

Kiiski, Matti

**Article — Published Version**

## The Riesz representation theorem and weak\* compactness of semimartingales

Finance and Stochastics

**Provided in Cooperation with:**

Springer Nature

*Suggested Citation:* Kiiski, Matti (2020) : The Riesz representation theorem and weak\* compactness of semimartingales, Finance and Stochastics, ISSN 1432-1122, Springer, Berlin, Heidelberg, Vol. 24, Iss. 4, pp. 827-870, <https://doi.org/10.1007/s00780-020-00432-5>

This Version is available at:

<https://hdl.handle.net/10419/288268>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>



# The Riesz representation theorem and weak\* compactness of semimartingales

Matti Kiiski<sup>1</sup>

Received: 25 September 2018 / Accepted: 18 March 2020 / Published online: 22 July 2020  
© The Author(s) 2020

**Abstract** We show that the sequential closure of a family of probability measures on the canonical space of càdlàg paths satisfying Stricker’s uniform tightness condition is a weak\* compact set of semimartingale measures in the dual pairing of bounded continuous functions and Radon measures, that is, the dual pairing from the Riesz representation theorem under topological assumptions on the path space. Similar results are obtained for quasi- and supermartingales under analogous conditions. In particular, we give a full characterisation of the strongest topology on the Skorokhod space for which these results are true.

**Keywords** Skorokhod space · Meyer–Zheng topology ·  $S$ -topology · Weak\* topology · Càdlàg semimartingale

**Mathematics Subject Classification (2010)** 28C05 · 54D30 · 60B05 · 60G05

**JEL Classification** C02 · G13

## 1 Introduction

The Riesz representation theorem states that the operation of integration defines a one-to-one correspondence between the continuous linear functionals on the bounded continuous functions and the Radon measures on a topological space. On the Skorokhod space, it provides a locally convex way of constructing all càdlàg stochastic processes on the canonical space as *tight* probability measures. On a conceptual level,

---

Research is partly supported by a Swiss National Foundation Grant SNF 200020-172815.

✉ M. Kiiski  
matti.kiiski@gmail.com

<sup>1</sup> Institute of Mathematics, University of Mannheim, B 6, 26, 68159 Mannheim, Germany

any criterion that characterises a certain object should give rise to some kind of compactness when applied uniformly to a family of objects. We relate Stricker's uniform tightness condition of semimartingales to the weak\* compactness in the dual pairing of bounded continuous functions and Radon measures, that is, the dual pairing from the Riesz representation theorem on the canonical space of càdlàg paths.

The weak topologies on the Skorohod space and weak convergence of stochastic processes, that is, the sequential convergence of laws of stochastic processes for the weak\* topology induced by the Riesz representation theorem, have been studied earlier in the works of Meyer and Zheng [44], Zheng [58], Stricker [55], Jakubowski et al. [34], Kurtz [40], Lowther [42] and Jakubowski [31, 33]. We rely heavily on these earlier results. In particular, we utilise the stability results of Kellerer [36], Meyer and Zheng [44], Jakubowski et al. [34], Jakubowski [31] and Lowther [42].

Weak topologies are rich in terms of convergent subsequences and have found various applications in studying convergence of financial markets, see Prigent [46, Chap. 2], time series analysis in econometrics, see Chan and Zhang [11], stochastic optimal control, see Kurtz and Stockbridge [41], Bahlali and Gherbal [2], Tan and Touzi [56], and martingale optimal transport, as introduced by Beiglböck et al. [3], and extended for a continuous-time parameter by Dolinsky and Soner [15] and Guo et al. [24]. We aim to provide a functional-analytic framework that unifies and elaborates these existing results and allows extending the analysis beyond the convergence of sequences. In particular, the framework allows studying the non-sequential compactness of families of semimartingales. The question of compactness arises naturally in the context of convex conjugate duality on functions and measures on the càdlàg path space. Many problems in quantitative risk management can be embedded in this framework as convex risk measures and their conjugates; see Föllmer and Schied [21, Chap. 4]. A classical example is the minimal superhedging cost of a derivative contract over a convex set of hedging positions. We recall the connection between compactness and the superhedging duality that yields model-independent price bounds for derivative contracts as originally observed by Beiglböck et al. [3], and extended to continuous time by Dolinsky and Soner [15] and Guo et al. [24].

The objective of the paper is to provide a weak\* compactness result for càdlàg semimartingales under the most general topological assumption on the path space. Our main contribution is to unify the previous results on the weak convergence of semimartingales and provide an easy method for constructing weak\* compact sets of semimartingales on the canonical space of càdlàg paths. We also give examples of such sets and show that the examples are consistent with earlier results for Banach spaces of stochastic processes defined over a common probability space.

We characterise the strongest topology on the Skorokhod space for which our main result is true. A natural candidate is Jakubowski's  $S$ -topology, due to its tightness criteria. However, it is an open problem whether the  $S$ -topology possesses a separation property so that the aforementioned Riesz representation theorem is true. We address the problem of regularity by introducing a new weak topology on the Skorokhod space that has the same continuous functions as the  $S$ -topology, suitable compact sets and additionally satisfies a strong separation axiom. The topological space is perfectly normal ( $T_6$ ) in comparison to the Hausdorff property ( $T_2$ ) that has been verified for the  $S$ -topology. The topology is obtained from Jakubowski's  $S$ -topology as a

result of a standard regularisation method that appears already in the classical works of Alexandroff [1, Chap. 2] and Knowles [38]. Our contribution is to carefully show that the important properties of the  $S$ -topology are preserved in the regularisation.

The rest of the paper is organised as follows.

In Sect. 2, we give rigorous definitions of semimartingale measures and related notions on the canonical space of càdlàg paths. We also provide a brief introduction to the aforementioned Riesz representation theorem that is the basis of our approach. The main results, examples and financial motivation are given in Sect. 3. The proofs of the main results and the required auxiliary results are provided in Sect. 4. In Sect. 5, we characterise the strongest topology on the Skorokhod space for which the results of the two previous sections are true. Some definitions and technical results are omitted in the main part of the article and are gathered in the [Appendix](#).

**Conventions and notations** Throughout, the comparatives ‘weaker’ and ‘stronger’ should be understood in the wide sense ‘weaker or equally strong’ and ‘stronger or equally strong’, respectively. We say that two topologies are ‘comparable’ if one is stronger than the other.

We fix the following notations:  $\mathbb{N}_0 := \{0\} \cup \mathbb{N}$  and  $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ . For any  $x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$ , we set  $|x| := \sum_{i=1}^d |x_i|$ ,  $x^+ = \sum_{i=1}^d x_i^+$ ,  $x^- = \sum_{i=1}^d x_i^-$  and  $\|x\|_\infty := |x_1| \vee |x_2| \vee \dots \vee |x_d|$ .

Further,  $\mathbb{D}$ ,  $\mathbb{D}(I)$  and  $\mathbb{D}(I; \mathbb{R}^d)$  denote the  $\mathbb{R}^d$ -valued càdlàg functions on  $I$ ;  $\mathbb{V}(I)$  denotes the functions of finite variation on  $I$ . We equip  $\mathbb{D}$  with a topology which is not yet specified at this point.

By  $\mathbb{C}(\mathbb{X})$  we denote the family of continuous functions on a topological space  $\mathbb{X}$ , e.g.  $\mathbb{X} = \mathbb{D}$ , while  $\mathbb{U}(\mathbb{D})$  denotes the family of upper semicontinuous functions on  $\mathbb{D}$ , and  $\mathbb{B}(\mathbb{D})$  the family of Borel functions on  $\mathbb{D}$ . We add a subscript  $b$  if we want the functions to be bounded as well. Also  $\mathbb{B}_0(\mathbb{D})$  denotes the family of bounded Borel functions on  $\mathbb{D}$  vanishing at infinity.

When talking about measures on a topological space, we follow Bogachev [8, Chap. 7]; so all measures are  $\sigma$ -additive and have values in  $\mathbb{R}$ , hence are signed measures of finite total variation. By  $\mathbb{M}_t(\mathbb{D})$  we denote the family of Radon measures (of finite total variation) on the Skorokhod space  $\mathbb{D}$ , while  $\mathbb{M}_\tau(\mathbb{D})$  (resp.  $\mathbb{M}_\sigma(\mathbb{D})$ ) denotes the family of  $\tau$ -additive (resp.  $\sigma$ -additive) Borel measures (of finite total variation) on  $\mathbb{D}$ ; see [8, Definition 7.2.1]. If  $\mathbb{M}_t(\mathbb{D}) = \mathbb{M}_\tau(\mathbb{D}) = \mathbb{M}_\sigma(\mathbb{D})$ , we denote all three families by  $\mathbb{M}(\mathbb{D})$ . Also  $\mathbb{M}_+(\mathbb{D})$  denotes the family of nonnegative elements of  $\mathbb{M}(\mathbb{D})$ .

By  $\mathbb{P}(\mathbb{D})$  (resp.  $\mathbb{P}(\mathbb{R}^d)$ ) we denote the family of all probability measures on the Skorokhod space  $\mathbb{D}$  (resp. the Euclidean space  $\mathbb{R}^d$ ).

Moreover,  $\mathcal{K}(\mathbb{D})$ ,  $\mathcal{B}(\mathbb{D})$  and  $\mathcal{B}a(\mathbb{D})$  denote the families of compact, Borel and Baire sets on the Skorokhod space  $\mathbb{D}$ , respectively. Similarly,  $\mathcal{B}(\mathbb{R}^d)$  denotes the Borel sets on the Euclidean space  $\mathbb{R}^d$ .

For a process  $X : \Omega \times I \rightarrow \mathbb{R}^d$  and a subset  $J \subseteq I$ , we write  $X_J$  for the restriction of  $X$  to  $J$ , i.e.,  $X_J : \Omega \times J \rightarrow \mathbb{R}^d$ . In the case of a singleton  $J = \{t\}$ ,  $t \in I$ , we suppress the dependence on the set  $J$  and simply write  $X_t$  for the value of  $X$  at  $t$ , whence  $X_t : \Omega \rightarrow \mathbb{R}^d$ . We set  $\|X\|_\infty(\omega) := \sup_{t \in I} \|X_t(\omega)\|_\infty$  and  $\|X\|_\mathbb{V}(\omega) := \sum_{i=1}^d \sup_\pi \{|X_0^i(\omega)| + \sum_{k=1}^n |X_{t_k}^i(\omega) - X_{t_{k-1}}^i(\omega)|\}$ , where the supremum is taken over all finite partitions  $\pi$  of  $I$ .

By  $E_Q[f] := \int f dQ$  we denote the integral, for a measurable function  $f : \Omega \rightarrow \mathbb{R}$  and a probability measure  $Q$  on  $(\Omega, \mathcal{F})$ , and by  $\|f\|_{L^p(Q)}$  the  $L^p$ -norm. For  $X : \Omega \times I \rightarrow \mathbb{R}^d$ , we set  $\|X\|_{L^{p,\infty}(Q)} := \|\|X\|_\infty\|_{L^p(Q)}$ .

For  $p \geq 1$ ,  $\|X\|_{\mathcal{H}^p(Q)} := \|\|M\|_\infty + \|A\|_{\mathbb{V}}\|_{L^p(Q)}$  is the (maximal)  $\mathcal{H}^p$ -norm for a semimartingale  $X$  whose canonical decomposition under  $Q$  is  $X = M + A$ .

Also  $\mathcal{E}(Q)$  denotes the family of elementary predictable processes which are  $Q$ -a.s. bounded by 1, for a probability measure  $Q$  on  $(\Omega, \mathcal{F})$ . Then

$$(H \bullet X)_t := \sum_{i=1}^d H_0^i X_0^i + \sum_{i=1}^d \sum_{k=1}^n H_{t_{k-1}}^i (X_{t \wedge t_k^i} - X_{t \wedge t_{k-1}^i}), \quad t \in I, \tag{1.1}$$

denotes the stochastic integral for  $H \in \mathcal{E}(Q)$  and the canonical process  $X$ ; the dependence on  $Q$  is omitted in this notation. For  $p \geq 1$ , we set

$$\|X\|_{\mathcal{E}^p(Q)} := \sup_{H \in \mathcal{E}(Q)} \|(H \bullet X)\|_{L^p(Q)}.$$

By  $N_\pi^{a,b}$  we denote the number of upcrossings of an interval  $[a, b]$  with respect to a partition  $\pi$  of  $I$ , and by  $N^{a,b} = \sup_\pi N_\pi^{a,b}$  the number of upcrossings of an interval  $[a, b]$ .

Also  $\Delta X := X_t - X_{t-}$  denotes the jump of  $X$  at  $t$ ;  $[\omega]^t$  denotes the restriction of  $\omega$  to  $[0, t]$ , while  $\rightarrow_Q$  denotes convergence in  $L^0(Q)$ .

**Terminology** We provide some frequently used terminology. Standard literature references for general topology and topological measure theory are [16, Chap. 1] and [8, Chap. 7].

Let  $X$  be a topological space. A subset  $Z \subseteq X$  is called *compact* if every cover of this set by open sets contains a finite subcover; *relatively compact* if  $Z$  is contained in a compact set; *sequentially compact* if every infinite sequence of elements of  $Z$  contains a subsequence converging to an element of  $Z$ ; *relatively sequentially compact* if every infinite sequence of elements of  $Z$  contains a subsequence converging in  $X$ .

*Remark 1.1* In contrast to the case of a metric space, in a general topological space, neither does compactness imply sequential compactness nor the other way round.

The *closure*  $\text{cl } Z$  of  $Z$  is the set of all points  $x \in X$  such that every neighbourhood of  $x$  contains at least one point of  $Z$ . The *sequential closure*  $[Z]_{\text{seq}}$  of  $Z$  is the set of all points  $x \in X$  for which there is a sequence in  $Z$  that converges to  $x$ .

*Remark 1.2* In contrast to the case of a metric space, in a general topological space, the sequential closure of a set is not necessarily a sequentially closed set.

All topological spaces considered are *Hausdorff* ( $T_2$ ), and a Hausdorff space  $X$  is called:

– *regular* ( $T_3$ ) if for every point  $x \in X$  and every closed set  $Z$  in  $X$  not containing  $x$ , there exist disjoint open sets  $U$  and  $V$  such that  $x \in U$  and  $Z \subseteq V$ ;

- *completely regular* ( $T_{3\frac{1}{2}}$ ) if for every point  $x \in X$  and every closed set  $Z$  in  $X$  not containing  $x$ , there exists a continuous function  $f : X \rightarrow [0, 1]$  such that  $f(x) = 1$  and  $f(z) = 0$  for all  $z \in Z$ ;
- *perfectly normal* ( $T_6$ ) if every closed set  $Z \subseteq X$  has the form  $Z = f^{-1}(0)$  for some continuous function  $f$  on  $X$ ;
- *paracompact* if every open cover of  $X$  has an open refinement that is locally finite;
- *k-space* if a set  $Z \subseteq X$  is closed in  $X$  provided that the intersection of  $Z$  with any compact subspace  $K$  of  $X$  is closed in  $K$ ;
- *sequential space* if every sequentially closed set is closed;
- *Fréchet–Urysohn space* if every subspace is a sequential space;
- *Polish space* if the space is homeomorphic to a complete separable metric space;
- *Lusin space* if the space is the image of a complete separable metric space under a continuous one-to-one mapping;
- *Souslin space* if the space is the image of a complete separable metric space under a continuous mapping;
- *Radon space* if every Borel measure on the space is a Radon measure;
- *perfect space* if every Borel measure on the space is perfect, i.e., for every Borel-measurable function  $f$  and every Borel measure  $Q$ , the set  $f(X)$  contains a Borel set  $B$  for which  $Q[f^{-1}(B)] = Q[X]$ ;
- *angelic space* if every set  $Z \subseteq X$  with the property that every infinite sequence of its elements has a limit point in  $X$  also possesses the following properties:  $Z$  is relatively compact and each point in the closure of  $Z$  is the limit of some sequence in  $Z$ .

**Remark 1.3** For closed subspaces, all these properties are hereditary, meaning that if the space has a property, then a closed subspace endowed with the relative topology has that property as well. So all discussion on these properties generalises as such for relative topologies on closed sets.

## 2 Càdlàg semimartingales as linear functionals

In this preliminary section, we define the canonical space for càdlàg semimartingales and related measures and continuous linear functionals.

### 2.1 Canonical space of càdlàg paths

We fix  $I$  to denote a usual time index set of a stochastic process, i.e.,  $I := [0, T]$  for  $0 < T \leq \infty$  or  $I := [0, \infty)$ . The Skorokhod space  $\mathbb{D}(I; \mathbb{R}^d)$ ,  $d \in \mathbb{N}$ , with the domain  $I$  consists of all  $\mathbb{R}^d$ -valued functions  $\omega$  on  $I$  that admit a limit  $\omega(t-)$  from the left for every  $t > 0$  and are continuous from right,  $\omega(t) = \omega(t+)$ , for every  $t < T$ . We call such functions càdlàg. The space  $\mathbb{D}([0, \infty]; \mathbb{R}^d)$  is regarded as the product space  $\mathbb{D}([0, \infty); \mathbb{R}^d) \times \mathbb{R}^d$ ; see Appendix A.1 and note the difference between  $[0, \infty]$  and  $[0, \infty)$ . We write  $\omega = (\omega^1, \dots, \omega^d)$  for  $\omega \in \mathbb{D}(I; \mathbb{R}^d)$  with  $\omega(t) = (\omega^1(t), \dots, \omega^d(t))$  for every  $t \in I$ . We denote by  $X$  the canonical process on  $\mathbb{D}(I; \mathbb{R}^d)$ , i.e.,  $X_t(\omega) = \omega(t)$  for all  $(t, \omega) \in I \times \mathbb{D}(I; \mathbb{R}^d)$ . We write  $X^i$  for each coordinate process of the canonical process  $X$  for  $i \leq d$ .

We endow the Skorokhod space  $\mathbb{D}(I; \mathbb{R}^d)$ ,  $d \in \mathbb{N}$ , with the right-continuous version  $\mathcal{F}_t := \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}^0$  of the raw canonical filtration  $\mathcal{F}_t^0 := \sigma(X_s : s \leq t)$  generated by the canonical process  $X$  on  $\mathbb{D}(I; \mathbb{R}^d)$ .

*Remark 2.1* The right-continuous version of the raw canonical filtration is needed in the proof of Proposition 4.15. Alternatively, we could use the universal completion of the raw canonical filtration; see Proposition 4.3 (b).

A càdlàg stochastic process is identified with a probability measure on the filtered canonical space  $(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T, (\mathcal{F}_t)_{t \in I})$ , where  $T = \sup_{t \in I} t \in (0, \infty]$  and  $\mathcal{F}_T := \bigvee_{t \in I} \mathcal{F}_t^0 = \bigvee_{t \in I} \mathcal{F}_t$ . The family of all probability measures, i.e., càdlàg processes, on  $(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T)$  is denoted by  $\mathbb{P}(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T)$ , and two elements of  $\mathbb{P}(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T)$  are identified as usual, i.e.,  $P = Q$ , if (and only if) one has  $P[F] = Q[F]$  for all  $F \in \mathcal{F}_T$ ; cf. Sect. 4.1.3.

### 2.2 Semimartingales on the Skorokhod space

We recall some basic concepts of semimartingale theory in the present setting. All semimartingales are assumed càdlàg. We adopt the terminology of Dolinsky and Soner [15] and Guo et al. [24] and call a probability measure  $Q$  on  $(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T)$  a *martingale measure* if the canonical process  $X$  is a martingale on  $(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T, (\mathcal{F}_t)_{t \in I}, Q)$  (up to  $\infty$  if  $I = [0, \infty]$ ). Note that  $X$  is a martingale on  $[0, \infty]$  if and only if  $X$  is a uniformly integrable martingale on  $[0, \infty)$ ; see e.g. [45, Theorem 1.1.2 (2)]. We say that the canonical process  $X$  is  $L^p$ -*bounded*, for some  $p \geq 1$ , on  $(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T, Q)$  if  $\sup_{t \in I} \|X_t\|_{L^p(Q)} < \infty$ .

(*Special semimartingale and supermartingale measures* are defined similarly to martingale measures. On  $(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T, (\mathcal{F}_t)_{t \in I}, Q)$ , for a fixed probability measure  $Q$ , let  $\mathcal{E}(Q)$  denote the family of elementary predictable integrands, i.e., the family of adapted càglàd processes of the form

$$H^i = H_0^i \mathbb{1}_{\{0\}} + \sum_{k=1}^n H_{t_{k-1}}^i \mathbb{1}_{(t_{k-1}^i, t_k^i]}, \quad i \leq d, \tag{2.1}$$

where  $n \in \mathbb{N}$ ,  $0 = t_0^i \leq t_1^i \leq \dots \leq t_n^i$  are in  $I$  and each  $H_{t_k^i}^i$  is an  $\mathcal{F}_{t_k^i}^i$ -measurable random variable satisfying  $|H_{t_k^i}^i| \leq 1$   $Q$ -a.s. For a family  $\mathcal{Q} \subseteq \mathbb{P}(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T)$ , consider the condition

$$\lim_{c \rightarrow \infty} \sup_{Q \in \mathcal{Q}} \sup_{H \in \mathcal{E}(Q)} Q[|(H \bullet X)_t| > c] = 0, \quad \forall t \in I, \tag{UT}$$

where  $H \bullet X$  is the elementary stochastic integral defined in (1.1). The condition (UT) was introduced by Stricker in [55]. By the classical result of Bichteler–Dellacherie–Mokobodzki, a probability measure  $Q$  on  $(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T)$  is an  $(\mathcal{F}_t)_{t \in I}$ -semimartingale measure if and only if  $Q = \{Q\}$  satisfies the condition (UT). The family (2.1) of processes generates the predictable  $\sigma$ -algebra, and the condition (UT) is sometimes called the *predictable uniform tightness condition* (P-UT). Remark that for

semimartingale measures, no integrability condition is imposed on  $X_0$ , i.e., the localisation of the local martingale in a semimartingale decomposition is understood in the sense of [25, Definition 7.1]; cf. [29, Remark 6.3]. We call a semimartingale measure  $Q$  of class  $\mathcal{H}^p$  if

$$\|X\|_{\mathcal{H}^p(Q)} = \inf \left\{ \|M\|_{\infty} + \|A\|_{\mathbb{V}} \Big|_{L^p(Q)} : X = M + A \right\} < \infty,$$

where the infimum is taken over all semimartingale decompositions of  $X$  into a local martingale  $M$  and a finite variation process  $A$  with  $A_0 = 0$ , all on the space  $(\mathbb{D}(I; \mathbb{R}^d), (\mathcal{F}_t)_{t \in I}, \mathcal{F}_T, Q)$ . In fact, if  $Q$  is a semimartingale measure of class  $\mathcal{H}^p$  for some  $p \geq 1$ , then  $X$  decomposes into a martingale  $M$  and a *predictable* finite variation process  $A$ , i.e.,  $X$  is a *special semimartingale* under  $Q$ .

To obtain compact statements for quasi- and supermartingales, we introduce two conditions. The first condition is

$$\sup_{Q \in \mathcal{Q}} \sup_{t \in I} \left( E_Q[|X_t|] + \sup_{H \in \mathcal{E}(Q)} E_Q[(H \bullet X)_t] \right) < \infty. \quad (\text{UB})$$

The second condition is the same condition, but the  $L^1$ -boundedness is strengthened to the uniform integrability of the negative parts, for every  $t \in I$ , i.e.,

$$Q \text{ satisfies } (\text{UB}) \quad \text{and} \quad \lim_{c \rightarrow \infty} \sup_{Q \in \mathcal{Q}} E_Q[X_t^- \mathbb{1}_{\{X_t^- > c\}}] = 0, \quad \forall t \in I. \quad (\text{UI})$$

The uniform integrability in (UI) yields the convergence of the first moments that preserves the supermartingale property; see Proposition 4.13. If we insist that  $t_n^i = t$  in (2.1), then the second supremum in (UB) is attained, by choosing

$$H_k^i = \text{sign}(E_Q[X_{t_k^i}^i - X_{t_{k-1}^i}^i \mid \mathcal{F}_{t_{k-1}^i}]), \quad 1 \leq k \leq n, \quad i \leq d,$$

for which the value of the integral is equal to the  $(\mathcal{F}_t)_{t \in I}$ -conditional variation of  $X^i$  on  $[0, t]$ ,

$$\text{Var}_t^Q(X^i) := \sup E_Q \left[ |X^i(0)| + \sum_{k=1}^n \left| E_Q[X_{t_k^i}^i - X_{t_{k-1}^i}^i \mid \mathcal{F}_{t_{k-1}^i}] \right| \right], \quad i \leq d,$$

where the supremum is taken over all partitions  $0 \leq t_0^i \leq t_1^i \leq \dots \leq t_n^i = t$ ,  $n \in \mathbb{N}$ ; see e.g. [14, Appendix II]. A probability measure  $Q$  on  $(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T)$  is a *quasi-martingale measure* if and only if  $Q = \{Q\}$  satisfies the condition (UB); see e.g. [25, Definition 8.12]. Moreover, a quasimartingale is an  $\mathcal{H}^1$ -semimartingale if and only if it is bounded in the  $L^{1,\infty}$ -norm; cf. [14, VII.(98.9)]. Finally, let us note that we have the hierarchy

$$(\text{UI}) \implies (\text{UB}) \implies (\text{UT}). \quad (2.2)$$

The first implication is obvious. The second implication follows from Lemma 2.2.



**Lemma 2.2** *There exists a constant  $b > 0$  such that for any  $Q \in \mathbb{P}(\mathbb{D}(I), \mathcal{F}_T)$ ,  $H \in \mathcal{E}(Q)$  and  $c > 0$ , we have*

$$Q[|(H \bullet X)_t| > c] \leq \frac{b}{c} \left( E_Q[|X_t|] + \sup_{H' \in \mathcal{E}(Q)} E_Q[(H' \bullet X)_t] \right), \quad t \in I, \quad (2.3)$$

where the right-hand side is possibly infinite.

Inequality (2.3) is well known, but we provide a proof for the convenience of the reader in Appendix A.2.

A family  $\mathcal{Q} \subseteq \mathbb{P}(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T)$  is called  $J^1$ -tight if it is exhausted by a sequence of  $J^1$ -compact sets; see Appendix A.1.4. Following the classical terminology [25, Definition 15.48], we say that a family  $\mathcal{Q} \subseteq \mathbb{P}(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T)$  is  $C$ -tight if it is  $J^1$ -tight and satisfies

$$\sup_{Q \in \mathcal{Q}} Q \left[ \sup_{s \leq t} |\Delta X_s| > c \right] = 0, \quad \forall t \in I, \forall c > 0.$$

The paths of the canonical process  $X$  of  $\mathbb{D}(I; \mathbb{R}^d)$  lie in  $\mathbb{C}(I; \mathbb{R}^d)$   $Q$ -almost surely if and only if  $\mathcal{Q} = \{Q\}$  is  $C$ -tight on  $\mathbb{D}(I; \mathbb{R}^d)$ . An analogous assertion is true for Hölder-continuity.

The Markov property is not preserved by the convergence of finite-dimensional marginal distributions, but a stronger property is needed; cf. [42, Sect. 2.3]. Following Lowther [42], we call a probability measure  $Q \in \mathbb{P}(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T)$  Lipschitz–Markov if for every  $s, t \in I$  with  $s < t$ , for every bounded Lipschitz-continuous function  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  with a Lipschitz constant  $L(g) \leq 1$ , there exists a bounded Lipschitz-continuous function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  with a Lipschitz constant  $L(f) \leq 1$  such that

$$f(X_s) = E_Q[g(X_t) | \mathcal{F}_s] \quad Q\text{-a.s.}$$

The Lipschitz–Markov property is indeed stronger than the Markov property; consider the sequence of functions  $g_n(x) = -n \vee |x| \wedge n$ ,  $n \in \mathbb{N}$ , on  $\mathbb{R}^d$ . The Lipschitz–Markov property was first studied by Kellerer [36].

### 2.3 The Riesz representation on the canonical space

The Riesz representation theorem for the laws of  $\mathbb{D}$ -valued random variables, i.e., stochastic processes, requires topological assumptions on the canonical space.

**Assumption 2.3** The Skorokhod space  $\mathbb{D}$  is endowed with a topology under which  $\mathbb{D}$  is a completely regular Radon space and the Borel  $\sigma$ -algebra coincides with the canonical  $\sigma$ -algebra.

The strict topology  $\beta_0$  on  $\mathbb{C}_b(\mathbb{D})$  is a locally convex topology generated by the family of seminorms

$$p_g(f) := \|fg\|_\infty, \quad f \in \mathbb{C}_b(\mathbb{D}), g \in \mathbb{B}_0(\mathbb{D}),$$

where

$$\mathbb{B}_0(\mathbb{D}) := \{f \in \mathbb{B}_b(\mathbb{D}) : \forall \varepsilon > 0 \exists K^\varepsilon \in \mathcal{K}(\mathbb{D}) \text{ such that } |f(x)| < \varepsilon, \forall x \notin K^\varepsilon\}.$$

The collection of finite intersections of the sets

$$V_{g,\varepsilon} := \{f \in \mathbb{C}_b(\mathbb{D}) : p_g(f) < \varepsilon\}, \quad g \in \mathbb{B}_0(\mathbb{D}), \varepsilon > 0,$$

forms a local basis at the origin for the topology  $\beta_0$ .

The linear space of all  $\beta_0$ -continuous linear functionals on  $\mathbb{C}_b(\mathbb{D})$  is isomorphic to the linear space  $\mathbb{M}(\mathbb{D})$  of all countably additive measures of finite total variation on  $\mathbb{D}$ . More precisely, under Assumption 2.3, we have the following Riesz representation theorem on the Skorokhod space. See Sect. 2.3.1 for a reference to a proof.

**Lemma 2.4** *Assume that the Skorokhod space  $\mathbb{D}$  satisfies Assumption 2.3. Then any  $\mu \in \mathbb{M}(\mathbb{D})$  induces a  $\beta_0$ -continuous linear functional on  $\mathbb{C}_b(\mathbb{D})$  by*

$$u_\mu(f) := \int f d\mu, \quad f \in \mathbb{C}_b(\mathbb{D}), \quad (2.4)$$

and any  $\beta_0$ -continuous linear functional on  $\mathbb{C}_b(\mathbb{D})$  is of the form (2.4) for some unique  $\mu \in \mathbb{M}(\mathbb{D})$ , and the one-to-one correspondence  $\mu \leftrightarrow u_\mu$  defined by (2.4) is linear.

For the elements of  $\mathbb{P}(\mathbb{D}) \subseteq \mathbb{M}_+(\mathbb{D})$ , we write

$$E_Q[f] := \int f dQ, \quad f \in \mathbb{C}_b(\mathbb{D}), Q \in \mathbb{P}(\mathbb{D}).$$

The *weak\* topology* on  $\mathbb{M}(\mathbb{D})$  is a locally convex topology generated by the family of seminorms

$$p_f(\mu) := \left| \int f d\mu \right|, \quad f \in \mathbb{C}_b(\mathbb{D}), \mu \in \mathbb{M}(\mathbb{D}). \quad (2.5)$$

We write  $\mu_\alpha \rightarrow_{w^*} \mu$  for a net  $(\mu_\alpha)_{\alpha \in \mathcal{A}}$  with a directed set  $\mathcal{A}$  in  $\mathbb{M}(\mathbb{D})$  converging in the weak\* topology to  $\mu \in \mathbb{M}(\mathbb{D})$ , i.e., if

$$\int f d\mu_\alpha \rightarrow \int f d\mu, \quad \forall f \in \mathbb{C}_b(\mathbb{D}). \quad (2.6)$$

In the case that the Skorokhod space  $\mathbb{D}$  is endowed with a metric topology, the weak\* topology on the nonnegative orthant  $\mathbb{M}_+(\mathbb{D})$  is metrisable, i.e., the family of seminorms (2.5) can be replaced with a single metric, and consequently it is sufficient to consider sequences  $(\mu_n)_{n \in \mathbb{N}}$  in (2.6) that define a topology; see e.g. [8, Sect. 8.3]. Thus the weak\* topology then coincides with the classical notion of the topology of weak convergence widely used in probability theory, that is, the convergence

$$E_{Q_n}[f] \rightarrow E_Q[f], \quad \forall f \in \mathbb{C}_b(\mathbb{D}),$$

for probability measures  $(Q_n)_{n \in \mathbb{N}}$  and  $Q$  on a metric space  $\mathbb{D}$ . Following Sentilles [50], we write ‘weak\*’ in place of ‘weak’ to distinguish the topology from the weak topology on the bounded continuous functions in the pairing of Lemma 2.4.

### 2.3.1 Background

The classical Riesz representation theorem is stated as a Banach space result for bounded continuous functions vanishing at infinity on a locally compact space. The strict topology  $\beta_0$ , introduced by Buck in [10] for locally compact spaces, gives up the Banach space structure, but allows relaxing the assumption that the bounded continuous functions are vanishing at infinity. Further observations in the 1970s by Giles [22] and Hoffmann-Jørgensen [28] led to a generalisation of the Riesz representation theorem for completely regular spaces; locally compact spaces are completely regular. The weak\* topology was thoroughly studied by Sentilles in [50] in the case of a general completely regular space. A streamlined proof for Lemma 2.4 can be found e.g. in the book of Jarchow [35, Sect. 7.6]. The proof relies on the fact that on a completely regular space, every continuous function admits a unique continuous extension to the Stone–Čech compactification of the space. The fact that the underlying topological space is completely regular ( $T_{3\frac{1}{2}}$ ) is also *necessary* for the Riesz representation theorem in the sense that the separation axiom cannot be relaxed to a weaker one as there exist examples of regular ( $T_3$ ) spaces on which every continuous function is a constant, and on such a space, the Riesz representation theorem cannot be true; see [26]. However, in our setting, it suffices to assume that the space is regular; see Sect. 4.1.

## 3 Main results and examples

A stochastic process is regarded as a probability measure on the canonical space, and the family of all probability measures (processes) on the canonical space is endowed with the weak\* topology (2.5) of the Riesz representation theorem (2.4). We impose the following assumption on the canonical space.

**Assumption 3.1** The Skorokhod space is endowed with a regular topology that is weaker than Jakubowski’s  $S$ -topology, but stronger than the Meyer–Zheng (MZ) topology.

The  $S^*$ -topology introduced in Sect. 5 meets the previous requirements and is the strongest topology on the Skorokhod space for which the results are true; see Theorem 5.13 and Remark 5.14.

### 3.1 Motivation by the analysis of a problem in finance

This work was initiated by investigations in robust pricing of derivative contracts by Guo et al. [24, Lemma 3.7] and the current author together with Cheridito et al. [12, Corollary 6.7]. In parallel work [12], we describe a general superhedging/sublinear-pricing paradigm and its relation to the compactness studied in the present work.

In our application, the canonical process  $X$  represents the value of an underlying asset that can consist of liquid options or common stocks. The objective is to determine the price for a derivative contract  $\xi$  that is a function of  $X$ , i.e.,  $\xi : \mathbb{D} \rightarrow \mathbb{R}$ . Every viable pricing model  $Q \in \mathbb{P}(\mathbb{D})$  for a class  $\mathcal{D}(\mathcal{F}_T)$  of  $\mathcal{F}_T$ -measurable derivative contracts on the underlying asset  $X$  must satisfy

$$E_Q[\xi] \leq \Phi(\xi), \quad \forall \xi \in \mathcal{D}(\mathcal{F}_T),$$

where  $\Phi(\xi)$  denotes the greatest lower bound for the initial capital requirement at time  $t = 0$  of all portfolios that produce a value greater than or equal to  $\xi$  at time  $t = T$ , for every possible realisation of the underlying  $X$ . Indeed, otherwise it is possible to make a sure profit by creating a portfolio that consists of a short position in a derivative contract  $\xi \in \mathcal{D}(\mathcal{F}_T)$  and a long position in a portfolio that produces a value greater than or equal to  $\xi$ . On an efficient market, this should not be possible; cf. the seminal work by Black and Scholes [6, Abstract]. On the other hand, for  $\xi \in \mathcal{D}(\mathcal{F}_T)$ , if the least upper bound  $\mathcal{V}(\xi)$  for the expected value of  $\xi$  over all viable pricing models is taken to be the lower bound for the price of  $\xi$ , then one may ask whether this lower bound coincides with the upper bound given by the superhedging price given as  $\Phi(\xi)$ , i.e., one seeks sufficient conditions for the equality

$$\mathcal{V}(\xi) = \Phi(\xi), \quad \forall \xi \in \mathcal{D}(\mathcal{F}_T).$$

In the case of an increasing sublinear  $\Phi$ , it turns out that the necessary and sufficient condition for the two values to be equal for  $\mathcal{D}(\mathcal{F}_T) = \mathbb{C}_b(\mathbb{D})$  is the lower  $\beta_0$ -semicontinuity of  $\Phi$  on  $\mathbb{C}_b(\mathbb{D})$ . The question whether the set of viable market models is weak\* compact then arises naturally. Indeed, the weak\* compactness allows extending the duality correspondence immediately to  $\mathcal{D}(\mathcal{F}_T) = \mathbb{U}_b(\mathbb{D})$  and is also a sufficient condition for the duality to hold on  $\mathcal{D}(\mathcal{F}_T) = \mathbb{B}_b(\mathbb{D})$ , under another continuity assumption on  $\Phi$ , namely the upper  $\sigma$ -order semicontinuity of  $\Phi$  on  $\mathbb{B}_b(\mathbb{D})$ , provided that the underlying space  $\mathbb{D}$  is perfectly normal ( $T_6$ ). These extension criteria are classical in measure transport; consult e.g. the seminal work by Strassen [52].

Although no probabilistic assumption on the stock dynamics is made a priori, it is reasonable to restrict to the class of semimartingale measures, which form the largest class of stock price models for which it is impossible in a frictionless market to make unbounded profits (or losses) by selling and buying the stock in a non-anticipative manner. Indeed, this is the statement of the Bichteler–Dellacherie–Mokobodzki theorem. Further, if one allows non-anticipative trading of the underlying asset without transaction costs and no interest rate, then all viable pricing models are martingale measures on the canonical space of càdlàg paths satisfying an additional half-space constraint posed by the prices of statistically traded European options. Indeed, if there are static positions available on the market, they translate to half-space constraints on the viable martingale measures. This is the so-called martingale optimal transport problem studied e.g. in the aforementioned works of [3, 15, 24, 12].

### 3.2 Main results

The following Theorem 3.2 is our main result that together with its corollaries and an auxiliary lemma provides an easy method of constructing weak\* sets of semimartin-

gale measures. The statement regarding sequential compactness in Theorem 3.2 refines the classical results of Meyer and Zheng [44], Stricker [55] and Jakubowski [31] for semimartingale measures, i.e., for semimartingales on the canonical space. The statement about (non-sequential) compactness is, to the best of our knowledge, a new result.

The proofs are postponed to Sect. 4.3.

**Theorem 3.2** *Let  $\mathcal{S}$  be a family of semimartingale measures satisfying the condition (UT). Under Assumption 3.1, the set  $[\mathcal{S}]_{\text{seq}}$  is a weak\* compact and sequentially weak\* compact set of semimartingale measures.*

**Corollary 3.3** *Let  $\mathcal{Q}$  be a family of quasimartingale measures satisfying the condition (UB). Under Assumption 3.1, the set  $[\mathcal{Q}]_{\text{seq}}$  is a weak\* compact and sequentially weak\* compact set of quasimartingale measures.*

**Corollary 3.4** *Let  $\mathcal{M}$  be a set of supermartingale measures satisfying the condition (UI). Under Assumption 3.1, the set  $[\mathcal{M}]_{\text{seq}}$  is a weak\* compact and sequentially weak\* compact set of supermartingale measures.*

By combining one or both of the assertions of the following Lemma 3.5 with Theorem 3.2, or one of its corollaries, one obtains compact sets of continuous and Markov semimartingales, quasimartingales and supermartingales.

**Lemma 3.5** *Let  $\mathcal{P}$  be a family of probability measures on the Skorokhod space. Under Assumption 3.1, we have the following:*

- (a) *If the set  $\mathcal{P}$  is  $C$ -tight, then the set  $[\mathcal{P}]_{\text{seq}}$  consists of continuous processes.*
- (b) *If each measure in the set  $\mathcal{P}$  is Lipschitz–Markov, then the set  $[\mathcal{P}]_{\text{seq}}$  consists of Lipschitz–Markov processes.*

### 3.3 Examples

The following Example 3.6, essentially an observation made by Guo et al. [24, Lemma 3.7], was our original motivation to study weak\* compactness in the present setting. In Example 3.6, we allow an infinite time horizon, i.e., the index set  $I = [0, T]$  for some  $T \in (0, \infty]$ .

*Example 3.6* Let  $\mathcal{M}^u$  be the family of uniformly integrable, hence  $L^1$ -bounded martingale measures and  $\mathcal{P}$  a weak\* compact subset of  $\mathbb{P}(\mathbb{R}^d)$ . Then the set

$$\mathcal{M}_{\mathcal{P}}^u = \{Q \in \mathcal{M}^u : Q \circ X_T^{-1} \in \mathcal{P}\}$$

is weak\* compact and sequentially weak\* compact.

*Proof* We adapt the proof of [24, Lemma 3.7]. For  $a > 0$ , we have

$$E_Q[|X_T| \mathbb{1}_{\{|X_T| \geq a\}}] \leq 2E_Q[(|X_T| - a/2)^+] \leq 2E_Q[(|X_T| - a/2)^+]$$

uniformly over  $(t, Q) \in I \times \mathcal{M}_{\mathcal{P}}^u$ , and

$$E_Q[(H \bullet X)_t] = 0$$

for every elementary predictable  $|H| \leq 1$ , every  $t \in I$  and every  $Q \in \mathcal{M}^u$ . Thus by the general form of Prokhorov's theorem, see e.g. [8, Theorem 8.6.2], the family  $\mathcal{M}_{\mathcal{P}}^u$  satisfies the condition (UI); cf. [29, Lemma IX.1.11]. By Example 5.10 (b) below, the evaluation mapping is (sequentially) continuous at the terminal time; so we have  $\mathcal{M}_{\mathcal{P}}^u = [\mathcal{M}_{\mathcal{P}}^u]_{\text{seq}}$ . A measure  $Q$  is a martingale measure for  $X$  on  $\mathbb{D}(I; \mathbb{R}^d)$  if and only if  $Q$  is a supermartingale measure for  $X^i$  and  $-X^i$  for every  $i \leq d$ ; so by Corollary 3.4, the set  $\mathcal{M}_{\mathcal{P}}^u$  is weak\* compact and sequentially weak\* compact.  $\square$

*Example 3.7* Let  $\mathcal{M}^p$  denote the family of  $L^p$ -bounded martingale measures. Then the sets

$$\mathcal{M}_r^p := \{Q \in \mathcal{M}^p : \|X\|_{L^{p,\infty}(Q)} \leq r\}, \quad r > 0, \quad (3.1)$$

are weak\* compact and sequentially weak\* compact for  $1 < p < \infty$ .

*Proof* The increasing continuous function  $y \mapsto y^p$  composed with the lower semicontinuous function  $y = \|\omega\|_{\infty}$  is lower semicontinuous, see Lemma A.12, and non-negative. So by [8, Proposition 8.9.8], the functional  $\|X\|_{L^{p,\infty}(Q)}$  is lower semicontinuous in the weak\* topology. Thus the set  $\mathcal{M}_r^p$  is weak\* closed for  $r > 0$  and  $p > 1$ . The  $L^p$ -boundedness for  $p > 1$  implies that the set  $\mathcal{M}_r^p$  satisfies the condition (UI) for  $r > 0$  and  $p > 1$  and hence is weak\* compact and sequentially weak\* compact; cf. Example 3.6.  $\square$

Assume that a probability measure  $Q$  is fixed and  $p > 1$ . Then the Hardy space of  $L^p(Q)$ -bounded (equivalence classes of indistinguishable) càdlàg martingales,  $\mathcal{M}^p(Q) := \mathcal{M}^p(\mathbb{D}(I; \mathbb{R}), \mathcal{F}_T, (\mathcal{F}_t)_{t \in I}, Q)$ , can be identified with the Lebesgue space  $L^p(Q) := L^p(\mathbb{D}(I; \mathbb{R}), \mathcal{F}_T, Q)$ . Indeed, there exists a linear one-to-one correspondence between the (uniformly integrable)  $L^p(Q)$ -bounded martingales on  $[0, T]$  and the random variables  $X_T$  of  $L^p(Q)$ ,  $p > 1$ , as each  $X_T \in L^p(Q)$  defines a  $Q$ -a.s. unique càdlàg  $L^p(Q)$ -bounded martingale on  $[0, T]$  via

$$X_t := E_Q[X_T | \mathcal{F}_t], \quad t \in [0, T],$$

and conversely, each such càdlàg martingale has a  $Q$ -a.s. unique terminal value  $X_T \in L^p(Q)$ . By Doob's maximal  $L^p$ -inequality, the  $L^{p,\infty}(Q)$ -norm on  $\mathcal{M}^p(Q)$  and the  $L^p(Q)$ -norm on  $L^p(Q)$  are equivalent for  $p > 1$ . In the case  $p = 1$ , there is no one-to-one correspondence, but we only have  $\mathcal{M}^1(Q) \subseteq L^1(Q)$  for  $I = [0, \infty)$ . Further discussion on this one-to-one correspondence can be found e.g. in [14, VII.64]. We mention [54, Theorem 3], where this one-to-one correspondence is used in the construction of semimartingale decompositions for quasimartingales, which forms a central step in the classical proofs for the Bichteler–Dellacherie–Mokobodzki theorem.

In the case of a fixed probability measure  $Q$ , due to the one-to-one correspondence, the weak\* compactness follows alternatively from classical results for Banach

spaces  $L^p(Q)$ ,  $p > 1$ . For  $p > 1$ , the space  $L^p(Q)$  is a reflexive Banach space; so the (sequential) weak\* compactness of the sets (3.1) follows from the Banach–Alaoglu theorem in conjunction with the Eberlein–Šmulian theorem; see [55]. The Dunford–Pettis theorem states that a uniformly integrable subset of the non-reflexive Banach space  $L^1(Q)$  is relatively sequentially compact in the weak topology; but the random variables in  $L^1(Q)$  are not in one-to-one correspondence with either the family of  $L^{1,\infty}(Q)$ -bounded or the family of  $L^1(Q)$ -bounded martingales for  $I = [0, \infty)$ .

*Example 3.8* Let  $\mathcal{H}^p$  denote the family of  $\mathcal{H}^p$ -semimartingale measures. Then the sets

$$\mathcal{S}_r^p := \{Q \in \mathcal{H}^p : \|X\|_{L^{p,\infty}(Q)} + \|X\|_{\mathcal{E}^p(Q)} \leq r\}, \quad r > 0,$$

are weak\* compact and sequentially weak\* compact for  $1 \leq p < \infty$ .

*Proof* The sets  $\mathcal{S}_r^p$ ,  $r > 0$ ,  $p \geq 1$ , satisfy the condition (UB); so by Corollary 3.3, the sets  $[\mathcal{S}_r^p]_{\text{seq}}$  are weak\* compact and sequentially weak\* compact sets of quasimartingales. Moreover, for any sequence  $(Q_n)_{n \in \mathbb{N}}$  in  $\mathcal{S}_r^p$  converging in the weak\* topology to some  $Q$ , we have

$$\|X\|_{L^{p,\infty}(Q)} \leq \liminf_{n \rightarrow \infty} \|X\|_{L^{p,\infty}(Q^n)} < \infty \tag{3.2}$$

for  $p \geq 1$ ; cf. Example 3.7. Thus we have  $[\mathcal{S}_r^p]_{\text{seq}} \subseteq \mathcal{H}^1$ ; cf. [14, VII.98]. To show that  $\mathcal{S}_r^p = [\mathcal{S}_r^p]_{\text{seq}}$  and  $[\mathcal{S}_r^p]_{\text{seq}} \subseteq \mathcal{H}^p$ , we introduce an auxiliary class  $\mathcal{A}$  of smooth elementary integrands of the form

$$A^i = \sum_{j=1}^k A_{t_{j-1}^i}^i \varphi_{t_{j-1}^i, t_j^i}, \quad i \leq d,$$

where  $k \in \mathbb{N}$ ,  $0 = t_0^i \leq t_1^i \leq \dots \leq t_k^i$  in  $I$ , each  $A_{t_j^i}^i$  is a continuous  $\mathcal{F}_{t_j^i-}$ -measurable function satisfying  $|A_{t_j^i}^i| \leq 1$ , and each  $\varphi_{t_{j-1}^i, t_j^i}$  is a smooth function on  $I$  vanishing outside  $(t_{j-1}^i, t_j^i + \varepsilon_j^i)$  for some  $\varepsilon_j^i \in (t_j^i, t_{j+1}^i)$  and satisfying  $|\varphi_{t_{j-1}^i, t_j^i}| \leq 1$ ; we allow non-zero  $\varphi_{t_{k-1}^i, t_k^i}(t_k^i) \in [-1, 1]$  for  $t_k^i = T$ , cf. (A.2). Now let  $Q \in [\mathcal{S}_r^p]_{\text{seq}}$  and assume we are given an elementary predictable integrand  $H = (H^i)_{i=1}^d$ , i.e., an element of  $\mathcal{E}(Q)$ , see (2.1), such that each  $H_{t_j^i}^i$  in (2.1) is  $\mathcal{F}_{t_j^i-}^0$ -measurable. By [13, IV.69 (c)], the domain of  $\mathcal{F}_{t_j^i-}^0$ -measurable functions is homeomorphic to a closed subset of  $\mathbb{D}(I; \mathbb{R}^d)$ ; cf. Corollary A.11. Thus by Lusin’s theorem in conjunction with Tietze’s extension theorem, for every  $\mathcal{F}_{t_j^i-}^0$ -measurable  $|H_{t_j^i}^i| \leq 1$ , there exists a sequence of continuous  $\mathcal{F}_{t_j^i-}^0$ -measurable functions  $(A_{t_j^i}^{i,n})_{n \in \mathbb{N}}$  with  $|A_{t_j^i}^{i,n}| \leq 1$  and such that  $A_{t_j^i}^{i,n} \rightarrow H_{t_j^i}^i$   $Q$ -a.s. as  $n \rightarrow \infty$ ; see e.g. [17] and [16, 2.1.8.]. Moreover, càglàd step functions can be approximated from the right with smooth functions, and smooth

functions can be approximated from the right with càglàd step functions; so there exists a sequence  $(A^n)_{n \in \mathbb{N}}$ ,  $A^n = (A^{i,n})_{i=1}^d$ , of elements of  $\mathcal{A}$  such that  $\|A^n\|_{\mathbb{V}} \leq \|H\|_{\mathbb{V}}$   $Q$ -a.s. for all  $n \in \mathbb{N}$  and  $(A^n \bullet X)_T \rightarrow (H \bullet X)_T$   $Q$ -a.s. as  $n \rightarrow \infty$ . Integrating by parts, we get

$$|(A^n \bullet X)_T| \leq (|X_T A_T^n| + \|X\|_{\infty} \|A^n\|_{\mathbb{V}}) \leq c \|X\|_{\infty}, \quad A_0^n = 0, n \in \mathbb{N}, \quad (3.3)$$

where  $c := 2\|H\|_{\mathbb{V}} < \infty$   $Q$ -a.s. and  $\|X\|_{\infty} \in L^p(Q)$  by (3.2). Thus by dominated convergence, for any  $Q \in [\mathcal{S}_r^p]_{\text{seq}}$ , we have

$$\lim_{n \rightarrow \infty} \|(A^n \bullet X)_T\|_{L^p(Q)} = \|(H \bullet X)_T\|_{L^p(Q)}.$$

The elements of  $\mathcal{A}$  can similarly be approximated with elements of  $\mathcal{E}(Q)$ . Moreover, due to the uniform bound (3.3), for any sequence of integrands bounded in total variation, the right-continuity of  $X = (X^1, X^2, \dots, X^d)$  allows to relax the  $\mathcal{F}_{t_j^-}^0$ -measurability of the random variables  $H_{t_j}^i$  to  $\mathcal{F}_{t_j}^0$ -measurability, and further to  $\mathcal{F}_{t_j^+}^0$ -measurability; cf. (4.17) below. Thus for any  $Q \in [\mathcal{S}_r^p]_{\text{seq}}$ , we have

$$\|X\|_{\mathcal{E}^p(Q)} = \|X\|_{\mathcal{A}^p(Q)} := \sup_{A \in \mathcal{A}} \|(A \bullet X)_T\|_{L^p(Q)}.$$

Now since each  $A^i$  of  $A = (A^i)_{i=1}^d$  in  $\mathcal{A}$  is continuously differentiable in  $t$  for every  $\omega \in \mathbb{D}(I; \mathbb{R}^d)$ , the function  $|(A \bullet X)_T|^p$  is continuous, see (A.2), and nonnegative. So by Proposition 4.5 and [8, Proposition 8.9.8], the functional  $\|(A \bullet X)_T\|_{L^p(Q)}$  is weak\* lower semicontinuous on  $\mathcal{S}_r^p$  for every  $A \in \mathcal{A}$ . Consequently, the functional  $\|X\|_{\mathcal{A}^p(Q)}$  is weak\* lower semicontinuous on  $\mathcal{S}_r^p$ , which in conjunction with the weak\* lower semicontinuity (3.2) of the functional  $\|X\|_{L^{p,\infty}(Q)}$  yields the weak\* closedness of the sets  $\mathcal{S}_r^p$  in  $\mathcal{E}^p$ . Indeed, for any  $r > 0$  and  $p \geq 1$ , for any sequence  $(Q_n)_{n \in \mathbb{N}}$  in  $\mathcal{S}_r^p$  converging in the weak\* topology to some  $Q$ , we have

$$\begin{aligned} \|X\|_{L^{p,\infty}(Q)} + \|X\|_{\mathcal{E}^p(Q)} &= \|X\|_{L^{p,\infty}(Q)} + \|X\|_{\mathcal{A}^p(Q)} \\ &\leq \liminf_{n \rightarrow \infty} (\|X\|_{L^{p,\infty}(Q_n)} + \|X\|_{\mathcal{A}^p(Q_n)}) \\ &= \liminf_{n \rightarrow \infty} (\|X\|_{L^{p,\infty}(Q_n)} + \|X\|_{\mathcal{E}^p(Q_n)}) \leq r, \end{aligned}$$

i.e.,  $Q \in \mathcal{S}_r^p$ . Thus we have  $\mathcal{S}_r^p = [\mathcal{S}_r^p]_{\text{seq}} \subseteq \mathcal{H}^p$ , i.e., the sets  $\mathcal{S}_r^p$  are weak\* compact for  $r > 0$  and  $p \geq 1$ . Every element in  $\mathcal{S}_r^p$  is indeed an  $\mathcal{H}^p$ -semimartingale measure; cf. (3.4) and (3.5) below.  $\square$

The pseudonorm in Example 3.8 given by the sum of the  $L^{p,\infty}$ -norm and the Emery pseudonorm

$$\|X\|_{\mathcal{E}^p(Q)} := \sup_{H \in \mathcal{E}(Q)} \|(H \bullet X)_T\|_{L^p(Q)}, \quad p \geq 1, \quad (3.4)$$

is equivalent to the (maximal)  $\mathcal{E}^p$ -norm

$$\|X\|_{\mathcal{E}^p(Q)} := \|\|M\|_{\infty} + \|A\|_{\mathbb{V}}\|_{L^p(Q)}, \quad p \geq 1, \quad (3.5)$$



where  $X = M + A$ ,  $A_0 = 0$ , denotes the canonical semimartingale decomposition of  $X$  under  $Q$ ; see [14, Theorem VII.104 and subsequent remark]. Fix a probability measure  $Q$  and  $p > 1$ . Then the Hardy space  $\mathcal{H}^p(Q) := \mathcal{H}^p(\mathbb{D}(I; \mathbb{R}), \mathcal{F}_T, (\mathcal{F}_t)_{t \in I}, Q)$  of  $\mathcal{H}^p(Q)$ -bounded (equivalence classes of indistinguishable) semimartingales is a Banach space; see [25, Problem 10.10]. For martingales in particular, the norm  $\|X\|_{\mathcal{E}^p(Q)}$  is equivalent to the norm  $\|X\|_{L^{p,\infty}(Q)}$ , and as mentioned in the context of Example 3.7, there is an analogous Banach pairing

$$(\mathcal{M}^p(Q))' = (L^p(Q))' = L^q(Q) = \mathcal{M}^q(Q)$$

for  $p, q > 1$  with  $1/p + 1/q = 1$ ; see [14, beginning of Sect. VII.3] and [25, Problem 10.29].

In contrast to the classical Banach space pairing, the pairing of the Riesz representation theorem makes it straightforward to construct compact sets of continuous and Markov processes by invoking the classical stability criteria for almost sure (Hölder-)continuity and the (Lipschitz-)Markov property.

*Example 3.9* Let  $\mathcal{S}$  be a set of semimartingale measures satisfying the condition (UT).

(a) Assume that there exist constants  $a, b, c > 0$  such that

$$\sup_{Q \in \mathcal{S}} E_Q[|X_t - X_s|^a] \leq b|s - t|^{1+c}, \quad \forall s, t \in I. \tag{3.6}$$

Then the set  $[\mathcal{S}]_{\text{seq}}$  is a weak\* compact set of continuous semimartingales.

(b) Let  $Q$  be the standard Wiener measure on the Skorokhod space  $\mathbb{D}(\mathbb{R}_+; \mathbb{R})$ . Assume that for every  $Q^\alpha$  in  $\mathcal{P}$ , there exists a function  $\sigma^\alpha : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$  that is continuous, continuously differentiable in the first component, and that  $Q^\alpha$  is the law of a (unique strong) solution  $X^\alpha$  of

$$X_0^\alpha \in \mathbb{R}, \quad X_t^\alpha = X_0^\alpha + \int_0^t \sigma^\alpha(u, X_u^\alpha) dX_u, \quad \forall t \geq 0, Q\text{-a.s.} \tag{3.7}$$

Then the set  $[\mathcal{S}]_{\text{seq}}$  is a weak\* compact set of Markov semimartingales.

Analogous assertions to (a) and (b) are true for the sequential closures  $[Q]_{\text{seq}}$  and  $[\mathcal{M}]_{\text{seq}}$  of a set  $\mathcal{Q}$  of quasimartingale measures and a set  $\mathcal{M}$  of supermartingale measures satisfying the condition (UB) and (UI), respectively.

*Proof* (a) The Kolmogorov continuity criterion (3.6) is a sufficient criterion for the  $C$ -tightness of  $\mathcal{S}$ ; see e.g. [48, XII.1.8]. Thus the statement follows from Theorem 3.2 in conjunction with Lemma 3.5 (a).

(b) We adapt the argument of [27, Proposition 5]. Under the given assumptions, an equation of the form (3.7) admits a unique strong solution. For fixed  $\alpha$ ,  $s \geq 0$  and  $x \in \mathbb{R}$ , let  $X^{\alpha,s,x} = (X_t^{\alpha,s,x})_{t \geq s}$  denote the solution to

$$X_s^{\alpha,s,x} = x, \quad X_t^{\alpha,s,x} = x + \int_s^t \sigma^\alpha(u, X_u^{\alpha,u,x}) dX_u, \quad \forall t \geq s, Q\text{-a.s.}$$

and define the process  $M^{\alpha,s,x} = (M_t^{\alpha,s,x})_{t \geq s}$  by  $M_t^{\alpha,s,x} := \frac{\partial}{\partial x} X_t^{\alpha,s,x}$ . Then for every  $t \geq s$ , we have

$$M_t^{\alpha,s,x} = \exp \left( \int_s^t \sigma_x^\alpha(u, X_u^{\alpha,s,x}) dX_u - \frac{1}{2} \int_s^t (\sigma_x^\alpha)^2(u, X_u^{\alpha,s,x}) du \right) \quad Q\text{-a.s.},$$

where  $\sigma_x^\alpha := \frac{\partial}{\partial x} \sigma^\alpha$ . The process  $M^{\alpha,s,x}$  is a nonnegative local martingale, so that  $M_t^{\alpha,s,x} \geq 0$   $Q$ -a.s. and  $E_Q[M_t^{\alpha,s,x}] \leq 1$  for every  $t \geq s$ . Assume now that  $g$  is a bounded continuously differentiable function with  $|g'(x)| \leq 1$  for every  $x \in \mathbb{R}$ , and fix  $t > s \geq 0$ . Then define the function  $f : \mathbb{R} \rightarrow \mathbb{R}$  by

$$f(x) := E_Q[g(X_t^{\alpha,s,x})], \quad x \in \mathbb{R}.$$

The function  $f$  is bounded and continuously differentiable with

$$|f'(x)| = E_Q[g'(X_t^{\alpha,s,x}) M_t^{\alpha,s,x}] \leq 1, \quad x \in \mathbb{R}.$$

Therefore  $X^\alpha = (X_t^\alpha)_{t \geq 0}$  satisfies the Lipschitz–Markov property on the space  $(\mathbb{D}(\mathbb{R}_+; \mathbb{R}), \mathcal{F}_\infty, (\mathcal{F}_t)_{t \geq 0}, Q)$  for every  $\alpha$ , i.e., the law  $Q^\alpha$  of  $X^\alpha$  is a Lipschitz–Markov measure on the space  $(\mathbb{D}(\mathbb{R}_+; \mathbb{R}), \mathcal{F}_\infty, (\mathcal{F}_t)_{t \geq 0})$  for every  $\alpha$ . Thus by Theorem 3.2 in conjunction with Lemma 3.5 (b), every element of  $[\mathcal{P}]_{\text{seq}}$  is a (Lipschitz–) Markov semimartingale measure.

The statements for quasimartingale measures and supermartingale measures are obtained by replacing Theorem 3.2 above with Corollary 3.3 and 3.4, respectively.  $\square$

## 4 Auxiliary results and the proofs for Sect. 3.2

The purpose of this section is to provide the proofs for Theorem 3.2 and its corollaries that we omitted in Sect. 3.2 and the required auxiliary results leading to the proofs. In Sect. 4.1, we establish three basic results for the weak\* topology deployed in the proof of Theorem 3.2. The required stability and tightness results for the weak\* topology are covered in Sect. 4.2. Finally, the proofs for the results of Sect. 3.2 are provided in Sect. 4.3.

In Sect. 4.1, in addition to that the underlying topological space  $X$  is regular and Souslin, we assume that it has the following separation property.

*Property 4.1* There exists a countable family of real-valued continuous functions  $f_k$ ,  $k \in \mathbb{N}$ , such that for all  $x, y \in X$ , we have

$$f_k(x) = f_k(y), \forall k \in \mathbb{N} \implies x = y. \quad (4.1)$$

*Remark 4.2* A topological space satisfying Property 4.1 is submetrisable, i.e., there exists a weaker topology that is metrisable. Indeed, such a topology is given e.g. by the metric

$$d(x, y) := \sum_{k=1}^{\infty} 2^{-k} \frac{|f_k(x) - f_k(y)|}{1 + |f_k(x) - f_k(y)|}, \quad x, y \in X, \quad (4.2)$$

where the  $f_k$  are given by (4.1). The author would like to thank Professor Jakubowski for pointing out this fact.

Property 4.1 was extensively studied in Jakubowski [30]; see also Sect. 5.2. In particular, we note that a Skorokhod space satisfying Assumption 3.1 has this property. Indeed, for the Skorokhod space  $\mathbb{D}(I; \mathbb{R}^d)$ , a countable family of continuous functions possessing the property (4.1) is given e.g. by the family of functions

$$\omega \mapsto \frac{1}{r} \int_q^{q+r} \omega^i(t) dt \quad \text{and} \quad \omega \mapsto \omega^i(T), \quad \text{for } I = [0, T],$$

where  $q$  and  $q + r$  run over the rationals in  $I$  and  $i$  over the spatial dimensions  $1, \dots, d$ ; cf. (A.2) and (A.3). Therefore, the results of this section are true for a Skorokhod space satisfying Assumption 3.1. Indeed, the Souslin property is verified in Proposition 5.12.

### 4.1 Weak\* topology

The results of this section are established under the assumption that **the space  $\mathbb{D}$  is a regular Souslin space satisfying Property 4.1**. Under this assumption, we obtain a stronger separation axiom than the required  $T_{3\frac{1}{2}}$ ; cf. Sect. 2.3. Indeed, combining the fact that  $\mathbb{D}$  is regular ( $T_3$ ) with the fact that it is a Souslin space, it follows from a result of Fernique [18, Proposition I.6.1] that *the space  $\mathbb{D}$  is perfectly normal ( $T_6$ )*.

Recall that the families  $\mathbb{M}_t(\mathbb{D})$ ,  $\mathbb{M}_\tau(\mathbb{D})$  and  $\mathbb{M}_\sigma(\mathbb{D})$  are defined in Sect. 1.

**Proposition 4.3** *The following characterise the dual space in the pairing given in Lemma 2.4.*

(a) *The dual of  $\mathbb{C}_b(\mathbb{D})$  under the strict topology  $\beta_0$  is isomorphic to the family of measures*

$$\mathbb{M}_t(\mathbb{D}) = \mathbb{M}_\tau(\mathbb{D}) = \mathbb{M}_\sigma(\mathbb{D}), \tag{4.3}$$

where the Skorokhod space  $\mathbb{D}$  is endowed with the canonical  $\sigma$ -algebra.

(b) *The assertion (a) remains true if the canonical  $\sigma$ -algebra is augmented with the universal nullsets.*

*Proof* (a) Every Lusin space is a Radon space; see e.g. [49, Theorem I.3.9]. Thus we have  $\mathbb{M}_\sigma(\mathbb{D}) \subseteq \mathbb{M}_t(\mathbb{D})$ . The equality (4.3) follows from the fact that the inclusions  $\mathbb{M}_t(\mathbb{D}) \subseteq \mathbb{M}_\tau(\mathbb{D})$  and  $\mathbb{M}_\tau(\mathbb{D}) \subseteq \mathbb{M}_\sigma(\mathbb{D})$  are true for an arbitrary topological space; see [8, Proposition 7.2.2]. By [35, Theorem 7.6.3], the dual space of  $\mathbb{C}_b(\mathbb{D})$  under the strict topology  $\beta_0$  is isomorphic to the family  $\mathbb{M}_t(\mathbb{D})$ , and hence isomorphic to the family (4.3).

(b) Let  $\widehat{\mathcal{B}}(\mathbb{D})$  denote the universal completion of the Borel  $\sigma$ -algebra  $\mathcal{B}(\mathbb{D})$ . For every  $\mu \in \mathbb{P}(\mathcal{B}(\mathbb{D}))$ , there exists a unique  $\widehat{\mu} \in \mathbb{P}(\widehat{\mathcal{B}}(\mathbb{D}))$  such that

$$\int f d\mu = \int f d\widehat{\mu}, \quad \forall f \in \mathbb{C}_b(\mathbb{D}).$$

Since any measure of finite variation is a linear combination of two probability measures, it suffices to observe that the mapping  $\mu \mapsto \widehat{\mu}$  is a bijection; see e.g. [13, Remark 32 (c)(1)]. The statement then follows from (a).  $\square$

*Remark 4.4* It also follows that every measure in the class (4.3) is perfect; see e.g. [8, Theorem 7.5.10 (i)].

**We use the equality (4.3) without mentioning it when we apply the results from the book of Bogachev [8].**

#### 4.1.1 The Eberlein–Šmulian properties

In this section, we show that the nonnegative orthant  $\mathbb{M}_+(\mathbb{D})$  endowed with the weak\* topology is *angelic*. In angelic spaces, the properties of compactness and sequential compactness coincide. In general, one does not imply the other; see e.g. [9, Example 3.4.1].

**Proposition 4.5** *The space  $\mathbb{M}_+(\mathbb{D})$  of nonnegative measures endowed with the weak\* topology is angelic. In particular, for any subset  $M \subseteq \mathbb{M}_+(\mathbb{D})$ , the following are equivalent:*

- (i) *Any sequence in  $M$  has a weak\* convergent subsequence in  $\mathbb{M}_+(\mathbb{D})$ .*
- (ii) *The weak\* closure of  $M$  is weak\* compact in  $\mathbb{M}(\mathbb{D})$ .*

*Moreover, under these conditions, the weak\* closure of  $M$  is metrisable.*

*Proof* Because the underlying topological space  $\mathbb{D}$  is a regular Souslin space, it admits a continuous injective mapping to a metric space; see [19, Theorem 2.25 (i)]. It is also known that if a regular space can be continuously injected into an angelic space, then this regular space is also angelic; see [20, Theorem 3.3]. Since the weak\* topology on the space  $\mathbb{M}_+(\mathbb{D})$  of nonnegative measures is metrisable for a metrisable topology on the underlying space  $\mathbb{D}$ , see [8, Theorem 8.3.2], and metric spaces are angelic, the space  $\mathbb{M}_+(\mathbb{D})$  endowed with the weak\* topology is angelic under our assumption that  $\mathbb{D}$  is a regular Souslin space. By [8, Theorem 8.10.4], the weak\* closure of a subset  $M$  of  $\mathbb{M}_+(\mathbb{D})$  satisfying (i) or (ii) is a compact metrisable subspace of  $\mathbb{M}(\mathbb{D})$ ; so (i) and (ii) are indeed equivalent for  $M$ .  $\square$

It is immediate from Proposition 4.5 that the properties of weak\* compactness and sequential weak\* compactness are equivalent for subsets of  $\mathbb{M}_+(\mathbb{D})$ . In fact, a stronger statement is true. In angelic spaces, the closure of a relatively compact set is completely exhausted by the limits of sequences of points in this set.

**Corollary 4.6** *Assume that  $M$  is a subset of  $\mathbb{M}_+(\mathbb{D})$  that satisfies the equivalent conditions of Proposition 4.5. Then the sequential weak\* closure of  $M$  in  $\mathbb{M}_+(\mathbb{D})$ , i.e., the set*

$$[M]_{\text{seq}} = \{\mu \in \mathbb{M}_+(\mathbb{D}) : \exists (\mu_n)_{n \in \mathbb{N}} \subseteq M \text{ such that } \mu^n \rightarrow_{w^*} \mu\},$$

*is weak\* closed.*

*Proof* By Proposition 4.5, the closure of  $M$  endowed with the relative topology is a first countable space. In particular, the space is a Fréchet–Urysohn space. By [16, Theorem 1.6.14], the sequential closure  $[M]_{\text{seq}}$  coincides with the closure of  $M$ .  $\square$

Various properties of a family of laws of stochastic processes, such as (UT), (UB), (UI) and  $C$ -tightness, that guarantee  $\mathcal{S}$ -tightness and stability of a family of processes as studied in Sect. 4.2.1 are not preserved in taking the weak\* closure; so the previous results are crucial for constructing weak\* compact sets of stochastic processes.

#### 4.1.2 Prokhorov’s theorem

We say that a subset  $M$  of  $\mathbb{M}(\mathbb{D})$  is  $\beta_0$ -equicontinuous if the corresponding family of linear functionals

$$\left\{ f \mapsto u_\mu(f) := \int f d\mu : \mu \in M \right\}$$

is equicontinuous in the  $\beta_0$ -topology on  $\mathbb{C}_b(\mathbb{D})$ , i.e., if for every  $\varepsilon > 0$ , there exists a  $\beta_0$ -neighbourhood  $V$  in  $\mathbb{C}_b(\mathbb{D})$  such that  $|u_\mu(f)| < \varepsilon$  for all  $(f, \mu) \in V \times M$ .

A measure  $\mu \in \mathbb{M}(\mathbb{D})$  is called *tight* if there exists an exhausting net of compact sets  $(K^\varepsilon)_{\varepsilon>0}$  for  $\mu$ , i.e.,  $|\mu|(\mathbb{D} \setminus K^\varepsilon) < \varepsilon$  for every  $\varepsilon > 0$ , where  $|\mu|$  is the total variation of  $\mu$ . A subset  $M$  of  $\mathbb{M}(\mathbb{D})$  is called *uniformly tight* if there exists a net of compact sets  $(K^\varepsilon)_{\varepsilon>0}$  which is uniformly exhausting for the total variation of  $M$ , i.e.,  $\sup_{\mu \in M} |\mu|(\mathbb{D} \setminus K^\varepsilon) < \varepsilon$  for every  $\varepsilon > 0$ .

**Proposition 4.7** *A subset  $M$  of  $\mathbb{M}(\mathbb{D})$  is  $\beta_0$ -equicontinuous if and only if it is bounded in total variation and uniformly tight. In addition, we have the following:*

(a) *If  $M$  is a  $\beta_0$ -equicontinuous subset of  $\mathbb{M}(\mathbb{D})$ , then  $M$  is relatively compact and relatively sequentially compact in the weak\* topology.*

(b) *If  $(\mu_n)_{n \in \mathbb{N}}$  is a uniformly tight sequence in  $\mathbb{M}(\mathbb{D})$  converging in the weak\* topology to  $\mu \in \mathbb{M}(\mathbb{D})$ , then for any  $f \in \mathbb{C}(\mathbb{D})$  satisfying*

$$\lim_{c \rightarrow \infty} \sup_{n \in \mathbb{N}} \int |f| \mathbb{1}_{\{|f| \geq c\}} d\mu_n = 0,$$

*we have*

$$\int f d\mu_n \longrightarrow \int f d\mu \quad \text{as } n \longrightarrow \infty.$$

*Proof* As the underlying topological space is completely regular, the characterisation follows directly from [50, Theorem 5.1]. The compact subsets of a completely regular Souslin space are metrisable, cf. Proposition 4.5 and [8, Lemma 8.9.2]. In conjunction with the fact the space is completely regular, this verifies the assumptions for both the sequential and the nonsequential Prokhorov theorem and hence gives (a); see [8, Theorem 8.6.7]. The convergence criterion in (b) is similarly a direct consequence of the fact that the underlying space is completely regular; see [7, Lemma 3.8.7].  $\square$

The characterisation of relative compactness in terms of  $\beta_0$ -equicontinuity yields also a criterion for compactness of closures (of convex (circled) hulls).

**Corollary 4.8** *The closed convex circled hull of a  $\beta_0$ -equicontinuous subset of  $\mathbb{M}(\mathbb{D})$  is  $\beta_0$ -equicontinuous, weak\* compact and sequentially weak\* compact. In particular, the closure and the closed convex hull of a  $\beta_0$ -equicontinuous set are weak\* compact and sequentially weak\* compact.*

*Proof* The weak\* compactness of the closed convex circled hull of an equicontinuous set follows from [37, 18.5]. Closure and closed convex hull are closed subsets of the closed convex circled hull, from which the second statement follows. The  $\beta_0$ -equicontinuous sets are bounded in total variation, in particular, from below; so the sequential statements are true by Proposition 4.5.  $\square$

#### 4.1.3 Skorokhod's representation theorem

Jakubowski's fundamental observation was that Property 4.1 yields a subsequential Skorokhod representation theorem.

**Proposition 4.9** *Let  $(Q_n)_{n \in \mathbb{N}}$  be a sequence converging in the weak\* topology to  $Q$  in  $\mathbb{P}(\mathbb{D})$ . Then there exist a subsequence  $(Q_{n_k})_{k \in \mathbb{N}}$ , a probability space  $(\Omega, \mathcal{F}, P)$  and  $\mathbb{D}$ -valued random variables  $(Y_k)_{k \in \mathbb{N}}$  and  $Y$  on  $(\Omega, \mathcal{F}, P)$  such that  $Q_{n_k} = P \circ (Y_k)^{-1}$  for  $k \in \mathbb{N}$ ,  $Q = P \circ Y^{-1}$  and*

$$f(Y_k(\omega)) \longrightarrow f(Y(\omega)), \quad \forall \omega \in \Omega, \forall f \in \mathbb{C}_b(\mathbb{D}). \quad (4.4)$$

*Proof* The Euclidean space endowed with its usual inner product is a Hilbert space; so by [32, Theorem 1], the existence of an a.s. convergent subsequence of  $\mathbb{D}$ -valued random variables  $(Y_k)_{k \in \mathbb{N}}$  as in the assertion follows from Property 4.1. The convergence in [32, Theorem 1] is the pointwise convergence in the topology of the underlying space, which for a sequence in a completely regular space is equivalent to the convergence (4.4); cf. (5.3) below. Moreover, modifying the  $\mathbb{D}$ -valued random variables  $Y_k$  and  $Y$  given by [32, Theorem 1] on a set of measure zero does not affect their weak\* convergence; so their almost sure convergence can be strengthened to pointwise convergence.  $\square$

In particular, by Proposition 4.9, every element of  $\mathbb{P}(\mathbb{D})$  can be regarded as the law of some  $\mathbb{D}$ -valued random variable. Conversely, any such random variable induces a probability measure on  $\mathbb{D}$ .

## 4.2 Stability and tightness

In this section, we cover the required stability and tightness results. We present the required multidimensional infinite-horizon extensions of the stability results of Meyer and Zheng [44], Jakubowski et al. [34] and Lowther [42], for the right-continuous version of the raw canonical filtration.

### 4.2.1 Stability

Under Assumption 3.1, it suffices to establish the required stability results for the Meyer–Zheng topology MZ; see Appendix A.1.3. The required stability results are classical and thoroughly studied in the aforementioned works [44, 34, 31] for scalar-valued processes. We demonstrate that after some slight modifications, they are true in the present setting. We utilise the following multidimensional extension of [44, Theorem 5], provided by Jakubowski’s subsequential Skorokhod representation theorem.

**Lemma 4.10** (Jakubowski) *If  $(Q_n)_{n \in \mathbb{N}}$  is a sequence converging in the weak\* topology to  $Q$  in  $\mathbb{P}(\mathbb{D}(I; \mathbb{R}^d))$ , then there exist a subsequence  $(Q_{n_k})_{k \in \mathbb{N}}$  and a set  $L \subseteq I$  of full Lebesgue measure such that  $T \in L$  if  $I = [0, T]$  and*

$$Q_n \circ X_F^{-1} \xrightarrow{w^*} Q \circ X_F^{-1} \quad \text{as } n \rightarrow \infty \tag{4.5}$$

for every finite subset  $F$  of  $L$ . In particular, there exists a (countable) dense set  $D \subseteq I$  such that  $T \in D$  if  $I = [0, T]$  and (4.5) is true for every finite subset  $F$  of  $D$ .

*Proof* By Proposition 4.9, we can find a subsequence  $(Q_{n_k})_{k \in \mathbb{N}}$  and  $\mathbb{D}$ -valued random variables  $(Y_k)_{k \in \mathbb{N}}$  and  $Y$  on some  $(\Omega, \mathcal{F}, P)$  such that  $Q_{n_k} = P \circ Y_k^{-1}$  for  $k \in \mathbb{N}$ ,  $Q = P \circ Y^{-1}$  and  $Y_k(\omega) \xrightarrow{\text{MZ}} Y(\omega)$  for every  $\omega \in \Omega$  as  $k \rightarrow \infty$ ; since the topology MZ is metrisable, (4.4) is equivalent to  $\xrightarrow{\text{MZ}}$ . By Lemma A.10, there exist a subsequence  $(Y_{k_m})_{m \in \mathbb{N}}$  and a set  $L$  of full Lebesgue such that  $T \in L$  if  $I = [0, T]$  and  $Y_{k_m,t}(\omega) \rightarrow Y_t(\omega)$  for every  $(t, \omega) \in L \times \Omega$  as  $m \rightarrow \infty$ . Hence the finite-dimensional distributions of the process  $(Y_{k_m,t})_{t \in L}$  converge to those of  $(Y_t)_{t \in L}$ . The complement of  $L$  is a  $\lambda$ -nullset; cf. Definition A.8, where the measure  $\lambda$  is given. Thus the set  $L$  contains a (countable) dense set  $D$  such that  $T \in D$  for  $I = [0, T]$ .  $\square$

In Proposition 4.11, we show that the required part of [34, Theorem 2.1], which is an extension of [54, Theorem 2] for a right-continuous canonical filtration, is true on a multidimensional Skorokhod space.

**Proposition 4.11** *Let  $(Q_n)_{n \in \mathbb{N}}$  be a sequence of semimartingale measures satisfying the condition (UT) and converging in the weak\* topology to  $Q$ . Then the weak\* limit  $Q$  is a semimartingale measure.*

*Proof* The proof is essentially a combination of [34, Lemmas 1.1 and 1.3]. By Lemma 4.10, there exist a subsequence  $(Q_{n_k})_{k \in \mathbb{N}}$  and a countable dense set  $D \subseteq I$  such that  $T \in D$  if  $I = [0, T]$  and  $(Q_{n_k})_{k \in \mathbb{N}}$  converges to  $Q$  in finite-dimensional distributions on the set  $D$ . For every finite collection  $t_1 < \dots < t_j$  in  $D$ , let  $\mathcal{A}_{t_1, \dots, t_j}$  denote the family of continuity sets of the marginal law of  $Q$  on  $t_1 < \dots < t_j$ , i.e.,  $\mathcal{A}_{t_1, \dots, t_j}$  consists of Borel sets  $B \in \bigotimes_{i \leq j} \mathcal{B}(\mathbb{R}^d)$  for which  $Q \circ X_{t_1, \dots, t_j}^{-1}[\partial B] = 0$ , where  $\partial B$  denotes the (Euclidean) topological boundary of  $B$  on  $\mathbb{R}^{d \times j}$ . Following [34], we introduce an auxiliary class  $\mathcal{J}(D)$  of integrands, determined by the weak\*

limit  $Q$  and the dense set  $D$ , that are of the form

$$J = \sum_{i=1}^d \sum_{k=1}^{n(i)} J_{t_{k-1}^i}^{t_i} \mathbb{1}_{(t_{k-1}^i, t_k^i]}, \quad n(i) \in \mathbb{N}, i \leq d,$$

where every  $t_0^i < t_1^i < \dots < t_{n(i)}^i$  is a finite collection of elements of  $D$  and every  $J_{t_{k-1}^i}^{t_i}$  is a finite linear combination of indicator functions of the continuity sets of the marginal law of  $Q$  on  $s_1 < \dots < s_j \leq t_{k-1}^i$ ,  $s_1, \dots, s_j \in D$ , embedded in  $\mathbb{D}(I; \mathbb{R}^d)$  and bounded by 1 in absolute value, i.e.,  $|J_{t_{k-1}^i}^{t_i}| \leq 1$  and each  $J_{t_{k-1}^i}^{t_i}$  is of the form

$$J_{t_{k-1}^i}^{t_i} = \sum_{\ell=1}^p \alpha^\ell \mathbb{1}_{A^\ell} \circ X_{s_1, \dots, s_j}^{-1}, \quad s_j \leq t_{k-1}^i, \alpha^\ell \in \mathbb{R}, A^\ell \in \mathcal{A}_{s_1, \dots, s_j},$$

for some elements  $s_1 < \dots < s_j$  of  $D$  and  $j$  and  $p$  finite, for every  $i \leq d$  and every  $k \leq n$ . Now, since  $(Q_{n_k})_{k \in \mathbb{N}}$  is converging to  $Q$  in finite-dimensional distributions on the set  $D$ , by the vectorial Portemanteau lemma, see e.g. [57, Lemma 2.2], we have

$$\begin{aligned} Q[|(J \bullet X)_t| > c] &= Q \circ X_D^{-1}[|(J \bullet X)_t \circ X_D| > c] \\ &\leq \liminf_{k \rightarrow \infty} Q_{n_k} \circ X_D^{-1}[|(J \bullet X)_t \circ X_D| > c] \\ &= \liminf_{k \rightarrow \infty} Q_{n_k}[|(J \bullet X)_t| > c], \quad c > 0, t \in D, \end{aligned} \quad (4.6)$$

where  $J \in \mathcal{J}(D) \subseteq \mathcal{E}(Q^{n_k})$  for all  $k \in \mathbb{N}$ . Due to the condition (UT), for every  $t \in D$ , the last term in (4.6) tends to zero uniformly over  $\mathcal{J}(D)$  as  $c \rightarrow \infty$ , i.e., the family  $\{(J \bullet X)_t : J \in \mathcal{J}(D)\}$  is  $Q$ -tight, i.e., bounded in  $Q$ -probability, for every  $t \in D$ . The topology of convergence in probability is metrisable; so for every  $t \in D$ , sets contained in the (sequential) closure of  $\{(J \bullet X)_t : J \in \mathcal{J}(D)\}$  are bounded, which in particular entails that for every  $t \in I$ , the set  $\{(H \bullet X)_t : H \in \mathcal{E}(Q)\}$  is bounded in  $Q$ -probability. Indeed, by adapting a sequence of approximation arguments from [34], we show that for every  $t \in I$ , the set  $\{(H \bullet X)_t : H \in \mathcal{E}(Q)\}$  is contained in the closure of  $\{(J \bullet X)_s : J \in \mathcal{J}(D)\}$ , for any  $s \geq t$ ,  $s \in D$ . First, fix  $s \geq t$ ,  $s \in D$ . Now, since  $D$  is dense in  $I$  and contains  $T$ , for every  $t_0 < t_1 < \dots < t_n = t$  in  $I$ ,  $n \in \mathbb{N}$ , there exist  $t_0^k \leq t_1^k \leq \dots \leq t_n^k \leq s$  in  $D$ ,  $k \in \mathbb{N}$ , such that  $t_j \leq t_j^k$  for every  $j \leq n$  and every  $k \geq 1$ , and  $t_j^k \downarrow t_j$  for every  $j \leq n$  as  $k \rightarrow \infty$ ; we allow  $t_n^k = T$  if  $t_n = T$ . Since  $d$  and  $n$  are finite, the right-continuity of  $X = (X^1, \dots, X^d)$  yields

$$X_{t_j^k}^i \rightarrow X_{t_j}^i, \quad \text{uniformly over } i = 1, \dots, d \text{ and } j = 1, \dots, n, \text{ as } k \rightarrow \infty. \quad (4.7)$$

Secondly, for every  $i \leq d$ , every  $j \leq n$  and every  $t_j < T$ , any  $\mathcal{F}_{t_j}$ -measurable  $|H_{t_j}^i| \leq 1$  is  $\mathcal{F}_{t_j^k}^0$ -measurable for all  $k \geq 1$  and can therefore be expressed as a uniform limit of simple  $\mathcal{F}_{t_j^k}^0$ -measurable functions bounded by 1 in absolute value for every  $k \geq 1$ , i.e., for every  $i \leq d$ , every  $j \leq n$  and every  $k \geq 1$ , there exist functions



$|S_{t_j^k}^{i,\ell}| \leq 1, \ell \in \mathbb{N}$ , such that each  $S_{t_j^k}^{i,\ell}$  is of the form

$$S_{t_j^k}^{i,\ell} = \sum_{h=1}^{q(\ell)} \beta_{i,j,k}^{h,\ell} \mathbb{1}_{F_{i,j,k}^{h,\ell}}, \quad \beta_{i,j,k}^{h,\ell} \in \mathbb{R}, F_{i,j,k}^{h,\ell} \in \mathcal{F}_{t_j^k}^0, 1 \leq q(\ell) < \infty, \tag{4.8}$$

and we have

$$\|H_{t_j^k}^i - S_{t_j^k}^{i,\ell}\|_\infty \longrightarrow 0 \quad \text{as } \ell \rightarrow \infty. \tag{4.9}$$

Further, since each  $\mathcal{A}_{t_1, \dots, t_j}$  is an algebra generating  $\otimes_{i \leq j} \mathcal{B}(\mathbb{R}^d)$  on  $\mathbb{R}^{d \times j}$  and the finite unions of the cylindrical sets  $X_{t_1, \dots, t_j}^{-1}(\otimes_{i \leq j} \mathcal{B}(\mathbb{R}^d))$  form an algebra generating the canonical  $\sigma$ -algebra on  $\mathbb{D}(I; \mathbb{R}^d)$ , for every  $0 < t \in I$ , the family

$$\mathcal{A}_{t-}^o = \left\{ \bigcup_{k=1}^n X_{t_1^k, \dots, t_{j(k)}^k}^{-1}(A_k) : A_k \in \mathcal{A}_{t_1^k, \dots, t_{j(k)}^k}, t_1^k < \dots < t_{j(k)}^k < t, j(k), n \in \mathbb{N} \right\}$$

is an algebra generating  $\mathcal{F}_{t-}^0 = \sigma(X_u : u < t)$  on  $\mathbb{D}(I; \mathbb{R}^d)$ ; cf. Corollary A.11. Thus for every  $F_{i,j,k}^{h,\ell} \in \mathcal{F}_{t_j^k}^0$  in (4.8), there exists a sequence  $(A_{i,j,k}^{h,\ell,m})_{m \in \mathbb{N}}$  in  $\mathcal{A}_{t_j^k-}^o$  with

$$\mathbb{1}_{A_{i,j,k}^{h,\ell,m}} \longrightarrow_Q \mathbb{1}_{F_{i,j,k}^{h,\ell}} \quad \text{as } m \rightarrow \infty. \tag{4.10}$$

Finally, by combining the approximations (4.7) and (4.9), and in (4.9) invoking the approximation (4.10) in the sums (4.8), we conclude that for every  $t \in I$  and  $s \geq t, s \in D$ , for every  $H \in \mathcal{E}(Q)$ , there exist a sequence  $(J_n)_{n \in \mathbb{N}}, n = n(k, \ell, m)$ , of elements in  $\mathcal{J}(Q)$  and a sequence  $(t_n)_{n \in \mathbb{N}}$  in  $D \cap [t, s]$  such that  $t_n \downarrow t$  and we have

$$(J_n \bullet X)_{t_n} = \sum_{i=1}^d (J_n^i \bullet X^i)_{t_n} \longrightarrow_Q \sum_{i=1}^d (H^i \bullet X^i)_t = (H \bullet X)_t \quad \text{as } k \wedge \ell \wedge m \rightarrow \infty.$$

Thus for every  $t \in I$  and for any  $s \geq t, s \in D$ , the family of simple integrals  $\{(H \bullet X)_t : H \in \mathcal{E}(Q)\}$  is contained in the closure of  $\{(J \bullet X)_s : J \in \mathcal{J}(D)\}$  that by (4.6) is bounded in  $Q$ -probability. More precisely, (4.6) is immediate only for  $t \in D$ . However, this is sufficient: because  $D$  is dense in  $I$ , the set of simple integrals up to  $t \in I$  is the intersection of the sets of simple integrals up to  $s \in D$  for  $s > t$ , and an intersection of bounded sets is bounded. Hence for every  $t \in I$ , the family of simple integrals  $\{(H \bullet X)_t : H \in \mathcal{E}(Q)\}$  is bounded in  $Q$ -probability, i.e., the weak\* limit  $Q$  is an  $(\mathcal{F}_t)_{t \in I}$ -semimartingale measure and consequently an  $(\mathcal{F}_t^0)_{t \in I}$ -semimartingale measure; see e.g. [47, Theorem II.4].  $\square$

The following Proposition 4.12 is essentially [44, Theorem 4].

**Proposition 4.12** *Let  $(Q_n)_{n \in \mathbb{N}}$  be a sequence of quasimartingale measures satisfying the condition (UB) and converging in the weak\* topology to  $Q$ . Then the weak\* limit  $Q$  is a quasimartingale measure.*

*Proof* Let  $i \leq d$  be fixed. We adapt the proof of [44, Theorem 4] and show that the coordinate process  $X^i$  is a quasimartingale under  $Q$  on  $\mathbb{D}(I; \mathbb{R}^d)$ . We have

$$E_Q \left[ \frac{1}{\varepsilon} \int_0^\varepsilon |X^i_{(t+u) \wedge T}| du \right] \leq \liminf_{n \rightarrow \infty} E_{Q_n} \left[ \frac{1}{\varepsilon} \int_0^\varepsilon |X^i_{(t+u) \wedge T}| du \right] \leq b^i$$

for every  $t \in I$  and every  $\varepsilon > 0$ , where  $b^i := \liminf_{n \rightarrow \infty} \sup_{t \in I} E_{Q_n} [|X^i_t|] < \infty$  by the condition (UB). Thus by Fatou's lemma, we get

$$E_Q [|X^i_t|] \leq \liminf_{\varepsilon \rightarrow 0} E_Q \left[ \frac{1}{\varepsilon} \int_0^\varepsilon |X^i_{(s+u) \wedge T}| du \right] < \infty \quad (4.11)$$

for every  $t \in I$ . The truncated coordinate process, that is,

$$X^{i,a,b} := (-a) \vee (X^i \wedge b) = (-a) \vee X^i - (X^i - b)^+, \quad a, b > 0,$$

is a difference of two convex 1-Lipschitz functions of  $X^i$  for every  $a, b > 0$ ; so we have

$$\text{Var}_t^{Q_n}(X^{i,a,b}) \leq 4\text{Var}_t^{Q_n}(X^i), \quad n \in \mathbb{N}, t \in I, \quad (4.12)$$

see e.g. [53]. Let  $0 = t_0 < t_1 < \dots < t_k = t$  and  $|f_j| \leq 1$ ,  $j < k$ , be continuous  $\mathcal{F}_{t_{j-}}^0$ -measurable functions. By (4.12), we have

$$E_{Q_n} \left[ \sum_{j=1}^k f_{j-1}(X) (X^{i,a,b}_{(u+t_j) \wedge T} - X^{i,a,b}_{(u+t_{j-1}) \wedge T}) \right] \leq 4\text{Var}_t^{Q_n}(X^i), \quad n \in \mathbb{N}; \quad (4.13)$$

so by Fubini's theorem, for every  $n \in \mathbb{N}$  and every  $\varepsilon > 0$ , we get

$$E_{Q_n} \left[ \frac{1}{\varepsilon} \int_0^\varepsilon \left( \sum_{j=1}^k f_{j-1}(X) (X^{i,a,b}_{(u+t_j) \wedge T} - X^{i,a,b}_{(u+t_{j-1}) \wedge T}) \right) du \right] \leq 4\text{Var}_t^{Q_n}(X^i).$$

The mappings

$$F^\varepsilon(X) := \frac{1}{\varepsilon} \int_0^\varepsilon \left( \sum_{j=1}^k f_{j-1}(X) (X^{i,a,b}_{(u+t_j) \wedge T} - X^{i,a,b}_{(u+t_{j-1}) \wedge T}) \right) du, \quad \varepsilon > 0,$$

are lower semicontinuous and bounded from below, see (A.2) and Lemma A.12, so that we have for all  $\varepsilon > 0$  that

$$E_Q \left[ \frac{1}{\varepsilon} \int_0^\varepsilon \left( \sum_{j=1}^k f_{j-1}(X) (X^{i,a,b}_{(u+t_j) \wedge T} - X^{i,a,b}_{(u+t_{j-1}) \wedge T}) \right) du \right] \leq 4v^i, \quad (4.14)$$

where  $v^i := \liminf_{n \rightarrow \infty} \sup_{t \in I} \text{Var}_t^{Q_n}(X^i) < \infty$  by the assumption (UB). Due to (4.11), letting  $\varepsilon \rightarrow 0$ , then  $a \rightarrow \infty$  and finally  $b \rightarrow \infty$  in (4.14), right-continuity

and applying the monotone convergence theorem twice yield

$$E_Q \left[ \sum_{j=1}^k f_{j-1}(X)(X_{t_j}^i - X_{t_{j-1}}^i) \right] \leq 4v^i, \tag{4.15}$$

for all  $\mathcal{F}_{t_j-}^0$ -measurable continuous functions  $|f_j| \leq 1, j < k$ . Furthermore, by choosing  $f_j(X) = f(X_{t_j+u})$  before (4.13) for a continuous function  $|f| \leq 1$  on  $\mathbb{R}^d$ , we conclude that the inequality (4.15) is true for a family of continuous functions that for every  $j < k$  generates the  $\sigma$ -algebra  $\mathcal{F}_{t_j}^0$ ; cf. Corollary A.11. Thus by the standard  $L^1$ -approximation via Lusin’s theorem and Tietze’s extension theorem, for any  $\mathcal{F}_{t_j}^0$ -measurable  $|H_{t_j}| \leq 1$ , for every  $j < k$ , there exists a sequence  $(f_j^n)_{n \in \mathbb{N}}, |f_j^n| \leq 1$ , of functions satisfying (4.15) such that  $f_j^n \rightarrow H_{t_j}$  in  $L^1(Q)$  as  $n \rightarrow \infty$ ; see e.g. [17] and [16, 2.1.8]. Thus the inequality (4.15) is true for all  $\mathcal{F}_{t_j}^0$ -measurable functions  $|H_{t_j}| \leq 1$ , and so the process  $X$  is an  $(\mathcal{F}_t^0)_{t \in I}$ -quasimartingale on  $(\mathbb{D}(I), \mathcal{F}_T, Q)$ ; see [14, Appendix 2, (3.5)]. By Rao’s decomposition theorem, this is a necessary and sufficient condition for  $X$  to be decomposable into a difference  $X = Y - Z$  of two càdlàg  $(\mathcal{F}_t^0)_{t \in I}$ -supermartingales  $Y$  and  $Z$  on  $(\mathbb{D}(I), \mathcal{F}_T, Q)$ ; see [25, Theorem 8.13]. On the other hand, by Föllmer’s lemma,  $Y$  and  $Z$  are  $(\mathcal{F}_t)_{t \in I}$ -supermartingales, see [25, Theorem 2.46]; so again by Rao’s decomposition theorem, the process  $X$  is an  $(\mathcal{F}_t)_{t \in I}$ -quasimartingale on  $(\mathbb{D}(I), \mathcal{F}_T, Q)$ .  $\square$

The following Proposition 4.13 is essentially [44, Theorem 11].

**Proposition 4.13** *Let  $(Q_n)_{n \in \mathbb{N}}$  be a sequence of supermartingale measures satisfying the condition (UI) and converging in the weak\* topology to  $Q$ . Then the weak\* limit  $Q$  is a supermartingale measure.*

*Proof* We adapt the proof of [44, Theorem 11] and show that each coordinate process  $X^i, i \leq d$ , is a supermartingale under  $Q$ . We have  $E_Q[|X_t|] < \infty$  for every  $t \in I$ ; cf. (4.11). Moreover, by Lemma 4.10, there exist a subsequence  $(Q_{n_k})_{k \in \mathbb{N}}$  and a countable dense set  $D \subseteq I$  such that  $T \in D$  if  $I = [0, T]$  and  $(Q_{n_k})_{k \in \mathbb{N}}$  converges to  $Q$  in finite-dimensional distributions on the set  $D$ . Let  $X^{i,c}$  denote the coordinate process  $X^i$  truncated from above at  $c > 0$ , i.e.,

$$X^{i,c} := X^i \wedge c, \quad c > 0.$$

By the condition (UI) and the fact that each  $Q_{n_k}$  is a supermartingale measure for  $X^{i,b}$ , we have for every  $0 < a < b$  that

$$E_Q[f(X)(X_t^{i,a} - X_s^{i,b})] \leq \liminf_{k \rightarrow \infty} E_{Q_{n_k}}[f(X)(X_t^{i,a} - X_s^{i,b})] \leq 0, \quad s < t, s, t \in D,$$

where

$$f(X) := f_1(X_{t_1})f_2(X_{t_2}) \cdots f_n(X_{t_n}), \quad t_j \in D, f_j \in \mathbb{C}_b(\mathbb{R}^d), j \leq n;$$

see e.g. [57, Theorem 2.20]. Consequently, by Corollary A.11, we have

$$E_Q[\mathbb{1}_F(X)(X_t^{i,a} - X_s^{i,b})] \leq 0, \quad 0 < a < b, \quad (4.16)$$

for every  $s < t$  in  $D$  and  $F \in \mathcal{F}_{s-}^0$ . Letting  $b \rightarrow \infty$  and then  $a \rightarrow \infty$  in (4.16), applying monotone convergence twice yields the same inequality for the coordinate process  $X^i$ . By Föllmer's lemma, the inequality extends immediately to the whole set  $I$ , and further to  $F \in \mathcal{F}_{s+}^0$ . Indeed, we have

$$E_Q[\mathbb{1}_F(X)(X_t^i - X_s^i)] = \lim_{n \rightarrow \infty} E_Q[\mathbb{1}_F(X)(X_t^i - X_{s+1/n}^i)] \leq 0 \quad (4.17)$$

for every  $F \in \mathcal{F}_{s+}^0$ ; cf. [25, Theorem 2.44].  $\square$

For the sake of completeness, we provide the following Proposition 4.14. Assertion (a) is the classical Kolmogorov criterion for almost sure (Hölder-)continuity, and assertion (b) is essentially [27, Proposition 6]; see also [42, Lemma 4.5].

**Proposition 4.14** *Let  $(Q_n)_{n \in \mathbb{N}}$  be a sequence of probability measures converging in the weak\* topology to  $Q$ . Then we have the following:*

- (a) *If the sequence  $(Q_n)_{n \in \mathbb{N}}$  is  $C$ -tight, then the limit  $Q$  is  $C$ -tight.*
- (b) *If each  $Q_n$  is Lipschitz–Markov, then the limit  $Q$  is Lipschitz–Markov.*

*Proof* (a) By [25, Lemma 15.49], the  $C$ -tightness of the sequence  $(Q_n)_{n \in \mathbb{N}}$  implies the convergence along a subsequence in the weak\* topology of the Skorokhod  $J^1$ -topology, and a fortiori in any weaker topology, to the law of a continuous process, which is the limit  $Q$ ; cf. Proposition 5.12 below.

(b) Let  $s < t$  in  $I$  be fixed and take a bounded Lipschitz-continuous function  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  with a Lipschitz constant  $L(g) \leq 1$ . For each  $n \in \mathbb{N}$ , there exists a bounded Lipschitz-continuous function  $f_n : \mathbb{R}^d \rightarrow \mathbb{R}$  such that

$$f_n(X_s) = E_{Q_n}[g(X_t) | \mathcal{F}_s^0] \quad Q\text{-a.s.} \quad (4.18)$$

Further, we can assume that  $L(f_n) \leq 1$  and  $\|f_n\|_\infty \leq \|g\|_\infty < \infty$  for all  $n \in \mathbb{N}$ . Thus by the Arzelà–Ascoli theorem, there exist a subsequence  $(f_{n_k})_{k \in \mathbb{N}}$  and a bounded Lipschitz-continuous function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  with  $L(f) \leq 1$  such that  $(f_{n_k})$  converges to  $f$  uniformly on compacts as  $k \rightarrow \infty$ . Further, by Lemma 4.10, there exists a further subsequence of  $(Q_{n_k})_{k \in \mathbb{N}}$ , that we again denote by  $(Q_{n_k})_{k \in \mathbb{N}}$ , converging in finite-dimensional distributions to  $Q$  on a dense set  $D \subseteq I$  such that  $T \in D$  if  $I = [0, T]$ . Let  $i \in \mathbb{N}$  and  $0 \leq s_1 < s_2 < \dots < s_i \leq s < t$  in  $D$  and take a bounded continuous compactly supported  $\alpha : \mathbb{R}^d \rightarrow \mathbb{R}$  and a bounded continuous  $\beta : \mathbb{R}^{d \times i} \rightarrow \mathbb{R}$ . By (4.18), we have

$$E_{Q_{n_k}}[(f_{n_k}(X_s) - g(X_t))\alpha(X_s)(X_{s_1}, \dots, X_{s_i})] = 0, \quad \forall k \in \mathbb{N}. \quad (4.19)$$

Since  $(f_{n_k})_{k \in \mathbb{N}}$  converges uniformly to  $f$  on the compact support of each  $\alpha$  and  $(Q_{n_k})_{k \in \mathbb{N}}$  converges to  $Q$  in finite-dimensional distributions on the set  $D$  as  $k \rightarrow \infty$ ,

from (4.19), by the vectorial Portemanteau lemma, see e.g. [57, Lemma 2.2], we get

$$E_Q[(f(X_s) - g(X_t))\alpha(X_s)\beta(X_{s_1}, \dots, X_{s_i})] = 0. \tag{4.20}$$

Because (4.20) holds for all bounded continuous  $\alpha$  with compact support, we get

$$E_Q[(f(X_s) - g(X_t))\beta(X_{s_1}, \dots, X_{s_i})] = 0, \tag{4.21}$$

and as (4.21) holds for all bounded continuous  $\beta$ , we obtain

$$E_Q[(f(X_s) - g(X_t))h(X)] = 0, \tag{4.22}$$

for every bounded  $\mathcal{F}_s^0$ -measurable function  $h$ , for every  $s < t$  in  $D$ . Since  $D$  is a dense subset of  $I$  and  $s \mapsto f(X_s)$  is bounded and right-continuous on  $I$ , (4.22) extends by dominated convergence to the whole set  $I$ , and further to every bounded  $\mathcal{F}_s$ -measurable function  $h$ ; cf. (4.17). □

### 4.2.2 Tightness

We say a family  $\mathcal{Q}$  of probability measures on  $(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T)$  satisfies *Jakubowski's uniform tightness criterion* if we have

$$\begin{aligned} \lim_{c \rightarrow \infty} \sup_{Q \in \mathcal{Q}} Q[\|X^{i,t}\|_\infty > c] &= 0, \\ \lim_{c \rightarrow \infty} \sup_{Q \in \mathcal{Q}} Q[N^{a,b}(X^{i,t}) > c] &= 0, \quad \forall a < b, \end{aligned} \tag{US}$$

for every finite  $t \in I$  and every  $i \leq d$ , where  $X^{i,t}$  denotes the coordinate process  $X^i$  restricted to  $[0, t]$ ; cf. Corollary 5.9. It was shown in [31] that a family of probability measures on  $(\mathbb{D}([0, T]; \mathbb{R}), \mathcal{F}_T)$ ,  $T < \infty$ , satisfies the condition (US) if and only if it is uniformly  $S$ -tight. In particular, we have the hierarchy, cf. (2.2),

$$(UT) \implies (US) \implies (US^*), \tag{4.23}$$

where (US\*) stands for the uniform tightness in the  $S^*$ -topology; see Sect. 5 below.

The second implication in (4.23) is immediate from the definition of the  $S^*$ -topology; see Proposition 5.4 (i). The first implication in (4.23) follows from Proposition 4.15 below, which is essentially the result of Stricker [55, Theorem 2] which states that a sequence satisfying the condition (UT) admits a convergent subsequence and the resulting limit law is the law of a semimartingale. Analogous results were obtained for the  $S$ -topology by Jakubowski in [31, Theorem 4.1]; see also [31, Proposition 3.1].

**Proposition 4.15** *A family of semimartingale measures satisfying the condition (UT) satisfies the condition (US).*

*Proof* Following Stricker [54, Theorem 2], we define the family of stopping times

$$\tau^{i,c} = \inf\{s \in I : |X_s^{i,t}| > c\}, \quad i \leq d, c > 0,$$

and approximate each  $\tau^{i,c}$  from the right with the sequence of *simple* stopping times

$$\tau_n^{i,c} = \min\{m/n : m \in \mathbb{N}, \tau^{i,c} \leq m/n\}, \quad n \in \mathbb{N}. \quad (4.24)$$

Since we are assuming a right-continuous filtration  $(\mathcal{F}_t)_{t \in I}$  and  $X^i$  is right-continuous, the hitting times  $\tau^{i,c}$  and consequently their approximations  $\tau_n^{i,c}$  are indeed stopping times. Moreover, since each  $\tau_n^{i,c}$  takes only finitely many values in  $[0, t]$ , every process  $|H^n| \leq 1$  of the form

$$H^n = \mathbb{1}_{\|0, \tau_n^{i,c} \wedge t\|}, \quad i \leq d, c > 0, n \in \mathbb{N}, t \in I, \quad (4.25)$$

is an elementary predictable integrand; see (2.1). Now due to the right-continuity of  $X^i$ , dominated convergence gives for every  $Q \in \mathcal{Q}$  that

$$Q[\|X^{i,t}\|_\infty > c] = Q[|(H^n \bullet X^i)_t| > c, \forall n \in \mathbb{N}] \quad (4.26)$$

for every  $t \in I$  and every  $c > 0$ . By the condition (UT), the left-hand side in (4.26) tends to 0 uniformly over  $Q \in \mathcal{Q}$  for every  $i \leq d$  and every  $t \in I$  as  $c \rightarrow \infty$ . Similarly, for  $a < b$ , we define recursively for all  $k \in \mathbb{N}_0$  the stopping times  $\sigma_0^{i,a} = \tau_0^{i,b} = 0$  and

$$\begin{aligned} \sigma_k^{i,a} &= \inf\{s > \tau_{k-1}^{i,b} : |X_s^{i,t}| < a\}, \\ \tau_k^{i,b} &= \inf\{s > \sigma_k^{i,a} : |X_s^{i,t}| > b\}, \end{aligned}$$

and the respective decreasing sequences  $(\sigma_{k,n}^{i,a})_{n \in \mathbb{N}}$  and  $(\tau_{k,n}^{i,b})_{n \in \mathbb{N}}$  of approximating stopping times taking only finitely many values on finite intervals; cf. (4.24). The processes  $|H^{m,n}| \leq 1$ ,  $m, n \in \mathbb{N}$ , of the form

$$H^{m,n} = \sum_{k=1}^m \mathbb{1}_{\|\sigma_{k,n}^{i,a} \wedge t, \tau_{k,n}^{i,b} \wedge t\|}$$

are finite linear combinations of processes of the form (4.25) so that each process  $|H^{m,n}| \leq 1$  is an elementary predictable integrand. Moreover, we have

$$Q[N^{a,b}(X^{i,t}) > c] \leq Q\left[\lim_{m \rightarrow \infty} |(H^{m,n} \bullet X^i)_t| > a^+ + c(b-a), \forall n \in \mathbb{N}\right]. \quad (4.27)$$

By the condition (UT), for every  $a < b$ , the right-hand side of (4.27) tends to zero uniformly over  $Q \in \mathcal{Q}$  as  $c \rightarrow 0$ ; cf. (4.6) and (4.7). Thus by Corollary 5.9 below, the family  $\mathcal{Q}$  satisfies the condition (US).  $\square$

### 4.3 The proofs of the main results

In this section, we provide the proofs for the results of Sect. 3.2 that we omitted there. We begin by proving Theorem 3.2 by invoking the results of Sect. 4.1 in conjunction with the stability and tightness results on the semimartingale property of Sect. 4.2. Then the rest of the results of Sect. 3.2 follow from the respective stability results of Sect. 4.2.

*Proof of Theorem 3.2* The condition (UT) is stronger than the condition (US\*); cf. (4.23). So by Proposition 4.7, the family  $\mathcal{S}$  is  $\beta_0$ -equicontinuous. Thus by Corollary 4.8, the closure of  $\mathcal{S}$  is compact and sequentially compact in the weak\* topology; see Proposition 4.5. By Corollary 4.6, the closure of  $\mathcal{S}$  coincides with the sequential closure of  $\mathcal{S}$ . It remains to show that every element in the sequential closure  $[\mathcal{S}]_{\text{seq}}$  is a semimartingale measure. This particular fact is the statement of Proposition 4.11.  $\square$

*Proof of Corollary 3.3* The condition (UB) is weaker than the condition (UT); see (2.2). By Proposition 4.12, the class of quasimartingale measures is stable in the weak\* convergence under (UB). Thus Corollary 3.3 follows from Theorem 3.2.  $\square$

*Proof of Corollary 3.4* The condition (UI) is weaker than the condition (UB); see (2.2). By Proposition 4.13, the class of supermartingale measures is stable in the weak\* convergence under (UI). Thus Corollary 3.4 follows from Corollary 3.3.  $\square$

*Proof of Lemma 3.5* Since every sequence in the set is  $C$ -tight, by Proposition 4.14 (a), every limit point in the sequential closure is  $C$ -tight. Thus Lemma 3.5 (a) is true. Likewise, Lemma 3.5 (b) is a direct consequence of the stability of the Lipschitz–Markov property under weak\* convergence; see Proposition 4.14 (b).  $\square$

## 5 The weak $\mathcal{S}$ -topology

We introduce the weak  $\mathcal{S}$ -topology and study its properties and relation to other topologies on the Skorokhod space.

### 5.1 Definition

A possibility of defining a completely regular (non-sequential)  $\mathcal{S}$ -topology is discussed already by Jakubowski in [31]; see [31, Sect. 3]. We describe here a general method for regularising any given topology. Our approach is inspired by the seminal work of Alexandroff [1, Chap. 2]. Let  $\mathcal{X} = (X, \tau)$  be an arbitrary topological space and  $\mathcal{V}$  an arbitrary subbase for the Euclidean topology on  $\mathbb{R}$ ; then the family

$$\{f^{-1}(V) : f \in \mathbb{C}_b(\mathcal{X}), V \in \mathcal{V}\} \quad (5.1)$$

is a subbase for a (unique) topology  $\tau^*$  on  $X$ . Indeed, the topology  $\tau^*$  generated by the subbase (5.1) on  $X$  is independent of the choice of the subbase  $\mathcal{V}$  on  $\mathbb{R}$ ; see e.g. [23, 3.4]. The topology  $\tau^*$  is generated by the family of pseudometrics

$$\{\rho_{f_1, f_2, \dots, f_k} : f_1, f_2, \dots, f_k \in \mathbb{C}_b(\mathcal{X})\}, \quad (5.2)$$

where

$$\rho_{f_1, f_2, \dots, f_k}(x, y) := \max\{|f_1(x) - f_1(y)|, |f_2(x) - f_2(y)|, \dots, |f_k(x) - f_k(y)|\}$$

for  $x, y \in X$ , and thus the convergence of a net  $(x_\alpha)$  to an element  $x$  in this topology  $\tau^*$  is equivalent to

$$f(x_\alpha) \longrightarrow f(x), \quad \forall f \in \mathbb{C}_b(\mathcal{X}); \quad (5.3)$$

see e.g. [16, Example 8.1.19]. We remark that by replacing  $\mathbb{C}_b(\mathcal{X})$  with  $\mathbb{C}(\mathcal{X})$  in (5.1)–(5.3), one obtains an equivalent characterisation. Any of these characterisations is necessary and sufficient for a topological space  $\mathcal{X}$  to be completely regular ( $T_{3\frac{1}{2}}$ ); see e.g. [23, 3.4].

**Definition 5.1** The *weak  $S$ -topology*  $S^*$  is the topology  $\tau^*$  constructed from the  $S$ -topology on the Skorokhod space  $\mathbb{D}$  by choosing  $\mathcal{X} = (\mathbb{D}, S)$  above.

*Remark 5.2* Convergence in the weak\* topology on the  $\beta_0$ -dual of  $\mathbb{C}_b(\mathbb{D})$  (as described in Lemma 2.4), traditionally called “weak convergence” for sequences of probability measures, is equivalent to the convergence (5.3) if the measures are Dirac measures; see [8, Lemma 8.9.2.].

*Remark 5.3* It should be emphasised that if one could show that the  $S$ -topology is regular (or linear), then the  $S$ - and the weak  $S$ -topology would coincide. It was communicated to the author by Professor Jakubowski that the regularity of the  $S$ -topology remains an open question.

## 5.2 Properties

Recall Property 4.1 from Sect. 4. A topological space satisfying Property 4.1 is submetrisable, from which various useful properties follow; see (4.2) and [30]. In fact, all key properties of the  $S$ -topology follow immediately from Property 4.1, and Property 4.1 is preserved in the regularisation (5.1).

Given a topological space  $\mathcal{X} = (X, \tau)$ , we write  $\mathcal{K}(\tau)$  for  $\mathcal{K}(\mathcal{X})$ ,  $\mathcal{B}(\tau)$  for  $\mathcal{B}(\mathcal{X})$  etc.

**Proposition 5.4** *The  $S$ -topology on  $\mathbb{D}$  has the following properties:*

- (a) *The Skorokhod space endowed with  $S$  is a Hausdorff space.*
- (b) *Each  $K \in \mathcal{K}(S)$  is metrisable.*
- (c) *A set is compact if and only if it is sequentially compact.*
- (d) *The Borel  $\sigma$ -algebra  $\mathcal{B}(S)$  and the canonical  $\sigma$ -algebra coincide.*
- (e) *The Skorokhod space endowed with  $S$  is a Lusin space.*

*The  $S^*$ -topology on  $\mathbb{D}$  has the properties (a)–(e) and additionally:*

- (f) *The Skorokhod space endowed with  $S^*$  is perfectly normal and paracompact.*
- (g) *The Borel  $\sigma$ -algebra  $\mathcal{B}(S^*)$  and the Baire  $\sigma$ -algebra  $\mathcal{B}_a(S^*)$  coincide.*
- (h)  $\mathbb{C}(S) = \mathbb{C}(S^*)$ .
- (i)  $\mathcal{K}(S) = \mathcal{K}(S^*)$ .



*Proof* The properties (a), (b) and (c) follow immediately from the fact that both  $S$  and  $S^*$  satisfy Property 4.1; see [30, pages 10–11]. Indeed, the mappings

$$\frac{1}{r} \int_q^{q+r} X_t^i dt \quad \text{and} \quad X_T^i, \quad \text{for } I = [0, T],$$

where  $q$  and  $q + r$  run over the rationals in  $I$  and  $i$  over the spatial dimensions  $1, \dots, d$ , constitute a countable family of continuous functions possessing the property (4.1); cf. Example 5.10 below.

(d) We prove the claim for  $I = [0, T]$ . The proof is completely similar for  $I = [0, \infty)$ . Fix a coordinate  $i \leq d$ . For all  $0 \leq t < T$ , we have

$$X_t^i = \lim_{\delta \rightarrow 0} \frac{1}{\delta} \int_t^{t+\delta} X_u^i du,$$

i.e., the mapping  $X_t^i$  is a limit of elements of  $\mathbb{C}(S)$  for every  $t$  in  $I$ , while for  $t = T$ , the mapping  $X_t^i$  is an element of  $\mathbb{C}(S)$ . Consequently, we have  $\sigma(X_u : u \in I) \subseteq \mathcal{B}(S^*)$  and since  $S^*$  is weaker than  $S$ , we have  $\mathcal{B}(S^*) \subseteq \mathcal{B}(S)$ . On the other hand, by Proposition 5.12 below,  $S$  is weaker than  $J^1$ ; so we have  $\mathcal{B}(S) \subseteq \mathcal{B}(J^1) = \sigma(X_u : u \in I)$ . Thus  $\mathcal{B}(S^*) = \mathcal{B}(S) = \sigma(X_u : u \in I)$ . By Proposition 5.12 below, we have  $S^* \subseteq S \subseteq J^1$  and the Skorokhod space endowed with  $J^1$  is a Polish space; so the Skorokhod space endowed with  $S$  or  $S^*$  is a Lusin space. Thus we have (e).

The Skorokhod space endowed with  $S^*$  is a (completely) regular Souslin space. By Fernique [18, Proposition I.6.1], every regular Souslin space is perfectly normal and paracompact. Thus we have (f). Now by (f), the Skorokhod space endowed with  $S^*$  is perfectly normal, and consequently, by [8, Proposition 6.3.4], we have  $\mathcal{B}(S^*) = \mathcal{B}a(S^*)$ , i.e., we have (g). The claim (h) follows directly from Definition 5.1.

To prove (i), we first observe that  $\mathcal{K}(S) \subseteq \mathcal{K}(S^*)$  by Definition 5.1. To prove the converse inclusion, we use Jakubowski’s  $\Sigma$ -topology; see Appendix A.1.2. By [33, Remark 3.6], we have  $\Sigma \subseteq S$  and hence  $\mathbb{C}(\Sigma) \subseteq \mathbb{C}(S) = \mathbb{C}(S^*)$ . Thus we have  $\Sigma \subseteq S^*$  since the topology  $\Sigma$  is completely regular. Indeed, topological vector spaces are completely regular; see e.g. [9, Theorem 1.6.5]. Consequently, [33, Remark 3.8] gives  $\mathcal{K}(S^*) \subseteq \mathcal{K}(\Sigma) = \mathcal{K}(S)$ . Thus we have shown  $\mathcal{K}(S) = \mathcal{K}(S^*)$ . □

*Remark 5.5* A countable product of regular Souslin spaces is a regular Souslin space. Thus by Fernique [18, Proposition I.6.1], the previous properties (after the obvious modifications) are inherited for (at most) countable products of  $S^*$ -topologies; cf. Sect. 4.1.

### 5.3 Compact sets and continuous functions

In this section, we recall compactness and continuity criteria for the  $S$ -topology from [31] and [33].

#### 5.3.1 Compactness criteria

The *necessity and sufficiency* of the condition (5.4) below for the relative (sequential) compactness in the  $S$ -topology was proved in [31] for  $I = [0, T]$ ,  $T < \infty$ , and the

multidimensional infinite-horizon extension in Proposition 5.8 below was provided in [33].

**Proposition 5.6** *A subset  $K$  of  $\mathbb{D}([0, T]; \mathbb{R})$ ,  $T < \infty$ , is relatively sequentially  $S$ -compact if and only if it satisfies the conditions*

$$\left\{ \begin{array}{l} \sup_{\omega \in K} \|\omega\|_\infty < \infty, \\ \sup_{\omega \in K} N^{a,b}(\omega) < \infty, \end{array} \quad \forall a < b, a, b \in \mathbb{R}. \right. \quad (5.4)$$

*Remark 5.7* A right-continuous function  $\omega : [0, T] \rightarrow \mathbb{R}$  is càdlàg if and only if it satisfies the conditions

$$\left\{ \begin{array}{l} \|\omega\|_\infty < \infty, \\ N^{a,b}(\omega) < \infty, \end{array} \quad \forall a < b, a, b \in \mathbb{R}. \right.$$

**Proposition 5.8** *A subset  $K$  of  $\mathbb{D}([0, \infty); \mathbb{R})$  is relatively sequentially  $S$ -compact if and only if the set  $K$  restricted to  $[0, t]$  satisfies (5.4) for every  $0 < t < \infty$ .*

The proof of Proposition 5.8 can be found in [33].

We have the following compactness criterion.

**Corollary 5.9** *Let  $K = K^1 \times \dots \times K^d$  be a Cartesian product set in the Skorokhod space  $\mathbb{D}(I; \mathbb{R}^d)$  endowed with  $S$  or  $S^*$ . Then the set  $K$  is compact if for each  $i \leq d$ , there exist a (non-decreasing) function  $C_{q,r}^i : I \rightarrow \mathbb{R}_+$  for all  $q < r$  in  $\mathbb{Q}$  and a (non-decreasing) function  $M^i : I \rightarrow \mathbb{R}_+$  such that*

$$K^i = \bigcap_{q < r} \{\omega^i : N^{q,r}([\omega^i]^t) \leq C_{q,r}^i(t) \text{ and } \|[\omega^i]^t\|_\infty \leq M^i(t), \forall t < \infty\},$$

where the intersection is taken over all rationals  $q < r$  and  $[\omega^i]^t$  denotes the restriction of  $\omega^i$  to  $[0, t]$ .

*Proof* The product set  $K = K^1 \times \dots \times K^d$  is compact in the product topology if each set  $K^i$  in the product is  $S$ -compact; cf. Proposition 5.4 (i). For any two real numbers  $a < b$ , one can find rationals  $r < q$  with  $a < r < q < b$ , and so it is sufficient to let  $a < b$  range through rationals in Proposition 5.6. Given that for each  $t < \infty$  and every  $i \leq d$ , one has  $N^{q,r}([\omega^i]^t) \leq C_{q,r}^i(t)$  and  $\|[\omega^i]^t\|_\infty \leq M^i(t)$  for some constants  $C_{q,r}^i(t) < \infty$  and  $M^i(t) < \infty$ , by Proposition 5.6, each set  $K^i$  is relatively  $S$ -compact. By Lemma A.12, the mappings  $N^{q,r}(\cdot)$  and  $\|\cdot\|_\infty$  are lower semicontinuous in the MZ-topology, and by Proposition 5.12 below in the  $S$ -topology as well; so the sets  $K^i$ ,  $i \leq d$ , are  $S$ -closed. Thus the sets  $K^i$  are  $S$ -compact, and hence the product set  $K$  is compact in the product topology.  $\square$

Remark that for  $I = [0, T]$ ,  $T < \infty$ , it suffices to consider constant  $C_{q,r}^i$  and  $M^i$  in Corollary 5.9.

### 5.3.2 Examples of (semi-)continuous functions

By Proposition 5.4 (h),  $S^*$ -continuous functions are precisely the  $S$ -continuous ones. In particular, we should like to emphasise that the evaluation mapping at  $t$  is not continuous for any  $t < T$ ; see [31, Sect. 2].

*Example 5.10* (a) The mappings

$$\omega \mapsto \int_I G(t, \omega^i(t)) d\mu(t), \quad i \leq d,$$

are  $S^*$ -continuous on  $\mathbb{D}(I; \mathbb{R}^d)$  whenever  $G$  is measurable as a mapping of  $(t, x)$ , continuous as a mapping of  $x$  for every  $t \in I$  and such that

$$\sup_{0 \leq t \leq c} \sup_{|x| \leq c} |G(t, x)| < \infty, \quad \forall c > 0,$$

and  $\mu$  is a diffusive (i.e., atomless) measure on  $I$ ; see [31, Corollary 2.11].

(b) The mapping

$$\omega \mapsto \omega(T)$$

is  $S^*$ -continuous on  $\mathbb{D}([0, T]; \mathbb{R}^d)$ ; see [31, Remark 2.4].

*Example 5.11* By Proposition 5.12 below, the  $S^*$ -topology is stronger than the Meyer–Zheng topology MZ; so the functions

$$\omega \mapsto \|\omega\|_\infty \quad \text{and} \quad \omega \mapsto N^{a,b}(\omega^i), \quad a < b, i \leq d,$$

are lower semicontinuous in the  $S^*$ -topology on  $\mathbb{D}(I; \mathbb{R}^d)$ . See Lemma A.12.

## 5.4 Relation to other topologies

The definitions of Jakubowski’s  $\Sigma$ -topology, Jakubowski’s  $S$ -topology, the Meyer–Zheng topology MZ and Skorokhod’s  $J^1$ -topology are given in Appendix A.1. The  $S^*$ -topology is related to these topologies as follows.

**Proposition 5.12** *We have  $MZ \subseteq S^*$ ,  $\Sigma \subseteq S^*$  and  $S^* \subseteq S \subseteq J^1$ .*

*Proof* The functions in (A.2) and (A.3) that generate the topology MZ are  $S^*$ -continuous; see Example 5.10. Moreover, the topology MZ is metrisable and hence in particular sequential and completely regular. Thus by Example 5.10, the inclusion  $MZ \subseteq S^*$  is true; cf. (5.3). The topology  $\Sigma$  is a completely regular topology weaker than  $S$ ; see [33, Remark 3.8] and [9, Theorem 1.6.5]. Since the topology  $S^*$  is the strongest completely regular topology weaker than  $S$ , we have  $\Sigma \subseteq S^* \subseteq S$ . The final inclusion  $S \subseteq J^1$  is proved in [31] for a finite compact interval and extends immediately for the infinite interval due to (A.5); cf. (A.1). □

The Skorokhod space endowed with the  $J^1$ -topology is a Polish space; so the space is a Lusin space for any topology that is weaker than the  $J^1$ -topology. The following Theorem 5.13 states that the  $S^*$ -topology, which is the strongest (completely) regular topology weaker than the  $S$ -topology, is the strongest (completely) regular Souslin topology on the Skorokhod space for which the sets (5.4) are compact, and consequently, Jakubowski's uniform tightness criterion (US) is a sufficient tightness criterion; cf. Sect. 4.2.2.

**Theorem 5.13** *Let  $T$  be a completely regular Souslin topology on the Skorokhod space, comparable to  $S$ , and such that  $\mathcal{K}(T) = \mathcal{K}(S)$ . Then*

$$T \subseteq S.$$

*Proof* Assume that  $S \subseteq T$  and let  $T_S$  denote the sequential topology generated by  $T$ . Since the compact sets of a completely regular Souslin space are metrisable, we have  $\mathcal{K}(T) \subseteq \mathcal{K}(T_S)$ ; see e.g. [8, Theorem 8.10.5]. Consequently, we have

$$\mathcal{K}(S) = \mathcal{K}(T) = \mathcal{K}(T_S), \quad (5.5)$$

where

$$S \subseteq T \subseteq T_S \quad (5.6)$$

and  $S$  and  $T_S$  are sequential; see Appendix A.1.1. By [16, Theorem 3.3.20], the Skorokhod space is a (Hausdorff)  $k$ -space for  $S$  and  $T_S$ ; so by (5.5) and (5.6), we have  $S = T$ .  $\square$

*Remark 5.14* For the Riesz representation theorem in Lemma 2.4 and the required auxiliary results in Sect. 4, it is necessary that the underlying topological space is completely regular and Souslin. Among all such topologies, the topology  $S^*$  is the strongest one that is weaker than  $S$ . Recall that in addition to Lemma 2.4 and the results of Sect. 4, the proof of the main result in Theorem 3.2 goes through the property (US), while the underlying relative compactness criterion (5.4) that gives rise to the tightness (US) is both necessary and sufficient for  $S$ ; see also Remark 5.7. On the other hand, as shown in Theorem 5.13, any topology with the previously cited properties and the appropriate compact closed sets (5.4) cannot be strictly stronger than  $S$ .

**Acknowledgements** Open Access funding provided by Projekt DEAL. The author wants to thank two anonymous referees and the Editor for their careful reading and comments that helped to improve the paper.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material.

If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix

The appendix collects the definitions of topologies and auxiliary results used in the main part of the article.

### A.1 Topologies on the Skorokhod space

We recall the definitions of Jakubowski’s  $S$ -topology and  $\Sigma$ -topology, the Meyer–Zheng pseudo-path topology and Skorokhod’s  $J^1$ -metric topology. We define each topology separately on  $\mathbb{D}([0, T]; \mathbb{R}^d)$  for  $T < \infty$  and on  $\mathbb{D}([0, \infty); \mathbb{R}^d)$ . In particular, we use a formal definition of the Meyer–Zheng pseudo-path topology MZ that takes into account the fluctuations of the terminal value in the case of a finite time horizon. The space  $\mathbb{D}([0, \infty]; \mathbb{R}^d)$  is regarded as the product space

$$\mathbb{D}([0, \infty]; \mathbb{R}^d) = \mathbb{D}([0, \infty); \mathbb{R}^d) \times \mathbb{R}^d,$$

where the space  $\mathbb{R}^d$  is endowed with the Euclidean topology; note the difference between  $[0, \infty]$  and  $[0, \infty)$ .

#### A.1.1 The $S$ -topology

Jakubowski’s  $S$ -topology, introduced in [31], is a sequential topology. The following definition of  $S$ -convergence on  $\mathbb{D}([0, T]; \mathbb{R})$  is taken from [31]; the multidimensional version can be found in [33].

**Definition A.1** On  $\mathbb{D}([0, T]; \mathbb{R}^d)$ , we write  $\omega_n \rightarrow_S \omega_0$  if for every  $i \leq d$  and every  $\varepsilon > 0$ , one can find  $(\nu_n^{i,\varepsilon})_{n \in \mathbb{N}_0} \subseteq \mathbb{V}([0, T])$  such that

$$\|\omega_n^i - \nu_n^{i,\varepsilon}\|_\infty \leq \varepsilon, \quad \forall n \in \mathbb{N}_0, \quad \text{and} \quad \nu_n^{i,\varepsilon} \xrightarrow{w^*} \nu_0^{i,\varepsilon} \quad \text{as } n \rightarrow \infty,$$

where the convergence ‘ $\xrightarrow{w^*}$ ’ is in the weak\* topology on  $\mathbb{V}([0, T])$ , which can be identified with the Banach dual of  $\mathbb{C}([0, T])$  under the uniform norm.

The following definition of  $S$ -convergence on  $\mathbb{D}([0, \infty); \mathbb{R}^d)$  is taken from [33].

**Definition A.2** On  $\mathbb{D}([0, \infty); \mathbb{R}^d)$ , we write  $\omega_n \rightarrow_S \omega_0$  if for every  $i \leq d$ , one can find a sequence of positive real numbers  $(T^r)_{r \in \mathbb{N}}$  increasing to  $\infty$  such that

$$[\omega_n^i]^{T^r} \rightarrow_S [\omega_0^i]^{T^r}, \quad \text{for every } r \in \mathbb{N}, \tag{A.1}$$

where  $[\omega^i]^{T^r}$  denotes the restriction of a path  $\omega^i \in \mathbb{D}([0, \infty); \mathbb{R})$  to  $\mathbb{D}([0, T^r]; \mathbb{R})$ .

A topological convergence is obtained by requiring that every subsequence admits a further  $S$ -convergent subsequence; see [33, Theorem 6.3]. The following definitions for the  $S$ -topology on the Skorokhod spaces  $\mathbb{D}([0, T]; \mathbb{R}^d)$  and  $\mathbb{D}([0, \infty); \mathbb{R}^d)$  are taken from [31] and [33], respectively.

**Definition A.3** The  $S$ -topology is the topology generated on the Skorokhod space by subsequential  $S$ -convergence.

The Skorokhod space endowed with the  $S$ -topology is known to be a Hausdorff ( $T_2$ ) space, and a stronger separation axiom is an open problem. A weak separation axiom is a well-known issue for topologies defined via subsequential convergence (Kantorovich–Vulih–Pinsker–Kisyński (KVPK) recipe); see [33, Appendix] for elaboration. The difficulties encountered in establishing the regularity of the  $S$ -topology are explained in [31, Remark 3.12].

### A.1.2 The $\Sigma$ -topology

Jakubowski's  $\Sigma$ -topology was introduced in [33]. The Skorokhod space endowed with  $\Sigma$  is a locally convex vector space. Following [33], we start by defining an auxiliary mode of convergence  $\rightarrow_\tau$  on the space

$$\mathbb{A}([0, T]; \mathbb{R}) := \mathbb{C}([0, T]; \mathbb{R}) \cap \mathbb{V}([0, T]; \mathbb{R})$$

of (bounded) continuous functions of finite variation. We write  $A_n \rightarrow_\tau A_0$  for a sequence  $(A_n)_{n \in \mathbb{N}_0} \subseteq \mathbb{A}([0, T]; \mathbb{R})$  if

$$\|A_n - A_0\|_\infty \longrightarrow 0 \quad \text{as } n \rightarrow \infty,$$

and

$$\sup_{n \in \mathbb{N}_0} \|A_n\|_\mathbb{V} < \infty,$$

where  $\|\cdot\|_\mathbb{V}$  denotes the total variation norm.

**Definition A.4** The topology  $\Sigma$  on  $\mathbb{D}([0, T]; \mathbb{R}^d)$  is the topology generated by the seminorms

$$\rho_{\mathcal{A}}^i = \sup_{A \in \mathcal{A}} \left| \int_{[0, T]} \omega^i(u) dA(u) \right|, \quad i \leq d,$$

where  $\mathcal{A}$  ranges over relatively  $\tau$ -compact subsets of  $\mathbb{A}([0, T]; \mathbb{R})$ .

**Definition A.5** The topology  $\Sigma$  on  $\mathbb{D}([0, T]; \mathbb{R}^d)$  is the topology generated by the seminorms  $\rho_T^i(\omega) = |\omega^i(T)|$ ,  $i \leq d$ , and the seminorms

$$\rho_{\mathcal{A}}^i = \sup_{A \in \mathcal{A}} \left| \int_{[0, T]} \omega^i(u) dA(u) \right|, \quad i \leq d,$$

where  $\mathcal{A}$  ranges over relatively  $\tau$ -compact subsets of  $\mathbb{A}([0, T]; \mathbb{R})$ . Note the difference between  $[0, T)$  and  $[0, T]$ .

The topology  $\Sigma$  was defined on the Skorokhod space  $\mathbb{D}([0, T]; \mathbb{R})$  for  $T = 1$  in [33]. The following properties were shown to be true for  $\Sigma$  on  $\mathbb{D}([0, T]; \mathbb{R})$ .

**Proposition A.6** *The  $\Sigma$ -topology has the following properties:*

- (i) *The Skorokhod space endowed with  $\Sigma$  is a locally convex vector space.*
- (ii) *The topology  $\Sigma$  is weaker than the topology  $S$ .*
- (iii) *A set is  $\Sigma$ -compact if and only if it is  $S$ -compact.*

*Remark A.7* It was communicated to the author by Professor Jakubowski that the properties of Proposition A.6 remain true for the infinite horizon extension of the  $\Sigma$ -topology.

### A.1.3 The Meyer–Zheng topology

The Meyer–Zheng topology, introduced in [44], is a relative topology on the image measures on the graphs  $(t, \omega(t))_{t \in [0, \infty]}$  of trajectories  $(\omega(t))_{t \in [0, \infty]}$  under the measure  $\lambda(dt) := e^{-t} dt$  (called pseudo-paths), induced by the weak topology on probability laws on the compactified space  $[0, \infty] \times \overline{\mathbb{R}}$ . We write MZ for the Meyer–Zheng topology, that is, the topology on the Skorokhod space  $\mathbb{R}(I; \mathbb{R}^d)$  generated by the coordinatewise convergence in measure; see (A.2). The following definition is adapted from [44, Lemma 1], which states that on  $\mathbb{D}([0, \infty); \mathbb{R})$ , the convergence in measure (A.2) is indeed equivalent to the convergence in the pseudo-path topology.

**Definition A.8** For  $I = [0, \infty)$ , the topology MZ on  $\mathbb{D}(I; \mathbb{R}^d)$  is the topology generated by the convergence

$$\int_I f(t, \omega_n^i(t)) \lambda(dt) \longrightarrow \int_I f(t, \omega^i(t)) \lambda(dt), \quad \forall f \in C_b(I \times \mathbb{R}), \forall i \leq d, \tag{A.2}$$

where  $\lambda(dt) := e^{-t} dt$ .

On  $\mathbb{D}([0, T]; \mathbb{R}^d)$ , we additionally require the convergence of the terminal value; see (A.3) below. Without this addition, the topology is not a Hausdorff topology on  $\mathbb{D}([0, T]; \mathbb{R}^d)$ .

**Definition A.9** For  $I = [0, T]$ , the topology MZ on  $\mathbb{D}(I; \mathbb{R}^d)$  is the topology generated by the convergence (A.2) in conjunction with the convergence

$$\omega_n(T) \longrightarrow \omega(T). \tag{A.3}$$

The key lemma of Meyer and Zheng [44, Lemma 1] extends to  $I = [0, T]$  for  $T$  finite and  $d > 1$  via a simple iterative argument; cf. Sect. 4.2.1.

**Lemma A.10** *Let  $(\omega_n)_{n \in \mathbb{N}}$  and  $\omega$  be paths in  $\mathbb{D}(I; \mathbb{R}^d)$  such that  $\omega_n \rightarrow_{MZ} \omega$ . Then  $\omega_n^i \rightarrow_\lambda \omega^i$  for every  $i \leq d$ . Moreover, there exist a subsequence  $(\omega_{n_k})$  and a set  $L \subseteq I$  of full Lebesgue measure such that  $T \in L$  if  $I = [0, T]$  and  $\omega_{n_k}^i(t) \rightarrow \omega^i(t)$  for every  $i \leq d$  and every  $t \in L$ . In particular, there exists a (countable) dense set  $D \subseteq I$  such that  $T \in D$  if  $I = [0, T]$  and  $\omega_{n_k}^i(t) \rightarrow \omega^i(t)$  for every  $i \leq d$  and every  $t \in D$ .*

*Proof* Let  $\omega_n \rightarrow_{\text{MZ}} \omega$ . By the definition (A.2), we have

$$\int_I f(t, \omega_n^i(t)) \lambda(dt) \longrightarrow \int_I f(t, \omega^i(t)) \lambda(dt), \quad \forall f \in \mathbb{C}_b(I \times \mathbb{R}^d), \forall i \leq d,$$

where the measure  $\lambda(dt) = e^{-t} dt$  is equivalent to Lebesgue measure on  $I$ . By taking  $f(t, x) := \alpha(t) \arctan(x)$ ,  $\alpha \in \mathbb{C}_b(I)$ , we deduce that the uniformly bounded sequence  $u_n^i := \arctan(\omega_n^i)$ ,  $n \in \mathbb{N}$ , converges to  $u^i := \arctan(\omega^i)$  in the weak topology of  $L^2(\lambda)$ , for every  $i \leq d$ . Then, by taking  $f(t, x) := \alpha(t) |\arctan(x)|^2$ ,  $\alpha \in \mathbb{C}_b(I)$ , we deduce that  $\|u_n^i\|_{L^2(\lambda)} \rightarrow \|u^i\|_{L^2(\lambda)}$  as  $n \rightarrow \infty$ , and hence  $(u_n^i)_{n \in \mathbb{N}}$  converges strongly in  $L^2(\lambda)$  to  $u^i$ , for every  $i \leq d$ . Indeed, in a Hilbert space, strong convergence is equivalent to weak convergence plus convergence of norms. Consequently,  $(\omega_n^i)$  converges in  $\lambda$ -measure to  $\omega^i$  in  $I$ , i.e.,  $\omega_n^i \rightarrow_\lambda \omega^i$ , for every  $i \leq d$ . Thus for  $i = 1$ , there exists a subsequence  $(\omega_{n_\ell})_{\ell \in \mathbb{N}} = (\omega_{n_\ell}^1, \dots, \omega_{n_\ell}^d)_{\ell \in \mathbb{N}}$  of  $(\omega_n)_{n \in \mathbb{N}}$  such that

$$\omega_{n_\ell}^1(t) \longrightarrow \omega^1(t) \tag{A.4}$$

for every  $t$  in some set  $L_1$  of full Lebesgue measure. By dominated convergence, we have

$$\int_I f(t, \omega_{n_\ell}^i(t)) \lambda(dt) \longrightarrow \int_I f(t, \omega^i(t)) \lambda(dt), \quad \forall f \in \mathbb{C}_b(I \times \mathbb{R}^d), \forall i \leq d.$$

Replacing  $i = 1$  with  $i = 2$  and  $(\omega_n)_{n \in \mathbb{N}}$  with  $(\omega_{n_\ell})_{\ell \in \mathbb{N}}$  before (A.4), we obtain a further subsequence  $(\omega_{n_{\ell m}})_{m \in \mathbb{N}} = (\omega_{n_{\ell m}}^1, \dots, \omega_{n_{\ell m}}^d)_{m \in \mathbb{N}}$  of  $(\omega_n)_{n \in \mathbb{N}}$  and a set  $L_2$  of full Lebesgue measure such that  $\omega_{n_{\ell m}}^2(t) \rightarrow \omega^2(t)$  for every  $t \in L_2$ . We therefore have

$$\omega_{n_{\ell m}}^1(t) \longrightarrow \omega^1(t) \quad \text{and} \quad \omega_{n_{\ell m}}^2(t) \longrightarrow \omega^2(t)$$

for every  $t \in L_1 \cap L_2$ , where the set  $L_1 \cap L_2$  is of full Lebesgue measure. By repeating the argument  $d - 2$  more times, we obtain a set  $L := L_1 \cap L_2 \cap \dots \cap L_d$  and a subsequence  $(\omega_{n_k})_{k \in \mathbb{N}} = (\omega_{n_k}^1, \dots, \omega_{n_k}^d)_{k \in \mathbb{N}}$  such that

$$\omega_{n_k}^i(t) \longrightarrow \omega^i(t), \quad \forall i \leq d,$$

for every  $t \in L$ , where the set  $L$  is of full Lebesgue measure. Moreover, by (A.3) for  $I = [0, T]$ , we have  $\omega_n(T) \rightarrow \omega(T)$ ; so the set  $L$  can be chosen to contain the terminal time  $T$  as well. The complement of  $L$  is a  $\lambda$ -nullset, and so the set  $L$  contains a (countable) dense set  $D$  such that  $T \in D$  if  $I = [0, T]$ .  $\square$

**Corollary A.11** *We have  $\mathcal{F}_{t-}^0 := \sigma(\omega(s) : s \in [0, t]) = \sigma(\omega(s) : s \in D \cap [0, t])$  for any countable dense subset  $D$  of  $[0, t]$  and every  $t \leq T$ . Moreover, we have*

$$\mathcal{F}_t^0 = \sigma(\mathcal{G}_t^0, \omega(t)), \quad t \leq T,$$

where  $\mathcal{G}_t^0$  denotes the  $\sigma$ -algebra generated by the family of  $\mathcal{F}_t^0$ -measurable MZ-continuous functions.



*Proof* Let  $\mathcal{G}_t^o$  denote the  $\sigma$ -algebra generated by the family of all  $\mathcal{F}_t^0$ -measurable MZ-continuous functions. We have  $\mathcal{G}_t^o \subseteq \mathcal{F}_t^0$  and  $\mathcal{F}_{t-}^0 \subseteq \mathcal{G}_t^o$  from Lemma A.10. Moreover, we have

$$\omega^i(t) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_0^\varepsilon \omega^i(t+u)du, \quad i \leq d, \varepsilon < T-t,$$

where each  $\omega \mapsto \frac{1}{\varepsilon} \int_0^\varepsilon \omega^i(t+u)du$  is an MZ-continuous function. Thus the assertion follows. □

**Lemma A.12** *The mappings*

$$\omega \mapsto \|\omega\|_\infty \quad \text{and} \quad \omega \mapsto N^{a,b}(\omega^i), \quad a < b, i \leq d,$$

*are MZ-lower semicontinuous.*

*Proof* The proof is adapted from [44]. Let  $i \leq d$  and  $\omega_n^i \rightarrow_{\text{MZ}} \omega^i$  with  $\sup_n \|\omega_n^i\| \leq c$ . If  $\|\omega^i\|_\infty > c$ , there either exists  $s < t$  such that  $\omega^i(u) > c$  for all  $u \in [s, t)$ , or we have  $\omega^i(T) > c$ . In both cases, there exists an MZ-continuous function  $F$  for which  $\lim_{n \rightarrow \infty} F(\omega_n^i) < F(\omega^i)$ , cf. (A.2) and (A.3), which is a contradiction. Thus the mapping  $\omega \mapsto \|\omega\|_\infty := \|\omega^1\|_\infty \vee \dots \vee \|\omega^d\|_\infty$  is MZ-lower semicontinuous. Similarly, one can show that the sets of the form  $\{\omega : \exists u \in [s, t) \text{ such that } \omega^i(u) > b\}$  and  $\{\omega : \exists u \in [s, t) \text{ such that } \omega^i(u) < a\}$ ,  $s < t, a < b$ , are open in the MZ-topology, from which the MZ-lower semicontinuity of the mappings  $N^{a,b}$ ,  $a < b$ , follows. Indeed, let  $a < b$  be fixed and consider a finite partition  $\pi := \{t_0 < t_1 < \dots < t_n\}$  of  $[0, t_n]$ . We write  $N_\pi^{a,b}(\omega^i) \geq k$  if one can find

$$\ell_1 < m_1 \leq \ell_2 < m_2 \leq \dots < \ell_k < m_k \leq n$$

such that for all  $j < k$ ,  $\omega^i(s) < a$  for some  $s \in [t_{\ell_{j-1}}, t_{\ell_j})$  and  $\omega^i(t) > b$  for some  $t \in [t_{m_{j-1}}, t_{m_j})$  (or, for  $t = T$ , if  $j = k$  and  $m_k = n$ ). The partition  $\pi$  is finite; so the sets

$$\{\omega : N_\pi^{a,b}(\omega^i) \geq k\} = \{\omega : N_\pi^{a,b}(\omega^i) > k - 1\}, \quad k \in \mathbb{N},$$

are open in the MZ-topology. Consequently, the mapping  $\omega \mapsto N_\pi^{a,b}(\omega^i)$  and the mapping  $N^{a,b}(\omega^i) := \sup_\pi N_\pi^{a,b}(\omega^i)$  are MZ-lower semicontinuous for every  $i \leq d$ . □

We refer the reader to the book by Dellacherie and Meyer [13, IV. 40–46] and the paper [44] by Meyer and Zheng for details on pseudo-paths and the Meyer–Zheng topology, respectively.

### A.1.4 The Skorokhod $J^1$ -topology

The following *complete* metric, generating a topology called Skorokhod’s  $J^1$ -topology, was introduced by Kolmogorov in [39]. The metric introduced by Skorokhod himself in [51] is not complete despite generating an equivalent topology.

**Definition A.13** The Skorokhod  $J^1$ -topology on  $\mathbb{D}([0, T]; \mathbb{R}^d)$  is the topology generated by the complete metric

$$J_T^1(\omega, \tilde{\omega}) := \inf_{\lambda \in \Lambda} \left\{ \sup_{s < t} \left| \log \frac{\lambda t - \lambda s}{t - s} \right| \vee \|\omega - \tilde{\omega} \circ \lambda\|_\infty \right\},$$

where  $\Lambda$  denotes the class of strictly increasing, continuous mappings of  $[0, T]$  onto itself.

**Definition A.14** The Skorokhod  $J^1$ -topology on  $\mathbb{D}([0, \infty); \mathbb{R}^d)$  is the topology generated by the complete metric

$$J^1(\omega, \tilde{\omega}) := \sum_{r=1}^{\infty} 2^{-r} (1 \wedge J_r^1([\omega]^r, [\tilde{\omega}]^r)), \quad (\text{A.5})$$

where  $[\omega]^r$  denotes the restriction of  $\omega$  to  $[0, r]$ .

Let us recall from Billingsley [5, Sects. 12, 16] the criterion for relative compactness in the  $J^1$ -topology. Let  $\delta > 0$  and denote

$$m_\delta(\omega) := \inf_{t_i \in \pi} \max_{i \leq n} \sup_{s, t \in (t_{i-1}, t_i]} |\omega(s) - \omega(t)|, \quad \omega \in \mathbb{D}([0, T]; \mathbb{R}^d),$$

where the infimum is taken over all finite partitions  $0 = t_0 < t_1 < \dots < t_n = T$  of  $[0, T]$  with mesh size  $t_i - t_{i-1} > \delta$  for all  $i \leq n$ .

**Lemma A.15** A subset  $K$  of  $\mathbb{D}([0, T]; \mathbb{R}^d)$ ,  $T < \infty$ , is relatively  $J^1$ -compact if and only if it is bounded and

$$\lim_{\delta \rightarrow 0} \sup_{\omega \in K} m_\delta(\omega) = 0.$$

A subset  $K$  of  $\mathbb{D}([0, \infty); \mathbb{R}^d)$  is relatively  $J^1$ -compact if and only if the restriction of  $K$  to  $\mathbb{D}([0, T]; \mathbb{R}^d)$  is relatively  $J^1$ -compact for every  $T < \infty$ .

We refer the reader to [5, Sects. 12, 16] for details on the Skorokhod  $J^1$ -metric on  $\mathbb{D}([0, T]; \mathbb{R}^d)$  and  $\mathbb{D}([0, \infty); \mathbb{R}^d)$ , respectively.

## A.2 The proof of Lemma 2.2

Recall that we claimed that there exists a uniform constant  $b > 0$  such that for any  $Q \in \mathbb{P}(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T)$ ,  $H \in \mathcal{E}(Q)$  and  $c > 0$ , we have

$$Q[|(H \bullet X)_t| > c] \leq \frac{b}{c} \left( E_Q[|X_t|] + \sup_{H \in \mathcal{E}(Q)} E_Q[(H \bullet X)_t] \right), \quad t \in I. \quad (\text{A.6})$$

*Proof* The inequality (A.6) is a generalisation of Burkholder's inequality which states that there exists a uniform constant  $a > 0$  such that for any  $H \in \mathcal{E}(Q)$ , any

$Q$ -martingale  $M$  and  $c > 0$ , we have

$$Q[|H \bullet M|_t > c] \leq \frac{a}{c} E_Q[|M_t|], \quad t \in I; \tag{A.7}$$

see e.g. [43, Theorem 47] or [4] for a proof of (A.7). For a fixed  $Q \in \mathbb{P}(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T)$  and  $H \in \mathcal{E}(Q)$ , by [14, Appendix 2, 3], we have

$$\sup_{H \in \mathcal{E}(Q)} E_Q[(H \bullet X)_t] = \sum_{i=1}^d \text{Var}_t^Q(X^i), \quad t \in I. \tag{A.8}$$

Let us fix  $i \leq d$  and assume that  $E_Q[|X_t^i|] + \text{Var}_t^Q(X^i)$  is finite, as otherwise the result is trivial. Let  $t_0^i < t_1^i < \dots < t_n^i = t$  and  $H = (H^i)_{i=1}^d$  be an element of  $\mathcal{E}(Q)$ , i.e.,

$$H^i = \sum_{k=1}^n H_{t_k^i}^i \mathbb{1}_{(t_{k-1}^i, t_k^i]}, \quad i \leq d,$$

where each  $|H_{t_k^i}^i| \leq 1$  is  $\mathcal{F}_{t_k^i}$ -measurable; cf. (2.1). Consider the Doob decomposition

$$X_{t_k^i}^i = M_{t_k^i}^i + A_{t_k^i}^i, \quad k = 1, 2, \dots, n,$$

where  $A_{t_k^i}^i = \sum_{j=1}^k E_Q[X_{t_j^i}^i - X_{t_{j-1}^i}^i | \mathcal{F}_{t_{j-1}^i}]$  and  $M^i$  is a  $Q$ -martingale on the index set  $\{t_0^i, t_1^i, \dots, t_n^i\}$ . We have

$$Q[|(H^i \bullet A^i)_t| > c] \leq \frac{1}{c} E_Q[|(H^i \bullet A^i)_t|] \leq \frac{1}{c} \text{Var}_t^Q(X^i), \quad i \leq d. \tag{A.9}$$

Similarly, for  $M^i$ , we have

$$E_Q[|M_t^i|] \leq E_Q[|X_t^i| + |A_t^i|] \leq E_Q[|X_t^i|] + \text{Var}_t^Q(X^i).$$

Hence by (A.7), we have

$$Q[|(H^i \bullet M^i)_t| > c] \leq \frac{a}{c} (E_Q[|X_t^i|] + \text{Var}_t^Q(X^i)), \quad i \leq d. \tag{A.10}$$

Combining (A.8)–(A.10), we get for  $H \in \mathcal{E}(Q)$ ,  $M = (M^i)_{i=1}^d$  and  $A = (A^i)_{i=1}^d$  that

$$\begin{aligned} Q[|(H \bullet X)_t| > c] &\leq Q[|(H \bullet M)_t + (H \bullet A)_t| > c] \\ &\leq \sum_{i=1}^d Q\left[|(H^i \bullet M^i)_t| > \frac{c}{2d}\right] + \sum_{i=1}^d Q\left[|(H^i \bullet A^i)_t| > \frac{c}{2d}\right] \\ &\leq \frac{2ad}{c} \sum_{i=1}^d (E_Q[|X_t^i|] + \text{Var}_t^Q(X^i)) + \frac{2d}{c} \sum_{i=1}^d \text{Var}_t^Q(X^i) \\ &\leq \frac{b}{c} \left( E_Q[|X_t|] + \sup_{H \in \mathcal{E}(Q)} E_Q[(H \bullet X)_t] \right), \quad t \in I, c > 0, \end{aligned}$$

where  $b := 2(a + 1)d$ . The proof for the filtration  $(\mathcal{F}_t^0)_{t \in I}$  is completely analogous.  $\square$

## References

1. Alexandroff, A.D.: Additive set functions in abstract spaces. *Mat. Sb.* **51**(NS 9), 563–621 (1941)
2. Bahlali, K., Gherbal, B., Mezerdi, B.: Existence of optimal controls for systems driven by FBSDEs. *Syst. Control Lett.* **60**, 344–349 (2011)
3. Beiglböck, M., Henry-Labordère, P., Penkner, F.: Model-independent bounds for option prices—a mass transport approach. *Finance Stoch.* **17**, 477–501 (2013)
4. Bichteler, K.: Stochastic integration and  $L^p$ -theory of semimartingales. *Ann. Probab.* **9**, 49–89 (1981)
5. Billingsley, P.: *Convergence of Probability Measures*. Wiley, New York (1968)
6. Black, F., Scholes, M.: The pricing of options and corporate liabilities. *J. Polit. Econ.* **81**, 637–654 (1973)
7. Bogachev, V.I.: *Gaussian Measures*. Am. Math. Soc., Providence (1998)
8. Bogachev, V.I.: *Measure Theory*. Springer, Berlin (2007)
9. Bogachev, V.I., Smolyanov, O.: *Topological Vector Spaces and Their Applications*. Springer, Berlin (2017)
10. Buck, R.C.: Bounded continuous functions on a locally compact space. *Mich. Math. J.* **5**, 95–104 (1958)
11. Chan, N.H., Zhang, R.M.: Inference for unit-root models with infinite variance GARCH errors. *Stat. Sin.* **20**, 1363–1393 (2010)
12. Cheridito, P., Kiiski, M., Prömel, D.J., Soner, H.M.: Martingale optimal transport duality. *Math. Ann.* (2020). <https://doi.org/10.1007/s00208-019-01952-y>
13. Dellacherie, C., Meyer, P.-A.: *Probabilities and Potential*. North-Holland Mathematics Studies, vol. 29 (1978)
14. Dellacherie, C., Meyer, P.-A.: *Probabilities and Potential B. Theory of Martingales*. North-Holland Mathematics Studies, vol. 72 (1982)
15. Dolinsky, Y., Soner, H.M.: Martingale optimal transport in the Skorokhod space. *Stoch. Process. Appl.* **125**, 3893–3931 (2015)
16. Engelking, R.: *General Topology*. Heldermann, Berlin (1977)
17. Feldman, M.B.: A proof of Lusin’s theorem. *Am. Math. Mon.* **88**, 191–192 (1981)
18. Fernique, X.: Processus linéaires, processus généralisés. *Ann. Inst. Fourier* **17**, 1–92 (1967)
19. Florescu, L.C., Godet-Thobie, C.: *Young Measures and Compactness in Measure Spaces*. de Gruyter, Berlin (2012)
20. Floret, K.: *Weakly Compact Sets*. Lecture Notes in Mathematics, vol. 801. Springer, Berlin (1980)
21. Föllmer, H., Schied, A.: *Stochastic Finance: An Introduction in Discrete Time*, 2nd edn. de Gruyter, Berlin (2011)
22. Giles, R.: A generalization of the strict topology. *Trans. Am. Math. Soc.* **161**, 467–474 (1971)
23. Gillman, L., Jerison, M.: *Rings of Continuous Functions*. Springer, New York (1976)
24. Guo, G., Tan, X., Touzi, N.: Tightness and duality of martingale transport on the Skorokhod space. *Stoch. Process. Appl.* **127**, 927–956 (2017)
25. He, S.W., Wang, J.G., Yan, J.A.: *Semimartingale Theory and Stochastic Calculus*. Kexue Chubanshe (Science Press), Beijing (1992)
26. Herrlich, H.: Wann sind alle stetigen Abbildungen in  $Y$  konstant? *Math. Z.* **90**, 152–154 (1965)
27. Hirsch, F., Roynette, B., Yor, M.: Kellerer’s theorem revisited. In: Dawson, D., et al. (eds.) *Asymptotic Laws and Methods in Stochastics. A Volume in Honour of Miklós Csörgő*. Fields Institute Communications, vol. 76, pp. 347–363. Springer, New York (2015)
28. Hoffmann-Jørgensen, J.: A generalization of the strict topology. *Math. Scand.* **30**, 313–323 (1972)
29. Jacod, J., Shiryaev, A.N.: *Limit Theorems for Stochastic Processes*. Springer, Berlin (1987)
30. Jakubowski, A.: Sequential topologies in probability theory: Four papers on functional convergence. Preprint (1995). Available online at <https://kpbk.umk.pl/Content/39961/PDF/4pap.pdf>
31. Jakubowski, A.: A non-Skorokhod topology on the Skorokhod space. *Electron. J. Probab.* **2**(4), 1–21 (1997)
32. Jakubowski, A.: The almost sure Skorokhod representation for subsequences in nonmetric spaces. *Theory Probab. Appl.* **42**, 167–174 (1998)

33. Jakubowski, A.: New characterizations of the  $S$  topology on the Skorokhod space. *Electron. Commun. Probab.* **23**(2), 1–16 (2018)
34. Jakubowski, A., Mémin, J., Pages, G.: Convergence en loi des suites d'intégrales stochastiques sur l'espace  $\mathbb{D}^1$  de Skorokhod. *Probab. Theory Relat. Fields* **81**, 111–137 (1989)
35. Jarchow, H.: *Locally Convex Spaces*. Teubner, Leipzig (1981)
36. Kellerer, H.G.: Markov-Komposition und eine Anwendung auf Martingale. *Math. Ann.* **198**, 99–122 (1972)
37. Kelley, J.L., Namioka, I.: *Linear Topological Spaces*. Springer, Berlin (1976)
38. Knowles, J.D.: Measures on topological spaces. *Proc. Lond. Math. Soc.* **s3-17**, 139–156 (1967)
39. Kolmogorov, A.N.: On the Skorokhod convergence. *Teor. Veroâtn. Primen.* **1**, 239–247 (1956)
40. Kurtz, T.: Random time changes and convergence in distribution under the Meyer–Zheng conditions. *Ann. Probab.* **19**, 1010–1034 (1991)
41. Kurtz, T., Stockbridge, R.: Stationary solutions and forward equations for controlled and singular martingale problems. *Electron. J. Probab.* **6**(17), 1–52 (2001)
42. Lowther, G.: Limits of one-dimensional diffusions. *Ann. Probab.* **37**, 78–106 (2009)
43. Meyer, P.A.: *Martingales and Stochastic Integrals*. Lecture Notes in Mathematics, vol. 284. Springer, Berlin (1972)
44. Meyer, P.A., Zheng, W.A.: Tightness criteria for laws of semimartingales. *Ann. Inst. Henri Poincaré Probab. Stat.* **20**, 353–372 (1984)
45. Pham, H.: *Continuous-Time Stochastic Control and Optimization with Financial Applications*. Springer, Berlin (2009)
46. Prigent, J.L.: *Weak Convergence of Financial Markets*. Springer, Berlin (2003)
47. Protter, P.E.: *Stochastic Integration and Differential Equations: A New Approach*, 2nd edn. Springer, Berlin (2005)
48. Revuz, D., Yor, M.: *Continuous Martingales and Brownian Motion*, 3rd edn. Springer, Berlin (2013)
49. Schwartz, L.: *Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures*. Oxford University Press, London (1973)
50. Senteilles, F.D.: Bounded continuous functions on a completely regular space. *Trans. Am. Math. Soc.* **168**, 311–336 (1972)
51. Skorokhod, A.V.: Limit theorems for stochastic processes. *Theory Probab. Appl.* **1**, 261–290 (1956)
52. Strassen, V.: The existence of probability measures with given marginals. *Ann. Math. Stat.* **36**, 423–439 (1965)
53. Stricker, C.: Semimartingales et valeur absolue. In: Dellacherie, C., et al. (eds.) *Séminaire de Probabilités 13*. LNM, vol. 721, pp. 472–477. Springer, Berlin (1979)
54. Stricker, C.: Sur la caractérisation des semimartingales. In: Azéma, J., Yor, M. (eds.) *Séminaire de Probabilités 15*. LNM, vol. 850, pp. 523–525. Springer, Berlin (1981)
55. Stricker, C.: Lois de semimartingales et critères de compacité. In: Azéma, J., Yor, M. (eds.) *Séminaire de Probabilités 19*. LNM, vol. 1123, pp. 209–217. Springer, Berlin (1985)
56. Tan, X., Touzi, N.: Optimal transportation under controlled stochastic dynamics. *Ann. Probab.* **41**, 3201–3240 (2013)
57. van der Vaart, A.W.: *Asymptotic Statistics*. Cambridge University Press, Cambridge (1998)
58. Zheng, W.A.: Tightness results for laws of diffusion processes – application to stochastic mechanics. *Ann. Inst. Henri Poincaré Probab. Stat.* **21**, 103–124 (1985)