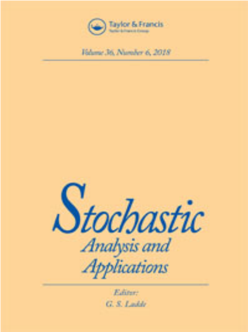

Stochastic Analysis and Applications

ISSN: 0736-2994 (Print) 1532-9356 (Online) Journal homepage: https://www.tandfonline.com/loi/lsaa20

Support of solutions of stochastic differential equations in exponential Besov–Orlicz spaces

Martin Ondreját, Prokop Šimon & Michal Kupsa

To cite this article: Martin Ondreját, Prokop Šimon & Michal Kupsa (2018) Support of solutions of stochastic differential equations in exponential Besov–Orlicz spaces, Stochastic Analysis and Applications, 36:6, 1037-1052, DOI: 10.1080/07362994.2018.1524304

To link to this article: https://doi.org/10.1080/07362994.2018.1524304

Published online: 18 Oct 2018.

STOCHASTIC ANALYSIS AND APPLICATIONS
2018, VOL. 36, NO. 6, 1037–1052
https://doi.org/10.1080/07362994.2018.1524304

Support of solutions of stochastic differential equations in exponential Besov–Orlicz spaces

Martin Ondreját, Prokop Šimon, and Michal Kupsa

Institute of Information Theory and Automation, The Czech Academy of Sciences, Praha 8, Czech Republic

ABSTRACT

The Besov–Orlicz space $B^{1/2}_{\Phi,\infty}(0,T;\mathbb{R}^d)$ with $\Phi(x)=\exp(x^2)-1$ is currently the smallest known classical function space to which paths of the Wiener process belong almost surely. We consider stochastic differential equations with no global growth condition on the non-linearities and we describe the topological support of the laws of trajectories of the solutions in every Polish subspace of continuous functions into which the Besov–Orlicz space $B^{1/2}_{\Phi,\infty}(0,T;\mathbb{R}^d)$ is embedded compactly.

ARTICLE HISTORY
Received 18 June 2018
Accepted 10 September 2018

KEYWORDS
Stochastic differential equation; topological support; Besov–Orlicz space

2010 MATHEMATICS SUBJECT CLASSIFICATION
60H10

1. Introduction

Let us consider a Stratonovich stochastic differential equation

$$dX = b(X)\ dt + \sigma(X)\circ dW, \quad X(0)=\xi \tag{1.1}$$

where $b:\mathbb{R}^d\to\mathbb{R}^d, \sigma:\mathbb{R}^d\to\mathbb{R}^d\otimes\mathbb{R}^m$, W is an $\mathbb{R}^m$-valued Wiener process and $\xi\in\mathbb{R}^d$. In the seminal paper [1] it was shown that if Y denotes the path space $C([0,T];\mathbb{R}^d)$ then

$$\text{the support of the law of } X \text{ in the path space } Y = \overline{\{x^w : w\in L^2(0,T;\mathbb{R}^m)\}}^Y \tag{1.2}$$

provided that b is bounded and Lipschitz continuous, σ is C^2-smooth and σ, σ' and σ'' are bounded, where by x^w the solution to the ordinary differential equation

$$x' = b(x)+\sigma(x)w, \quad x(0)=\xi \tag{1.3}$$

is denoted.

If Equation (1.2) is established in a path space Y smaller and finer than $C([0,T];\mathbb{R}^d)$ then one gets more precise information about the law of the solution of Equation (1.1); a considerable attention has been paid to this problem as well as to weakening hypotheses on b and σ. Before we provide an overview of known results, we mention that the same characterization was proved for stochastic equations driven by little more general semimartingales S of the form $dS = Adt + QdW$ where W is a Wiener process and the

CONTACT Martin Ondreját ondrejat@utia.cas.cz Institute of Information Theory and Automation, The Czech Academy of Sciences, Pod Vodárenskou věží 4, Praha 8 CZ-182 08, Czech Republic.

Table 1. Survey.

Year	Ref.	Drift term b	Diffusion term σ	State space	Global existence
1972	[1]	Bounded Lipschitz	C_b^2	C	Implied
1985	[3]	Not present	C_b^3	C	Implied
1986	[4]	C_b^3	C_b^3	C	Implied
1990	[2]	Lipschitz	Lipschitz & C^2	C	Implied
1994	[8]	Bounded Lipschitz & C^∞	$C_b^1 \cap C^\infty$	$C^{(1/2)-}$	Implied
1994	[14]	Bounded Lipschitz	C_b^2	$C^{(1/2)-}$	Implied
1994	[13]	Lipschitz	C_b^2	$B_{\Phi,\infty}^{\lvert r \log r \rvert^{1/2}}$	Implied
1995	[11]	Locally mononotone Semimononotone Continuous	Linear growth & $C^{1,1}$	$(B_{p,\infty}^{1/2} \cap C^{\lvert r \log r \rvert^{1/2}})^-$	Implied
2018		Locally mononotone	$C^{1,1}$	$(B_{\Phi,\infty}^{1/2})^-$	Assumed

Here $C^{(1/2)-}$ means C^α for some $\alpha < 1/2$ and the symbol $(Z)^-$ stands for *"for every Banach space Y such that Z is embedded in Y compactly."*

processes A, Q and Q^{-1} are typically bounded, see e.g. [2–4] for precise assumptions, and we also refer the reader to [5] for support theorems for stochastic equations driven by general Gaussian and Markov processes, approached via the rough paths theory.

In the table below and the comments following it, we list results on the support of the law of the solution to Equation (1.1) from [1–14]. First we have to introduce the function spaces used in these works, although we do not need most of them in the sequel, to have the possibility to compare the results. We will denote by C^k the space of k-times differentiable functions, by C_b^k the space of continuous bounded k-times differentiable functions such that all derivatives up to the order k are continuous and bounded, by $C^{k,\alpha}$ the space of k-times differentiable functions such that k^{th}-derivatives are locally α-Hölder continuous and by $C_b^{k,\alpha}$ the space of continuous bounded k-times differentiable functions such that all derivatives up to the order k are continuous and bounded and the k^{th}-derivatives are globally α-Hölder continuous. If h is a continuous positive function on some interval $(0, \varepsilon]$ such that $h(0+) = 0$, the generalized Hölder space C^h corresponding to the modulus of continuity h, with the norm

$$||f||_{C^h([0,T];\mathbb{R}^d)} = \sup_{t \in [0,T]} |f(t)| + \sup_{0 \le s < t \le T \wedge \varepsilon} \frac{|f(t) - f(s)|}{h(t-s)} \tag{1.4}$$

will be considered. In particular, if $h(t) = t^\alpha$ for some $\alpha \in (0, 1)$, then we get the classical Hölder space C^α. If $A : [0, \infty) \to [0, \infty)$ is a continuous, non-decreasing, convex function such that $A(0) = 0$ and $A(\infty-) = \infty$ then we denote by L^A the Orlicz space with the Luxemburg norm

$$||f||_{L^A(0,T;\mathbb{R}^d)} = \inf \left\{ \lambda > 0 : \int_0^T A\left(\lambda^{-1}|f(s)|\right) ds \le 1 \right\}.$$

Only two cases of A will be needed in the sequel. First $A(x) = \Phi(x) = \exp\ (x^2) - 1$ and second $A(t) = t^p, p \in [1, \infty)$. In the latter case, $L^A(0, T; \mathbb{R}^d)$ is the standard Lebesgue space $L^p(0, T; \mathbb{R}^d)$ with the norm

$$||f||_{L^p(0,T;\mathbb{R}^d)} = \left(\int_0^T |f(s)|^p ds \right)^{\frac{1}{p}}.$$

$$
\begin{array}{ccccccccc}
 & & B^{1/2}_{\Phi,\infty} & \rightarrow & B^{1/2}_{p,\infty} \cap C^{|r\log r|^{1/2}} & \twoheadrightarrow & B^{|\log r|^{\gamma} r^{1/2}}_{p,\infty} \cap C^{|\log r|^{\beta} r^{1/2}} & & \\
 & \nearrow & & & & & & \searrow\!\!\!\!\searrow & \\
C^{1/2} & & \downarrow\!\!\!\downarrow & & \downarrow\!\!\!\downarrow & & \downarrow\!\!\!\downarrow & & C^{\alpha} \\
 & \searrow\!\!\!\!\searrow & & & & & & \nearrow\!\!\!\!\nearrow & \\
 & & B^{r^{1/2}|\log r|^{\gamma}}_{\Phi,\infty} & \twoheadrightarrow & B^{|r\log r|^{1/2}}_{\Phi,\infty} & \rightarrow & C^{|\log r| r^{1/2}} & &
\end{array}
$$

but

$$\bigcap_{1\le q<\infty} B^{1/2}_{q,\infty} \cap C^{|r\log r|^{1/2}} \not\subseteq B^{r^{1/2}|\log r|^{\gamma}}_{\Phi,\infty}.$$

Figure 1. Embeddings.

Now modulus Besov–Orlicz spaces $B^h_{A,\infty}$, with h as above, can be introduced as spaces with the norm

$$||f||_{B^h_{A,\infty}\left(0,T;\mathbb{R}^d\right)} = ||f||_{L^A\left(0,T;\mathbb{R}^d\right)} + \sup_{0<t\le T\wedge\varepsilon} \frac{||f(t+\cdot)-f(\cdot)||_{L^A\left([0,T-t];\mathbb{R}^d\right)}}{h(t)}. \tag{1.5}$$

The dependence of the norms in Equations (1.4) and (15.5) on ε is not important (the norms are equivalent for different ε). Let us realize that if $A(t) = t^p$ and $h(t) = t^\alpha$ for some $p \in [1,\infty)$ and $\alpha \in (0,1)$ then $B^h_{A,\infty}$ is the classical Besov space $B^\alpha_{p,\infty}$.

Table 1 surveys chronologically, as far as we know, the progress on sufficient conditions that guarantee the equality (11.2), whereas the last line is referring to the present paper (we do not list those papers from the bibliogrpahy that provided different proofs of already known results). To better understand the table, e.g. that

$$\text{present paper} \quad \Rightarrow \quad [12] \quad \Rightarrow \quad [19]$$

but [11] does not cover the present paper, realize that the embeddings in Figure 1 hold for every $p \in [1,\infty), \alpha \in [0,1/2), \beta \in (1/2,1), \gamma \in (0,1/2)$ (here $\rightarrow$ and $\twoheadrightarrow$ denote continuous non-compact and compact embeddings respectively), cf. Corollary 5.5 and Example 3.5.

The goal of the present paper is twofold. Firstly, we want to prove the support theorem for equations with minimal regularity and no global growth assumption on b and σ, e.g. for broken polynomials, and secondly, as already indicated in the table, we want to prove it for any path space into which the Besov–Orlicz space $B^{1/2}_{\Phi,\infty}(0,T;\mathbb{R}^d)$ is compactly embeded; note that $B^{1/2}_{\Phi,\infty}(0,T;\mathbb{R}^d)$ seems to be the smallest Banach space such that the paths of a Wiener process are known to belong to almost surely (see e.g. [15] or [16]). Apparently, to achieve the former goal, we need to separate the problem of existence of solutions from the problem of characterization of the topological support of the solutions, and so our assumptions on b and σ do not imply existence of solutions (which must be thus assumed). In particular, we generalize all the above cited papers [1–5, 8–14].

As far as the proof of the main result is concerned, we extend the idea of approximations of SDEs from [2] to exponential Besov–Orlicz path spaces using a compactness argument as in [11]. A Lenglart-type Lemma 4.1 for continuous local martinagales is our main tool. We also modify the method to avoid global growth assumptions on the non-linearities (as these are imposed in the papers [1–5, 8, 10–14]). Then we proceed via a change-of-measure argument as in [4] and hence, we could consider equations driven by semimartingales S of the form $dS = Adt + QdW$ as in [4]. Yet we restrict ourselves to $S = W$ for simplicity.

We add, for completeness, that the seminal paper [1] was extended to stochastic equations in Hilbert spaces (but not to SPDEs) by [7], that a description of the support of laws of solutions to stochastic ordinary Stratonovich equations with smooth non-linearities with bounded derivatives of order 1 (and more) was also treated in [6] and that support theorems were proved, using the theory of rough paths, also in [9, 10] in modulus Hölder path spaces C^h for every modulus of continuity h such that

$$\lim_{\varepsilon\to 0+} \frac{h(\varepsilon)}{|\varepsilon \log \varepsilon|^{1/2}} = \infty. \tag{1.6}$$

The interesting proofs in [5, 9, 10, 12] based on rough paths, in particular on the continuity of the It Lyons map (which exists also for various different metrics), are shorter comparing to the standard approach but do not reach the generality of the path space considered in the earlier paper [11].

2. General notation and conventions, part I

We convene that, throughput this paper,

- every filtration on a given probability space is assumed to be complete, i.e. every σ-algebra in the filtration contains all measurable sets of zero probability,
- $L^\Phi(a,b;\mathbb{R}^d)$ is the Orlicz with the Luxemburg norm for $\Phi(x) = \exp(x^2) - 1$,
- $\mathcal{Y}$ is a complete metric space embedded continuously in $C([0,T];\mathbb{R}^d)$ such that bounded sets in $B^{1/2}_{\Phi,\infty}(0,T;\mathbb{R}^d)$ are relatively compact in $\mathcal{Y}$,
- $\mathcal{X} = \{g \in C^\infty(\mathbb{R};\mathbb{R}^m) : \ g = 0 \text{ on } (-\infty, 0)\}$,
- if $X_{\mathbf{b}}$ is a continuous adapted process of bounded variation, $X_{\mathbf{m}}$ is a continuous local martingale and $X_{\mathbf{b}}(0) = X_{\mathbf{m}}(0) = 0$ then $X = X_{\mathbf{b}} + X_{\mathbf{m}}$ is called a continuous semimartingale,
- $|R|_t = \sup\ \{|R(s)| : \ s \in [0,t]\}$,
- $||R||(t)$ denotes the variation of a scalar or vector function R on $[0,t]$,
- $||R||_{\mathrm{Lip}[0,t]} = \inf\ \{c \geq 0 : |R(b) - R(a)| \leq c|b-a|,\ \ 0 \leq a \leq b \leq t\}, t \geq 0$,
- $\Delta_t(R) = \inf\ \{c : ||R(h+\cdot) - R||_{L^\Phi[0,t-h]} \leq ch^{1/2},\ \ \forall h \in [0,t]\}, t \geq 0$,
- $\langle N \rangle = \mathrm{Trace}\ (\langle N_i, N_j \rangle)_{ij}$ for a vector continuous semimartingale N,
- $\inf \emptyset = \infty$.

3. The main result

Recall that

$$B^{1/2}_{\Phi,\infty}(0,T;\mathbb{R}^d) = \{f \in L^\Phi(0,T) : \Delta_T(f) < \infty\},$$

equipped with the norm $||f|| = ||f||_{L^\Phi(0,T)} + \Delta_T(f)$, is a Banach space (see Section 5 for basic properties).

Theorem 3.1. *Let $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$ be a stochastic basis, $T > 0$, let W be an $\mathbb{R}^m$-valued $(\mathscr{F}_t)$-Wiener process, let $\xi \in \mathbb{R}^d$ and assume that*

(H0) *$b : \mathbb{R}^d \to \mathbb{R}^d$ is locally bounded and, for every $R > 0$, there exists C_R such that*

$$\big(b(x) - b(y)\big) \cdot (x - y) \le C_R |x - y|^2$$

holds for every $|x| \le R$ and $|y| \le R$,

(H1) *$\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^m$ is C^1-smooth and σ' is locally Lipschitz continuous,*

(H2) *$x' = b(x) + \sigma(x)w, x(0) = \xi$ has a solution $\mathbf{x}^w$ on $[0, T]$ for every $w \in \mathcal{X}$.*

Consider the set $\mathcal{S} = \{\mathbf{x}^w : \ w \in \mathcal{X}\}$ and let X solve

$$dX = b(X) \ dt + \sigma(X)^\circ \ dW, \quad X(0) = \xi \tag{3.1}$$

on $[0, T]$. Then $X \in B^{1/2}_{\Phi,\infty}(0, T; \mathbb{R}^d)$ a.s., $\mathcal{S} \subseteq B^{1/2}_{\Phi,\infty}(0, T; \mathbb{R}^d)$, X is a separably valued Borel measurable random variable in $\mathcal{Y}$ and if K is a closed set in $\mathcal{Y}$ then $\mathbb{P}\ [X \in K] = 1$ holds if and only if $\mathcal{S} \subseteq K$.

Example 3.2 (Generality of non-linearities). The hypotheses **(H0)** - **(H2)** of Theorem 3.1 are satisfied e.g. for the equation

$$dX = -\mathrm{sgn}(X) X^2 |\log(X)| \ dt + (X_+)^2 \ dW, \quad X(0) = \xi$$

that has a global solution. In particular, Theorem 3.1 applies whereas the support theorems in [1–14] do not, even in the path space $C([0, T])$.

Example 3.3 (The path space $\mathcal{Y}$] To give an example of a path space where Theorem 3.1 holds but the support theorems in [1–14] do not, we can think of e.g. the intersection Fréchet space

$$\mathcal{Y} = \bigcap_{n \ge 1} B^{r^{1/2} \log^{(n)}(1/r)}_{\Phi,\infty}\big(0, T; \mathbb{R}^d\big)$$

with the inductive topology where $\log^{(n)}(x) = \log(\log^{(n-1)}(x))$ for $n \ge 2$, see Example 3.5. This space is embedded in $C^{|r \log r|^{1/2} \log^{(n)}(1/r)}([0, T]; \mathbb{R}^d)$ for every $n \ge 1$ by Lemma 5.1.

Example 3.4 (Counterexample). The characterization Equation (1.2) does not hold in the space $Y = C^{|r \log r|^{1/2}}([0, T]; \mathbb{R}^d)$ nor in any smaller and finer space, in particular in $B^{1/2}_{\Phi,\infty}(0, T; \mathbb{R}^d)$. To this end, denote by

$$\mathbf{m}(f, \varepsilon) = \sup\big\{|f(b) - f(a)| : \ a, b \in [0, 1], \ |b - a| \le \varepsilon\big\} \tag{3.2}$$

the modulus of continuity of f and consider a simple equation $dX = \circ \ dW, X(0) = 0$ solved by the Wiener process W. Then

$$\lim_{\varepsilon \to 0+} \frac{\mathbf{m}(W, \varepsilon)}{|2\varepsilon \log \varepsilon|^{1/2}} = 1 \qquad \text{a.s.} \tag{3.3}$$

by [17, Theorem 1.1.1]. Hence $\sqrt{2} \le ||W - w||_Y$ a.s. for every $w \in \mathcal{S}$. In particular, $\mathbb{P}\ [W \in \bar{\mathcal{S}}^Y] = 0$, and the same is true for $Y = B^{1/2}_{\Phi,\infty}(0, T; \mathbb{R}^d)$ by Corollary 5.5.

Example 3.5 (Comparison with [11]). If $0<\lambda<1/2$ and $\mu\in\mathbb{R}$ then

$$\bigcap_{1\le p<\infty} B^{1/2}_{p,\infty}(0,1)\cap C^{|r\log r|^{1/2}}([0,1])\not\subseteq B^{r^{1/2}|\log r|^{\lambda}}_{\Phi,\infty}(0,1)$$

and the embedding

$$B^{r^{1/2}|\log r|^{\mu}}_{\Phi,\infty}(0,1),\to C^{r^{1/2}|\log r|^{\mu+1/2}}([0,1])$$

is not compact. In particular, whereas [11] does not imply the validity of the support theorem in $B^{r^{1/2}|\log r|^{\lambda}}_{\Phi,\infty}$ for $0<\lambda<1/2$ due to the above relation, Theorem 3.1 hereabove does.

For let $\{f_n : n\ge 0\}$ denote the Franklin system of orthonormal splines (see e.g. [18, Section III.C]) and, for every real continuous function F on $[0,1]$, consider the expansion

$$F=\sum_{n=0}^{\infty}\tilde{F}_n f_n \quad \text{where} \quad \tilde{F}_n=\int_0^1 F(t)f_n(t)dt.$$

There exist strictly positive constants c_λ and C_λ for $\lambda\in\mathbb{R}$ such that

$$c_\lambda||F||_{B^{r^{1/2}|\log r|^{\lambda}}_{p,\infty}[0,1]}\le\sup_{j\ge 0}\left\{|\tilde{F}_0|,|\tilde{F}_1|,\frac{2^j2^{-j/p}||\tilde{F}||_{\ell_p[2^j<n\le 2^{j+1}]}}{(j+1)^{\lambda}}\right\}\le C_\lambda||F||_{B^{r^{1/2}|\log r|^{\lambda}}_{p,\infty}[0,1]}$$

holds for every $p\in[1,\infty),\lambda\in\mathbb{R}$ and every continuous function F, see [18, Théorèm III.1]. Consequently,

$$c_\lambda||F||_{C^{r^{1/2}|\log r|^{\lambda}}[0,1]}\le\sup_{j\ge 0}\left\{|\tilde{F}_0|,|\tilde{F}_1|,\frac{2^j||\tilde{F}||_{\ell_\infty[2^j<n\le 2^{j+1}]}}{(j+1)^{\lambda}}\right\}\le C_\lambda||F||_{C^{r^{1/2}|\log r|^{\lambda}}[0,1]}$$

$$c_\lambda||F||_{B^{r^{1/2}|\log r|^{\lambda}}_{\Phi,\infty}[0,1]}\le\sup_{j\ge 0,p\ge 1}\left\{|\tilde{F}_0|,|\tilde{F}_1|,\frac{2^j2^{-j/p}||\tilde{F}||_{\ell_p[2^j<n\le 2^{j+1}]}}{p^{1/2}(j+1)^{\lambda}}\right\}\le C_\lambda||F||_{B^{r^{1/2}|\log r|^{\lambda}}_{\Phi,\infty}[0,1]}$$

holds for every $\lambda\in\mathbb{R}$ and every real continuous function F on $[0,1]$.

Now fix $0<\beta<1-2\lambda$ and define $\tilde{F}_{2^j+k}=j^{1/2}2^{-j}$ for $j\ge 1$ and $1\le k\le 2^{j-j^{\beta}}$ and set $\tilde{F}_n=0$ for all remaining indices. Hence, the function F with the coefficients $\tilde{F}_n$ belongs to $B^{1/2}_{p,\infty}[0,1]\cap C^{|r\log r|^{1/2}}[0,1]$ for every $1\le p<\infty$ but not in $B^{r^{1/2}|\log r|^{\lambda}}_{\Phi,\infty}[0,1]$.

To disprove the compactness of the embedding, consider the sequence $F^{(j)}$ with the coefficients $\tilde{F}^{(j)}_{2^j+1}=j^{\mu+\frac{1}{2}}2^{-j}$ and $\tilde{F}^{(j)}_n=0$ for all remaining indices, for $j\ge 1$. Then $\{F^{(j)}\}$ is bounded in $B^{r^{1/2}|\log r|^{\mu}}_{\Phi,\infty}$ but

$$||F^{(j)}-F^{(k)}||_{C^{r^{1/2}|\log r|^{\mu+1/2}}}\ge\varepsilon,\quad j\ne k.$$

4. The key tools

We will often need to estimate exponential Besov–Orlicz pseudonorms $\Delta_T(\cdot)$ of indefinite Stieltjes and stochastic integrals in the sequel. For that purpose, we present the following estimates where the first one is the main and the most important ingredient of the paper. In case of the standard Besov spaces $B^{1/2}_{p,\infty}$, the origins of Proposition 4.1 can be found already in [19, Remarque 4] and the first realization of the idea (having been applied to Brownian local times) was given, to our best knowledge, in [20]. Below, we prove the result for continuous local martingales in the form of a Lenglart-type estimate, showing that uniform convergence of derivatives of the quadratic variation in $L^0(\Omega; L^\infty(0,T))$ turns to convergence of the local martingales in $L^0(\Omega; B^{1/2}_{\Phi,\infty}(0,T))$.

Theorem 4.1. *Let $T>0$. There exists a function $G_T : [0,\infty) \to [0,1]$ such that $\lim_{t\to\infty} G_T(t) = 0$ and*

$$\mathbb{P}\left[\Delta_\tau(V) \geq a,\ ||\langle V\rangle||_{\mathrm{Lip}[0,\tau]} \leq c\right] \leq G_T(ac^{-1/2}) \tag{4.1}$$

holds for every $a \in [0,\infty), c \in (0,\infty)$, every $[0,T]$-valued random variable τ and every real continuous local martingale V with $V(0)=0$.

Proof. Extend the stochastic basis such that there exists a Wiener process U independent of V. Then $N = U + V$ is a local martingale with quadratic variation $\langle N\rangle(t) = t + \langle V\rangle(t)$. By the Dambis-Dubins-Schwarz theorem, there exists a Wiener process W such that $N(t) = W(\langle N\rangle(t))$ for every $t \geq 0$ a.s. In particular, on the set $[||\langle V\rangle||_{\mathrm{Lip}[0,\tau]} \leq 1]$,

$$\Delta_\tau(V) \leq \Delta_\tau(U) + \Delta_\tau(N) \leq \Delta_\tau(U) + 192\Delta_{2\tau}(W) \leq \Delta_T(U) + 192\Delta_{2T}(W)$$

by Lemma 4.3 and we get the result by Lemma 4.4 and by scaling. □

Proposition 4.2. Let $T \geq 0$. Then, for every measurable bounded function A and every continuous function B with bounded variation,

$$\Delta_T\left(\int_0^{\cdot} A\ dB\right) \leq |A|_T \Delta_T(||B||). \tag{4.2}$$

Proof. The inequality follows directly from the definition of Δ_T. □

Lemma 4.3. *Let $u \geq 1$, $T>0$ and let ϱ be an increasing function on $[0,T]$ such that $u^{-1}|t-s| \leq |\varrho(t)-\varrho(s)| \leq u|t-s|$ holds for every $s,t \in [0,T]$. Then*

$$(48u^2)^{-1}\Delta_{\varrho(T)}(f) \leq \Delta_T(f\circ\varrho) \leq (48u^2)\Delta_{\varrho(T)}(f) \tag{4.3}$$

holds for every measurable vector or scalar function f on $(0,\varrho(T))$.

Proof. Except for the explicit constants in Equation (4.3), this result was essentially proved in [20, Lemma 2.2]. The constants can be derived by a careful revision of the proof therein. □

Lemma 4.4. Let W be a Wiener process and $T>0$. Then $\Delta_T(W) < \infty$ a.s.

Proof. See e.g. the original paper [15] or a simplified proof valid also for Banach space valued Wiener processes in [16], Theorem 4.1]. □

5. Basic properties of exponential Besov–Orlicz spaces

Below, let X be a Banach space. We are going to investigate embeddings of exponential Besov–Orlicz spaces to modulus Hölder spaces and existence of continuous extension operators that we will need in the sequel.

Lemma 5.1 (Garsia, Rodemich, Rumsey [21]). *Let $p:\mathbb{R}_+\to\mathbb{R}_+$ be a continuous, strictly increasing function with $p(0)=0$, let I be an interval and $f:I\to X$ measurable such that*

$$K=\int_{I\times I}\Phi\left(\frac{|f(a)-f(b)|}{p(|a-b|)}\right)\,dadb<\infty.$$

Then

$$|f(x)-f(y)|\le 8\int_0^{|x-y|}\Phi_{-1}(Ku^{-2})dp(u)$$

holds for all points of Lebesgue density $x,y\in I$.

Remark 5.2. The proof of Lemma 5.1 in [21] is carried out for scalar-valued functions f but such a restriction is not important and the proof holds for Banach space valued functions f mutatis mutandis.

Let ζ_T be the primitive function to $4\Phi_{-1}(2Tu^{-2})u^{-1/2}$ such that $\zeta_T(0)=0$. Applying Lemma 5.1 to a given function f and to $p(u)=\Delta_T(f)u^{1/2}$, we get the following result.

Corollary 5.3. *Let $T>0$ and let $f:[0,T]\to X$ be integrable. Then*

$$|f(x)-f(y)|\le\Delta_T(f)\zeta_T(|x-y|)$$

holds for all points of Lebesgue density $x,y\in I$.

Remark 5.4. Observe that $\zeta_T(x)$ is increasing in x and T and $|r\log r|^{-1/2}\zeta_T(r)$ converges to $2^{7/2}$ if $r\downarrow 0$.

Corollary 5.5. *Let $T>0$. Then*

$$B^{1/2}_{\Phi,\infty}(0,T;X)\subseteq C^{|r\log r|^{1/2}}([0,T];X),\quad B^{|r\log r|^{1/2}}_{\Phi,\infty}(0,T;X)\subseteq C^{|\log r|r^{1/2}}([0,T];X)$$

and the embeddings are continuous.

Proof. The first embedding follows from Corollary 5.3 whereas the second one follows from Lemma 5.1 for $p(u)=c|u\log u|^{1/2}$ for small u where c is the norm of f in the space $B^{|r\log r|^{1/2}}_{\Phi,\infty}(0,T;X)$. □

The next result is going to be applied to extended or stopped processes.

Corollary 5.6. *There exists an increasing function $\kappa:\mathbb{R}_+\to\mathbb{R}_+$ such that*

$$||f(h+\cdot)-f||_{L^\Phi(\mathbb{R};X)}\le\kappa(T)h^{1/2}\Delta_T(f),\quad\forall h\ge 0$$

holds for every $T>0$ and every continuous function $f:\mathbb{R}\to X$ constant on $(-\infty,0)$ and on (T,∞).

Proof. Since $f(h+x)-f(x)$ can be non-null only for $-h<x<T$ we observe that

$$||f(h+\cdot)-f||_{L^\Phi(\mathbb{R};X)}\le\Delta_T(f)\left[h^{1/2}+2\zeta_T(h\wedge T)||1||_{L^\Phi(0,h)}\right]$$

holds for every $h\ge 0$ by Corollary 5.3 by splitting $(-h,T)=(-h,0]\cup(0,T-h)\cup$

$[T-h, T)$ if $h \leq T$ and $(-h, T) = (-h, 0] \cup (0, T]$ if $h > T$. Now we get the claim as

$$\lim_{t\to 0} |\log t|^{1/2} ||1||_{L^\Phi(0,t)} = \lim_{t\to\infty} t^{-1/2} ||1||_{L^\Phi(0,t)} = 1. \qquad \square$$

The following simple lemma, whose proof is omitted, yields that $\{\Delta_t(Y)\}_{t\geq 0}$ is a left-continuous adapted process whenever Y is a progressively measurable process.

Lemma 5.7. *Let $f : [0, T] \to X$ be Bochner measurable. Then $h \mapsto ||f(h+\cdot)-f||_{L^\Phi(0,T-h;X)}$ is lower-semicontinuous on $[0, T]$.*

6. General notation, part II

Often in this paper, we are going to work on probability spaces $(\Omega^\delta, \mathscr{F}^\delta, \mathbb{P}^\delta)$ for $\delta \in \mathbb{N}$. So, to shorten the notation for convergence in law to the Dirac measure, we are going to use the notation below, cf. the symbol $\sim$ in [2].

Definition 6.1. *Let Z^δ be a random variable on a probability space $(\Omega^\delta, \mathscr{F}^\delta, \mathbb{P}^\delta)$ for $\delta \in \mathbb{N}$. We are going to write shortly $Z^\delta \rightsquigarrow 0$ instead of*

$$\lim_{\delta\to\infty} \mathbb{P}^\delta \left[|Z^\delta| \geq \varepsilon\right] = 0 \quad \text{for every } \varepsilon > 0.$$

We also convene to denote the $\mathbb{R}^d \otimes \mathbb{R}^m \otimes \mathbb{R}^m$-valued function

$$(\sigma'\sigma)_{ijk} = \sum_{l=1}^{d} \frac{\partial \sigma_{ij}}{\partial x_l} \sigma_{lk}$$

where σ is our diffusion non-linearity in Equation (1.1).

7. A generalization of approximation results of Gyöngy, Pröhle, Nualart, and Sanz-Solé

In [2], Gyöngy and Pröhle have proved, for equations driven by continuous semimartingales, that the equality (11.2) holds in $Y = C([0, T]; \mathbb{R}^d)$ provided that b is Lipschitz continuous, $\sigma \in C^2(\mathbb{R}^d)$ and $\nabla\sigma \in C_b(\mathbb{R}^d)$. Subsequently, in [11], Gyöngy, Nualart and Sanz-Solé have improved this result for equations driven by Wiener processes by showing the equality (11.2) in Banach spaces that embed compactly intersections of modulus Hölder and modulus Besov spaces, provided that b is continuous, locally monotone and of semilinear growth, σ' is locally Lipschitz and σ, $(\sigma'\sigma)$ grow at most linearly. In particular, in [11], the path space was finer and the assumptions on b and σ less restrictive than in [2]. In both papers, the proofs were based on an approximation result for equations

$$dx^\delta = b(x^\delta)\ dt + \sigma(x^\delta) \circ dM^\delta \tag{7.1}$$

$$dy^\delta = b(y^\delta)\ dt + \sigma(y^\delta) \circ dN^\delta \tag{7.2}$$

on $[0, T]$ where $\delta \in \mathbb{N}$, cf. [2, Theorem 2.2] and [11, Theorem 2.1]. We are now going to prove an analogous approximation result in exponential Besov–Orlicz spaces while removing the assumptions on the global growth of b and σ. In other words, comparing to [2, 11], the path space is finer and the assumptions on b and σ are less restrictive.

The global growth conditions on b and σ in [2, 11] guaranteed not only existence of global solutions to Equations (7.1), (7.2) but also their tightness. While existence of global solutions is an independent problem, tightness of $\{x^\delta, y^\delta\}_{\delta\in\mathbb{N}}$ in the exponential Besov–Orlicz spaces is the main ingredient for the proof of the support theorem. We will prove below that tightness does not require global growth assumptions upon b and σ whatsoever. Observe also, the we removed the assumption on continuity of b, comparing to [2, 11].

Theorem 7.1. *Let $(\Omega^\delta, \mathscr{F}^\delta, (\mathscr{F}^\delta_t), \mathbb{P}^\delta)$ be stochastic bases for $\delta\in\mathbb{N}$, $T>0$, M^δ, N^δ continuous $\mathbb{R}^m$-valued $(\mathscr{F}^\delta_t)$-semimartingales on $[0,T], x^\delta, y^\delta$ $(\mathscr{F}^\delta_t)$-adapted solutions to Equations (7.1), (7.2) and assume that*

a. the hypotheses (**H0**) and (**H1**) hold,
b. the laws of the processes $\{y^\delta\}_{\delta\in\mathbb{N}}$ are tight in $C([0,T];\mathbb{R}^d)$,
c. $\Delta_T(||M^\delta_{\mathbf{b}}||)$ is finite a.s. for every $\delta\in\mathbb{N}$,
d. $|x^\delta(0)-y^\delta(0)|\rightsquigarrow 0$, $|M^\delta-N^\delta|_T\rightsquigarrow 0$ and $|R^\delta|_T\rightsquigarrow 0$ where

$$R^\delta_{ij}(t) = \int_0^t \left(N^\delta_i - M^\delta_i\right)\, d\left(M^\delta_{\mathbf{b}}\right)_j + \langle N^\delta_i - M^\delta_i, M^\delta_j\rangle(t) + \frac{1}{2}\left(\langle M^\delta_i, M^\delta_j\rangle(t) - \langle N^\delta_i, N^\delta_j\rangle(t)\right),$$

e. the laws of $\Delta_T(M^\delta_{\mathbf{b}}), \Delta_T(||N^\delta_{\mathbf{b}}||), ||\langle M^\delta\rangle||_{\mathrm{Lip}[0,T]}, ||\langle N^\delta\rangle||_{\mathrm{Lip}[0,T]}$ and

$$\Delta_T\left(\int_0^{\cdot} |M^\delta - N^\delta|\ d||M^\delta_{\mathbf{b}}||\right)$$

are tight with respect to $\delta\in\mathbb{N}$

Then $\Delta_T(x^\delta), \Delta_T(y^\delta), \Delta_T(M^\delta), \Delta_T(N^\delta)$ are tight with respect to $\delta\in\mathbb{N}$ and $|x^\delta - y^\delta|_T\rightsquigarrow 0$.

The following simple lemma based on the Prokhorov theorem and a Borel isomorphism theorem (see e.g. [22], Chap. 3, Par. 39, Sect. IV, p. 487]) demonstrates the real strength of Theorem 7.1.

Corollary 7.2. *If conditions in Theorem 7.1 are satisfied then x^δ, y^δ are separably valued Borel random variables in $\mathcal{Y}$ and $d_{\mathcal{Y}}(x^\delta, y^\delta)\rightsquigarrow 0$.*

Proof. The proof of Theorem 7.1 will be carried out in several steps.

I. Tightness of $\Delta_T(y^\delta), \Delta_T(M^\delta), \Delta_T(N^\delta)$ follows from (b), (e), 4.1, 4.2. Additionally, from (c), one gets that $\Delta_T(x^\delta)$ is finite a.s. for every $\delta\in\mathbb{N}$ in the same vein.
II. Introducing, analogously to [2], an adapted, continuous, non-decreasing process

$$Q^\delta(t) = |x^\delta(0)| + |y^\delta|_t + |M^\delta - N^\delta|_t + \langle M^\delta\rangle(t) + \langle N^\delta\rangle(t) + ||N^\delta_{\mathbf{b}}||(t) + \int_0^t |M^\delta - N^\delta|\ d||M^\delta_{\mathbf{b}}||,$$

define stopping times

$$\rho_r^\delta = \inf\ \{t \in [0,T] : Q^\delta(t) \geq r\}, \quad \varrho_c^\delta = \inf\ \left\{t \in [0,T] : |x^\delta - y^\delta|_t \geq c\right\}$$

and observe (as in [2]) that

$$|x^\delta - y^\delta|_T \rightsquigarrow 0 \qquad \Longleftrightarrow \qquad |x^\delta - y^\delta|_{T\wedge\rho_r^\delta\wedge\varrho_c^\delta} \rightsquigarrow 0 \quad \text{for every } r>0 \text{ and } c>0$$

as $Q^\delta(T)$ is tight by (b), (d), (e), (I) and Corollary 5.3. So we fix $r, c \in (0,\infty)$, define $\nu^\delta = T\wedge\rho_r^\delta\wedge\varrho_c^\delta$ and proceed to prove that $|x^\delta - y^\delta|_{\nu^\delta} \rightsquigarrow 0$.

III. The laws of $\{\Delta_{\nu^\delta}(x^\delta)\}_{\delta\in\mathbb{N}}$ are tight. To prove this, let ν be a smooth density compactly supported around the origin in $\mathbb{R}^d$, define mollifiers $\nu_l(x) = l^d\nu(lx)$ and $\sigma^l = \sigma * \nu_l$, and apply the integration-by-parts formula $XdY = d(XY) - YdX - \langle X, Y\rangle$ to

$$\int_0^\cdot \sigma_{ij}^l(x^\delta)\ d\left(M_j^\delta - N_j^\delta\right)$$

in the same way as in the proof of [2, Theorem 2.2]. Using local boundedness of b, σ, σ', σ'', (e), Equations (4.1) and (4.2), we get that

$$\left\{\Delta_{\nu^\delta}\left(\sigma^l(x^\delta)(M^\delta - N^\delta) - \int_0^\cdot \sigma^l(x^\delta)\ d(M^\delta - N^\delta)\right)\right\}_{(\delta,l)\in\mathbb{N}^2}$$

is tight, hence

$$\left\{\Delta_{\nu^\delta}\left(\sigma(x^\delta)(M^\delta - N^\delta) - \int_0^\cdot \sigma(x^\delta)\ d(M^\delta - N^\delta)\right)\right\}_{\delta\in\mathbb{N}}$$

is tight by lower-semicontinuity of $\Delta_t(\cdot)$ with respect to a.e.-convergence, for every $t \geq 0$. Now, using again local boundedness of b, σ, σ', (e), Equations (4.1) and (4.2), we get that

$$\left\{\Delta_{\nu^\delta}\left(x^\delta - \sigma(x^\delta)(M^\delta - N^\delta)\right)\right\}_{\delta\in\mathbb{N}}$$

is tight. The inequality $\Delta_t(AB) \leq |A|_t\Delta_t(B) + |B|_t\Delta_t(A)$, local boundedness and local Lipschitzianity of σ, (I), as in [11, Proposition 2.4], yield existence of a constant C independent of δ and tight non-negative random variables $\{S_\delta\}_{\delta\in\mathbb{N}}$ such that

$$\Delta_{\nu^\delta}(x^\delta) \leq S_\delta + C|M^\delta - N^\delta|_{\nu_\delta}\Delta_{\nu^\delta}(x^\delta)$$

hence $\Delta_{\nu^\delta}(x^\delta)$ is tight by (d) and (I).

IV. $\{x^\delta(\cdot\wedge\nu^\delta), y^\delta(\cdot\wedge\nu^\delta)\}_{\delta\in\mathbb{N}}$ are tight in $C([0,T];\mathbb{R}^d)$ by (III), Corollary 5.3 and (b).

V. If $\{J_{ij}^\delta\}$ are continuous adapted processes with tight laws in $C([0,T])$ then

$$\left|\int_0^\cdot J^\delta\ dR^\delta\right|_T \rightsquigarrow 0.$$

This is so since $|R^\delta|_T \rightsquigarrow 0$ by (d), $||R^\delta||(T)$ is tight by (e) and J^δ, being tight in $C([0,T])$, can be uniformly approximated by step processes.

VI. The follownig equality holds

$$\int_0^t \left(x_i^\delta - y_i^\delta\right)\sigma_{ij}(x^\delta)\ d\left(M_j^\delta - N_j^\delta\right) = \sum_{k=1}^m \int_0^t \sigma_{ij}(x^\delta)\sigma_{ik}(y^\delta)\ d\langle M_j^\delta - N_j^\delta, N_k^\delta\rangle$$
$$+\sum_{k=1}^m \int_0^t \left[\left(x_i^\delta - y_i^\delta\right)(\sigma'\sigma)_{ijk}(x^\delta) + \sigma_{ij}(x^\delta)\sigma_{ik}(x^\delta)\right]\ dA_{jk}^\delta + K_{ij}^\delta(t) \tag{7.3}$$

where

$$A_{ij}^\delta(t) = \int_0^t \left(N_i^\delta - M_i^\delta\right)\ d\left(M_{\mathbf{b}}^\delta\right)_j + \langle N_i^\delta - M_i^\delta, M_j^\delta\rangle(t) \qquad \text{and} \qquad \mathbb{E}^\delta\ |K^\delta|_{\nu^\delta}^2 \le C\mathbb{E}^\delta\ |M^\delta - N^\delta|_{\nu^\delta}^2,$$

the constant C here not depending on δ. To prove this, apply the Itô integration-by-parts formula to the product of the semimartingales $x_i^\delta - y_i^\delta, \sigma_{ij}^l(x^\delta)$ and $M_j^\delta - N_j^\delta$, and let $l \to \infty$, using local boundedness of b, σ, σ' and σ''. Realize that $\mathbb{E}^\delta\ |K^\delta|_{\nu^\delta}^2 \to 0$ as $\delta \to \infty$.

VII. Defining

$$\psi^\delta(t) = \sum_{j=1}^m \sum_{k=1}^m \int_0^t \left[\sum_{i=1}^d \left(x_i^\delta - y_i^\delta\right)(\sigma'\sigma)_{ijk}(x^\delta) + (\sigma^*\sigma)_{jk}(x^\delta)\right]\ dR_{jk}^\delta,$$

we have that $|\psi^\delta|_{\nu^\delta} \le C$ by the definition of ν^δ, where C does not depend on δ, and $|\psi^\delta|_{\nu^\delta} \rightsquigarrow 0$ by (IV) and (V). In particular, $\mathbb{E}^\delta|\psi^\delta|_{\nu^\delta}^2 \to 0$ as $\delta \to \infty$.

VIII. Let us define the processes

$$S^\delta(t) = (t\wedge\nu^\delta) + \langle N^\delta\rangle(t\wedge\nu^\delta) + ||N_{\mathbf{b}}^\delta||(t\wedge\nu^\delta), \quad z^\delta(t) = |x^\delta(t\wedge\nu^\delta) - y^\delta(t\wedge\nu^\delta)|^2$$
$$U^\delta(t) = 2\int_0^{t\wedge\nu^\delta} \left(x^\delta - y^\delta\right)^*[\sigma(x^\delta) - \sigma(x^\delta)]\ dN_{\mathbf{m}}^\delta, z^{\delta,\alpha}(t) = z^\delta(t)1_{[z^\delta(0)\le\alpha]}.$$

By the Itô formula and the hypotheses **(H0)** and **(H1)**, we get for every $t \ge 0$ that

$$|z^{\delta,\alpha}|_t \le \alpha + 2|\psi^\delta|_{\nu^\delta} + C|K^\delta|_{\nu^\delta} + C\int_0^t |z^{\delta,\alpha}|_s\ dS^\delta(s) + 1_{[z^\delta(0)\le\alpha]}|U^\delta|_t$$

hence

$$\mathbb{E}^\delta|z^{\delta,\alpha}|_{\tau^\delta}^2 \le c\mathbb{E}^\delta\left[\alpha^2 + |\psi^\delta|_{\nu^\delta}^2 + |K^\delta|_{\nu^\delta}^2\right] + c\mathbb{E}^\delta \int_0^{\tau^\delta} |z^{\delta,\alpha}|_s^2\ dS^\delta(s)$$

holds by the Burkholder-Davies-Gundy inequality for every finite stopping time τ^δ, whereas the constants c and C do not depend on $\delta \in \mathbb{N}, \alpha > 0$ nor τ^δ. So, by the stochastic Gronwall inequality (see e.g. [3, Lemma 4]), we get

$$\mathbb{E}^\delta|z^{\delta,\alpha}|_\infty^2 \le \kappa\mathbb{E}^\delta\left[\alpha^2 + |\psi^\delta|_{\nu^\delta}^2 + |K^\delta|_{\nu^\delta}^2\right]$$

for some κ independent of $\delta \in \mathbb{N}$ and $\alpha > 0$. Thus, by the Chebyshev inequality,

$$\mathbb{P}^\delta\ \left[|x^\delta - y^\delta|_{\nu^\delta} \ge \varepsilon\right] \le \frac{\kappa}{\varepsilon^2}\mathbb{E}^\delta\left[\alpha^2 + |\psi^\delta|_{\nu^\delta}^2 + |K^\delta|_{\nu^\delta}^2\right] + \mathbb{P}^\delta\ \left[|x^\delta(0) - y^\delta(0)| \ge \sqrt{\alpha}\right]$$

which implies $|x^\delta - y^\delta|_{\nu^\delta} \rightsquigarrow 0$ by (d), (VI) and (VII). In particular, $|x^\delta - y^\delta|_T \rightsquigarrow 0$ by (II).

IX. Since $|x^\delta - y^\delta|_T$ is tight by (VIII), $|x^\delta|_{h_1,T}$ is tight by (II) and (III). □

8. Sufficient conditions for Theorem 7.1

Here we derive conditions that guarantee that the assumptions (d) and (e) in Theorem 7.1 are satisfied for suitable transformations of the Wiener process, introduced in [4]. In conformity with [3, 4], we let $\varphi : \mathbb{R} \to [0, \infty)$ be a smooth density supported in (0, 1) and we define $\varphi^\delta(t) = \delta\varphi(\delta t), \delta \in \mathbb{N}$ and

$$C_\delta(t) = \text{sign}(t)\min\ \{\delta, |t|\}, \quad t \in \mathbb{R}.$$

If V is a scalar continuous adapted process, we extend it as $V = 0$ on $(-\infty, 0)$ and we realize that $V_\delta = V * \varphi^\delta$ and $V^\delta = C_\delta(V) * \varphi^\delta$ are smooth adapted processes vanishing on $(-\infty, 0]$ and

$$|(V^\delta)'|_t \leq |\varphi'|_{L^1(0,1)}\delta\min\ \{\delta, |V|_t\}, \quad t \geq 0.$$

Below, we revisit the estimates in [3, 4] to cover the exponential Besov–Orlicz spaces. Towards this end, we fix $T \in (0, \infty)$.

Proposition 8.1. *Let V be an adapted continuous process. Then $\Delta_T(V^\delta) \leq \kappa(T)\Delta_T(V)$ for every $\delta \in \mathbb{N}$ and $|V^\delta - V|_T \to 0$.*

Proof. By Corollary 5.6, for every $h \in [0, T]$,

$$||V^\delta(h + \cdot) - V^\delta||_{L^\Phi[0,T-h]} \leq ||V(h + \cdot) - V||_{L^\Phi(-\infty,T-h]} \leq \kappa(T)h^{1/2}\Delta_T(V).$$

Lemma 8.2. *Let V be a standard Wiener process on $[0, T]$. Then, for every $\delta \in \mathbb{N}$,*

$$\mathbb{E}\int_0^T |V - V_\delta|^4\ dr \leq 3T\delta^{-2}, \quad \mathbb{E}\int_0^T |(V_\delta)'|^4\ dr \leq 3T\delta^2|\varphi|^4_{L^2(0,1)}.$$

Proof. We know that $\mathbb{E}\ |V(t) - V(s)|^4 = 3|t - s|^2$ for every $t, s \in [0, T]$. So

$$\mathbb{E}\ |V(t) - V_\delta(t)|^4 \leq \mathbb{E}\int_0^{1/\delta} |V(t) - V(t-u)|^4\varphi^\delta(u)\ du \leq 3\delta^{-2}.$$

By the Itô formula, for every semimartingale Y starting from zero and for every $t \geq 0$,

$$\mathbb{P}\left[(Y * \varphi^\delta)'(t) = \int_0^t \varphi^\delta(t - s)\ dY(s)\right] = 1 \tag{8.1}$$

so, by the Itô formula applied on $x \mapsto x^4$,

$$\mathbb{E}\ |(V_\delta)'(t)|^4 \leq 3\left(\int_0^t \left[\varphi^\delta(t-s)\right]^2\ ds\right)^2 \leq 3\delta^2|\varphi|^4_{L^2(0,1)}.$$ □

Proposition 8.3. Let V and J be standard Wiener processes on $[0,T]$. Then

$$\mathbb{E}\left[\Delta_T\left(\int_0^{\cdot}|V-V^\delta|\ d||J^\delta||\right)\right]^2 \le C_{T,\varphi}$$

Proof. Using the Cauchy-Schwarz inequality, we get that

$$\int_t^{t+h}|V-V_\delta||(J_\delta)'|\ dr \le h^{1/2}||V-V_\delta||_{L^4(0,T)}||(J_\delta)'||_{L^4(0,T)}$$

for every $0\le h\le T$ and $0\le t\le T-h$ so, by Lemma 8.2,

$$\mathbb{E}\left[\Delta_T\left(\int_0^{\cdot}|V-V_\delta|\ d||J_\delta||\right)\right]^2 \le C_T||\varphi||^2_{L^2(0,1)}.$$

Now $V^\delta=V_\delta$ and $J^\delta=J_\delta$ on the set $[|V|_T\vee|J|_T\le\delta]$, whereas, on its complement,

$$\int_t^{t+h}|V-V^\delta||\left(J^\delta\right)'|\ dr \le 2h\delta||\varphi'||_{L^1(0,1)}|V|_T|J|_T \le 2h||\varphi'||_{L^1(0,1)}\left[|V|_T\vee|J|_T\right]^3$$

for every $0\le h\le T$ and $0\le t\le T-h$. □

Proposition 8.4. Let V and J be standard Wiener processes on $[0,T]$. Then

$$\lim_{\delta\to\infty}\mathbb{E}\left|\int_0^{\cdot}\left(V-V^\delta\right)\ dJ^\delta-\frac{1}{2}\langle V,J\rangle\right|_T^2=0.$$

Proof. See [4, Lemma 3 (3)]. □

9. Proof of the main result

We devide the proof to four steps, the last two being within the framework of [2, 4].

1. $\Delta_T(X)<\infty$ a.s. and $\Delta_T(\mathbf{x}^w)<\infty$ for every $w\in\mathcal{X}$ by Equation 4.1, (4.2).
2. If $t\in(0,T]$ and $w^l,w\in\mathcal{X}$ satisfy $\int_0^t|w^l-w|\ ds\to0$ as $l\to\infty$ then $|\mathbf{x}^{w_l}-\mathbf{x}^w|_t\to0$ by **(H0)** and **(H1)**. In particular, the trace space $\{f|_{[0,t]}:f\in\mathcal{X}\}$ equipped with the $L^1(0,t)$-norm is a separable normed space and the assignment $\mathcal{X}_t\to\mathbb{R}^d$: $f|_{[0,t]}\mapsto\mathbf{x}^f$ is well-defined (by uniqueness of the equation in **(H2)**), continuous, hence Borel measurable. For if Z is an adapted process with paths in $\mathcal{X}$ then $\mathbf{x}^Z(t)$ is also an adapted process.
3. We follow [2, 4] in the rest of the proof. Set $\mathbb{P}^\delta=\mathbb{P},M_i^\delta=C_\delta(W_i)*\varphi^\delta$ as in Section 8, $N_i^\delta=W_i,x^\delta=\mathbf{x}^{(M^\delta)\prime}$ and $y^\delta=X$ for $\delta\in\mathbb{N}$. The omega-wise defined process x^δ is $(\mathscr{F}_t)$-adapted by **(2)**. By the results of Section 8, the assumptions of Theorem 7.1 are satisfied and so $d_{\mathcal{Y}}(x^\delta,X)\to0$ in probability by Corollary 7.2. Since paths of x^δ belong to $\mathcal{S}$, paths of X belong to the closure of $\mathcal{S}$ in $\mathcal{Y}$ a.s. by the Portmanteau theorem.
4. For the converse, we also follow [2, 4]. For let $w\in\mathcal{X}$ and define $W=0$ on $(-\infty,0)$. Unique solutions to

$$Z_i^\delta=W_i-w_i+C_\delta\left(Z_i^\delta\right)*\varphi^\delta,\quad Z_i^\delta=0\quad\text{on}\quad(-\infty,0],\quad 1\le i\le m \tag{9.1}$$

can be constructed omega-wise on $(-\infty, T]$ by Picard iterations[1] that converge uniformly on $(-\infty, T]$, hence the solutions Z_i^δ are continuous and $(\mathscr{F}_t)$-adapted. We define

$$Y^\delta = \int_0^\cdot [O^\delta]' \, dW, \quad L^\delta = \exp\left\{ Y^\delta(T) - \frac{1}{2}\langle Y^\delta\rangle(T)\right\}$$

where every derivative of $O_i^\delta := w_i - C_\delta(Z_i^\delta) * \varphi^\delta$ is uniformly bounded. The Novikov condition is satisfied for Y^δ at T, hence $\mathbb{E}L^\delta = 1$ and we define $d\mathbb{P}^\delta := L^\delta \, d\mathbb{P}$ for which Z^δ is an $((\mathscr{F}_t), \mathbb{P}^\delta)$-Wiener process on $[0, T]$ by the Girsanov theorem. We set $M^\delta = W, N^\delta = w, x^\delta = X, y^\delta = \mathbf{x}^w$ and consider the probability measures $\{\mathbb{P}^\delta\}_{\delta\in\mathbb{N}}$ under which $M_{\mathbf{m}}^\delta = Z^\delta$ and $M_{\mathbf{b}}^\delta = O^\delta$. Using the results of Section 8, we check that the assumptions of Theorem 7.1 are satisfied and so $d_{\mathcal{Y}}(\mathbf{x}^w, X) \rightsquigarrow 0$ by Corollary 7.2. Since $\mathbb{P}^\delta$ and $\mathbb{P}$ are equivalent, this means that $\mathbb{P}\,[d_{\mathcal{Y}}(X, \mathbf{x}^w) < \varepsilon] > 0$ for every $\varepsilon > 0$. If K is a closed set in Y such that $\mathbb{P}\,[X \in K] = 1$ then $\mathbf{x}^w \in K$, which is what we needed to prove.

Note

1. Since paths of W_i and w_i are bounded on $[0, T]$, one can consider a suitable $\alpha > 0$ and the complete norm $\sup_{t\leq T} e^{-\alpha t}|f(t)|$ for f bounded, continuous and null on $(-\infty, 0]$, and apply the Banach fixed point theorem.

Acknowledgments

The authors wish to thank Jan Seidler for valuable discussions and helpful comments.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The research of M. Ondreját was supported by the Czech Science Foundation grant no. 15-08819S.

References

[1] Stroock, D. W., Varadhan, S. R. S. (1972). *On the Support of Diffusion Processes with Applications to the Strong Maximum Principle*. Berkeley,CA: University of California Press, 333–359.

[2] Gyöngy, I., Pröhle, T. (1990). *On the approximation of stochastic differential equation and on Stroock-Varadhan's support theorem*. *Comput. Math. Appl.* 19(1):65–70.

[3] Matskyavichyus, V. (1985). *Sp-stability of solutions of symmetric stochastic differential equations*. *Litovsk. Mat. Sb.* 25(4):72–84.

[4] Matskyavichyus, V. (1986). *The support of the solution of a stochastic differential equation*. *Litovsk. Mat. Sb.* 26(1):91–98.

[5] Friz, P. K., Victoir, N. B. (2010). *Multidimensional Stochastic Processes as Rough Paths, Theory and applications. Cambridge Studies in Advanced Mathematics.*, vol. 120, Cambridge: Cambridge University Press.
[6] Aida, S., Kusuoka, S., Stroock, D. (1993). *On the Support of Wiener Functionals, Asymptotic Problems in Probability Theory: Wiener Functionals and Asymptotics (Sanda/ Kyoto, 1990), Pitman Research Notes in Mathematics Series*, vol. 284, Harlow: Longman Sci. Tech., pp. 3–34.
[7] Aida, S. (1990). *Support theorem for diffusion processes on Hilbert spaces. Publ. Res. Inst. Math. Sci.* 26(6):947–965.
[8] Arous, G. B., Grădinaru, M., Ledoux, M. (1994). *Hölder norms and the support theorem for diffusions. Ann. Inst. H. Poincaré Probab. Statist.* 30(3):415–436.
[9] Dereich, S., Dimitroff, G. (2012). *A support theorem and a large deviation principle for Kunita flows. Stoch. Dyn.* 12(3):1150022.
[10] Friz, P., Victoir, N. (2005). *Approximations of the Brownian rough path with applications to stochastic analysis. Ann. Inst. H. Poincaré Probab. Statist.* 41(4):703–724.
[11] Gyöngy, I., Nualart, D., Sanz-Solé, M. (1995). *Approximation and support theorems in modulus spaces. Probab. Theory Related Fields* 101(4):495–509.
[12] Ledoux, M., Qian, Z., Zhang, T. (2002). *Large deviations and support theorem for diffusion processes via rough paths. Stochastic Process. Appl*102(2):265–283. no. MR 1935127
[13] Mellouk, M. (1994). *Support des diffusions dans les espaces de Besov-Orlicz. C. R. Acad. Sci. Paris Sér. I Math.* 319(3):261–266.
[14] Millet, A., Sanz-Solé, D. M. (1994). *A Simple Proof of the Support Theorem for Diffusion Processes, Séminaire de Probabilités, XXVIII, Lecture Notes in Mathematics.*, vol. 1583, Berlin: Springer, pp. 36–48.
[15] Ciesielski, Z. (1993). *Orlicz spaces, spline systems, and Brownian motion. Constr. Approx.* 9(2–3):191–208.
[16] Hytönen, T. P., Veraar, M. C. (2008). *On Besov regularity of Brownian motions in infinite dimensions. Probab. Math. Statist.* 28(1):143–162.
[17] Csörgő, M., Révész, P. (1981). *Strong Approximations in Probability and Statistics, Probability and Mathematical Statistics*. New York: Academic Press, Inc.
[18] Ciesielski, Z., Kerkyacharian, G., Roynette, B. (1993). *Quelques espaces fonctionnels associés à des processus Gaussiens. Studia Math.* 107(2):171–204.
[19] Roynette, B. (1993). *Mouvement Brownien et espaces de besov. Stochastics Stochastics Rep.* 43(3–4):221–260.
[20] Hu, Y., Mellouk, M. (2000). *Régularité Besov-Orlicz du temps local brownien. Studia Math.* 139(1):1–7.
[21] Garsia, A. M., Rodemich, E., Rumsey, H. Jr., (1970/1971). *A real variable lemma and the continuity of paths of some Gaussian processes. Indiana Univ. Math. J.* 20:565–578.
[22] Kuratowski, K. (1966). *Topology. Vol. I, New Edition, Revised and Augmented*. Translated from the French by J. Jaworowski, New York: Academic Press.