Modern Stochastics: Theory and Applications 13 (3) (2026) 275–287
https://doi.org/10.15559/26-VMSTA293

Submartingale condition for weak convergence for semi-Markov processes

Vitaliy Golomoziy

Taras Shevchenko National University of Kyiv,
Faculty of Mechanics and Mathematics,
Department of Probability Theory, Statistics and Actuarial Mathematics
60, Volodymyrska Street,
City of Kyiv, Ukraine, 01033

vitaliy.golomoziy@knu.ua (V. Golomoziy)

Received: 21 September 2025, Revised: 27 December 2025, Accepted: 10 February 2026,
Published online: 24 February 2026

Abstract In this paper, we consider a modified version of a well-known submartingale condition for the weak convergence of probability measures, adapted to the semi-Markov case. In this setting, it is convenient to work with an embedded Markov chain and the filtration generated by jump times. We demonstrate that a straightforward restatement of the classical result is not valid, and that an additional condition is required.

Keywords Weak convergence of probability measures, semi-Markov processes, submartingale condition for weak convergnce, LaTeX
2010 MSC 60B10, 60K15

1 Introduction

The submartingale condition for weak convergence was introduced in the celebrated book [4], Theorem 1.4.6, and is stated as follows:

Theorem 1. *[D. Strook, S. Varadhan] Let* $\Omega = C([0,\infty);\mathbb{R}^d)$ *be the space of continuous functions on* $[0,\infty)$ *with values in* $\mathbb{R}^d$*, and let* $(\mathcal{M}_t)_{t\geq 0}$ *be the corresponding canonical filtration. Let* $\mathcal{P}$ *be a family of probability measures on* $(\Omega,(\mathcal{M}_t)_{t\geq 0})$*, such that for any non-negative* $f \in C_0^\infty(\mathbb{R}^d)$ *there exists a constant* A_f *such that, for all* $P\in\mathcal{P}$*, the stochastic process*

$$\big(f(x(t)) + A_f t, \mathcal{M}_t\big)_{t\geq 0}$$

© 2026 The Author(s). Published by VTeX. Open access article under the CC BY license.

is a non-negative P-submartingale. Assume also that A_f can be selected in such a way that it works for all translations of f (i.e. if $g(x) = f(x-a)$, $a \in \mathbb{R}^d$, then $A_g = A_f$). Assume further that

$$\lim_{c\to\infty} \sup_{P\in\mathcal{P}} P(|x(0)| \geq c) = 0. \tag{1}$$

Then the family $\mathcal{P}$ is weakly precompact (and thus tight).

This condition is useful in many situations, especially in connection with diffusion processes and the associated martingale problem. However, in the theory of semi-Markov processes we typically face a discrete-time martingale problem.

For instance, it is used as a main tool for establishing weak convergence in [2] (see Preface, page vii). Next, we give some basic definitions and recall the main facts about semi-Markov processes. We will use [2] as the main source, and we restrict our attention to processes with values in $\mathbb{R}^d$.

Definition 1. Function $Q\colon \mathbb{R}^d \times \mathcal{B} \times [0,\infty) \to [0,1]$, where $\mathcal{B}$ is a Borel sigma-field in $\mathbb{R}^d$, called a *semi-Markov kernel* on $(\mathbb{R}^d, \mathcal{B})$ if

- For every $x \in \mathbb{R}^d$ and $B \in \mathcal{B}$, function $Q(x, B, \cdot)$ is non-decreasing, right-continuous real function, such that $Q(x, B, 0) = 0$.
- For every $t \geq 0$, $Q(\cdot, \cdot, t)$ is a sub-Markov kernel on $(\mathbb{R}^d, \mathcal{B})$.
- $P(\cdot, \cdot) = Q(\cdot, \cdot, \infty)$ is a Markov kernel on $(\mathbb{R}^d, \mathcal{B})$.

Definition 2. An $\mathbb{R}^d$-valued *Markov renewal process* is a two-component, time-homogeneous Markov chain (x_n, τ_n), $n \geq 0$ taking values in $\mathbb{R}^d \times [0,\infty)$, with transition probability defined by a semi-Markov kernel Q as follows

$$\mathbb{P}(x_{n+1} \in B,\ \tau_{n+1} - \tau_n \leq t \mid \mathcal{G}_n) = Q(x_n, B, t),$$

for any integer $n \geq 0$, real $t \geq 0$, and Borel set $B \subset \mathbb{R}^d$. We assume that $\tau_0 = 0$. Here $(\mathcal{G}_n)_{n\geq 0}$ is the natural filtration generated by $\{(x_n, \tau_n) : n \geq 0\}$.

In what follows we assume that semi-Markov kernel $Q(x, B, t)$ admits a representation

$$Q(x, B, t) = P(x, B)F_x(t), \tag{2}$$

where $P(x, B)$ is a Markov kernel, and F_x is a distribution function for every x.

Thus, $\tau_{n+1} - \tau_n$ is a holding time, and when (2) holds, it has distribution F_x, where $x_n = x_{\tau_n} = x$. We denote its mean by

$$m(x) = \int_0^\infty t\, F_x(dt) = \int_0^\infty \big(1 - F_x(t)\big)\, dt. \tag{3}$$

Definition 3. A *semi-Markov process* associated with the Markov renewal process (x_n, τ_n), $n \geq 0$, is the stochastic process

$$x(t) = x_{\nu(t)}, \quad t \geq 0,$$

where

$$\nu(t) = \sup\{n \geq 0 : \tau_n \leq t\}, \quad t \geq 0,$$

is the counting process of jumps.

We also introduce the continuous version of τ_n by

$$\tau(s) = \tau_{\nu(s)},$$

and the continuous filtration generated by the semi-Markov process by

$$\mathcal{F}_t = \sigma\big(x(s), \tau(s), 0 \le s \le t\big).$$

Definition 4. Let $q(x) = 1/m(x)$. We define the compensating operator $\mathbb{L}$ of the Markov renewal process (x_n, τ_n), $n \ge 0$ (or of the associated semi-Markov process $x(t)$, $t \ge 0$) by

$$\mathbb{L}\varphi(x,t) = q(x)\left[\int_0^\infty F_x(ds) \int_{\mathbb{R}^d} P(x,dy)\,\varphi(y,t+s) - \varphi(x,t)\right],$$

where $\varphi(x,t)$ is a function from an appropriate class of test functions. By convention, we set $\mathbb{L}\varphi(x,t) = 0$ when $q(x) = 0$.

When $\varphi(x,t) = \varphi(x)$ does not depend on t, this expression reduces to

$$\mathbb{L}\varphi(x) = q(x)\left(\int_{\mathbb{R}^d} P(x,dy)\,\varphi(y) - \varphi(x)\right).$$

Finally, from Proposition 1.4 in [2], we know that the discrete-time process

$$Z_n^\varphi := \varphi(x_n,\tau_n) - \sum_{i=1}^n (\tau_i - \tau_{i-1})\,\mathbb{L}\varphi(x_{i-1},\tau_{i-1}), \quad n \ge 0,$$

is a martingale with respect to the discrete filtration $\mathcal{G}_n = \sigma(x_k, \tau_k : 0 \le k \le n)$, $n \ge 0$.

In applications, we typically use this fact to establish a discrete-time version of the submartingale condition of Theorem 1. Namely, we can establish the following:

Condition D. *Let $\mathcal{U}$ be an index set, and let $x^u(t)$, $u \in \mathcal{U}$, be a family of semi-Markov processes (with associated Markov renewal processes (x_n^u, τ_n^u), $n \ge 0$). Assume that condition* (1) *holds, and that for any non-negative $\varphi \in C_0^\infty(\mathbb{R}^d)$ there exists a constant $A_\varphi \ge 0$, the same for all translations of φ, such that for every $u \in \mathcal{U}$ the* ***discrete-time*** *process*

$$\big(\varphi(x_n^u) + A_\varphi \tau_n^u\big)_{n\ge 0}$$

is a non-negative submartingale with respect to its natural filtration $(\mathcal{G}_n^u)_{n\ge 0}$.

The question is whether **Condition D** is sufficient for tightness. The answer is no, as we show in Section 3, so an additional condition is required. We introduce and discuss this condition in Theorem 2. We will see that it is essential to ensure that there is no positive time interval during which no jumps occur for all semi-Markov processes in the family. In other words, the frequency of jumps must increase. We also show that in some important and typical applications this condition indeed holds.

This paper is organized as follows. Section 2 contains the main result, and Section 3 presents a counterexample that justifies the additional condition in Theorem 2. Finally, Section 4 is devoted to a special case—families of semi-Markov processes obtained by scaling a single given process in space and time—which are important examples in the theory of semi-Markov approximations.

2 Main result

In this section we assume that $\Omega = D[0,\infty)$ is the Skorokhod space equipped with the standard Borel σ-field $\mathcal{B}(\Omega)$ (with respect to the Skorokhod topology; see [1], Chapter 16 for details).

Note that any semi-Markov process defined through a Markov renewal process $(x_n, \tau_n)_{n\geq 0}$ has trajectories in a subspace Ω^* consisting of functions that have at most one accumulation point of jump times. In other words, $\omega \in \Omega^*$ if and only if there exists a non-decreasing sequence

$$0 \leq t_0 \leq t_1 \leq \cdots \leq t_n \leq \cdots$$

such that $\omega(t-) \neq \omega(t)$ if and only if $t \in \{t_n : n \geq 0\}$. For such $\omega \in \Omega^*$, we define a non-decreasing sequence of jump times by $\tau_n(\omega) = t_n$. For all other $\omega \in \Omega \setminus \Omega^*$, we set $\tau_n(\omega) = 0$. Since Ω^* is Borel-measurable, this construction yields a sequence of non-decreasing stopping times with respect to the natural coordinate filtration.

To simplify notation, we adopt the following convention. Let $(\xi_n)_{n\geq 0}$ be a sequence of random variables and let $\tau \geq 0$ be an integer-valued random variable. We will write $\xi_\tau(\omega)$ for $\xi_{\tau(\omega)}(\omega)$.

Theorem 2. *Let $(\zeta^u(t))_{t\geq 0}$, $u \in \{0,1,2,\ldots\}$, be a sequence of semi-Markov processes with values in $\mathbb{R}^d$, and let $\mathcal{P} = \{P^u\}_{u\geq 0}$ be the corresponding family of distributions on the Skorokhod space $(\Omega, \mathcal{B}(\Omega))$. Let $\{\tau_n(\omega) : n \geq 0\}$ be the non-decreasing sequence of jump times defined above, and define the corresponding value at jump time τ_n by*

$$X_n(\omega) = \omega(\tau_n(\omega)).$$

Assume the following conditions hold.

(i) For every $T > 0$,

$$\lim_{a\to\infty} \limsup_u P^u\left(\left\{\omega \in \Omega : \sup_{t\in[0,T]} \left|\omega(t)\right| \geq a\right\}\right) = 0. \tag{4}$$

(ii) Assume that for every non-negative $f \in C_0^\infty(\mathbb{R}^d)$ there exists a constant $A_f \geq 0$ such that the discrete-time process

$$\left(f(X_n) + A_f\tau_n\right)_{n\geq 0}$$

is a non-negative submartingale with respect to the filtration $\mathcal{G}_n = \sigma(X_k, \tau_k : 0 \leq k \leq n)$. Assume also that the choice of A_f can be made so that it works for all translates of f.

(iii) For any n and t, denote the next jump after $\tau_n + t$ by

$$\hat{\tau}_n(t) = \hat{\tau}_n(t;\omega) = \inf_{m>n}\{\tau_m(\omega) : \tau_m(\omega) > \tau_n(\omega) + t\}, \tag{5}$$

and define the conditional expectation of the time between $\tau_n + t$ and the next jump by

$$d_x^u(t) = E^u\left[\hat{\tau}_n(t) - \tau_n - t \,\middle|\, X_n = x\right], \quad x \in \mathbb{R}^d. \tag{6}$$

Assume that

$$\lim_{t\to 0} \limsup_{u\to\infty} \sup_{x\in\mathbb{R}^d} d_x^u(t) = 0. \tag{7}$$

Then the family $\mathcal{P}$ is tight in $D[0,\infty)$.

Remark 1. *Note that $d_x^u(t)$ does not depend on n, since the pair (X_n, τ_n) forms a time-homogeneous, discrete-time Markov chain, and $\hat{\tau}_n(t) - \tau_n$ is independent of τ_k, $k \le n$, and depends only on the state X_n.*

Proof. The proof follows the steps of the original proof of Theorem 1.4.6 from [4]. First, for a function $y \in \Omega = D[0,\infty)$ and $\delta > 0$, we define

$$w'_y(\delta; T) = \inf_{\{0=t_0<t_1<\cdots<t_n=T\}} \max_{1\le i\le n} \sup_{s,t\in[t_{i-1},t_i)} |y(t) - y(s)|,$$

where the infimum is taken over all sets $\{t_0, \dots, t_n\}$ such that $0 = t_0 < t_1 < \cdots < t_n = T$ (n is arbitrary) and $\min_{1\le i<n}(t_i - t_{i-1}) > \delta$ (see [1], p. 171 for details).

From [1], Theorem 16.8, we know that the family of distributions $\mathcal{P}$ is tight if and only if condition (4) holds, together with the following condition:

$$\lim_{\delta\to 0} \limsup_u P^u\left(\left\{y \in \Omega : w'_y(\delta; T) \ge \rho\right\}\right) = 0, \tag{8}$$

for all $\rho > 0$ and $T > 0$.

Thus, our goal is to prove (8). In what follows we assume that both T and ρ are fixed positive numbers. Following the proof of Theorem 1.4.6 from [4], we define for $\omega \in \Omega$:

$$s_0 = 0, \qquad s_n(\omega) = \inf\left\{t \ge s_{n-1}(\omega) : \left|\omega(t) - \omega(s_{n-1})\right| \ge \rho/4\right\}. \tag{9}$$

Denote

$$N = N(\omega) = \min\{n : s_{n+1}(\omega) > T\},$$

and

$$\Delta_\rho(T;\omega) = \min\{s_n(\omega) - s_{n-1}(\omega) : 1 \le n \le N(\omega)\}.$$

Next, we make a crucial observation: each P^u is concentrated on piecewise constant functions in $D[0,\infty)$, so the moments s_n necessarily coincide with some τ_m. Hence we can define $\nu_n(\omega)$ as the integer such that

$$s_n(\omega) = \tau_{\nu_n}(\omega),$$

where we use the convention $\tau_{\nu_n}(\omega) = \tau_{\nu_n(\omega)}(\omega)$ introduced at the beginning of this section. Note that ν_n is a stopping time with respect to the filtration $(\mathcal{G}_m)_{m\ge 0}$.

We then observe that

$$P^u\left(\left\{y \in \Omega : w'_y(\delta, T) \ge \rho\right\}\right) \le P^u\left(\Delta_\rho(T) \le \delta\right),$$

which follows directly from the definitions (the argument literally repeats the proof of Lemma 1.4.1 in [4]).

For each $\tilde{\omega} \in \Omega$, let $Q^u_{\tilde{\omega}}$ be a regular conditional probability,

$$Q^u_{\tilde{\omega}}(\cdot) = P^u(\cdot \mid \mathcal{G}_{\nu_n})(\tilde{\omega}),$$

whose existence is guaranteed by Theorems 1.1.6 and 1.1.8 in [4]. The corresponding expectation will be denoted by $E^u_{\tilde\omega}$, so that

$$E^u_{\tilde\omega}[\xi] = \int_\Omega \xi(\omega)\, Q^u_{\tilde\omega}(d\omega) = E^u\left[\xi \,\middle|\, \mathcal{G}_{\nu_n}\right](\tilde\omega). \tag{10}$$

Let us now choose $f \in C_0^\infty(\mathbb{R})$ such that $f(0) = 1$, $f(x) = 0$ for $|x| \geq \rho/4$, and $0 \leq f \leq 1$. Define

$$\tilde f(x) = \tilde f(x, \tilde\omega) = f\left(x - X_{\nu_n}(\tilde\omega)\right),$$

that is, a random translation of f by $X_{\nu_n}(\tilde\omega)$.

Let $\gamma_{n,\delta}$ be the index of the first jump after $\tau_{\nu_n} + \delta$, so that $\hat\tau_{\nu_n}(\delta) = \tau_{\gamma_{n,\delta}}$, where $\hat\tau_n(t)$ was defined in (5). Note that $\gamma_{n,\delta}$ is also a stopping time with respect to the filtration $(\mathcal{G}_m)_{m\geq 0}$.

Fix an arbitrary $q \in \mathbb{Q}^n$, an n-dimensional vector of rational numbers. Put $f_q(x) = f(x - q)$ and note that $A_{f_q} = A_f$ by the translation-invariance property of A_f (see condition (ii)).

Define

$$\kappa_{n,\delta} = \kappa_{n,\delta}(\omega) = \nu_{n+1}(\omega) \wedge \gamma_{n,\delta}(\omega).$$

Note that $\kappa_{n,\delta}$ is a $(\mathcal{G}_m)_{m\geq 0}$-stopping time and that $\kappa_{n,\delta} \geq \nu_n$ P^u-a.s. for all u. In what follows, we will write κ and γ to mean $\kappa_{n,\delta}$ and $\gamma_{n,\delta}$ respectively, as n and δ are fixed.

By Proposition IV.5.5 from [3] and condition (ii) of the theorem (namely, the submartingale property of the process $f(X_n) + A_f\tau_n$), we may conclude that there exists a null set $F_q \in \mathcal{B}(\Omega)$ such that for all $\omega' \notin F_q$,

$$E^u\left[f_q(X_\kappa) + A_f\tau_\kappa \,\middle|\, \mathcal{G}_{\nu_n}\right](\omega') \geq f_q\left(X_{\nu_n}(\omega')\right) + A_f\tau_{\nu_n}(\omega'). \tag{11}$$

Let us define the null set

$$F = \bigcup_q F_q,$$

and fix an arbitrary $\tilde\omega \notin F$. For this fixed $\tilde\omega$ and arbitrary $\varepsilon > 0$, we can find a rational vector $\tilde q = \tilde q(\varepsilon, \tilde\omega) \in \mathbb{Q}^n$ such that

$$\sup_{x\in\mathbb{R}^d}\left|\tilde f(x) - f_{\tilde q}(x)\right| = \sup_{x\in\mathbb{R}^d}\left|f\left(x - X_{\nu_n}(\tilde\omega)\right) - f(x - \tilde q)\right| < \varepsilon.$$

This is possible because f has compact support and is therefore uniformly continuous. It is clear that $\tilde q(\cdot, \varepsilon)$ is $\mathcal{G}_{\nu_n}$-measurable as a function of ω, since it is completely determined by $X_{\nu_n}(\omega)$.

Now, using (11), we obtain

$$\begin{aligned} E^u_{\tilde\omega}\left[\tilde f(X_\kappa) + A_f\tau_\kappa\right] &= E^u_{\tilde\omega}\left[\left(\tilde f(X_\kappa) - f_{\tilde q}(X_\kappa)\right) + f_{\tilde q}(X_\kappa) + A_f\tau_\kappa\right] \\ &\geq E^u_{\tilde\omega}\left[f_{\tilde q}(X_\kappa) + A_f\tau_\kappa\right] - \varepsilon \\ &\geq f_{\tilde q}\left(X_{\nu_n}(\tilde\omega)\right) + A_f\tau_{\nu_n}(\tilde\omega) - \varepsilon \\ &\geq \tilde f\left(X_{\nu_n}(\tilde\omega)\right) + A_f\tau_{\nu_n}(\tilde\omega) - 2\varepsilon, \end{aligned}$$

where $E^u_{\tilde\omega}$ is defined in (10).

Note that $\tilde{\omega}$ is fixed, so $\nu_n(\tilde{\omega})$ is a non-random integer. Also observe that $\tilde{f}(X_{\nu_n}(\tilde{\omega})) = f(0) = 1$. Finally, it is impossible that $\tau_{\nu_{n+1}} \in (\tau_{\nu_n} + \delta, \hat{\tau}_{\nu_n}(\delta))$, since $\tau_{\nu_{n+1}}$ is a jump time, and $\hat{\tau}_{\nu_n}(\delta)$ is the first jump time after $\tau_{\nu_n} + \delta$.

Thus, we have a.s.

$$E^u_{\tilde{\omega}}\left[\tilde{f}(X_\kappa) + A_f\left(\tau_\kappa - \tau_{\nu_n}(\tilde{\omega})\right)\right] \geq 1 - 2\varepsilon.$$

Hence,

$$\begin{aligned} E^u_{\tilde{\omega}}\left[1 - \tilde{f}(X_\kappa)\right] &\leq E^u_{\tilde{\omega}}\left[A_f\left(\tau_\kappa - \tau_{\nu_n}(\tilde{\omega})\right)\right] + 2\varepsilon \\ &\leq A_f\left(\delta + E^u_{\tilde{\omega}}\left[\hat{\tau}_{\nu_n}(\delta) - \tau_{\nu_n} - \delta\right]\right) + 2\varepsilon \\ &= A_f\delta + A_f d^u_{X_{\nu_n}(\tilde{\omega})}(\delta) + 2\varepsilon \\ &\leq A_f\left(\delta + \sup_x d^u_x(\delta)\right) + 2\varepsilon, \end{aligned}$$

where $d^u_x(\delta)$ is defined in (6), and we used the fact that

$$\tau_\kappa - \tau_{\nu_n} = \tau_{\nu_{n+1}\wedge\gamma} - \tau_{\nu_n} \leq \tau_\gamma - \tau_{\nu_n} = \hat{\tau}_{\nu_n}(\delta) - \tau_{\nu_n}.$$

Note that $0 \leq 1 - \tilde{f} \leq 1$.

Define

$$B_{\tilde{\omega}} = \left\{\omega \in \Omega : \tau_{\nu_{n+1}}(\omega) \leq \tau_{\nu_n}(\tilde{\omega}) + \delta\right\} \subset \mathcal{G}_{\nu_{n+1}}.$$

We know that for a regular conditional probability, $Q^u_{\tilde{\omega}}(C) = 1$ for $C \in \mathcal{B}(\Omega)$ if and only if $\tilde{\omega} \in C$ (see [4], p. 16). Put

$$C_{\tilde{\omega}} = \{\omega \in \Omega : \nu_n(\omega) = \nu_n(\tilde{\omega})\},$$

and note that $\tilde{\omega} \in C_{\tilde{\omega}}$.

Using this fact and the definition of f, we obtain

$$\begin{aligned} E^u_{\tilde{\omega}}\left[\tilde{f}(X_\kappa)\mathbb{1}_{B_{\tilde{\omega}}}\right] &= \int_{B_{\tilde{\omega}}} f\left(X_{\nu_{n+1}}(\omega) - X_{\nu_n}(\tilde{\omega})\right)\, Q^u_{\tilde{\omega}}(d\omega) \\ &= \int_{B_{\tilde{\omega}}\cap C_{\tilde{\omega}}} f\left(X_{\nu_{n+1}}(\omega) - X_{\nu_n}(\tilde{\omega})\right)\, Q^u_{\tilde{\omega}}(d\omega) \\ &= \int_{B_{\tilde{\omega}}\cap C_{\tilde{\omega}}} f\left(X_{\nu_{n+1}}(\omega) - X_{\nu_n}(\omega)\right)\, Q^u_{\tilde{\omega}}(d\omega) = 0, \end{aligned}$$

since

$$\left|X_{\nu_{n+1}}(\omega) - X_{\nu_n}(\omega)\right| \geq \rho/4.$$

Thus, we obtain

$$\begin{aligned} A_f\left(\delta + \sup_x d^u_x(\delta)\right) + 2\varepsilon &\geq E^u_{\tilde{\omega}}\left[(1 - \tilde{f}(X_\kappa))\left(\mathbb{1}_{B_{\tilde{\omega}}} + \mathbb{1}_{\Omega\setminus B_{\tilde{\omega}}}\right)\right] \\ &= Q^u_{\tilde{\omega}}(B_{\tilde{\omega}}) + E^u_{\tilde{\omega}}\left[\left(1 - f(X_\gamma)\right)\mathbb{1}_{\Omega\setminus B_{\tilde{\omega}}}\right] \\ &\geq Q^u_{\tilde{\omega}}(B_{\tilde{\omega}}), \end{aligned}$$

so we arrive at the inequality

$$A_f\left(\delta+\sup_x d_x^u(\delta)\right)+2\varepsilon\geq Q_{\tilde\omega}^u(B_{\tilde\omega})=P^u\left(\tau_{\nu_{n+1}}\leq\tau_{\nu_n}(\tilde\omega)+\delta\,\middle|\,\mathcal{G}_{\nu_n}\right)(\tilde\omega). \quad (12)$$

Next, for every $k>0$ we can write

$$\begin{aligned}P^u\left(\{\omega:\Delta_\rho(T;\omega)\leq\delta\}\right)&\leq P^u\left(\min_{1\leq i\leq k}\tau_{\nu_i}-\tau_{\nu_{i-1}}\leq\delta\right)+P^u(N>k)\\&\leq\sum_{i=1}^k P^u\left(\tau_{\nu_i}-\tau_{\nu_{i-1}}\leq\delta\right)+P^u(N>k)\\&\leq\sum_{i=1}^k E^u\left[P^u\left(\tau_{\nu_i}-\tau_{\nu_{i-1}}\leq\delta\,\middle|\,\mathcal{G}_{\nu_{i-1}}\right)\right]+P^u(N>k)\\&\leq k\left[A_f\left(\delta+\sup_x d_x^u(\delta)\right)+2\varepsilon\right]+P^u(N>k),\end{aligned}$$

where we used inequality (12) and the fact that A_f can be chosen so that it works for all translations of f.

Next, we write for any $t_0>0$:

$$\begin{aligned}E^u\left[e^{-(\tau_{\nu_{i+1}}-\tau_{\nu_i})}\,\middle|\,\mathcal{G}_{\nu_i}\right]&\leq P^u\left(\tau_{\nu_{i+1}}-\tau_{\nu_i}\leq t_0\,\middle|\,\mathcal{G}_{\nu_i}\right)+e^{-t_0}P^u\left(\tau_{\nu_{i+1}}-\tau_{\nu_i}>t_0\,\middle|\,\mathcal{G}_{\nu_i}\right)\\&\leq e^{-t_0}+(1-e^{-t_0})\,P^u\left(\tau_{\nu_{i+1}}-\tau_{\nu_i}\leq t_0\,\middle|\,\mathcal{G}_{\nu_i}\right)\\&\leq e^{-t_0}+(1-e^{-t_0})\left[A_f\left(t_0+\sup_x d_x^u(t_0)\right)+2\varepsilon\right],\end{aligned}$$

where we used (12).

From condition (iii) of the theorem we know that

$$\limsup_{u\to\infty}\sup_x d_x^u(t_0)\to 0,\qquad t_0\to 0,$$

so we can choose u_0, t_0 and ε such that

$$\lambda:=e^{-t_0}+(1-e^{-t_0})\left[A_f\left(t_0+\sup_x d_x^u(t_0)\right)+2\varepsilon\right]<1,$$

for all $u>u_0$.

Next, by Lemma 1.4.5 from [4], we conclude that for $u>u_0$,

$$P^u(N>k)\leq e^T\lambda^k,\qquad k\geq 0.$$

Thus, in order to prove the theorem, it remains to show that condition (8), namely

$$\lim_{\delta\to 0}\limsup_{u\to\infty}P^u\left(\{\omega:\Delta_\rho^u(T;\omega)\leq\delta\}\right)=0,$$

holds.

Indeed, we have

$$\begin{aligned}\lim_{\delta\to 0}\limsup_{u\to\infty} & P^u\left(\{\omega : \Delta^u_\rho(T;\omega) < \delta\}\right)\\ &\le \lim_{\delta\to 0}\limsup_{u\to\infty}\left(k\left[A_f(\delta + \sup_x d^u_x(\delta)) + 2\varepsilon\right] + P^u(N > k)\right)\\ &\le 2k\varepsilon + e^T\lambda^k,\end{aligned}$$

where in the last step we used condition (iii).

Now, we set $\varepsilon \to 0$ and then $k \to \infty$. Thus, condition (8) holds, and the theorem is proven. □

Remark 2. *Condition (iii) can be replaced with a somewhat opposite assumption:*

Condition (iv). *Assume that there exists a real number $a > 0$ such that for every $x \in \mathbb{R}^d$,*

$$\limsup_u P^u(\tau_{n+1} - \tau_n < a \mid X_n = x) = 0. \tag{13}$$

This condition means that asymptotically there are no jumps on the interval $(\tau_n, \tau_n{+}a)$. The proof of tightness is trivial in this case, since equation (8) takes the form (here we follow the notation from the proof)

$$\begin{aligned}\lim_{\delta\to 0}\limsup_{u\to\infty} & P^u\left(\{\omega : \Delta^u_\rho(T;\omega) \le \delta\}\right)\\ &\le \lim_{\delta\to 0}\limsup_{u\to\infty}\left(\sum_{i=1}^{k} E^u\left[P^u\left(\tau_{\nu_i} - \tau_{\nu_{i-1}} \le \delta \,\middle|\, \mathcal{G}_{\nu_{i-1}}\right)\right] + P^u(N > k)\right)\\ &= \limsup_{u\to\infty} P^u(N > k) \longrightarrow 0, \qquad k \to \infty.\end{aligned}$$

3 Discussion and examples

In the previous section we saw that each of Conditions (iii) or (iv) guarantees tightness (assuming the submartingale conditions (i) and (ii)). At the same time, it is clear that neither Condition (iii) nor Condition (iv) is sufficient on its own. The following example shows two things:

- These conditions cannot be omitted entirely;
- Tightness fails in situations where Conditions (iii) and (iv) are mixed in some sense.

Consider the sequence of semi-Markov processes $\zeta_n(t)$, $n \in \{1, 2, 3, \ldots\}$. Fix n. Let $\zeta_n(t)$ start at 0 and make a jump to 1 at one of two times, $1/n$ or 1, each with probability $1/2$. After that, $\zeta_n(t)$ remains equal to 1 forever.

To show that tightness does not hold, we use the same criterion from [1], Theorem 16.8, that was applied in the proof of Theorem 2. Theorem 16.8 states that a family of probability measures is tight if and only if both conditions (4) and (8) hold. We will show that condition (8) fails in this case.

Recall the definition of w' from the proof of Theorem 2. For $y \in D[0,\infty)$ and $\delta > 0$ define

$$w'_y(\delta;T) = \inf_{\{0=t_0<t_1<\cdots<t_n=T\}} \max_{1\le i\le n} \sup_{s,t\in[t_{i-1},t_i)} |y(t)-y(s)|,$$

where the infimum is taken over all partitions $\{t_0,\dots,t_n\}$ such that $0 = t_0 \le t_1 < \cdots < t_n = T$ (n arbitrary) and $\min_{1\le i<n}(t_i - t_{i-1}) > \delta$.

Our goal is to show that

$$\lim_{\delta\to 0} \limsup_{n\to\infty} P^n\big(\{y\in\Omega : w'_y(\delta;T) > \rho\}\big) > 0.$$

Indeed, fix arbitrary $\rho \in (0,1)$ and $\delta > 0$. We can choose n large enough so that $1/n < \delta$, which gives

$$\limsup_{n\to\infty} P^n\big(\{y\in\Omega : w'_y(\delta;T) > \rho\}\big) = \tfrac{1}{2}.$$

On the other hand, the submartingale property does hold. It suffices to check it at the first jump:

$$f(0) \le E^n[f(1) + A_f\tau_1] = f(1) + \tfrac{1}{2}\big(A_f/n + A_f\big),$$

which is true if $A_f = 4\sup|f|$.

Next, we demonstrate how Condition (iii) can be verified in a situation where a family of semi-Markov processes is obtained via a time-scale and space-scale transformation of a single process, which is typical in diffusion and averaging approximations.

4 Space-time scaled semi-Markov processes

Assume the semi-Markov process $x(t)$ is given in the sense of Definition 2, with associated Markov renewal process $(x_n,\tau_n)_{n\ge 0}$ and an associated family of holding-time distributions $\{F_x : x\in\mathbb{R}^d\}$. Let $\theta_n = \tau_{n+1} - \tau_n$. Note that θ_n depends only on x_n and does not depend on n or τ_n.

We then generate a family of processes $(\zeta^n(t))_{n\ge 1}$ defined by

$$\zeta^n(t) = \frac{x(a_n t)}{b_n}, \qquad t \ge 0,$$

where $\{a_n\}_{n\ge 1}$ and $\{b_n\}_{n\ge 1}$ are two increasing sequences of positive numbers such that $a_n \wedge b_n \to \infty$ as $n\to\infty$. Typical examples are $b_n = n$ and $a_n = n^2$ for diffusion schemes, or $a_n = n$ for averaging schemes (see [2] for details). We denote by (X_j^n,τ_j^n) a Markov renewal process that is associated with the semi-Markov process $(\zeta^n(t))_{t\ge 0}$.

It follows from the construction that all processes $(x(t))_{t\ge 0}$, $(x_j,\tau_j)_{j\ge 0}$, $(\zeta^n(t))_{t\ge 0}$, and $(X_j^n,\tau_j^n)_{j\ge 0}$ are defined on the same probability space. The next theorem gives a condition that allows one to verify condition (iii).

Theorem 3. *Let $(\zeta^n)_{n\ge 1}$ be the family of space-time-scaled semi-Markov processes defined above. Assume the following condition holds:*

$$\lim_{t\to 0} \limsup_{n\to\infty} \frac{1}{a_n} \sup_x \int_{a_n t}^{\infty} \frac{\bar F_x(r)}{\bar F_x(a_n t)}\, dr = 0, \tag{14}$$

where $\bar{F}_x = 1 - F_x$ is the tail distribution function associated with F_x. Then condition (iii) of Theorem 2 holds.

Proof. Let $P_{x,s}(\mathrm{d}y, \mathrm{d}t)$ be the transition probability of the Markov chain $(x_n, \tau_n)_{n\geq 0}$, and let $P^j_{x,s}(\mathrm{d}y, \mathrm{d}t)$ denote the corresponding j-step transition probability.

Consider the measurable space $(\Omega, \mathcal{F})$, where

$$\Omega = \left(\mathbb{R}^d \times [0, \infty)\right)^\infty$$

is a countable product space and $\mathcal{F}$ is the corresponding cylinder σ-field. Let

$$\omega = \left((\omega_{00}, \omega_{01}), (\omega_{10}, \omega_{11}), \ldots\right) \in \Omega.$$

Define the random variables

$$X_j(\omega) = \omega_{j0}, \qquad \tau_j(\omega) = \omega_{j1}, \qquad X^n_j = \frac{X_j}{b_n}, \qquad \tau^n_j = \frac{\tau_j}{a_n}.$$

For every pair $(x, s) \in \mathbb{R}^d \times [0, \infty)$ we may define a probability measure $\mathbb{P}_{x,s}$ on $(\Omega, \mathcal{F})$ such that the finite-dimensional distributions of the sequence $(X_j, \tau_j)_{j\geq 0}$ coincide with those of the original sequence $(x_j, \tau_j)_{j\geq 0}$ starting at (x, s). The same is true for the scaled sequence $(X^n_j, \tau^n_j)_{j\geq 0}$. In fact, not only the finite-dimensional distributions coincide, but also the distributions of the entire sequences as random elements of $(\Omega, \mathcal{F})$; however, for our purposes it suffices to work with finite-dimensional distributions.

Condition (iii) may then be restated as

$$\limsup_n \sup_{x,s} \mathbb{E}_{x,s}\left[\hat{\tau}^n_0(t) - s - t\right] \to 0, \quad t \to 0. \tag{15}$$

Moreover, it is enough to prove the convergence for $s = 0$

$$\limsup_n \sup_{x} \mathbb{E}_{x,0}\left[\hat{\tau}^n_0(t) - t\right] \to 0, \quad t \to 0, \tag{16}$$

where $\hat{\tau}^n_0(t) = \inf_{m>0}\{\tau^n_m : \tau^n_m > \tau^n_0 + t\}$.

In what follows we will simplify the notation by denoting

$$\mathbb{E}_x = \mathbb{E}_{x,0}, \; d^n_x(t) = \mathbb{E}_x\left[\hat{\tau}^n_0(t) - t\right].$$

We can write

$$\begin{aligned} d^n_x(t) = \mathbb{E}_x\left[\hat{\tau}^n_0(t) - t\right] &= \sum_{j=0}^{\infty} \mathbb{E}_x\left[\hat{\tau}^n_0(t) - t;\ \tau^n_j \leq t < \tau^n_{j+1}\right] \\ &= \sum_{j=0}^{\infty} \mathbb{E}_x\left[\tau^n_{j+1} - t;\ \tau^n_j \leq t < \tau^n_{j+1}\right] \\ &= \sum_{j=0}^{\infty} \mathbb{E}_x\left[\mathbb{E}_x\left[\tau^n_{j+1} - t;\ \tau^n_j \leq t < \tau^n_{j+1} \,\middle|\, X^n_j, \tau^n_j\right]\right] \end{aligned}$$

$$= \sum_{j=0}^{\infty} \mathbb{E}_x \left[\mathbb{E}_x \left[\theta_j^n - (t - \tau_j^n);\, \theta_j^n > (t - \tau_j^n) \,\middle|\, X_j^n, \tau_j^n \right] \mathbb{1}_{\{\tau_j^n \le t\}} \right],$$

where we used the relation $\theta_j^n = \tau_{j+1}^n - \tau_j^n$. Thus, by the Markov property,

$$\begin{aligned}
d_x^n(t) &= \sum_{j=0}^{\infty} \mathbb{E}_x \left[\mathbb{E}_{X_j^n, \tau_j^n} \left[\theta_0^n - (t - \tau_j^n);\, \theta_0^n > t - \tau_j^n \right] \mathbb{1}_{\{\tau_j^n \le t\}} \right] \\
&= \sum_{j=0}^{\infty} \mathbb{E}_x \left[\mathbb{E}_{X_j^n, \tau_j^n} \left[\frac{\theta_0}{a_n} - \left(t - \frac{\tau_j}{a_n} \right);\, \frac{\theta_0}{a_n} > \left(t - \frac{\tau_j}{a_n} \right) \right] \mathbb{1}_{\{\tau_j \le a_n t\}} \right] \\
&= \frac{1}{a_n} \sum_{j=0}^{\infty} \mathbb{E}_x \left[\mathbb{E}_{X_j^n, \tau_j^n} \left[\theta_0 - (a_n t - \tau_j);\, \theta_0 > a_n t - \tau_j \right] \mathbb{1}_{\{\tau_j \le a_n t\}} \right] \\
&= \frac{1}{a_n} \sum_{j=0}^{\infty} \int_{\mathbb{R}^d} \int_0^{a_n t} \mathbb{E}_{y/b_n} [\theta_0 - (a_n t - s);\, \theta_0 > a_n t - s]\, P_{x,0}^j(dy, ds),
\end{aligned}$$

where we used the equalities $X_j^n = x_j / b_n$ and $\tau_j^n = \tau_j / a_n$, together with the fact that θ_0 does not depend on τ_0. Hence we may omit the second subscript and write E_{y/b_n} instead of $E_{y/b_n, s/a_n}$.

Finally, since θ_0 is a nonnegative random variable, for every $z \in \mathbb{R}^d$

$$\begin{aligned}
&\mathbb{E}_z [\theta_0 - (a_n t - s);\, \theta_0 > a_n t - s] \\
&\quad = \left(\int_0^{\infty} \mathbb{P}_z (\theta_0 > r + (a_n t - s) \mid \theta_0 > a_n t - s)\, dr \right) \mathbb{P}_z (\theta_0 > a_n t - s) \\
&\quad = \left(\int_0^{\infty} \frac{1 - F_z(r + (a_n t - s))}{1 - F_z(a_n t - s)}\, dr \right) \left(1 - F_z(a_n t - s) \right) \\
&\quad = \left(\int_0^{\infty} \frac{\bar{F}_z(r + (a_n t - s))}{\bar{F}_z(a_n t - s)}\, dr \right) \bar{F}_z(a_n t - s).
\end{aligned}$$

Thus, we can continue

$$\begin{aligned}
d_x^n(t) &= \frac{1}{a_n} \sum_{j=0}^{\infty} \int_{\mathbb{R}^d} \int_0^{a_n t} \int_0^{\infty} \frac{\bar{F}_{y/b_n}(r + a_n t - s)}{\bar{F}_{y/b_n}(a_n t - s)}\, dr\, \bar{F}_{y/b_n}(a_n t - s)\, P_{x,0}^j(dy, ds) \\
&= \frac{1}{a_n} \sum_{j=0}^{\infty} \int_{\mathbb{R}^d} \int_0^{a_n t} \int_{a_n t - s}^{\infty} \frac{\bar{F}_{y/b_n}(r)}{\bar{F}_{y/b_n}(a_n t - s)}\, dr\, \bar{F}_{y/b_n}(a_n t - s)\, P_{x,0}^j(dy, ds) \\
&\le \left(\frac{1}{a_n} \sup_y \int_{a_n t}^{\infty} \frac{\bar{F}_y(r)}{\bar{F}_y(a_n t)}\, dr \right) \sum_{j=0}^{\infty} \int_{\mathbb{R}^d} \int_0^{a_n t} \bar{F}_{y/b_n}(a_n t - s)\, P_{x,0}^j(dy, ds) \\
&= \left(\frac{1}{a_n} \sup_y \int_{a_n t}^{\infty} \frac{\bar{F}_y(r)}{\bar{F}_y(a_n t)}\, dr \right) \sum_{j=0}^{\infty} \mathbb{E}_x \left[\mathbb{P}_{X_j, \tau_j} (\theta_0 > a_n t - \tau_j)\, \mathbb{1}_{\{\tau_j \le a_n t\}} \right] \\
&= \left(\frac{1}{a_n} \sup_y \int_{a_n t}^{\infty} \frac{\bar{F}_y(r)}{\bar{F}_y(a_n t)}\, dr \right) \sum_{j=0}^{\infty} \mathbb{E}_x \left[\mathbb{P}_x (\tau_{j+1} > a_n t \mid X_j, \tau_j)\, \mathbb{1}_{\{\tau_j \le a_n t\}} \right]
\end{aligned}$$

$$= \left(\frac{1}{a_n} \sup_y \int_{a_n t}^{\infty} \frac{\bar{F}_y(r)}{\bar{F}_y(a_n t)}\, dr \right) \sum_{j=0}^{\infty} \mathbb{E}_x \left[\mathbb{P}_x(\tau_j \le a_n t < \tau_{j+1}) \right]$$

$$= \frac{1}{a_n} \sup_y \int_{a_n t}^{\infty} \frac{\bar{F}_y(r)}{\bar{F}_y(a_n t)}\, dr.$$

Thus,

$$\lim_{t\to 0} \limsup_{n\to\infty} \sup_x d_x^n(t) = \lim_{t\to 0} \limsup_{n\to\infty} \frac{1}{a_n} \sup_x \int_{a_n t}^{\infty} \frac{\bar{F}_x(r)}{\bar{F}_x(a_n t)}\, dr = 0,$$

by condition (14). □

Remark 3. *As an example, consider the case when $F_x = F$ does not depend on x (so that all holding times are identically distributed). In this situation, condition* (14) *holds if the tails of F are exponential or polynomial.*

Indeed, if

$$c_1 e^{-\alpha u} \le \bar{F}(u) \le c_2 e^{-\alpha u},$$

for some constants $c_1, c_2 > 0$, and $\alpha > 0$ then

$$\frac{1}{a_n} \int_{a_n t}^{\infty} \frac{\bar{F}(r)}{\bar{F}(a_n t)}\, dr \le \frac{c_2}{c_1} \frac{1}{a_n} \int_{a_n t}^{\infty} e^{-\alpha r + \alpha a_n t}\, dr = \frac{c_2}{\alpha c_1} \frac{1}{a_n} \to 0, \quad n \to \infty.$$

If instead for $c_1, c_2 > 0$ and $\alpha > 1$

$$c_1 u^{-\alpha} \le \bar{F}(u) \le c_2 u^{-\alpha},$$

then

$$\frac{1}{a_n} \int_{a_n t}^{\infty} \frac{\bar{F}(r)}{\bar{F}(a_n t)}\, dr \le \frac{c_2 (a_n t)^{\alpha}}{c_1 a_n} \int_{a_n t}^{\infty} r^{-\alpha} dr = \frac{c_2}{(\alpha - 1) c_1} t \to 0, \quad t \to 0.$$

References

[1] Billingsley, P.: Convergence of Probability Measures. Second Edition. John Wiley and Sons, Inc., New York (1999) MR1700749. https://doi.org/10.1002/9780470316962

[2] Koroliuk, V., Limnios, N.: Stochastic Systems in Merging Phase Space. World Scientific Publishing Co. Pte. Ltd., Singapore (2005) MR2205562. https://doi.org/10.1142/9789812703125

[3] Neveu, J.: Mathematical Foundations of the Calculus of Probability. Holden-Day (1965) MR0198505

[4] Strook, D., Varadhan, S.: Multidimensional Diffusion Processes. Springer, Berlin Heidelberg (2006) MR2190038