Electronic Journal of Probability

Electron. J. Probab. **28** (2023), article no. 19, 1–46.
ISSN: 1083-6489 https://doi.org/10.1214/23-EJP902

The martingale problem method revisited*

David Criens[†] Peter Pfaffelhuber[‡] Thorsten Schmidt[§]

Abstract

We use the abstract method of (local) martingale problems in order to give criteria for convergence of stochastic processes. Extending previous notions, the formulation we use is neither restricted to Markov processes (or semimartingales), nor to continuous or càdlàg paths. We illustrate our findings both, by finding generalizations of known results, and proving new results. For the latter, we work on processes with fixed times of discontinuity.

Keywords: martingale problem; path space; fixed times of discontinuity; limit theorems; semimartingales; weak-strong convergence; stable convergence; Skorokhod topology; local uniform topology; Volterra equations.
MSC2020 subject classifications: Primary 60G07, Secondary 60F17; 60H15; 60G17.
Submitted to EJP on August 26, 2021, final version accepted on January 11, 2023.

Contents

*DC acknowledges financial support from the DFG project No. SCHM 2160/15-1.
[†]Albert-Ludwigs University of Freiburg, Germany. E-mail: david.criens@stochastik.uni-freiburg.de
[‡]Albert-Ludwigs University of Freiburg, Germany.
E-mail: peter.pfaffelhuber@stochastik.uni-freiburg.de
[§]Albert-Ludwigs University of Freiburg, Germany.
E-mail: thorsten.schmidt@stochastik.uni-freiburg.de

1 Introduction

1.1 Background

This article deals with one of the classical questions in probability theory: Limit theorems for stochastic processes. Starting with the work of Prokhorov [32], limit theorems are formulated via weak convergence of probability measures on function spaces, such as the Wiener or the Skorokhod space. Prokhorov's method for proving weak convergence consists of three steps: verifying tightness, showing the convergence of the finite dimensional distributions and establishing that these determine the limit. The first part is well-studied and the final part is essentially trivial. In contrast, verifying convergence of the finite dimensional distributions is often hard and sometimes even impossible. This difficulty motivated the necessity to develop more techniques for proving limit theorems. One of the most successful strategies is the *martingale problem method* initiated by Stroock and Varadhan [37]. Instead of studying the finite dimensional distributions, the idea is to deduce the martingale property of certain test processes from weak convergence and to show that these martingale properties characterize the limiting law uniquely.

Originally, Stroock and Varadhan developed their method for Itô diffusions, see Example 3.7 below. Later, Ethier and Kurtz [12] generalized it to a Markovian framework with Polish state space E, i.e. they considered a martingale problem described by test processes of the type

$$f(X) - \int_0^{\cdot} g(X_s)ds, \quad (f, g) \in A \subset C_b(E) \times C_b(E), \tag{1.1}$$

where A is sometimes referred to as the (pre-)generator of X, see Example 3.8 below. In the spirit of the general theory of stochastic processes, Jacod and Shiryaev [25] extended the martingale problem method to the class of semimartingales, which have not necessarily Markovian dynamics, by relating martingale properties to the characteristics of the semimartingale.

Recently, there is growing interest in processes which are not covered by [12, 25]. Examples of such are solutions to stochastic partial differential equations (SPDEs) and Volterra equations (VSDEs), or certain processes with fixed times of discontinuity. The latter generate increasing attention in applied probability (see, e.g. [3, 29]) and in finance (see, e.g. [7, 14, 17, 30]). To give a concrete example, fixed times of discontinuity arise naturally in the context of random environments via the quenched perspective, i.e. when the random environment is fixed. Here, we think of stochastic equations of the type

$$dX_t = \sigma(X_{t-})dL_t + \gamma(X_{t-})dS_t,$$

where σ and γ are suitable coefficients, and L and S are independent semimartingales with jumps. The random environment is represented by one of the drives, say S. Roughly speaking, fixing the environment means to freeze a path of S which gives rise to fixed times of discontinuity.

Limit theorems for certain types of SPDEs and VSDEs were proved in [1, 8, 35]. In contrast, for general processes with fixed times of discontinuity we are not aware of any systematic study. The natural state space for distributions of stochastic processes with discontinuities is the Skorokhod space of càdlàg functions endowed with the Skorokhod J_1 topology (which turns it into a Polish space). For this natural setup the presence of fixed times of discontinuity turns out to be a major difficulty, because typical test processes then lack continuity (in the Skorokhod J_1 topology). Below we will explain this issue in more detail.

1.2 Purpose of the current article

The aim of this article is to develop a general version of the martingale problem method which on one side is flexible enough to cover existing convergence results and on the other side can be used to establish new results, e.g. for processes with fixed times of discontinuity. We generalize the three main ingredients in the standard theory: (i) the state space, (ii) the set of test processes, i.e. we allow different test processes as for instance in (1.1), and (iii) we work with an extended type of weak convergence called *weak-strong convergence* in the sequel, see, e.g. [2, 23, 34] or Section 2 below.

Let us motivate the novelties (i), (ii) and (iii) in some detail. Almost all general results in the literature are formulated for continuous or càdlàg processes. However, some processes of recent interest have less regular paths. For example, the Volterra processes studied in [1] only have paths in (local) L^p spaces. To include these, and more general cases, we work with the minimal assumption that paths can be viewed as random variables in some Polish path space.

Next, we comment on our extension of the set of test processes. In the classical literature on the martingale problem method for Itô diffusions ([37]), Markov processes ([12]) or semimartingales ([25]), the test processes only depend on the paths of the process X whose dynamics should be described by the martingale problem. For instance, in the classical monograph of Ethier and Kurtz [12] the test processes are of the form (1.1). For some applications it turns out to be useful to introduce a second variable L, which we call *control variable*. In a generalized version of the Ethier and Kurtz [12] setting, for instance, this leads to test processes of the form

$$f(t, L, X_t) - \int_0^t g(s, L, X_s)ds, \quad t \in \mathbb{R}_+, \tag{1.2}$$

where f and g are suitable functions such that this process is well-defined. Let us emphasis that we do not stick to processes of this type but allow more flexibility. One reason to introduce L is to include stochastic equations with random coefficients into our framework. A natural example is the class of *change point processes* of the type

$$dX_t = (\sigma(X_t)\mathbb{1}_{\{t\le\tau\}} + \gamma(X_t)\mathbb{1}_{\{t>\tau\}})dW_t,$$

where W is a Brownian motion, σ and γ are suitable measurable functions and τ is an exogenous random time. In financial settings the random time τ could be interpreted as the time of interest rate adjustments, the default time of a major financial institution or the release time of political news, see [13] for applications of such processes in a financial context. Due to the additional randomness given through τ, the dynamics of X cannot be described by a martingale problem with test processes of the type (1.1). However, they can be described by test processes of the type (1.2) when L either is taken to be τ or to be the identity on the underlying probability space. The latter choice of L is even flexible enough to derive limit theorems for *arbitrary* semimartingales which are defined on the same stochastic basis, irrespective whether their characteristics are canonical (in the sense of [21]) or not.

Besides this extension, we also generalize the mode of convergence in the martingale problem method. More precisely, next to the convergent sequence $X^1, X^2, \dots$, we introduce an auxiliary sequence $L^1, L^2, \dots$ of *control variables* such that the bivariate sequence $(X^1, L^1), (X^2, L^2), \dots$ converges in the weak-strong sense. Roughly speaking, for weak-strong convergence, the first variable converges in the usual weak topology, while the second variable has to convergence setwise. We add the sequence $L^1, L^2, \dots$ for two reasons. First, it is natural that the new control variable L in the limiting martingale problem should also get reflected by the approximation sequence. This is, for instance, useful when we consider sequences of processes with random coefficients. Second, the control variables $L^1, L^2, \dots$ are useful technical tools to overcome the limitation that certain types of test processes on the Skorokhod space are not continuous in the Skorokhod J_1 topology. To see where continuity might gets lost, consider a natural generalization of (1.1) when fixed time discontinuities are present:

$$f(X_t) - \int_0^t g(X_s)ds - \int_0^t h(X_{s-})q(ds), \quad t \in \mathbb{R}_+, \tag{1.3}$$

where f, g, h are bounded and continuous and q is some deterministic locally finite measure. As projections to fixed times are not continuous in the Skorokhod J_1 topology, these random variables are not necessarily continuous when q has point masses. To overcome this problem we would like to relax the continuity assumption by replacing the Skorokhod J_1 topology with a stronger topology (in which more functions are continuous). At the same time we might not want to drop the Skorokhod J_1 topology in general, as, for instance, tightness is well-studied for this topology. At this stage the sequence $L^1, L^2, \dots$ and the concept of weak-strong convergence come into play. An important and powerful result for weak-strong convergence is the continuous mapping theorem of Jacod and Mémin [23] (see Theorem 2.9) where it suffices to ask for continuity conditions when the values of the controls are fixed. Roughly speaking, this means that we can restrict our attention to a (randomized) subset of the underlying space on which continuity holds conditionally. If we have enough control on the jump times of the sequence $X^1, X^2, \dots$ via the controls $L^1, L^2, \dots$, this allows us, for the continuity assumptions on the test processes, to replace the Skorokhod J_1 topology by the local uniform topology. This is a major improvement, as, for instance, processes of the form (1.3) are clearly continuous in the local uniform topology. Using this strategy we generalize convergence theorems from the monographs [12, 25] to a setting with fixed times of discontinuity. We learned about the idea to replace the Skorokhod J_1 topology by the local uniform topology via weak-strong convergence from proofs of Jacod and Mémin [22, 24] for certain existence and stability results for stochastic differential equations driven by semimartingales.

1.3 Main contributions and structure of the article

Let us summarize our contributions. The main abstract results are the Theorems 3.14 and 3.20. In the former, we show the martingale property of test processes of the limiting martingale problem directly and in the latter, we verify the martingale property using approximating sequences of martingales.

To illustrate applications of our theory we discuss a variety of examples. First of all, we show that our results cover, or even extend, several known limit theorems. More precisely, we reprove a classical theorem of Ethier and Kurtz [12] (see Section 4.1) and the stability result for Volterra processes from [1] (see Section 4.2). Furthermore, we localize conditions by Jacod and Shiryaev [25] which identify a weak limit as a semimartingale via its characteristics (see Section 4.3). We think that this extension is of interest for future applications.

Besides recovering results from the literature, we also present new results. First, we prove a version of the Ethier–Kurtz theorem for test processes of the type (1.3) and we present a tightness condition which is tailored to such processes, see Section 5.3. Second, we derive a stability result for semimartingales under a continuity assumption on the characteristics in the local uniform topology, see Section 5.5.1. As the latter is stronger than the classical Skorokhod J_1 topology, our result has a different scope than its counterpart from [25]. Furthermore, in Section 5.5.3 we specify our results to the annealed case where all processes are defined on the same probability space and the limit is allowed to have characteristics which also depend on the underlying space. Finally, in Section 5.5.4 we present an application to Itô processes with fixed times of discontinuity.

1.4 Notation

In our paper, we use the following notation.

- An inequality up to a multiplicative positive constant is denoted by $\lesssim$.
- The extended real line is denoted by $\overline{\mathbb{R}} \triangleq \mathbb{R} \cup \{\pm\infty\}$. The set of symmetric non-negative definite real-valued $d \times d$ matrices is denoted by $\mathbb{S}^d_+$.
- The Lebegue measure is denoted by $\lambda\!\!\!\lambda$.
- For a topological space E we write $C(E)$ for the space of continuous functions $E \to \mathbb{R}$, $B(E)$ for the space of bounded Borel functions $E \to \mathbb{R}$ and $C_b(E) \triangleq C(E) \cap B(E)$.
- For $p \geq 1$ and a Banach space $(E, \|\cdot\|)$ we denote by $L^p_{\text{loc}}(\mathbb{R}_+, E)$ the space of equivalence classes of locally p-integrable functions from $\mathbb{R}_+$ into E, i.e. of (strongly) measurable functions $f\colon \mathbb{R}_+ \to E$ such that $\int_0^t \|f(s)\|^p ds < \infty$ for all $t > 0$. We endow $L^p_{\text{loc}}(\mathbb{R}_+, E)$ with the local L^p-norm topology.
- We write $C^2(\mathbb{R}^d)$ for the space of twice continuously differentiable functions $\mathbb{R}^d \to \mathbb{R}$, $C^2_c(\mathbb{R}^d)$ for its subspace of functions with compact support, and $C^2_b(\mathbb{R}^d)$ for the set of bounded functions $f \in C^2(\mathbb{R}^d)$ with bounded gradient ∇f and bounded Hessian matrix $\nabla^2 f$.
- For an operator σ we write $\operatorname{tr}(\sigma)$ for its trace and σ^* for its adjoint.
- For a Polish space E we denote the space of continuous functions $\mathbb{R}_+ \to E$ by $\mathbb{C}(E)$ and the space of càdlàg functions $\mathbb{R}_+ \to E$ by $\mathbb{D}(E)$.
- On a space F the identity is denoted by $\mathsf{X}\colon F \to F$. In case $F = \mathbb{C}(E)$ or $\mathbb{D}(E)$, the identity X is called *coordinate process*.
- For a càdlàg process $Z = (Z_t)_{t \geq 0}$ we write $\Delta Z_t \triangleq Z_t - Z_{t-}$ for its time t jump. Moreover, we denote the quadratic variation process by $[\cdot, \cdot]$.
- For an integer-valued random measure $\mathfrak{p}$ with compensator $\mathfrak{q}$ and a suitable measurable function $H = H(\omega, t, y)$ we write

$$H * \mathfrak{p}_t \triangleq \int_0^t \int H(s, y) \mathfrak{p}(ds, dy), \quad t \in \mathbb{R}_+,$$

and $H * (\mathfrak{p} - \mathfrak{q})$ for the integral process of H w.r.t. the compensated random measure $\mathfrak{p} - \mathfrak{q}$, cf. [25, Section II.1.d]. Furthermore, we denote by $G_{\mathrm{loc}}(\mathfrak{p})$ the set of functions which are integrable w.r.t. $\mathfrak{p} - \mathfrak{q}$, see [25, Definition II.1.27]. For a semimartingale $Z = (Z_t)_{t \geq 0}$ we denote the set of Z-integrable processes by $L(Z)$, cf. [25, Section III.6]. For all unexplained terminology related to the *general theory of stochastic processes* we refer to [25, Chapter I].

2 Weak-strong convergence

In this section we recall the notion of *weak-strong convergence* of probability measures on a product space, which was introduced in [34] and systematically studied in [2, 23, 34], see also [22, 24]. This type of convergence will be crucial for our treatment of the martingale problem method.

2.1 Definition and first properties

Let $(U, \mathcal{U})$ be a measurable space and let $(F, \mathcal{B}(F))$ be a Polish space with its Borel σ-field. We define the product space $S \triangleq U \times F$ and the corresponding product σ-field $\mathcal{S} \triangleq \mathcal{U} \otimes \mathcal{B}(F)$. Let C_S be the set of bounded $\mathcal{S}/\mathcal{B}(\mathbb{R})$-measurable functions $f\colon S \to \mathbb{R}$ such that $\omega \mapsto f(\alpha, \omega)$ is continuous (as a function on F) for every $\alpha \in U$.

Definition 2.1. *Let $P, P^1, P^2, \dots$ be probability measures on $(S, \mathcal{S})$. We say that the sequence $(P^n)_{n \in \mathbb{N}}$* converges in the weak-strong sense *to P, written $P^n \to_{ws} P$, if*

$$E^{P^n}[f] \to E^P[f] \text{ as } n \to \infty \text{ for all } f \in C_S.$$

Remark 2.2. (i) Let $P, P^1, P^2, \dots$ be probability measures on F, let U be a singleton and extend $P, P^1, P^2, \dots$ to the product space $S = U \times F$ in the obvious manner. Then, it is clear that $P^n \to_{ws} P$ if and only if $P^n \to P$ weakly in the usual sense. This simple observation explains that weak-strong convergence is a natural extension of the usual weak convergence with an additional *control variable*. In particular, any strategy to identify weak-strong limits is also a strategy to identify weak limits in the usual sense.

(ii) Weak-strong convergence has a close relation to the notion of *stable convergence*, which is more commonly known in the probability literature, see [25, Section VIII.5.c]. To be more precise, if E is a Polish space and $(\Omega', \mathcal{F}', P')$ is a probability space which supports E-valued random variables $Z^1, Z^2, \dots$, then $(Z^n)_{n \in \mathbb{N}}$ converges *stably* (in the sense of [25, Definition VIII.5.28]) if and only if the sequence

$$P_n(d\omega, dz) \triangleq \delta_{Z^n(\omega)}(dz) P'(d\omega), \quad n \in \mathbb{N}, \tag{2.1}$$

converges in the weak-strong sense as a probability measure on $(\Omega' \times E, \mathcal{F}' \otimes \mathcal{B}(E))$, see also [23, Proposition 2.4]. In certain cases stable (and therefore also weak-strong) convergence is equivalent to convergence in probability. More precisely, the sequence $(P_n)_{n \in \mathbb{N}}$ as given in (2.1) converges in the weak-strong sense to

$$P(d\omega, dz) = \delta_{Z(\omega)}(dz) P'(d\omega)$$

if and only if $Z^n \to Z$ in probability, see [23, Proposition 3.5].

(iii) Let us provide another point of view on weak-strong convergence. Take a sequence $P_1, P_2, \dots$ of probability measures on $(S, \mathcal{S})$ with the same U-marginal μ. It is well-known that there exist transition kernel $K_1, K_2, \dots$ such that

$$P_n(du, df) = K_n(u, df) \mu(du), \quad n \in \mathbb{N}.$$

Then, the sequence $P_1, P_2, \dots$ converges in the weak-strong sense if and only if for every $f \in C_b(F)$ the sequence $K_1 f, K_2 f, \dots$ converges weakly (in the Banach space sense) in $L^1(U,\mathcal{U},\mu)$ (cf. also [23, Proposition 2.4]).

Let $M_{mc}(S)$ be the space of all probability measures on $(S,\mathcal{S})$ endowed with the weakest topology such that the map $P \mapsto E^P[f]$ is continuous for every $f \in C_S$. The acronym mc reflects the structure of the functions from C_S, as they are measurable in the first and continuous in the second variable. Of course, $P^n \to P$ in $M_{mc}(S)$ if and only if $P^n \to_{ws} P$.

Let $M_m(U)$ be the space of probability measures on $(U,\mathcal{U})$ endowed with the weakest topology such that the map $P \mapsto E^P[f]$ is continuous for every bounded $\mathcal{U}/\mathcal{B}(\mathbb{R})$-measurable function $f \colon U \to \mathbb{R}$.

Remark 2.3. By [23, Proposition 2.4], the above topology on $M_m(U)$ is also the weakest topology such that the map $P \mapsto P(A)$ is continuous for every $A \in \mathcal{U}$. In other words, it is the topology of *setwise convergence*.

Furthermore, let $M_c(F)$ be the space of all probability measures on $(F,\mathcal{B}(F))$ endowed with the usual weak topology, i.e. the weakest topology such that the map $P \mapsto E^P[f]$ is continuous for every bounded continuous function $f \colon F \to \mathbb{R}$.

For $P \in M_{mc}(S)$ we write P_U for its U-marginal and P_F for its F-marginal, i.e.

$$P_U(du) \triangleq P(du \times F), \qquad P_F(df) \triangleq P(U \times df).$$

To get a better understanding of the concept of weak-strong convergence, we recall the following result, which shows that weak-strong and weak convergence on S distinguish by relative compactness of the first marginal in $M_m(U)$.

Theorem 2.4 (Corollary 2.9 in [23])**.** *Suppose that U is Polish and that $\mathcal{U}$ is its Borel σ-field. Then, $P^n \to_{ws} P$ if and only if $\{P^n_U \colon n \in \mathbb{N}\}$ is relatively compact in $M_m(U)$ and $P^n \to P$ in $M_c(S)$.*

The following result relates relative (sequential) compactness in $M_{mc}(S)$ to relative (sequential) compactness of the marginals.

Theorem 2.5. *For a set $\Pi \subset M_{mc}(S)$ the following are equivalent:*

(i) *Π is relatively compact in $M_{mc}(S)$.*

(ii) *$\Pi_U \triangleq \{P_U \colon P \in \Pi\}$ is relatively compact in $M_m(U)$ and $\Pi_F \triangleq \{P_F \colon P \in \Pi\}$ is relatively compact in $M_c(F)$.*

(iii) *Π is relatively sequentially compact in $M_{mc}(S)$.*

(iv) *Π_U is relatively sequantially compact in $M_m(U)$ and Π_F is relatively sequentially compact in $M_c(F)$.*

Proof. Recall that $M_c(F)$ is metrizable and, in particular, sequential. Thus, the equivalence of (ii) and (iv) follows from [2, Proposition 2.2]. Moreover, the equivalence of (i) and (ii) follows from [2, Theorem 5.2] and the equivalence of (ii) and (iii) follows from [2, Theorem 2.5]. $\square$

Remark 2.6. In [23, Proposition 2.10] it is claimed that $M_{mc}(S)$ and $M_m(U)$ are metrizable in case $\mathcal{U}$ is separable. It was pointed out in [2, Remark 2.1] that the proof in [23] is not convincing and a short counterexample for the validity of the argument is given. Indeed, [23, Proposition 2.10] is incorrect, since, in case $(U,\mathcal{U})$ is a Polish space with its Borel σ-field, the space $M_m(U)$ is *not* metrizable by [18, Proposition 2.2.1]. If the σ-field $\mathcal{U}$ is separable and the U-marginal of some set $\Pi \subset M_{mc}(S)$ is relatively compact in $M_m(U)$, [2, Proposition 2.3] provides the positive result that Π is metrizable for the weak-strong topology.

2.2 The continuous mapping theorem of Jacod and Mémin

Next, we recall the continuous mapping theorem of Jacod and Mémin [23] for weak-strong convergence. For $A \in \mathcal{S}$ and $\alpha \in U$, we write

$$A_\alpha \triangleq \big\{\omega \in F \colon (\alpha, \omega) \in A\big\} \in \mathcal{B}(F).$$

Definition 2.7. *An $\mathcal{S}/\mathcal{B}(\mathbb{R})$-measurable function $g \colon S \to \mathbb{R}$ is called (P^n, P)-*continuous *if there exists a set $A \in \mathcal{S}$ such that*

(i) *$P^n(A) \to 1$ as $n \to \infty$, and $P(A) = 1$;*

(ii) *the set*
$$\big\{(\alpha, \omega) \in A \colon A_\alpha \ni \zeta \mapsto g(\alpha, \zeta) \text{ is discontinuous at } \omega\big\}$$
is P-null.

The following partial version of the Portmanteau theorem can be used to check part (i) in Definition 2.7.

Proposition 2.8 (Proposition 2.11 in [23])**.** *If $P^n \to_{ws} P$, then $\limsup_{n\to\infty} P^n(G) \le P(G)$ for all $G \in \mathcal{S}$ such that G_α is closed in F for every $\alpha \in U$.*

Theorem 2.9 (Theorem 2.16 in [23])**.** *Suppose that $P^n \to_{ws} P$ and let $g \colon S \to \mathbb{R}$ be (P^n, P)-continuous such that*

$$\sup_{n \in \mathbb{N}} E^{P^n}\big[|g| \mathbb{1}_{\{|g| > a\}}\big] \to 0$$

as $a \to \infty$. Then, $E^{P^n}[g] \to E^P[g]$ as $n \to \infty$.

At the end of this section we introduce a useful component to build a set A as in the definition of (P^n, P)-continuity. In the following, let (E, r) be a Polish space[1] and let $k \colon \mathbb{R}_+ \to \mathbb{R}_+$ be a Borel function such that, for every $t > 0$,

$$\lim_{\varepsilon \searrow 0} \sup \big\{k(s) \colon s \ne t, t - \varepsilon \le s \le t + \varepsilon\big\} = 0. \tag{2.2}$$

Lemma 2.10. *The property* (2.2) *holds if and only if for every $T, a > 0$ there exists no $t > 0$ such that the set $\{s \in [0, T] \colon k(s) \ge a\}$ contains a sequence $(t_n)_{n \in \mathbb{N}}$ with $t_n \ne t$ and $t_n \to t$ as $n \to \infty$. In particular,* (2.2) *holds if $\{s \in [0, T] \colon k(s) \ge a\}$ is finite for all $T, a > 0$.*

Proof. Let us start with the *if* implication. Take $t > 0$ and assume for contradiction that there exists a sequence $\varepsilon_n \searrow 0$ and a constant $a > 0$ such that $\sup\{k(s) \colon s \ne t, t - \varepsilon_n \le s \le t + \varepsilon_n\} > a$ for all $n \in \mathbb{N}$. There exists an $s_1 \ne t, t - \varepsilon_1 \le s_1 \le t + \varepsilon_1$ such that $k(s_1) \ge a$. Next, take $s_2 \ne t, t - \varepsilon_2 \le s_2 \le t + \varepsilon_2$ such that $k(s_2) \ge a$. Proceeding in this manner, we get a sequence $s_1, s_2, \dots$ with $s_n \ne t, s_n \to t$ and $k(s_n) \ge a$. This is a contradiction and the if implication follows.

We now prove the *only if* implication. For contradiction, assume that $T, a, t > 0$ are such that there exists a sequence $t_1, t_2, \dots \in [0, T]$ such that $t_n \ne t, t_n \to t$ and $k(t_n) \ge a$. Note that $|t - t_n| \ne 0$ and that $t - |t - t_n| \le t_n \le t + |t - t_n|$. Hence,

$$\liminf_{n \to \infty} \sup \big\{k(s) \colon s \ne t, t - |t - t_n| \le s \le t + |t - t_n|\big\} \ge \liminf_{n \to \infty} k(t_n) \ge a.$$

As this is a contradiction, the only if implication is also proved. □

Remark 2.11. It is possible that (2.2) holds while $\{t \in [0, T] \colon k(t) \ge a\}$ is infinite for some $T, a > 0$. Indeed, take for instance $k(t) = \sum_{k=1}^\infty \mathbb{1}_{\{t = 1/k\}}$.

[1] r is the corresponding metric.

To motivate what comes next, suppose that $\omega_1, \omega_2, \dots \in \mathbb{D}(E)$ is a sequence whose jumps are controlled by k, i.e. $r(\omega_n(t), \omega_n(t-)) \leq k(t)$ for all $t > 0$ and $n \in \mathbb{N}$. Furthermore, suppose that $\omega \in \mathbb{D}(E)$ is such that $\omega_n \to \omega$ in the Skorokhod J_1 topology. By standard properties of this topology, for every $t > 0$ with $r(\omega(t), \omega(t-)) > 0$ there exists a sequence $t_n \to t$ such that $r(\omega_n(t_n), \omega_n(t_n-)) \to r(\omega(t), \omega(t-))$. W.l.o.g. we can assume that there exists an $a > 0$ such that $r(\omega_n(t_n), \omega_n(t_n-)) \geq a$ for all $n \in \mathbb{N}$. By hypothesis, $k(t_n) \geq r(\omega_n(t_n), \omega_n(t_n-)) \geq a$. Recalling Lemma 2.10, we must have $t_n = t$ for all large enough n. In summary, we obtain convergence of the jumps, i.e. $r(\omega_n(t), \omega_n(t-)) \to r(\omega(t), \omega(t-))$. This property is certainly necessary for local uniform convergence and, as we will see below, it is even sufficient. Summarizing, for sequences whose jumps are controlled via k we obtain equivalence of the Skorokhod J_1 and the local uniform topology. As we are mainly interested in continuity properties, this is quite useful. In the following we fill in the remaining details.

Let $\kappa\colon \mathbb{R}_+ \to \mathbb{R}_+$ be increasing and continuous. Notice that we do *not* assume that $\kappa(0) = 0$. In fact, typical examples for κ could be $\kappa \equiv 1$ or $\kappa(x) = 1 + x$. Finally, fix a reference point $x_0 \in E$. Define

$$G \triangleq \Big\{\omega \in \mathbb{D}(E)\colon r(\omega(t), \omega(t-)) \leq k(t)\kappa\Big(\sup_{s\leq t} r(\omega(s), x_0)\Big) \text{ for all } t > 0\Big\}.$$

Using an exhausting sequence for the jumps of the coordinate process on $\mathbb{D}(E)$, we immediately see that $G \in \mathcal{B}(\mathbb{D}(E))$, where $\mathbb{D}(E)$ is endowed with the Skorokhod J_1 topology. In fact, we can say more, as the following proposition shows. For $E = \mathbb{R}^d$ and $\kappa \equiv 1$ or $\kappa(x) = 1 + x$, Proposition 2.12 is given by [22, Lemma 4.2] and [24, Lemma 3.6]. We adapt the proof to our abstract framework.

Proposition 2.12. *The set G is closed in $\mathbb{D}(E)$ for the local uniform and the Skorokhod J_1 topology. Moreover, on G the Skorokhod J_1 topology coincides with the local uniform topology.*

Proof. First of all, as κ is continuous, it is easy to see that G is closed in the local uniform topology. Hence, it suffices to prove the second claim, i.e. that the Skorokhod J_1 and the local uniform topology coincide on G. Of course, we only need to show that Skorokhod J_1 convergence implies local uniform convergence. Take $\omega_1, \omega_2, \dots \in G$ and $\omega \in \mathbb{D}(E)$ such that $\omega_n \to \omega$ in the Skorokhod J_1 topology. By virtue of [36, Theorem 2.6.2] and [25, Propositions VI.2.1, VI.2.7], it suffices to prove that $r(\omega_n(t), \omega_n(t-)) \to r(\omega(t), \omega(t-))$ for all $t > 0$ such that $r(\omega(t), \omega(t-)) > 0$. We fix such a $t > 0$. Thanks to [12, Problem 16, p. 152] (or [36, Theorem 2.7.1]), there exists a compact set $K = K_t \subset E$ such that $\omega_n(s) \in K$ for all $s \leq t + 1$ and $n \in \mathbb{N}$. Hence, taking into account that κ is increasing, there exists a constant $C > 0$ such that

$$\sup_{n\in\mathbb{N}} \kappa\Big(\sup_{s\leq t+1} r(\omega_n(s), x_0)\Big) \leq \kappa\Big(\sup_{x\in K} r(x, x_0)\Big) \leq C.$$

It is well-known ([25, Proposition VI.2.1]) that there exists a sequence $t_n \to t$ with $r(\omega_n(t_n), \omega_n(t_n-)) \to r(\omega(t), \omega(t-))$. Now, for large enough n we get $r(\omega_n(t_n), \omega_n(t_n-)) \leq Ck(t_n)$ and Lemma 2.10 yields that $t_n = t$ when n is large enough. This implies $r(\omega_n(t), \omega_n(t-)) \to r(\omega(t), \omega(t-))$ and the proof is complete. ☐

In Section 5 below we use a randomized version of the set G and the continuous mapping theorem for weak-strong convergence to relax the continuity assumptions in certain stability results for semimartingales, and to derive a version of the Ethier–Kurtz stability theorem associated to test processes of the type (1.3). The randomization is important as it allows for a much more flexible jump structure than the set G might suggest. We learned about the idea to use a set similar to G to relax continuity assumptions from

Jacod and Mémin [22, 24] and their work on stability and existence results for SDEs with semimartingale drivers.

3 The martingale problem method revisited

The martingale problem method is a powerful tool to identify the weak limit X of a weakly convergent sequence of stochastic processes $X^1, X^2, \dots$ via martingale properties of a family $\mathfrak{X}$ of test processes. In the following we present a general version of this method. More precisely, in Section 3.1 we start our program with a formal introduction of a general martingale problem and in Section 3.2 we discuss various examples appearing in the literature. The main results are given in Section 3.3.

3.1 Abstract martingale problems: the setting

Let us start with an introduction of an *abstract martingale problem*.

Setting. Let $(\Omega, \mathcal{F}, \mathbf{F} = (\mathcal{F}_t)_{t\geq 0})$ be a filtered space which supports a family $\mathfrak{X}$ of $\overline{\mathbb{R}}$-valued right-continuous adapted processes. Furthermore, let $(U, \mathcal{U})$ be a measurable space and let $(E, \mathcal{B}(E))$ be a Polish space with its Borel σ-field. Take also a subset $F \subset E^{\mathbb{R}_+}$ and endow it with a topology which turns it into a Polish space. We emphasize that the choice of the topology is rather flexible. Let $X = (X_t)_{t\geq 0}$ be an E-valued process such that for each $\omega \in \Omega$ the path $t \mapsto X_t(\omega)$ is an element of the path space F and the map $\Omega \times \mathbb{R}_+ \ni (\omega, t) \mapsto X_t(\omega) \in F$ is $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+)/\mathcal{B}(F)$-measurable. Finally, let L be a U-valued random variable on $(\Omega, \mathcal{F})$ and, as in Section 2, we define the product space $S \triangleq U \times F$ and $\mathcal{S} \triangleq \mathcal{U} \otimes \mathcal{B}(F)$.

Example 3.1. (i) In many classical cases X has càdlàg or even continuous paths and it is natural to take $F = \mathbb{D}(E)$ or $\mathbb{C}(E)$ (endowed with the Skorokhod J_1 topology[2] which renders $\mathbb{D}(E)$ and $\mathbb{C}(E)$ into Polish spaces).

(ii) In some frameworks of recent interest X has less path regularity. For example, this is the case for solutions to VSDEs as introduced in Example 3.13 below, where $F = L^p_{\text{loc}}(\mathbb{R}_+, \mathbb{R}^d) \times \mathbb{D}(\mathbb{R}^k)$ is a natural state space for solutions.

Definition 3.2. *We call a probability measure P on $(\Omega, \mathcal{F})$ a* solution to the (local) martingale problem (MP) $\mathfrak{X}$*, if all processes in $\mathfrak{X}$ are (local) $(\mathbf{F}, P)$-martingales. The sets of solutions to the martingale problem and the local martingale problem are denoted by $\mathcal{M}(\mathfrak{X})$ and $\mathcal{M}_{\text{loc}}(\mathfrak{X})$, respectively.*

In many cases of interest, the set $\mathfrak{X}$ has a canonical structure as described in the following definition.

Definition 3.3. *We call $\mathfrak{X}$* canonical for (L, X)*, or simply* canonical*, if the pair (L, X) is clear from the context, if for every $t \in \mathbb{R}_+$ there exists a set $\mathfrak{X}^\circ_t$ of $\mathcal{S}/\mathcal{B}(\overline{\mathbb{R}})$-measurable functions $Y^\circ_t \colon S \to \overline{\mathbb{R}}$ such that for every $Y \in \mathfrak{X}$ there exists a $Y^\circ_t \in \mathfrak{X}^\circ_t$ such that $Y_t = Y^\circ_t(L, X)$. We call $(Y^\circ_t)_{t\geq 0}$* a canonical version *of Y.*

In this paper we restrict our attention to canonical test processes.

Standing Assumption 3.4. The set $\mathfrak{X}$ is canonical for (L, X).

A prototype example for a test process from $\mathfrak{X}$ is given by

$$f(X_t) - \int_0^t g(s, L, X) q(ds), \quad t \in \mathbb{R}_+,$$

[2] On $\mathbb{C}(E)$ the Skorokhod J_1 coincides with the local uniform topology.

where f and g are sufficiently measurable and q is a deterministic locally finite measure. In Section 5.3 below, we investigate the martingale problem associated to such test processes and specify our theorems to this setting.

Besides the set of test processes $\mathfrak{X}$, which are candidates for (local) martingales, we further introduce a set of test functions for the *martingale property*.

Definition 3.5. *We call a family $\mathcal{Z}^\circ = \{\mathcal{Z}^\circ_t, t \in \mathbb{R}_+\}$ a* determining set *(for $\mathfrak{X}$), if it has the following properties:*

(i) *for every $t \in \mathbb{R}_+$, $\mathcal{Z}^\circ_t$ consists of bounded $\mathcal{S}/\mathcal{B}(\mathbb{R})$-measurable functions $S \to \mathbb{R}$;*
(ii) *for every probability measure P on $(\Omega, \mathcal{F})$, $Y \in \mathfrak{X}$ and $s < t$ with $Y_t, Y_s \in L^1(P)$ the following implication holds:*

$$E^P\big[Y_t Z^\circ_s(L, X)\big] = E^P\big[Y_s Z^\circ_s(L, X)\big] \text{ for all } Z^\circ_s \in \mathcal{Z}^\circ_s \implies P\text{-a.s. } E^P\big[Y_t|\mathcal{F}_s\big] = Y_s.$$

Let us present three typical determining sets for important settings. Notice that for product settings it is possible to combine the determining sets from these examples.

Example 3.6. (i) Suppose that $\mathcal{F}_t = \sigma(X_s, s \le t)$ for $t \in \mathbb{R}_+$. Take $F = \mathbb{D}(E)$ or $\mathbb{C}(E)$ and denote the corresponding coordinate process by X. Then, for any dense set $D \subset \mathbb{R}_+$, the family $\mathcal{Z}^\circ = \{\mathcal{Z}^\circ_t, t \in \mathbb{R}_+\}$ defined by

$$\mathcal{Z}^\circ_t \triangleq \Big\{ \prod_{i=1}^n h_i(\mathsf{X}_{t_i})\colon n \in \mathbb{N}, t_1, \dots, t_n \in D \cap [0, t], h_1, \dots, h_n \in C_b(E) \Big\}$$

is determining. As $\mathcal{B}(F) = \sigma(\mathsf{X}_t, t \in \mathbb{R}_+)$ (see [12, Proposition 3.7.1]), part (i) of Definition 3.5 is obvious and part (ii) follows from the monotone class theorem.

(ii) Take $E = \mathbb{R}, F = L^p_{\text{loc}}(\mathbb{R}_+, \mathbb{R})$, for $p \ge 1$, and let $\mathsf{X}\colon F \to F$ be the identity. Furthermore, assume that X is progressively measurable w.r.t. $\mathcal{F}_t \triangleq \sigma(X_s, s \le t)$ and

$$X_t = \hat{X}_t \triangleq \liminf_{n\to\infty} \Big(n \int_t^{t+\frac{1}{n}} X_s ds \Big), \quad t \in \mathbb{R}_+. \tag{3.1}$$

By Lebesgue's differentiation theorem we always have $X_t = \hat{X}_t$ for a.a. $t \in \mathbb{R}_+$ and hence, as X should be considered as an $L^p_{\text{loc}}(\mathbb{R}_+, \mathbb{R})$-valued random variable, the last assumption is essentially without loss of generality. Then, for any dense set $D \subset \mathbb{R}_+$, the family $\mathcal{Z}^\circ = \{\mathcal{Z}^\circ_t, t \in \mathbb{R}_+\}$ defined by

$$\mathcal{Z}^\circ_t \triangleq \Big\{ \prod_{i=1}^n h_i\Big(\int_0^{t_i} \mathsf{X}_s ds \Big)\colon n \in \mathbb{N}, t_1, \dots, t_n \in D \cap [0, t], h_1, \dots, h_n \in C_b(\mathbb{R}) \Big\}$$

is determining. Part (i) in Definition 3.5 follows from the fact that the maps

$$L^p_{\text{loc}}(\mathbb{R}_+, \mathbb{R}) \ni f = (f(s))_{s \ge 0} \mapsto \int_0^t f(s) ds, \quad t \in \mathbb{R}_+,$$

are continuous. For (ii) it suffices to use the monotone class theorem together with the observation that

$$\mathcal{F}_t = \mathcal{G}_t \triangleq \sigma\Big(\int_0^s X_u du, s \le t \Big).$$

Here, the inclusion $\mathcal{G}_t \subset \mathcal{F}_t$ is clear and the converse inclusion follows from $X = \hat{X}$. To see this, note that $\hat{X}$ is $(\mathcal{G}_{t+})_{t \ge 0} \triangleq (\mathcal{H}_t)_{t \ge 0}$-predictable (as pointwise limit of continuous processes) and thus $(\mathcal{H}_{t-})_{t \ge 0}$ adapted. As $\mathcal{H}_{t-} \subset \mathcal{G}_t$, $\hat{X}$ is $(\mathcal{G}_t)_{t \ge 0}$-adapted and $X = \hat{X}$ implies $\mathcal{F}_t \subset \mathcal{G}_t$.

(iii) The determining sets from (i) and (ii) are rather classical in the sense that they only depend on X and not on the control variable L. We now present an example which only depends on L and which turns out to be useful for annealed settings, see Section 5.5.3 below. Suppose that $U = \Omega$ and that $L\colon \Omega \to U = \Omega$ is the identity on Ω. In this case, if Π_t is an intersection stable collection of subsets of Ω such that $\Omega \in \Pi_t$ and $\mathcal{F}_t = \sigma(\Pi_t)$, then the family $\mathscr{Z}^\circ_t = \{\mathbb{1}_G \colon G \in \Pi_t\}$ is determining. Of course, this fact is an immediate consequence of the monotone class theorem. We emphasis that, quite contrary to (i) and (ii), for this example solely L is needed.

Before we turn to our abstract main results, we collect a variety of important examples for martingale problems.

3.2 Examples for martingale problems

The first example is the most classical martingale problem as introduced by Stroock and Varadhan.

Example 3.7 (Martingale Problem of Stroock and Varadhan)**.** Let $\Omega = \mathbb{C}(\mathbb{R}^d)$ and let X be the coordinate process on Ω. Furthermore, let $b\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times r}$ be locally bounded Borel functions. To obtain the martingale problem of Stroock and Varadhan [37], define $\mathfrak{X}$ to be the set of all test processes of the form

$$f(\mathsf{X}) - \int_0^\cdot \big(\langle b(s,\mathsf{X}_s), \nabla f(\mathsf{X}_s)\rangle + \tfrac{1}{2}\operatorname{tr}(\sigma\sigma^*(s,\mathsf{X}_s)\nabla^2 f(\mathsf{X}_s))\big)ds$$

where $f \in C^2_c(\mathbb{R}^d)$. It is classical (see, e.g. [28, Section 5.4]) that the set $\mathcal{M}(\mathfrak{X})$ coincides with the set of solution measures (i.e. laws of solution processes) for the SDE

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t,$$

where W is an r-dimensional standard Brownian motion.

Example 3.8 (Martingale Problem of Ethier and Kurtz)**.** Let E be a Polish space and take $\Omega = \mathbb{D}(E)$ or $\mathbb{C}(E)$. Again, let X be the coordinate process. Fix a set $A \subset C_b(E) \times B(E)$. To obtain the martingale problem of Ethier and Kurtz [12], define $\mathfrak{X}$ to be the set of all test processes of the form

$$f(\mathsf{X}) - \int_0^\cdot g(\mathsf{X}_s)ds, \quad (f,g) \in A.$$

The setting from Example 3.7 is a special case of this framework.

Example 3.9 (Semimartingale Problems)**.** In the following we discuss two ways to characterize the laws of semimartingales via martingale problems. The first one is given by [25, Theorem III.2.7]. Set $\Omega = \mathbb{D}(\mathbb{R}^d)$ and let (B, C, ν) be a candidate triplet for semimartingale characteristics corresponding to a fixed truncation function $h\colon \mathbb{R}^d \to \mathbb{R}^d$, see [25, Definition II.2.6] for a precise definition including the technical requirements. Let X be the coordinate process and define

$$\begin{aligned}
\mathsf{X}(h) &\triangleq \mathsf{X} - \sum_{s\le\cdot} \big(\Delta\mathsf{X}_s - h(\Delta\mathsf{X}_s)\big),\\
M(h) &\triangleq \mathsf{X}(h) - \mathsf{X}_0 - B,\\
\widetilde{C}^{ij} &\triangleq C^{ij} + \int h^i(x)h^j(x)\nu([0,\cdot]\times dx) - \sum_{s\le\cdot}\Delta B^i_s \Delta B^j_s.
\end{aligned}$$

Further, let $\mathscr{C}^+(\mathbb{R}^d)$ be a family of bounded real-valued Borel functions on $\mathbb{R}^d$ vanishing around the origin, which is measure determining for the class of Borel measures η on $\mathbb{R}^d$ with the properties $\eta(\{0\}) = 0$ and $\eta(\{x \in \mathbb{R}^d \colon \|x\| > \varepsilon\}) < \infty$ for all $\varepsilon > 0$, cf. [25, II.2.20] for more details. Let $\mathfrak{X}$ consist of the following processes:

(i) $M(h)^{(i)}, i = 1, \dots, d$;
(ii) $M(h)^{(i)} M(h)^{(j)} - \widetilde{C}^{(ij)}, i, j = 1, \dots, d$;
(iii) $\sum_{s \leq \cdot} g(\Delta \mathsf{X}_s) - \int g(x) \nu([0, \cdot] \times dx), g \in \mathscr{C}^+(\mathbb{R}^d)$.

Then, $\mathcal{M}_{\text{loc}}(\mathfrak{X})$ is the set of laws of semimartingales with characteristics (B, C, ν).

Next, we discuss an alternative characterization which can be seen as a reformulation of [25, Theorem II.2.42]. It is well-known ([25, Proposition II.2.9]) that (B, C, ν) admits a decomposition of the form

$$dB_t = b_t dA_t, \quad dC_t = c_t dA_t, \quad \nu(dt, dx) = F_t(dx) dA_t,$$

where A is an increasing right-continuous predictable process, and b, c and F are the predictable densities of (B, C, ν) w.r.t. the induced measure dA_t. For $f \in C^2_b(\mathbb{R}^d)$, set

$$\begin{aligned} \mathcal{L} f(s) \triangleq \langle b_s, \nabla f(\mathsf{X}_{s-}) \rangle + \tfrac{1}{2} \operatorname{tr}(c_s \nabla^2 f(\mathsf{X}_{s-})) \\ + \int \big(f(\mathsf{X}_{s-} + x) - f(\mathsf{X}_{s-}) - \langle \nabla f(\mathsf{X}_{s-}), h(x) \rangle \big) F_s(dx). \end{aligned}$$

Let $\mathfrak{X}^*$ be the set of all test processes of the form

$$f(\mathsf{X}) - \int_0^\cdot \mathcal{L} f(s) dA_s, \quad f \in C^2_b(\mathbb{R}^d).$$

Then, $\mathcal{M}_{\text{loc}}(\mathfrak{X}^*) = \mathcal{M}_{\text{loc}}(\mathfrak{X})$.

Finally, let us relate the local MP $(\mathfrak{X})$ to the class of Itô diffusions and the martingale problem of Stroock and Varadhan as explained in Example 3.7. For Brownian motion, more generally for diffusions, it is well-known that it suffices to consider linear and quadratic test functions, i.e. $f(x) = x^{(i)}$ and $f(x) = x^{(i)} x^{(j)}$ for $i, j = 1, \dots, d$, provided one asks in addition for continuous paths, cf. [28, Proposition 5.4.6]. For Brownian motion this observation is precisely Lévy's characterization. Namely, using $f(x) = x^{(i)}$ yields that X is a continuous[3] local martingale, and using in addition $f(x) = x^{(i)} x^{(j)}$ implies that $[\mathsf{X}, \mathsf{X}] = \mathrm{Id}$. The set $\mathfrak{X}$ generalizes this idea to general semimartingales. Thereby, the processes in (iii) take care of the jump structure. To see this, assume that $\nu = 0$, which means that (iii) consists of the processes $\sum_{s \leq \cdot} g(\Delta \mathsf{X}_s)$ with $g \in \mathscr{C}^+(\mathbb{R}^d)$. It is clear that this class consists of local martingales if and only if X is a.s. continuous. This observation relates the processes in (iii) above to the requirement of continuous paths in Lévy's characterization.

Remark 3.10. It may happen that a probability measure solves the martingale problems from Examples 3.8 and 3.9 but not both in a unique manner. For instance, suppose that X is a Brownian motion sticky at the origin, i.e. X solves the system

$$dX_t = \mathbb{1}_{\{X_t \neq 0\}} dW_t, \qquad \mathbb{1}_{\{X_t = 0\}} dt = \tfrac{1}{\mu} dL^0_t(X), \qquad \mu > 0,$$

where W is a standard Brownian motion and $L^0(X)$ denotes the semimartingale (right) local time of X in the origin. This characterization of a sticky Brownian motion is taken from [11]. It is obvious that X is a continuous local martingale (and hence a semimartingale) with quadratic variation

$$[X, X] = \int_0^\cdot \mathbb{1}_{\{X_s \neq 0\}} ds.$$

Thus, independent of the parameter μ, the law of X solves the (semi)martingale problem from Example 3.9 with $(0, C, 0)$ where

$$C(\omega) = \int_0^\cdot \mathbb{1}_{\{\omega(s) \neq 0\}} ds, \quad \omega \in \mathbb{D}(\mathbb{R}).$$

[3] Here, the additional requirement of continuous paths has to be taken into consideration.

In fact, even the Wiener measure solves this martingale problem. We conclude that the law of X cannot be captured in a unique manner by the semimartingale problem but, as X is a one-dimensional diffusion in the sense of Itô and McKean [20], its law is a unique solution to the martingale problem of Example 3.8 when $A \subset C_b(\mathbb{R}) \times C_b(\mathbb{R})$ is chosen appropriately,[4] see the discussion on p. 994 in [11] and [6, Remark 5.3] for more details.

Example 3.11 (Martingale Problem for Diffusions with a Change Point)**.** Let $(\Omega, \mathcal{F}, \mathbf{F}, P)$ be a filtered probability space which supports a one-dimensional standard Brownian motion W and a finite stopping time τ. Take two Borel functions $\sigma, \gamma \colon \mathbb{R} \to \mathbb{R}$ and let X be a continuous adapted process whose dynamics solve the equation

$$dX_t = (\sigma(X_t)\mathbb{1}_{\{t \leq \tau\}} + \gamma(X_t)\mathbb{1}_{\{t > \tau\}})dW_t.$$

The process X is a so-called *change point process* with change time τ. Processes of this type (possibly with drift, which we left out for simplicity) were, for instance, studied in [13] from a financial perspective. The dynamics of X can be captured by a canonical martingale problem. Let $\mathfrak{X}$ be the set of all processes of the form

$$f(X) - \int_0^\cdot \tfrac{1}{2} f''(X_s)(\sigma^2(X_s)\mathbb{1}_{\{s \leq \tau\}} + \gamma^2(X_s)\mathbb{1}_{\{s > \tau\}})ds, \quad f \in C^2(\mathbb{R}).$$

We take $U = \mathbb{R}_+, F = \mathbb{C}(\mathbb{R})$ and we set $L \triangleq \tau$. For these choices, it is easy to see that the set $\mathfrak{X}$ is canonical for (X, L). It follows from Itô's formula that $P \in \mathcal{M}(\mathfrak{X})$. Conversely, if $Q \in \mathcal{M}(\mathfrak{X})$, then, by the same arguments as used in the proof of [28, Proposition 5.4.6], on a standard extension of the underlying basis $(\Omega, \mathcal{F}, \mathbf{F}, Q)$, there exists a one-dimensional standard Brownian motion B such that

$$dX_t = (\sigma(X_t)\mathbb{1}_{\{t \leq \tau\}} + \gamma(X_t)\mathbb{1}_{\{t > \tau\}})dB_t.$$

We emphasis that in this example, the underlying space Ω does not necessarily coincide with $F = \mathbb{C}(\mathbb{R})$ and that the MP $(\mathfrak{X})$ is canonical for (X, L) but that the test processes do not solely depend on X.

Example 3.12 (Martingale Characterization for SPDEs)**.** We now describe a martingale problem for the semigroup approach to semilinear stochastic partial differential equations (SPDEs). The standard reference for this framework is the monograph of Da Prato and Zabczyk [10].

Let $E = (E, \langle \cdot, \cdot \rangle_E)$ be a separable real Hilbert space and set $\Omega = \mathbb{C}(E)$. Take another separable real Hilbert space $(H, \langle \cdot, \cdot \rangle_H)$ and denote by $L(H, E)$ the space of linear bounded operators $H \to E$. Moreover, let $\mu \colon \mathbb{R}_+ \times \Omega \to E$ and $\sigma \colon \mathbb{R}_+ \times \Omega \to L(H, E)$ be progressively measurable processes. To be precise, we mean that $\sigma h \colon \mathbb{R}_+ \times \Omega \to E$ is progressively measurable for every $h \in H$. Finally, let $A \colon D(A) \to E$ be the generator of a C_0-semigroup on E and let $A^* \colon D(A^*) \to E$ be its adjoint. Define Σ to be the set of all functions $g(\langle \cdot, y^* \rangle_E)$ where $y^* \in D(A^*)$ and $g \in C^2(\mathbb{R})$. For $f = g(\langle \cdot, y^* \rangle_E) \in \Sigma$ we set

$$\begin{aligned}(\mathcal{L}f)_s \triangleq g'(\langle \mathsf{X}_s, y^* \rangle_E)(\langle \mathsf{X}_s, A^* y^* \rangle_E + \langle \mu_s(\mathsf{X}), y^* \rangle_E) \\ + \tfrac{1}{2} g''(\langle \mathsf{X}_s, y^* \rangle_E) \langle \sigma_s^*(\mathsf{X}) y^*, \sigma_s^*(\mathsf{X}) y^* \rangle_H.\end{aligned}$$

Let $\mathfrak{X}$ be the set of all test processes of the form

$$f(\mathsf{X}) - \int_0^\cdot (\mathcal{L}f)_s ds, \quad f \in \Sigma.$$

[4] see Section 2.7 in [15] for details on how $A \subset C_b(\mathbb{R}) \times C_b(\mathbb{R})$ can be taken to capture diffusions in the sense of Itô and McKean.

Then, under suitable assumptions on the coefficients A, b and σ, the set $\mathcal{M}_{\mathrm{loc}}(\mathfrak{X})$ coincides with the set of laws of mild solutions to the SPDE

$$dX_t = (AX_t + \mu_t(X))dt + \sigma_t(X)dW_t,$$

where W is a standard cylindrical Brownian motion. We refer to [9, Proposition 2.6, Lemma 3.6] for more details.

Example 3.13 (Local Martingale Problem for SDEs of Volterra type)**.** Let (X, Z) be a measurable process with paths in $L^p_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^d) \times \mathbb{D}(\mathbb{R}^k)$ such that, on its underlying filtered probability space, X is predictable, Z is a semimartingale with characteristics

$$B^Z = \int_0^\cdot b(X_s)ds, \quad C^Z = \int_0^\cdot a(X_s)ds, \quad \nu^Z(dx, dt) = \nu(X_t, dx)dt,$$

corresponding to a fixed truncation function $h\colon \mathbb{R}^k \to \mathbb{R}^k$, and

$$X_t = g_0(t) + \int_{[0,t)} K_{t-s} dZ_s, \quad P \otimes \lambda\text{-a.e.},$$

where K is a convolution kernel $\mathbb{R}_+ \to \mathbb{R}^{d\times k}$. We call such a process (X, Z) a solution to a *Volterra SDE (VSDE)*. Recently, it was proven in [1] that solutions to VSDEs have a martingale characterization. For $f \in C^2_b(\mathbb{R}^k)$ and $(x, z) \in \mathbb{R}^d \times \mathbb{R}^k$ we set

$$\begin{aligned}\mathcal{L}f(x, z) \triangleq \langle b(x), \nabla f(z)\rangle &+ \tfrac{1}{2}\operatorname{tr}(a(x)\nabla^2 f(z)) \\ &+ \int \big(f(z + y) - f(z) - \langle h(y), \nabla f(z)\rangle\big)\nu(x, dy).\end{aligned}$$

Then, (X, Z) is a solution to the VSDE described above if and only if the processes

$$f(Z) - \int_0^\cdot \mathcal{L}f(X_s, Z_s)ds, \quad f \in C^2_b(\mathbb{R}^k),$$

are local martingales and

$$\int_0^t X_s ds = \int_0^t g_0(s)ds + \int_0^t K_{t-s}Z_s ds, \quad t \in \mathbb{R}_+.$$

In Section 4.2 below we take a closer look at VSDEs.

3.3 Identifying weak limits via abstract martingale problems

The classical martingale problems from Examples 3.7, 3.8 and 3.9 proved themselves as valuable tools to identify weak limits of stochastic processes. In the following we discuss such an application for the abstract martingale problem as introduced in Definition 3.2. We extend the setting from the beginning of Section 3.1 as follows:

Setting. For every $n \in \mathbb{N}$, let $\mathbb{B}^n \triangleq (\Omega^n, \mathcal{F}^n, \mathbf{F}^n = (\mathcal{F}^n_t)_{t\geq 0}, P^n)$ be a filtered probability space, which supports a U-valued random variable L^n and an E-valued measurable processes $X^n = (X^n_t)_{t\geq 0}$ such that, for every $\omega \in \Omega$, the process $X^n(\omega)$ is an element of F and the map $\Omega^n \ni \omega \mapsto X^n(\omega) \in F$ is $\mathcal{F}^n/\mathcal{B}(F)$-measurable. Moreover, we fix a probability measure P on $(\Omega, \mathcal{F})$ and denote $Q^n \triangleq P^n \circ (L^n, X^n)^{-1}$ and $Q \triangleq P \circ (L, X)^{-1}$.

Recall Definition 2.1 for the concept of weak-strong convergence, and Definition 2.7 for the concept of (Q^n, Q)-continuity. Furthermore, recall that $\mathfrak{X}$ is assumed to be canonical (see Definition 3.3).

Theorem 3.14. *Let $D \subset \mathbb{R}_+$ be dense, and assume the following:*

(A1) $Q^n \to_{ws} Q$;

(A2) *there exists a determining set $\mathcal{Z}^\circ = \{\mathcal{Z}^\circ_t, t \in \mathbb{R}_+\}$ for $\mathfrak{X}$;*

(A3) *for every $Y \in \mathfrak{X}$ there exists a canonical version $(Y^\circ_t)_{t\geq 0}$ such that for every $t \in D$, $s \in D \cap [0,t]$ and $Z^\circ_s \in \mathcal{Z}^\circ_s$ the following hold: Y°_t and $Y^\circ_t Z^\circ_s$ are (Q^n, Q)-continuous, the set $\{Y^\circ_r(L^n, X^n) \colon r \in D \cap [0,t], n \in \mathbb{N}\}$ is uniformly integrable, and*

$$\lim_{n\to\infty} E^{P^n}\big[(Y^\circ_t(L^n,X^n) - Y^\circ_s(L^n,X^n))Z^\circ_s(L^n,X^n)\big] = 0. \tag{3.2}$$

Then, P solves the MP $(\mathfrak{X})$, i.e. $P \in \mathcal{M}(\mathfrak{X})$.

Before we prove this theorem, we briefly recall [25, Lemma IX.1.11], which is very useful in the following.

Lemma 3.15. *A family $\{Z_i \colon i \in I\}$ is uniformly integrable if and only if*

$$\sup_{i\in I} E\big[|Z_i| - |Z_i| \wedge z\big] \to 0 \text{ as } z \to \infty.$$

Proof of Theorem 3.14. We have to prove that every $Y \in \mathfrak{X}$ is a martingale on $(\Omega, \mathcal{F}, \mathbf{F}, P)$. Fix $Y \in \mathfrak{X}$ with canonical version $Y^\circ = (Y^\circ_t)_{t\geq 0}$ and take $s, t \in D$ such that $s < t$ and $Z^\circ_s \in \mathcal{Z}^\circ_s$. Using the canonical property of Y in the first, the (Q^n, Q)-continuity of $Y^\circ Z^\circ$ and Theorem 2.9 in the second, and (3.2) in the third equality, we find

$$\begin{aligned} E^P\big[(Y_t - Y_s)Z^\circ_s(L,X)\big] &= E^P\big[(Y^\circ_t(L,X) - Y^\circ_s(L,X))Z^\circ_s(L,X)\big] \\ &= \lim_{n\to\infty} E^{P^n}\big[(Y^\circ_t(L^n,X^n) - Y^\circ_s(L^n,X^n))Z^\circ_s(L^n,X^n)\big] = 0. \end{aligned} \tag{3.3}$$

Thus, by (A2) and part (ii) of Definition 3.5, we conclude that P-a.s. $E^P\big[Y_t|\mathcal{F}_s\big] = Y_s$.

Next, we show this identity for general $s < t$. We start by showing that the set $\{Y^\circ_s(L,X) \colon\ s \in D \cap [0,t]\}$ is uniformly integrable for every $t \in \mathbb{R}_+$. Let $s \in D$. The (Q^n, Q)-continuity of Y°_s and Theorem 2.9 yield that

$$\begin{aligned} E^P\big[|Y^\circ_s(L,X)| - |Y^\circ_s(L,X)| \wedge N\big] &= \lim_{n\to\infty} E^{P^n}\big[|Y^\circ_s(L^n,X^n)| - |Y^\circ_s(L^n,X^n)| \wedge N\big] \\ &\leq \sup_{n\in\mathbb{N}} E^{P^n}\big[|Y^\circ_s(L^n,X^n)| - |Y^\circ_s(L^n,X^n)| \wedge N\big]. \end{aligned}$$

Thus,

$$\begin{aligned} &\sup_{s\in D\cap[0,t]} E^P\big[|Y^\circ_s(L,X)| - |Y^\circ_s(L,X)| \wedge N\big] \\ &\qquad \leq \sup_{s\in D\cap[0,t]} \sup_{n\in\mathbb{N}} E^{P_n}\big[|Y^\circ_s(L^n,X^n)| - |Y^\circ_s(L^n,X^n)| \wedge N\big] \to 0 \end{aligned}$$

as $N \to \infty$ by Lemma 3.15. Another application of Lemma 3.15 implies that the set $\{Y_s \colon s \in D \cap [0,t]\} = \{Y^\circ_s(L,X) \colon s \in D \cap [0,t]\}$ is uniformly integrable.

Now, let $s < t$ be arbitrary, i.e. not necessarily in the set D. As D is dense in $\mathbb{R}_+$, there are sequences $t_n \searrow t$ and $s_n \searrow s$ in D such that $s_n < t_n$ for all $n \in \mathbb{N}$. The right-continuity of Y and Vitali's theorem yield that for every $G \in \mathcal{F}_s$ we have

$$E^P\big[Y_t \mathbb{1}_G\big] = \lim_{n\to\infty} E^P\big[Y_{t_n}\mathbb{1}_G\big] = \lim_{n\to\infty} E^P\big[Y_{s_n}\mathbb{1}_G\big] = E^P\big[Y_s\mathbb{1}_G\big]. \tag{3.4}$$

We conclude the P-martingale property of Y. The proof is complete. □

Remark 3.16. In case (A1) holds and $P \in \mathcal{M}(\mathfrak{X})$, (3.2) has to hold under the remaining assumptions in (A3), cf. (3.3).

Let us also comment on the case without control variables, which can be captured with the assumption that U is a singleton. We will simplify our notation for this situation and remove L^n and L. To clarify our terminology, we write $X^n \to X$ weakly when the laws of X^n converge in $M_c(F)$ to the law of X. Moreover, we call a Borel function $f\colon F \to \mathbb{R}$ to be *P-continuous at X*, if there exists a set $C \in \mathcal{B}(F)$ such that $P(X \in C) = 1$ and $f(s_n) \to f(s)$ whenever $s_n \to s \in C$. The following is an immediate consequence of Theorem 3.14.

Corollary 3.17. *Suppose that U be a singleton, $D \subset \mathbb{R}_+$ be dense, and assume the following:*

(S1) *$X^n \to X$ weakly;*

(S2) *there exists a determining set $\mathcal{Z}^\circ = \{\mathcal{Z}^\circ_t, t \in \mathbb{R}_+\}$ for $\mathfrak{X}$;*

(S3) *for every $Y \in \mathfrak{X}$ there exists a canonical version $(Y^\circ_t)_{t\geq 0}$ such that for all $t \in D$, $s \in D \cap [0,t]$ and $Z^\circ_s \in \mathcal{Z}^\circ_s$ the following hold: Y°_t and $Y^\circ_t Z^\circ_s$ are P-continuous at X, the set $\{Y^\circ_r(X^n)\colon r \in D \cap [0,t], n \in \mathbb{N}\}$ is uniformly integrable, and*

$$\lim_{n\to\infty} E^{P^n}\big[(Y^\circ_t(X^n) - Y^\circ_s(X^n))Z^\circ_s(X^n)\big] = 0.$$

Then, P solves the MP $(\mathfrak{X})$, i.e. $P \in \mathcal{M}(\mathfrak{X})$.

As the following proposition shows, (3.2) in Theorem 3.14 holds in case $Y^\circ(L^n, X^n)$ can be approximated by a sequence of martingales on $\mathbb{B}^n$.

Proposition 3.18. *Let all assumptions from Theorem 3.14 hold, except* (3.2)*. Suppose that for every $s \in D, Z^\circ_s \in \mathcal{Z}^\circ_s$ the random variable $Z^\circ_s(L^n, X^n)$ is $\mathcal{F}^n_s$-measurable, and that there exists a sequence $(Y^n)_{n\in\mathbb{N}}$ such that Y^n is a martingale on $\mathbb{B}^n$. If*

$$\lim_{n\to\infty} E^{P^n}\big[(Y^n_t - Y^\circ_t(L^n, X^n))Z^\circ_s(L^n, X^n)\big] = 0, \quad s, t \in D, s \leq t, Z^\circ_s \in \mathcal{Z}^\circ_s, \tag{3.5}$$

then (3.2) *holds. In particular,* (3.2) *holds in case*

$$\lim_{n\to\infty} E^{P^n}\big[|Y^n_t - Y^\circ_t(L^n, X^n)|\big] = 0. \tag{3.6}$$

Proof. Using the martingale property of Y^n in the first, and (3.5) in the second equality, the hypothesis yields that

$$\begin{aligned}&\lim_{n\to\infty} E^{P^n}\big[(Y^\circ_t(L^n, X^n) - Y^\circ_s(L^n, X^n))Z^\circ_s(L^n, X^n)\big]\\ &\qquad = \lim_{n\to\infty} E^{P^n}\big[(Y^\circ_t(L^n, X^n) - Y^n_t + Y^n_s - Y^\circ_s(L^n, X^n))Z^\circ_s(L^n, X^n)\big] = 0.\end{aligned}$$

The second claim follows from the first. □

Remark 3.19. By Vitali's theorem, (3.6) can be replaced by uniform integrability and convergence in probability. More precisely, if the family $\{|Y^n_t - Y^\circ_t(L^n, X^n)|\colon n \in \mathbb{N}\}$ is uniformly integrable and, for all $\varepsilon > 0$,

$$P^n(|Y^n_t - Y^\circ_t(L^n, X^n)| \geq \varepsilon) \to 0 \text{ as } n \to \infty,$$

then (3.6) holds.

Discussion. By virtue of Proposition 3.18, it seems natural to assume that also the sequence $P^1, P^2, \ldots$ solves an (abstract) martingale problem. Although we are only interested in limiting martingale problems which are canonical for (L, X), it is *not* always the case that the martingale problems associated to $P^1, P^2, \ldots$ are *canonical for* (L^n, X^n). The reason for this is that we ask the sequence $(L^1, X^1), (L^2, X^2), \ldots$ to converge in the weak-strong sense, which in particular means that the sequence $L^1, L^2, \dots$ converges setwise. In some cases this requirement does not hold for arbitrary choices of $L^1, L^2, \dots$. To illustrate this issue, let us discuss an explicit example (inspired by [19]). For every $n = 1, 2, \dots$, take a sequence $Z^n_1, Z^n_2, \dots$ of real-valued i.i.d. random variables with mean μ^n and variances $v^n/2$. Furthermore, let $N = (N_t)_{t\geq 0}$ be a standard Poisson process which is independent of all $Z^n_k, n, k = 1, 2, \dots$. We are interested in the limit of the sequence

$$X^n \triangleq n \int_0^\cdot Z^n_{N_{n^2 s}} ds \equiv n \int_0^\cdot Y^n_s ds, \quad n = 1, 2, \dots.$$

In case $n\mu^n \to 0$ and $v^n \to 1$, it follows from the Lindeberg–Feller theorem that the finite dimensional distributions of $X^1, X^2, \dots$ converge to those of Brownian motion if and only if the Lindeberg–Feller condition holds, i.e., for every $\varepsilon > 0$, $E[(Z^n_0)^2 \mathbb{1}_{\{|Z^n_0| > \varepsilon n\}}] \to 0$ as $n \to \infty$. Clearly, this shows that the only candidate for a functional limit is Brownian motion. To use the martingale problem method to identify Brownian motion as limiting process, one can consider the approximating martingale problems given by

$$\mathfrak{X}^n \triangleq \Big\{ f(X^n) + \tfrac{1}{n} Y^n f'(X^n) - \int_0^\cdot \big(f''(X^n_s)(Y^n_s)^2 + f'(X^n_s) n\mu_n \big) ds \colon f \in C^2_c(\mathbb{R}) \Big\}.$$

As shown in [19], in case $\sup_{n\in\mathbb{N}} E[(Z^n_0)^{2+\varepsilon}] < \infty$ for some $\varepsilon > 0$, approximately

$$\begin{aligned} f(X^n) + \tfrac{1}{n} Y^n f'(X^n) - \int_0^\cdot \big(f''(X^n_s)(Y^n_s)^2 + f'(X^n_s) n\mu_n \big) ds & \\ \approx f(X^n) - \int_0^\cdot \tfrac{1}{2} f''(X^n_s) ds, & \end{aligned} \tag{3.7}$$

where the latter test process corresponds to Brownian motion. Evidently, the martingale problem $\mathfrak{X}^n$ is canonical for (Y^n, X^n) and therefore, it would be natural to take $L^n = Y^n$. However, the sequence $(Y^1, X^1), (Y^2, X^2), \dots$ seems not to converge in the weak-strong sense. To overcome this problem, we can take the sequence $L^1, L^2, \dots$ to be deterministic and constant (or U to be a singleton), for instance, and solely work with the sequence $X^1, X^2, \dots$. Finally, (3.7) shows that randomness gets lost in the limit in the sense that the approximating martingale problems $\mathfrak{X}^1, \mathfrak{X}^2, \dots$ are not only depending on $X^1, X^2, \dots$, while the limiting martingale problem does only depend on the limit X.

We now replace the uniform integrability assumption in (A3) by a uniform integrability assumption on the approximating martingales.

Theorem 3.20. *Let all assumptions from Theorem 3.14 hold, except* (A3). *Suppose that for every $s \in D, Z^\circ_s \in \mathcal{Z}^\circ_s$ the random variable $Z^\circ_s(L^n, X^n)$ is $\mathcal{F}^n_s$-measurable, and that the following holds:*

(A4) *for every $Y \in \mathfrak{X}$ there exists a canonical version $(Y^\circ_t)_{t\geq 0}$ such that for every $t \in D$, $s \in D \cap [0, t]$ and $Z^\circ_s \in \mathcal{Z}^\circ_s$ the following hold: Y°_t and $Y^\circ_t Z^\circ_s$ are (Q^n, Q)-continuous. Moreover, there exists a sequence $(Y^n)_{n\in\mathbb{N}}$ such that Y^n is a martingale on $\mathbb{B}^n$, the set $\{Y^n_s \colon s \in D \cap [0, t], n \in \mathbb{N}\}$ is uniformly integrable and*

$$\lim_{n\to\infty} P^n(|Y^n_t - Y^\circ_t(L^n, X^n)| \geq \varepsilon) = 0, \quad \varepsilon > 0. \tag{3.8}$$

Then, P solves the MP $(\mathfrak{X})$, i.e. $P \in \mathcal{M}(\mathfrak{X})$.

Proof. It is not hard to see that the proof of Theorem 3.14 remains valid in case the following two properties hold:

$$\{Y^\circ_s(L, X) \colon s \in D \cap [0, t]\} \text{ is uniformly integrable for all } t \in D, \tag{3.9}$$

$$E^{P^n}\big[Y^n_t Z^\circ_s(L^n,X^n)\big] \to E^P\big[Y^\circ_t(L,X)Z^\circ_s(L,X)\big] \text{ for all } s,t\in D, s\le t, Z^\circ_s\in\mathcal{Z}^\circ_s. \quad (3.10)$$

Indeed, (3.9) suffices for (3.4), and if (3.10) holds, we write for $s,t\in D, s<t, Z^\circ_s\in\mathcal{Z}^\circ_s$

$$\begin{aligned} E^P\big[(Y_t-Y_s)Z^\circ_s(L,X)\big] &= E^P\big[(Y^\circ_t(L,X)-Y^\circ_s(L,X))Z^\circ_s(L,X)\big] \\ &= \lim_{n\to\infty} E^{P_n}\big[(Y^n_t-Y^n_s)Z^\circ_s(L^n,X^n)\big]=0, \end{aligned}$$

i.e. the conclusion of (3.3) holds as well.

For (3.9), note that (3.8) and Theorem 2.9 yield that, for all $N>0$ and $t\in D$,

$$\begin{aligned} &\lim_{n\to\infty}\big|E^{P^n}\big[|Y^n_t|\wedge N\big]-E^P\big[|Y^\circ_t(L,X)|\wedge N\big]\big| \\ &\quad\le \lim_{n\to\infty} E^{P^n}\big[|Y^n_t-Y^\circ_t(L^n,X^n)|\wedge N\big] \\ &\qquad\quad + \lim_{n\to\infty}\big|E^{P_n}\big[|Y^\circ_t(L^n,X^n)|\wedge N\big]-E^P\big[|Y^\circ_t(L,X)|\wedge N\big]\big| = 0. \end{aligned}$$

Hence, for every $N>0$ and $t\in D$ we have

$$\begin{aligned} E^P\big[|Y^\circ_t(L,X)|-|Y^\circ_t(L,X)|\wedge N\big] &= \lim_{m\to\infty} E^P\big[|Y^\circ_t(L,X)|\wedge m-|Y^\circ_t(L,X)|\wedge N\big] \\ &= \lim_{m\to\infty}\lim_{n\to\infty} E^{P^n}\big[|Y^n_t|\wedge m-|Y^n_t|\wedge N\big] \\ &\le \sup_{n\in\mathbb{N}} E^{P^n}\big[|Y^n_t|-|Y^n_t|\wedge N\big]. \end{aligned}$$

Together with uniform integrability of $\{Y^n_s : s\in D\cap[0,t], n\in\mathbb{N}\}$ and Lemma 3.15, this inequality yields (3.9). Next, we verify (3.10). For $s,t\in D, s\le t, Z^\circ_s\in\mathcal{Z}^\circ_s$ and $N>0$, we obtain

$$\begin{aligned} &\big|E^P\big[Y^\circ_t(L,X)Z^\circ_s(L,X)\big]-E^{P^n}\big[Y^n_t Z^\circ_s(X^n)\big]\big| \\ &\qquad\lesssim E^P\big[|Y^\circ_t(L,X)|-|Y^\circ_t(L,X)|\wedge N\big] \\ &\qquad\quad + \big|E^P\big[(Y^\circ_t(L,X)\vee(-N)\wedge N)Z^\circ_s(L,X)\big] \\ &\qquad\qquad\quad - E^{P^n}\big[(Y^\circ_t(L^n,X^n)\vee(-N)\wedge N)Z^\circ_s(L^n,X^n)\big]\big| \\ &\qquad\quad + E^{P^n}\big[\big|Y^\circ_t(L^n,X^n)\vee(-N)\wedge N-Y^n_t\vee(-N)\wedge N\big|\big] \\ &\qquad\quad + E^{P^n}\big[|Y^n_t|-|Y^n_t|\wedge N|\big] \\ &\qquad\triangleq I_1+I_2+I_3+I_4. \end{aligned}$$

Theorem 2.9 yields that $I_2\to0$ as $n\to\infty$. Moreover, (3.8) implies that $I_3\to0$ as $n\to\infty$. Finally, uniform integrability and Lemma 3.15 yield that $I_1+I_4\to0$ as $N\to\infty$ uniformly in n. In summary, we conclude that (3.10) holds and hence the proof is complete. □

In Theorem 3.20 we do not impose integrability assumptions on the elements of $\mathfrak{X}$ but on its approximation sequences. Hence, Theorems 3.14 and 3.20 have different scopes and do not imply each other.

Finally, let us again comment on the case without control variables. The following is an immediate consequence of Theorem 3.20.

Corollary 3.21. *Let all assumptions from Corollary 3.17 hold, except* (S3). *Suppose that for every $s\in D$ and $Z^\circ_s\in\mathcal{Z}^\circ_s$ the random variable $Z^\circ_s(X^n)$ is $\mathcal{F}^n_s$-measurable, and that the following holds:*

(S4) *For every $Y\in\mathfrak{X}$ there exists a canonical version $(Y^\circ_t)_{t\ge0}$ such that for every $t\in D$, $s\in D\cap[0,t]$ and $Z^\circ_s\in\mathcal{Z}^\circ_s$ the following hold: Y°_t and $Y^\circ_tZ^\circ_s$ are P-continuous at X. Moreover, there exists a sequence $(Y^n)_{n\in\mathbb{N}}$ such that Y^n is a martingale on $\mathbb{B}^n$, the set $\{Y^n_s\colon s\in D\cap[0,t], n\in\mathbb{N}\}$ is uniformly integrable and*

$$\lim_{n\to\infty} P^n(|Y^n_t-Y^\circ_t(X^n)|\ge\varepsilon)=0,\quad \varepsilon>0.$$

Then, P solves the MP $(\mathfrak{X})$, i.e. $P \in \mathcal{M}(\mathfrak{X})$.

In the next section we relate the results above to known theorems from the literature. Thereafter, in Section 5 we present new results which are tailored to processes with fixed times of discontinuity. For these results it is crucial that we can work with the concept of weak-strong convergence.

4 Relation to existing results

The purpose of this section is to specialize the terminologies introduced in the previous section to three examples taken from the literature: in Section 4.1 we recover the classical convergence theorem for Markovian martingale problems as presented in the monograph [12] by Ethier and Kurtz, and in Section 4.2 we prove a mild generalization of a stability result for Volterra SDEs from [1]. Finally, in Section 4.3 we localize a theorem by Jacod and Shiryaev [25] for semimartingales by replacing a global with a local boundedness hypothesis on the semimartingale characteristics. Such a generalization has been announced in [25], but it was not stated in a precise manner. We believe it to be useful for future applications and therefore of independent interest.

4.1 Relation to a theorem by Ethier and Kurtz

Let E be a Polish space, for every $n \in \mathbb{N}$, let $\mathbb{B} = (\Omega, \mathcal{F}, \mathbf{F}, P)$ and $\mathbb{B}^n = (\Omega^n, \mathcal{F}^n, \mathbf{F}^n, P^n)$ be filtered probability spaces which support E-valued càdlàg adapted processes X and X^n, respectively. Moreover, suppose that the filtration $\mathbf{F}$ on $\mathbb{B}$ is generated by X. Let $A \subset C_b(E) \times C_b(E)$ and define $\mathfrak{X}$ to be the set of all test processes of the form

$$f(X) - \int_0^{\cdot} g(X_s)ds, \quad (f, g) \in A. \tag{4.1}$$

Moreover, for every $n \in \mathbb{N}$, let $\mathfrak{X}^n$ be a set of pairs (ξ, ϕ) consisting of real-valued progressively measurable processes on $\mathbb{B}^n$ such that

$$\sup_{s \le T} E^{P^n}\big[|\xi_s| + |\phi_s|\big] < \infty, \quad T > 0,$$

and such that

$$\xi - \int_0^{\cdot} \phi_s ds$$

is a martingale on $\mathbb{B}^n$. The following theorem is a version of the implication (c′) ⇒ (a′) from [12, Theorem 4.8.10].

Theorem 4.1. *Suppose that $X^n \to X$ weakly on $\mathbb{D}(E)$ endowed with the Skorokhod J_1 topology, and assume that there exists a set $\Gamma \subset \mathbb{R}_+$ with countable complement such that, for each $(f, g) \in A$ and $T > 0$, there exists a sequence $(\xi^n, \phi^n) \in \mathfrak{X}^n$ such that*

$$\sup_{n \in \mathbb{N}} \sup_{s \le T} E^{P^n}\big[|\xi^n_s| + |\phi^n_s|\big] < \infty, \tag{4.2}$$

$$\lim_{n \to \infty} E^{P^n}\Big[(\xi^n_t - f(X^n_t)) \prod_{i=1}^k h_i(X^n_{t_i})\Big] = 0, \tag{4.3}$$

$$\lim_{n \to \infty} E^{P^n}\Big[\int_s^t (\phi^n_u - g(X^n_u))du \prod_{i=1}^k h_i(X^n_{t_i})\Big] = 0, \tag{4.4}$$

for all $k \in \mathbb{N}, t_1, \dots, t_k \in \Gamma \cap [0, t], t \in \Gamma \cap [0, T]$, $h_1, \dots, h_k \in C_b(E)$. Then, $P \in \mathcal{M}(\mathfrak{X})$.

Proof. We check (S1) – (S3) in Corollary 3.17. Of course, (S1) holds by hypothesis. Let $\mathcal{Z}^\circ = \{\mathcal{Z}^\circ_t, t \in \mathbb{R}_+\}$ be as in part (i) of Example 3.6 with

$$D \triangleq \big\{t \in \Gamma \colon P(X_t \neq X_{t-}) = 0\big\}.$$

As Γ^c and $\{t \in \mathbb{R}_+ \colon P(X_t \neq X_{t-}) > 0\}$ are countable (see [12, Lemma 3.7.7] for the countability of the second set), the set D^c is also countable. Consequently, D is dense in $\mathbb{R}_+$. As explained in Example 3.6, $\mathcal{Z}^\circ$ is a determining set for $\mathfrak{X}$. Thus, (S2) holds. Finally, we check (S3). It is clear that any canonical version $(Y^\circ_t)_{t\geq 0}$ from $\mathfrak{X}$ is bounded on finite time intervals. This implies that $\{Y^\circ_r(X^n) \colon r \in D \cap [0,t], n \in \mathbb{N}\}$ is uniformly integrable. Moreover, as $\omega \mapsto \omega(t)$ is continuous at ω whenever $\omega(t) = \omega(t-)$, for every $t \in D$ any $Z^\circ_t \in \mathcal{Z}^\circ_t$ is P-a.s. continuous at X by definition of D. Similarly, again by definition of D, for every $t \in D$ the random variable Y°_t is P-a.s. continuous at X. It is left to verify the final part of (S3). Take $Y \in \mathfrak{X}$ such that

$$Y = f(X) - \int_0^\cdot g(X_s) ds,$$

and set

$$Y^n \triangleq \xi^n - \int_0^\cdot \phi^n_s ds,$$

where ξ^n and ϕ^n are as in (4.2), (4.3) and (4.4). Let $s, t \in D \subset \Gamma$ with $s < t$ and take $Z^\circ_s \in \mathcal{Z}^\circ_s$. Clearly, we have

$$\begin{aligned}
&(Y^\circ_t(X^n) - Y^n_t + Y^n_s - Y^\circ_s(X^n)) Z^\circ_s(X^n) \\
&\qquad = (f(X^n_t) - \xi^n_t + \xi^n_s - f(X^n_s)) \prod_{i=1}^k h_i(X^n_{t_i}) \\
&\qquad\qquad + \int_s^t (\phi^n_u - g(X^n_u)) du \prod_{i=1}^k h_i(X^n_{t_i})
\end{aligned}$$

for certain $k \in \mathbb{N}, t_1, \dots, t_k \in \Gamma \cap [0,s], h_1, \dots, h_k \in C_b(E)$ related to Z°_s, cf. Example 3.6 (i). The P^n-expectation of the first term converges to zero by (4.3), and the P^n-expectation of the second term converges to zero by (4.4). As Y^n is a martingale on $\mathbb{B}^n$ we have

$$E^{P^n}\big[(Y^n_s - Y^n_t) Z^\circ_s(X^n)\big] = 0,$$

and consequently,

$$\begin{aligned}
&\lim_{n\to\infty} E^{P^n}\big[(Y^\circ_t(X^n) - Y^\circ_s(X^n)) Z^\circ_s(X^n)\big] \\
&\qquad = \lim_{n\to\infty} E^{P^n}\big[(Y^\circ_t(X^n) - Y^n_t + Y^n_s - Y^\circ_s(X^n)) Z^\circ_s(X^n)\big] = 0.
\end{aligned}$$

We conclude that (S3) holds. Hence, the claim follows from Corollary 3.17. □

Remark 4.2. In [12, Theorem 4.8.10, (a′) ⇒ (c′)] it is shown that in case $X^n \to X$ weakly and $P \in \mathcal{M}(\mathfrak{X})$, there exist processes $(\xi^n, \phi^n) \in \mathfrak{X}^n$ with the properties (4.2), (4.3) and (4.4).

In Section 5.3 below, we derive a version of Theorem 4.1 where we add another integral term w.r.t. a general locally finite measure to the class of test processes from (4.1). At this point we stress that the proof (and the result itself) requires substantial adjustments, as in this case the test processes have no Skorokhod J_1 continuous canonical versions in general. More comments on this issue are given at the end of Section 4.3.

4.2 A stability result for Volterra equations

In this section we discuss a stability result for Volterra SDEs (VSDEs) of the type

$$X_t = g_0(t) + \int_{[0,t)} K_{t-s} dZ_s, \quad t \in \mathbb{R}_+, \tag{4.5}$$

where X is an $\mathbb{R}^d$-valued predictable process and Z is an $\mathbb{R}^k$-valued semimartingale with differential characteristics $(b(X), a(X), \nu(X))$, i.e. whose characteristics (B^Z, C^Z, ν^Z) are of the form

$$B^Z = \int_0^{\cdot} b(X_s) ds, \quad C^Z = \int_0^{\cdot} a(X_s) ds, \quad \nu^Z(dx, dt) = \nu(X_t, dx) dt.$$

Throughout this section we suppose them to correspond to a fixed continuous truncation function $h\colon \mathbb{R}^k \to \mathbb{R}^k$. A version of Theorem 4.4 below has recently been proven in [1, Theorem 3.4]. The purpose of this section is to illustrate an application of Corollary 3.17 beyond the classical continuous or càdlàg setting.

We now provide a precise definition for solutions to VSDEs and introduce its parameters. The space $L^p_{\text{loc}}(\mathbb{R}_+, \mathbb{R}^d)$, for $p \geq 2$ and $d \in \mathbb{N}$, endowed with the local L^p-norm topology, will serve as state space of the process X in (4.5). Now, we introduce the following coefficients:

(D1) an initial value $g_0 \in L^p_{\text{loc}}(\mathbb{R}_+, \mathbb{R}^d)$;

(D2) a convolution kernel $K\colon \mathbb{R}_+ \to \mathbb{R}^{d\times k}$ in $L^p_{\text{loc}}(\mathbb{R}_+, \mathbb{R}^d)$;

(D3) a characteristic triplet (b, a, ν), consisting of two Borel functions $b\colon \mathbb{R}^d \to \mathbb{R}^k$ and $a\colon \mathbb{R}^d \to \mathbb{S}^k_+$ and a Borel transition kernel ν from $\mathbb{R}^d$ into $\mathbb{R}^k$ which does not charge the set $\{0\}$, such that there exists a constant $c > 0$ with

$$\|b(x)\| + \|a(x)\| + \int (1 \wedge \|y\|^2) \nu(x, dy) \leq c(1 + \|x\|^p)$$

for all $x \in \mathbb{R}^d$.

We are in the position to define solutions to the VSDE (4.5).

Definition 4.3. *A triplet $(\mathbb{B}, X, Z)$ is called a* weak solution *to the* Volterra SDE (VSDE) *associated to (g_0, K, b, a, ν), if $\mathbb{B} = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ is a stochastic basis which supports two processes X and Z, where X is $\mathbb{R}^d$-valued, predictable and has paths in $L^p_{\text{loc}}(\mathbb{R}_+, \mathbb{R}^d)$, Z is an $\mathbb{R}^k$-valued càdlàg semimartingale with differential characteristics $(b(X), a(X), \nu(X))$, and (4.5) holds $P \otimes \lambda$-almost everywhere.*

Let $(g^n_0, K^n, b^n, a^n, \nu^n)$ and (g_0, K, b, a, ν) be coefficients for VSDEs. Moreover, for every $f \in C^2_c(\mathbb{R}^k)$ and $(x, z) \in \mathbb{R}^d \times \mathbb{R}^k$ we set

$$\begin{aligned} \mathcal{L}f(x, z) \triangleq{}& \langle b(x), \nabla f(z)\rangle + \tfrac{1}{2} \operatorname{tr}(a(x)\nabla^2 f(z)) \\ &+ \int \big(f(z + y) - f(z) - \langle h(y), \nabla f(z)\rangle\big)\nu(x, dy). \end{aligned}$$

Similar to $\mathcal{L}$, we define $\mathcal{L}^n$ with (b, a, ν) replaced by (b^n, a^n, ν^n). Set $F = L^p_{\text{loc}}(\mathbb{R}_+, \mathbb{R}^d) \times \mathbb{D}(\mathbb{R}^k)$ endowed with the product topology, where $\mathbb{D}(\mathbb{R}^k)$ is endowed with the Skorokhod J_1 topology.

Theorem 4.4. *Let $(\mathbb{B}^n, X^n, Z^n)$ be a weak solution to the VSDE $(g^n_0, K^n, b^n, a^n, \nu^n)$ for every $n \in \mathbb{N}$, and let $(\Omega, \mathcal{F}, P)$ be a probability space which supports a measurable process (X, Z) with paths in F. Set $\mathcal{F}_t \triangleq \sigma(X_s, Z_s, s \leq t)$ for $t \in \mathbb{R}_+$ and $\mathbb{B} \triangleq (\Omega, \mathcal{F}, \mathbf{F} \triangleq (\mathcal{F}_t)_{t\geq 0}, P)$. Suppose that X is $\mathbf{F}$-progressively measurable, that (3.1) holds and assume the following:*

(L1) $(X^n, Z^n) \to (X, Z)$ *weakly in* F*;*

(L2) *for every* $f \in C^2_c(\mathbb{R}^k)$ *there exists a constant* $c_f > 0$ *such that*

$$|\mathcal{L}^n f(x,z)| \le c_f(1 + \|x\|^p), \quad (n,x,z) \in \mathbb{N} \times \mathbb{R}^d \times \mathbb{R}^k;$$

(L3) $\mathcal{L}f$ *is continuous for every* $f \in C^2_c(\mathbb{R}^k)$*;*

(L4) $g^n_0 \to g_0$ *and* $K^n \to K$ *in* $L^p_{\text{loc}}(\mathbb{R}_+, \mathbb{R}^d)$*, and*

$$\begin{aligned} f \in C^2_c(\mathbb{R}^k),\ \mathbb{R}^{d+k} \ni (x_n, z_n) \to (x,z) \in \mathbb{R}^{d+k} \\ \Longrightarrow \quad \big|\mathcal{L}^n f(x_n,z_n) - \mathcal{L}f(x_n,z_n)\big| \to 0. \end{aligned} \tag{4.6}$$

Then, $(\mathbb{B}, X, Z)$ *is a weak solution to the VSDE* (g_0, K, b, a, ν)*.*

Remark 4.5. (i) The assumption that X is $\mathbf{F}$-progressively measurable comes without loss of generality, see [31, Theorem 0.1].

(ii) It is clear that (4.6) holds whenever $\mathcal{L}^n f \to \mathcal{L}f$ locally uniformly for every $f \in C^2_c(\mathbb{R}^k)$. This condition and the continuity of all $\mathcal{L}^1 f, \mathcal{L}^2 f, \dots$ are assumed in [1]. Of course, this already implies (L3). In Theorem 4.4 we only ask for the weaker assumptions (L3) and (4.6) and consequently, Theorem 4.4 can also be applied when $\mathcal{L}^n$ has discontinuous coefficients.

Proof. Using standard extensions, we can assume that $\mathbb{B}^n$ and $\mathbb{B}$ support random variables U^n and U which are uniformly distributed on $[0,1]$ such that U^n is independent of (X^n, Z^n), U is independent of (X,Z) and $(U^n, X^n, Z^n) \to (U, X, Z)$ weakly as $n \to \infty$. Furthermore, we can redefine $\mathcal{F}^n_0 \triangleq \mathcal{F}^n_0 \vee \sigma(U^n)$ and $\mathcal{F}_0 \triangleq \mathcal{F}_0 \vee \sigma(U)$. Let $(\mathsf{U}, \mathsf{X}, \mathsf{Z})$ be the identity map on $[0,1] \times F$ and set

$$T_m \triangleq \inf\Big(t \in \mathbb{R}_+ \colon \int_0^t (1 + \mathsf{U} + \|\mathsf{X}_s\|^p) ds \ge m\Big), \quad m > 0.$$

Moreover, let $\mathfrak{X}^\circ$ be the set of all processes

$$f(\mathsf{Z}_{\cdot \wedge T_m}) - \int_0^{\cdot \wedge T_m} \mathcal{L}f(\mathsf{X}_s, \mathsf{Z}_s) ds, \quad m > 0, f \in C^2_c(\mathbb{R}^k),$$

and define $\mathfrak{X}^n$ to be the set of all processes

$$f(Z^n_{\cdot \wedge T_m(U^n, X^n)}) - \int_0^{\cdot \wedge T_m(U^n, X^n)} \mathcal{L}f(X^n_s, Z^n_s) ds, \quad m > 0, f \in C^2_c(\mathbb{R}^k).$$

The following lemma is a direct consequence of [1, Lemma 3.3].

Lemma 4.6. (i) $(\mathbb{B}, X, Z)$ *is a weak solution to the VSDE* (g_0, K, b, a, ν)*, if for every* $Y^\circ \in \mathfrak{X}^\circ$ *the process* $Y^\circ(U, X, Z)$ *is a martingale, and*

$$\int_0^t X_s ds = \int_0^t g_0(s) ds + \int_0^t K_{t-s} Z_s ds, \quad t \in \mathbb{R}_+. \tag{4.7}$$

(ii) *All processes in* $\mathfrak{X}^n$ *are martingales on* $\mathbb{B}^n$*.*

Equation (4.7) follows from (L1) and (L4), see [1, Lemma 3.5] for details. Thus, to conclude the claim of the theorem, it suffices to show that all processes in $\mathfrak{X} \triangleq \{Y^\circ(U, X, Z) \colon Y^\circ \in \mathfrak{X}^\circ\}$ are martingales. To show this we use Corollary 3.17.

First of all, note that (S1) in Corollary 3.17 coincides with (L1). Thus, we only need to verify (S2) and (S3), where we take

$$D \triangleq \big\{t > 0 \colon P(\Delta Z_t \neq 0) = 0\big\}.$$

Recall that D is dense in $\mathbb{R}_+$ (see [12, Lemma 3.7.7]).

For $t \in \mathbb{R}_+$, define $\mathcal{Z}^\circ_t$ to be the set of functions

$$\prod_{i=1}^m h_i\Big(\mathsf{U}, \int_0^{t_i} \mathsf{X}_s ds, \mathsf{Z}_{t_i}\Big),$$

where $m \in \mathbb{N}, t_1, \dots, t_m \in D \cap [0,t]$ and $h_1, \dots, h_m \in C_b(\mathbb{R}^{1+d+k})$. Recalling both parts of Example 3.6, we note that $\mathcal{Z}^\circ = \{\mathcal{Z}^\circ_t, t \in \mathbb{R}_+\}$ is a determining set for $\mathfrak{X}$. Thus, (S2) holds.

Finally, we verify (S3). By (D3) and the definition of T_m, all processes in $\mathfrak{X}^\circ$ are bounded on finite time intervals. Furthermore, by the definition of D, for every $t \in \mathbb{R}_+$, all elements of $\mathcal{Z}^\circ_t$ are P-a.s. continuous at (U, X, Z). Take $t \in D$ and $Y^\circ \in \mathfrak{X}^\circ$ such that

$$Y^\circ_t = f(\mathsf{Z}_{t\wedge T_m}) - \int_0^{t \wedge T_m} \mathcal{L} f(\mathsf{X}_s, \mathsf{Z}_s) ds.$$

Note that

$$\{T_m > s\} = \Big\{ \int_0^s (1 + \mathsf{U} + \|\mathsf{X}_u\|^p) du < m \Big\},$$
$$\{T_m < s\} = \Big\{ \int_0^s (1 + \mathsf{U} + \|\mathsf{X}_u\|^p) du > m \Big\},$$

because $s \mapsto \int_0^s (1 + \mathsf{U} + \|\mathsf{X}_u\|^p) du$ is strictly increasing. Hence, using the continuity of $(\mathsf{U}, \mathsf{X}) \mapsto \int_0^s (1 + \mathsf{U} + \|\mathsf{X}_u\|^p) du$, the map T_m is upper and lower semicontinuous and consequently, continuous. With this observation at hand, it follows easily from the continuity of $\mathcal{L} f$ that

$$(\mathsf{U}, \mathsf{X}, \mathsf{Z}) \mapsto \int_0^{t \wedge T_m(\mathsf{U},\mathsf{X})} \mathcal{L} f(\mathsf{X}_s, \mathsf{Z}_s) ds$$

is continuous, too. Thanks to the randomization given by U, a.s. Z does not jump at time $T_m(U, X)$, see [1] for details. Thus, $(\mathsf{U}, \mathsf{X}, \mathsf{Z}) \mapsto \mathsf{Z}_{t \wedge T_m(\mathsf{U},\mathsf{X})}$ is P-continuous at (U, X, Z) for every $t \in D$, by the definition of the set D and [25, Proposition VI.2.1]. In summary, Y°_t and $Y^\circ_t Z^\circ_s$ for $Z^\circ_s \in \mathcal{Z}^\circ_s$ and $s \le t$ are P-continuous at (U, X, Z) for every $t \in D$.

It is left to verify the final part of (S3). Let $Y^n \in \mathfrak{X}^n$ be given by

$$f(Z^n_{\cdot \wedge T_m(U^n, X^n)}) - \int_0^{\cdot \wedge T_m(U^n, X^n)} \mathcal{L}^n f(X^n_s, Z^n_s) ds.$$

Thanks to part (ii) of Lemma 4.6, Y^n is a martingale on $\mathbb{B}^n$. By Skorokhod's coupling theorem ([27, Theorem 3.30]), we can and will assume that (U^n, X^n, Z^n) and (U, X, Z) are defined on the same probability space and that a.s. $(U^n, X^n, Z^n) \to (U, X, Z)$ as $n \to \infty$. We now show that

$$E\big[\big|Y^n_t - Y^\circ_t(U^n, X^n, Z^n)\big|\big] \to 0 \text{ as } n \to \infty, \tag{4.8}$$

i.e. (3.6) in Proposition 3.18. This implies the last part in (S3) and thereby completes the proof. Notice that

$$Y^n_t - Y^\circ_t(U^n, X^n, Z^n) = \int_0^{t \wedge T_m(U^n, X^n)} (\mathcal{L}^n f(X^n_s, Z^n_s) - \mathcal{L} f(X^n_s, Z^n_s)) ds.$$

The implication (4.6) yields that a.s. for a.a. $s \in [0, t]$

$$\big|\mathcal{L}^n f(X^n_s, Z^n_s) - \mathcal{L} f(X^n_s, Z^n_s)\big| \to 0 \text{ as } n \to \infty.$$

As a.s. $X^n \to X$ in $L^p_{\text{loc}}(\mathbb{R}_+, \mathbb{R}^k)$, the family $\{\|X^n\|^p : n \in \mathbb{N}\}$ restricted to any finite time interval is a.s. uniformly integrable w.r.t. the Lebesgue measure. Thus, using (D3) and (L2), we deduce from Vitali's theorem that a.s.

$$\int_0^t \big|\mathcal{L}^n f(X^n_s, Z^n_s) - \mathcal{L} f(X^n_s, Z^n_s)\big| ds \to 0 \text{ as } n \to \infty.$$

Finally, since

$$\int_0^{t \wedge T_m(U^n, X^n)} \big|\mathcal{L}^n f(X^n_s, Z^n_s) - \mathcal{L} f(X^n_s, Z^n_s)\big| ds \lesssim 1 + m$$

by (D3), (L2) and the definition of T_m, the dominated convergence theorem yields (4.8). The proof is complete. □

4.3 An extension of a theorem by Jacod and Shiryaev

Let (B, C, ν) be a candidate triplet for semimartingale characteristics defined on the canonical space $\mathbb{D}(\mathbb{R}^d)$ endowed with the Skorokhod J_1 topology, see [25, III.2.3] for the technical requirements. Here, we assume that (B, C, ν) corresponds to a continuous truncation function $h \colon \mathbb{R}^d \to \mathbb{R}^d$.

For $m > 0$, we set

$$T_m(\omega) \triangleq \inf(t \in \mathbb{R}_+ : \|\omega(t)\| \geq m \text{ or } \|\omega(t-)\| \geq m), \quad \omega \in \mathbb{D}(\mathbb{R}^d), \tag{4.9}$$

$$\Theta_{m,t} \triangleq \Big\{\omega \in \mathbb{D}(\mathbb{R}^d) \colon \sup_{s \leq t} \|\omega(s)\| \leq m\Big\}.$$

Let $C_1(\mathbb{R}^d)$ be a subset of the set of non-negative bounded continuous functions vanishing around the origin as described in [25, VII.2.7].

The following theorem generalizes [25, Theorem IX.2.11] for the quasi-left continuous case as outlined on p. 533 in [25].

Theorem 4.7. *Let $(\Omega, \mathcal{F}, \mathbf{F}, P)$ and $(\Omega^n, \mathcal{F}^n, \mathbf{F}^n, P^n)$ be filtered probability spaces which support $\mathbb{R}^d$-valued càdlàg adapted processes X and X^n such that each X^n is a semimartingale with semimartingale characteristics (B^n, C^n, ν^n) corresponding to the (continuous) truncation function h. Assume that $X^n \to X$ weakly on $\mathbb{D}(\mathbb{R}^d)$ and that the following hold:*

(i) *there exists a set $\Gamma \subset \mathbb{R}_+$ with countable complement such that, for every $t \in \Gamma$, $m, \varepsilon > 0$ and $g \in C_1(\mathbb{R}^d)$, we have*

$$P^n(\|B^n_{t \wedge T_m(X^n)} - B_{t \wedge T_m(X^n)}(X^n)\| \geq \varepsilon) \to 0,$$
$$P^n(\|\widetilde{C}^n_{t \wedge T_m(X^n)} - \widetilde{C}_{t \wedge T_m(X^n)}(X^n)\| \geq \varepsilon) \to 0,$$
$$P^n(|g * \nu^n_{t \wedge T_m(X^n)} - g * \nu_{t \wedge T_m(X^n)}(X^n)| \geq \varepsilon) \to 0,$$

as $n \to \infty$;

(ii) *for all $t \in \mathbb{R}_+, m > 0$ and $g \in C_1(\mathbb{R}^d)$, we have*

$$\sup_{\omega \in \Theta_{m,t}} \big(\|\widetilde{C}_t(\omega)\| + |(g * \nu_t)(\omega)|\big) < \infty;$$

(iii) *there exists a dense set $\Gamma^* \subset \mathbb{R}_+$ such that, for all $t \in \Gamma^*$ and $g \in C_1(\mathbb{R}^d)$, the maps*

$$\mathbb{D}(\mathbb{R}^d) \ni \omega \mapsto B_t(\omega), \widetilde{C}_t(\omega), (g * \nu_t)(\omega)$$

are Skorokhod J_1 continuous;

(iv) *for every $m > 0$, there exists a continuous increasing function $F^m \colon \mathbb{R}_+ \to \mathbb{R}_+$ such that the processes*

$$F^m - \sum_{i=1}^d \operatorname{Var}(B^{(i)}_{\cdot\wedge T_m}); \quad F^m - \Big(\sum_{i=1}^d C^{(ii)}_{\cdot\wedge T_m} + (\|x\|^2 \wedge 1) * \nu_{\cdot\wedge T_m} \Big)$$

are increasing.

Then, X is a semimartingale with semimartingale characteristics $(B(X), C(X), \nu(X))$.

Proof. We deduce the result from Corollary 3.21, applied with a localized version of $\mathfrak{X}$ as defined in Example 3.9, see (i) – (iii) in Example 3.9. In the following we will verify (S4) in Corollary 3.21 for a localized version of the processes in (i). The argument for the processes from (ii) can be found in the proof of Theorem 5.8 below. For the processes in (iii) the argument is similar to those in (i), see the proof of [25, Theorem IX.2.11] for some details.

Due to [25, Propositions VI.2.11, VI.2.12] and the arguments in the proof of [25, Proposition IX.1.17], there exists an increasing sequence $(k_n)_{n\in\mathbb{N}} \subset \mathbb{R}_+$ with $k_n \to \infty$ such that the maps

$$\mathbb{D}(\mathbb{R}^d) \ni \omega \mapsto T^n(\omega) \triangleq T_{k_n}(\omega), \omega(\cdot \wedge T^n(\omega))$$

are P-a.s. Skorokhod J_1 continuous at X. Let Γ be as in (i) and set

$$D \triangleq \big\{ t \in \Gamma \colon P(\Delta X^{T^m(X)}_t \neq 0) = 0 \text{ for all } m \in \mathbb{N} \big\}, \quad X^{T^m(X)} \triangleq X_{\cdot\wedge T^m(X)}.$$

As D^c is countable, D is dense in $\mathbb{R}_+$.

Fix $m \in \mathbb{N}, T \in D$, and let $K = K(m,T) > 0$ be such that

$$\sup_{\omega\in\Theta_{k_m,T}} \|\widetilde{C}_T(\omega)\| \leq K, \tag{4.10}$$

see hypothesis (ii). We define

$$S^n \triangleq \inf(t \in \mathbb{R}_+ \colon \|\widetilde{C}^n_{t\wedge T^m(X^n)}\| \geq K+1).$$

Let us recall that

$$\omega(h) \triangleq \omega - \sum_{s\leq\cdot} (\Delta\omega(s) - h(\Delta\omega(s))), \quad \omega \in \mathbb{D}(\mathbb{R}^d),$$

where h is the continuous truncation function we have fixed in the beginning of this section. We take

$$Y^\circ(\omega) \triangleq \omega(h)_{\cdot\wedge T^m(\omega)\wedge T} - \omega(0) - B_{\cdot\wedge T^m(\omega)\wedge T}(\omega),$$

and

$$Y^n \triangleq X^n(h)_{\cdot\wedge T^m(X^n)\wedge S^n\wedge T} - X^n_0 - B^n_{\cdot\wedge T^m(X^n)\wedge S^n\wedge T}.$$

Recalling (4.10), thanks to hypothesis (i), we obtain

$$\begin{aligned} P^n(S^n \leq T) &= P^n(\|\widetilde{C}^n_{T\wedge T^m(X^n)}\| \geq K+1) \\ &\leq P^n(\|\widetilde{C}^n_{T\wedge T^m(X^n)} - \widetilde{C}_{T\wedge T^m(X^n)}(X^n)\| \geq 1) \to 0 \end{aligned}$$

as $n \to \infty$. Thus, using (i) again, for all $t \in \Gamma$ and $\varepsilon > 0$, we get

$$P^n(\|Y^\circ_t(X^n) - Y^n_t\| \geq \varepsilon) \leq P^n(\|B^n_{t\wedge T\wedge T_m(X^n)} - B_{t\wedge T\wedge T_m(X^n)}(X^n)\| \geq \varepsilon)$$

$$+ P^n(S^n \le T) \to 0$$

as $n \to \infty$. Due to hypothesis (iii) and (iv) and the P-a.s. Skorokhod J_1 continuity of $\omega \mapsto T^m(\omega)$ at X, it follows from [25, IX.3.42] that, for every $t \in \mathbb{R}_+$, the map $\omega \mapsto B_{t \wedge T^m(\omega)}(\omega)$ is P-a.s. Skorokhod J_1 continuous at X. Moreover, whenever $t \in D$, [25, VI.2.3, Corollary VI.2.8] show that the map $\omega \mapsto \omega(h)_{t \wedge T^m(\omega)}$ is P-a.s. Skorokhod J_1 continuous at X. Consequently, for every $t \in D$, the map $\omega \mapsto Y^\circ_t(\omega)$ is P-a.s. Skorokhod J_1 continuous at X.

By the martingale problem for semimartingales (see Example 3.9 or [25, Theorem II.2.21]), Y^n is a locally square-integrable P^n-martingale whose predictable quadratic variation process is given by $\widetilde{C}^n_{\cdot \wedge T^m(X^n) \wedge S^n \wedge T}$. Hence, it follows from Doob's inequality ([25, Theorem I.1.43]) that, for all $a > 0$,

$$E^{P^n}\Big[\sup_{s \le a} |Y^{n,(i)}_s|^2\Big] \le 4 E^{P^n}\Big[\widetilde{C}^{n,(ii)}_{a \wedge T^m(X^n) \wedge S^n \wedge T}\Big]. \tag{4.11}$$

As $|\Delta \widetilde{C}^{n,(ij)}| \le 2\|h\|^2_\infty$, the definition of S^n yields that the r.h.s. of (4.11) is bounded uniformly in n. Consequently, $\{Y^n_t \colon t \in [0,a], n \in \mathbb{N}\}$ is uniformly integrable.

In summary, Y° and $(Y^n)_{n \in \mathbb{N}}$ have the properties as in (S4). As mentioned at the beginning of this proof, similar arguments work for suitably localized versions of the processes defined in (ii) and (iii) of Example 3.9. We omit the remaining details. $\square$

The third part of hypothesis (i) plus hypothesis (iv) yield quasi-left continuity of the limit, see the proof of [25, Theorem IX.3.21]. In [25, Theorem IX.2.11] this assumption is not needed, but it is assumed that part (iii) holds P-a.s. at X. Although the dependence on P is sort of minimal, one can only benefit from it when the limit is more or less known, see [25, Remark IX.2.13]. The monograph [25] suggests two deterministic versions of this condition. Namely, a version of (iii) ([25, IX.2.14]) and [25, IX.2.16], i.e. continuity of $\omega \mapsto B^i(\omega), \widetilde{C}^{ij}(\omega), (g * \nu)(\omega)$ from $\mathbb{D}(\mathbb{R}^d)$ into $\mathbb{D}(\mathbb{R})$. As functions of the type

$$\omega \mapsto \int_0^t f(\omega(s-)) q(ds) \tag{4.12}$$

are not necessarily continuous in the Skorokhod J_1 topology when q is allowed to have point masses, both of these assumptions might be too stringent for applications with fixed times of discontinuity. To give an example, consider $d = 1$ and

$$\omega \mapsto F_t(\omega) \triangleq \int_0^t \omega(s-)\delta_1(ds) = \begin{cases} 0, & t < 1, \\ \omega(1-), & t \ge 1, \end{cases} \quad \omega \in \mathbb{D}(\mathbb{R}).$$

The function $\omega \mapsto F_t(\omega)$ is obviously continuous if $t < 1$, but it is discontinuous for all $t \ge 1$, as is easily seen by taking $\omega_n = \mathbb{1}_{[1-1/n,\infty)} \to \omega = \mathbb{1}_{[1,\infty)}$. Thus, for this example there is no dense set $\Gamma \subset \mathbb{R}_+$ such that F_t is continuous for all $t \in \Gamma$. Moreover, $\omega \mapsto F(\omega)$ is also not continuous from $\mathbb{D}(\mathbb{R})$ into $\mathbb{D}(\mathbb{R})$. Indeed, if F would be continuous we must have $\omega_n \to \omega \Rightarrow F_t(\omega_n) \to F_t(\omega)$ for all $t \ne 1$ as $\{s > 0 \colon \Delta F_s(\omega) \ne 0\} \subset \{1\}$, which is not true. Therefore, we note that functions of the type (4.12) do not necessarily have the continuity properties from [25, IX.2.14, IX.2.16].

In Section 5.5 below we discuss versions of Theorem 4.7 where in (iii) the Skorokod J_1 topology is replaced by the local uniform topology, which seems to us more suitable for applications to semimartingales with fixed times of discontinuity.

5 Stability results for processes with fixed times of discontinuity

In this section we establish stability results which are tailored to processes with fixed times of discontinuity. To be more precise, in Section 5.3 we derive a version of

Theorem 4.1 which applies to test processes of the type

$$f(X_t) - \int_0^t g(s, L, X)ds - \int_0^t h(s, L, X)q(ds), \quad t \in \mathbb{R}_+,$$

where q is a locally finite Borel measure on $\mathbb{R}_+$ which is allowed to have point masses, and f, g and h are suitable function of the processes to be well-defined. In Section 5.5 we prove a version of Theorem 4.7 for semimartingales whose characteristics are only assumed to be continuous in the local uniform instead of the Skorokhod J_1 topology. In both cases we work with *control variables* $L^1, L^2, \ldots$ and the notion of weak-strong convergence. Before we present our results, we motivate the presence of fixed times of discontinuities.

5.1 Motivation

Continuous state branching processes (CSBP) are analogues of Galton–Watson processes in continuous time with continuous state spaces. Typically, CSBP are modeled as strong solutions to SDEs driven by a Brownian motion and a Poisson random measure. More recently, there is an increasing interest in CSBPs in random environments (CSBPRE), where the random environments are modeled by additional independent, multiplicative and (sometimes) discontinuous noise, see, e.g. [4, 5]. Leaving the environment random is often called the *annealed* perspective. In contrast, fixing the random environment corresponds to the so-called *quenched* perspective. In case the environment is represented by discontinuous noise, taking a specific path introduces fixed times of discontinuity, which therefore arise in a natural manner in the context of CSBPRE. To the best of our knowledge, the literature contains only selected stability results for quenched dynamics of CSBPRE, see, e.g. [3, 5].

Fixed times of discontinuity also occur naturally in mathematical finance such as in interest rate markets in the post-crisis environment. Indeed, a closer look on historical data of European reference interest rates (see [14, Figure 1]) shows jumps at prescheduled dates. As a consequence, the financial literature shows an increasing interest in stochastic models for interest rates which allow for fixed times of discontinuity, see, for instance, [14, 29].

5.2 A short example to keep in mind

Before we start our theoretical program, let us explain one explicit situation to keep in mind when reading the remainder of this section. Suppose we are interested in identifying the limiting process for a sequence $X^1, X^2, \ldots$ of one-dimensional processes whose dynamics are given by the equations

$$dX^n_t = \sigma^n(X^n_t)dW^n_t - \int_0^1 (1-\theta)X^n_{t-}q^n(dt, d\theta),$$

where $W^1, W^2, \ldots$ is a sequence of (one-dimensional) Brownian motions and $q^1, q^2, \ldots$ is a sequence of deterministic measures on $\mathbb{R}_+ \times [0, 1]$ such that

$$q^n(dt, d\theta) = \sum_{i=1}^{m_n} \delta_{(t_i, w_i)}(dt, d\theta)$$

for some $m_n \in \mathbb{Z}_+, 0 \le t_1 < t_2 < \cdots < t_{m_n} < \infty$ and $w_1, \ldots, w_{m_n} \in [0, 1]$. As in the paper [3], one might think of q^n as a realization of a Poisson random measure with intensity measure $dt \otimes P(C \in d\theta)$, where C is some random variable with values in the unit interval. Fixing such a realization corresponds to the quenched perspective as explained in the previous motivating section.

In case the sequence $\sigma^1, \sigma^2, \ldots$ converges in a suitable sense to a limiting coefficient σ and in case the sequence $q^1, q^2, \ldots$ converges in a suitable sense to a limiting measure q, it is natural to expect that the limiting object of the sequence $X^1, X^2, \ldots$ is a process X whose dynamics are given by the equation

$$dX_t = \sigma(X_t)dW_t - \int_0^1 (1-\theta)X_{t-}q(dt, d\theta), \tag{5.1}$$

where W is a Brownian motion. Indeed, by an application of Itô's formula, for $f \in C_b^2(\mathbb{R})$, the process

$$f(X^n) - \int_0^\cdot \tfrac{1}{2}(\sigma^n(X^n_s))^2 f''(X^n_s)ds - \int_0^\cdot \int_0^1 \big(f(\theta X^n_{s-}) - f(X^n_{s-})\big)q^n(ds, d\theta) \tag{5.2}$$

is a local martingale. Under assumptions of the type $\sigma^n \to \sigma$ and $q^n \to q$, where the convergence is meant to be in a suitable sense, it is reasonable to expect that the local martingale property of the process (5.2) transfers to the process

$$f(X) - \int_0^\cdot \tfrac{1}{2}(\sigma(X_s))^2 f''(X_s)ds - \int_0^\cdot \int_0^1 \big(f(\theta X_{s-}) - f(X_{s-})\big)q(ds, d\theta),$$

which relates the limiting process X to the dynamics given by (5.1).

Although this conjecture seems intuitively reasonable, it cannot be deduced from classical results (as in [25], for instance) due to the lack of certain continuity properties for the Skorokhod J_1 topology, see the discussion at the end of Section 4.3 for more details. In the following we will derive some results to overcome such technical difficulties. In Section 5.4 below we will return to (a slightly more general version of) this example and explain more precisely how our theoretical results can be used.

5.3 A version of the Ethier–Kurtz theorem with fixed times of discontinuity

In this section we derive a version of Theorem 4.1 which allows fixed times of discontinuity. Let (E, r) be a Polish space, let $(U, \mathcal{U})$ be a measurable space and let $\mathbb{B} = (\Omega, \mathcal{F}, \mathbf{F}, P)$ and $\mathbb{B}^n = (\Omega^n, \mathcal{F}^n, \mathbf{F}^n, P^n)$, $n = 1, 2, 3, \ldots$, be filtered probability spaces which support E-valued càdlàg adapted processes X and X^n and U-valued random variables L and L^n, respectively. The laws of $X, X^1, X^2, \ldots$ are considered to be Borel probability measures on $\mathbb{D}(E)$, which we endow with the Skorokhod J_1 topology in this section.

We now introduce the martingale problem of our current interest. Let $\mathscr{P}_{luc}$ be the set of all bounded measurable functions $f\colon \mathbb{R}_+ \times U \times \mathbb{D}(E) \to \mathbb{R}$ such that $\Omega \times \mathbb{R}_+ \ni (\omega, t) \mapsto f(t, L(\omega), X(\omega))$ is $\mathbf{F}$-predictable and $\mathbb{D}(E) \ni \alpha \mapsto f(t, u, \alpha)$ is continuous in the local uniform topology for every $(t, u) \in \mathbb{R}_+ \times U$. We parameterize the martingale problem by a set $D \subset C_b(E) \times \mathscr{P}_{luc} \times \mathscr{P}_{luc}$ and a locally finite Borel measure q on $\mathbb{R}_+$. Namely, we define $\mathfrak{X}$ to be the set of all processes

$$f(X) - \int_0^\cdot g(s, L, X)ds - \int_0^\cdot h(s, L, X)q(ds), \quad (f, g, h) \in D. \tag{5.3}$$

Next, we also introduce a set of approximating martingales on $\mathbb{B}^1, \mathbb{B}^2, \ldots$. Let $q^1, q^2, \ldots$ be a sequence of locally finite Borel measures on $\mathbb{R}_+$. Moreover, for every $n \in \mathbb{N}$, let $\mathfrak{X}^n$ be a set of triplets (ξ, ϕ, ψ) consisting of real-valued predictable processes on $\mathbb{B}^n$ such that

$$\sup_{s \le T} E^{P^n}\big[|\xi_s| + |\phi_s| + |\psi_s|\big] < \infty, \quad T > 0,$$

and such that

$$\xi - \int_0^\cdot \phi_s ds - \int_0^\cdot \psi_s q^n(ds)$$

is a martingale on $\mathbb{B}^n$.

Finally, we introduce some technical ingredients. Take a $\mathcal{U}\otimes\mathcal{B}(\mathbb{R}_+)/\mathcal{B}(\mathbb{R}_+)$-measurable function $\mathfrak{u}\colon U \times \mathbb{R}_+ \to \mathbb{R}_+$ such that, for every $(u,t) \in U \times (0,\infty)$,

$$\lim_{\varepsilon \searrow 0} \sup \big\{ \mathfrak{u}(u,s)\colon s \neq t, t - \varepsilon \leq s \leq t + \varepsilon \big\} = 0, \tag{5.4}$$

take an increasing and continuous function $\kappa\colon \mathbb{R}_+ \to \mathbb{R}_+$ and fix a reference point $x_0 \in E$. We define

$$\begin{aligned} A &\triangleq \Big\{ (u,\omega) \in U \times \mathbb{D}(E)\colon r(\omega(t), \omega(t-)) \leq \mathfrak{u}(u,t)\kappa\Big(\sup_{s\leq t} r(\omega(s), x_0)\Big) \text{ for all } t > 0 \Big\} \\ &= \bigcap_{n\in\mathbb{N}} \Big\{ (u,\omega)\colon r(\omega(S_n(\omega)), \omega(S_n(\omega)-)) \leq \mathfrak{u}(u, S_n(\omega))\kappa\Big(\sup_{s\leq S_n(\omega)} r(\omega(s), x_0)\Big)\Big\}, \end{aligned}$$

where $(S_n)_{n\in\mathbb{N}}$ is an exhausting sequences for the jumps of the coordinate process. Clearly, the second line shows that $A \in \mathcal{U} \otimes \mathcal{B}(\mathbb{D}(E))$.

Assumption 5.1. We have $P^n \circ (L^n, X^n)^{-1} \to_{ws} P \circ (L, X)^{-1}$ and

$$P^n((L^n, X^n) \in A) \to 1 \tag{5.5}$$

as $n \to \infty$.

Remark 5.2. If $X^n \to X$ weakly and $L, L^1, L^2, \ldots$ have the same distribution, then Theorem 2.5 shows that, along a subsequence, $(L^n, X^n) \to_{ws} (L, X)$.

Discussion. The main idea behind Assumption 5.1 is inspired by proofs from [22, 24] for stability and existence results of SDEs with semimartingale drivers. We think of (5.5) as an assumption on the jump structure of $X^1, X^2, \ldots$. More precisely, it means that the jump times of $X^1, X^2, \ldots$ can be controlled via $L^1, L^2, \ldots$ in an uniform manner and that the latter sequence behaves nicely in the sense that it converges in a rather strong sense.

To get a better understanding of (5.5), suppose that $L^1, L^2, \ldots$ are càdlàg E-valued processes which can be seen as drivers for the processes $X^1, X^2, \ldots$. Further, suppose that $\mathfrak{u}(L^n, t) = r(L^n_t, L^n_{t-})$ for $t > 0$ and $n = 1, 2, \ldots$, see also Examples 5.12 and 5.13 below. In this case $\mathfrak{u}$ has all required properties (see Lemma 2.10) and

$$(L^n, X^n) \in A \quad \Longleftrightarrow \quad r(X^n_t, X^n_{t-}) \leq r(L^n_t, L^n_{t-})\, \kappa\Big(\sup_{s\leq t} r(X^n_s, x_0)\Big) \ \ \forall t > 0.$$

In other words, (5.5) means that, with probability tending to one, the jump times of $X^1, X^2, \ldots$ are controlled by those of $L^1, L^2, \ldots$.

When it comes to applications, the crucial point behind (5.5) is the choice of $\mathfrak{u}$ and $L^1, L^2, \ldots$. In Section 5.4 below we explain how these objects can be chosen for a slight extension of the example from Section 5.2.

Let us also provide some technical comments. For notational convenience, set $Q^n \triangleq P^n \circ (L^n, X^n)^{-1}$ and $Q \triangleq P \circ (L, X)^{-1}$. By virtue of Proposition 2.12, the continuity assumption in the definition of $\mathscr{P}_{luc}$ implies that the test processes in (5.3) are (Q_n, Q)-continuous. More precisely, for (Q_n, Q)-continuity we can treat the control variables as deterministic, i.e. we only need to verify the Skorokhod J_1 continuity of the time t values of the processes (5.3) on the sections A_u for each $u \in U$. On A_u the Skorokhod J_1 and the local uniform topology coincide by Proposition 2.12. Roughly speaking, the paths of $X^1, X^2, \dots$ take values in a randomized subset of the Skorokhod space $\mathbb{D}(E)$ with conditionally nice topological properties which can be used thanks to the concept of weak-strong convergence.

Assumption 5.3. Let $\Gamma \subset \mathbb{R}_+$ be a dense set and let $\mathcal{Z}^\circ = \{\mathcal{Z}^\circ_t, t \in \mathbb{R}_+\}$ be a determining set for $\mathfrak{X}$ such that, for every $t \in \Gamma$ and all $n \in \mathbb{N}, u \in U$ and $Z^\circ_t \in \mathcal{Z}^\circ_t$, $Z^\circ_t(L^n, X^n)$ is $\mathcal{F}^n_t$-measurable and $\mathbb{D}(E) \ni \alpha \mapsto Z_t(u, \alpha)$ is continuous in the local uniform topology. For each $(f, g, h) \in D$ and $T > 0$, there exists a sequence $(\xi^n, \phi^n, \psi^n) \in \mathfrak{X}^n$ such that

$$\sup_{n \in \mathbb{N}} \sup_{r \leq T} E^{P^n}\big[|\xi^n_r| + |\phi^n_r| + |\psi^n_r|\big] < \infty, \tag{5.6}$$

$$\lim_{n \to \infty} E^{P^n}\big[(\xi^n_t - f(X^n_t))\, Z^\circ_s(L^n, X^n)\big] = 0, \tag{5.7}$$

$$\lim_{n \to \infty} E^{P^n}\Big[\int_s^t (\phi^n_u - g(u, L^n, X^n))du\, Z^\circ_s(L^n, X^n)\Big] = 0, \tag{5.8}$$

$$\lim_{n \to \infty} E^{P^n}\Big[\int_s^t (\psi^n_u - h(u, L^n, X^n))q(du)\, Z^\circ_s(L^n, X^n)\Big] = 0, \tag{5.9}$$

$$\lim_{n \to \infty} E^{P^n}\Big[\int_s^t \psi^n_u (q^n - q)(du)\, Z^\circ_s(L^n, X^n)\Big] = 0, \tag{5.10}$$

for all $s, t \in \Gamma \cap [0, T], s < t$, and $Z^\circ_s \in \mathcal{Z}^\circ_s$.

The following is a version of Theorem 4.1 which allows fixed times of discontinuity.

Theorem 5.4. *Suppose that the Assumptions 5.1 and 5.3 hold. Then,* $P \in \mathcal{M}(\mathfrak{X})$.

Proof. We apply Theorem 3.14 with the product space $(S, \mathcal{S}) = (U \times \mathbb{D}(E), \mathcal{U} \otimes \mathcal{B}(\mathbb{D}(E)))$. We set $Q^n \triangleq P^n \circ (L^n, X^n)^{-1}$ and $Q \triangleq P \circ (L, X)^{-1}$. Evidently, (A1) and (A2) hold by virtue of Assumption 5.1. Thus, it suffices to show that (A3) holds. Each $Y^\circ \in \mathfrak{X}$ is bounded on compact time intervals, which yields the uniform integrability assumption from (A3). Simply by hypothesis, for each $t \in \Gamma, s \in \Gamma \cap [0, t], u \in U$ and $Z^\circ_s \in \mathcal{Z}^\circ_s$, the maps $\alpha \mapsto Y^\circ_t(u, \alpha)$ and $\alpha \mapsto Y^\circ_t(u, \alpha) Z^\circ_s(u, \alpha)$ are continuous in the local uniform topology. Thus, (5.5) and Propositions 2.8 and 2.12 show the (Q^n, Q)-continuity assumption in (A3). Finally, it remains to show (3.2). Take $Y \in \mathfrak{X}$ such that

$$Y = f(X) - \int_0^\cdot g(s, L, X)ds - \int_0^\cdot h(s, L, X)q(ds),$$

and set

$$Y^n \triangleq \xi^n - \int_0^\cdot \phi^n_s ds - \int_0^\cdot \psi^n_s q^n(ds),$$

where ξ^n, ϕ^n and ψ^n are as in (5.6) – (5.10). Let $s, t \in \Gamma$ with $s < t$ and take $Z^\circ_s \in \mathcal{Z}^\circ_s$. We have

$$\begin{aligned}
(Y^\circ_t(L^n, X^n) - Y^n_t + Y^n_s - Y^\circ_s(L^n, X^n))Z^\circ_s(L^n, X^n)& \\
= (f(X^n_t) - \xi^n_t + \xi^n_s - f(X^n_s))Z^\circ_s(L^n, X^n)& \\
+ \int_s^t (\phi^n_u - g(u, L^n, X^n))du Z^\circ_s(L^n, X^n)&
\end{aligned}$$

$$+ \int_s^t (\psi_u^n - h(u, L^n, X^n)) q(du) Z_s^\circ(L^n, X^n)$$
$$+ \int_s^t \psi_u^n (q^n - q)(du) Z_s^\circ(L^n, X^n).$$

The P^n-expectation of the first term converges to zero by (5.7), the P^n-expectation of the second term converges to zero be (5.8), the P^n-expectation of the third term converges to zero by (5.9), and the P^n-expectation of the last term convergences to zero by (5.10). We now can proceed as in the proof of Theorem 4.1. As Y^n is a martingale on $\mathbb{B}^n$, we have

$$E^{P^n}\big[(Y_s^n - Y_t^n) Z_s^\circ(L^n, X^n)\big] = 0,$$

and consequently,

$$\lim_{n\to\infty} E^{P^n}\big[(Y_t^\circ(L^n, X^n) - Y_s^\circ(L^n, X^n)) Z_s^\circ(L^n, X^n)\big]$$
$$= \lim_{n\to\infty} E^{P^n}\big[(Y_t^\circ(L^n, X^n) - Y_t^n + Y_s^n - Y_s^\circ(L^n, X^n)) Z_s^\circ(L^n, X^n)\big] = 0.$$

We conclude that (A3) holds. Hence, the claim follows from Theorem 3.14. □

Example 5.5. An important example is the natural extension of the setting from Ethier and Kurtz [12], where $\mathfrak{X}$ is given through processes as in (1.3), i.e.

$$f(X_t) - \int_0^t g(X_s) ds - \int_0^t h(X_{s-}) q(ds), \quad t \in \mathbb{R}_+,$$

for $f, g, h \in C_b(E)$. Clearly, for every $t \in \mathbb{R}_+$, the map

$$\mathbb{D}(E) \ni \alpha \mapsto f(\alpha(t)) - \int_0^t g(\alpha(s)) ds - \int_0^t h(\alpha(s-)) q(ds)$$

is continuous in the local uniform topology. Thus, the above continuity assumptions on $\mathfrak{X}$ are fulfilled. Furthermore, in this case we can take $\mathbf{F}$ to be the natural filtration generated by X and consequently, $\mathcal{Z}^\circ$ can be taken as in part (i) of Example 3.6. It is also not hard to see that these processes have the necessary continuity properties when D (from Example 3.6) is defined to be the set of all times $t \in \mathbb{R}_+$ such that $P(\Delta X_t \neq 0) = 0$. Finally, we stress that in this setup the control variable L can be constructed from the sequence $L^1, L^2, \dots$. More precisely, in case $X^n \to X$ weakly on $\mathbb{D}(E)$ and the distributions of $L^1, L^2, \dots$ are relatively (sequentially) compact in $M_m(U)$, then, by Theorem 2.5, there exists a weak-strong convergent subsequence of $Q^n = P^n \circ (L^n, X^n)^{-1}, n = 1, 2, \dots$. In this case, as we are only interested in the law of X, we can start with a stochastic basis $\mathbb{B}$ which supports a random variable (X, L) which is distributed according to a weak-strong accumulation point.

We emphasis that, although the test processes for the limiting martingale problem $\mathfrak{X}$ are independent of L, the sequence $L^1, L^2, \dots$ is crucial for the proof to replace the Skorokhod J_1 topology by the local uniform topology.

We now also consider the problem of verifying tightness of the family $\{X^n \colon n \in \mathbb{N}\}$. A quite general criterion for tightness, which can be viewed as a version of Aldous' criterion for processes with fixed times of discontinuity, has recently been proved in [3]. In the following we present an application of this tightness criterion in the spirit of [12, Theorem 3.9.4]. For a compact set $K \subset E$, we set

$$T_K^n \triangleq \inf(t \in \mathbb{R}_+ \colon X_t^n \notin K), \quad n \in \mathbb{N}.$$

At least when $\mathbf{F}^n$ is right-continuous, which we assume without loss of generality, it is well-known that T^n_K is an $\mathbf{F}^n$-stopping time. We also define the stochastic interval

$$[\![0, T^n_K]\!] \triangleq \{(\omega, t) \in \Omega^n \times \mathbb{R}_+ : 0 \le t \le T^n_K(\omega)\}.$$

We have the following tightness condition for $\{X^n : n \in \mathbb{N}\}$.

Theorem 5.6. *Assume the following:*

(i) *for every $\eta, T > 0$ there exists a compact set $K = K(\eta, T) \subset E$ such that*

$$\inf_{n \in \mathbb{N}} P^n\big(X^n_t \in K \text{ for all } 0 \le t \le T\big) \ge 1 - \eta;$$

(ii) *let $H \subset C_b(E)$ be a subalgebra which is dense for the uniform topology on compact subsets of E. For every $f \in H$ and any compact set $K \subset E$, there exists sequences $\phi^1, \phi^2, \dots$ and $\psi^1, \psi^2, \dots$, which of course might depend on f and K, such that $(f(X^n), \phi^n, \psi^n) \in \mathfrak{X}^n$ and*

$$\sup\big\{|\phi^n_s(\omega)| + |\psi^n_s(\omega)| : n \in \mathbb{N}, (\omega, s) \in [\![0, T^n_K]\!]\big\} < \infty, \quad T > 0;$$

(iii) *there exists an increasing càdlàg function $Q : \mathbb{R}_+ \to \mathbb{R}_+$ with $Q(0) = 0$ such that $q^n([0, \cdot]) \to Q$ as $n \to \infty$ in the Skorokhod J_1 topology.*

Then, the family $\{X^n : n \in \mathbb{N}\}$ is tight (in the Skorokhod space with the Skorokhod J_1 topology).

Remark 5.7. Thanks to [25, Theorem VI.2.15], $q^n([0, \cdot]) \to Q$ in the Skorokhod J_1 topology if and only if there exists a dense set $I \subset \mathbb{R}_+$ such that, for all $t \in I$,

$$q^n([0, t]) \to Q(t), \qquad \sum_{0 < s \le t} |q^n(\{s\})|^2 \to \sum_{0 < s \le t} |\Delta Q(s)|^2.$$

Here, we note that the l.h.s. means that q^n converges weakly to the measure induced by Q. The r.h.s. is an additional requirement.

Proof of Theorem 5.6. Step 1: Tightness of $\{f(X^n) : n \in \mathbb{N}\}$. Take $f \in H$ and fix a compact set $K \subset E$. As H is a subalgebra, $f^2 \in H$. Denote by $\phi^1, \phi^2, \dots$ and $\psi^1, \psi^2, \dots$ the sequences from (ii) for f and K, and let $\hat{\phi}^1, \hat{\phi}^2, \dots$ and $\hat{\psi}^1, \hat{\psi}^2, \dots$ be the sequences for f^2 and K. Fix $0 \le s \le t$. We compute

$$\begin{aligned}
&E\big[(f(X^n_{t \wedge T^K_n}) - f(X^n_{s \wedge T^n_K}))^2 | \mathcal{F}^n_s\big] \\
&\quad = E\big[f^2(X^n_{t \wedge T^n_K}) - 2 f(X^n_{t \wedge T^n_K}) f(X^n_{s \wedge T^n_K}) + f^2(X^n_{s \wedge T^n_K}) | \mathcal{F}^n_s\big] \\
&\quad = E\Big[f^2(X^n_{t \wedge T^n_K}) - \int_0^{t \wedge T^n_K} \hat{\phi}^n_r dr - \int_0^{t \wedge T^n_K} \hat{\psi}^n_r q^n(dr) \big| \mathcal{F}^n_s\Big] \\
&\qquad\quad + E\Big[\int_0^{t \wedge T^n_K} \hat{\phi}^n_r dr + \int_0^{t \wedge T^n_K} \hat{\psi}^n_r q^n(dr) \big| \mathcal{F}^n_s\Big] \\
&\qquad\quad - 2 f(X^n_{s \wedge T^n_K}) E\Big[f(X^n_{t \wedge T^n_K}) - \int_0^{t \wedge T^n_K} \phi^n_r dr - \int_0^{t \wedge T^n_K} \psi^n_r q^n(dr) \big| \mathcal{F}^n_s\Big] \\
&\qquad\quad - 2 f(X^n_{s \wedge T^n_K}) E\Big[\int_0^{t \wedge T^n_K} \phi^n_r dr + \int_0^{t \wedge T^n_K} \psi^n_r q^n(dr) \big| \mathcal{F}^n_s\Big] + f^2(X^n_{s \wedge T^n_K}) \\
&\quad = f^2(X^n_{s \wedge T^n_K}) - \int_0^{s \wedge T^n_K} \hat{\phi}^n_r dr - \int_0^{s \wedge T^n_K} \hat{\psi}^n_r q^n(dr) \\
&\qquad\quad + E\Big[\int_0^{t \wedge T^n_K} \hat{\phi}^n_r dr + \int_0^{t \wedge T^n_K} \hat{\psi}^n_r q^n(dr) \big| \mathcal{F}^n_s\Big]
\end{aligned}$$

$$
\begin{aligned}
&\quad - 2f^2(X^n_{s\wedge T^n_K}) + 2f(X^n_{s\wedge T^n_K})\Big(\int_0^{s\wedge T^n_K} \phi^n_r dr + \int_0^{s\wedge T^n_K} \psi^n_r q^n(dr)\Big) \\
&\quad - 2f(X^n_{s\wedge T^n_K}) E\Big[\int_0^{t\wedge T^n_K} \phi^n_r dr + \int_0^{t\wedge T^n_K} \psi^n_r q^n(dr)\big|\mathcal{F}^n_s\Big] + f^2(X^n_{s\wedge T^n_K}) \\
&= E\Big[\int_{s\wedge T^n_K}^{t\wedge T^n_K} \hat{\phi}^n_r dr + \int_{s\wedge T^n_K}^{t\wedge T^n_K} \hat{\psi}^n_r q^n(dr)\big|\mathcal{F}^n_s\Big] \\
&\quad - 2f(X^n_{s\wedge T^n_K}) E\Big[\int_{s\wedge T^n_K}^{t\wedge T^n_K} \phi^n_r dr + \int_{s\wedge T^n_K}^{t\wedge T^n_K} \psi^n_r q^n(dr)\big|\mathcal{F}^n_s\Big].
\end{aligned}
$$

Now, using the (local) boundedness of $\phi^1, \phi^2, \dots \psi^1, \psi^2, \dots, \hat{\phi}^1, \hat{\phi}^2, \dots$ and $\hat{\psi}^1, \hat{\psi}^2, \dots$ as assumed in (ii), we obtain the existence of a constant $C > 0$, which might depend on f, K and the sequences $\phi^1, \phi^2, \dots \psi^1, \psi^2, \dots, \hat{\phi}^1, \hat{\phi}^2, \dots$ and $\hat{\psi}^1, \hat{\psi}^2, \dots$, such that

$$
\begin{aligned}
E\big[(f(X^n_{t\wedge T^K_n}) - f(X^n_{s\wedge T^n_K}))^2|\mathcal{F}^n_s\big] &\le CE\Big[t\wedge T^n_K - s\wedge T^n_K + \int_{s\wedge T^n_K}^{t\wedge T^n_K} q^n(dr)\big|\mathcal{F}^n_s\Big] \\
&\le C\big((t + q^n([0,t])) - (s + q^n([0,s]))\big).
\end{aligned}
$$

By virtue of assumption (iii), we conclude that the assumption A2') from [3, Corollary 1.2] holds. Next, we explain that the family $\{f(X^n) \colon n \in \mathbb{N}\}$ also satisfies the compact containment condition given by A1) in [3]. Take $\eta, T > 0$ and let $K = K(\eta, T) \subset E$ be the compact set as in assumption (i). Then, by the continuity of f, the set $f(K) \subset \mathbb{R}$ is compact. Furthermore, by (i), we obtain

$$
\inf_{n\in\mathbb{N}} P\big(f(X^n_t) \in f(K) \text{ for all } 0 \le t \le T\big) \ge \inf_{n\in\mathbb{N}} P\big(X^n_t \in K \text{ for all } 0 \le t \le T\big) \ge 1 - \eta.
$$

Hence, A1) from [3] holds. Now, [3, Corollary 1.2] yields that $\{f(X^n) \colon n \in \mathbb{N}\}$ is tight.

Step 2: Conclusion. Due to the fact that we assume the compact containment condition (i.e. (i)) and the properties of H, [12, Theorem 3.9.1] yields that tightness of the family $\{X^n \colon n \in \mathbb{N}\}$ is equivalent to tightness of the families $\{f(X^n) \colon n \in \mathbb{N}\}$ for every $f \in H$. As latter is the case thanks to Step 1, the claim of the theorem follows. □

5.4 A short example to keep in mind: revisited

Let us relate Theorem 5.4 to the example from Section 5.2. More precisely, we explain how $\mathfrak{u}$ and $L^1, L^2, \dots$ from Section 5.3 can be chosen.

Suppose that $X^1, X^2, \dots$ are (one-dimensional) processes whose dynamics are given by

$$
dX^n_t = \sigma^n(X^n_{t-})dS^n_t - \int_0^1 (1-\theta)X^n_{t-} q^n(dt, d\theta),
$$

where $S^1, S^2, \dots$ is a sequence of semimartingales with equal laws and $q^1, q^2, \dots$ is a sequence of deterministic measures on $\mathbb{R}_+ \times [0,1]$ such that $q^n(dt, d\theta) = \sum_{i=1}^{m_n} \delta_{t_i}(dt) F_i(d\theta)$ with $F_i([0,1]) = 1$ for $i = 1, \dots, m_n$. Let $\mathfrak{g} \colon \mathbb{R}_+ \to \mathbb{R}_+$ be continuous and increasing, and assume that

$$
|\sigma^n(x)| \le \mathfrak{g}(|x|), \quad n = 1, 2, \dots, \ x \in \mathbb{R}.
$$

As explained in the discussion below Assumption 5.3, the purpose of the sequence $L^1, L^2, \dots$ is to control the jump times of $X^1, X^2, \dots$. Notice that a.s.

$$
\Delta X^n_t = \sigma^n(X^n_{t-})\Delta S^n_t - X^n_{t-} \int_0^1 (1-\theta) q^n(\{t\} \times d\theta), \quad t > 0.
$$

Motivated by this computation, we take

$$
U \triangleq \mathbb{D}(\mathbb{R}),\ L^n \triangleq S^n,\ x_0 \triangleq 0,\ \kappa(x) \triangleq \mathfrak{g}(x) + x, x \in \mathbb{R}_+,
$$

and

$$\mathfrak{u}(u,t) \triangleq |\Delta u(t)| + \sup_{n\in\mathbb{N}} \int_0^1 (1-\theta) q^n(\{t\}\times d\theta), \quad (u,t)\in\mathbb{D}(\mathbb{R})\times\mathbb{R}_+.$$

With these definitions at hand, it is clear that, for all $n = 1, 2, \dots$, $(L^n, X^n) \in A$ up to a null set, i.e. (5.5) holds. Furthermore, as the semimartingales $S^1, S^2, \dots$ are assumed to have the same law, the first part of Assumption 5.1 can be treated as in Remark 5.2. Finally, we discuss the property (5.4). By virtue of Lemma 2.10, a sufficient condition for $\mathfrak{u}$ to satisfy (5.4) is that the set $\{t \in [0,T] \colon \mathfrak{u}(u,t) \geq a\}$ is finite for arbitrary $T, a > 0$ and $u \in \mathbb{D}(\mathbb{R})$. By standard properties of càdlàg functions, $\{t\in[0,T]\colon |\Delta u(t)| \geq a/2\}$ is finite for such T, a and u and hence, the same is true for $\{t \in [0,T]\colon \mathfrak{u}(u,t)\geq a\}$ once the set

$$\Big\{t \in [0,T]\colon \sup_{n\in\mathbb{N}} \int_0^1 (1-\theta) q^n(\{t\}\times d\theta) \geq a/2\Big\}$$

is finite. Latter holds for instance when the sequence

$$t \mapsto \int_0^1 (1-\theta) q^n([0,t]\times d\theta), \quad n = 1, 2, \dots,$$

converges in the local uniform topology, see Example 5.12 below for more details.

5.5 Theorems for semimartingales

In this section we derive stability results for semimartingales which are tailored to the presence of fixed times of discontinuity. We start with a general result in Section 5.5.1, which we specify further for the annealed case in Section 5.5.3, i.e. the case where all processes are defined on the same measurable space.

5.5.1 The main result

We pose ourselves into the setting of Section 4.3. To be precise, let (B, C, ν) be a candidate triplet for semimartingale characteristics (corresponding to a fixed continuous truncation function $h\colon \mathbb{R}^d \to \mathbb{R}^d$) defined on the canonical space $\mathbb{D}(\mathbb{R}^d)$. Except stated otherwise, we endow $\mathbb{D}(\mathbb{R}^d)$ with the Skorokhod J_1 topology. We write $C_1(\mathbb{R}^d)$ for a subset of the set of non-negative bounded continuous functions vanishing in a neighborhood of the origin as described in [25, VII.2.7].

Let $(U, \mathcal{U})$ be a measurable space and fix a $\mathcal{U}\otimes\mathcal{B}(\mathbb{R}_+)/\mathcal{B}(\mathbb{R}_+)$-measurable function $\mathfrak{u}\colon U\times\mathbb{R}_+\to\mathbb{R}_+$ such that for every $(u,t)\in U\times(0,\infty)$

$$\lim_{\varepsilon\searrow 0} \sup\big\{\mathfrak{u}(u,s)\colon s\neq t, t-\varepsilon\leq s\leq t+\varepsilon\big\} = 0,$$

let $\kappa\colon\mathbb{R}_+\to\mathbb{R}_+$ be increasing and continuous, and define

$$A \triangleq \Big\{(u,\omega)\in U\times\mathbb{D}(\mathbb{R}^d)\colon \|\Delta\omega(t)\| \leq \mathfrak{u}(u,t)\kappa\Big(\sup_{s\leq t}\|\omega(s)\|\Big) \text{ for all } t>0\Big\}.$$

The following is the main result of this section.

Theorem 5.8. *Let $\mathbb{B} = (\Omega, \mathcal{F}, \mathbf{F}, P)$ and $\mathbb{B}^n = (\Omega^n, \mathcal{F}^n, \mathbf{F}^n, P^n)$ be filtered probability spaces which support $\mathbb{R}^d$-valued càdlàg adapted processes X and X^n, respectively, such that each X^n is a semimartingale with semimartingale characteristics (B^n, C^n, ν^n) corresponding to the (continuous) truncation function h. Moreover, for each $n\in\mathbb{N}$, let L^n be a U-valued random variable on $\mathbb{B}^n$ such that $\{P^n\circ(L^n)^{-1}\colon n\in\mathbb{N}\}$ is relatively (sequentially)[5] compact in $M_m(U)$. Assume that $X^n\to X$ weakly on $\mathbb{D}(\mathbb{R}^d)$ and the existence of a dense set $\Gamma\subset\mathbb{R}_+$ such that the following hold:*

[5] By [16, Theorem 2.6], relative compactness and sequential relative compactness are equivalent in $M_m(U)$.

(i)′ *for every* $t \in \Gamma, \varepsilon > 0$ *and* $g \in C_1(\mathbb{R}^d)$*, we have*

$$\begin{aligned} P^n(\|B^n_t - B_t(X^n)\| \geq \varepsilon) \to 0, \\ P^n(\|\widetilde{C}^n_t - \widetilde{C}_t(X^n)\| \geq \varepsilon) \to 0, \\ P^n(|g * \nu^n_t - g * \nu_t(X^n)| \geq \varepsilon) \to 0, \end{aligned}$$

as $n \to \infty$*;*

(ii)′ *for all* $T \in \Gamma$ *and* $g \in C_1(\mathbb{R}^d)$*, there is a sequence* $S_1, S_2, \dots$ *of stopping times on* $\mathbb{B}^1, \mathbb{B}^2, \dots$*, i.e.* S_n *is an* $\mathbf{F}^n$*-stopping time, such that*

$$P^n(S_n < T) \to 0, \quad n \to \infty,$$

and

$$\sup_{n \in \mathbb{N}} E^{P^n}\Big[\|\widetilde{C}^n_{T \wedge S_n}\|^2 + g^2 * \nu^n_{T \wedge S_n}\Big] < \infty;$$

(iii)′ *for all* $t \in \Gamma$ *and* $g \in C_1(\mathbb{R}^d)$*, the maps*

$$\mathbb{D}(\mathbb{R}^d) \ni \omega \mapsto B_t(\omega), \widetilde{C}_t(\omega), (g * \nu_t)(\omega)$$

are continuous in the local uniform topology, and

$$P^n((L^n, X^n) \in A) \to 1 \text{ as } n \to \infty. \tag{5.11}$$

Then, X *is a semimartingale for its canonical filtration and its semimartingale characteristics are given by* $(B(X), C(X), \nu(X))$*.*

Proof. Let Y° to be any of the processes in (i) – (iii) from Example 3.9. We show that Y° is a P-martingale for the (right-continuous) canonical filtration on $\mathbb{D}(\mathbb{R}^d)$. For simplicity, we restrict our attention to the process in (ii) of Example 3.9. More precisely, let Y° be defined by

$$Y^\circ \triangleq V^{(i)} V^{(j)} - \widetilde{C}^{(ij)},$$

where $V = (V^{(1)}, \dots, V^{(d)})$ is given by

$$V \triangleq \mathsf{X}(h) - \mathsf{X}_0 - B,$$

with

$$\mathsf{X}(h) = \mathsf{X} - \sum_{s \leq \cdot} \big(\Delta \mathsf{X}_s - h(\Delta \mathsf{X}_s)\big).$$

Our strategy is to apply Theorem 3.20. We define probability measures $Q_1, Q_2, \dots$ on the product space $(U \times \mathbb{D}(\mathbb{R}^d), \mathcal{U} \otimes \mathcal{B}(\mathbb{D}(\mathbb{R}^d)))$ via

$$Q_n \triangleq P^n \circ (L^n, X^n)^{-1}, \quad n \in \mathbb{N}.$$

As we assume that $X^n \to X$ weakly and that the distributions of $L^1, L^2, \dots$ are relatively compact in $M_m(U)$, Theorem 2.5 yields the existence of a subsequence of $(Q_n)_{n \in \mathbb{N}}$ which converges in $M_{mc}(U \times \mathbb{D}(\mathbb{R}^d))$ to some probability measure Q. To keep our notation simple, we denote the subsequence again by $(Q_n)_{n \in \mathbb{N}}$. Clearly, we have $Q_{\mathbb{D}(\mathbb{R}^d)} = P \circ X^{-1}$. In the following we show that Y°_t is (Q_n, Q)-continuous for every $t \in \Gamma$.

Thanks to Proposition 2.12, for every $u \in U$ the set $A_u = \{\omega \in \mathbb{D}(\mathbb{R}^d) \colon (u, \omega) \in A\}$ is closed in the Skorokhod J_1 topology and on A_u the Skorokhod J_1 topology coincides with the local uniform topology. In (iii)′ we assume that $Q_n(A) \to 1$ as $n \to \infty$. Hence, we deduce from Proposition 2.8 that $Q(A) = 1$. The first part of assumption (iii)′ yields that $B_t|_{A_u}$ and $\widetilde{C}_t|_{A_u}$ are continuous in the Skorokhod J_1 topology for every $t \in \Gamma$.

Lemma 5.9. *Let $g\colon \mathbb{R}^d \to \mathbb{R}$ be a continuous function which vanishes in a neighborhood of the origin. For every $t > 0$, the map $\omega \mapsto \sum_{s\leq t} g(\Delta\omega(s))$ is continuous in the local uniform topology.*

Proof. For $(\omega, u) \in \mathbb{D}(\mathbb{R}^d) \times (0, \infty)$, we set

$$t^0(\omega, u) \triangleq 0, \quad t^{p+1}(\omega, u) \triangleq \inf(t > t^p(\omega, u)\colon \|\Delta\omega(t)\| > u), \quad p \in \mathbb{Z}_+.$$

Furthermore, we set

$$U(\omega) \triangleq \big\{u > 0\colon \exists t > 0 \text{ such that } \|\Delta\omega(t)\| = u\big\}, \quad \omega \in \mathbb{D}(\mathbb{R}^d).$$

Now, suppose that $\omega_n \to \omega$ in the local uniform topology and take some $t > 0$. As $U(\omega)$ is at most countable, there is a $0 < u \notin U(\omega)$ such that $g(x) = 0$ for $\|x\| \leq u$. Let $p' \triangleq \max\{p \in \mathbb{Z}_+\colon t^p(\omega, u) \leq t\}$. Then, thanks to [36, Theorem 2.6.2], there exists an $N \in \mathbb{N}$ such that

$$\sum_{s\leq t} g(\Delta\omega_{n+N}(s)) = \sum_{k=1}^{p'} g(\Delta\omega_{n+N}(t^k(\omega, u))), \quad n \in \mathbb{N}.$$

As $n \to \infty$ the r.h.s. converges to

$$\sum_{k=1}^{p'} g(\Delta\omega(t^k(\omega, u))) = \sum_{s\leq t} g(\Delta\omega(s)).$$

This completes the proof. $\square$

By virtue of this lemma, for every $t \in \Gamma$, we conclude that the set

$$\{(u, \omega) \in A\colon A_u \ni \xi \mapsto Y^\circ_t(\xi) \text{ is discontinuous at } \omega\}$$

is Q-null and consequently, that Y°_t is (Q_n, Q)-continuous.

Let $\mathcal{Z}^\circ$ be the determining set from part (i) of Example 3.6 with $D = \Gamma$. Then, it is clear that for every $Z^\circ_s \in \mathcal{Z}^\circ_s$ with $s \leq t$ the random variable $Y^\circ_t Z^\circ_s$ is also (Q_n, Q)-continuous. It remains to verify the final two parts of (A4) from Theorem 3.20. We fix $T \in \Gamma$. Let $S_1, S_2, \dots$ be as in (ii)' and set

$$Y^n \triangleq \big(X^n(h)_{\cdot\wedge T\wedge S_n} - X^n_0 - B^n_{\cdot\wedge T\wedge S_n}\big)^{(i)}\big(X^n(h)_{\cdot\wedge T\wedge S_n} - X^n_0 - B^n_{\cdot\wedge T\wedge S_n}\big)^{(j)} - \widetilde{C}^{n,(ij)}_{\cdot\wedge T\wedge S_n},$$

which is a local martingale on $\mathbb{B}^n$. First of all, as $|\Delta(X^n(h) - X_0 - B^n)^{(i)}| \leq 2\|h\|_\infty$, we deduce from [25, Lemma VII.3.34] that

$$E^{P^n}\Big[\sup_{s\leq T\wedge S_n} \big|\big(X^n(h)_{s\wedge S_n} - X^n_0 - B^n_{s\wedge S_n}\big)^{(i)}\big|^4\Big] \lesssim E^{P^n}\Big[\big|\widetilde{C}^{n,(ii)}_{T\wedge S_n}\big|^2\Big]^{\frac{1}{2}} + E^{P^n}\Big[\big|\widetilde{C}^{n,(ii)}_{T\wedge S_n}\big|^2\Big].$$

Consequently, hypothesis (ii)' yields that

$$\sup_{n\in\mathbb{N}} E^{P^n}\Big[\sup_{s\leq T} |Y^n_s|^2\Big] < \infty.$$

Hence, Y^n is a true martingale on $\mathbb{B}^n$ and the set $\{Y^n_s\colon s \in [0, T], n \in \mathbb{N}\}$ is uniformly integrable. It remains to verify (3.8). Notice that on $[\![0, T \wedge S_n]\!]$

$$\begin{aligned} Y^n - Y^\circ(X^n) &= \big(X^n(h) - X^n_0 - B^n\big)^{(i)}\big(B(X^n) - B^n\big)^{(j)} \\ &\quad + \big(X^n(h) - X^n_0 - B(X^n)\big)^{(j)}\big(B(X^n) - B^n\big)^{(i)} - \widetilde{C}^{n,(ij)} + \widetilde{C}^{(ij)}(X^n). \end{aligned}$$

Let us recall an elementary fact ([27, Exercise 3.5, p. 58]). If $\xi_1, \xi_2, \dots$ and $\eta_1, \eta_2, \dots$ are sequences of random variables such that $(\xi_n)_{n\in\mathbb{N}}$ is uniformly integrable and $\eta_n \to 0$ in probability, then $\xi_n\eta_n \to 0$ in probability. Using this fact and assumptions (i)' and (ii)', for all $t \in \Gamma \cap [0, T]$ and $\varepsilon > 0$, we obtain

$$P^n\big(|Y^n_t - Y^\circ_t(X^n)| \geq \varepsilon\big) \leq P^n\big(|Y^n_t - Y^\circ_t(X^n)| \geq \varepsilon, t \leq S_n\big) + P^n\big(T > S_n\big) \to 0$$

as $n \to \infty$. As $T \in \Gamma$ was arbitrary and $\Gamma \subset \mathbb{R}_+$ is dense, we conclude that (A4) holds and consequently, the claim follows. □

Remark 5.10. The literature contains several conditions for tightness of processes with fixed times of discontinuity. Conditions for semimartingales are given in [25, Theorems VI.5.10, IX.3.20]. We also refer to the recent article [3] where a version of Aldous's tightness criterion for processes with fixed times of discontinuity is proved.

Remark 5.11. Hypothesis (ii)' holds for instance under the following uniform boundedness assumption: for all $T > 0$ and $g \in C_1(\mathbb{R}^d)$, we have

$$\sup_{\omega\in\mathbb{D}(\mathbb{R}^d)} \big(\|\widetilde{C}_T(\omega)\| + |(g * \nu_T)(\omega)|\big) < \infty.$$

This follows from arguments used in the proof of Theorem 4.7. In practice (ii)' seems to be more flexible than this boundedness condition. For instance, the first part of the assumption also holds in case

$$\|\widetilde{C}^n\| \lesssim 1 + \sup_{s\leq\cdot} \|X^n_s\|^2$$

and

$$\sup_{n\in\mathbb{N}} E^{P^n}\Big[\sup_{s\leq T} \|X^n_s\|^4\Big] < \infty, \quad T > 0.$$

Under suitable linear growth assumptions on the characteristics (B^n, C^n, ν^n), and a suitable integrability condition on the initial distributions, the fourth moment condition can be verified by Gronwall's lemma.

5.5.2 Examples for $\mathfrak{u}$ and $L^1, L^2, \dots$

In this section we explain how the function $\mathfrak{u}$ and the control variables $L^1, L^2, \dots$ can be chosen such that (5.11) holds when $X^1, X^2, \dots$ are stochastic integrals. Hereby, we use ideas from [22, 24].

Example 5.12. As we only want to fix ideas, suppose that all semimartingales $X^1, X^2, \dots$ are one-dimensional, defined on the same stochastic basis $\mathbb{B} = (\Omega, \mathcal{F}, \mathbf{F}, P)$ and are stochastic integrals of the form

$$dX^n_t = \sigma^n_t dZ^n_t,$$

for a one-dimensional semimartingale Z^n and a predictable process $\sigma^n \in L(Z^n)$. In this case we have $\Delta X^n = \sigma^n \Delta Z^n$. Assume that there exists a non-negative predictable process γ such that $\gamma \in L(Z^n)$ and $|\sigma^n| \leq \gamma(1 + \sup_{s\leq\cdot} |X^n_s|)$ for all $n \in \mathbb{N}$. The linear growth condition can be relaxed. Then,

$$|\Delta X^n| \leq \gamma|\Delta Z^n|\Big(1 + \sup_{s\leq\cdot} |X^n_s|\Big) = |\Delta L^n|\Big(1 + \sup_{s\leq\cdot} |X^n_s|\Big), \qquad L^n \triangleq \int_0^\cdot \gamma_s dZ^n_s.$$

Now, we set $U \triangleq \mathbb{D}(\mathbb{R})$ and $\mathfrak{u}(u, t) \triangleq |\Delta u(t)|$ for $(u, t) \in \mathbb{D}(\mathbb{R}) \times \mathbb{R}_+$. By standard properties of càdlàg functions, the set $\{t \in [0, T] \colon \mathfrak{u}(u, t) \geq a\} = \{t \in [0, T] \colon |\Delta u(t)| \geq a\}$ is finite for every $(a, T, u) \in (0, \infty) \times (0, \infty) \times \mathbb{D}(\mathbb{R})$. Consequently, by Lemma 2.10, $\mathfrak{u}$ has the desired properties and (5.11) is satisfied with $\kappa(x) \triangleq 1 + x$ for $x \in \mathbb{R}_+$. Often enough, it holds that $Z^n \equiv Z$, which implies that the law of L^n is independent of n.

Alternatively, suppose that there exists a càdlàg measurable process Z such that $Z^n \to Z$ in the ucp[6] topology, i.e., for all $t \in \mathbb{R}_+$,

$$\sup_{s \le t} |Z^n_s - Z_s| \to 0$$

in probability as $n \to \infty$. For what follows, fix $T > 0$. Up to passing to a subsequence, which we ignore for simplicity, the set

$$\Omega^o \triangleq \Big\{ \omega \in \Omega \colon \lim_{n \to \infty} \sup_{s \le T} |Z^n_s(\omega) - Z_s(\omega)| = 0 \Big\}$$

is full. Now, when we define

$$\mathfrak{u}(\omega, t) \triangleq \begin{cases} \sup_{n \in \mathbb{N}} |\Delta Z^n_t(\omega)|, & \omega \in \Omega^o, \\ 0, & \text{otherwise}, \end{cases}$$

for $(\omega, t) \in \Omega \times [0, T]$, the set $\{t \in [0, T] \colon \mathfrak{u}(\omega, t) \ge a\}$ is finite for every $\omega \in \Omega$ and $a > 0$.

To see this, take $\omega \in \Omega^o$ and let $N = N(\omega) \in \mathbb{N}$ be such that

$$\sup_{n \ge N} \sup_{s \le T} |Z^n_s(\omega) - Z_s(\omega)| \le \frac{a}{3}.$$

Then, for every $t \in [0, T]$,

$$\begin{aligned} \sup_{n \ge N} |\Delta Z^n_t(\omega)| \ge a \implies |\Delta Z_t(\omega)| &\ge \sup_{n \ge N} |\Delta Z^n_t(\omega)| - \sup_{n \ge N} |\Delta Z^n_t(\omega) - \Delta Z_t(\omega)| \\ &\ge a - \frac{2a}{3} = \frac{a}{3}. \end{aligned}$$

As there are only finitely many $t \in [0, T]$ such that $|\Delta Z_t(\omega)| \ge a/3$, there are also only finitely many $t \in [0, T]$ such that $\sup_{n \ge N} |\Delta Z^n_t(\omega)| \ge a$. Now, since

$$\Big\{ t \colon \sup_{n \in \mathbb{N}} |\Delta Z^n_t(\omega)| \ge a \Big\} \subset \Big(\bigcup_{k=1}^{N-1} \Big\{ t \colon |\Delta Z^k_t(\omega)| \ge a \Big\} \Big) \cup \Big\{ t \colon \sup_{n \ge N} |\Delta Z^n_t(\omega)| \ge a \Big\},$$

we conclude that there are at most finitely many $t \in [0, T]$ such that $\sup_{n \in \mathbb{N}} |\Delta Z^n_t(\omega)| \ge a$, which was the claim.

Up to a pasting argument, if γ is, for instance, constant, we can take $U \triangleq \Omega$ and $L^n \equiv \mathrm{Id}$ such that (5.11) holds. In particular, as we assume that all processes are defined on the same stochastic basis, the distributions of $L^1, L^2, \dots$ are trivially (relatively) compact in $M_m(U)$. At the cost of slightly more complicated conditions, this argument can be transferred to the more general case where $\mathbb{B}^n = (\Omega, \mathcal{F}, \mathbf{F}^n, P^n)$. More details on this strategy are given in the proof of Corollary 5.19 below.

Example 5.13. In this example we explain how (iii)' can be checked in case $X^1, X^2, \dots$ are stochastic integrals w.r.t. a (compensated) random measure. As in Example 5.12, for simplicity assume that all $X^1, X^2, \dots$ are one-dimensional and defined on the same stochastic basis. Moreover, we assume that

$$X^n = X^n_0 + \int_0^{\cdot} \int H^n(s, y)(\mathfrak{p}^n - \mathfrak{q}^n)(ds, dy),$$

[6] ucp = uniformly on compacts in probability.

where $\mathfrak{p}^n - \mathfrak{q}^n$ is a compensated integer-valued random measure on a Blackwell space $(E, \mathcal{E})$ and $H^n \in G_{\text{loc}}(\mathfrak{p}^n)$. Suppose that γ is a non-negative predictable process such that a.s. $\gamma * \mathfrak{q}^n < \infty$ and $|H^n| \leq \gamma(1 + \sup_{s \leq \cdot} |X^n_s|)$ for all $n \in \mathbb{N}$. The linear growth condition can be relaxed. We now set

$$L^n \triangleq \int_0^{\cdot} \int \gamma(s, y)(\mathfrak{p}^n + \mathfrak{q}^n)(ds, dy),$$

and we obtain that

$$\begin{aligned} |\Delta X^n_t| &= \Big| \int H^n(t, y)\mathfrak{p}^n(\{t\} \times dy) - \int H^n(t, y)\mathfrak{q}^n(\{t\} \times dy) \Big| \\ &\leq \int |H^n(t, y)|\mathfrak{p}^n(\{t\} \times dy) + \int |H^n(t, y)|\mathfrak{q}^n(\{t\} \times dy) \\ &\leq \int \gamma(t, y)(\mathfrak{p}^n + \mathfrak{q}^n)(\{t\} \times dy)\Big(1 + \sup_{s \leq t} |X^n_s|\Big) \\ &= |\Delta L^n_t|\Big(1 + \sup_{s \leq t} |X^n_s|\Big) \end{aligned}$$

for all $t > 0$. Now, we can define $U = \mathbb{D}(\mathbb{R})$, $\mathfrak{u}(u, t) = |\Delta u(t)|$ and $\kappa(x) = 1 + x$ such that (5.11) holds. Often enough the law of L^n is independent of n. The strategy outlined in the second part of Example 5.12 can also be transferred to this setting, see the proof of Corollary 5.19 below.

5.5.3 The annealed setting

In this section we assume that $X^1, X^2, \dots$ are defined on the same filtered probability space, which can be viewed as an annealed setting, see Section 5.1. In this case we allow the limiting characteristics to be random.

We fix a filtered probability space $\mathbb{B} = (\Omega, \mathcal{F}, \mathbf{F}, P)$ which supports $\mathbb{R}^d$-valued cádlág adapted processes $X^1, X^2, \dots$. Moreover, we define an *extension* $(\Omega', \mathcal{F}', \mathbf{F}')$ of the filtered space $(\Omega, \mathcal{F}, \mathbf{F})$ by

$$\Omega' \triangleq \Omega \times \mathbb{D}(\mathbb{R}^d), \quad \mathcal{F}' \triangleq \mathcal{F} \otimes \mathcal{D}(\mathbb{R}^d), \quad \mathcal{F}'_t \triangleq \bigcap_{s > t} \big(\mathcal{F}_s \otimes \mathcal{D}_s(\mathbb{R}^d)\big),$$

where $\mathcal{D}(\mathbb{R}^d)$ and $(\mathcal{D}_t(\mathbb{R}^d))_{t \geq 0}$ are the canonical σ-field and the canonical (right-continuous) filtration on $\mathbb{D}(\mathbb{R}^d)$. With little abuse of terminology, we denote the *canonical process* on Ω' by $\mathsf{X}(\omega, \alpha) = \alpha$ for $(\omega, \alpha) \in \Omega'$.

Let (B, C, ν) be a candidate triplet on $(\Omega', \mathcal{F}', \mathbf{F}')$ relative to a fixed continuous truncation function $h\colon \mathbb{R}^d \to \mathbb{R}^d$, cf. [25, III.2.3]. Let $\mathfrak{u}\colon \Omega \times \mathbb{R}_+ \to \mathbb{R}_+$ be an $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+)/\mathcal{B}(\mathbb{R}_+)$-measurable function such that, for every $(\omega, t) \in \Omega \times (0, \infty)$,

$$\lim_{\varepsilon \searrow 0} \sup \big\{\mathfrak{u}(\omega, s)\colon s \neq t, t - \varepsilon \leq s \leq t + \varepsilon\big\} = 0,$$

let $\kappa\colon \mathbb{R}_+ \to \mathbb{R}_+$ be increasing and continuous, and define

$$A^n \triangleq \Big\{\omega \in \Omega\colon \|\Delta X^n_t(\omega)\| \leq \mathfrak{u}(\omega, t)\kappa\Big(\sup_{s \leq t} \|X^n_s(\omega)\|\Big) \text{ for all } t > 0\Big\}.$$

The set $\bigcap_{n \in \mathbb{N}} A^n$ can be interpreted as follows: the jumps of the processes $X^1, X^2, \dots$ are controlled by a process $\mathfrak{u}$ which roughly behaves like the jump process $|\Delta Z|$ of some one-dimensional càdlàg process Z.

The following is the main result of this section.

Theorem 5.14. *Suppose that each X^n is a semimartingale with semimartingale characteristics $(B^n, C^n,\ \nu^n)$ corresponding to the (continuous) truncation function h. Assume that there exists a probability measure $P \circ X^{-1}$ on $(\mathbb{D}(\mathbb{R}^d), \mathcal{D}(\mathbb{R}^d))$ such that $P \circ (X^n)^{-1} \to P \circ X^{-1}$ weakly (where $\mathbb{D}(\mathbb{R}^d)$ is endowed with the Skorokhod J_1 topology) and that there exists a dense set $\Gamma \subset \mathbb{R}_+$ such that the following hold:*

(i) *for every $t \in \Gamma, \varepsilon > 0$ and $g \in C_1(\mathbb{R}^d)$, we have*

$$\begin{aligned} P(\|B^n_t - B_t(X^n)\| \geq \varepsilon) &\to 0, \\ P(\|\widetilde{C}^n_t - \widetilde{C}_t(X^n)\| \geq \varepsilon) &\to 0, \\ P(|g * \nu^n_t - g * \nu_t(X^n)| \geq \varepsilon) &\to 0, \end{aligned}$$

as $n \to \infty$;

(ii) *for all $T \in \Gamma$ and $g \in C_1(\mathbb{R}^d)$, there is a sequence of stopping times $(S_n)_{n \in \mathbb{N}}$ such that*

$$P(S_n < T) \to 0, \quad n \to \infty,$$

and

$$\sup_{n \in \mathbb{N}} E^P \Big[\|\widetilde{C}^n_{T \wedge S_n}\|^2 + g^2 * \nu^n_{T \wedge S_n} \Big] < \infty;$$

(iii) *for all $\omega \in \Omega, t \in \Gamma$ and $f \in C_1(\mathbb{R}^d)$, the maps*

$$\mathbb{D}(\mathbb{R}^d) \ni \alpha \mapsto B_t(\omega, \alpha), \widetilde{C}_t(\omega, \alpha), (f * \nu_t)(\omega, \alpha)$$

are continuous in the local uniform topology. Moreover,

$$P(A^n) \to 1 \text{ as } n \to \infty.$$

Then, there exists a probability measure Q on $(\Omega', \mathcal{F}')$, which is a weak-strong accumulation point of $\{P \circ (\mathrm{Id}, X^n)^{-1} \colon n \in \mathbb{N}\}$ with $Q_\Omega = P$ and $Q_{\mathbb{D}(\mathbb{R}^d)} = P \circ X^{-1}$, such that on $(\Omega', \mathcal{F}', \mathbf{F}', Q)$ the canonical process X is a semimartingale with characteristics (B, C, ν).

Proof. The proof is similar to those of Theorem 5.8 where we take $(U, \mathcal{U}) = (\Omega, \mathcal{F})$ and $L^n(\omega) \equiv L(\omega) = \omega$ for all $\omega \in \Omega$. Let us emphasis that in this case we can use a mixture of the determining sets described in part (i) and (iii) of Example 3.6. The details are left to the reader. □

In the context of SDEs with semimartingale drivers, Theorem 5.14 is related to [24, Theorem 3.16].

Example 5.15. We provide a short example for an application of Theorem 5.14 in a setting without jumps. A more detailed exposition of a closely related setting with fixed times of discontinuity is given in Section 5.5.4 below. Furthermore, a related discussion for a setting with jumps is given on p. 205 in [24]. We take $\mathbb{B}$ as underlying filtered space. Let $\tau^0, \tau^1, \tau^2, \dots$ be stopping times on $\mathbb{B}$, which we think to be change points of economic scenarios, and let W be a one-dimensional standard Brownian motion on the stochastic basis $\mathbb{B}$. Moreover, take $b, b^\circ \colon \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ and $\sigma, \sigma^\circ \colon \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ to be sufficiently regular functions such that, for each $n \in \mathbb{Z}_+$, the SDE

$$\begin{aligned} dX^n_t = \big(b(t, X^n_t) \mathbb{1}_{\{t \leq \tau^n\}} + b^\circ(t, X^n_t) \mathbb{1}_{\{t > \tau^n\}} \big) dt \\ + \big(\sigma(t, X^n_t) \mathbb{1}_{\{t \leq \tau^n\}} + \sigma^\circ(t, X^n_t) \mathbb{1}_{\{t > \tau^n\}} \big) dW_t, \quad X^n_0 = x_0, \end{aligned}$$

has a solution process X^n. It is well-known that (local) Lipschitz (or monotonicity) and linear growth conditions on b, b° and σ, σ° imply existence (and uniqueness in a strong sense), see, e.g. [21, Chapter 14] or [24, Section 4]. We also stress that the above SDEs have random coefficients, where the randomness enters in terms of the sequence $\tau^0, \tau^1, \tau^2, \dots$. Part (i) of Theorem 5.14 holds if $\tau^n \to \tau^0$ in probability. Under this condition and under suitable assumptions on the coefficients (see [24, Section 4]), if the laws of $X^1, X^2, \dots$ converge weakly, then[7] the measure Q in Theorem 5.14 is given by

$$Q(d\omega, d\alpha) = \delta_{X^0(\omega)}(d\alpha)P(d\omega).$$

Since $Q_{\mathbb{D}(\mathbb{R})} = P \circ (X^0)^{-1}$, this shows that the laws of $X^1, X^2, \dots$ converge weakly to the law of X^0. In fact, we can say more: by Remark 2.2, we can even conclude that $X^n \to X^0$ in the ucp topology. This observation can be compared to classical ucp stability results for semimartingale SDEs as for instance given in [33]. We stress that the above argument does not rely on the Lipschitz continuity of the coefficients as the argument in [33], but on strong existence and uniqueness. Therefore, we think it is more flexible when it comes to the regularity of the coefficients. However, in the presence of jumps the argument only yields convergence in probability for the Skorokhod J_1 topology, which is weaker than ucp convergence.

5.5.4 Application: Itô processes with fixed times of discontinuity

In this section we specify Theorem 5.8 for solutions to SDEs driven by a Gaussian continuous local martingale and a Poisson random measure.

Let $(E, \mathcal{E})$ be a Blackwell space and let $\mathcal{P}$ be the predictable σ-field on $\mathbb{R}_+ \times \mathbb{D}(\mathbb{R}^d)$ when $\mathbb{D}(\mathbb{R}^d)$ is equipped with the canonical filtration. Let $\sigma \colon \mathbb{R}_+ \times \mathbb{D}(\mathbb{R}^d) \to \mathbb{R}^{d\times r}$ be $\mathcal{P}/\mathcal{B}(\mathbb{R}^{d\times r})$-measurable and let $v, b \colon \mathbb{R}_+ \times \mathbb{D}(\mathbb{R}_+) \times E \to \mathbb{R}^d$ be $\mathcal{P}\otimes\mathcal{E}/\mathcal{B}(\mathbb{R}^d)$-measurable. Let C^n and C be covariance functions for r-dimensional continuous Gaussian martingales such that

$$C^n = \int_0^\cdot c^n_s ds, \qquad C = \int_0^\cdot c_s ds.$$

Moreover, let $\mathfrak{q}$ and $\mathfrak{q}^n$ be intensity measures of Poisson random measures on $(E, \mathcal{E})$, let q^n and q be σ-finite measures on $(\mathbb{R}_+ \times E, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E})$ and let $h \colon \mathbb{R}^d \to \mathbb{R}^d$ be a continuous truncation function. For each $n \in \mathbb{N}$, we fix a stochastic basis $\mathbb{B}^n \triangleq (\Omega, \mathcal{F}, \mathbf{F}^n, P)$ which supports the following: a continuous Gaussian martingale W^n with covariance function C^n, a Poisson random measure $\mathfrak{p}^n$ with intensity measure $\mathfrak{q}^n$ and a semimartingale X^n with dynamics

$$\begin{aligned} X^n = X^n_0 + \int_0^\cdot \int b^n(t,y) q^n(dt,dy) + \int_0^\cdot \sigma^n_t dW^n_t + \int_0^\cdot \int h(v^n(t,y))(\mathfrak{p}^n - \mathfrak{q}^n)(dt,dy) \\ + \int_0^\cdot \int (v^n(t,y) - h(v^n(t,y)))\mathfrak{p}^n(dt,dy), \end{aligned}$$

where b^n, σ^n and v^n are suitable processes such that the integrals are well-defined. We now also formulate some technical conditions.

Assumption 5.16. There exists a càdlàg process X on a probability space $(\Omega^*, \mathcal{F}^*, P^*)$ such that $X^n \to X$ weakly on $\mathbb{D}(\mathbb{R}^d)$ endowed with the Skorokhod J_1 topology. Let $\mathbf{F}^X$ be the canonical (right-continuous) filtration generated by X and set $\mathbb{B} \triangleq (\Omega^*, \mathcal{F}^*, \mathbf{F}^X, P^*)$.

[7] More precisely, under (local) Lipschitz or monotonicity conditions as given in [24, Section 4], by [24, Corollary 2.26] there exists a unique solution measure (in the sense of [24, Definition 1.6]) to the SDE of X^0 and it is strong (see [24, Definition 2.21]). Thus, by [24, Theorem 2.22], the solution measure has the form $\delta_{X^0(\omega)}(d\alpha)P(d\omega)$. As Q in Theorem 5.14 is a solution measure to the SDE of X^0 by [24, Theorem 2.10], the claim follows.

Assumption 5.17. Part (i)' and (ii)' of Theorem 5.8 hold for a dense set $\Gamma \subset \mathbb{R}_+$ and the following characteristics (B^n, C^n, ν^n) and (B, C, ν):

$$B^n \triangleq \int_0^\cdot \int b^n(t,y) q^n(dt,dy),$$
$$C^n \triangleq \int_0^\cdot \sigma_t^n c_t^n \sigma_t^{n*} dt,$$
$$\nu^n([0,t] \times G) \triangleq \int_0^t \int \mathbb{1}_G(v^n(s,y)) \mathfrak{q}^n(ds,dy), \quad t \in \mathbb{R}_+, G \in \mathcal{B}(\mathbb{R}^d \backslash \{0\}),$$

and

$$B \triangleq \int_0^\cdot \int b(t,y) q(dt,dy),$$
$$C \triangleq \int_0^\cdot \sigma_t c_t \sigma_t^* dt,$$
$$\nu([0,t] \times G) \triangleq \int_0^t \int \mathbb{1}_G(v(s,y)) \mathfrak{q}(ds,dy), \quad t \in \mathbb{R}_+, G \in \mathcal{B}(\mathbb{R}^d \backslash \{0\}).$$

It is implicit[8] that (B^n, C^n, ν^n) and (B, C, ν) are candidate triplets in the sense of [25, III.2.3]. Moreover, for every $t \in \Gamma$ and $g \in C_1(\mathbb{R}^d)$, the functions $B_t, \widetilde{C}_t$ and $g * \nu_t$ are continuous on $\mathbb{D}(\mathbb{R}^d)$ endowed with the local uniform topology.

Assumption 5.18. Let $\kappa \colon \mathbb{R}_+ \to \mathbb{R}_+$ be increasing and continuous. For each $T \in \mathbb{N}$, there exists a sequence $\gamma^n = \gamma^{n,T} \colon \mathbb{R}_+ \times \Omega \times E \to \mathbb{R}_+$ of non-negative $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F} \otimes \mathcal{E}/\mathcal{B}(\mathbb{R}_+)$-measurable functions with the following properties:

(i) a.s. $\gamma^n * \mathfrak{p}_T^n < \infty$ for all $n \in \mathbb{N}$;
(ii) $\|v^n\| \le \gamma^n \kappa(\sup_{s \le \cdot} \|X_s^n\|)$ on $\Omega \times (T-1, T] \times E$;
(iii) there exists a càdlàg measurable process $Z = Z^T$ such that

$$\sup_{s \le T} |\gamma^n * \mathfrak{p}_s^n - Z_s| \to 0$$

in probability $n \to \infty$.

Corollary 5.19. *Suppose that Assumptions 5.16, 5.17 and 5.18 hold. Then, X is a semimartingale on $\mathbb{B}$ whose semimartingale characteristics are given by $(B(X), C(X), \nu(X))$. Possibly on a standard extension of $\mathbb{B}$, there is a Gaussian continuous martingale W with covariance function C and a Poisson random measure $\mathfrak{p}$ with intensity measure $\mathfrak{q}$ such that*

$$X = X_0 + \int_0^\cdot \int b(t,X,y) q(dt,dy) + \int_0^\cdot \sigma_t(X) dW_t + \int_0^\cdot \int h(v(t,X,y))(\mathfrak{p} - \mathfrak{q})(dt,dy)$$
$$+ \int_0^\cdot \int (v(t,X,y) - h(v(t,X,y)))\mathfrak{p}(dt,dy).$$

Proof. Our strategy is to apply Theorem 5.8. By hypothesis, referring to Theorem 5.8, (i)' and (ii)' and the continuity assumption from (iii)' hold. Via passing to a subsequence, which we ignore in our notation for simplicity, we can assume that a.s.

$$\sup_{s \le T} |\gamma^{n,T} * \mathfrak{p}_s^n - Z_s^T| \to 0, \qquad T = 1, 2, \ldots,$$

[8] In particular, $\Delta B_t^n = \int h(x) \nu^n(\{t\} \times dx)$, which means $\int b^n(t,y) q^n(\{t\} \times dy) = \int h(v^n(t,y)) \mathfrak{q}^n(\{t\} \times dy)$. As a consequence, $\Delta X_t^n = \int v^n(t,y) \mathfrak{p}^n(\{t\} \times dy)$.

as $n \to \infty$. Define

$$\Omega^o \triangleq \Big\{\omega \in \Omega \colon \lim_{n\to\infty} \sup_{t\le T} |\gamma^{n,T} * \mathfrak{p}^n_t - Z^T_t| = 0,\ T = 1, 2, \dots \Big\},$$

which then is a full set. Let $(U, \mathcal{U}) = (\Omega, \mathcal{F})$ and let $L^n(\omega) = \omega$ be the corresponding identity map. The distributions of $L^1, L^2, \dots$ are trivially (relatively) compact in $M_m(U)$. For $T \in \mathbb{N}$ and $(\omega, t) \in \Omega \times (T-1, T]$, we define

$$\mathfrak{u}(\omega, t) \triangleq \begin{cases} \sup_{n\in\mathbb{N}} \int \gamma^{n,T}(\omega; t, y)\mathfrak{p}^n(\omega; \{t\} \times dy), & \omega \in \Omega^o, \\ 0, & \text{otherwise.} \end{cases}$$

We also set $\mathfrak{u}(\omega, 0) \equiv 0$ for all $\omega \in \Omega$. By definition of Ω^o, the fact that Z^T has càdlàg paths and standard properties of càdlàg functions, for every $\omega \in \Omega$ and $a > 0$, the set $\{t \in (T-1, T] \colon \mathfrak{u}(\omega, t) \ge a\}$ is finite, see Example 5.12 for more details. Consequently, by Lemma 2.10, $\mathfrak{u}$ is as in Section 5.5 and it remains to verify (5.11). For every $T-1 < t \le T$, we have a.s.

$$\begin{aligned} \|\Delta X^n_t\| &\le \int \|v^n(t, y)\| \mathfrak{p}^n(\{t\} \times dy) \\ &\le \int \gamma^{n,T}(t, y)\mathfrak{p}^n(\{t\} \times dy)\kappa\Big(\sup_{s\le t} \|X^n_s\|\Big) \\ &\le \mathfrak{u}(L^n, t)\kappa\Big(\sup_{s\le t} \|X^n_s\|\Big), \end{aligned}$$

which shows (5.11). In summary, we conclude from Theorem 5.8 that X is a semimartingale (for its canonical filtration) with characteristics $(B(X), C(X), \nu(X))$.

The final claim, i.e. the representation of X as stochastic integrals, follows from classical representation theorems as given in [26] and [21, Section XIV.3]. □

There is also a version of Corollary 5.19 for the case $\mathbb{B}^n = (\Omega, \mathcal{F}, \mathbf{F}^n, P^n)$, i.e. with varying probability measures. Before we present this version, let us emphasis that even if $P^1, P^2, \dots$ are allowed to be different, we ask them to be quite close in the sense that they converge to each other in *total variation*.

Corollary 5.20. *Suppose that*

$$\sup_{G\in\mathcal{F}} |P^n(G) - P(G)| \to 0 \tag{5.12}$$

as $n \to \infty$, that the Assumptions 5.16 and 5.17 hold, and that Assumption 5.18 holds with (i) *and* (iii) *replaced by*

(i)' P^n*-a.s.* $\gamma^n * \mathfrak{p}^n_T < \infty$ *for all* $n \in \mathbb{N}$*;*

(iii)' *there exists a càdlàg measurable process* $Z = Z^T$ *such that for all* $\varepsilon > 0$

$$P^n\Big(\sup_{s\le T} |\gamma^n * \mathfrak{p}^n_s - Z_s| \ge \varepsilon\Big) \to 0$$

as $n \to \infty$.

Then, Corollary 5.19 holds for $\mathbb{B}^n = (\Omega, \mathcal{F}, \mathbf{F}^n, P^n)$ *and* $\mathbb{B} = (\Omega, \mathcal{F}, \mathbf{F}^X, P)$*.*

Proof. Thanks to Remark 2.3, (5.12) implies $P^n \to P$ in $M_m(\Omega)$. Thus, the set $\{P^n \colon n \in \mathbb{N}\}$ is relatively sequentially compact in $M_m(\Omega)$. Under (5.12), we have, for every $\varepsilon > 0$,

$$P^n\Big(\sup_{s\le T} |\gamma^n * \mathfrak{p}^n_s - Z_s| \ge \varepsilon\Big) \to 0 \quad \Longleftrightarrow \quad P\Big(\sup_{s\le T} |\gamma^n * \mathfrak{p}^n_s - Z_s| \ge \varepsilon\Big) \to 0.$$

Furthermore, with Ω^o as in the proof of Corollary 5.19, $P^n(\Omega^o) \to P(\Omega^o) = 1$. With these observations at hand, the proof of Corollary 5.19 needs no further change. □

Finally, we remark that Theorem 5.14 can also be transferred to the current setting. This yields a stability result for random coefficients b, σ and v. We leave the precise statement to the reader.

References

[1] E. Abi Jaber, C. Cuchiero, M. Larsson, and S. Pulido. A weak solution theory for stochastic Volterra equations of convolution type. *The Annals of Applied Probability*, 31(6):2924–2952, 2021. MR4350978

[2] E. J. Balder. On ws-convergence of product measures. *Mathematics of Operations Research*, 28(3):494–518, 2001. MR1849882

[3] V. Bansaye, T. G. Kurtz, and F. Simatos. Tightness for processes with fixed points of discontinuities and applications in varying environment. *Electronic Communications in Probability*, 21(81):1–9, 2016. MR3580450

[4] V. Bansaye, J. C. Pardo Millan, and C. Smadi. On the extinction of continuous state branching processes with catastrophes. *Electronic Journal of Probability*, 18(106):1–31, 2013. MR3151726

[5] V. Bansaye and F. Simatos. On the scaling limits of Galton-Watson processes in varying environments. *Electronic Journal of Probability*, 20(75):1–36, 2015. MR3371434

[6] R. Bass. A stochastic differential equation with a sticky point. *Electronic Journal of Probability*, 19(32):1–22, 2014. MR3183576

[7] A. Bélanger, S. E. Shreve, and D. Wong. A general framework for pricing credit risk. *Mathematical Finance*, 14(3):317–350, 2004. MR2070167

[8] D. Criens. Limit theorems for cylindrical martingale problems associated with Lévy generators. *Journal of Theoretical Probability*, 33(2):866–905, 2020. MR4091582

[9] D. Criens and M. Ritter. On a theorem by A.S. Cherny for semilinear stochastic partial differential equations. *Journal of Theoretical Probability*, 35:2052–2067, 2022. MR4488571

[10] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*. Cambridge University Press, 2nd edition, 2014. MR3236753

[11] H.-J. Engelbert and G. Peskir. Stochastic differential equations for sticky Brownian motion. *Stochastics*, 86(6):993–1021, 2014. MR3271518

[12] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, 2005. MR0838085

[13] C. Fontana, Z. Grbac, M. Jeanblanc, and Q. Li. Information, no-arbitrage and completeness for asset price models with a change point. *Stochastic Processes and their Applications*, 124(9):3009–3030, 2014. MR3217431

[14] C. Fontana, Z. Grbac, S. Gümbel, and T. Schmidt. Term structure modelling for multiple curves with stochastic discontinuities. *Finance and Stochastics*, 24(2):465–511, 2020. MR4078341

[15] D. Freedman. *Brownian motion and diffusion*. Springer New York, 1983. MR0686607

[16] P. Gänssler. Compactness and sequential compactness in spaces of measures. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 17:124–146, 1971. MR0283562

[17] F. Gehmlich and T. Schmidt. Dynamic defaultable term structure modeling beyond the intensity paradigm. *Mathematical Finance*, 28(1):211–239, 2018. MR3758922

[18] J. K. Ghosh and R. V. Ramamoorthi. *Bayesian Nonparametrics*. Springer New York, 2003. MR1992245

[19] M. Hutzenthaler, P. Pfaffelhuber and C. Printz. Stochastic averaging for multiscale Markov processes with an application to a Wright-Fisher model with fluctuating selection. Arxiv preprint arXiv:1504.01508v2, 2018.

[20] K. Itô and H. P. McKean. *Diffusion processes and their sample paths*. Springer Berlin Heidelberg, 1974. MR0345224

[21] J. Jacod. *Calcul stochastique et problèmes de martingales*. Springer Berlin Heidelberg New York, 1979. MR0542115

[22] J. Jacod and J. Mémin. Existence of weak solutions for stochastic differential equations with driving semimartingales. *Stochastics*, 4(4):317–337, 1981. MR0609691

[23] J. Jacod and J. Mémin. Sur un type de convergence intermédiaire entre la convergence en loi et la convergence en probabilité. *Séminaire de probabilités de Strasbourg*, 15:529–546, 1981. MR0622586

[24] J. Jacod and J. Mémin. Weak and strong solutions of stochastic differential equations: Existence and stability. In D. Williams, editor, *Stochastic Integrals*, volume 851 of *Lecture Notes in Mathematics*, pages 169–212. Springer Berlin Heidelberg, 1981. MR0620991

[25] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer Berlin Heidelberg, 2nd edition, 2003. MR1943877

[26] Y. M. Kabanov, R. S. Liptser, and A. N. Shiryaev. On the representation of integral-valued random measures and local martingales by means of random measures with deterministic compensators. *Mathematics of the USSR-Sbornik*, 39:267–280, 1981. MR0564354

[27] O. Kallenberg. *Foundations of Modern Probability*. Springer New York Berlin Heidelberg, 1997. MR1464694

[28] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer Science+Business Media New York, 2nd edition, 1991. MR1121940

[29] M. Keller-Ressel, T. Schmidt, and R. Wardenga. Affine processes beyond stochastic continuity. *The Annals of Applied Probability*, 29(6):3387–3437, 2019. MR4047984

[30] R. Merton. On the pricing of corporate debt: the risk structure of interest rates. *The Journal of Finance*, 29(2):449–470, 1974.

[31] M. Ondreját and J. Seidler. On the existence of progressively measurable modifications. *Electronic Communications in Probability*, 18(20):1–6, 2013. MR3037218

[32] Y. V. Prokhorov. Convergence of random processes and limit theorems in probability theory. *Theory of Probability & Its Applications*, 1(2):157–214, 1956. MR0084896

[33] P. Protter. $\mathcal{H}^p$ stability of solutions of stochastic differential equations. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 44(4):337–352, 1978.

[34] M. Schäl. On dynamic programming: compactness of the space of policies. *Stochastic Processes and their Applications*, 3:345–364, 1975. MR0386706

[35] J. Seidler. Weak convergence of infinite-dimensional diffusions. *Stochastic Analysis and Applications*, 15(3):399–417, 1997. MR1454096

[36] A. V. Skorokhod. Limit theorems for stochastic processes. *Theory of Probability & Its Applications*, 1(3):261–290, 1956. MR0084897

[37] D. W. Stroock and S. R. S. Varadhan. *Multidimensional Diffussion Processes*. Springer Berlin Heidelberg, 1979. MR0532498

Acknowledgments. The authors are grateful to an anonymous referee for many helpful comments and suggestions. DC acknowledges financial support from the DFG project No. SCHM 2160/15-1.