm{rlims}(\mathcal{T})$ and $\epsilon \in \mathbb{R}_{>0}$. Fix any $n \in \mathbb{N}$ such that $|t| < n$. Then by the condition in the statement, there is some $\delta' \in \mathbb{R}_{>0}$ such that for all $r \in \mathcal{T} \cap]t, t+\delta'[$

$$\mu_{(t,r)}(\mathcal{X}_{\neq}^2) < (r-t)(\lambda_n + \epsilon).$$

Let $\delta := \min\{\delta', \epsilon/(\lambda_n + \epsilon)\}$. Then for all $r \in \mathcal{T} \cap]t, t+\delta[$

$$\mu_{(t,r)}(\mathcal{X}_{=}^2) = 1 - \mu_{(t,r)}(\mathcal{X}_{\neq}^2) > 1 - (r-t)(\lambda_n + \epsilon)$$
$$\ge 1 - \delta(\lambda_n + \epsilon) \ge 1 - \epsilon.$$

Since ϵ was an arbitrary strictly positive real number, we conclude that $\lim_{\mathcal{T} \ni r \searrow t} \mu_{(t,r)}(\mathcal{X}_{=}^2) = 1$, as required.

To verify (R2), we fix any $n \in \mathbb{N}$ such that $[-n,n] \cap \mathcal{T} \neq \varnothing$. Furthermore, we fix any $\epsilon \in \mathbb{R}_{>0}$. Then by the condition in the statement, for all $t \in \mathcal{T} \cap [-n,n]$ there is some $\delta_t \in \mathbb{R}_{>0}$ such that

$$(20) \qquad (\forall s, r \in \mathcal{T} \cap]t - \delta_t, t + \delta_t[\colon s < r) \; \mu_{(s,r)}(\mathcal{X}_{\neq}^2) < (r-s)(\lambda_n + \epsilon).$$

Because $\mathcal{T} \cap [-n,n]$ is closed by assumption and clearly bounded, it follows from Lemma 22 that there is some tuple of time points $(s_1, \ldots, s_m) \in \mathcal{U}_{[-n,n] \cap \mathcal{T}}$ such that

$$(21) \qquad \mathcal{T} \cap [-n,n] \subseteq \bigcup_{\ell=1}^{m}]s_\ell - \delta_{s_\ell}, s_\ell + \delta_{s_\ell}[$$

and, for all $u = (t_1, \ldots, t_k) \in \mathcal{U}_{[-n,n] \cap \mathcal{T}}$ with $k \ge 2$,

$$(22) \quad \left| \{i \in \{2, \ldots, k\} \colon (\nexists \ell \in \{1, \ldots, m\}) \; \{t_{i-1}, t_i\} \subseteq]s_\ell - \delta_{s_\ell}, s_\ell + \delta_{s_\ell}[\} \right| < m.$$

Fix any grid $u = (t_1, \ldots, t_k) \in \mathcal{U}_{[-n,n] \cap \mathcal{T}}$, and recall from Eqn. (5) that

$$E_{\mu_u}(\hat{\eta}_u) = \sum_{i=2}^{k} \mu_{(t_{i-1}, t_i)}(\mathcal{X}_{\neq}^2).$$

Let us split the index set $\{2, \ldots, k\}$ into two parts:

$$\mathcal{I}_1 := \{i \in \{2, \ldots, k\} \colon (\exists \ell \in \{1, \ldots, m\}) \; \{t_{i-1}, t_i\} \subseteq]s_\ell - \delta_{s_\ell}, s_\ell + \delta_{s_\ell}[\}$$

and $\mathcal{I}_2 := \{2, \ldots, k\} \setminus \mathcal{I}_2$.

On the one hand, it follows from Eqn. (20) that for all $i \in \mathcal{I}_1$,

$$\mu_{(t_{i-1}, t_i)}(\mathcal{X}_{\neq}^2) < (t_i - t_{i-1})(\lambda_n + \epsilon).$$

On the other hand, we know from Eqn. (22) that there are at most $m-1$ indices in $\mathcal{I}_2$, and for these indices i clearly

$$\mu_{(t_{i-1}, t_i)}(\mathcal{X}_{\neq}^2) \le 1 < (t_i - t_{i-1})(\lambda_n + \epsilon) + 1.$$

Thus, we find that

$$E_{\mu_u}(\hat{\eta}_u) = \sum_{i=2}^{k} \mu_{(t_{i-1},t_i)}\big(\mathcal{X}_{\neq}^2\big)$$

$$= \sum_{i\in\mathcal{I}_1} \mu_{(t_{i-1},t_i)}\big(\mathcal{X}_{\neq}^2\big) + \sum_{i\in\mathcal{I}_2} \mu_{(t_{i-1},t_i)}\big(\mathcal{X}_{\neq}^2\big)$$

$$\leq m - 1 + \sum_{i=2}^{k}(t_i - t_{i-1})(\lambda_n + \epsilon)$$

$$\leq m - 1 + (t_k - t_1)(\lambda_n + \epsilon).$$

This inequality holds for all $u = (t_1,\dots,t_k) \in \mathcal{U}_{[-n,n]\cap\mathcal{T}}$ – so with $-n \leq t_1 \leq t_k \leq n$ – and therefore

$$\sup\big\{E_{\mu_u}(\hat{\eta}_u)\colon u \in \mathcal{U}_{[-n,n]\cap\mathcal{T}}\big\} \leq m - 1 + 2n(\lambda_n + \epsilon) < +\infty.$$

A similar argument as the one in the second part of the proof of Proposition 12 shows that this implies (R2). □

For the third sufficient condition, we assume that for all $n \in \mathbb{N}$, $\mathcal{T} \cap [-n,n]$ is some finite union of intervals – bounded, closed and convex subsets of $\mathbb{T}$ – and we impose a condition on the limit superior of the rate of change of the probability of changing states. The argument in the second part of our proof generalises the proof of Lemma 5.49 in [12].

Proposition 24. *Suppose that for all $n \in \mathbb{N}$ such that $\mathcal{T} \cap [-n,n] \neq \varnothing$, there are time points $s_{n,1} \leq r_{n,1} < \cdots < s_{n,m_n} \leq r_{n,m_n}$ such that $\mathcal{T} \cap [-n,n] = \bigcup_{k=1}^{m_n}[s_{n,k},r_{n,k}]$. Consider a consistent collection $\mu_\bullet$ of finite-dimensional distributions. If for all $n \in \mathbb{N}$ such that $\mathcal{T} \cap [-n,n] \neq \varnothing$ there is some $\lambda_n \in \mathbb{R}_{\geq 0}$ such that for all $k \in \{1,\dots,m_n\}$,*

$$\limsup_{r\searrow t} \frac{\mu_{(t,r)}\big(\mathcal{X}_{\neq}^2\big)}{r-t} \leq \lambda_n \quad \text{for all } t \in [s_{n,k},r_{n,k}[$$

and

$$\limsup_{s\nearrow t} \frac{\mu_{(s,t)}\big(\mathcal{X}_{\neq}^2\big)}{t-s} \leq \lambda_n \quad \text{for all } t \in \,]s_{n,k},r_{n,k}],$$

then $\mu_\bullet$ is regular.

Proof. The proof for (R1) is almost exactly the same as the one in our proof for Proposition 21, although it does become a bit simpler. Fix any $t \in \mathcal{T}\cap\mathrm{rlims}(\mathcal{T})$ and some $n \in \mathbb{N}$ such that $|t| < n$. Then it follows from the condition in the statement that there are time points $s^\star, r^\star \in \mathcal{T}$ and a non-negative real number $\lambda_n \in \mathbb{R}_{\geq 0}$ such that $t \in [s^\star,r^\star[\subseteq \mathcal{T}$ and

$$\limsup_{r\searrow t} \frac{\mu_{(t,r)}\big(\mathcal{X}_{\neq}^2\big)}{r-t} \leq \lambda_n.$$

Clearly, this can only be the case if $\lim_{r\searrow t}\mu_{(t,r)}\big(\mathcal{X}_{\neq}^2\big) = 0$. Since $\mu_{(t,r)}\big(\mathcal{X}_{=}^2\big) = 1 - \mu_{(t,r)}\big(\mathcal{X}_{\neq}^2\big)$ for all $r \in \,]t,r^\star]$, we infer from this that

$$\lim_{r\searrow t}\mu_{(t,r)}\big(\mathcal{X}_{=}^2\big) = \lim_{r\searrow t} 1 - \mu_{(t,r)}\big(\mathcal{X}_{\neq}^2\big) = 1,$$

as required.

To verify (R2), we fix any $n \in \mathbb{N}$ such that $[-n,n] \cap \mathcal{T} \neq \varnothing$. Then by the conditions in the statement there are time points $s_{n,1} \leq r_{n,1} < \cdots < s_{n,m_n} \leq$

$r_{n,m_n} \in \mathcal{T}$ such that $[-n, n] \cap \mathcal{T} = \bigcup_{k=1}^{m_n} [s_{n,k}, r_{n,k}]$. Let $u = (t_1, \ldots, t_m)$ be an arbitrary tuple of time points in $\mathcal{U}_{[-n,n]\cap\mathcal{T}}$, and recall from Eqn. (5) that

$$\mu_u(\hat{\eta}_u) = \sum_{\ell=2}^{m} \mu_{(t_{\ell-1}, t_\ell)}(\mathcal{X}_{\neq}^2).$$

For all $\ell \in \{1, \ldots, m\}$, there is a unique index $k_\ell \in \{1, \ldots, m_n\}$ such that $t_\ell \in [s_{n,k_\ell}, r_{n,k_\ell}]$. It is obvious that $k_1 \leq \cdots \leq k_m$ because $t_1 < \cdots < t_m$, so the index set

$$\mathcal{L}_1 := \{\ell \in \{2, \ldots, m\} \colon k_{\ell-1} < k_\ell\}$$

contains at most $m_n - 1$ indices. For these $m_n - 1$ indices $\ell \in \mathcal{L}_1$, the most we can say is that

$$\mu_{(t_{\ell-1}, t_\ell)}(\mathcal{X}_{\neq}^2) \leq 1.$$

However, we can say more for those indices ℓ for which $t_{\ell-1}$ and t_ℓ belong to the same interval $[s_{n,k}, r_{n,k}]$; that is, for those indices in the index set

$$\mathcal{L}_2 := \{2, \ldots, m\} \setminus \mathcal{L}_1 = \{\ell \in \{2, \ldots, m\} \colon k_{\ell-1} = k_\ell\}.$$

To ease our notation, we fix two time points $s, r \in \mathcal{T}$ such that $s < r$ and $[s, r] \subseteq [s_{n,k}, r_{n,k}]$ for some $k \in \{1, \ldots, m_n\}$. We set out to show that

$$(23) \qquad \mu_{(s,r)}(\mathcal{X}_{\neq}^2) \leq (r - s)\lambda_n.$$

To this end, we recall from Eqn. (5) that

$$(24) \qquad \mu_{(s,r)}(\mathcal{X}_{\neq}^2) = E_{\mu_{(s,r)}}(\hat{\eta}_{(s,r)}).$$

By the assumptions in the statement, there is some $\lambda_n \in \mathbb{R}_{\geq 0}$ such that for all $t \in [s, r]$ and $\epsilon \in \mathbb{R}_{>0}$, there are positive real numbers $\delta_t^+, \delta_t^- \in \mathbb{R}_{>0}$ such that

$$(25) \qquad \left(\forall r' \in [s, r] \cap \,]t, t + \delta_t^+[\right) \frac{\mu_{(t,r')}(\mathcal{X}_{\neq}^2)}{r' - t} < \lambda_n + \epsilon$$

and

$$(26) \qquad \left(\forall s' \in [s, r] \cap \,]t - \delta_t^-, t[\right) \frac{\mu_{(s',t)}(\mathcal{X}_{\neq}^2)}{t - s'} < \lambda_n + \epsilon.$$

Fix any $\epsilon \in \mathbb{R}_{>0}$, and for all $t \in [s, r]$, let $\delta_t := \min\{\delta_t^+, \delta_t^-\}$. Since $[s, r]$ is clearly a closed, bounded and convex subset of $\mathcal{T}$, it follows from Lemma 22 that there is a tuple of time points $(s_1, \ldots, s_p) \in \mathcal{U}_{[s,r]}$ such that, with $\Delta_i := \delta_{s_i}$ for all $i \in \{1, \ldots, p\}$,

    (1) $[s, r] \subseteq \bigcup_{i=1}^{p}]s_i - \Delta_i, s_i + \Delta_i[$,
    (2) $s_i - \Delta_i < s_j - \Delta_j$ and $s_i + \Delta_i < s_j + \Delta_j$ for all $i, j \in \{1, \ldots, p\}$ such that $i < j$, and
    (3) $s_{i-1} + \Delta_{i-1} > s_i - \Delta_i$ for all $i \in \{2, \ldots, p\}$.

Let $v = (r_1, \ldots, r_q) \in \mathcal{U}_{[s,r]}$ be a tuple of time points that (i) starts in $r_1 = s$, (ii) ends in $r_q = r$, (iii) contains all the time points s_1, $\ldots$, s_p, and (iv) for all $i \in \{2, \ldots, p\}$, contains one time point in $]s_{i-1}, s_i[\cap \,]s_i - \Delta_i, s_{i-1} + \Delta_{i-1}[$ – to see that this intersection is non-empty, recall the three properties above. Note that $q = 2p + 1$ if $s < s_1$ and $s_p < r$, $q = 2p$ if either $s = s_1$ or $s_p = r$, and $q = 2p - 1$ if $s = s_1$ and $s_p = r$. Then by construction $(s, r) \sqsubseteq v$ and, due to Eqns. (25) and (26),

$$\mu_{(r_{i-1}, r_i)}(\mathcal{X}_{\neq}^2) < (r_i - r_{i-1})(\lambda_n + \epsilon) \quad \text{for all } i \in \{2, \ldots, q\}.$$

From this and Eqn. (5), we infer that

$$E_{\mu_v}(\hat{\eta}_v) = \sum_{i=2}^{q} \mu_{(r_{i-1}, r_i)}(\mathcal{X}_{\neq}^2) < \sum_{i=2}^{q}(r_i - r_{i-1})(\lambda_n + \epsilon) = (r - s)(\lambda_n + \epsilon).$$

From this, Eqn. (24) and Lemma 23, it now follows that

$$\mu_{(s,r)}(\mathcal{X}^2_{\neq}) = E_{\mu_{(s,r)}}(\hat{\eta}_{(s,r)}) \leq E_{\mu_v}(\hat{\eta}_v) < (r - s)(\lambda_n + \epsilon).$$

Since this inequality holds for arbitrary $\epsilon \in \mathbb{R}_{>0}$, we have proven Eqn. (23).

Let us return to where we were before. Due to Eqn. (23), we know that for all $\ell \in \mathcal{L}_2$,

$$\mu_{(t_{\ell-1}, t_\ell)}(\mathcal{X}^2_{\neq}) \leq (t_\ell - t_{\ell-1})\lambda_n.$$

For this reason, and because $\mu_{(t_{\ell-1}, t_\ell)}(\mathcal{X}^2_{\neq}) \leq 1$ for all $\ell \in \mathcal{L}_1 = \{2, \ldots, m\} \setminus \mathcal{L}_2$,

$$\begin{aligned}
\mu_u(\hat{\eta}_u) &= \sum_{\ell=2}^{m} \mu_{(t_{\ell-1}, t_\ell)}(\mathcal{X}^2_{\neq}) \\
&= \sum_{\ell \in \mathcal{L}_1} \mu_{(t_{\ell-1}, t_\ell)}(\mathcal{X}^2_{\neq}) + \sum_{\ell \in \mathcal{L}_2} \mu_{(t_{\ell-1}, t_\ell)}(\mathcal{X}^2_{\neq}) \\
&\leq \sum_{\ell \in \mathcal{L}_1} 1 + \sum_{\ell \in \mathcal{L}_2} (t_\ell - t_{\ell-1})\lambda_n \\
&\leq m_n - 1 + (t_m - t_1)\lambda_n \\
&\leq m_n - 1 + 2n\lambda_n,
\end{aligned}$$

where for the second inequality we used that $|\mathcal{L}_1| \leq m_n - 1$ and for the final inequality we used that $u \in \mathcal{U}_{[-n,n] \cap \mathcal{T}}$. This inequality holds for any $u = (t_1, \ldots, t_m) \in \mathcal{U}_{[-n,n] \cap \mathcal{T}}$, so we conclude that

$$\sup\{E_{\mu_u}(\hat{\eta}_u) \colon u \in \mathcal{U}_{[-n,n] \cap \mathcal{T}}\} \leq m_n - 1 + 2n\lambda_n < +\infty.$$

Here too, (R2) follows from the same argument as that in the second part of the proof of Proposition 12. $\qquad\square$

FOUNDATIONS LAB, GHENT UNIVERSITY, BELGIUM