Topics in Stochastic Differential Equations and Rough Path Theory

vorgelegt von
Dipl.-Math. Joscha Diehl

Von der Fakultät II – Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
– Dr. rer. nat. –

genehmigte Dissertation

Promotionsausschuss

Vorsitzender:	Prof. Dr. Dietmar Hömberg
Berichter:	Prof. Dr. Peter Friz
	Prof. Dr. Terry Lyons

Tag der wissenschaftlichen Aussprache: 22. März 2012

Berlin 2012
D 83

Acknowledgment

I want to express my deep gratitude towards my advisor Prof. Peter Friz, who set me on an interesting collection of problems. He was always ready to discuss any piece of mathematics and his unyielding enthusiasm was a constant source of joy.

I want to thank Prof. Terry Lyons for accepting to be co-examiner of this thesis.

Financial support by the DFG through the IRTG "Stochastic Models of Complex Processes" as well as the SPP1324 is gratefully acknowledged.

Contents

1 Introduction

Rough path theory, introduced by Lyons [52], is a recent development to treat solutions of differential equations driven by paths that are very irregular in time. More precisely, the theory gives meaning to differential equation of the form

$$Y_t = Y_0 + \int_0^t V(Y_r) dX_r, \tag{1}$$

under appropriate assumptions on the deterministic path X and the vector field V, which go way beyond the case $X \in C^1([0,T], \mathbb{R}^d)$.

Heuristically, the theory can deal with α-Hölder continuous paths of some order $\alpha > 0$. The regime $\alpha > 1/2$ is the well understood Young theory ([62, 49]). For $\alpha \in (1/3, 1/2]$ it turns out that the information of the $\mathbb{R}^d$-valued path alone is not sufficient to build a satisfying theory. An α-Hölder rough path then takes values in a larger metric space which, in addition to the path itself, contains "second order information" which can be though of as the iterated integrals of the path against itself. [1] This information has to be given exogenously, since it is not canonically determined, and has to satisfy a certain algebraic compatibility condition. On the space of all such paths a rough path "norm" [2] (similar to the usual α-Hölder norm of paths in $\mathbb{R}^d$) is introduced, and all elements with finite norm are denoted *α-Hölder rough paths*. Distance between paths is measured using this "norm", giving rise to a *rough path metric.* [3]

There are then at least three different approaches to solutions of rough differential equations (RDEs) of the form (1), which lead essentially to the same results. Lyons' original idea was to define some sort of rough integration first and consider the integral equation as a fixed point problem. One of the key results is stability: with respect to rough path metric the solution mapping of RDEs is continuous in the driving signal (Lyons' limit theorem). In particular, if X is a *geometric rough path*, which by definition means that there exists a sequence of smooth paths X^n converging to X in rough path metric, then in particular the solution to the RDE driven by X is the limit of classical ODE solutions driven by the approximating smooth paths X^n . The approach by Davie [21] defines a solution locally, taking inspiration from (higher order) Euler approximations. In this thesis we shall adopt the approach put forward by Friz and Victoir [33]. It takes Lyons' limit theorem as starting point, and *defines* a solution to an RDE driven by a geometric rough path as the limit (if it exists uniquely) of ODE solutions driven by approximating smooth paths.

Let us note that the theory can treat many stochastic processes as driving signals, continuous semimartingales being the prime example (see e.g. [51, 33]). Owing to stability properties of rough path theory, many known results from stochastic analysis can then be easily recovered (existence and regularity of stochastic flows, support theorems, large deviation principles, see e.g. [46, 33]). A plethora of new results have also been made possible. Let us mention new numerical algorithms for SDEs [35], SDEs driven by non-semimartingales [17, 33, 40], Hörmander type theorems for SDEs driven by Gaussian (rough) paths [11, 3] and ergodicity for SDEs driven by fractional Brownian motion [39].

As we pointed out, an important (or even defining) feature of rough path theory is, that it establishes continuity in the driving signal for time inhomogeneous differential equations with

[1] For $\alpha < 1/3$, more and more additional information (more "iterated integrals") has to be given a priori, related to the size of $1/\alpha$.

[2] The space is not a vector space.

[3] We note that one can actually choose from a family of metrics; this shall be of no concern here.

respect to a certain metric. This thesis investigates whether the principle of stability can be extended to other objects that are classically well-defined when taking a smooth path η as input. In particular we consider the solution mapping for SDEs with an additional drift

$$S_t = S_0 + \int_0^t a(S_r)dr + \int_0^t b(S_r)dB_r + \int_0^t c(S_r)d\eta_r, \tag{2}$$

BSDEs with an additional driver

$$Y_t = \xi + \int_t^T f(r, Y_r, Z_r)dr + \int_t^T H(Y_r)d\eta_r - \int_t^T Z_r dW_r, \tag{3}$$

and parabolic PDES of the form

$$\partial_t u + F(t, u, Du, D^2u) + H(u, Du)\dot{\eta}_t = 0. \tag{4}$$

Note that if η is smooth, these equations possess a unique solution (under appropriate assumptions on the coefficients).

Section 2 covers our results on equation (2). We establish stability in rough path metric with respect to η, which makes it possible to define a solution to an *SDE with rough drift* for arbitrary geometric rough paths $\boldsymbol{\eta}$. Deterministic rough path estimates do not easily mix with stochastic estimates and we tackle the problem using three different approaches. The first one consist in the application of a flow transform, which converts the equation into one that depends on the driving signal in a harmless way; in particular there is no explicit dependence on the derivative anymore. A similar technique is used in the section on BSDEs with rough driver, but the flow transform leads to more technical problems there. The second approach interprets equation (2) directly as an RDE. For this to work one needs to understand $\boldsymbol{\eta}$ and the Brownian motion B together as a *joint* rough path. There is a canonical way of doing this though using stochastic integration and formal integration by parts, at least for the case $\alpha > 1/3$. The main problem in this approach is showing exponential integrability of solutions, a property that we need for applications. Lastly, in the Young regime $\alpha > 1/2$ we are able to solve the integral equation directly. The main technical difficulty here stems from the fact that the deterministic estimates available for the Young integral are not easily combined with the stochastic (not pathwise) estimates available for the stochastic integral.

The motivation and main application for equations of type (2) is a problem in the theory of nonlinear filtering. As conjectured by Lyons [52], we show that in order to do *robust filtering* for multidimensional observation, correlated with the signal, one needs to treat the observation as an element of a rough path space. This is shown in Section 2.4.1 using SDEs with rough drift. We refer to the introduction of that section for a thorough explanation and a historical account concerning the problem of robust filtering.

In Section 2.4.2 we consider an application to (pathwise) stochastic control. The resulting Hamilton-Jacobi-Bellman equation is of main interest here, since it gives a stochastic representation for a class of "rough partial differential equations" (4). We note that rough path theory has been applied to the field of partial differential equations (PDEs) in several papers [36, 27, 61]. We work with the approach put forward by Friz et al [9, 10, 32], which treats rough PDEs as the limit of PDEs in which the rough path is replaced by an approximating sequence of smooth paths. The aforementioned stability of rough path theory is also a key feature in the theory of *viscosity solutions to PDEs* ([18]). It is then natural to interpret solutions to the approximating PDEs in the viscosity sense, and this is excatly what Friz et al did. This also turns out to be convenient for this particular application, since viscosity solutions are the right solution concept for PDEs originating from optimal control. Let us mention that we can, in a Brownian setting (that is, when formally replacing η by a Wiener process), recover results by Buckdahn and Ma [8], under somewhat weaker assumptions and significantly shorter proofs.

We note that (formally), when plugging in a fractional Brownian motion with Hurst parameter $H > 1/2$ for the path η in (2), one gets the solution to a mixed Brownian-fractional Brownian SDE. Such equations are treated for example in [63] (with only one-dimensional Brownian motion and stronger assumptions on the vector fields). This statement is made rigorous in Theorem 33.

Backward stochastic differential equations (BSDEs) were introduced by Bismut in 1973. In [4] he applied linear BSDEs to stochastic optimal control. In 1990 Pardoux and Peng [57] then considered non-linear equations. A solution to a BSDE with driver f and random variable $\xi \in L^2(\mathcal{F}_T)$ is an adapted pair of processes (Y, Z) in suitable spaces, satisfying

$$Y_t = \xi + \int_t^T f(r, Y_r, Z_r) ds - \int_t^T Z_r dW_r, \quad t \leq T.$$

Under appropriate conditions on f and ξ they showed the existence of a unique solution to such an equation.

Again we are interested in the (formal) equation (3), when $\boldsymbol{\eta}$ is a rough path. In the case of η beeing a smooth path, one is back in the classical setting. In Section 3 we will show that under appropriate conditions, if the smooth path converge to a rough path, then the corresponding solutions also converge.

We use a flow transformation to prove the limit theorem. As opposed to the case of aforementioned SDEs, this leads to technical difficulties. The transformed BSDE falls outside the classical Lipschitz framework, since it is quadratic in the control variable, and we have to use stability results by Kobylanski [43].

We note that formally, when plugging in (the rough path lift of) a Brownian motion for the rough path in (3), one gets the solution to a backward doubly stochastic differential equation (BDSDE), a notion introduced in [55]. This statement is made rigorous in Section 3.4.

When the randomness of the parameters (f, ξ) of a BSDE comes from the state of a standard SDE, the system it is called a forward-backward stochastic differential equation (FBSDE). The solutions of FBSDE's are connected to (viscosity) solutions of partial differential equations (PDEs) in what is sometimes called a nonlinear Feynman-Kac formula, see e.g. [54]. We extend this result to BSDEs with rough drift, thereby arriving at a nonlinear Feynman-Kac formula for a class of rough RDEs in Section 3.3. At the same time we get existence, uniqueness and stability for a class of equations, that has not been treated before in the literature.

2 Stochastic differential equations with rough drift

The aim of this part is to give meaning to an equation of the form

$$S_t = S_0 + \int_0^t a(S_r)dr + \int_0^t b(S_r)d\bar{B}_r + \int_0^t c(S_r)d\boldsymbol{\eta}_r,$$

in the case where $\boldsymbol{\eta}$ is a rough path. As mentioned in the introduction, we proceed by continuously extending the solution mapping, which is of course classically well defined if $\boldsymbol{\eta}$ is (the lift of) a smooth path.

After introducing some notation, we present and prove the main results in Section 2.2. In Section 2.3 we improve, in the case of a Young driving path, on the regularity assumptions. A specific feature of this case is, that the SDE with rough drift actually satisfies the integral equation. Finally, in Section 2.4 we present some applications.

2.1 Notation

Lip^γ is the set of γ-Lipschitz [4] functions $a : \mathbb{R}^m \to \mathbb{R}^n$ where m and n are chosen according to the context.

$G^{[p]}(\mathbb{R}^{d_Y}) \cong \mathbb{R}^d \oplus so(d_Y)$ is the free nilpotent group of step $[p]$ over $\mathbb{R}^d$. Here $[p]$ denotes the largest integer not larger then p. [5]

$\mathcal{C}^{0,\alpha} := C_0^{0,\alpha-\text{Höl}}([0,t], G^{[1/\alpha]}(\mathbb{R}^{d_Y}))$ is the set of geometric α-Hölder rough paths $\eta : [0,t] \to G^{[1/\alpha]}(\mathbb{R}^{d_Y})$ starting at 0. We shall use the non-homogeneous metric $\rho_{\alpha-\text{Höl}}$ on this space.

$\mathcal{C}^{0,q-var} := C_0^{0,q-var}([0,t], G^{[q]}(\mathbb{R}^{d_Y}))$ is the set of geometric q-variation rough paths $\eta : [0,t] \to G^{[q]}(\mathbb{R}^{d_Y})$ starting at 0.

$(\bar{\Omega}, \bar{\mathcal{F}}, (\bar{F}_t)_{t\geq 0}, \bar{\mathbb{P}})$ will be a filtered probability space. Let $\mathcal{S}^0 = \mathcal{S}^0(\bar{\Omega})$ denote the space of adapted, continuous processes in $\mathbb{R}^{d_S}$, with the topology of uniform convergence in probability.

For $q \geq 1$ we denote by $\mathcal{S}^q = \mathcal{S}^q(\bar{\Omega})$ the space of processes $X \in \mathcal{S}^0$ such that

$$||X||_{\mathcal{S}^q} := \left(\bar{\mathbb{E}}[\sup_{s\leq t} |X_s|^q]\right)^{1/q} < \infty.$$

2.2 Main results

Let $(\bar{\Omega}, \bar{\mathcal{F}}, (\bar{F}_t)_{t\geq 0}, \bar{\mathbb{P}})$ be a filtered probability space carrying a d_B-dimensional Brownian motion $\bar{B}$ and a bounded d_S-dimensional random vector S_0 independent of $\bar{B}$.

In the following $\alpha \in (0,1]$. Let $\eta^n : [0,t] \to \mathbb{R}^{d_Y}$ be smooth paths, such that $\eta^n \to \boldsymbol{\eta}$ in α-Hölder, for some $\boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}$ and let S^n be a d_S-dimensional process which is the unique solution to the classical SDE

$$S_t^n = S_0 + \int_0^t a(S_r^n)dr + \int_0^t b(S_r^n)d\bar{B}_r + \int_0^t c(S_r^n)d\eta_r^n,$$

[4] In the sense of E. Stein, i.e. bounded k-th derivative for $k = 0, \ldots, \lfloor\gamma\rfloor$ and $\gamma - \lfloor\gamma\rfloor$-Hölder continuous $\lfloor\gamma\rfloor$-th derivative, where $\lfloor\gamma\rfloor$ is the largest integer strictly smaller then γ.

[5] This is the "correct" state space for a geometric $1/p$-Hölder rough path; the space of such paths subject to $1/p$-Hölder regularity (in rough path sense) yields a complete metric space under $1/p$-Hölder rough path metric. Technical details of geometric rough path spaces can be found e.g. in Section 9 of [33].

where we assume that

(a1) $\alpha \in (0,1]$, $\gamma > 1/\alpha$, $a \in \mathrm{Lip}^1(\mathbb{R}^{d_S})$, $b_1, \dots, b_{d_B} \in \mathrm{Lip}^1(\mathbb{R}^{d_S})$ and $c_1, \dots, c_{d_Y} \in \mathrm{Lip}^{\gamma+2}(\mathbb{R}^{d_S})$

(a1') $\alpha \in (0,1]$, $\gamma > 1/\alpha$, $a \in \mathrm{Lip}^1(\mathbb{R}^{d_S})$, $b_1, \dots, b_{d_B} \in \mathrm{Lip}^1(\mathbb{R}^{d_S})$ and $c_1, \dots, c_{d_Y} \in \mathrm{Lip}^{\gamma+3}(\mathbb{R}^{d_S})$).

(a2) $\alpha \in (1/3, 1/2)$, $\gamma > 1/\alpha$, $a \in \mathrm{Lip}^1(\mathbb{R}^{d_S})$, $b_1, \dots, b_{d_B} \in \mathrm{Lip}^{\gamma}(\mathbb{R}^{d_S})$ and $c_1, \dots, c_{d_Y} \in \mathrm{Lip}^{\gamma}(\mathbb{R}^{d_S})$).

Theorem 1. *Under assumption* **(a1)**, **(a1')** *or* **(a2)**, *there exists a d_S-dimensional process $S^\infty \in \mathcal{S}^0$ such that*

$$S^n \to S^\infty, \qquad \textit{in } \mathcal{S}^0.$$

In addition, the limit $\Xi(\boldsymbol{\eta}) := S^\infty$ only depends on $\boldsymbol{\eta}$ and not on the approximating sequence. Moreover, for all $q \geq 1$, $\boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}$ it holds that $\Xi(\boldsymbol{\eta}) \in \mathcal{S}^q$ and the corresponding mapping $\Xi : \mathcal{C}^{0,\alpha} \to \mathcal{S}^q$ is locally uniformly continuous (and locally Lipschitz under assumption **(a1')** *or* **(a2)***).*

Following Theorem 1, we say that $\Xi(\boldsymbol{\eta})$ is a solution of the *SDE with rough drift*

$$\Xi(\boldsymbol{\eta})_t = S_0 + \int_0^t a(\Xi(\boldsymbol{\eta})_r)dr + \int_0^t b(\Xi(\boldsymbol{\eta})_r)d\bar{B}_r + \int_0^t c(\Xi(\boldsymbol{\eta})_r)d\boldsymbol{\eta}_r. \tag{5}$$

The following result establishes some of the salient properties of solutions of SDEs with rough drift. Let $(\Omega, \mathcal{F}, \mathbb{P}_0)$ carry, a d_Y-dimensional Brownian motion Y and let $\hat{\Omega} = \Omega \times \bar{\Omega}$ be the product space, with product measure $\hat{\mathbb{P}} := \mathbb{P}_0 \otimes \bar{\mathbb{P}}$. Let S be the unique solution on this probability space to the SDE

$$S_t = S_0 + \int_0^t a(S_r)dr + \int_0^t b(S_r)d\bar{B}_r + \int_0^t c(S_r) \circ dY_r. \tag{6}$$

Denote by $\boldsymbol{Y}$ the rough path lift of Y (i.e. the enhanced Brownian Motion over Y).

Theorem 2. *Under assumption* **(a1)**, **(a1')** *or* **(a2)** *we have that*

- *For every $R > 0, q \geq 1$*

$$\sup_{||\boldsymbol{\eta}||_{\alpha-H\ddot{o}l}<R} \bar{\mathbb{E}}[\exp(q|\Xi(\boldsymbol{\eta})|_{\infty;[0,t]})] < \infty. \tag{7}$$

- *For $\mathbb{P}_0 - a.e.\ \omega$*

$$\bar{\mathbb{P}}[S_s(\omega, \cdot) = \Xi(\boldsymbol{Y}(\omega))_s(\cdot), \quad s \leq t] = 1. \tag{8}$$

We split the proof in two cases. The first case uses a flow transformation, similar to what is done in Section 3 for BSDEs with rough drift, and works for assumption (a1) or (a1').

In the second case, under assumption (a2), we create a joint rough path lift of the Brownian motion $\bar{B}$ and the deterministic rough path $\boldsymbol{\eta}$, and then solve a classical RDE with this driving signal. The major problem here, is to show some integrability properties.

Proof of Case 1

Proof of Theorem 1. Let $\boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}$ be the lift of a smooth path η. Let $S^{\boldsymbol{\eta}}$ be the unique solution of the SDE

$$S^{\boldsymbol{\eta}}_t = S_0 + \int_0^t a(S^{\boldsymbol{\eta}}_r)dr + \int_0^t b(S^{\boldsymbol{\eta}}_r)d\bar{B}_r + \int_0^t c(S^{\boldsymbol{\eta}}_r)d\eta_r.$$

Define $\tilde{S}^{\boldsymbol{\eta}} := (\phi^{\boldsymbol{\eta}})^{-1}(t, S^{\boldsymbol{\eta}}_t)$, where $\phi^{\boldsymbol{\eta}}$ is the ODE flow

$$\phi^{\boldsymbol{\eta}}(t,x) = x + \int_0^t c(\phi^{\boldsymbol{\eta}}(r,x))d\boldsymbol{\eta}_r. \tag{9}$$

By Lemma 3, we have that $\tilde{S}^{\boldsymbol{\eta}}$ satisfies the SDE

$$\tilde{S}^{\boldsymbol{\eta}}_t = S_0 + \int_0^t \tilde{a}^{\boldsymbol{\eta}}(r, \tilde{S}^{\boldsymbol{\eta}}_r)dr, + \int_0^t \tilde{b}^{\boldsymbol{\eta}}(r, \tilde{S}^{\boldsymbol{\eta}}_r)d\bar{B}_r \tag{10}$$

with $\tilde{a}^{\boldsymbol{\eta}}, \tilde{b}^{\boldsymbol{\eta}}$ defined as in Lemma 3.

This equation makes sense, even if $\boldsymbol{\eta}$ is a generic rough path in $\mathcal{C}^{0,\alpha}$, (in which case (9) is now really an RDE). Indeed, since the first two derivatives of $\phi^{\boldsymbol{\eta}}$ and its inverse are bounded (Proposition 11.11 in [33]) we have that $\tilde{a}^{\boldsymbol{\eta}}(t,\cdot), \tilde{b}^{\boldsymbol{\eta}}(t,\cdot)$ are also in Lip^1. Hence by Theorem V.7 in [59], there exists a unique strong solution to (10).

We define the mapping introduced in Theorem 1 as

$$\Xi(\boldsymbol{\eta})_t := \phi^{\boldsymbol{\eta}}(t, \tilde{S}^{\boldsymbol{\eta}}_t).$$

To show continuity of the mapping we restrict ourselves to the case $q = 2$. Moreover we shall assume $c_1, \dots, c_{d_Y} \in \mathrm{Lip}^{\gamma+3}(\mathbb{R}^{d_S})$, and we will hence prove the local Lipschitz property of the respective maps.

Let $\boldsymbol{\eta}^1, \boldsymbol{\eta}^2 \in \mathcal{C}^{0,\alpha}$ with $|\boldsymbol{\eta}^1|_{\alpha-\text{Höl}}, |\boldsymbol{\eta}^2|_{\alpha-\text{Höl}} < R$. By Lemma 6 we have

$$\bar{\mathbb{E}}[\sup_{s\le t} |\tilde{S}^1_s - \tilde{S}^2_s|^2]^{1/2} \le C_{Lem6}(R)\rho_{\alpha-\text{Höl}}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2).$$

Hence

$$\begin{aligned}
\bar{\mathbb{E}}[\sup_{s\le t} |\Xi(\boldsymbol{\eta}^1)_s - \Xi(\boldsymbol{\eta}^2)_s|^2]^{1/2} &= \bar{\mathbb{E}}[\sup_{s\le t} |\phi^1(s,\tilde{S}^1_s) - \phi^2(s,\tilde{S}^2_s)|^2]^{1/2} \\
&\le \bar{\mathbb{E}}[\sup_{s\le t} |\phi^1(s,\tilde{S}^1_s) - \phi^1(s,\tilde{S}^2_s)|^2]^{1/2} + \bar{\mathbb{E}}[\sup_{s\le t} |\phi^1(s,\tilde{S}^2_s) - \phi^2(s,\tilde{S}^2_s)|^2]^{1/2} \\
&\le \bar{\mathbb{E}}[\sup_{s\le t} |\phi^1(s,\tilde{S}^1_s) - \phi^1(s,\tilde{S}^2_s)|^2]^{1/2} + \sup_{s\le t, x\in\mathbb{R}^{d_S}} |\phi^1(s,x) - \phi^1(s,x)| \\
&\le K(R)\bar{\mathbb{E}}[\sup_{s\le t} |\tilde{S}^1_s - \tilde{S}^2_s|^2]^{1/2} + C_{Lem7}(R)\rho_{\alpha-\text{Höl}}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2) \\
&\le C_1 \rho_{\alpha-\text{Höl}}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2),
\end{aligned}$$

as desired, where

$$C_1 = K_{Lem7}(R)C_{Lem6}(R) + C_{Lem7}(R),$$

where $K_{Lem7}(R)$ and $C_{Lem7}(R)$ are the constants from Lemma 7. □

Proof of Theorem 2. In order to show (7), pick $k \in 1, \dots, d_S$. We first note that by simply scaling (the coefficients of) $S^{\boldsymbol{\eta}}$ it is sufficient to argue for $q = 1$. And consider the k-th component of Ξ.

Then

$$
\begin{aligned}
\mathbb{E}[\exp(|\Xi^{(k)}(\boldsymbol{\eta})|_{\infty;[0,t]})] &\le \mathbb{E}[\exp(|D\psi^{\boldsymbol{\eta}}|_\infty \left(|\phi^{\boldsymbol{\eta}}(0,S_0)| + |\tilde{S}^{(k);\boldsymbol{\eta}}|_{\infty;[0,t]}\right))] \\
&\le \exp(|D\psi^{\boldsymbol{\eta}}|_\infty \sup_{|x|\le|S_0|_{L^\infty}} |\phi^{\boldsymbol{\eta}}(0,x))\mathbb{E}[\exp(|D\psi^{\boldsymbol{\eta}}|_\infty \sup_{s\le t} |\tilde{S}_t^{(k);\boldsymbol{\eta}}|)] \\
&\le \exp(|D\psi^{\boldsymbol{\eta}}|_\infty \sup_{|x|\le|S_0|_{L^\infty}} |\phi^{\boldsymbol{\eta}}(0,x)) \\
&\quad \times \left(\mathbb{E}[\exp(|D\psi^{\boldsymbol{\eta}}|_\infty \sup_{s\le t} \tilde{S}_s^{(k);\boldsymbol{\eta}})] + \mathbb{E}[-\exp(|D\psi^{\boldsymbol{\eta}}|_\infty \sup_{s\le t} \tilde{S}_s^{(k);\boldsymbol{\eta}})]\right) \\
&= \exp(|D\psi^{\boldsymbol{\eta}}|_\infty \sup_{|x|\le|S_0|_{L^\infty}} |\phi^{\boldsymbol{\eta}}(0,x)) \\
&\quad \times \left(\mathbb{E}[\sup_{s\le t} \exp(|D\psi^{\boldsymbol{\eta}}|_\infty \tilde{S}_s^{(k);\boldsymbol{\eta}})] + \mathbb{E}[-\sup_{s\le t} \exp(|D\psi^{\boldsymbol{\eta}}|_\infty \tilde{S}_s^{(k);\boldsymbol{\eta}})]\right).
\end{aligned}
$$

Now, only the boundedness of the last two terms remains to be shown, for $\boldsymbol{\eta}$ bounded.

By applying Itô's formula we get that

$$
\begin{aligned}
\exp(\tilde{S}_t^{(k);\boldsymbol{\eta}}) &= 1 + \int_0^t \exp(\tilde{S}_r^{(k);\boldsymbol{\eta}}) d\tilde{S}_r^{(k);\boldsymbol{\eta}} + \int_0^t \exp(\tilde{S}_r^{(k);\boldsymbol{\eta}}) d\langle \tilde{S}^{(k);\boldsymbol{\eta}}\rangle_r \\
&= 1 + \int_0^t \exp(\tilde{S}_r^{(k);\boldsymbol{\eta}}) \tilde{a}_k^{\boldsymbol{\eta}}(\tilde{S}_r^{(k);\boldsymbol{\eta}}) dr + \sum_{i=1}^{d_B} \int_0^t \exp(\tilde{S}_r^{(k);\boldsymbol{\eta}}) \tilde{b}_{ki}^{\boldsymbol{\eta}}(\tilde{S}_r^{(k);\boldsymbol{\eta}}) d\bar{B}_r^i \\
&\quad + \sum_{i=1}^{d_B} \int_0^t \exp(\tilde{S}_r^{(k);\boldsymbol{\eta}}) |\tilde{b}_{ki}^{\boldsymbol{\eta}}(\tilde{S}_r^{(k);\boldsymbol{\eta}})|^2 dr.
\end{aligned}
$$

Hence the process $\exp(\tilde{S}^{(k);\boldsymbol{\eta}})$ satisfies a SDE with Lipschitz coefficients and by an application of Gronwall's Lemma and Burkholder-Davis-Gundy inequality (see also Lemma V.2 in [59]) one arrives at

$$
\sup_{|\boldsymbol{\eta}|_{\alpha-\text{Höl}}<R} \sup_{s\le t} \bar{\mathbb{E}}[\exp(|D\psi^{\boldsymbol{\eta}}|_\infty \tilde{S}_t^{(k);\boldsymbol{\eta}})] \le C\exp(C_2), \tag{11}
$$

where C is universal and

$$
C_2 := \sup_{\boldsymbol{\eta}:||\boldsymbol{\eta}||_{\alpha-\text{Höl}}<R} |\tilde{a}^{\boldsymbol{\eta}}|_\infty + |\tilde{b}^{\boldsymbol{\eta}}|_\infty^2,
$$

which is finite because of Lemma 4. One argues analogously for $\sup_{s\le t} \bar{\mathbb{E}}[-\exp(|D\psi^{\boldsymbol{\eta}}|_\infty \tilde{S}_t^{(k);\boldsymbol{\eta}})]$, which then gives (7).

Now, for the correspondence to an SDE solution let Ω be the additional probability space as given in the statement. Let S be the solution to the SDE (6).

In Section 3 in [8] it was shown (see also Theorem 2 in [1]), that if we let Θ be the stochastic (Stratonovich) flow

$$
\Theta(\omega; t, x) = x + \int_0^t c(\Theta(\omega; r, x)) \circ dY_r(\omega),
$$

then with $\hat{S}_t := \Theta^{-1}(t, S_t)$ we have $\hat{\mathbb{P}}$-a.s.

$$\hat{S}_s(\omega, \omega^{\bar{B}}) = S_0 + \int_0^s \hat{a}(r, \hat{S}_r)dr + \int_0^s \hat{b}(r, \hat{S}_r)d\bar{B}_r, \qquad s \in [0,t], \qquad \hat{\mathbb{P}} \ a.e. \ (\omega, \omega^{\bar{B}}). \quad (12)$$

Here, componentwise,

$$\hat{a}(t,x)_i := \sum_k \partial_{x_k}\Theta_i^{-1}(t,\Theta(t,x))a_k(\Theta(t,x)) + \frac{1}{2}\sum_{j,k}\partial_{x_j x_k}\Theta_i^{-1}(t,\Theta(t,x))\sum_l b_{jl}(\Theta(t,x))b_{kl}(\Theta(t,x)),$$

$$\hat{b}(t,x)_{ij} := \sum_k \partial_{x_k}\Theta_i^{-1}(t,x)b_{kj}(\Theta(t,x)).$$

Especially, by a Fubini type theorem (e.g. Theorem 3.4.1 in [5]), there exists Ω_0 with $\mathbb{P}_0(\Omega_0) = 1$ such that for $\omega \in \Omega_0$ equation (12) holds true $\bar{\mathbb{P}}$ *a.s.*.

Let $\mathbf{Y} \in \mathcal{C}^{0,\alpha}$ be the enhanced Brownian motion over Y. We can then construct ω-wise the rough flow $\phi^{\mathbf{Y}(\omega)}$ as given in (9). By the very definition of Ξ we know that $\tilde{S}_t^{\mathbf{Y}(\omega)}(\omega) := (\phi^{\mathbf{Y}(\omega)})^{-1}(\omega; t, \Xi(\omega)_t)$ satisfies the SDE

$$\tilde{S}_t^{\mathbf{Y}(\omega)} = S_0 + \int_0^t \tilde{b}^{\mathbf{Y}(\omega)}(r, \tilde{S}_r^{\mathbf{Y}(\omega)})dr + \int_0^t \tilde{b}^{\mathbf{Y}(\omega)}(r, \tilde{S}_r^{\mathbf{Y}(\omega)})d\bar{B}_r, \qquad \bar{\mathbb{P}}a.e. \ \omega^{\bar{B}}, \quad (13)$$

where

$$\begin{aligned}\tilde{a}^{\mathbf{Y}(\omega)}(t,x)_i := &\sum_k \partial_{x_k}(\phi^{\mathbf{Y}(\omega)})_i^{-1}(t,\phi^{\mathbf{Y}(\omega)}(t,x))a_k(t,\phi^{\mathbf{Y}(\omega)}(t,x)) \\ &+ \frac{1}{2}\sum_{j,k}\partial_{x_j x_k}(\phi^{\mathbf{Y}(\omega)})_i^{-1}(t,\phi^{\mathbf{Y}(\omega)}(t,x))\sum_l b_{jl}(t,\phi^{\mathbf{Y}(\omega)}(t,x))b_{kl}(t,\phi^{\mathbf{Y}(\omega)}(t,x)), \\ \tilde{b}^{\mathbf{Y}(\omega)}(t,x)_{ij} := &\sum_k \partial_{x_k}(\phi^{\mathbf{Y}(\omega)})_i^{-1}(t,x)b_{kj}(t,\phi^{\mathbf{Y}(\omega)}(t,x)).\end{aligned}$$

It is a classical rough path result (see e.g. section 17.5 in [33]), that there exists Ω_1 with $\mathbb{P}^Y(\Omega_1) = 1$ such that for $\omega \in \Omega_1$, we have

$$\phi^{\mathbf{Y}(\omega)}(\cdot,\cdot) = \Theta(\omega;\cdot,\cdot).$$

Hence for $\omega \in \Omega_1$ we have that $\hat{a} = \tilde{a}^{\mathbf{Y}(\omega)}, \hat{b} = \tilde{b}^{\mathbf{Y}(\omega)}$. Hence for for $\omega \in \Omega_0 \cap \Omega_1$ the processes $\hat{S}_t(\omega,\cdot), \tilde{S}_t^{\mathbf{Y}(\omega)}(\cdot)$ satisfy the same Lipschitz SDE (with respect to $\bar{\mathbb{P}}$). [6] By strong uniqueness we hence have for $\omega \in \Omega_0 \cap \Omega_1$ that $\bar{\mathbb{P}}$ *a.s.*

$$\hat{S}_s(\omega,\cdot) = \tilde{S}_s^{\mathbf{Y}(\omega)}(\cdot), \quad s \le t.$$

Hence for $\omega \in \Omega_0 \cap \Omega_1$

$$S_s(\omega,\cdot) = \Xi(\mathbf{Y}(\omega))(\cdot)_s, \quad s \le t, \qquad \bar{\mathbb{P}} \ a.s.$$

□

[6]Here one has to argue, that fixing ω in equation (13) gives ($\mathbb{P}_0$-a.s.) the solution to the respective SDE on $\Omega^{\bar{B}}$. This is similar to the arguments given in Case 2 below, so we omit them.

Lemma 3. *Let η be a smooth d_Y-dim path η and S be the solution of the the following classical SDE*

$$S_t = S_0 + \int_0^t a(S_r)dr + \int_0^t b(S_r)d\bar{B}_r + \int_0^t c(S_r)d\eta_r$$

where $\bar{B}$ is a d_B-dimensional Brownian motion, $\int_0^t c(S_r)d\eta_r := \sum_{i=1}^{d_S}\int_0^t c_i(S_r)\dot{\eta}_r^i dr$, $a \in \mathrm{Lip}^1(\mathbb{R}^{d_S})$, $b_1,\dots,b_{d_B} \in \mathrm{Lip}^1(\mathbb{R}^{d_S})$, $c_1,\dots,c_{d_Y} \in \mathrm{Lip}^3(\mathbb{R}^{d_S})$, and $S_0 \in L^\infty(\bar{\Omega};\mathbb{R}^{d_S})$ independent of $\bar{B}$. Consider the flow

$$\phi(t,x) = x + \int_0^t c(\phi(r,x))d\eta_r. \tag{14}$$

Then $\tilde{S}_t := \phi^{-1}(t,S_t)$ satisfies the following SDE

$$\tilde{S}_t = S_0 + \int_0^t \tilde{a}(r,\tilde{S}_r)dr + \int_0^t \tilde{b}(r,\tilde{S}_r)d\bar{B}_r,$$

where we define componentwise

$$\tilde{a}(t,x)_i := \sum_k \partial_{x_k}\phi_i^{-1}(t,\phi(t,x))a_k(\phi(t,x)) + \frac{1}{2}\sum_{j,k}\partial_{x_jx_k}\phi_i^{-1}(t,\phi(t,x))\sum_l b_{jl}(\phi(t,x))b_{kl}(\phi(t,x)),$$
$$\tilde{b}(t,x)_{ij} := \sum_k \partial_{x_k}\phi_i^{-1}(t,\phi(t,x))b_{kj}(\phi(t,x)).$$

Proof. Denote $\psi(t,x) := \phi^{-1}(t,x)$. Then

$$\psi(r,x) = x - \int_0^t \partial_x\psi(r,x)c(x)d\eta_r.$$

By Itô's formula

$$\begin{aligned}
\psi_i(t,S_t) - \psi_i(0,S_0) &= \int_0^t \partial_t\psi_i(r,S_r)dr + \sum_j \int_0^t \partial_{x_j}\psi_i(r,S_r)dS_j(r) \\
&\quad + \sum_{j,k}\frac{1}{2}\int_0^t \partial_{x_jx_k}\psi_i(r,S_r)d\langle S_k,S_j\rangle_r \\
&= \sum_j \int_0^t \partial_{x_j}\psi_i(r,S_r)a_j(S_r)dr + \sum_j \int_0^t \partial_{x_j}\psi_i(r,S_r)\sum_k b_{jk}(S_r)d\bar{B}_k(r) \\
&\quad + \sum_{j,k}\frac{1}{2}\int_0^t \partial_{x_jx_k}\psi_i(S_r)\sum_l b_{kl}(S_r)b_{jl}(S_r)dr.
\end{aligned}$$

□

Lemma 4. *Consider for a rough path $\boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}$ the analogously to Lemma 3 transformed coefficients, $\tilde{a}^{\boldsymbol{\eta}}$, $\tilde{b}^{\boldsymbol{\eta}}$, i.e. consider the rough flow*

$$\phi(t,x) = \phi^{\boldsymbol{\eta}}(t,x) = x + \int_0^t c(\phi(r,x))d\boldsymbol{\eta}_r. \tag{15}$$

and define

$$\tilde{a}^{\boldsymbol{\eta}}(t,x)_i := \sum_k \partial_{x_k}\phi_i^{-1}(t,\phi(t,x))a_k(\phi(t,x)) + \frac{1}{2}\sum_{j,k}\partial_{x_j x_k}\phi_i^{-1}(t,\phi(t,x))\sum_l b_{jl}(\phi(t,x))b_{kl}(\phi(t,x)),$$
$$\tilde{b}^{\boldsymbol{\eta}}(t,x)_{ij} := \sum_k \partial_{x_k}\phi_i^{-1}(t,\phi(t,x))b_{kj}(\phi(t,x)).$$

Then for every $R > 0$ there exists $K_{Lem4} = K_{Lem4}(R) < \infty$ such that

$$\sup_{\boldsymbol{\eta}:|\boldsymbol{\eta}|_{\alpha-H\ddot{o}l}<R} |\tilde{a}^{\boldsymbol{\eta}}|_\infty \le K_{Lem4},$$
$$\sup_{\boldsymbol{\eta}:|\boldsymbol{\eta}|_{\alpha-H\ddot{o}l}<R} |\tilde{b}^{\boldsymbol{\eta}}|_\infty \le K_{Lem4},$$
$$\sup_{\boldsymbol{\eta}:|\boldsymbol{\eta}|_{\alpha-H\ddot{o}l}<R}\ \sup_{s\le t} |D\tilde{a}^{\boldsymbol{\eta}}(s,\cdot)|_\infty \le K_{Lem4},$$
$$\sup_{\boldsymbol{\eta}:|\boldsymbol{\eta}|_{\alpha-H\ddot{o}l}<R}\ \sup_{s\le t} |D\tilde{b}^{\boldsymbol{\eta}}(s,\cdot)|_\infty \le K_{Lem4},$$

and such that if $\boldsymbol{\eta}, \tilde{\boldsymbol{\eta}}$ are two rough paths with $|\boldsymbol{\eta}|_{\alpha-H\ddot{o}l}, |\tilde{\boldsymbol{\eta}}|_{\alpha-H\ddot{o}l} < R$ we have

$$\sup_{t,x} |\tilde{a}^1(t,x) - \tilde{a}^2(t,x)| \le K_{Lem4}(R)\rho_{\alpha-H\ddot{o}l}(\boldsymbol{\eta}^1,\boldsymbol{\eta}^2),$$
$$\sup_{t,x} |\tilde{b}^1(t,x) - \tilde{b}^2(t,x)| \le K_{Lem4}(R)\rho_{\alpha-H\ddot{o}l}(\boldsymbol{\eta}^1,\boldsymbol{\eta}^2).$$

Proof. This is a straightforward calculation using Lemma 7 and the properties of a, b. □

The following is a standard result for continuous dependence of SDEs on parameters.

Lemma 5. *Let $\tilde{a}^i(t,x)$, $\tilde{b}^i(t,x), i = 1,2$ be bounded and uniformly Lipschitz in x. Let $\tilde{S}^i$ be the corresponding unique solutions to the SDEs*

$$\tilde{S}^i_t = S_0 + \int_0^t \tilde{a}^i(r,\tilde{S}^i_r)dr + \int_0^t \tilde{b}^i(r,\tilde{S}^i_r)d\bar{B}_r, \qquad i = 1,2.$$

Assume

$$\sup_{s\le t}|D\tilde{a}^1(s,\cdot)|_\infty, \sup_{s\le t}|D\tilde{b}^1(s,\cdot)|_\infty < K < \infty,$$
$$\sup_{r,x}\left(\tilde{a}^1(r,x) - \tilde{a}^2(r,x)\right), \sup_{r,x}\left(\tilde{b}^1(r,x) - \tilde{b}^2(r,x)\right) < \varepsilon < \infty.$$

Then there exists $C_{Lem5} = C_{Lem5}(K)$ such that

$$\mathbb{E}[\sup_{s\le t}|\tilde{S}^1_s - \tilde{S}^2_s|^2]^{1/2} \le C_{Lem5}\ \varepsilon,$$

Proof. This is a straightforward application of Ito's formula and Burkholder-Davis-Gundy inequality. □

We now apply the previous Lemma to our concrete setting.

Lemma 6. *Let $\boldsymbol{\eta}^1, \boldsymbol{\eta}^2 \in \mathcal{C}^{0,\alpha}$ and let $\tilde{S}^1, \tilde{S}^2$ be the corresponding unique solutions to the SDEs*

$$\tilde{S}_t^i = S_0 + \int_0^t \tilde{a}^i(r, \tilde{S}_r^i)dr + \int_0^t \tilde{b}^i(r, \tilde{S}_r^i)d\bar{B}_r, \qquad i = 1, 2,$$

where $\tilde{a}^i, \tilde{b}^i$ are given as in Lemma 4.
Assume $R > \max\{|\boldsymbol{\eta}^1|_{\alpha-H\ddot{o}l}, |\boldsymbol{\eta}^2|_{\alpha-H\ddot{o}l}\}$. Then there exists $C_{Lem6} = C_{Lem6}(R)$.

$$\mathbb{E}[\sup_{s\le t} |\tilde{S}_s^1 - \tilde{S}_s^2|^2]^{1/2} \le C_{Lem6}\rho_{\alpha-H\ddot{o}l}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2).$$

Proof. Fix $R > 0$. Let $||\boldsymbol{\eta}^1||_{\alpha-\text{Höl}}, ||\boldsymbol{\eta}^2||_{\alpha-\text{Höl}} < R$.
From Lemma 4 we know that

$$\sup_{t,x} |\tilde{b}^1(t,x) - \tilde{b}^2(t,x)| \le K_{Lem4}(R)\rho_{\alpha-\text{Höl}}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2),$$

Analogously, we get

$$\sup_{t,x} |\tilde{a}^1(t,x) - \tilde{a}^2(t,x)| \le L_2\rho_{\alpha-\text{Höl}}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2),$$

for a $L_2 = L_2(R)$. □

Lemma 7. *Let $\alpha \in (0,1)$. Let $\gamma > \frac{1}{\alpha} \ge 1$, $k \in \{1, 2, \dots\}$ and assume that $V = (V_1, \dots, V_d)$ is a collection of $\text{Lip}^{\gamma+k}$-vector fields on $\mathbb{R}^e$. Write $n = (n_1, \dots, n_e) \in \mathbb{N}^e$ and assume $|n| := n_1 + \dots + n_e \le k$.*
Then, for all $R > 0$ there exist $C = C(R, |V|_{\text{Lip}^{\gamma+k}}), K = K(R, |V|_{\text{Lip}^{\gamma+k}})$ such that if $\mathbf{x}^1, \mathbf{x}^2 \in C^{\alpha-H\ddot{o}l}([0,t], G^{[p]}(\mathbb{R}^d))$ with $\max_i ||\mathbf{x}^i||_{\alpha-H\ddot{o}l;[0,t]} \le R$ then

$$\sup_{y_0\in\mathbb{R}^e} \left|\partial_n \pi_{(V)}(0, y_0; \mathbf{x}^1) - \partial_n \pi_{(V)}(0, y_0; \mathbf{x}^2)\right|_{\alpha-H\ddot{o}l;[0,t]} \le C\rho_{\alpha-H\ddot{o}l}(\mathbf{x}^1, \mathbf{x}^2),$$
$$\sup_{y_0\in\mathbb{R}^e} \left|\partial_n \pi_{(V)}(0, y_0; \mathbf{x}^1)^{-1} - \partial_n \pi_{(V)}(0, y_0; \mathbf{x}^2)^{-1}\right|_{\alpha-H\ddot{o}l;[0,t]} \le C\rho_{\alpha-H\ddot{o}l}(\mathbf{x}^1, \mathbf{x}^2),$$
$$\sup_{y_0\in\mathbb{R}^e} \left|\partial_n \pi_{(V)}(0, y_0; \mathbf{x}^1)\right|_{\alpha-H\ddot{o}l;[0,t]} \le K,$$
$$\sup_{y_0\in\mathbb{R}^e} \left|\partial_n \pi_{(V)}(0, y_0; \mathbf{x}^1)^{-1}\right|_{\alpha-H\ddot{o}l;[0,t]} \le K.$$

Proof. The fact that $V \in \text{Lip}^{\gamma+k}$ (instead of just $\text{Lip}^{\gamma+k-1}$) entails that the derivatives up to order k are unique non-explosive solutions to RDEs with $\text{Lip}^{\gamma}_{loc}$ vector fields (see Section 11 in [33]). Localization (uniform for driving paths bounded in α-Hölder norm) then yields the desired results. □

Proof of Case 2

Proof of Theorem 1. We will first define the solution mapping and only in the end we will see that it is the limit of classical SDEs.

Write $\eta = \pi_1(\boldsymbol{\eta})$ for the projection of $\boldsymbol{\eta}$ to the first level. We define a path $\boldsymbol{\lambda} = \boldsymbol{\lambda}\left(\boldsymbol{\eta}, \omega^{\bar{B}}\right)$ taking values in $G^2(\mathbb{R}^d)$, $d := d_Y + d_B$. The first level is given as

$$\boldsymbol{\lambda}^{(1)} = \left(\eta, \bar{B}\left(\omega^{\bar{B}}\right)\right)$$

The elements $\boldsymbol{\lambda}^{(2);i,j}$ of the second level are defined as $\boldsymbol{\eta}^{(2);i,j}$ for $i,j \in \{1,\ldots,d_1\}$, $\bar{\boldsymbol{B}}^{(2);i,j}\left(\omega^{\bar{B}}\right)$ for $i,j \in \{d_1+1,\ldots,d_1+d_2\}$ and the remaining second-level entries as

$$\begin{aligned}
\boldsymbol{\lambda}_t^{(2);i,j} &= \int_0^t \eta_u^i d\bar{B}_u^j \qquad , \text{for } i \in \{1,\ldots,d_1\}, j \in \{1+d_1,\ldots,d_1+d_2\}, \\
\boldsymbol{\lambda}^{(2);i,j} &= \eta_t^j \bar{B}_t^i - \int_0^t \eta_u^j d\bar{B}_u^i \qquad , \text{otherwise.}
\end{aligned}$$

The Campbell-Baker-Hausdorff formula (Theorem 7.24 in [33]) then shows that $\boldsymbol{\lambda}_{s,t} := \boldsymbol{\lambda}_s^{-1} \otimes \boldsymbol{\lambda}_t = \exp\left[\left(\eta_{s,t}, \bar{B}_{s,t}\right) + A_{s,t}\right]$ where

$$so(d) \ni A_{s,t}^{i,j} := \begin{cases}
\frac{1}{2}\left(\int_s^t \eta_{s,u}^i d\eta_u^j - \int_s^t \eta_{s,u}^j d\eta_u^i\right) & , \text{for } i,j \in (1,\cdots,d_Y) \\
\frac{1}{2}\left(\int_s^t \bar{B}_{s,u}^{i-d_Y} d\bar{B}_u^{j-d_Y} - \int_s^t \bar{B}_{s,u}^{j-d_Y} d\bar{B}_u^{i-d_Y}\right) & , \text{for } i,j \in (d_Y+1,\cdots,d_Y+d_B) \\
\left(\int_s^t \eta_{s,u}^i d\bar{B}_u^{j-d_Y} - \frac{1}{2}\eta_{s,t}^i \bar{B}_{s,t}^{j-d_Y}\right) & , \text{for } i \in (1,\cdots,d_Y), j \in (d_Y+1,\cdots,d_Y+d_B) \\
\left(-\int_0^t \eta_u^j d\bar{B}_u^{i-d_Y} + \frac{1}{2}\eta_{s,t}^j \bar{B}_{s,t}^{i-d_Y}\right) & , \text{for } i \in (d_Y+1,\cdots,d_Y+d_B), j \in (1,\cdots,d_Y)
\end{cases}$$

That is, $\boldsymbol{\lambda}_t$ takes values in $G^2(\mathbb{R}^d)$. We demonstrate that $\boldsymbol{\lambda}\left(\boldsymbol{\eta}, \omega^{\bar{B}}\right) \in C^{0,\alpha'-\text{Höl}}\left(G^2(\mathbb{R}^d)\right)$ ($\mathbb{P}$-a.s.) for $\alpha' < \alpha$ by showing that $|\boldsymbol{\lambda}|_{\alpha'}$ even has Gaussian tails.

It is clear that the first level of $\boldsymbol{\lambda}$ is α-Hölder; to deal with the second level note that $|A_{st}| \sim \sum_{ij}\left|A_{st}^{ij}\right|$ and that $2\alpha-$Hölder regularity already holds for the first two cases, that is $i,j \in \{1,\ldots,d_1\}$ and $i,j \in \{d_1+1,\ldots,d_1+d_2\}$ (in fact even a Gauss tail via the Fernique estimate for rough path norms). Now for the remaining case we have by definition that

$$\left|A_{s,t}^{ij}\right| \quad \leq \quad \left|\int_s^t \eta_{s,u}^i d\overline{B}_u^j\right| + \frac{1}{2}\left|\eta_{s,t}^i \overline{B}_{s,t}^j\right|$$

and since $\left|\eta_{s,t}^i \overline{B}_{s,t}^j\right| \leq |\eta|_{\alpha-\text{Höl}}\left|\overline{B}\right|_{\alpha-\text{Höl}}|t-s|^{2\alpha}$ it just remains to treat $\left|\int_s^t \eta_{s,u}^i d\overline{B}_u^j\right|$. But since for each $s<t$, $\int_s^t \eta_{s,u}^i d\overline{B}_u^j$ has the same distribution as $\sqrt{\int_s^t \left(\eta_{s,u}^i\right)^2 du}.Z$ for some fixed $Z \sim \mathcal{N}(0,1)$, we also have

$$\exp\left[\eta\left(\frac{\left|\int_s^t \left(\eta_{s,u}^i\right) d\overline{B}_u^j\right|}{(t-s)^{2\alpha}}\right)^2\right] =^{Law} \exp\left[\eta Z^2\left(\frac{\sqrt{\left|\int_s^t \left(\eta_{s,u}^i\right)^2 du\right|}}{(t-s)^{2\alpha}}\right)^2\right]$$

and by the elementary estimate $\left|\int_s^t \left(\eta_{s,u}^i\right)^2 du\right| \leq |\eta|_{\alpha-\text{Höl}}(t-s)^{2\alpha+1}$ we can conclude by taking $\sup_{s<t}\mathbb{E}$ in above expression and using the Gaussian integrability for Z. This then reads as

$$\sup_{s<t}\mathbb{E}\exp\left[\kappa Z^2 |\eta|_{\alpha-\text{Höl}}(t-s)^{1/2-\alpha}\right] < \infty$$

for some $\kappa > 0$. By Theorem A.19 in [33] this already yields the desired $2\alpha'$-Hölder regularity of $\left|\int_s^t \eta^i_{s,u} d\bar{B}^j_u\right|$ for any $\alpha' < \alpha$. Putting everything together, we have shown that $\|\boldsymbol{\lambda}\|_{\alpha'-\text{Höl}} < \infty$.

We can actually deduce for every $R > 0$ the existence of $b = b(\alpha', \alpha, T, R) > 0$ such that

$$\sup_{|\boldsymbol{\eta}|_\alpha \le R} \mathbb{E}\exp(b|\boldsymbol{\lambda}|^2_{\alpha'-\text{Höl}}) < \infty, \tag{16}$$

So indeed $\boldsymbol{\lambda} \in C^{\alpha'-\text{Höl}}\left(G^2(\mathbb{R}^d)\right), \mathbb{P} - a.s.$ for all $\alpha' < \alpha$. Hence (see e.g. Corollary 8.24 in [33]) $\boldsymbol{\lambda} \in C^{0,\alpha'-\text{Höl}}\left(G^2(\mathbb{R}^d)\right), \mathbb{P} - a.s.$ for all $\alpha' < \alpha$.

Using $\alpha' < \alpha$ large enough, such that $\gamma > 1/\alpha'$, we can then use standard existence and uniqueness for the solution of the RDE

$$S^{\boldsymbol{\lambda}}_t = S_0 + \int_0^t \mathring{a}(S^{\boldsymbol{\lambda}}_r)dr + \int_0^t V(S^{\boldsymbol{\lambda}}_r)d\boldsymbol{\lambda}_r, \tag{17}$$

where $V = (c, b)$ and $\mathring{a}^k(x) := a^k(x) + \sum_{i=1}^{d_B} \langle Db^k_i(x), b_i(x)\rangle$. [7] We then define the mapping as $\Xi(\boldsymbol{\eta}) := S^{\boldsymbol{\lambda}}_t$.

We proceed to show continuity of the mapping of a rough path $\boldsymbol{\eta}$ to the lift $\boldsymbol{\lambda}(\boldsymbol{\eta})$. Fix $\alpha' < \alpha$ such that $\gamma > 1/\alpha'$. We shall use Theorem A.13 in [33]. Let then $q \ge q_0(\alpha, \alpha')$, as given in this Theorem. We show that for $\boldsymbol{\eta}, \bar{\boldsymbol{\eta}} \in \mathcal{C}^{0,\alpha}$ with $|\boldsymbol{\eta}|_{\alpha-\text{Höl}}, |\bar{\boldsymbol{\eta}}|_{\alpha-\text{Höl}} \le R$ we have for the corresponding lifts $\boldsymbol{\lambda}, \bar{\boldsymbol{\lambda}}$ that

$$\mathbb{E}[\rho_{\alpha'}(\boldsymbol{\lambda}, \bar{\boldsymbol{\lambda}})^q]^{1/q} \le C_{Lip}\rho_\alpha(\boldsymbol{\eta}, \bar{\boldsymbol{\eta}}),$$

where $C_{Lip} = C_{Lip}(q, R, \alpha, \alpha')$.

Denote $\varepsilon := \rho_{\alpha'}(\boldsymbol{\lambda}, \bar{\boldsymbol{\lambda}})$. By (16) exists a constant $C_1 = C_1(q, R)$ such that

$$\mathbb{E}[|d(\boldsymbol{\lambda}_s, \boldsymbol{\lambda}_t)|^q], \mathbb{E}[|d(\bar{\boldsymbol{\lambda}}_s, \bar{\boldsymbol{\lambda}}_t)|^q] \le C_1|t-s|^{\alpha q}.$$

Moreover

$$\begin{aligned}\mathbb{E}[|\pi_1(\boldsymbol{\lambda}_{s,t} - \bar{\boldsymbol{\lambda}}_{s,t})|^q] &= |\eta_{s,t} - \bar{\eta}_{s,t}|^q \\ &\le \varepsilon^q|t-s|^{\alpha q}.\end{aligned}$$

and (again the constants C may only depend on R and q)

$$\begin{aligned}\mathbb{E}[|\pi_2(\boldsymbol{\lambda}_{s,t} - \bar{\boldsymbol{\lambda}}_{s,t})|^{q/2}] &= \mathbb{E}[|\frac{1}{2}\pi_1(\boldsymbol{\lambda}_{s,t} - \bar{\boldsymbol{\lambda}}_{s,t}) \otimes \pi_1(\boldsymbol{\lambda}_{s,t} - \bar{\boldsymbol{\lambda}}_{s,t}) + A_{s,t} - \bar{A}_{s,t}|^{q/2}] \\ &\le C\varepsilon^{q/2}|t-s|^{\alpha q} + C\mathbb{E}[|A_{s,t} - \bar{A}_{s,t}|^{q/2}]\end{aligned}$$

Also, since $n \ge N$

$$|\int_s^t \eta^i_{s,u}d\eta^j_u - \int_s^t \eta^j_{s,u}d\eta^i_u - \left(\int_s^t \bar{\eta}^i_{s,u}d\bar{\eta}^j_u - \int_s^t \bar{\eta}^j_{s,u}d\bar{\eta}^i_u\right)|^{q/2} \le \varepsilon^{q/2}|t-s|^{\alpha q}.$$

[7] This is a plus the Stratonovich corrector for the $d\bar{B}$ integral. Note that we only have $\mathring{a} \in \text{Lip}^1$ so we have to use results on RDEs with drift (e.g. Theorem 12.6 and Theorem 12.10 in [33]) to get existence of a unique solution.

And finally

$$\begin{aligned}
&\mathbb{E}[|\int_s^t \eta^i_{s,u} d\bar{B}^{j-d_Y}_u - \frac{1}{2}\eta^i_{s,t}\bar{B}^{j-d_Y}_{s,t} - \left(\int_s^t \bar{\eta}^i_{s,u} d\bar{B}^{j-d_Y}_u - \frac{1}{2}\bar{\eta}^i_{s,t}\bar{B}^{j-d_Y}_{s,t}\right)|^{q/2}] \\
&\le C\mathbb{E}[|\int_s^t \eta^i_{s,u} - \bar{\eta}^i_{s,u} d\bar{B}^{j-d_Y}_u|^{q/2}] + C\mathbb{E}[|\eta^i_{s,t}\bar{B}^{j-d_Y}_{s,t} - \bar{\eta}^i_{s,t}\bar{B}^{j-d_Y}_{s,t}|^{q/2}] \\
&\le C|\int_s^t |\eta^i_{s,u} - \bar{\eta}^i_{s,u}|^2 du|^{q/4} + C|\eta^i_{s,t} - \eta^i_{s,t}|^{q/2}\mathbb{E}[|\bar{B}^{j-d_Y}_{s,t}|^{q/2}] \\
&\le C\varepsilon^{q/2}(t-s)^{\alpha q/2+q/4} + C\varepsilon^{q/2}(t-s)^{\alpha q/2}(t-s)^{q/4} \\
&\le C\varepsilon^{q/2}(t-s)^{\alpha q}.
\end{aligned}$$

Hence

$$\mathbb{E}[|\pi_2(\boldsymbol{\lambda}_{s,t} - \bar{\boldsymbol{\lambda}}_{s,t})|^{q/2}] \le C_2\varepsilon^{q/2}|t-s|^{\alpha q}$$

So with $M = M(R,q) := \max\{1, C_1^{1/q}, C_2^{1/(2q)}\}$ we have for all $q \ge 1$

$$\begin{aligned}
&\mathbb{E}[|d(\boldsymbol{\lambda}_s, \bar{\boldsymbol{\lambda}}_t)|^q]^{1/q} \le M|t-s|^{\alpha}, \\
&\mathbb{E}[|\pi_1(\boldsymbol{\lambda}_{s,t} - \bar{\boldsymbol{\lambda}}_{s,t})|^q]^{1/q} \le \varepsilon M|t-s|^{\alpha}, \\
&\mathbb{E}[|\pi_2(\boldsymbol{\lambda}_{s,t} - \bar{\boldsymbol{\lambda}}_{s,t})|^{q/2}]^{2/q} \le \varepsilon M^2|t-s|^{2\alpha}.
\end{aligned}$$

Hence by Theorem A.13 (i) in [33] there exists q large enough and a $K = K(\alpha, \alpha', T, q)$ such that

$$\begin{aligned}
&\mathbb{E}[\left\{\sup_{s<t} \frac{|\pi_1(\boldsymbol{\lambda}_{s,t} - \bar{\boldsymbol{\lambda}}_{s,t}|}{|t-s|^{\alpha'}}\right\}^q]^{1/q} \le \varepsilon KM, \\
&\mathbb{E}[\left\{\sup_{s<t} \frac{|\pi_2(\boldsymbol{\lambda}_{s,t} - \bar{\boldsymbol{\lambda}}_{s,t}|}{|t-s|^{\alpha'}}\right\}^{q/2}]^{2/q} \le \varepsilon(KM)^2.
\end{aligned}$$

Now using this result for q and $2q$ we get in combination

$$\mathbb{E}[\rho_{\alpha'-\text{Höl}}(\boldsymbol{\lambda}, \bar{\boldsymbol{\lambda}})^q]^{1/q} \le \mathbb{E}[\left(\sup_{s\neq t} \frac{\pi_1(\boldsymbol{\lambda}_{s,t} - \bar{\boldsymbol{\lambda}}_{s,t})}{|t-s|^{\alpha'}}\right)^q]^{1/q} + \mathbb{E}[\left(\sup_{s\neq t} \frac{\pi_2(\boldsymbol{\lambda}_{s,t} - \bar{\boldsymbol{\lambda}}_{s,t})}{|t-s|^{2\alpha'}}\right)^q]^{1/q} \le C_{Lip}\varepsilon,$$

as desired. This consideration was for q large enough, but since L^r is Lipschitz continuously embedded in L^q for $r > q$, the result follows for all q.

We can now proceed to show local Lipschitzness of the solution mapping Ξ.

Let still $\boldsymbol{\eta}, \bar{\boldsymbol{\eta}} \in \mathcal{C}^{0,\alpha}$, Let $\alpha' < \alpha$ with $\gamma > 1/\alpha'$ and let $p' := 1/\alpha'$. By Theorem 8 in [13] we have the deterministic estimate [8], that for some C^*, any $\beta \in (0,1)$ there exists $C = C(C^*, \beta)$ such that

$$\rho_{p'-var}(\Xi(\boldsymbol{\eta}), \Xi(\bar{\boldsymbol{\eta}})) \le C\rho_{p'-var}(\boldsymbol{\lambda}, \bar{\boldsymbol{\lambda}})\exp(-N(\beta, C^*, \upsilon)\log(1-\beta)),$$

where υ, N are given as in Lemma 8 below.

[8]This estimate modifies the estimates in [33] in a way that is essential to get integrability in our context.

Hence, for every $q \geq 1$,

$$\begin{aligned}
\mathbb{E}[\sup_{t\leq T}|\Xi(\boldsymbol{\eta})_t - \Xi(\bar{\boldsymbol{\eta}})_t|^q]^{1/q} &\leq \mathbb{E}[\rho_{p'-var}(\Xi(\boldsymbol{\eta}),\Xi(\bar{\boldsymbol{\eta}}))^q]^{1/q} \\
&\leq C\mathbb{E}[\left[\rho_{p'-var}(\boldsymbol{\lambda},\bar{\boldsymbol{\lambda}})\right]^{2q}]^{1/(2q)}\mathbb{E}[[\exp(-N(\beta,C^*,\upsilon)\log(1-\beta))]^{2q}]^{1/(2q)} \\
&\leq C\mathbb{E}[\left[\rho_{\alpha'-\text{Höl}}(\boldsymbol{\lambda},\bar{\boldsymbol{\lambda}})\right]^{2q}]^{1/(2q)}\mathbb{E}[[\exp(-N(\beta,C^*,\upsilon)\log(1-\beta))]^{2q}]^{1/(2q)} \\
&\leq C\rho_{\alpha'-\text{Höl}}(\boldsymbol{\eta},\bar{\boldsymbol{\eta}})\mathbb{E}[[\exp(-N(\beta,C^*,\upsilon)\log(1-\beta))]^{2q}]^{1/(2q)}.
\end{aligned}$$

By Lemma 8 we know that for every $q \geq 1, R > 0$ there exists $\beta_0(q,R) > 0$ such that for $|\boldsymbol{\eta}|_{\alpha-\text{Höl}}, |\bar{\boldsymbol{\eta}}|_{\alpha-\text{Höl}} < R$, we have for $\beta \leq \beta_0$ that

$$\mathbb{E}[[\exp(-N(\beta,C^*,\upsilon)\log(1-\beta))]^q]^{1/q}$$

is bounded. This yields the desired local Lipschitzness of Ξ.

□

Proof of Theorem 2. Fix $R > 0$. We show

$$\sup_{||\boldsymbol{\eta}||_{\alpha-\text{Höl}}<R} \bar{\mathbb{E}}[\exp(q\Xi^{(k)}(\boldsymbol{\eta})_t)] < \infty. \tag{18}$$

By Theorem 10.14 in [33] we have

$$|S|_{p'-var;[s,t]} \leq C\left(||\boldsymbol{\lambda}||_{p'-var;[s,t]} \vee ||\boldsymbol{\lambda}||^{p'}_{p'-var;[s,t]}\right).$$

Define the control

$$\upsilon(s,t) := \left(||\boldsymbol{\lambda}||_{p'-var;[s,t]} \vee ||\boldsymbol{\lambda}||^{p'}_{p'-var;[s,t]}\right)^{p'}.$$

Then for any partition D of $[0,t]$

$$\begin{aligned}
||S||_{\infty;[0,t]} &\leq |S_0| + \sum_{t_i\in D}||S - S_{t_i}||_{\infty;[t_i,t_{i+1}]} \\
&\leq |S_0| + \sum_{t_i\in D}||S||_{p-var;[t_i,t_{i+1}]} \\
&\leq |S_0| + C\sum_{t_i\in D}\upsilon(t_i,t_{i+1})^{1/p'}.
\end{aligned}$$

Then, using for $a > 0$ the partition

$$\begin{aligned}
t_0(a) &:= 0, \\
t_{i+1}(a) &:= \inf\{s \geq t_i(a) : \upsilon(\sigma_i(a),s) \geq a^{p'}\} \vee t,
\end{aligned}$$

we get

$$||S||_{\infty;[0,t]} \leq |S_0| + Ca^{p'}\left(\sup\{n : t_n < t\} + 1\right).$$

The exponential integrability (18) then follows as in the proof of Lemma 8.

We now show correspondence to the SDE solution. On $\hat{\Omega}$ let $EBM(Y,\bar{B})$ be the enhanced Brownian motion lift of $(Y,\bar{B})$. We want to compare it to the mapping $\boldsymbol{\lambda}^{EBM(Y)}$, where we plug the enhanced Brownian motion lift of Y into the above constructed mapping. Note, that this composition is in general not a random variable on $\hat{\Omega}$, since joint measurability is not clear.

We know by definition that $\pi_1(\boldsymbol{\lambda}(EBM(Y)))$ is measurable and coincides with $\pi_1(EBM(Y,\bar{B}))$ $\hat{\mathbb{P}}$-a.s.

We now consider only one component of the second level, the others follow similarly. So without loss of generality let $i \in \{1,\dots,d_Y\}, j \in \{d_Y+1,\dots,d_Y+d_B\}$.

Then

$$EBM(Y,\bar{B})_t^{(2);i,j} = \int_0^t Y_r^i d\bar{B}_r^{j-d_Y} \qquad \hat{\mathbb{P}} - a.s.$$

Now by results on stochastic integrals we have

$$\left(\int_0^t Y_r^i d\bar{B}_r^{j-d_Y}\right)(\omega,\omega^{\bar{B}}) = \lim_{n\to\infty}\sum_{k=1}^{2^n} Y^k_{(k-1)/2^n t}(\omega)\left[B^{j-d_Y}_{k/2^n t}(\omega^{\bar{B}}) - B^{j-d_Y}_{(k-1)/2^n t}(\omega^{\bar{B}})\right]$$

on $\hat{A}\subset\hat{\Omega}$ with $\hat{\mathbb{P}}[\hat{A}]=1$. By a Fubini type theorem (e.g. Theorem 3.4.1 in [5]), there exists a $\mathring{\Omega}\subset\Omega$ such that for $\omega\in\mathring{\Omega}$ we have that $\hat{A}_\omega := \{\omega^{\bar{B}} : (\omega,\omega^{\bar{B}})\in\hat{A}\}$ satisfy $\mathbb{P}^{\bar{B}}[\hat{A}_\omega]=1$.

On the other hand for fixed $\omega\in\Omega$

$$\boldsymbol{\lambda}_t^{(2);i,j}(EBM(Y)(\omega)) = \int_0^t Y_r^i(\omega) d\bar{B}_r^{j-d_1}.$$

Now

$$\left(\int_0^t Y_r^i(\omega) d\bar{B}_r^{j-d_Y}\right)(\omega^{\bar{B}}) = \lim_{m\to\infty}\sum_{k=1}^{2^{n_m}} Y^k_{(k-1)/2^{n_m} t}(\omega)\left[B^{j-d_Y}_{k/2^{n_m} t}(\omega^{\bar{B}}) - B^{j-d_Y}_{(k-1)/2^{n_m} t}(\omega^{\bar{B}})\right]$$

for $\omega^B \in D_\omega \subset \Omega^{\hat{B}}$ with $\mathbb{P}^{\hat{B}}[D_\omega]=1$. Here, D_ω as well as the subsequence $(n_m)_m$ will depend on ω.

So for $\omega\in\mathring{\Omega}, \omega^B\in\hat{A}_\omega\cap D_\omega$ we have that

$$\begin{aligned}\left(\int_0^t Y_r^i d\bar{B}_r^{j-d_Y}\right)(\omega,\omega^{\bar{B}}) &= \lim_{n\to\infty}\sum_{k=1}^{2^n} Y^k_{(k-1)/2^n t}(\omega)\left[B^{j-d_Y}_{k/2^n t}(\omega^{\bar{B}}) - B^{j-d_Y}_{(k-1)/2^n t}(\omega^{\bar{B}})\right]\\ &= \lim_{m\to\infty}\sum_{k=1}^{2^{n_m}} Y^k_{(k-1)/2^{n_m} t}(\omega)\left[B^{j-d_Y}_{k/2^{n_m} t}(\omega^{\bar{B}}) - B^{j-d_Y}_{(k-1)/2^{n_m} t}(\omega^{\bar{B}})\right]\\ &= \left(\int_0^t Y_r^i(\omega) d\bar{B}_r^{j-d_Y}\right)(\omega^{\bar{B}}).\end{aligned}$$

Here, the second equality follows from the fact that the sum already converges along n, so it will converge along any subsequence $(n_m)_m$.

Noting that for $\omega\in\mathring{\Omega}$ we have $\mathbb{P}^{\bar{B}}[\hat{A}_\omega\cap D_\omega]=1$ we get that for $\mathbb{P}^Y$-a.e. $\omega\in\Omega$

$$EBM(Y,\bar{B})_\cdot = \boldsymbol{\lambda}_\cdot(EBM(Y)(\omega)), \qquad \mathbb{P}^{\bar{B}} - a.s. \tag{19}$$

Now by classical rough paths results (see e.g. Section 17.5 in [33]), the RDE solution to (17) driven by $EBM(Y,\bar{B})$ corresponds $\hat{\mathbb{P}}$-a.s. to the SDE solution of (6). Combining with (19) we arrive at (8).

The results in Section 17.5 in [33] also yield, that if $\boldsymbol{\eta}$ is (the lift of) a smooth path η, then $\Xi(\eta)$ actually solves (5) as a classical SDE. Together with continuity of the mapping Ξ this yields the first claim of Theorem 1. $\square$

Lemma 8. *Let $\boldsymbol{\eta}, \bar{\boldsymbol{\eta}} \in \mathcal{C}^{0,\alpha}$, with corresponding lifts (as in the proof of Case 2 above) $\boldsymbol{\lambda}, \bar{\boldsymbol{\lambda}}$. Let $|\boldsymbol{\eta}|_\alpha, |\bar{\boldsymbol{\eta}}|_\alpha < R$ and Let $C > 0$. Let $\alpha' < \alpha$ with $\gamma > 1/\alpha'$ and let $p' := 1/\alpha'$. Define*

$$\upsilon(s,t) := ||\boldsymbol{\lambda}||^{p'}_{p'-var;[s,t]} + ||\bar{\boldsymbol{\lambda}}||^{p'}_{p'-var;[s,t]},$$
$$N(x,y,\upsilon) := \left\lceil f(x,y) \sup_{\upsilon(t_i,t_{i+1}) \le f(x,y)^{-1}} \sum_i \upsilon(t_i,t_{i+1}) \right\rceil,$$

where

$$f(x,y) := \frac{y^{p'}}{\psi^{-1}(x)^p},$$
$$\psi(x) := x \exp(x).$$

For every $q \ge 1$ there exists $\beta_0 = \beta_0(R,C,q) > 0, K = K(R,C,q) > 0$ such that for $\beta < \beta_0$

$$|| \exp(-N(\beta,C,\upsilon) \log(1-\beta))||_{L^q} \le K.$$

Proof. Define

$$M(\beta,C,\upsilon) := \sup\{n : \tau_1(f(\beta,C)^{-1}) + \dots + \tau_n(f(\beta,C)^{-1}) \le T\},$$

where

$$\tau_0(a) := 0,$$
$$\tau_{i+1}(a) := \inf\{t \ge 0 : \upsilon(\sigma_i(a), \sigma_i(a)+t) \ge a\},$$

with the convention $\inf \emptyset = \infty$ and $\infty - \infty = \infty$, where

$$\sigma_i(a) := \sum_{j=1}^{i} \tau_j(a).$$

By the argument in Proposition 4.6 in [12] one sees that

$$N(\beta,C,\upsilon) \le 2M(\beta,C,\upsilon) + 1,$$

so we can focus on M.

First

$$\begin{aligned}\mathbb{P}[M(\beta,C,\upsilon) > n] &= \mathbb{P}[\tau_1(f(\beta,C)^{-1}) + \dots + \tau_n(f(\beta,C)^{-1}) \le T] \\ &\le \mathbb{P}[\tau_1(f(\beta,C)^{-1}) \wedge \phi(\beta) + \dots + \tau_n(f(\beta,C)^{-1}) \wedge \phi(\beta) \le T],\end{aligned}$$

where we take ϕ to be the function given in Lemma 9.

Define

$$m_i := m_i(\beta) := \mathbb{E}[\tau_i(f(\beta,C)^{-1}) \wedge \phi(\beta)],$$
$$\bar{S}_n := \frac{1}{n} \sum_{i=1}^{n} \left[\tau_i(f(\beta,C)^{-1}) \wedge \phi(\beta) - m_i\right].$$

then

$$\mathbb{P}[M(\beta, C, v) > n] \leq \mathbb{P}[\bar{S}_n \leq \frac{T}{n} - \frac{1}{n}\sum_{i=1}^{n} m_i].$$

Now Lemma 9 shows, that $0 < m(\beta) \leq \frac{1}{n}\sum_{i=1}^{n} m_i$ for all n, where $m(\beta) := \kappa\phi(\beta)$.
Then, if $n \geq n_0(T, \beta) := 2T/m(\beta)$,

$$\cdots \leq \mathbb{P}[\bar{S}_n \leq -\frac{1}{2}m(\beta)].$$

If $\varepsilon_0 = \frac{1}{2}m(\beta)$, then

$$\begin{aligned} \ldots &\leq \mathbb{P}[\bar{S}_n \leq -\varepsilon_0] \\ &\leq 2\exp(-\frac{2\varepsilon_0^2 n^2}{n\phi(\beta)^2}). \end{aligned}$$

by Hoeffding's inequality (see e.g. Theorem 2 in [41]).
Now

$$\mathbb{E}[\exp(-q\log(1-\beta)M(\beta, C, v))] \leq \sum_{n=1}^{\infty} \mathbb{P}[M(\beta, C, v) \geq \lfloor \frac{\log n}{q(-\log(1-\beta))} \rfloor]$$

Now for

$$n \geq n_1 := \exp((n_0 + 1)(-\log(1-\beta))q),$$

one has by the above,

$$\begin{aligned} \mathbb{P}[M(\beta, C, v) \geq \lfloor \frac{\log n}{q(-\log(1-\beta))} \rfloor] &\leq 2\exp(-\frac{2\varepsilon_0^2}{\phi(\beta)^2}\lfloor \frac{\log n}{q(-\log(1-\beta))} \rfloor) \\ &\leq 2\exp(\frac{2\varepsilon_0^2}{\phi(\beta)^2})\exp(-\frac{2\varepsilon_0^2}{\phi(\beta)^2}\frac{\log n}{q(-\log(1-\beta))}) \\ &= 2\exp(\frac{2\varepsilon_0^2}{\phi(\beta)^2})n^{-\frac{2\varepsilon_0^2}{\phi(\beta)^2 q(-\log(1-\beta))}} \\ &= 2\exp(\frac{1}{2}\kappa)n^{-\frac{\frac{1}{2}\kappa}{q(-\log(1-\beta))}} \end{aligned}$$

Finally, pick $\beta_0 = \beta_0(q)$ such that

$$\frac{\frac{1}{2}\kappa}{q(-\log(1-\beta))} \geq 2,$$

for $\beta \leq \beta_0$. Then for $\beta \leq \beta_0$

$$\mathbb{E}[\exp(-q\log(1-\beta)M(\beta, C, v))] \leq n_1(\beta) + 2\exp(\frac{1}{2}\kappa)\zeta(2),$$

and $n_1(\beta)$ is bounded for $|\boldsymbol{\eta}|_{\alpha-\text{Höl}}, |\bar{\boldsymbol{\eta}}|_{\alpha-\text{Höl}}$ bounded, by Lemma 9. This finishes the proof. □

Lemma 9. *In the setting of the previous Theorem, there exists $\phi = \phi_{\boldsymbol{\eta},\bar{\boldsymbol{\eta}}} : \mathbb{R}_+ \to \mathbb{R}_+$, such that $\phi(x) > 0$ for $x \neq 0$, and $\kappa > 0$ such that for all $n \geq 1$*

$$\frac{1}{n}\sum_{i=1} \mathbb{E}[\tau_i(f(\beta, C)^{-1}) \wedge \phi(\beta)] \geq \kappa\phi(\beta),$$

and such that for $x > 0$, $\phi(x)$ is bounded away from 0 for $|\boldsymbol{\eta}|_{\alpha-H\ddot{o}l}, |\bar{\boldsymbol{\eta}}|_{\alpha-H\ddot{o}l}$ bounded.

Proof. First for every i

$$\begin{aligned}\mathbb{E}[\tau_i(f(\beta, C)^{-1}) \wedge \phi(\beta)] &\geq \mathbb{P}[\tau_i(f(\beta, C)^{-1}) \geq \phi(\beta)] \cdot \phi(\beta)\\ &= \Big(1 - \mathbb{P}[\tau_i(f(\beta, C)^{-1}) < \phi(\beta)]\Big) \cdot \phi(\beta)\\ &\geq \Big(1 - \mathbb{P}[\tau_i(f(\beta, C)^{-1}) \leq \phi(\beta)]\Big) \cdot \phi(\beta)\\ &= \Big(1 - \mathbb{P}[\upsilon(\sigma_i(f(\beta, C)^{-1}), \sigma_i(f(\beta, C)^{-1}) + \phi(\beta)) \geq f(\beta, C)^{-1}]\Big) \cdot \phi(\beta).\end{aligned}$$

Now, by Markov inequality,

$$\begin{aligned}&\mathbb{P}[\upsilon(\sigma_i(f(\beta, C)^{-1}), \sigma_i(f(\beta, C)^{-1}) + \phi(\beta)) \geq f(\beta, C)^{-1}]\\ &\quad\leq f(\beta, C)\mathbb{E}[\upsilon(\sigma_i(f(\beta, C)^{-1}), \sigma_i(f(\beta, C)^{-1}) + \phi(\beta))].\end{aligned}$$

Hence

$$\mathbb{E}[\tau_i(f(\beta, C)^{-1}) \wedge \phi(\beta)] \geq \Big(1 - f(\beta, C)\mathbb{E}[\upsilon(\sigma_i(f(\beta, C)^{-1}), \sigma_i(f(\beta, C)^{-1}) + \phi(\beta))]\Big) \cdot \phi(\beta).$$

Summing up we get

$$\begin{aligned}&\frac{1}{n}\sum_{i=1} \mathbb{E}[\tau_i(f(\beta, C)^{-1}) \wedge \phi(\beta)]\\ &\quad\geq \frac{1}{n}\sum_{i=1}^{n}\Big(1 - f(\beta, C)\mathbb{E}[\upsilon(\sigma_i(f(\beta, C)^{-1}), \sigma_i(f(\beta, C)^{-1}) + \phi(\beta))]\Big) \cdot \phi(\beta) \qquad (20)\\ &\quad= \Big(1 - f(\beta, C)\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[\upsilon(\sigma_i(f(\beta, C)^{-1}), \sigma_i(f(\beta, C)^{-1}) + \phi(\beta))]\Big) \cdot \phi(\beta).\end{aligned}$$

Now

$$\begin{aligned}&\Big(1 - f(\beta, C)\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[\upsilon(\sigma_i(f(\beta, C)^{-1}), \sigma_i(f(\beta, C)^{-1}) + \phi(\beta))]\Big)\\ &\quad\geq \Big(1 - f(\beta, C)\sup_{\tau} \mathbb{E}[\upsilon(\tau, \tau + \phi(\beta))]\Big),\end{aligned}$$

where the supremum is over all stopping times τ.

We define

$$\phi(\beta) := \frac{1}{2}\sup\{a : \sup_{\tau} \mathbb{E}[\upsilon(\tau, \tau + a)] \leq \frac{1}{2f(\beta, C)}\}.$$

We claim that $\phi > 0$ for $\beta > 0$.

Indeed, note that

$$\begin{aligned}\sup_\tau \mathbb{E}[\upsilon(\tau,\tau+a)] &= \sup_\tau \mathbb{E}[||\boldsymbol{\lambda}||^{p'}_{p'-var;[\tau,\tau+a]} + ||\bar{\boldsymbol{\lambda}}||^{p'}_{p'-var;[\tau,\tau+a]}] \\ &\le \sup_\tau \mathbb{E}[a||\boldsymbol{\lambda}||^{p'}_{\alpha'-\text{Höl};[0,T]} + a||\bar{\boldsymbol{\lambda}}||^{p'}_{\alpha'-\text{Höl};[0,T]}] \\ &= \mathbb{E}[||\boldsymbol{\lambda}||^{p'}_{\alpha'-\text{Höl};[0,T]} + ||\bar{\boldsymbol{\lambda}}||^{p'}_{\alpha'-\text{Höl};[0,T]}]a,\end{aligned}$$

which shows that $\phi(\beta) > 0$ for $\beta > 0$. By (16) we have in particular that $\phi(\beta)$ is bounded away from 0 for $|\boldsymbol{\eta}|_{\alpha-\text{Höl}}, |\bar{\boldsymbol{\eta}}|_{\alpha-\text{Höl}}$ bounded.

Then by (20)

$$\begin{aligned}\frac{1}{n}\sum_{i=1}\mathbb{E}[\tau_i(f(\beta,C)^{-1}) \wedge \phi(\beta)] &\ge \Big(1 - f(\beta,C)\sup_\tau \mathbb{E}[\upsilon(\tau,\tau+\phi(\beta))]\Big)\cdot\phi(\beta) \\ &\ge \frac{1}{2}\phi(\beta),\end{aligned}$$

so $\kappa := 1/2$ does the job. □

2.3 The Young case

In Section 2.2 we showed existence, uniqueness and stability of an SDE with rough drift under the assumption that the vector fields in front of the Brownian motion are Lip^1 and the vector fields in front of the rough path are $\text{Lip}^{\gamma+2}$, if the rough path is α-Hölder and $\gamma > 1/\alpha$. The aim of this section is to show that in the case of $\alpha > 1/2$ actually Lip^γ regularity suffices in front of the rough path, which is the regularity one would need when solving a standard RDE. Moreover the vector fields in front of the Brownian motion do not have to be bounded anymore, as one would hope when thinking of the standard assumption for classical SDEs.

The results are most conveniently derived when working in variation spaces. In this section $q \in (1,2)$, so of course $G^{[q]}(\mathbb{R}^{d_Y}) = \mathbb{R}^{d_Y}$ and we are dealing with classical paths in $\mathbb{R}^{d_Y}$. Let $\gamma_q > q$.

Assume

(a3) *a Lipschitz, $b_1,\dots,b_{d_B}$ Lipschitz and $c_1,\dots,c_{d_Y} \in \text{Lip}^{\gamma_q}(\mathbb{R}^{d_S})$*

Let as before $(\bar{\Omega}, \bar{\mathcal{F}}, (\bar{F}_t)_{t\ge 0}, \bar{\mathbb{P}})$ be a filtered probability space carrying a d_B-dimensional Brownian motion $\bar{B}$ and a bounded d_S-dimensional random vector S_0 independent of $\bar{B}$. Let $\eta^n : [0,t] \to \mathbb{R}^{d_Y}$ be smooth paths, such that $\eta^n \to \boldsymbol{\eta}$ in q-variation for some $\boldsymbol{\eta} \in \mathcal{C}^{0,q-var}$ and let S^n be a d_S-dimensional process which is the unique solution to the classical SDE

$$S^n_t = S_0 + \int_0^t a(S^n_r)dr + \int_0^t b(S^n_r)d\bar{B}_r + \int_0^t c(S^n_r)d\eta^n_r,$$

Theorem 10. *Under assumption* **(a3)** *there exists a d_S-dimensional process $S^\infty \in \mathcal{S}^0$ such that*

$$S^n \to S^\infty, \qquad \textit{in } \mathcal{S}^0.$$

In addition, the limit $\Xi(\boldsymbol{\eta}) := S^\infty$ only depends on $\boldsymbol{\eta}$ and not on the approximating sequence.

Moreover, for all $r \geq 1$, $\boldsymbol{\eta} \in \mathcal{C}^{0,q-var}$ *it holds that* $\Xi(\boldsymbol{\eta}) \in \mathcal{S}^r$ *and the corresponding mapping* $\Xi : \mathcal{C}^{0,q-var} \to \mathcal{S}^r$ *is locally uniformly continuous.* [9]

Moreover, $S^\infty \in \mathcal{C}^{0,p-var}$, *for all* $p > 2$ *and satisfies the integral equation*

$$S_t = S_0 + \int_0^t a(S_r)dr + \int_0^t b(S_r)d\bar{B}_r + \int_0^t c(S_r)d\boldsymbol{\eta}_r, \tag{21}$$

where the last integral is a (pathwise) Young integral. It is the only process that satisfies this equation.

We prove this in four steps. First we show uniqueness of a solution to the integral equation, if it exists (Theorem 12). Then we show weak existence of a solution under the assumption $c \in \text{Lip}^1$. (Theorem 13). Using standard methods from strong existence for standard SDEs, we can then deduce from these two results the existence of a unique strong solution (Theorem 14). We finally show stability of the solution with respect to the Young path $\boldsymbol{\eta}$, which yields also the identification as a limit, as in the statement of the theorem. For notational convenience we assume $a \equiv 0$.

Remark 11. *The main reason why we were not able to employ a contraction mapping argument directly, is the fact that the deterministic Young inequality only gives the locally Lipschitz estimate*

$$\begin{aligned}&|| \int H(Y)d\boldsymbol{\eta} - \int H(\bar{Y})d\boldsymbol{\eta}||_{q-var;[0,T]} \\ &\quad \leq C\left(||Y||_{p-var;[0,T]} + ||\bar{Y}||_{p-var;[0,T]}\right) ||Y - \bar{Y}||_{p-var;[0,T]} ||\boldsymbol{\eta}||_{q-var;[0,T]}.\end{aligned}$$

If we apply the expectation operator to both sides of this inequality it becomes evident that a naive Gronwall argument will fail in a stochastic setting.

Theorem 12 (Pathwise Uniqueness)**.** *Let* $T > 0$. *Let* $q \in (1,2)$ *and* $\boldsymbol{\eta} \in \mathcal{C}^{0,q-var}$.

Let $\gamma_q > q$ *and* $c = (c_1, \ldots, c_{d_Y})$ *be a collection of* Lip^{γ_q} *vector fields, let* $b = (b_1, \ldots, b_{d_B})$ *be a collection of Lipschitz (not necessarily bounded) vector fields.*

Let ξ *be a random variable independent of* B.

Then there exists at most one solution S *on* $[0,T]$ *of* (21).

Proof. Let $p \in (2,3)$ such that $1/p + 1/q > 1$.

Let S, $\bar{S}$ be two solutions of (21) on $[0,T]$ (if the solutions are only valid up to a stopping time, the following argument shows that they coincide up to this stopping time).

For $\mathfrak{T} > 0$, $n \geq 0$ define the stopping times

$$\begin{aligned}\tau_0^{\mathfrak{T}} &:= 0, \\ \tau_{n+1}^{\mathfrak{T}} &:= \inf\{t \geq \tau_n^{\mathfrak{T}} : ||S||_{p-var;[\tau_n^{\mathfrak{T}},t]} \geq 1 \text{ or } ||\bar{S}||_{p-var;[\tau_n^{\mathfrak{T}},t]} \geq 1\} \wedge T \wedge (\tau_n^{\mathfrak{T}} + \mathfrak{T}).\end{aligned}$$

Since S, $\bar{S}$ solve (21), their p-variation is a.s. finite (this follows e.g. from Theorem 14.9 in [33]). Hence for every $\mathfrak{T} > 0$, the stopping times exhaust the interval $[0,T]$.

[9] Actually we show that the mapping is continuous into the space of all processes X with norm $||X||_{p,r} = \mathbb{E}[||X||_{p-var;[0,T]^r}]^{1/r} < \infty$ for all $p > 2, r \geq 1$.

Assume that $S_{\tau_n^{\mathfrak{T}}} = \bar{S}_{\tau_n^{\mathfrak{T}}}$. First (here it is essential that $c \in \mathrm{Lip}^{\gamma_q}$)

$$\begin{aligned}
&||\int c(S)d\boldsymbol{\eta} - \int c(\bar{S})d\boldsymbol{\eta}||_{p-var;[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]} \\
&\leq ||\int c(S)d\boldsymbol{\eta} - \int c(\bar{S})d\boldsymbol{\eta}||_{q-var;[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]} \\
&\leq C\left(1+\max\{||S||_{p-var;[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]}, ||\bar{S}||_{p-var;[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]}\}\right)||\boldsymbol{\eta}||_{q-var;[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]}||S-\bar{S}||_{p-var;[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]} \\
&\leq C2\sup_{t\leq T}\Big(||\boldsymbol{\eta}||_{q-var;[t,t+\mathfrak{T}]}\Big)||S-\bar{S}||_{p-var;[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]}.
\end{aligned}$$

Choosing $\mathfrak{T}$ small enough we get

$$\mathbb{E}[||\int_0^\cdot c(S)d\boldsymbol{\eta} - \int_0^\cdot c(\bar{S})d\boldsymbol{\eta}||^r_{p-var;[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]}]^{1/r} \leq 1/3\mathbb{E}[||S-\bar{S}||^r_{p-var;[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]}]^{1/r}.$$

On the other hand, define the martingale $\tilde{B}_t := 1_{[\tau_n^{\mathfrak{T}},\infty)}(t)\left[B_{t\wedge\tau_{n+1}} - B_{\tau_n}\right]$. Using the BDG inequality for p-variation (Theorem 14.12 in [33]), we get

$$\begin{aligned}
\mathbb{E}[||\int b(S)-b(\bar{S})dB||^r_{p-var;[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]}]^{1/r} &= \mathbb{E}[||\int b(S)-b(\bar{S})d\tilde{B}||^r_{p-var;[0,T]}]^{1/r} \\
&\leq C\mathbb{E}[\left(\int_0^T |b(S_r)-b(\bar{S}_r)|^2dr\right)^{r/2}]^{1/r} \\
&= C\mathbb{E}[\left(\int_{\tau_i^n}^{\tau_{i+1}^n} |b(S_r)-b(\bar{S}_r)|^2dr\right)^{r/2}]^{1/r} \\
&\leq C\mathfrak{T}^{1/2}\mathbb{E}[\sup_{\tau_i^n\leq v\leq\tau_{i+1}^n}|S_v-\bar{S}_v|^r]^{1/r} \\
&\leq \mathfrak{T}^{1/2}C\mathbb{E}[||S-\bar{S}||^r_{p-var;[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]}]^{1/r},
\end{aligned}$$

(where we need $\mathfrak{T}\leq 1$ to bound sup with p-var; also we use the fact that $S_{\tau_n^{\mathfrak{T}}} = \bar{S}_{\tau_n^{\mathfrak{T}}}$).
Choosing $\mathfrak{T}$ small enough we get

$$\mathbb{E}[||\int b(S)dB - \int b(\bar{S})dB||^r_{p-var;[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]}]^{1/r} \leq 1/3\mathbb{E}[||S-\bar{S}||^r_{p-var;[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]}]^{1/r}.$$

Hence

$$\begin{aligned}
&\mathbb{E}[||S-\bar{S}||^r_{p-var;[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]}]^{1/r} \\
&\quad= \mathbb{E}[||\int_0^\cdot c(S_r)d\boldsymbol{\eta}_r + \int_0^\cdot b(S_r)dB_r - \int_0^\cdot c(\bar{S}_r)d\boldsymbol{\eta}_r - \int_0^\cdot b(\bar{S}_r)dB_r||^r_{p-var;[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]}]^{1/r} \\
&\quad\leq 2/3\mathbb{E}[||S-\bar{S}||^r_{p-var;[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]}]^{1/r}.
\end{aligned}$$

Hence (note that the right hand side is finite, because of the stopping time) $S=\bar{S}$ on $[\tau_n^{\mathfrak{T}},\tau_{n+1}^{\mathfrak{T}}]$.
This is true for all n, hence $S=\bar{S}$ on $[0,T]$. □

Theorem 13 (Weak Existence)**.** *Let $T>0$. Let $q\in(1,2)$ and $\boldsymbol{\eta}\in\mathcal{C}^{0,q-var}$.*

Let $\gamma_q > q$ and $c=(c_1,\ldots,c_{d_Y})$ be a collection of Lip^{γ_q} vector fields, let $b=(b_1,\ldots,b_{d_B})$ be a collection of Lipschitz Lip^1 (i.e. Lipschitz and bounded) vector fields.

Let μ be a probability measure on $\mathbb{R}^n$.

Then there exists a probability space carrying a Brownian motion $\tilde{B}$ and an adapted process $\tilde{S}$ such that on $[0,T]$

$$\tilde{S}_t = \tilde{S}_0 + \int_0^t b(\tilde{S}_r) d\tilde{B}_r + \int_0^t c(\tilde{S}_r) d\boldsymbol{\eta}_r, \tag{22}$$

and $\tilde{S}_0 \sim \mu$.

Proof. Let c^n be a sequence of collections of Lip^4 vector fields such that $|c^n - c|_{\mathrm{Lip}^{\gamma_q}} \to 0$.

Let a Brownian motion B be given. Let ξ be a random variable, independent of B, with $\xi \sim \mu$. For every n there exists a solution S^n to

$$S^n_t = \xi + \int_0^t b(S^n_r) dB_r + \int_0^t c^n(S^n_r) d\boldsymbol{\eta}_r.$$

Indeed, by Theorem 1 there exists a solution to corresponding SDE with rough drift, which so far only formally satisfies the integral equation. But looking at the proof, the solution is the back-transformation of an SDE with transformed coefficients. Applying Theorem 19 then gives, that S^n actually satisfies the integral equation.

Let $p_- < p \in (2,3)$ such that $1/p + 1/q > 1$. We will show that S^n is tight in $C^{0,p}$. Let $\mathfrak{T} > 0$. First (using BDG, exploiting boundedness of b, when applying Kolmogorov for the stochastic integral)

$$\begin{aligned}
&\mathbb{E}[\big(||S^n||_{p_- -var;[i\mathfrak{T},(i+1)\mathfrak{T}]}\big)^r]^{1/r} \\
&\quad \le \mathbb{E}[\Big(||\int c^n(S^n) d\boldsymbol{\eta}||_{p_- -var;[i\mathfrak{T},(i+1)\mathfrak{T}]}\Big)^r]^{1/r} + \mathbb{E}[\Big(||\int b(S^n_r) dB_r||_{p_- -var;[i\mathfrak{T},(i+1)\mathfrak{T}]}\Big)^r]^{1/r} \\
&\quad \le \mathbb{E}[\big(C(1 + ||S^n||_{p_- -var;[i\mathfrak{T},(i+1)\mathfrak{T}]})||\boldsymbol{\eta}||_{q-var;[i\mathfrak{T},(i+1)\mathfrak{T}]}\big)^r]^{1/r} \\
&\qquad + \mathfrak{T}^{1/p_-}\mathbb{E}[\Big(||\int b(S^n_r) dB_r||_{1/p_- -\text{Höl};[i\mathfrak{T},(i+1)\mathfrak{T}]}\Big)^r]^{1/r} \\
&\quad \le C||\boldsymbol{\eta}||_{q-var;[i\mathfrak{T},(i+1)\mathfrak{T}]} + \mathbb{E}[\big(||S^n||_{p_- -var;[i\mathfrak{T},(i+1)\mathfrak{T}]})\big)^r]^{1/r}||\boldsymbol{\eta}||_{q-var;[i\mathfrak{T},(i+1)\mathfrak{T}]} + C\mathfrak{T}^{1/p_-}.
\end{aligned}$$

Choosing $\mathfrak{T}$ so small such that $\sup_t ||\boldsymbol{\eta}||_{q-var;[t,t+\mathfrak{T}]} < 1/2$, we get

$$\mathbb{E}[\big(||S^n||_{p_- -var;[i\mathfrak{T},(i+1)\mathfrak{T}]}\big)^r]^{1/r} \le 2C||\boldsymbol{\eta}||_{q-var;[i\mathfrak{T},(i+1)\mathfrak{T}]} + 2C\mathfrak{T}^{1/p_-}, \tag{23}$$

independent of n. [10] An application of Lemma 17 then yields

$$\sup_n \mathbb{E}[\big(||S^n||_{p_- -var;[0,T]}\big)^r]^{1/r} < \infty.$$

Hence S^n is tight as random variable in $C^{0,p-var}$. Indeed, let $K_R := \{x \in C^{p_- -var} : |x_0| \le R, ||x||_{p_-} \le R\}$, which is compact in C^{p-var}. Then

$$\mathbb{P}[S^n \notin K_R] \le \mathbb{P}[|\xi| > R] + \mathbb{P}[||S^n||_{p_- -var} > R].$$

[10] Note that, for this reasoning to work, we need to know - a priori - that $\mathbb{E}[\Big(||S^n||^{p_-}_{p_- -var;[i\mathfrak{T},(i+1)\mathfrak{T}]}\Big)^r]^{1/r} < \infty$. But this follows e.g. by using the former inequality for the stopping times $\sigma_n := \inf\{t \ge i\mathfrak{T} : ||S^n||^{p_-}_{p_- -var;[i\mathfrak{T},t]} > n\}$ and taking the limit.

Now the first term goes to zero, trivially uniformly in n, as $R \to \infty$. The second term goes to zero, uniformly in n, as $R \to \infty$, by the Chebychev inequality using (23). This gives tightness.

Hence there exists a probability space with Brownian motions $\tilde{B}^n$ and random processes $\tilde{S}^n$ such that $(S^n, B) \sim (\tilde{S}^n, \tilde{B}^n)$ and (by relabeling a subsequence if necessary)

$$(\tilde{S}^n, \tilde{B}^n) \to (\tilde{S}, \tilde{B}) \qquad \text{a.s. in } C^{0,p-var},$$

for some limit processes $(\tilde{S}, \tilde{B})$. It is easy to see that

$$\tilde{S}^n_t = \tilde{S}^n_0 + \int_0^t b(\tilde{S}^n_r) d\tilde{B}_r + \int_0^t c^n(\tilde{S}^n_r) d\boldsymbol{\eta}_r,$$

and $\tilde{S}^n_0 \sim \mu$. But (e.g. by Theorem 10.47 in [33])

$$\int c^n(\tilde{S}^n) d\boldsymbol{\eta} \to \int c(\tilde{S}) d\boldsymbol{\eta} \qquad \text{a.s. in } \mathcal{C}^{0,q-var}.$$

Moreover by Lemma 5.2 in [38] (or Theorem 7.10 in [44], where one does not need that b is bounded) we have

$$\int b(\tilde{S}^n) d\tilde{B}^n \to \int b(\tilde{S}) d\tilde{B} \qquad \text{in ucp.}$$

Hence

$$\tilde{S}_t = \tilde{S}_0 + \int_0^t b(\tilde{S}_r) d\tilde{B}_r + \int_0^t c(\tilde{S}_r) d\boldsymbol{\eta}_r,$$

and $\tilde{S}_0 \sim \mu$, i.e. we have found a weak solution to (22) as desired. □

Theorem 14 (Strong Existence)**.** *Let* $T > 0$. *Let* $q \in (1,2)$ *and* $\boldsymbol{\eta} \in \mathcal{C}^{0,q-var}$.

Let $\gamma_q > q$ *and* $c = (c_1, \ldots, c_{d_Y})$ *be a collection of* Lip^{γ_q} *vector fields, let* $b = (b_1, \ldots, b_{d_B})$ *be a collection of Lipschitz vector fields.*

Let B *a Brownian motion on some probability space. Let* $S_0 \in L^\infty(\mathcal{F}_0)$.

Then there exists a unique solution S *on* $[0,T]$ *of* (21).

Proof. Assume b is actually a collection of Lip^1 (i.e. Lipschitz and bounded) vector fields. Then we get a strong solution using Theorem 12, Theorem 13, and the standard arguments by Yamada/Watanabe that derives, in a classical SDE setting, strong existence from pathwise uniqueness and weak existence (see e.g. Lemma 18.17 in [42]). [11]

Now let b just be Lipschitz and not necessarily bounded. Let ϕ_n be Lipschitz, with Lipschitz constant 1, such that for $|x| \le n$ we have $\phi_n(x) = x$ and $|\phi_n|_\infty \le n$. Define $b^n(x) := b(\phi^n(x))$. Then by the above, there exists a unique solution to

$$S^n_t = \xi + \int_0^t b^n(S^n_r) dB_r + \int_0^t c(S^n_r) d\boldsymbol{\eta}_r, \tag{24}$$

[11] Let us briefly note another approach to arrive at a strong solution, that might be more instructive. Using Lemma 1.1 in [37] and pathwise uniqueness given by Theorem 12 one can show that the sequence used in Theorem 13 to construct a weak solution actually converges to a strong solution.

Define the stopping time

$$\tau_n := \inf\{t : |S^n_t| \geq n\}.$$

Then, on $[0, \tau_n]$, S^n solves

$$S^n_t = \xi + \int_0^t b(S^n_r)dB_r + \int_0^t c(S^n_r)d\boldsymbol{\eta}_r.$$

Claim: $\tau_n \leq \tau_{n+1}$ and $S^n = S^{n+1}$ on $[0, \tau_n]$. This follows from uniqueness for equation (24). Moreover there exists $C > 0$ such that for all n

$$\mathbb{E}[||S^n||^r_{p-var;[0,T]}]^{1/r} \leq C + C\mathbb{E}[|\xi|^r]^{1/r}. \tag{25}$$

Indeed, using the BDG inequality for p-variation (Theorem 14.12 in [33])

$$\begin{aligned}
&\mathbb{E}[\big(||S^n||_{p-var;[i\mathfrak{T},(i+1)\mathfrak{T}]}\big)^r]^{1/r} \\
&\quad\leq \mathbb{E}[\left(||\int_0^{\cdot} c(S^n)d\boldsymbol{\eta}||_{p-var;[i\mathfrak{T},(i+1)\mathfrak{T}]}\right)^r]^{1/r} + \mathbb{E}[\left(||\int_0^{\cdot} b^n(S^n_r)dB_r||_{p-var;[i\mathfrak{T},(i+1)\mathfrak{T}]}\right)^r]^{1/r} \\
&\quad\leq \mathbb{E}[\big(C(1+||S^n||_{p-var;[i\mathfrak{T},(i+1)\mathfrak{T}]})^p||\boldsymbol{\eta}||_{q-var;[i\mathfrak{T},(i+1)\mathfrak{T}]}\big)^r]^{1/r} \\
&\qquad + C\mathbb{E}[\left(\langle\int_0^{\cdot} b^n(S^n_r)dB_r\rangle_{[i\mathfrak{T},(i+1)\mathfrak{T}]}\right)^{r/2}]^{1/r} \\
&\quad\leq C||\boldsymbol{\eta}||_{q-var;[i\mathfrak{T},(i+1)\mathfrak{T}]} + \mathbb{E}[\big(||S^n||_{p-var;[i\mathfrak{T},(i+1)\mathfrak{T}]})\big)^r]^{1/r}||\boldsymbol{\eta}||_{q-var;[i\mathfrak{T},(i+1)\mathfrak{T}]} \\
&\qquad + C\mathbb{E}[\left(\int_{i\mathfrak{T}}^{(i+1)\mathfrak{T}} |b^n(S^n_r)|^2dr\right)^{r/2}]^{1/r}
\end{aligned}$$

Now

$$\begin{aligned}
\mathbb{E}[\left(\int_{i\mathfrak{T}}^{(i+1)\mathfrak{T}} |b^n(S^n_r)|^2dr\right)^{r/2}]^{1/r} &\leq \mathbb{E}[\left(\int_{i\mathfrak{T}}^{(i+1)\mathfrak{T}} (|b^n(S^n_0)| + |Db^n|_\infty|S^n_{0,r}|)^2dr\right)^{r/2}]^{1/r} \\
&\leq \mathfrak{T}^{1/2}\mathbb{E}[(|b^n(S^n_0)| + |Db^n|_\infty|S^n|_0)^r]^{1/r} \\
&\leq \mathfrak{T}^{1/2}|Db^n|_\infty\left[\mathbb{E}[|S^n_0|^r]^{1/r} + \mathbb{E}[|S^n|^r_0]^{1/r}\right] \\
&\leq \mathfrak{T}^{1/2}|Db^n|_\infty\left[\mathbb{E}[|S^n_0|^r]^{1/r} + \mathbb{E}[|S^n|^r_{p-var}]^{1/r}\right].
\end{aligned}$$

Choosing $\mathfrak{T}$ so small (independent of n and $S^n_{i\mathfrak{T}}$), such that

$$\sup_{t\leq T}||\boldsymbol{\eta}||_{q-var;[t,t+\mathfrak{T}]} + \mathfrak{T}^{1/2}|Db^n|_\infty \leq 1/2,$$

we get

$$\mathbb{E}[\big(||S^n||_{p-var;[i\mathfrak{T},(i+1)\mathfrak{T}]}\big)^r]^{1/r} \leq C + C\mathbb{E}[|S^n_{i\mathfrak{T}}|^r]^{1/r}.$$

An application of Lemma 17 then yields

$$\mathbb{E}[\big(||S^n||_{p-var;[0,T]}\big)^r]^{1/r} \leq C + C\mathbb{E}[|\xi|^r]^{1/r}, \tag{26}$$

where C does not depend on the initial value.

But then $\tau_n \to T$ almost surely. Indeed assume there exists $\mathbb{P}[A] = \varepsilon > 0$ such that on A we have $\tau_n \leq T - \delta$ for all n. But then $\mathbb{P}[|S^n|_\infty \geq n] \geq \varepsilon$, for all n, which contradicts (25).

Hence $S^n_{\cdot \wedge \tau_n}$ converges to a process S that solves (21) on $[0, T]$, and, by an application of Fatou's Lemma, S necessarily satisfies (26).

Now a (unique, by Theorem 12) solution in the case of $\xi \in L^0$, is constructed by solving with initial condition $-n \wedge \xi \vee n$ and letting $n \to \infty$. □

Theorem 15 (Stability)**.** *Let $q \in (1, 2)$. Assume $\boldsymbol{\eta}, \boldsymbol{\eta}^1, \boldsymbol{\eta}^2, \dots \in \mathcal{C}^{0,q-var}$.*

Let the assumptions of Theorem 14 be fulfilled.

If

$$\boldsymbol{\eta}^n \to \boldsymbol{\eta}, \qquad \textit{in } \mathcal{C}^{0,q-var},$$

then for every $r \geq 1$ we have for the corresponding solution to SDE with rough drift,

$$\mathbb{E}[||S^n - S||^r_{p-var;[0,T]}]^{1/r} \to 0.$$

Proof. We want to apply Lemma 17 and Lemma 18.

1. For the first one let $\mathfrak{T} > 0$. Using the BDG inequality for p-variation (Theorem 14.12 in [33])

$$\begin{aligned}
&\mathbb{E}[\big(||S^n||_{p-var;[i\mathfrak{T},(i+1)\mathfrak{T}]}\big)^r]^{1/r} \\
&\quad \leq \mathbb{E}[\left(||\int c(S^n) d\boldsymbol{\eta}^n||_{p-var;[i\mathfrak{T},(i+1)\mathfrak{T}]}\right)^r]^{1/r} + \mathbb{E}[\left(||\int b(S^n_r) dB_r||_{p-var;[i\mathfrak{T},(i+1)\mathfrak{T}]}\right)^r]^{1/r} \\
&\quad \leq \mathbb{E}[\big(C(1 + ||S^n||_{p-var;[i\mathfrak{T},(i+1)\mathfrak{T}]})^p ||\boldsymbol{\eta}^n||_{q-var;[i\mathfrak{T},(i+1)\mathfrak{T}]}\big)^r]^{1/r} \\
&\qquad + C\mathbb{E}[\left(\langle \int b(S^n_r) dB_r \rangle_{[i\mathfrak{T},(i+1)\mathfrak{T}]}\right)^{r/2}]^{1/r} \\
&\quad \leq C||\boldsymbol{\eta}^n||_{q-var;[i\mathfrak{T},(i+1)\mathfrak{T}]} + \mathbb{E}[\big(||S^n||_{p-var;[i\mathfrak{T},(i+1)\mathfrak{T}]})\big)^r]^{1/r} ||\boldsymbol{\eta}^n||_{q-var;[i\mathfrak{T},(i+1)\mathfrak{T}]} \\
&\qquad + C\mathbb{E}[\left(\int_{i\mathfrak{T}}^{(i+1)\mathfrak{T}} |b(S^n_r)|^2 dr\right)^{r/2}]^{1/r}
\end{aligned}$$

Now

$$\begin{aligned}
\mathbb{E}[\left(\int_{i\mathfrak{T}}^{(i+1)\mathfrak{T}} |b(S^n_r)|^2 dr\right)^{r/2}]^{1/r} &\leq \mathbb{E}[\left(\int_{i\mathfrak{T}}^{(i+1)\mathfrak{T}} (|b(S^n_{i\mathfrak{T}})| + |Db|_\infty |S^n_{0,r}|)^2 dr\right)^{r/2}]^{1/r} \\
&\leq \mathfrak{T}^{1/2} \mathbb{E}[(|b(S^n_{i\mathfrak{T}})| + |Db^n|_\infty |S^n|_0)^r]^{1/r} \\
&\leq \mathfrak{T}^{1/2} |Db|_\infty \left[\mathbb{E}[|S^n_{i\mathfrak{T}}|^r]^{1/r} + \mathbb{E}[|S^n|^r_0]^{1/r}\right] \\
&\leq \mathfrak{T}^{1/2} |Db|_\infty \left[\mathbb{E}[|S^n_{i\mathfrak{T}}|^r]^{1/r} + \mathbb{E}[||S^n||^r_{p-var;[i\mathfrak{T},(i+1)\mathfrak{T}]}]^{1/r}\right].
\end{aligned}$$

Choosing $\mathfrak{T}$ so small (independent of n and $S^n_{i\mathfrak{T}}$!), such that

$$\sup_n \sup_{t \leq T} ||\boldsymbol{\eta}^n||_{q-var;[t,t+\mathfrak{T}]} + \mathfrak{T}^{1/2} |Db^n|_\infty \leq 1/2,$$

we get

$$\mathbb{E}[\big(||S^n||_{p-var;[i\mathfrak{T},(i+1)\mathfrak{T}]}\big)^r]^{1/r} \leq C + C\mathbb{E}[|S^n_{i\mathfrak{T}}|^r]^{1/r}.$$

An application of Lemma 17 yields

$$\mathbb{E}[\left(||S^n||_{p-var;[0,T]}\right)^r]^{1/r} \leq C.$$

Hence

$$\mathbb{E}[||S||^r_{p-var;[0,T]}]^{1/r} + \sup_n \mathbb{E}[||S^n||^r_{p-var;[0,T]}]^{1/r} =: R < \infty.$$

2. Define the controls

$$\upsilon^n(s,t) := ||S||^p_{p;[s,t]} + ||S^n||^p_{p;[s,t]}.$$

For $i \geq 1$, $n \geq 0$, $\mathfrak{T} > 0$ define the stopping times (Note, that because of this we have to work in p-var space.)

$$\begin{aligned} \tau_0^n &:= 0, \\ \tau_{i+1}^n &:= \inf\{t \geq \tau_i^n : \upsilon^n(\tau_i^n, t) \geq 1\} \wedge T \wedge (\tau_i^n + \mathfrak{T}). \end{aligned}$$

Then

$$\begin{aligned} &\mathbb{E}[||S - S^n||^r_{p;[\tau_i^n,\tau_{i+1}^n]}]^{1/r} \\ &\quad \leq \mathbb{E}[||\int_0^\cdot c(S_r) - c(S_r^n) d\boldsymbol{\eta}_r||^r_{p;[\tau_i^n,\tau_{i+1}^n]}]^{1/r} + \mathbb{E}[||\int_0^\cdot c(S_r^n) d\left[\boldsymbol{\eta}_r - \boldsymbol{\eta}_r^n\right]||^r_{p;[\tau_i^n,\tau_{i+1}^n]}]^{1/r} \\ &\qquad + \mathbb{E}[||\int_0^\cdot b(S_r) - b(S_r^n) dB_r||^r_{p;[\tau_i^n,\tau_{i+1}^n]}]^{1/r}. \end{aligned}$$

Now, as usual (see e.g. the proof of Theorem 12), the last term is estimated with

$$C\mathfrak{T}^{1/2}\mathbb{E}[||S - S^n||^r_{\infty;[\tau_i^n,\tau_{i+1}^n]}]^{1/r} \leq C\mathfrak{T}^{1/2}\mathbb{E}[|S_{\tau_i^n} - S^n_{\tau_i^n}|^r]^{1/r} + C\mathfrak{T}^{1/2}\mathbb{E}[||S - S^n||^r_{p;[\tau_i^n,\tau_{i+1}^n]}]^{1/r}.$$

The second term is estimated by

$$C\mathbb{E}[||S^n||^r_{p;[\tau_i^n,\tau_{i+1}^n]}]^{1/r}||\boldsymbol{\eta} - \boldsymbol{\eta}^n||_{q;[0,T]} \leq C||\boldsymbol{\eta} - \boldsymbol{\eta}^n||_{q;[0,T]}.$$

The first term is estimated by (here we use the definition of the stopping times)

$$C\mathbb{E}[|S_{\tau_i^n} - S^n_{\tau_i^n}|^r]^{1/r} \sup_{t \leq T} ||\boldsymbol{\eta}||_{q;[t,t+\mathfrak{T}]} + C\mathbb{E}[||S - S^n||^r_{p;[\tau_i^n,\tau_{i+1}^n]}]^{1/r} \sup_{t \leq T} ||\boldsymbol{\eta}||_{q;[t,t+\mathfrak{T}]}.$$

Choosing $\mathfrak{T}$ small enough we arrive at

$$\begin{aligned} &\mathbb{E}[||S - S^n||^r_{p;[\tau_i^n,\tau_{i+1}^n]}]^{1/r} \\ &\quad \leq C\mathbb{E}[|S_{\tau_i^n} - S^n_{\tau_i^n}|^r]^{1/r} + C||\boldsymbol{\eta} - \boldsymbol{\eta}^n||_{q;[0,T]}. \end{aligned}$$

Hence

$$\begin{aligned} \mathbb{E}[||S - S^n||^r_{\infty;[\tau_i^n,\tau_{i+1}^n]}]^{1/r} &\leq (1 + C)\mathbb{E}[|S_{\tau_i^n} - S^n_{\tau_i^n}|^r]^{1/r} + C||\boldsymbol{\eta} - \boldsymbol{\eta}^n||_{q;[0,T]} \\ &= C\mathbb{E}[|S_{\tau_i^n} - S^n_{\tau_i^n}|^r]^{1/r} + C||\boldsymbol{\eta} - \boldsymbol{\eta}^n||_{q;[0,T]}. \end{aligned}$$

An application of Lemma 18 then yields the desired convergence (using, that the first part did hold true for all $r \geq 1$). □

We now present the lemmas that were necessary for above reasoning.

Lemma 16. *Let* $0 = t_0 < t_1 < \cdots < t_n = T$. *Let* $p \geq 1$.
Then for every $X \in \mathcal{C}^{0,p-var}$

$$||X||_{p-var;[0,T]} \leq n^{p-1} \left(||X||_{p-var;[t_0,t_1]} + ||X||_{p-var;[t_1,t_2]} + \cdots + ||X||_{p-var;[t_{n-1},t_n]} \right).$$

Proof. Straightforward calculation. □

Lemma 17. *Let* S *be a stochastic process in* $\mathcal{C}^{0,p-var}$.
Let $r \geq 1$.
i). Let

$$\mathbb{E}[|S_0|^r]^{1/r} = C_0.$$

If there exists $\mathfrak{T} > 0$ *such that for all* i

$$\mathbb{E}[||S||^r_{p-var;[i\mathfrak{T},(i+1)\mathfrak{T}]}]^{1/r} \leq C_1 + C_1 \mathbb{E}[|S_{i\mathfrak{T}}|^r]^{1/r}.$$

Then there exiss $C_2 = C_2(C_1, \mathfrak{T}, T, r, p)$ *such that*

$$\mathbb{E}[||S||^r_{p-var;[0,T]}]^{1/r} \leq C_2,$$
$$\mathbb{E}[||S||^r_{\infty;[0,T]}]^{1/r} \leq C_2.$$

ii). If there exists $\mathfrak{T} > 0$ *such that for all* i

$$\mathbb{E}[||S||^r_{p-var;[i\mathfrak{T},(i+1)\mathfrak{T}]}]^{1/r} \leq C_1,$$

then there exists $C_3 = C_3(C_1, \mathfrak{T}, T, r, p)$ *such that*

$$\mathbb{E}[||S||^r_{p-var;[0,T]}]^{1/r} \leq C_3.$$

Proof. Without loss of generality assume $T = i\mathfrak{T}$, for some $i \geq 1$.
By Lemma 16 we have

$$\begin{aligned}
\mathbb{E}[||S||^r_{p-var;[0,i\mathfrak{T}]}]^{1/r} &\leq 2^{p-1}\mathbb{E}[||S||^r_{p-var;[0,(i-1)\mathfrak{T}]}]^{1/r} + 2^{p-1}\mathbb{E}[||S||^r_{p-var;[(i-1)\mathfrak{T},i\mathfrak{T}]}]^{1/r} \\
&\leq 2^{p-1}\mathbb{E}[||S||^r_{p-var;[0,(i-1)\mathfrak{T}]}]^{1/r} + 2^{p-1}\left[C_1 + C_1\mathbb{E}[|S_{(i-1)\mathfrak{T}}|^r]^{1/r}\right] \\
&\leq 2^{p-1}\mathbb{E}[||S||^r_{p-var;[0,(i-1)\mathfrak{T}]}]^{1/r} \\
&\quad + 2^{p-1}\left[C_1 + C_1\left(\mathbb{E}[|S_0|^r]^{1/r} + \mathbb{E}[||S||^r_{p-var;[0,(i-1)\mathfrak{T}]}]^{1/r}\right)\right].
\end{aligned}$$

Iterating up to time 0 yields the desired result. □

Lemma 18. *Let* S, S^n *be a sequence of stochastic processes in* $\mathcal{C}^{0,p-var}$.
Define the controls

$$v^n(s,t) := ||S||^p_{p;[s,t]} + ||S^n||^p_{p;[s,t]}.$$

For $i,n \geq 0$, $\mathfrak{T} > 0$ define the stopping times

$$\tau_0^n := 0,$$
$$\tau_{i+1}^n := \inf\{t \geq \tau_i^n : \upsilon^n(\tau_i^n, t) \geq 1\} \wedge T \wedge (\tau_i^n + \mathfrak{T}).$$

Assume that for some $\mathfrak{T} > 0$, $r > 1$ one has for all i, n

$$\mathbb{E}[||S - S^n||^r_{p-var;[\tau_i^n,\tau_{i+1}^n]}]^{1/r} \leq C\mathbb{E}[|S_{\tau_i^n} - S^n_{\tau_i^n}|^r]^{1/r} + h(n),$$

where $h(n) \to 0$ as $n \to \infty$.
Assume

$$\mathbb{E}[||S||^r_{p-var;[0,T]}] + \sup_n \mathbb{E}[||S^n||^r_{p-var;[0,T]}] =: R < \infty. \tag{27}$$

Assume

$$\mathbb{E}[|S_0 - S_0^n|^r]^{1/r} \to 0.$$

Then for every $1 \leq r_- < r$ we have

$$\mathbb{E}[||S - S^n||^{r_-}_{p-var;[0,T]}]^{1/r_-} \to_{n\to\infty} 0.$$

Proof. 1. For every $\varepsilon > 0$ there exists $M(\varepsilon)$ such that uniformly for all n

$$\mathbb{P}[\tau_M^n = T] \geq 1 - \varepsilon.$$

Indeed,

$$\mathbb{P}[\tau_M^n < T] \leq \mathbb{P}[\upsilon^n(0,T) \geq M - \lceil T/\mathfrak{T}\rceil] \leq \frac{1}{M - \lceil T/\mathfrak{T}\rceil}\mathbb{E}[\upsilon^n(0,T)],$$

which can be chosen small, uniformly in n, because of (27).
2. We estimate

$$\begin{aligned}
&\mathbb{E}[||S - S^n||^r_{p-var;[0,\tau_m^n]}]^{1/r} \\
&\quad\leq 2^{p-1}\mathbb{E}[||S - S^n||^r_{p-var;[0,\tau_{m-1}^n]}]^{1/r} + 2^{p-1}\mathbb{E}[||S - S^n||^r_{p-var;[\tau_{m-1}^n,\tau_m^n]}]^{1/r} \\
&\quad\leq 2^{p-1}\mathbb{E}[||S - S^n||^r_{p-var;[0,\tau_{m-1}^n]}]^{1/r} + 2^{p-1}\left[C\mathbb{E}[||S - S^n||^r_{\infty;[0,\tau_{m-1}^n]}]^{1/r} + h(n)\right] \\
&\quad\leq 2^{p-1}\mathbb{E}[||S - S^n||^r_{p-var;[0,\tau_{m-1}^n]}]^{1/r} \\
&\qquad + 2^{p-1}\left[C\mathbb{E}[|S - S_0^n|^r]^{1/r} + C\mathbb{E}[||S - S^n||^r_{p-var;[0,\tau_{m-1}^n]}]^{1/r} + h(n)\right].
\end{aligned}$$

Iterating yields

$$\mathbb{E}[||S - S^n||_{p-var;[0,\tau_m^n]}]^r]^{1/r} \leq C(m,p)\left[\mathbb{E}[|S - S_0^n|^r]^{1/r} + h(n)\right].$$

3. We show convergence. Let $1 \leq r_- < r$.

For every $n, m \geq 1$ let $A_{n,m} := \{\tau_m^n = T\}$. We have

$$\begin{aligned}
\mathbb{E}[||S - S^n||_{p-var;[0,T]}^{r-}] &= \mathbb{E}[||S - S^n||_{p-var;[0,T]}^{r-} 1_{A_{n,m}}] + \mathbb{E}[||S - S^n||_{p-var;[0,T]}^{r-} 1_{A_{n,m}^C}] \\
&= \mathbb{E}[||S - S^n||_{p-var;[0,\tau_m^n]}^{r-} 1_{A_{n,m}}] + \mathbb{E}[||S - S^n||_{p-var;[0,T]}^{r-} 1_{A_{n,m}^C}] \\
&\leq \left(C(m,p) \left[\mathbb{E}[|S - S_0^n|^{r-}]^{1/r-} + h(n) \right] \right)^{r-} \\
&\quad + \mathbb{E}[||S - S^n||_{p-var;[0,T]}^{r}]^{1/r} \mathbb{P}[A_n^C]^{1/\theta} \\
&\leq \left(C(m,p) \left[\mathbb{E}[|S - S_0^n|^{r-}]^{1/r-} + h(n) \right] \right)^{r-} + R^{1/r} \mathbb{P}[A_n^C]^{1/\theta},
\end{aligned}$$

where $\theta = \frac{r}{r - r_-}$.

Let $\varepsilon > 0$ be given. Pick m so large, such that the second term is smaller then ε (uniformly in n, by Step 1). Then pick n so large that the first term is smaller then ε. This yields convergence. $\square$

Lemma 19. *Let $q \in (1, 2)$, let $\gamma_q > q$. Let $\boldsymbol{\eta} \in \mathcal{C}^{0,q-var}$. Let ϕ be the Young flow*

$$\phi(t, y) = y + \int_0^t c(\phi(r, y)) d\boldsymbol{\eta}_r.$$

Assume $c = (c_1, \ldots, c_{d_{\boldsymbol{\eta}}})$ is a collection of $\mathrm{Lip}^{\gamma_q + 2}$ vector fields. Then ϕ is a C^3 diffeomorphism.

Let Y be a continuous semimartingale with values in $C^{0,p-var}$, where $1/q + 1/p > 1$. Then

$$\phi(t, Y_t) - \phi(0, Y_0) = \int_0^t c(\phi(r, Y_r)) d\boldsymbol{\eta}_r + \int_0^t \partial_y \phi(r, Y_r) dY_r. + \frac{1}{2} \int_0^t \partial_{yy} \phi(r, Y_r) d\langle Y \rangle_r.$$

Remark 20. *By Proposition 14.9 in [33] every continuous semimartingale has values in $C^{0,p-var}$ for every $p > 2$.*

Proof. For notational convenience we assume $d_Y = 1$.

1. Obviously $\phi(\cdot, Y_\cdot) \in C^{0,p-var}$ so the Young integral is well defined.

2. Let $t_i^n := i 2^{-n} t$.

Now

$$\phi(t, Y_t) - \phi(t, Y_0) = \sum_{i=0}^{2^n - 1} \phi(t_{i+1}^n, Y_{t_i^n}) - \phi(t_{i+1}^n, Y_{t_{i+1}^n}) + \phi(t_{i+1}^n, Y_{t_i^n}) - \phi(t_i^n, Y_{t_i^n})$$

Now classically

$$\begin{aligned}
&\sum_{i=0}^{2^n - 1} \phi(t_{i+1}^n, Y_{t_{i+1}^n}) - \phi(t_{i+1}^n, Y_{t_i^n}) \\
&= \sum_{i=0}^{2^n - 1} \partial_y \phi(t_{i+1}^n, Y_{t_i^n}) \left[Y_{t_{i+1}^n} - Y_{t_i^n} \right] + \frac{1}{2} \partial_{yy} \phi(t_{i+1}^n, Y_{t_{i+1}^n}) \left[Y_{t_{i+1}^n} - Y_{t_i^n} \right]^2 \\
&\quad + \frac{1}{6} \partial_{yyy} \phi(t_{i+1}^n, \xi_{t_{i+1}^n}) \left[Y_{t_{i+1}^n} - Y_{t_i^n} \right]^3 \\
&\to_{n \to \infty} \int_0^t \partial_y \phi(r, Y_r) dY_r + \frac{1}{2} \int_0^t \partial_{yy} \phi(r, Y_r) d\langle Y \rangle_r,
\end{aligned}$$

in probability, say.

On the other hand

$$\sum_{i=0}^{2^n-1} \phi(t_{i+1}^n, Y_{t_i^n}) - \phi(t_i^n, Y_{t_i^n})$$
$$= \sum_{i=0}^{2^n-1} c(\phi(t_i^n, Y_{t_i^n})) \left[\boldsymbol{\eta}_{t_{i+1}^n} - \boldsymbol{\eta}_{t_i^n}\right] + \left(\phi(t_{i+1}^n, Y_{t_i^n}) - \phi(t_i^n, Y_{t_i^n}) - c(\phi(t_i^n, Y_{t_i^n})) \left[\boldsymbol{\eta}_{t_{i+1}^n} - \boldsymbol{\eta}_{t_i^n}\right]\right)$$

Let $2 > q' > q$. Now by Theorem 6.8 in [33] we have

$$\begin{aligned}
&\left|\phi(t_{i+1}^n, Y_{t_i^n}) - \phi(t_i^n, Y_{t_i^n}) - c(\phi(t_i^n, Y_{t_i^n})) \left[\boldsymbol{\eta}_{t_{i+1}^n} - \boldsymbol{\eta}_{t_i^n}\right]\right| \\
&= \left|\int_{t_i^n}^{t_{i+1}^n} c(\phi(r, Y_{t_i^n})) d\boldsymbol{\eta}_r - \phi(t_i^n, Y_{t_i^n}) \left[\boldsymbol{\eta}_{t_{i+1}^n} - \boldsymbol{\eta}_{t_i^n}\right]\right| \\
&\le C |c(\phi(\cdot, Y_{t_i^n}))|_{q'-var;[t_i^n, t_{i+1}^n]} |\boldsymbol{\eta}|_{q'-var;[t_i^n, t_{i+1}^n]} \\
&\le C |\phi(\cdot, Y_{t_i^n})|^2_{q'-var;[t_i^n, t_{i+1}^n]} + C |\boldsymbol{\eta}|^2_{q'-var;[t_i^n, t_{i+1}^n]}.
\end{aligned}$$

Now by Theorem 10.14 in [33], we have

$$\begin{aligned}
\sup_y |\phi(\cdot, y)|^2_{q'-var;[t_i^n, t_{i+1}^n]} &\le C \left(||\boldsymbol{\eta}||^2_{q'-var;[t_i^n, t_{i+1}^n]} + ||\boldsymbol{\eta}||^{2q'}_{q'-var;[t_i^n, t_{i+1}^n]}\right) \\
&\le C \left(||\boldsymbol{\eta}||^{q'}_{q'-var;[t_i^n, t_{i+1}^n]} + ||\boldsymbol{\eta}||^{(q')^2}_{q'-var;[t_i^n, t_{i+1}^n]}\right),
\end{aligned}$$

where the last line holds for n large enough, since $q' < 2$.

Since also $|\boldsymbol{\eta}|^2_{q'-var;[t_i^n, t_{i+1}^n]} \le |\boldsymbol{\eta}|^{q'}_{q-var;[t_i^n, t_{i+1}^n]}$ for n large enough, we get by Wiener's characterization of $C^{0,p-var}$, (Theorem 8.22 (i.2a) in [33])

$$\sum_{i=0}^{2^n-1} \left(\phi(t_{i+1}^n, Y_{t_i^n}) - \phi(t_i^n, Y_{t_i^n}) - c(\phi(t_i^n, Y_{t_i^n})) \left[\boldsymbol{\eta}_{t_{i+1}^n} - \boldsymbol{\eta}_{t_i^n}\right]\right) \to_{n\to\infty} 0.$$

By Exercise 6.9 in [33] we know that

$$\sum_{i=0}^{2^n-1} c(\phi(t_i^n, Y_{t_i^n})) \left[\boldsymbol{\eta}_{t_{i+1}^n} - \boldsymbol{\eta}_{t_i^n}\right] \to_{n\to\infty} \int_0^t c(\phi(r, Y_r)) d\boldsymbol{\eta}_r,$$

Hence

$$\sum_{i=0}^{2^n-1} \phi(t_{i+1}^n, Y_{t_i^n}) - \phi(t_i^n, Y_{t_i^n}) \to_{n\to\infty} \int_0^t c(\phi(r, Y_r)) d\boldsymbol{\eta}_r,$$

which finishes the proof. □

2.4 Applications

We present three applications for stochastic differential equations wit drift. The first one concerns problem of robustness in the theory of nonlinear filtering. With the help of SDEs with rough drift, we solve the case in which the signal and multidimensional observation are correlated.

The second application concerns the stochastic representation of certain rough PDEs, a topic on which we will also elaborate in the section on BSDEs with rough drift.

Lastly, the results on the Young case enable us to solve mixed Brownian-fractional Brownian SDEs, under significantly weaker assumptions then those that seem to be available in the literature.

2.4.1 Robust filtering

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space on which we have defined a two component diffusion process (X, Y) solving a stochastic differential equation driven by a multidimensional Brownian motion. One assumes that the first component X is unobservable and the second component Y is observed. The filtering problem consists in computing the conditional distribution of the unobserved component, called the *signal* process, given the *observation* process Y. Equivalently, one is interested in computing

$$\pi_t(f) = \mathbb{E}[f(X_t, Y_t)|\mathcal{Y}_t],$$

where $\mathcal{Y} = \{\mathcal{Y}_t, t \geq 0\}$ is the observation filtration and f is a suitably chosen test function. An elementary measure theoretic result tells us[12] that there exists a Borel-measurable map $\theta_t^f : C([0,t], \mathbb{R}^{d_Y}) \to \mathbb{R}$, such that

$$\pi_t(\varphi) = \theta_t^f(Y_\cdot) \qquad \mathbb{P}\text{-a.s.}, \tag{28}$$

where d_Y is the dimension of the observation state space and $Y_\cdot$ is the path-valued random variable

$$Y_\cdot : \Omega \to C([0,t], \mathbb{R}^{d_Y}),\ Y_\cdot(\omega) = (Y_s(\omega),\ 0 \leq s \leq t).$$

Of course, θ_t^f is not unique. Any other function $\bar{\theta}_t^f$ such that

$$\mathbb{P} \circ Y_\cdot^{-1}\left(\bar{\theta}_t^f \neq \theta_t^f\right) = 0,$$

where $\mathbb{P} \circ Y_\cdot^{-1}$ is the distribution of $Y_\cdot$ on the path space $C([0,t], \mathbb{R}^{d_Y})$ can replace θ_t^f in (28). It would be desirable to solve this ambiguity by choosing a suitable representative from the class of functions that satisfy (28). A *continuous* version, if it exists, would enjoy the following uniqueness property: if the law of the observation $\mathbb{P} \circ Y_\cdot^{-1}$ positively charges all non-empty open sets in $C([0,t], \mathbb{R}^{d_Y})$ then there exists a unique continuous function θ_t^f that satisfies (28). In this case, we call $\theta_t^f(Y_\cdot)$ the *robust version* of $\pi_t(\varphi)$ and equation (28) is the robust representation formula for the solution of the stochastic filtering problem.

The need for this type of representation arises when the filtering framework is used to model and solve 'real-life' problems. As explained in a substantial number of papers (e.g. [15, 16, 22, 23, 24, 25, 26, 45]) the model chosen for the "real-life" observation process $\bar{Y}$ may not be a perfect one. However, as long as the distribution of $\bar{Y}_\cdot$ is close in a weak sense to that of $Y_\cdot$ (and some integrability assumptions hold), the estimate $\theta_t^f(\bar{Y}_\cdot)$ computed on the actual observation will still be reasonable, as $\mathbb{E}[(f(X_t, Y_t) - \theta_t^f(\bar{Y}_\cdot))^2]$ is well approximated by the idealized error $\mathbb{E}[(f(X_t, Y_t) - \theta_t^f(Y_\cdot))^2]$.

[12] See, for example, Proposition 4.9 page 69 in [6].

Moreover, even when Y and $\bar{Y}$ actually coincide, one is never able to obtain and exploit a continuous stream of data as modeled by the continuous path $Y_{\cdot}(\omega)$. Instead the observation arrives and is processed at discrete moments in time

$$0 = t_0 < t_1 < t_2 < \cdots < t_n = t.$$

However the continuous path $\hat{Y}_{\cdot}(\omega)$ obtained from the discrete observations $(Y_{t_i}(\omega))_{i=1}^n$ by linear interpolation is close to $Y_{\cdot}(\omega)$ (with respect to the supremum norm on $C([0,t], \mathbb{R}^{d_Y})$; hence, by the same argument, $\theta_t^f(\hat{Y}_{\cdot})$ will be a sensible approximation to $\pi_t(\varphi)$.

In the following, we will assume that the pair of processes (X, Y) satisfy the equation

$$dX_t = l_0(X_t, Y_t)dt + \sum_k Z_k(X_t, Y_t)dW_t^k + \sum_j L_j(X_t, Y_t)dB_t^j, \tag{29}$$

$$dY_t = h(X_t, Y_t)dt + dW_t, \tag{30}$$

with X_0 being a bounded random variable and $Y_0 = 0$. In (29) and (30), the process X is the d_X-dimensional signal, Y is the d_Y-dimensional observation, B and W are independent d_B-dimensional, respectively, d_Y-dimensional Brownian motions independent of X_0. Suitable assumptions on the coefficients $l_0, L_1, \ldots, L_{d_B} : \mathbb{R}^{d_X+d_Y} \to \mathbb{R}^{d_X}$, $Z_1, \ldots, Z_{d_Y} : \mathbb{R}^{d_X+d_Y} \to \mathbb{R}^{d_X}$ and $h = (h^1, \ldots, h^{d_Y}) : \mathbb{R}^{d_X+d_Y} \to \mathbb{R}^{d_Y}$ will be introduced later on. This framework covers a wide variety of applications of stochastic filtering (see, for example, [20] and the references therein) and has the added advantage that, within it, $\pi_t(\varphi)$ admits an alternative representation that is crucial for the construction of its robust version. Let us detail this representation first.

Let $u = \{u_t, t > 0\}$ be the process defined by

$$u_t = \exp\left(-\sum_{i=1}^{d_Y} \int_0^t h^i(X_s, Y_s)dW_s^i - \frac{1}{2}\int_0^t (h^i(X_s, Y_s))^2 ds\right). \tag{31}$$

Then, under suitable assumptions[13], u is a martingale which is used to construct the probability measure $\mathbb{P}_0$ equivalent to $\mathbb{P}$ on $\bigcup_{0 \le t < \infty} \mathcal{F}_t$ whose Radon–Nikodym derivative with respect to $\mathbb{P}$ is given by u, viz

$$\left.\frac{\mathrm{d}\mathbb{P}_0}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t} = u_t.$$

Under $\mathbb{P}_0$, Y is a Brownian motion under independent of B. Moreover the equation for the signal process X becomes

$$dX_t = \bar{l}_0(X_t, Y_t)dr + \sum_k Z_k(X_t, Y_t)dY_t^k + \sum_j L_j(X_t, Y_t)dB_t^j. \tag{32}$$

Observe that equation (32) is now written in terms of the pair of Brownian motions (Y, B) and the coefficient $\bar{l}_0$ is given by $\bar{l}_0 = l_0 + \sum_k Z_k h_k$. Moreover, for any measurable, bounded function $f : \mathbb{R}^{d_X+d_Y} \to \mathbb{R}$, we have the following formula called the Kallianpur-Striebel's formula,

$$\pi_t(f) = \frac{p_t(f)}{p_t(1)}, \quad p_t(f) := \mathbb{E}_0[f(X_t, Y_t)v_t|\mathcal{Y}_t] \tag{33}$$

[13] For example, if Novikov's condition is satisfied, that is, if $\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^t \|h^i(X_s, Y_s)\|^2 ds\right)\right] < \infty$ for all $t > 0$, then u is a martingale. In particular it will be satisfied in our setting, in which h is bounded.

where $v = \{v_t, t > 0\}$ is the process defined as $v_t := \exp(I_t)$, $t \geq 0$ and

$$I_t := \sum_{i=1}^{d_Y} \left(\int_0^t h^i(X_r, Y_r) dY_r^i - \frac{1}{2} \int_0^t (h^i(X_r, Y_r))^2 dr \right), \quad t \geq 0. \tag{34}$$

The representation (33) suggests the following three-step methodology to construct a robust representation formula for π_t^f:

Step 1 . We construct the triplet of processes (X^y, Y^y, I^y) [14] corresponding to the pair (y, B) where y is now a *fixed* observation path $y_. = \{y_s, s \in [0,t]\}$ belonging to a suitable class of continuous functions and prove that the random variable $f(X^y, Y^y)exp(I^y)$ is $\mathbb{P}_0$-integrable.

Step 2 . We prove that the function $y. \to g_t^f(y.)$ defined as

$$g_t^f(y.) = \mathbb{E}_0 \left[f(X_t^y, Y_t^y) \exp(I_t^y) \right] \tag{35}$$

is continuous.

Step 3 . We prove that $g_t^f(Y.)$ is a version of $p_t(f)$. Then, following (33), the function, $y. \to \theta_t^f(y.)$ defined as

$$\theta_t^f = \frac{g_t^f}{g_t^1} \tag{36}$$

provides the robust version of of $\pi_t(f)$.

We emphasize that **Step 3** cannot be omitted from the methodology. Indeed one has to prove that $g_t^f(Y.)$ is a version of $p_t(f)$ as this fact is not immediate from the definition of g_t^f.

Step 1 is immediate in the particular case when only the Brownian motion B drives X (i.e. the coefficient $Z = 0$) and X is itself a diffusion, i.e., it satisfies an equation of the form

$$dX_t = l_0(X_t)dr + \sum_j L_j(X_t) dB_t^j, \tag{37}$$

and h does only depend on X. In this case the process (X^y, Y^y) can be taken to be the pair (X, y). Moreover, we can define I^y by the formula

$$I_t^y := \sum_{i=1}^{d_Y} \left(h^i(X_t) y_t^i - \int_0^t y_r^i dh^i(X_r) - \frac{1}{2} \int_0^t (h^i(X_r, Y_r))^2 dr \right), \quad t \geq 0. \tag{38}$$

provided the processes $h^i(X)$ are semi-martingales. In (38), the integral $\int_0^t y_r^i dh^i(X_r)$ is the Itô integral of the non-random process y^i with respect to $h^i(X)$. Note that the formula for I_t^y is obtained by applying integration by parts to the stochastic integral in (34)

$$\int_0^t h^i(X_r) dY_r^i = h^i(X_t) Y_r^i - \int_0^t Y_r^i dh^i(X_r), \tag{39}$$

and replacing the process Y by the fixed path y in (39). This approach has been successfully used to study the robustness property for the filtering problem for the above case in a number of papers ([15, 16, 45]).

[14] As we shall see momentarily, in the uncorrelated case the choice of Y^y will trivially be y. In the correlated case we make it part of the *SDE with rough drift*, for (notational) convenience.

The construction of the process (X^y, Y^y, I^y) is no longer immediate in the case when $Z \neq 0$, i.e. when the signal is driven by both B and W (the correlated noise case). In the case when the observation is one dimensional one can solve this problem by using a method akin with Doss-Sussmann's "pathwise solution" of a stochastic differential equation (see [29],[60]). This approach has been employed by Davis to extend the robustness result to the correlated noise case with scalar observation (see, [22, 23, 24, 26]). In this case one constructs first a diffeomorphism which is a pathwise solution of the equation[15]

$$\phi(t,x) = x + \int_0^t Z(\phi(s,x)) \circ dY_t. \tag{40}$$

The diffeomorphism is used to express the solution X of equation (32) as a composition between the diffeomorphism ϕ and the solution of a stochastic differential equation driven by B only and whose coefficients depend continuously on Y. As a result, we can make sense of X^y. I^y is then defined by a suitable (formal) integration by parts that produces a pathwise interpretation of the stochastic integral appearing in (34) and Y^y is chosen to be y, as before. The robust representation formula is then introduced as per (36).

To our knowledge, the correlated noise and multidimensional observation case has not been studied and it is the subject of the current work. In this case in turns out that we cannot hope to have robustness in the sense advocated by Clark. More precisely, there may not exists a map continuous map $\theta_t^f : C([0,t], \mathbb{R}^{d_Y}) \to \mathbb{R}$, such that the representation (28) holds almost surely. The following is a simple example that illustrates this.

Example 21. *Consider the filtering problem where the signal and the observation process solve the following pair of equations*

$$X_t = X_0 + \int_0^t X_r dY_r + \frac{1}{2}\int_0^t X_r dr$$
$$Y_t = \int_0^t h(X_r)dr + W_t,$$

where Y is 2-dimensional and $\mathbb{P}(X_0 = 0) = \mathbb{P}(X_0 = 1) = \frac{1}{2}$. Then with f, h such that $f(0) = h_1(0) = h_2(0) = 0$ one can explicitly compute

$$\mathbb{E}[f(X_t)|\mathcal{Y}_t] = \frac{f(\exp(Y_t))}{1 + \exp\Big(-\sum_{k=1,2}\int_0^t h^k(\exp(Y_r))dY_r^k + \frac{1}{2}\int_0^t ||h(\exp(Y_r)||^2 dr\Big)}. \tag{41}$$

Following the findings of rough path theory (see, eg, [33, 49, 50, 51]) the expression on the right hand side of (41) *is not continuous in supremum norm (nor in any other metric on path space) because of the stochastic integral. Explicitly, this follows, for example, from Theorem 1.1.1 in [51] by rewriting the exponential term as the solution to a stochastic differential equation driven by Y.*

Nevertheless, we can show that a variation of the robustness representation formula still exists in this case. For this we need to "enhance" the original process Y by adding a second component to it which consists of its iterated integrals (that, knowing the path, is in a one-to-one correspondence with the Levỳ area process). Explicitly we consider the process

[15] Here $d^Y = 1$ and Y is scalar.

$\boldsymbol{Y} = \{\boldsymbol{Y}_t, t \geq 0\}$ defined as

$$\boldsymbol{Y}_t = \left(Y_t, \begin{pmatrix} \int_0^t Y_r^1 \circ dY_r^1 & \cdots & \int_0^t Y_r^1 \circ dY_r^{d^Y} \\ \cdots & \cdots & \cdots \\ \int_0^t Y_r^{d^Y} \circ dY_r^1 & \cdots & \int_0^t Y_r^{d^Y} \circ dY_r^{d^Y} \end{pmatrix} \right), \quad t \geq 0. \tag{42}$$

The stochastic integrals in (42) are Stratonovich integrals. The state space of $\boldsymbol{Y}$ is $G^2(\mathbb{R}^{d_Y}) \cong \mathbb{R}^{d_Y} \oplus so(d_Y)$, where $so(d_Y)$ is the set of anti-symmetric matrices of dimension d_Y.[16] Over this state space we consider not the space of continuous function, but a subspace $\mathcal{C}^{0,\alpha}$ that contains paths $\eta : [0,t] \to G^2(\mathbb{R}^{d_Y})$ that are α-Hölder in the $\mathbb{R}^{d_Y}$-component and 2α-Hölder in the $so(d_Y)$-component, where α is a suitably chosen constant $\alpha < 1/2$. Note that there exists a modification of $\boldsymbol{Y}$ such that $\boldsymbol{Y}(\omega) \in \mathcal{C}^{0,\alpha}$ for all ω (Corollary 13.14 in [33]).

The space $\mathcal{C}^{0,\alpha}$ is endowed with the α-Hölder rough path metric under which $\mathcal{C}^{0,\alpha}$ becomes a complete metric space. The main result of the paper is that there exists a continuous map $\theta_t^f : \mathcal{C}^{0,\alpha} \to \mathbb{R}$, such that

$$\pi_t(\varphi) = \theta_t^f(\boldsymbol{Y}_\cdot) \qquad \mathbb{P}\text{-a.s.} \tag{43}$$

Even though the map is defined on a slightly more abstract space, it nonetheless enjoys the desirable properties described above for the case of a continuous version on $C([0,t],\mathbb{R}^d)$. Since $\mathbb{P} \circ \mathbf{Y}^{-1}$ positively charges all non-empty open sets of $\mathcal{C}^{0,\alpha}$,[17] the continuous version we construct will be unique. Also, it provides a certain model robustness, in the sense that $\mathbb{E}[(\varphi(X_t) - \theta_t^f(\bar{\mathbf{Y}}_\cdot))^2]$ is well approximated by the idealized error $\mathbb{E}[(\varphi(X_t) - \theta_t^f(\mathbf{Y}_\cdot))^2]$, if $\bar{\mathbf{Y}}_\cdot$ is close in distribution to $\mathbf{Y}_\cdot$. The problem of discrete observation is a little more delicate. It is true, that the rough path lift $\hat{\boldsymbol{Y}}$ calculated from the linearly interpolated Brownian motion $\hat{Y}$ will converge to the true rough path $\boldsymbol{Y}$ in probability as the mesh goes to zero (Corollary 13.21 in [33]), which implies that $\theta_t^f(\hat{\boldsymbol{Y}})$ is close in probability to $\theta_t^f(\boldsymbol{Y})$ (we provide local Lipschitz estimates for θ^f). In effect, most sensible approximations will do, as is for example shown in Chapter 13 in [33] (although, contrary to the uncorrelated case, not all interpolations that converge in uniform topology will work, see e.g. Theorem 13.24 ibid). But these are probabilistic statements, that somehow miss the pathwise stability that one wants to provide with θ_t^f. If on the other hand, one is able to observe (at discrete time points), instead of only the process itself, also the second level, i.e. the area, one can construct an interpolating rough path by using geodesic interpolation (see e.g. Chapter 13.3.1 in [33]), which is close to the true (lifted) observation path $\mathbf{Y}$ in the relevant metric *for all realizations* $\mathbf{Y} \in \mathcal{C}^{0,\alpha}$.

In the following we will make use of the Stratonovich version of equation (32), i.e. we will consider that the signal satisfies the equation

$$\begin{aligned} X_t &= X_0 + \int_0^t L_0(X_r, Y_r)dr + \sum_k \int_0^t Z_k(X_r, Y_r) \circ dY_r^k + \sum_j \int_0^t L_j(X_r, Y_r)dB_r^j, \\ Y_t &= \int_0^t h(X_r, Y_r)dr + W_t. \end{aligned} \tag{44}$$

where $L_0^j(x,y) = \bar{l}_0^j(x,y) - \frac{1}{2}\sum_k \sum_i \partial_{x_i} Z_k^j(x,y) Z_k^i(x,y) - \frac{1}{2}\sum_k \partial_{y_k} Z_k^j(x,y)$. We remind that under $\mathbb{P}_0$ the observation Y is a Brownian motion independent of B.

[16] More generally, $G^{[1/\alpha]}(\mathbb{R}^d)$ is the "correct" state space for a geometric α-Hölder rough path; the space of such paths subject to α-Hölder regularity (in rough path sense) yields a complete metric space under α-Hölder rough path metric. Technical details of geometric rough path spaces (as found e.g. in section 9 of [33]) are not required for understanding the results of the present paper.

[17] This fact is a consequence of the support theorem of Brownian motion in Hölder rough path topology [34], see also Chapter 13 in [33].

We will assume that f is a bounded Lipschitz function and we fix $\alpha \in (\frac{1}{3}, \frac{1}{2})$, $\gamma > 1/\alpha$, $t > 0$ and X_0 is a bounded random vector independent of B and Y. We will use one of the following assumptions

(A1) $Z_1, \dots, Z_{d_Y} \in \mathrm{Lip}^{\gamma+2}$, $h^1, \dots, h^{d_Y} \in \mathrm{Lip}^{\gamma+2}$ and $L_0, L_1, \dots, L_{d_B} \in \mathrm{Lip}^1$

(A1') $Z_1, \dots, Z_{d_Y} \in \mathrm{Lip}^{\gamma+3}$, $h^1, \dots, h^{d_Y} \in \mathrm{Lip}^{\gamma+3}$ and $L_0, L_1, \dots, L_{d_B} \in \mathrm{Lip}^1$

(A2) $Z_1, \dots, Z_{d_Y} \in \mathrm{Lip}^{\gamma}$, $h^1, \dots, h^{d_Y} \in \mathrm{Lip}^{\gamma}$ and $L_0 \in \mathrm{Lip}^1$, $L_1, \dots, L_{d_B} \in \mathrm{Lip}^{\gamma}$

Remark 22. *Assumption (A1), (A1') and (A2) lead to the existence of a solution of an SDEs with rough driver (Theorem 1). Under (A1) the solution mapping is locally uniformly continuous, under (A1'), (A2) it is locally Lipschitz.*

Assume either (A1), (A1') or (A2). For $\boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}$ there exists by Theorem 1 a solution $(X^{\boldsymbol{\eta}}, I^{\boldsymbol{\eta}})$ to the following *SDE with rough drift*

$$\begin{aligned} X_t^{\boldsymbol{\eta}} &= X_0 + \int_0^t L_0(X_r^{\boldsymbol{\eta}}, Y_r^{\boldsymbol{\eta}})dr + \int_0^t Z(X_r^{\boldsymbol{\eta}}, Y_r^{\boldsymbol{\eta}})d\boldsymbol{\eta}_r + \sum_j \int_0^t L_j(X_r^{\boldsymbol{\eta}}, Y_r^{\boldsymbol{\eta}})d\bar{B}_r^j, \\ Y_t^{\boldsymbol{\eta}} &= \int_0^t d\boldsymbol{\eta}_r, \\ I_t^{\boldsymbol{\eta}} &= \int_0^t h(X_r^{\boldsymbol{\eta}}, Y_r^{\boldsymbol{\eta}})d\boldsymbol{\eta}_r - \frac{1}{2}\sum_k \int_0^t D_k h^k(X_r^{\boldsymbol{\eta}}, Y_r^{\boldsymbol{\eta}})dr. \end{aligned} \tag{45}$$

Remark 23. *Note that formally (!) when replacing the rough path $\boldsymbol{\eta}$ with the process Y, $X^{\boldsymbol{\eta}}, Y^{\boldsymbol{\eta}}$ yields the solution to the SDE (44) and $\exp(I_t^{\boldsymbol{\eta}})$ yields the (Girsanov) multiplicator in* (33). *This observation is made precise in the statement of Theorem 2.*

We introduce the functions $g^f, g^1, \theta : \mathcal{C}^{0,\alpha} \to \mathbb{R}$ defined as

$$g^f(\boldsymbol{\eta}) := \bar{\mathbb{E}}[f(X_t^{\boldsymbol{\eta}}, Y_t^{\boldsymbol{\eta}}) \exp(I_t^{\boldsymbol{\eta}})], \quad g^1(\boldsymbol{\eta}) := \bar{\mathbb{E}}[\exp(I_t^{\boldsymbol{\eta}})], \quad \theta(\boldsymbol{\eta}) := \frac{g^f(\boldsymbol{\eta})}{g^1(\boldsymbol{\eta})}, \quad \boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}.$$

Theorem 24. *Assume that (A1) holds, then θ is locally uniformly continuous. If (A1') or (A2) hold, then θ is locally Lipschitz.*

Proof. From Theorem 1 we know that for $\boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}$ the SDE with rough drift (45) has a unique solution $(X^{\boldsymbol{\eta}}, Y^{\boldsymbol{\eta}}, I^{\boldsymbol{\eta}})$ belonging to $\mathcal{S}^2$.

Let now $\boldsymbol{\eta}, \boldsymbol{\eta}' \in \mathcal{C}^{0,\alpha}$. Denote $X = X^{\boldsymbol{\eta}}, Y = Y^{\boldsymbol{\eta}}, I = I^{\boldsymbol{\eta}}$ and analogously for $\boldsymbol{\eta}'$.

Then

$$\begin{aligned} |g^f(\boldsymbol{\eta}) - g^f(\boldsymbol{\eta})| &\le \mathbb{E}[|f(X_t, Y_t)\exp(I_t) - f(X'_t, Y"_t)\exp(I'_t)|] \\ &\le \mathbb{E}[|f(X_t, Y_t)||\exp(I_t) - \exp(I'_t)|] + \mathbb{E}[|f(X_t, Y_t) - f(X'_t, Y'_t)|\exp(I'_t)] \\ &\le |f|_\infty \mathbb{E}[|\exp(I_t) - \exp(I'_t)|] + \mathbb{E}[|f(X_t, Y_t) - f(X'_t, Y'_t)|^2]^{1/2}\mathbb{E}[|\exp(I'_t)|^2]^{1/2} \\ &\le |f|_\infty \mathbb{E}[|\exp(I_t) + \exp(I'_t)|^2]^{1/2}\mathbb{E}[|I_t - I'_t|]^{1/2} \\ &\qquad + \mathbb{E}[|f(X_t, Y_t) - f(X'_t, Y'_t)|^2]^{1/2}\mathbb{E}[|\exp(I'_t)|^2]^{1/2} \end{aligned}$$

Hence, using from Theorems 1+2 the continuity statements as well as the boundedness of exponential moments, we see that g^f is locally uniformly continuous under (A1) and it is locally Lipschitz under (A1') or (A2).

The same then holds true for g^1 and moreover $g^1(\boldsymbol{\eta}) > 0$. Hence θ is locally uniformly continuous under (A1) and locally Lipschitz under (A1') or (A2). □

Denote by $\boldsymbol{Y}_\cdot$, as before, the canonical rough path lift of Y to $\mathcal{C}^{0,\alpha}$. We then have

Theorem 25. *Assume either (A1), (A1') or (A2). Then $\theta(\boldsymbol{Y}_\cdot) = \pi_t(f)$, $\mathbb{P} - a.s.$*

Proof. To prove the statement it is enough to show that

$$g^f(\boldsymbol{Y}_\cdot) = p_t(f) \qquad \mathbb{P} - a.s.$$

which is equivalent to

$$g^f(\boldsymbol{Y}_\cdot) = p_t(f) \qquad \mathbb{P}_0 - a.s.$$

For that, it suffices to show that

$$\mathbb{E}_0[p_t(f)\Upsilon(Y_\cdot)] = \mathbb{E}_0[g^f(\boldsymbol{Y}_\cdot)\Upsilon(Y_\cdot)], \tag{46}$$

for an arbitrary continuous bounded function $\Upsilon : C([0,t], \mathbb{R}^{d_Y}) \to \mathbb{R}$.

Let $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ be the auxiliary probability space from before, carrying an d_B-dimensional Brownian motion $\bar{B}$. Let $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}) := (\Omega \times \bar{\Omega}, \mathcal{F} \otimes \bar{\mathcal{F}}, \mathbb{P}_0 \otimes \bar{\mathbb{P}})$. By Y and X_0 we denote also the 'lift' of Y to $\hat{\Omega}$, i.e. $Y(\omega, \bar{\omega}) = Y(\omega)$, $X_0(\omega, \bar{\omega}) = X_0(\omega)$. Then (Y, B) (on Ω under $\mathbb{P}_0$) has the same distribution as $(Y, \bar{B})$ (on $\hat{\Omega}$ under $\hat{\mathbb{P}}$).

Denote by $(\hat{X}, \hat{I})$ the solution on $(\hat{\Omega}, \hat{F}, \hat{\mathbb{P}})$ to the SDE

$$\hat{X}_t = X_0 + \int_0^t L_0(\hat{X}_r, Y_r)dr + \sum_k \int_0^t Z_k(\hat{X}_r, Y_r) \circ dY_r^k + \sum_j \int_0^t L_j(\hat{X}_r, Y_r)d\bar{B}_r^j,$$
$$\hat{I}_t = \sum_k \int_0^t h^k(\hat{X}_r, Y_r) \circ dY_r^k - \frac{1}{2}\sum_k \int_0^t D_k h^k(\hat{X}_r, Y_r)dr.$$

Then

$$(Y, \hat{X}, \hat{I})_{\hat{\mathbb{P}}} \sim \left(Y, X, \sum_k \int_0^\cdot h^k(X_r, Y_r) \circ dY_r^k - \frac{1}{2}\sum_k \int_0^\cdot D_k h^k(X_r, Y_r)dr\right)_{\mathbb{P}_0}.$$

Hence, for the left hand side of (46),

$$\mathbb{E}_0[p_t(f)\Upsilon(Y_\cdot)] = \mathbb{E}_0[f(X_t, Y_t)\exp\left(\sum_k \int_0^t h^k(X_r, Y_r) \circ dY_r^k - \frac{1}{2}\sum_k \int_0^t D_k h^k(X_r, Y_r)dr\right)\Upsilon(Y_\cdot)]$$
$$= \hat{\mathbb{E}}[f(\hat{X}_t, Y_t)\exp\left(\hat{I}_t\right)\Upsilon(Y_\cdot)]$$

On the other hand, from Theorem 2 we know that for $\mathbb{P}_0 - a.e\ \omega$

$$X^{\boldsymbol{Y}_\cdot(\omega)}(\bar{\omega})_t = \hat{X}_t(\omega, \bar{\omega}) \qquad \text{for } \bar{\mathbb{P}} - a.e.\ \bar{\omega},$$
$$Y^{\boldsymbol{Y}_\cdot(\omega)}(\bar{\omega})_t = \hat{Y}_t(\omega, \bar{\omega}) \qquad \text{for } \bar{\mathbb{P}} - a.e.\ \bar{\omega},$$
$$I^{\boldsymbol{Y}_\cdot(\omega)}(\bar{\omega})_t = \hat{I}_t(\omega, \bar{\omega}) \qquad \text{for } \bar{\mathbb{P}} - a.e.\ \bar{\omega}.$$

Hence, for the right hand side of (46) we get (using Fubini for the last equality)

$$\begin{aligned}\mathbb{E}_0[g^f(\boldsymbol{Y}_\cdot)\Upsilon(Y_\cdot)] &= \mathbb{E}_0[\bar{\mathbb{E}}[f(X_t^{\boldsymbol{Y}_\cdot}, Y_t^{\boldsymbol{Y}_\cdot})\exp\left(I_t^{\boldsymbol{Y}_\cdot}\right)]\Upsilon(Y_\cdot)] \\ &= \mathbb{E}_0[\bar{\mathbb{E}}[f(\hat{X}_t, Y_t)\exp\left(\hat{I}_t\right)]\Upsilon(Y_\cdot)] \\ &= \hat{\mathbb{E}}[f(\hat{X}_t, Y_t)\exp\left(\hat{I}_t\right)\Upsilon(Y_\cdot)],\end{aligned}$$

which yields (46). □

2.4.2 Pathwise stochastic control

In this section we consider *controlled* SDEs with rough drift. Similar to classical stochastic optimal control, we show that in this framework, the value function solves a certain PDE, which is now a rough PDE (in the sense of [10]); this is the content of Theorem 29. Without the theory of rough paths, Buckdahn and Ma ([8]) considered *pathwise stochastic control* in a purely Brownian setting. Heuristically, they fix the path of a component of the multidimensional Brownian motion driving the controlled SDEs and only then they optimize. We investigate the connection to our results in Theorem 31.

As in the previous section, let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}_0)$ be a filtered probability space carrying a d_Y dimensional Brownian motion Y and and independent d_B dimensional Brownian motion B. As before, $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathcal{F}}_t, \bar{\mathbb{P}})$ will be an auxiliary probability space, carrying a d_B dimensional Brownian motion $\bar{B}$, which serves as the solution space to the SDEs with rough drift.

We shall need a version of Theorem 1 and Theorem 2 with stochastic and time dependent coefficients. The proofs are identical to the proofs under assumption (a1) above, so we omit them.

Theorem 26. *Let $\alpha \in (0,1]$, $\gamma > 1/\alpha$.*

Let $a(\omega,t,x), b_1(\omega,t,x), \cdots, b_{d_B}(\omega,t,x)$ be adapted and Lip^1 with respect to x, uniformly in t and ω. Let $c_1,\ldots,c_{d_Y} \in \mathrm{Lip}^{\gamma+2}(\mathbb{R}^{d_S})$.

Let $S_0 \in L^\infty(\mathcal{F}_0)$. Let $\eta^n : [0,t] \to \mathbb{R}^{d_Y}$ be smooth paths, such that $\eta^n \to \boldsymbol{\eta}$ in α-Hölder, for some $\boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}$ and let S^n be a d_S-dimensional process which is the unique solution to the classical SDE

$$S_t^n = S_0 + \int_0^t a(r, S_r^n)dr + \int_0^t b(r, S_r^n)d\bar{B}_r + \int_0^t c(S_r^n)d\eta_r^n,$$

There exists a d_S-dimensional process $S^\infty \in \mathcal{S}^0$ such that

$$S^n \to S^\infty, \qquad \text{in } \mathcal{S}^0.$$

In addition, the limit $\Xi(\boldsymbol{\eta}) := S^\infty$ only depends on $\boldsymbol{\eta}$ and not on the approximating sequence.

Moreover, for all $q \geq 1$, $\boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}$ it holds that $\Xi(\boldsymbol{\eta}) \in \mathcal{S}^q$ and the corresponding mapping $\Xi : \mathcal{C}^{0,\alpha} \to \mathcal{S}^q$ is locally uniformly continuous.

Let $\hat{\Omega} = \Omega \times \bar{\Omega}$ be the product space, with product measure $\hat{\mathbb{P}} := \mathbb{P}_0 \otimes \bar{\mathbb{P}}$. Let S be the unique solution on this probability space to the SDE

$$S_t = S_0 + \int_0^t a(r, S_r)dr + \int_0^t b(r, S_r)d\bar{B}_r + \int_0^t c(S_r) \circ dY_r.$$

Denote by $\boldsymbol{Y}$ the rough path lift of Y (i.e. the enhanced Brownian Motion over Y).

Theorem 27. *Under the assumptions of the previous theorem*

- *For every* $R > 0, q \geq 1$

$$\sup_{||\boldsymbol{\eta}||_{\alpha-H\ddot{o}l} < R} \bar{\mathbb{E}}[\exp(q|\Xi(\boldsymbol{\eta})|_{\infty;[0,t]})] < \infty.$$

- *For* $\mathbb{P}_0 - a.e.\ \omega$

$$\bar{\mathbb{P}}[S_s(\omega,\cdot) = \Xi(\boldsymbol{Y}(\omega))_s(\cdot), \quad s \leq t] = 1.$$

Consider for a smooth path η the controlled SDE $X := X^{s,x,\nu}$

$$X_t = x + \int_s^t a(r, X_r, \nu_r) dr + \int_s^t b(r, X_r, \nu_r) d\bar{B}_r + \int_s^t c(X_r) d\eta_r,$$

where conditions an a, b, c will be given later on. Here $\nu \in \bar{\mathcal{U}} = \bar{\mathcal{U}}(\bar{\Omega}, a, b)$, the progressively measurable processes such that

$$\bar{\mathbb{E}}[\int_0^T |b(r,x,\nu_r)|^2 dr + \int_0^T |a(r,x,\nu_r)| dr] < \infty, \qquad x \in \mathbb{R}^n.$$

We define the value function

$$v(s,x) := \inf_{\nu \in \bar{\mathcal{U}}} \bar{\mathbb{E}}[g(X_T^{s,x})].$$

We have the following standard result in stochastic optimal control (remember that η is assumed to be smooth).

Lemma 28. *Let* a *and* b *be continuous, bounded and Lipschitz continuous in* x*, uniformly in* $t \in [0,T], u \in \mathbb{R}^l$*. Let* c *be bounded an Lipschitz. Let* g *be bounded and uniformly continuous. Then* v *is a viscosity solution to*

$$\partial_t v^n(t,x) + H(t,x,v^n(t,x),\partial_x v^n(t,x),\partial_{xx} v^n(t,x)) + \partial_x v^n(t,x)\cdot c(x)\dot{\eta}_t^n = 0, \quad (t,x) \in [0,T) \times \mathbb{R}^n \tag{47}$$

$$v^n(T,x) = g(x).$$

where

$$H(t,x,r,p,A) := \inf_u \{a(t,x,u)p + \frac{1}{2}\operatorname{Tr}[b(t,x,u)b(t,x,u)^T A]\}. \tag{48}$$

Moreover $v \in BUC([0,T] \times \mathbb{R}^n$*, and it is the only solution to* (47) *in this space.*

Proof. By Theorem 4.3.1 and Remark 4.3.4 in [58] we have that v is a (discontinuous) viscosity solution to

$$\partial_t v^n(t,x) + H(t,x,v^n(t,x),\partial_x v^n(t,x),\partial_{xx} v^n(t,x)) + \partial_x v^n(t,x)\cdot c(x)\dot{\eta}_t^n = 0, \quad (t,x) \in [0,T) \times \mathbb{R}^n.$$

Since g is bounded, v is bounded, so the Comparison Theorem 4.4.4 in [58] applies and we get that v is actually continuous and the unique BUC solution to (47) (here we used also Remark 4.3.5 for the terminal condition). ☐

Now we get the corresponding statement, in the case where $\boldsymbol{\eta}$ is a rough path.

Theorem 29. *Let $\alpha \in (0,1]$, $\gamma > 1/\alpha$. Let $\boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}$. Let $c \in \mathrm{Lip}^{\gamma+2}$. Let a and b be Lip^1 with respect to x, uniformly in $t \in [0,T], u \in \mathbb{R}^l$.*

Let

$$v(s,x) = \inf_{\nu \in \bar{\mathcal{U}}} \bar{\mathbb{E}}[g(X_T^{s,x,\nu})]$$

where $X = X^{s,x,\nu}$ is the solution to the following controlled SDE with rough drift

$$X_t = x + \int_s^t a(r, X_r, \nu_r) dr + \int_s^t b(r, X_r, \nu_r) d\bar{B}_r + \int_s^t c(X_r) d\boldsymbol{\eta}_r.$$

Then v is the unique solution to the rough partial differential equation [18]

$$[\partial_t v(t,x) + H(t,x,v(t,x), Dv(t,x), D^2 v(t,x))]dt + Dv(t,x) \cdot c(x) d\boldsymbol{\eta}_t = 0, \tag{49}$$
$$v(T,x) = g(x). \tag{50}$$

Proof. By Theorem 1 and Example 5 in [10] there exists a unique solution v^0 to the rough HJB equation (49).

For the sake of unified notation, write $\boldsymbol{\eta}$ as η^0. Let now η^k, $k = 1, 2, \ldots$ be a sequence of smooth path, such that $\eta^k \to \eta^0$ in p-variation.

By Theorem 1 in [10] we have that $v^0(t,x) = \hat{v}^0(t, \psi^0(t,x))$ where ϕ^0 is the flow (14) corresponding to the rough path η^0, ψ^0 is its x-inverse and $\hat{v}^0$ is the unique BUC solution to the PDE

$$\partial_t \hat{v}^0(t,x) + \hat{H}^0(t,x,\hat{v}^0(t,x), D\hat{v}^0(t,x), D^2\hat{v}^0(t,x)) = 0,$$
$$\hat{v}^0(T,x) = g(\psi^0(T,x)).$$

Here we define for $k \geq 0$

$$\begin{aligned}
&\hat{H}^k(t,x,r,p,A) \\
&:= \inf_u \Big\{ a(t, \phi^k(t,x), u) \cdot \langle p, \partial_x \psi^k(t, \phi^k(t,x)) \rangle \\
&\quad \frac{1}{2} + \mathrm{Tr}[b(t, \phi^k(t,x), u) b(t, \phi^k(t,x), u)^T \\
&\qquad \times \Big(\langle A, \partial_x \psi^k(t, \phi^k(t,x)) \otimes \partial_x \psi(t, \phi^k(t,x)) \rangle + \langle p, \partial_{xx} \psi^k(t, \phi^k(t,x)) \rangle \Big)] \Big\} \\
&= \inf_u \Big\{ \tilde{b}(t,x,u) p + \frac{1}{2} \mathrm{Tr}[\tilde{\sigma}(t,x,u) \tilde{\sigma}(t,x,u)^T A] \Big\},
\end{aligned}$$

where $\tilde{b}, \tilde{\sigma}$ were defined in Lemma 3.

By Lemma 28 we have that

$$\hat{v}^0(s,x) = \inf_{\nu \in \bar{\mathcal{U}}} \bar{\mathbb{E}}[g(\psi^0(T, \hat{X}_T^{s,x,\nu}))]$$

where $\hat{X}^{s,x,\nu}$ is the controlled diffusion

$$\hat{X}_t^{s,x,\nu} = x + \int_s^t \tilde{a}(r, \hat{X}^{s,x,\nu}, \nu_r) dr + \int_s^t \tilde{b}(r, \hat{X}^{s,x,\nu}, \nu_r) dW_r. \tag{51}$$

[18] This type of rough partial differential equation has been studied in [10].

Now by Lemma 3 we know that $\tilde{X}_t^{s,x,\nu} := \phi(t, X_t^{s,x,\nu})$ satsifies the SDE (51), with starting point $\phi(s,x)$. Hence we can conclude

$$\begin{aligned} v^0(s,x) &= \hat{v}^0(s,\phi(s,x)) = \inf_{\nu\in\bar{\mathcal{U}}} \bar{\mathbb{E}}[g(\psi^0(T,\hat{X}_T^{s,\phi(s,x),\nu}))] \\ &= \inf_{\nu\in\bar{\mathcal{U}}} \bar{\mathbb{E}}[g(\psi^0(T,\tilde{X}_T^{s,x,\nu}))] = \inf_{\nu\in\bar{\mathcal{U}}} \bar{\mathbb{E}}[g(X_T^{s,x,\nu})]. \end{aligned}$$

□

We now investigate the connection to the results in [8]. Denote now $X^{rp} := X^{rp,s,x,\nu,\boldsymbol{\eta}}$, $\nu \in \bar{\mathcal{U}}$, as the solution to the controlled SDE with rough drift

$$X_t^{rp} = x + \int_s^t a(r, X_r^{rp}, \nu_r)dr + \int_s^t b(r, X_r^{rp}, \nu_r)d\bar{B}_r + \int_s^t c(X_r^{rp})d\boldsymbol{\eta}_r.$$

Let

$$L(s,x,\boldsymbol{\eta}) := \inf_{\nu} \bar{\mathbb{E}}[g(X_T^{rp,s,x,\nu,\boldsymbol{\eta}})].$$

Lemma 30. *Let $\alpha \in (0,1]$, $\gamma > 1/\alpha$.*

Let a and b be in Lip^1 with respect to x, uniformly in $t \in [0,T], u \in \mathbb{R}^l$. Let $c \in \mathrm{Lip}^{\gamma+2}$. Let $g \in \mathrm{Lip}^1(\mathbb{R}^n)$.

For fixed s, x and $R > 0$ the mapping

$$\begin{aligned} \mathcal{C}^{0,\alpha} &\to \mathbb{R} \\ \eta &\mapsto L(s,x,\eta) \end{aligned}$$

is uniformly continuous on $\{||\boldsymbol{\eta}||_{\mathcal{C}^{0,\alpha}} < R\}$.

Proof. By Theorem 26 the mapping that takes a rough path to the corresponding solution of SDE with rough drift is uniformly continuous on bounded sets. Inspection of the proof, using the fact that a and b are in Lip^1 with respect to x, uniformly in $t \in [0,T]$, $u \in \mathbb{R}^l$, yields that the solution mapping is locally uniform continuous, uniform over all controls ν.

The trivial inequality

$$|\inf_{\nu} \bar{\mathbb{E}}[g(X_T^{rp,s,x,\nu,\boldsymbol{\eta}^1})] - \inf_{\nu} \bar{\mathbb{E}}[g(X_T^{rp,s,x,\nu,\boldsymbol{\eta}^2})]| \leq \sup_{\nu} |\bar{\mathbb{E}}[g(X_T^{rp,s,x,\nu,\boldsymbol{\eta}^1})] - \bar{\mathbb{E}}[g(X_T^{rp,s,x,\nu,\boldsymbol{\eta}^2})]|$$

then yields the desired continuity. □

Let $\hat{\Omega} = \Omega \times \bar{\Omega}$ be the product space, with product measure $\hat{\mathbb{P}} := \mathbb{P}_0 \otimes \bar{\mathbb{P}}$. Consider, following [8], the controlled SDE $X := X^{s,x,\nu}$ on $\hat{\Omega}$ where ν is $\mathcal{F}_t^B$ adapted

$$X_t = x + \int_s^t a(r, X_r, \nu_r)dr + \int_s^t b(r, X_r, \nu_r)d\bar{B}_r + \int_s^t c(X_r) \circ dY_r.$$

where conditions an a, b, c will be given later on.

Here $\nu \in \hat{\mathcal{U}} = \hat{\mathcal{U}}(\hat{\Omega}, a, b)$, the progressively measurable processes such that

$$\bar{\mathbb{E}}[\int_0^T |b(r,x,\nu_r)|^2 dr + \int_0^T |a(r,x,\nu_r)|dr] < \infty, \qquad x \in \mathbb{R}^n.$$

Denote

$$v(s,x) := \operatorname*{ess\,inf}_{\nu \in \hat{\mathcal{U}}} \hat{\mathbb{E}}[g(X_T^{s,x,\nu})|\mathcal{F}^Y].$$

Let $\mathbf{Y}$ be the rough path lift of the Brownian motion Y.

Theorem 31. *Let σ and b be in Lip^1 with respect to x, uniformly in $t \in [0,T], u \in \mathbb{R}^l$. Let $g \in \mathrm{Lip}^1(\mathbb{R}^n)$.*

For $\mathbb{P}_0$-a.e. ω^Y we have

$$L(s,x,\mathbf{Y}(\omega^Y)) = v(s,x)(\omega^Y).$$

Remark 32. *Let us shortly remark, that in [8], a and b are assumed to be uniformly continuous (also in the control state). The assumption on c is essentially the same, which comes as no surprise, as they also use a kind of Doss-Sussmann transform in order to deal with the problem. Our approach leads to the same solution, as Theorem 31 shows, but avoids many of the technical problems in [8].*

Proof. "$\leq$":

We have seen already in Theorem 25, that for every ν

$$\bar{\mathbb{E}}[g(X_T^{rp,s,x,\nu,\mathbf{Y}(\omega^Y)})] = \hat{\mathbb{E}}[g(X_T^{s,x,\nu})|Y = \omega^Y], \qquad \text{for } \mathbb{P}_0 - \text{a.e. } \omega^Y. \tag{52}$$

Now by definition

$$L(s,x,\mathbf{Y}(\omega^Y)) \leq \bar{\mathbb{E}}[g(X_T^{rp,s,x,\nu,\mathbf{Y}(\omega^Y)})], \qquad \forall \nu, \forall \omega^Y.$$

Hence for every ν

$$L(s,x,\mathbf{Y}(\omega^Y)) \leq \hat{\mathbb{E}}[g(X_T^{s,x,\nu})|Y = \omega^Y], \qquad \text{for } \mathbb{P}^Y - \text{a.e. } \omega^Y.$$

By using Lemma 30 we see that $\omega^Y \mapsto L(s,x,\mathbf{Y}(\omega^Y))$ is measurable.

Hence by definition of the essential infimum, we have

$$L(s,x,\mathbf{Y}(\omega^Y)) \leq v(s,x)(\omega^Y), \qquad \text{for } \mathbb{P}^Y - \text{a.e. } \omega^Y.$$

"$\geq$":

Using (52), it is enough to show that for every $\delta > 0$ there exists $\Omega_1, \Omega_2, \dots \in \mathcal{F}$ and $\nu_1, \nu_2, \cdots$ such that $\cup_{i=1}^{\infty} \Omega_i = \Omega$ and

$$\bar{\mathbb{E}}[g(X_T^{s,x,\nu_i,\mathbf{Y}}] \leq L(s,x,\mathbf{Y}) + \delta, \qquad \text{on } \Omega_i.$$

For all $\boldsymbol{\eta} \in \mathcal{C}^{0,\alpha}$ let $\nu_{\boldsymbol{\eta}}$ be such that

$$\bar{\mathbb{E}}[g(X_T^{s,x,\nu_{\boldsymbol{\eta}},\boldsymbol{\eta}})] \leq L(s,x,\boldsymbol{\eta}) + \delta/2.$$

Define

$$A_{\boldsymbol{\eta}} := \{\boldsymbol{\xi} \in \mathcal{C}^{0,\alpha} : \bar{\mathbb{E}}[g(X_T^{s,x,\nu_{\boldsymbol{\eta}},\boldsymbol{\xi}})] \leq L(s,x,\boldsymbol{\xi}) + \delta\}.$$

Since $\boldsymbol{\eta} \mapsto \bar{\mathbb{E}}[g(X_T^{s,x,\nu,\boldsymbol{\eta}})]$ and $\boldsymbol{\eta} \mapsto L(s,x,\boldsymbol{\eta})$ are continuous, the $A_{\boldsymbol{\eta}}$ give an open cover of $\mathcal{C}^{0,\alpha}$. Since $\mathcal{C}^{0,\alpha}$ is separable, there exists $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots$ such that

$$\cup_{i=1}^{\infty} A_{\boldsymbol{\eta}_1} = \mathcal{C}^{0,\alpha}.$$

Then $\Omega_i := \mathbf{Y}^{-1}(A_i)$ and $\nu_i := \nu_{\boldsymbol{\eta}_i}$ yield the desired properties. □

2.4.3 Mixed stochastic differential equations

Let $(\bar{\Omega}, \bar{\mathcal{F}}, (\bar{F}_t)_{t\geq 0}, \bar{\mathbb{P}})$ be a filtered probability space carrying a d_B dimensional Brownian motion B. Let $(\Omega, \mathcal{F}, (F_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space carrying a d_Y dimensional fractional Brownian motion Y with Hurst parameter $H > 1/2$. Denote the product space $\hat{\Omega} := \Omega \times \bar{\Omega}$, with corresponding filtration and the product measure $\hat{\mathbb{P}} := \mathbb{P} \otimes \bar{\mathbb{P}}$.

On this space we consider the mixed Brownian-fractional Brownian motion SDE

$$X_t = X_0 + \int_0^t a(X_r)dr + \int_0^t b(X_r)d\bar{B}_r + \int_0^t c(X_r)dY_r. \tag{53}$$

Such equations have for example been considered by Zaehle [63], although only for one-dimensional Brownian motion $\bar{B}$ and essentially Lip^2 vector fields. We also mention the monograph [53], which considers applications to financial mathematics for equations of this type. Let us note, that owing to $H > 1/2$, a joint rough path lift of $\bar{B}, Y$ can easily be constructed (even without the results of Theorem 1, assumption (a2)). Hence (53) could be solved directly as RDE, if b is essentially Lip^2. The main appeal of the following theorem is then, that we get away with a lot less regularity.

Theorem 33. *Let $\gamma > 1/H$. Assume a Lipschitz, $b_1, \dots, b_{d_B}$ Lipschitz and $c_1, \dots, c_{d_Y} \in \mathrm{Lip}^\gamma(\mathbb{R}^{d_S})$. Let $X_0 \in L^\infty(\hat{\mathcal{F}}_0)$.*

Then there exists a unique solution to the mixed SDE (53)*.*

Proof. Uniqueness follows from Theorem 12, we show existence.

On $\hat{\Omega}$ denote the classical solution to the SDE

$$X_t^n = X_0 + \int_0^t a(X_r^n)dr + \int_0^t b(X_r^n)d\bar{B}_r + \int_0^t c(X_r^n)dY_r^n, \tag{54}$$

where Y^n denotes the piecewise linear interpolation of Y.

Now for every n there exists Ω_n, $\mathbb{P}[\Omega_n = 1]$ such that for $\omega \in \Omega_n$, $X^n(\omega, \cdot)$ solves (54) on $\bar{\Omega}$. Let $\Omega' := \cap_n \Omega_n$.

Let $p > 2$ such that $1/p + H > 1$. Define the measurable set

$$A := \{\hat{\omega} \in \hat{\Omega} : X^n(\hat{\omega}) \text{ converges in } \mathcal{C}^{0,p-var}\}.$$

Now by stability of the integrals we have that $X := \lim_n X^n$ solves (53) on A. We show $\hat{\mathbb{P}}[A] = 1$.

For every $\omega \in \Omega'$, X^n solves the SDE with rough drift $Y^n(\omega)$ on $\bar{\Omega}$ for every n. Hence there exists M_ω such that $X^n(\omega, \bar{\omega})$ converges for $\bar{\omega} \in M_\omega$, and $\bar{\mathbb{P}}[M_\omega] = 1$.

Then

$$\{(\omega, \bar{\omega}) : \omega \in \Omega', \bar{\omega} \in M_\omega\} \subset A,$$

an application of Fubini's theorem yields $\hat{\mathbb{P}}[A] = 1$, as desired. □

3 Backward stochastic differential equations with rough driver

We recall that *backward stochastic differential equations* (BSDEs) are stochastic equations of the type

$$Y_t = \xi + \int_t^T f(r, Y_r, Z_r)\, dr - \int_t^T Z_r dW_r. \tag{55}$$

Here, W is an m-dimensional Brownian motion on some filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, \mathbb{P}\right)$. The *terminal data* ξ is assumed to be $\mathcal{F}_T$-measurable, the *driver* $f : \Omega \times [0,T] \times \mathbb{R}\times\mathbb{R}^m \to \mathbb{R}$ is a predictable random field; a solution to this equation is a $(1+m)$-dimensional adapted solution process of the form $(Y_t, Z_t)_{0\le t\le T}$; subject to some integrability properties depending on the framework imposed by the type of assumptions on f. Equation (55) can also be written in differential form

$$-dY_t = f(t, Y_t, Z_t)\, dt - Z_t dW_t.$$

The aim of this section, partially motivated from the recent progress on partial differential equations driven by rough path [9, 10, 36, 27, 61], is to consider

$$-dY_t = f(t, Y_t, Z_t)\, dt + H(Y_t) d\eta_t - Z_t dW_t,$$

where η is (at first) a smooth d-dimensional driving signal - accordingly $H = (H_1, \dots, H_d)$ - followed by a discussion in which we establish *rough path stability* of the solution process (Y, Z) as a function of η. Note that we do *not* establish any sort of rough path stability in W. Indeed when $f \equiv 0$ in (55), BSDE theory reduces to martingale representation, an intrinsically stochastic result which does not seem amenable to a rough pathwise approach. [19] We are able to carry out our analysis in a framework in which the ω-dependence of the terms driven by η factorizes through an Itô process. That is, we consider, for fixed $(t_0, x_0) \in [0,T] \times \mathbb{R}^n$,

$$\begin{aligned} dX_t &= b(\omega;t)\, dt + \sigma(\omega;t)\, dW_t, \quad t_0 \le t \le T; \quad X_{t_0} = x_0 \in \mathbb{R}^n, \\ -dY_t &= f(\omega; t, Y_t, Z_t)\, dt + H(X_t, Y_t)\, d\eta - Z_t dW, \quad t_0 \le t \le T; \quad Y_T = \xi \in L^\infty(\mathcal{F}_T). \end{aligned}$$

Our main-result is, under suitable conditions on f and $H = (H_1, \dots, H_d)$, that any sequence (η^n) which is Cauchy in rough path metric gives rise to a solution (Y, Z) of the *BSDE with rough driver*

$$-dY_t = f(\omega; t, Y_t, Z_t)\, dt + H(X_t, Y_t)\, d\eta - Z_t dW_t, \tag{56}$$

where η denotes the (rough path) limit of (η^n) and where indeed (Y, Z) depends only on η and not on the particular approximating sequence. An interesting feature of this result, which somehow encodes the particular structure of the above equation, is that one does *not* need to construct or understand the iterated integrals of η and W; but only those of η which is tantamount to speak of the rough path η. This is in strict contrast to the usual theory of rough differential equations in which both $d\eta$ and dW figure as driving differentials, e.g. in equations of the form $dy = V_1(y)d\eta + V_2(y)dW$.

If we specialize to a fully Markovian setting, say $\xi = g(X_T)$, $\sigma(\omega;t) = \sigma(t, X_t(\omega))$, $b(\omega;t) = b(t, X_t(\omega))$, $f(\omega; t, y, z) = f(t, X_t(\omega), y, z)$, $H = H(X_t, Y_t)$, we find that the solution to

[19] See however the recent work of Liang et al. [47] in which martingale representation is replaced by an abstract transformation.

(56), evaluated at $t = t_0$, yields a solution to the (terminal value problem of the) *rough partial differential equation*

$$-du = (\mathcal{L}u)\,dt + f\,(t, x, u, Du\;\sigma(t,x))\,dt + H\,(x, u)\,d\eta, \;\; u_T\,(x) = g\,(x)\,,$$

where $\mathcal{L}$ denotes the generator of X. If one is interested in the Cauchy problem, $\tilde{u}\,(t, x) = u\,(T - t, x)$ satisfies,

$$d\tilde{u} = (\mathcal{L}\tilde{u})\,dt + f\,(x, \tilde{u}, D\tilde{u}\;\sigma(t,x))\,dt + H\,(x, \tilde{u})\,d\tilde{\eta}, \;\; \tilde{u}_0\,(x) = g\,(x)\,, \tag{57}$$

where $\tilde{\eta} = \eta\,(T - \cdot)$.

To the best of our knowledge, (56) is the first attempt to introduce rough path methods [49, 51, 50, 33] in the field of backward stochastic differential equations [56, 30, 43]. Of course, there are many hints in the literature towards the possibility of doing so: we mention in particular the Pardoux-Peng [55] theory of *backward doubly stochastic differential equations* (BDSDEs) which amounts to replacing $d\eta$ in (56) by another set of Brownian differentials, say dB, independent of W. This theory was then employed by Buckdahn and Ma [7] to construct (stochastic viscosity) solutions to (57) with $d\eta$ replaced by a Brownian differential and the assumption that the vector fields $H_1\,(x, \cdot)\,, \ldots, H_d\,(x, \cdot)$ commute.

In Section 3.2 we state and prove our main result concerning the existence and uniqueness of BSDEs with rough drivers. Section 3.3 specializes the setting to a purely Markovian one. In this context BSDEs with rough drivers are connected to rough partial differential equations, which we analyze in their own right. In Section 3.4 we establish the connection to BDSDEs.

3.1 Notation

We fix once and for all a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_t, \mathbb{P})$, which carries an m-dimensional Brownian motion W. Let $\mathcal{F}_t$ be the usual filtration of W. Denote by $H^2_{[0,T]}(\mathbb{R}^m)$ the space of predictable processes X in $\mathbb{R}^m$ such that $||X||^2 := \mathbb{E}[\int_0^T |X|_r^2 dr] < \infty$. Denote by $H^\infty_{[0,T]}(\mathbb{R})$ the space of predictable processes that are almost surely bounded. We will say a sequence converges in H^∞ if it converges uniformly on $[0, T]$, $\mathbb{P}$-a.s. For a random variable ξ we denote by $||\xi||_\infty$ its essential supremum, for a process Y we denote by $||Y||_\infty$ the essential supremum of $\sup_{0 \le t \le T} |Y_t|$.

$BC([0, T] \times \mathbb{R}^n)$ (resp. $BC(\mathbb{R}^n)$) denotes the space of bounded continuous functions on $[0, T] \times \mathbb{R}^n$ (resp. $\mathbb{R}^n$) with the topology of uniform convergence on compacta. Similarly $BUC([0, T] \times \mathbb{R}^n)$ (resp. $BUC(\mathbb{R}^n)$) denotes the space of bounded uniformly continuous functions on $[0, T] \times \mathbb{R}^n$ (resp. $\mathbb{R}^n$) with the topology of uniform convergence on compacta.

3.2 Main results

For a smooth path η in $\mathbb{R}^d$ and $\xi \in L^\infty(\mathcal{F}_T)$ we consider the BSDE

$$Y_t = \xi + \int_t^T f(r, Y_r, Z_r)dr + \int_t^T H(X_r, Y_r)d\eta(r) - \int_t^T Z_r dW_r, \quad t \le T, \tag{58}$$

where the $\mathbb{R}^n$-valued semimartingale X has the form

$$X_t = x + \int_0^t \sigma_r dW_r + \int_0^t b_r dr.$$

Here, $H = (H_1, \dots, H_d)$ with $H_k : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}, k = 1, \dots, d$ and $\int_t^T H(X_r, Y_r) d\eta(r) := \sum_{k=1}^d \int_t^T H_k(X_r, Y_r)\dot{\eta}^k(r)dr$. W is an m-dimensional Brownian motion (hence Z is a row vector taking values in $\mathbb{R}^{m\times 1}$ identified with $\mathbb{R}^m$). $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$ is a predictable random function, $x \in \mathbb{R}^n$, σ is a predictable process taking values in $\mathbb{R}^{n\times m}$, b is a predictable process taking values in $\mathbb{R}^n$.

Definition 34. We call equation (58) *BSDE with data* (ξ, f, H, η). We call (Y, Z) a solution if $Y \in H^\infty_{[0,T]}(\mathbb{R}), Z \in H^2_{[0,T]}(\mathbb{R}^m)$ and (58) is satisfied.

For a vector x we denote the Euclidean norm as usual by $|x|$. For a matrix X we denote by $|X|$, depending on the situation, either the 1-norm (operator norm), the 2-norm (Euclidean norm) or the ∞-norm (operator norm of the transpose). This slight abuse of notation will not lead to confusion, as all inequalities will be valid up to multiplicative constants.

We introduce the following assumptions:

(A1) There exists a constant $C_\sigma > 0$ such that $\mathbb{P} - a.s.$ for $t \in [0,T]$

$$|\sigma_t(\omega)| \le C_\sigma.$$

(A2) There exists a constant $C_b > 0$ such that $\mathbb{P} - a.s.$ for $t \in [0,T]$

$$|b_t(\omega)| \le C_b.$$

(F1) There exists a constant $C_{1,f} > 0$ such that $\mathbb{P} - a.s.$ for $(t, y, z) \in [0,T] \times \mathbb{R} \times \mathbb{R}^m$ [20]

$$|f(\omega; t, y, z)| \le C_{1,f} + C_{1,f}|z|^2,$$
$$|\partial_z f(\omega; t, y, z)| \le C_{1,f} + C_{1,f}|z|.$$

(F2) There exists a constant $C_{2,f} > 0$ such that $\mathbb{P} - a.s.$ for $(t, y, z) \in [0,T] \times \mathbb{R} \times \mathbb{R}^m$

$$\partial_y f(\omega; t, y, z) \le C_{2,f}.$$

For given real numbers $\gamma > p \ge 1$ we have the following assumption:

$(H_{p,\gamma})$ Let $H(x, \cdot) = (H_1(x, \cdot), \dots, H_d(x, \cdot))$ be a collection of vector fields on $\mathbb{R}$, parameterized by $x \in \mathbb{R}^n$. Assume that for some $C_H > 0$, we have joint regularity of the form

$$\sup_{i=1,\dots,d} |H_i|_{\mathrm{Lip}^{\gamma+2}(\mathbb{R}^{n+1})} \le C_H.$$

As a consequence of Theorem 2.3 and Theorem 2.6 in [43], we get the following

Lemma 35. *Assume (A1), (A2), (F1), (F2) and let H be Lipschitz on $\mathbb{R}^n \times \mathbb{R}$. Let $\xi \in L^\infty(\mathcal{F}_T)$ and a smooth path η be given. Then there exists a unique solution to the BSDE with data (ξ, f, H, η).*

[20] When we use partial derivatives, we assume implicitly that the function in question is continuously differentiable in the respective variable. In fact, throughout, it would suffice to assume (local) Lipschitzness and bound the Lipschitz constant analogously.

We want to give meaning to equation (58), where the smooth path η is replaced by a general geometric rough path $\eta \in C^{0,p-\mathrm{var}}([0,T], G^{[p]}(\mathbb{R}^d))$, where $G^{[p]}(\mathbb{R}^d)$ is the free step-$[p]$ nilpotent group over $\mathbb{R}^d$, realized as subset of $\mathbb{R} \oplus \mathbb{R}^d \oplus \cdots \oplus (\mathbb{R}^d)^{[p]}$, equipped with Carnot-Caratheodory metric We give our main result, the proof of which we present at the end of the section.

Theorem 36. *Let $p \geq 1$, $\gamma > p$ and $\eta^n, n = 1, 2, \ldots,$ be smooth paths in $\mathbb{R}^d$. Assume $\eta^n \to \eta$ in p-variation, for a path $\eta \in C^{0,p-\mathrm{var}}([0,T], G^{[p]}(\mathbb{R}^d))$. Let $\xi \in L^\infty(\mathcal{F}_T)$. Let f be a random function satisfying (F1) and (F2). Moreover, assume (A1), (A2) and ($H_{p,\gamma}$). For $n \geq 1$ denote by (Y^n, Z^n) the solutions to the BSDE with data (ξ, f, H, η^n).*

Then there exists a process $(Y, Z) \in H^\infty_{[0,T]} \times H^2_{[0,T]}$ such that

$$\begin{aligned} Y^n &\to Y \quad \text{uniformly on } [0,T]\ \mathbb{P}-a.s., \\ Z^n &\to Z \quad \text{in } H^2_{[0,T]}. \end{aligned}$$

The process is unique in the sense, that it only depends on the limiting rough path η and not on the approximating sequence. We write (formally [21])

$$Y_t = \xi + \int_t^T f(r, Y_r, Z_r) dr + \int_t^T H(X_r, Y_r) d\eta(r) - \int_t^T Z_r dW_r. \tag{59}$$

Moreover, the solution mapping

$$\begin{aligned} C^{0,p-\mathrm{var}}([0,T], G^{[p]}(\mathbb{R}^d)) \times L^\infty(\mathcal{F}_T) &\to H^\infty_{[0,T]} \times H^2_{[0,T]}, \\ (\eta, \xi) &\mapsto (Y, Z) \end{aligned}$$

is continuous.

The problem in showing convergence of the processes (Y^n, Z^n) in the statement of the theorem lies in the fact, that in general the Lipschitz constants for the corresponding BSDEs will tend to infinity as $n \to \infty$. It does not seem possible then, to directly control the solutions via a priori bounds, a standard tool in the theory of BSDEs (see e.g. [30]). We will take another approach and transform the BSDEs corresponding to the smooth paths η^n into BSDEs which are easier to analyze.

We start by defining the flow (parametrized by x)

$$\phi(t, x, y) = y + \int_t^T \sum_{k=1}^d H_k(x, \phi(r, x, y)) d\eta^k(r). \tag{60}$$

Let ϕ^{-1} be the y-inverse of ϕ, then

$$\phi^{-1}(t, x, y) = y - \int_t^T \sum_{k=1}^d \partial_y \phi^{-1}(r, x, y) H_k(x, y) d\eta^k(r).$$

We have the following

[21]The "integral" $\int H(X,Y)d\eta$ is *not* a rough integral defined in the usual rough path theory (e.g. [51] or [33]); regularity issues aside one misses the iterated integrals of X (and thus W) against those of η. For what it's worth, in the present context (59) can be taken as an implicit definition of $\int H(X,Y)d\eta$. (Somewhat similar in spirit: Föllmer's Itô's integral which appears in his Itô formula *sans probabilité*.) More pragmatically, notation (59) is justified *a posteriori* through our uniqueness result; in addition it is consistent with standard BSDE notation when η happens to be a smooth path.

Lemma 37. *Assume (A1), (A2), (F1), (F2) and let H be Lipschitz on $\mathbb{R}^n \times \mathbb{R}$. Let $\xi \in L^\infty(\mathcal{F}_T)$ and a smooth path η be given and let ϕ be the corresponding flow defined in* (60)*. Let (Y, Z) be the unique solution to the BSDE with data (ξ, f, H, η).*

Then, the process $(\tilde{Y}, \tilde{Z})$ defined as

$$\tilde{Y}_t := \phi^{-1}(t, X_t, Y_t), \quad \tilde{Z}_t := -\frac{\partial_x \phi(t, X_t, \tilde{Y}_t)}{\partial_y \phi(t, X_t, \tilde{Y}_t)} \sigma_t + \frac{1}{\partial_y \phi(t, X_t, \tilde{Y}_t)} Z_t,$$

satisfies the BSDE

$$\tilde{Y}_t = \xi + \int_t^T \tilde{f}(r, X_r, \tilde{Y}_r, \tilde{Z}_r) dr - \int_t^T \tilde{Z}_r dW_r, \tag{61}$$

where (throughout, ϕ and all its derivatives will always be evaluated at $(t, x, \tilde{y})$)

$$\tilde{f}(t, x, \tilde{y}, \tilde{z}) := \frac{1}{\partial_y \phi} \Big\{ f\left(t, \phi, \partial_y \phi \tilde{z} + \partial_x \phi \sigma_t\right) + \langle \partial_x \phi, b_t \rangle + \frac{1}{2} \operatorname{Tr}\left[\partial_{xx} \phi \sigma_t \sigma_t^T\right] \\ + \langle \tilde{z}, (\partial_{xy} \phi \sigma_t)^T \rangle + \frac{1}{2} \partial_{yy} \phi |\tilde{z}|^2 \Big\}.$$

Remark 38. *This ("Doss-Sussman") transformation is well known and has been recently applied to BDSDEs [7] and rough partial differential equations [32]. We include details for the reader's convenience.*

Proof. Denoting $\psi := \phi^{-1}$ and $\theta_r := (r, X_r, Y_r)$, we have by Itô formula

$$\begin{aligned}
\psi(t, X_t, Y_t) = \xi &- \int_t^T \sum_{k=1}^d \partial_y \psi(\theta_r) H_k(X_r, Y_r) \dot{\eta^k}(r) dr - \int_t^T \langle \partial_x \psi(\theta_r), b_r \rangle dr - \int_t^T \langle \partial_x \psi(\theta_r), \sigma_r dW_r \rangle \\
&+ \int_t^T \partial_y \psi(\theta_r) f(r, Y_r, Z_r) dr + \int_t^T \sum_{k=1}^d \partial_y \psi(\theta_r) H_k(X_r, Y_r) \dot{\eta^k}(r) dr - \int_t^T \partial_y \psi(\theta_r) Z_r dW_r \\
&- \frac{1}{2} \int_t^T \operatorname{Tr}\left[\partial_{xx} \psi(\theta_r) \sigma_r \sigma_r^T \rangle\right] dr - \frac{1}{2} \int_t^T \partial_{yy} \psi(\theta_r) |Z_r|^2 dr - \int_t^T \langle \partial_{xy} \psi(\theta_r), \sigma_r Z_r^T \rangle dr \\
= \xi &+ \int_t^T \Big[\partial_y \psi(\theta_r) f(r, Y_r, Z_r) - \langle \partial_x \psi(\theta_r), b_r \rangle - \frac{1}{2} \operatorname{Tr}\left[\partial_{xx} \psi(\theta_r) \sigma_r \sigma_r^T\right] \\
&\qquad - \frac{1}{2} \partial_{yy} \psi(\theta_r) |Z_r|^2 - \langle \partial_{xy} \psi(\theta_r), \sigma_r Z_r^T \rangle \Big] dr \\
&- \int_t^T \langle \partial_x \psi(\theta_r) \sigma_r + \partial_y \psi(\theta_r) Z_r, dW_r \rangle.
\end{aligned}$$

Now, by deriving the identity $\psi(t, x, \phi(t, x, \tilde{y})) = \tilde{y}$ we get

$$\begin{aligned}
0 &= \partial_x \psi + \partial_y \psi \partial_x \phi, \\
0 &= \partial_{xx} \psi + \partial_{yx} \psi \otimes \partial_x \phi + [\partial_{xy} \psi + \partial_{yy} \psi \partial_x \phi] \otimes \partial_x \phi + \partial_y \psi \partial_{xx} \phi \\
&= \partial_{xx} \psi + 2 \partial_{xy} \psi \otimes \partial_x \phi + \partial_{yy} \psi \partial_x \phi \otimes \partial_x \phi + \partial_y \psi \partial_{xx} \phi, \\
1 &= \partial_y \psi \partial_y \phi, \\
0 &= \partial_{xy} \psi \partial_y \phi + \partial_{yy} \psi \partial_x \phi \partial_y \phi + \partial_y \psi \partial_{xy} \phi, \\
0 &= \partial_{yy} \psi (\partial_y \phi)^2 + \partial_y \psi \partial_{yy} \phi.
\end{aligned}$$

And hence

$$\partial_{yy}\psi = -\frac{\partial_{yy}\phi}{(\partial_y\phi)^3}, \quad \partial_x\psi = -\frac{\partial_x\phi}{\partial_y\phi}, \quad \partial_{xy}\psi = \frac{\partial_{yy}\phi}{(\partial_y\phi)^3}\partial_x\phi - \frac{\partial_{xy}\phi}{(\partial_y\phi)^2},$$
$$\partial_{xx}\psi = 2\left[\frac{\partial_{yy}\phi}{(\partial_y\phi)^3}\partial_x\phi - \frac{\partial_{xy}\phi}{(\partial_y\phi)^2}\right]\otimes\partial_x\phi + \frac{\partial_{xx}\phi}{(\partial_y\phi)^3}\partial_x\phi\otimes\partial_x\phi - \frac{1}{\partial_y\phi}\partial_{xx}\phi.$$

If we define

$$\begin{aligned}\tilde{Y}_t &:= \psi(t, X_t, Y_t) = \phi^{-1}(t, X_t, Y_t),\\ \tilde{Z}_t &:= \partial_x\psi(t, X_t, Y_t)\sigma_t + \partial_y\psi(t, X_t, Y_t)Z_t\\ &= -\frac{\partial_x\phi(t, X_t, \tilde{Y}_t)}{\partial_y\phi(t, X_t, \tilde{Y}_t)}\sigma_t + \frac{1}{\partial_y\phi(t, X_t, \tilde{Y}_t)}Z_t,\end{aligned}$$

and (ψ and its derivatives are always evaluated at $(t, x, \phi(t, x, \tilde{y}))$, ϕ and its derivatives are evaluated at $(t, x, \tilde{y})$)

$$\begin{aligned}\tilde{f}(t, x, \tilde{y}, \tilde{z}) &:= \partial_y\psi f\left(t, \phi, \partial_y\phi(\tilde{z} + \frac{\partial_x\phi\sigma_t}{\partial_y\phi})\right) - \langle\partial_x\psi, b_t\rangle - \frac{1}{2}\operatorname{Tr}\left[\partial_{xx}\psi\sigma_t\sigma_t^T\right]\\ &\quad - \frac{1}{2}\partial_{yy}\psi\left|\frac{\tilde{z} - \partial_x\psi\sigma_t}{\partial_y\psi}\right|^2 - \langle\partial_{xy}\psi, \sigma_t\left(\frac{\tilde{z} - \partial_x\psi\sigma_t}{\partial_y\psi}\right)^T\rangle\\ &= \frac{1}{\partial_y\phi}\Big\{f\left(t, \phi, \partial_y\phi\tilde{z} + \partial_x\phi\sigma_t\right) + \langle\partial_x\phi, b_t\rangle + \frac{1}{2}\operatorname{Tr}\left[\partial_{xx}\phi\sigma_t\sigma_t^T\right]\\ &\quad + \langle\tilde{z}, (\partial_{xy}\phi\sigma_t)^T\rangle + \frac{1}{2}\partial_{yy}\phi|\tilde{z}|^2\Big\},\end{aligned}$$

we therefore obtain

$$\tilde{Y}_t = \xi + \int_t^T \tilde{f}(r, x, \tilde{Y}_r, \tilde{Z}_r)dr - \int_t^T \tilde{Z}_r dW_r.$$

□

Definition 39. We call equation (61) *BSDE with data* $(\xi, \tilde{f}, 0, 0)$.

The BSDE (58) only makes sense for a smooth path η. On the other hand, equation (60) yields a flow of diffeomorphisms for a general geometric rough path $\eta \in C^{0,p-\mathrm{var}}([0,T], G^{[p]}(\mathbb{R}^d)), p \geq 1$. Hence we can, also in this case, consider the function $\tilde{f}$ from the previous lemma. We now record important properties for this induced function.

Lemma 40. *Let $p \geq 1$, $\eta \in C^{0,p-\mathrm{var}}([0,T], G^{[p]}(\mathbb{R}^d))$ and $\gamma > p$. Assume (A1), (A2), (F1), (F2) and $(H_{p,\gamma})$. Let ϕ be the flow corresponding to equation* (60) *(now solved as a rough differential equation). Then the function*

$$\begin{aligned}\tilde{f}(t, x, \tilde{y}, \tilde{z}) := \frac{1}{\partial_y\phi}\Big\{f\left(t, \phi, \partial_y\phi\tilde{z} + \partial_x\phi\sigma_t\right) + \langle\partial_x\phi, b_t\rangle + \frac{1}{2}\operatorname{Tr}\left[\partial_{xx}\phi\sigma_t\sigma_t^T\right]\\ + \langle\tilde{z}, (\partial_{xy}\phi\sigma_t)^T\rangle + \frac{1}{2}\partial_{yy}\phi|\tilde{z}|^2\Big\}\end{aligned} \tag{62}$$

satisfies the following properties:

- *There exists a constant $\tilde{C}_{1,f} > 0$ depending only on C_σ, C_b, $C_{1,f}$, C_H and $||\eta||_{p-\text{var};[0,T]}$ such that*

$$|\tilde{f}(t,x,\tilde{y},\tilde{z})| \le \tilde{C}_{1,f} + \tilde{C}_{1,f}|\tilde{z}|^2,$$
$$|\partial_{\tilde{z}}\tilde{f}(t,x,\tilde{y},\tilde{z})| \le \tilde{C}_{1,f} + \tilde{C}_{1,f}|\tilde{z}|.$$

- *There exists a constant $\tilde{C}_{\text{unif}} > 0$ that only depends on C_σ, C_b, $C_{2,f}$, C_H and $||\eta||_{p-\text{var};[0,T]}$ such that for every ε there exists an $h_\varepsilon > 0$ that only depends on C_σ, C_b, C_H and $||\eta||_{p-\text{var};[0,T]}$ such that on $[T-h_\varepsilon, T]$ we have*

$$\partial_{\tilde{y}}\tilde{f}(t,x,\tilde{y},\tilde{z}) \le \tilde{C}_{\text{unif}} + \varepsilon|\tilde{z}|^2.$$

Proof. (i). Note that

$$\begin{aligned}
|\tilde{f}(t,x,\tilde{y},\tilde{z})| &\le |\frac{1}{\partial_y\phi}|\Big(|f\left(t,\phi,\partial_y\phi\tilde{z}+\partial_x\phi\sigma_t\right)| + |\langle\partial_x\phi, b_t\rangle| + |\frac{1}{2}\operatorname{Tr}\left[\partial_{xx}\phi\sigma_t\sigma_t^T\right]| \\
&\qquad + |\langle\tilde{z}, (\partial_{xy}\phi\sigma_t)^T\rangle| + |\frac{1}{2}\partial_{yy}\phi||\tilde{z}|^2\Big) \\
&\le |\frac{1}{\partial_y\phi}|\Big(C_{1,f} + C_{1,f}|\partial_y\phi\tilde{z}+\partial_x\phi\sigma_t|^2 + |\partial_x\phi||b_t| + \frac{1}{2}|\partial_{xx}\phi||\sigma_t\sigma_t^T| \\
&\qquad + |\tilde{z}||\partial_{xy}\phi\sigma_t| + \frac{1}{2}|\partial_{yy}\phi||\tilde{z}|^2\Big) \\
&\le |\frac{1}{\partial_y\phi}|\Big(C_{1,f} + C_{1,f}2(|\partial_y\phi|^2|\tilde{z}| + |\partial_x\phi||\sigma_t^T|) + |\partial_x\phi||b_t| + \frac{1}{2}|\partial_{xx}\phi||\sigma_t|^2 \\
&\qquad + |\tilde{z}||\partial_{xy}\phi||\sigma_t^T| + \frac{1}{2}|\partial_{yy}\phi||\tilde{z}|^2\Big) \\
&\le \tilde{C}_{1,f} + \tilde{C}_{1,f}|\tilde{z}|^2.
\end{aligned}$$

Here we have used (A1), (A2) and (F1). For the boundedness of the flow and its derivatives we have used Lemma 3.4. Note that $\tilde{C}_{1,f}$ hence only depends on C_σ, C_b, $C_{1,f}$, C_H and $||\eta||_{p-\text{var};[0,T]}$.
(ii). Note that

$$\begin{aligned}
|\partial_{\tilde{z}}\tilde{f}(t,x,\tilde{y},\tilde{z})| &= |\partial_z f\left(t,\phi,\partial_y\phi\tilde{z}+\partial_x\phi\sigma_t\right) + \frac{1}{\partial_y\phi}\Big(\partial_{xy}\phi\sigma_t + \partial_{yy}\phi\tilde{z}\Big)| \\
&\le C_{1,f} + C_{1,f}(|\partial_y\phi||\tilde{z}| + |\partial_x\phi||\sigma_t|) + |\frac{\partial_{xy}\phi}{\partial_y\phi}||\sigma_t| + |\frac{\partial_{yy}\phi}{\partial_y\phi}||\tilde{z}| \\
&\le \tilde{C}_{1,f} + \tilde{C}_{1,f}|\tilde{z}|.
\end{aligned}$$

Here we have used (A1), (A2) and (F1). For the boundedness of the flow and its derivatives we have used Lemma 3.4. Note that again, $\tilde{C}_{1,f}$ hence only depends on C_σ, C_b, $C_{1,f}$, C_H and $||\eta||_{p-\text{var};[0,T]}$. Without loss of generality we can choose it to be the same constant as in the estimate for (i).

(iii). Note that

$$\begin{aligned}\partial_{\tilde y}\tilde f(t,x,\tilde y,\tilde z) = &-\frac{\partial_{yy}\phi}{(\partial_y\phi)^2}\Big\{f\left(t,\phi,\partial_y\phi\tilde z+\partial_x\phi\sigma_t\right)+\langle\partial_x\phi,b_t\rangle+\frac{1}{2}\operatorname{Tr}\left[\partial_{xx}\phi\sigma_t\sigma_t^T\right]\\ &\quad+\langle\tilde z,(\partial_{xy}\phi\sigma_t)^T\rangle+\frac{1}{2}\partial_{yy}\phi|\tilde z|^2\Big\}\\ &+\frac{1}{\partial_y\phi}\Big\{\partial_y\phi\partial_y f\left(t,\phi,\partial_y\phi\tilde z+\partial_x\phi\sigma_t\right)+\langle\partial_{yx}\phi,b_t\rangle+\frac{1}{2}\operatorname{Tr}\left[\partial_{yxx}\phi\sigma_t\sigma_t^T\right]\\ &\quad+\langle\tilde z,(\partial_{yxy}\phi\sigma_t)^T\rangle+\frac{1}{2}\partial_{yyy}\phi|\tilde z|^2\Big\}.\end{aligned}$$

Hence using our assumptions on f we get

$$\begin{aligned}\partial_{\tilde y}\tilde f(t,x,\tilde y,\tilde z) \le\ &|\frac{\partial_{yy}\phi}{(\partial_y\phi)^2}|\Big\{C_{2,f}+C_{2,f}|\partial_y\phi\tilde z+\partial_x\phi\sigma_t|^2+|\partial_x\phi||b_t|+\frac{1}{2}|\partial_{xx}\phi||\sigma_t|^2\\ &\quad+|\tilde z||\partial_{xy}\phi||\sigma_t|+\frac{1}{2}|\partial_{yy}\phi||\tilde z|^2\Big\}\\ &+\partial_y f\left(t,\phi,\partial_y\phi\tilde z+\partial_x\phi\sigma_t\right)+\frac{1}{\partial_y\phi}\Big\{|\partial_{yx}\phi||b_t|+\frac{1}{2}|\partial_{yxx}\phi||\sigma_t|\\ &\quad+(1+|\tilde z|^2)|\partial_{yxy}\phi||\sigma_t|_{op}+\frac{1}{2}\partial_{yyy}\phi|\tilde z|^2\Big\}\\ \le\ &|\frac{\partial_{yy}\phi}{(\partial_y\phi)^2}|\Big\{C_{2,f}+C_{2,f}2|\partial_x\phi|^2|\sigma_t|^2+|\partial_x\phi||b_t|+\frac{1}{2}|\partial_{xx}\phi||\sigma_t|^2+|\partial_{xy}\phi||\sigma_t|\Big\}\\ &+\partial_y f\left(t,\phi,\partial_y\phi\tilde z+\partial_x\phi\sigma_t\right)\\ &+\frac{1}{\partial_y\phi}\Big\{|\partial_{yx}\phi||b_t|+\frac{1}{2}|\partial_{yxx}\phi||\sigma_t|+|\partial_{yxy}\phi||\sigma_t|\Big\}\\ &+\Big\{|\frac{\partial_{yy}\phi}{(\partial_y\phi)^2}|C_{2,f}2|\partial_y\phi|^2+|\frac{\partial_{yy}\phi}{(\partial_y\phi)^2}||\partial_{xy}\phi||\sigma_t|+|\frac{\partial_{yy}\phi}{(\partial_y\phi)^2}|\frac{1}{2}|\partial_{yy}\phi|\\ &\quad+\frac{1}{\partial_y\phi}|\partial_{yxy}\phi||\sigma_t^T|+\frac{1}{\partial_y\phi}\frac{1}{2}\partial_{yyy}\phi\Big\}|\tilde z|^2\\ \le\ &\tilde C_{\text{unif}}+\Big\{|\frac{\partial_{yy}\phi}{(\partial_y\phi)^2}|C_{2,f}2|\partial_y\phi|^2+|\frac{\partial_{yy}\phi}{(\partial_y\phi)^2}||\partial_{xy}\phi||\sigma_t|+|\frac{\partial_{yy}\phi}{(\partial_y\phi)^2}|\frac{1}{2}|\partial_{yy}\phi|\\ &\quad+\frac{1}{\partial_y\phi}|\partial_{yxy}\phi||\sigma_t^T|+\frac{1}{\partial_y\phi}\frac{1}{2}\partial_{yyy}\phi\Big\}|\tilde z|^2,\end{aligned}$$

where $\tilde C_{\text{unif}}$ only depends on C_σ, C_b, C_H and $||\eta||_{p-\text{var};[0,T]}$ (here we have used Lemma 3.4 to bound the flow and its derivatives).

By (A1), (A2) σ and b are bounded. Then, by the properties of the flow, the term in front of $|\tilde z|^2$ goes uniformly to zero as t approaches T. To be specific: using $(H_{p,\gamma})$ we obtain, again by Lemma 3.4, that for every $\varepsilon > 0$ there exists an $h_\varepsilon > 0$, depending on C_σ, C_b, C_H and $||\eta||_{p-\text{var};[0,T]}$ such that on $[T-h_\varepsilon, T]$ we have

$$\partial_{\tilde y}\tilde f(t,x,\tilde y,\tilde z) \le \tilde C_{\text{unif}} + \varepsilon|\tilde z|^2.$$

□

We are now ready to prove Theorem 36.

Proof of Theorem 36. For the sake of unified notation, we write (Y^0, Z^0) for the (rough BSDE) solution (Y, Z); similarly, we write η^0 for the rough path η.

1. **Existence**

Let $\phi^n, n \geq 0$ be the (ODE, for $n \geq 1$ and RDE, when $n = 0$) solution flow (parametrized by x)

$$\phi^n(t,x,y) = y + \int_t^T H(x, \phi^n(r,x,y)) d\eta^n(r).$$

By Lemma 3.4, we have for all $n \geq 0$, $x \in \mathbb{R}^n$, that $\phi^n(t,x,\cdot)$ is a flow of C^3-diffeomorphisms. Let $\psi^n(t,x,\cdot)$ be its y-inverse. We have that $\phi^n(t,\cdot,\cdot)$ and its derivatives up to order three are bounded (Lemma 3.4). The same holds true for $\psi^n(t,\cdot,\cdot)$ and its derivatives up to order three.

Moreover, by Lemma 3.5 we have that locally uniformly on $[0,T] \times \mathbb{R}^n \times \mathbb{R}$

$$(\phi^n, \frac{1}{\partial_y \phi^n}, \partial_y \phi^n, \partial_{yy}\phi^n, \partial_x \phi^n, \partial_{xx}\phi^n, \partial_{yx}\phi^n) \to (\phi^0, \frac{1}{\partial_y \phi^0}, \partial_y \phi^0, \partial_{yy}\phi^0, \partial_x \phi^0, \partial_{xx}\phi^0, \partial_{yx}\phi^0). \tag{63}$$

Denote for $n \geq 0$ the function

$$\begin{aligned}\tilde{f}^n(t,x,\tilde{y},\tilde{z}) := \frac{1}{\partial_y \phi^n}\Big\{ & f\left(t, \phi^n, \partial_y \phi^n \tilde{z} + \partial_x \phi^n \sigma_t\right) + \langle \partial_x \phi^n, b_t \rangle + \frac{1}{2} \operatorname{Tr}\left[\partial_{xx}\phi^n \sigma_t \sigma_t^T\right] \\ & + \langle \tilde{z}, (\partial_{xy}\phi^n \sigma_t)^T \rangle + \frac{1}{2}\partial_{yy}\phi^n |\tilde{z}|^2 \Big\}.\end{aligned}$$

Now, we have seen above that for $n \geq 1$, the process

$$\begin{aligned}(\tilde{Y}^n, \tilde{Z}^n) &:= L^n(Y^n, Z^n) \\ &:= ((\phi^n)^{-1}(\cdot, X_\cdot, Y_\cdot^n), -\frac{\partial_x \phi^n(\cdot, X_\cdot, (\phi^n)^{-1}(\cdot, X_\cdot, Y_\cdot^n))}{\partial_y \phi^n(\cdot, X_\cdot, (\phi^n)^{-1}(\cdot, X_\cdot, Y_\cdot^n))}\sigma_\cdot + \frac{1}{\partial_y \phi^n(\cdot, X_\cdot, (\phi^n)^{-1}(\cdot, X_\cdot, Y_\cdot^n))} Z_\cdot^n).\end{aligned}$$

solves the BSDE with data $(\xi, \tilde{f}^n, 0, 0)$.

Note that although $(\xi, \tilde{f}^n, 0, 0)$ is a quadratic BSDE, existence and uniqueness of a solution are guaranteed for $n \geq 1$ by the fact that the mapping L^n is one to one and by the existence of a unique solution to the untransformed BSDE (Lemma 35).

For $n = 0$, using the good properties of $\tilde{f}^0$ demonstrated in Lemma 40, there exists a solution $(\tilde{Y}^0, \tilde{Z}^0) \in H^\infty_{[0,T]} \times H^2_{[0,T]}$ to the BSDE with data $(\xi, \tilde{f}^0, 0, 0)$ by Theorem 2.3 in [43]. Note that it is a priori not unique, but we will show that it is at least unique on a small time interval up to T.

We now construct the process (Y^0, Z^0) of the statement on subintervals of $[0,T]$. First of all notice that we can choose the constant $\tilde{C}_{1,f}$ appearing in Lemma 40 uniformly for all $n \geq 0$. Let $M := ||\xi||_\infty + T\tilde{C}_{1,f}$. By Corollary 2.2 in [43] we have

$$||\tilde{Y}^n||_\infty \leq M, \quad n \geq 0. \tag{64}$$

Now by Lemma 40

- there exists $\tilde{C}_{1,f} > 0$ that only depends on C_σ, C_b, $C_{1,f}$, C_H and $||\eta||_{p-\text{var};[0,T]}$ such that

$$\begin{aligned}|\tilde{f}^0(t,x,y,z)| &\leq \tilde{C}_{1,f} + \tilde{C}_{1,f}|z|^2, \\ |\partial_z \tilde{f}^0(t,x,y,z)| &\leq \tilde{C}_{1,f} + \tilde{C}_{1,f}|z|.\end{aligned}$$

- There exists a constant $\tilde{C}_{\text{unif}} > 0$ that only depends on C_σ, C_b, $C_{2,f}$, C_H and $||\eta||_{p-\text{var};[0,T]}$ such that for every ε there exists an $h_\varepsilon > 0$ that only depends on C_σ, C_b, C_H and $||\eta||_{p-\text{var};[0,T]}$ such that on $[T-h_\varepsilon, T]$ we have

$$\partial_y \tilde{f}^0(t,x,y,z) \leq \tilde{C}_{\text{unif}} + \varepsilon |z|^2.$$

Hence we can choose $h = h_{\delta(\tilde{C}_{1,f},M)}$, such that for $t \in [T-h, T]$ we have

$$\partial_y \tilde{f}(t,x,y,z) \leq \tilde{C}_{\text{unif}} + \delta(\tilde{C}_{1,f}, M)|z|^2.$$

Here δ is the universal function given in the statement of Theorem 3.2. We can then apply Theorem 3.2 to get uniqueness of our solution $(\tilde{Y}^0, \tilde{Z}^0)$ on $[T-h,T]$. Now, as a consequence of (63) we have

$$\tilde{f}^n \to \tilde{f}^0 \qquad \text{uniformly on compacta.}$$

Hence, by the argument of Theorem 2.8 in [43] we have that

$$\begin{aligned} &\tilde{Y}^n \to \tilde{Y}^0 \quad \text{uniformly on } [T-h,T]\ \mathbb{P}-a.s., \\ &\tilde{Z}^n \to \tilde{Z}^0 \quad \text{in } H^2_{[T-h,T]}. \end{aligned} \tag{65}$$

Moreover, if we define

$$\begin{aligned} Y_t^0 &:= \phi^0(t, X_t, \tilde{Y}_t^0), \quad t \in [T-h, T], \\ Z_t^0 &:= \partial_y \phi^0(t, X_t, \tilde{Y}_t^0) \left[\tilde{Z}_t^0 + \frac{\partial_x \phi^0(t, X_t, \tilde{Y}_t^0)}{\partial_y \phi^0(t, X_t, \tilde{Y}_t^0)} \sigma_t \right], \quad t \in [T-h, T], \end{aligned}$$

and remembering that by construction

$$\begin{aligned} Y_t^n &= \phi^n(t, X_t, \tilde{Y}_t^n), \\ Z_t^n &= \partial_y \phi^n(t, X_t, \tilde{Y}_t^n) \left[\tilde{Z}_t^n + \frac{\partial_x \phi^n(t, X_t, \tilde{Y}_t^n)}{\partial_y \phi^n(t, X_t, \tilde{Y}_t^n)} \sigma_t \right], \end{aligned}$$

and using (63) we get

$$\begin{aligned} &Y^n \to Y^0 \quad \text{uniformly on } [T-h,T]\ \mathbb{P}-a.s., \\ &Z^n \to Z^0 \quad \text{in } H^2_{[T-h,T]}. \end{aligned} \tag{66}$$

Let us proceed to the next subinterval. To make the rough path disappear in the BSDE, we will use a similar transformation via a flow as above. As before we need to control the driver of the transformed BSDE, this time near $T-h$. For this reason we have to start the flow anew. First, we rewrite the BSDEs for $n \geq 1$ as

$$Y_t^n = Y_{T-h}^n + \int_t^{T-h} f(r, Y_r^n, Z_r^n) dr - \int_t^{T-h} H(X_r, Y_r^n) d\eta_r^n - \int_t^{T-h} Z_r^n dW_r.$$

Then define the flow $\phi^{n,T-h}$ started at time $T-h$, i.e.

$$\phi^{n,T-h}(t,x,y) = y + \int_t^{T-h} H(x, \phi^{n,T-h}(r,x,y)) d\eta^n(r), \quad t \leq T-h.$$

On $[0, T-h]$ define

$$\tilde{Y}^{n,T-h}_{\cdot}, := \phi^{n,T-h})^{-1}(\cdot, X_{\cdot}, Y^n_{\cdot}),$$

$$\tilde{Z}^{n,T-h}_{\cdot} := -\frac{\partial_x \phi^{n,T-h}(\cdot, X_{\cdot}, (\phi^{n,T-h})^{-1}(\cdot, X_{\cdot}, Y^n_{\cdot}))}{\partial_y \phi^{n,T-h}(\cdot, X_{\cdot}, (\phi^{n,T-h})^{-1}(\cdot, X_{\cdot}, Y^n_{\cdot}))}\sigma_{\cdot} + \frac{1}{\partial_y \phi^{n,T-h}(\cdot, X_{\cdot}, (\phi^{n,T-h})^{-1}(\cdot, X_{\cdot}, Y^n_{\cdot}))} Z^n_{\cdot}.$$

Then

$$\tilde{Y}^{n,T-h}_t = Y^n_{T-h} + \int_t^{T-h} \tilde{f}^{n,T-h}(r, X_r, \tilde{Y}^{n,T-h}_r, \tilde{Z}^{n,T-h}_r) dr - \int_t^{T-h} \tilde{Z}^{n,T-h}_r dW_r,$$

where

$$\tilde{f}^{n,T-h}(t,x,\tilde{y},\tilde{z}) := \frac{1}{\partial_y \phi^{n,T-h}} \Big\{ f\left(t, \phi^{n,T-h}, \partial_y \phi^{n,T-h} \tilde{z} + \partial_x \phi^{n,T-h} \sigma_t\right) + \langle \partial_x \phi^{n,T-h}, b_t \rangle$$
$$+ \frac{1}{2} \operatorname{Tr}\left[\partial_{xx} \phi^{n,T-h} \sigma_t \sigma_t^T\right] + \langle \tilde{z}, \left(\partial_{xy} \phi^{n,T-h} \sigma_t\right)^T \rangle + \frac{1}{2} \partial_{yy} \phi^{n,T-h} |\tilde{z}|^2 \Big\}.$$

This BSDE is also defined for $n = 0$ and as before we get via Lemma 40 for the same h and the same $\tilde{C}_{1,f}$ and $\tilde{C}_{\text{unif}}$ as before (here the explicit dependence of these constants is crucial), that on $[T-2h, T-h]$ we have

$$\partial_y \tilde{f}^{0,T-h}(t,x,y,z) \leq \tilde{C}_{\text{unif}} + \delta(\tilde{C}_{1,f}, M)|z|^2.$$

Hence we can apply Comparison Theorem 3.2 to get uniqueness of our solution $(\tilde{Y}^{0,T-h}, \tilde{Z}^{0,T-h})$ on $[T-2h, T-h]$. Now, also note that for the terminal value we have from (66) and (64)

$$\begin{aligned} &Y^n_{T-h} \to Y^0_{T-h} \quad \mathbb{P}-a.s., \\ &|Y^n_{T-h}| \leq M, \quad n \geq 1. \end{aligned}$$

Hence, again by the argument of Theorem 2.8 in [43] [22]

$$\begin{aligned} &\tilde{Y}^{n,T-h} \to \tilde{Y}^{0,T-h} \quad \text{uniformly on } [T-2h, T-h]\ \mathbb{P}-a.s., \\ &\tilde{Z}^{n,T-h} \to \tilde{Z}^{0,T-h} \quad \text{in } H^2_{[T-2h,T-h]}. \end{aligned}$$

Finally, reversing the transformation, we get as above

$$\begin{aligned} &Y^n \to Y^0 \quad \text{uniformly on } [T-2h, T-h]\ \mathbb{P}-a.s., \\ &Z^n \to Z^0 \quad \text{in } H^2_{[T-2h,T-h]}. \end{aligned}$$

Then, we can iterate this procedure on subintervals of length h up to time 0. Without loss of generality we can assume that $T = Nh$ for an $N \in \mathbb{N}$. Then, patching the results together we get

$$\sup_{t \leq T} |Y^n_t - Y^0_t| \leq \sum_{k=1}^{N} \sup_{(k-1)h \leq t \leq kh} |Y^n_t - Y^0_t| \to 0 \qquad \mathbb{P}-a.s.$$

[22] Note that Theorem 2.8 in [43] demands convergence in L^∞ of the terminal value. A closer look at the proof though, reveals that $\mathbb{P}$-a.s. convergence combined with a uniform deterministic bound (M in our case) is enough. To be specific: the convergence of the terminal value is only used at two instances for Theorem 2.8 and this is in the proof of Proposition 2.4 (which is the main ingredient for Theorem 2.8). Firstly, it is used on p. 568, right before Step 2 where it reads "By Lebesgue's dominated ...". Secondly, it is used on p. 570, before the end of the proof where it reads "from which we deduce that ...". In both cases, the above stated requirement is enough.

and

$$\mathbb{E}\left[\int_0^T |Z_r^n - Z_r^0|^2 dr\right] = \sum_{k=1}^N \mathbb{E}\left[\int_{(k-1)h}^{kh} |Z_r^n - Z_r^0|^2 dr\right] \to 0.$$

2. **Uniqueness**

Let $\bar{\zeta}^n, n \geq 1$ be another sequence of smooth paths that converges to η in p-variation. Let $(\bar{Y}^n, \bar{Z}^n)$ be the solutions to BSDEs with data $(\xi, f, H, \bar{\zeta}^n)$. Then, as above

$$\begin{aligned} \tilde{\bar{Y}}^n &\to \tilde{Y}^0 \quad \text{uniformly on } [T-h, T]\ \mathbb{P}-a.s., \\ \tilde{\bar{Z}}^n &\to \tilde{Z}^0 \quad \text{in } H^2_{[T-h,T]}. \end{aligned}$$

And hence

$$\begin{aligned} \bar{Y}^n &\to Y^0 \quad \text{uniformly on } [T-h, T]\ \mathbb{P}-a.s., \\ \bar{Z}^n &\to Z^0 \quad \text{in } H^2_{[T-h,T]}. \end{aligned}$$

Note that the choice of h in the proof of existence only depended on properties of the limiting function $\tilde{f}^0$, so we can use the same value here. One can now iterate this argument up to time 0 to get

$$\begin{aligned} \bar{Y}^n &\to Y^0 \quad \text{uniformly on } [0, T]\ \mathbb{P}-a.s., \\ \bar{Z}^n &\to Z^0 \quad \text{in } H^2_{[0,T]}, \end{aligned}$$

as desired.

3. **Continuity of the solution map**

We note that for a given $B > 0$, all terminal values ξ such that $|\xi| \leq B$ and all geometric p-rough paths with $||\eta||_{p-\text{var};[0,T]} \leq B$ we can choose an $h = h(B) > 0$ such that the above constructed unique solution (Y^0, Z^0) to the BSDE (59) is given by

$$Y_t^0 = \begin{cases} \phi^{0,T}(t, X_t, \tilde{Y}_t^T), & t \in [T-h, T], \\ \phi^{0,T-h}(t, X_t, \tilde{Y}_t^{T-h}), & t \in [T-2h, T-h], \\ \dots & \\ \phi^{0,h}(t, X_t, \tilde{Y}_t^h), & t \in [0, h], \end{cases}$$

$$Z_t^0 = \begin{cases} \partial_y \phi^{0,T}(t, X_t, \tilde{Y}_t^{0,T}) \left[\tilde{Z}_t^{0,T} + \frac{\partial_x \phi^{0,T}(t, X_t, \tilde{Y}_t^{0,T})}{\partial_y \phi^0(t, X_t, \tilde{Y}_t^{0,T})} \sigma_t\right], & t \in [T-h, T], \\ \partial_y \phi^{0,T-h}(t, X_t, \tilde{Y}_t^{0,T-h}) \left[\tilde{Z}_t^{0,T-h} + \frac{\partial_x \phi^{0,T-h}(t, X_t, \tilde{Y}_t^{0,T-h})}{\partial_y \phi^{0,T-h}(t, X_t, \tilde{Y}_t^{0,T-h})} \sigma_t\right], & t \in [T-2h, T-h], \\ \dots & \\ \partial_y \phi^{0,h}(t, X_t, \tilde{Y}_t^{0,h}) \left[\tilde{Z}_t^{0,h} + \frac{\partial_x \phi^{0,h}(t, X_t, \tilde{Y}_t^{0,h})}{\partial_y \phi^{0,h}(t, X_t, \tilde{Y}_t^{0,h})} \sigma_t\right], & t \in [0, h], \end{cases}$$

where we used the unique solutions to the following BSDEs

$$\begin{aligned} \tilde{Y}_t^{0,T} &= \xi + \int_t^T \tilde{f}^{0,T}(r, X_r, \tilde{Y}_r^{0,T}, \tilde{Z}_r^{0,T}) dr - \int_t^T \tilde{Z}_r^{0,T} dW_r, \\ \tilde{Y}_t^{0,T-h} &= \phi^{0,T}(T-h, X_{T-h}, \tilde{Y}_{T-h}^{0,T}) + \int_t^{T-h} \tilde{f}^{0,T-h}(r, X_r, \tilde{Y}_r^{0,T-h}, \tilde{Z}_r^{0,T-h}) dr - \int_t^{T-h} \tilde{Z}_r^{0,T-h} dW_r, \\ &\dots \\ \tilde{Y}_t^{0,h} &= \phi^{0,2h}(h, X_h, \tilde{Y}_h^{0,2h}) + \int_t^h \tilde{f}^{0,h}(r, X_r, \tilde{Y}_r^{0,h}, \tilde{Z}_r^{0,h}) dr - \int_t^h \tilde{Z}_r^{0,h} dW_r. \end{aligned}$$

From this representation and stability results on BSDEs (Theorem 2.8 in [43]) it easily follows that the solution map

$$C^{0,p-\mathrm{var}}([0,T],G^{[p]}(\mathbb{R}^d))\times L^\infty(\mathcal{F}_T)\to H^\infty_{[0,T]}\times H^2_{[0,T]}$$

is continuous in balls of radius B. Since this is true for every $B>0$ we get the desired result. □

3.3 The Markovian Setting - Connection to rough PDEs

We now specialize to a Markovian model. We are interested in solving the following forward backward stochastic differential equation for $(t_0,x_0)\in[0,T]\times\mathbb{R}^n$

$$\begin{aligned} X_t^{t_0,x_0} &= x_0+\int_{t_0}^t \sigma(r,X_r^{t_0,x_0})dW_r+\int_{t_0}^t b(r,X_r^{t_0,x_0})dr, \quad t\in[t_0,T],\\ Y_t^{t_0,x_0} &= g(X_T^{t_0,x_0})+\int_t^T f(r,X_r^{t_0,x_0},Y_r^{t_0,x_0},Z_r^{t_0,x_0})dr \\ &\quad +\int_t^T H(X_r^{t_0,x_0},Y_r^{t_0,x_0})d\eta_r-\int_t^T Z_r^{t_0,x_0}dW_r, \quad t\in[t_0,T]. \end{aligned} \tag{67}$$

Here $\sigma:[0,T]\times\mathbb{R}^n\to\mathbb{R}^{n\times m}, b:[0,T]\times\mathbb{R}^n\to\mathbb{R}^n, f:[0,T]\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m$, $g:\mathbb{R}^n\to\mathbb{R}$, $H=(H_1,\dots,H_d):\mathbb{R}^n\times\mathbb{R}\to\mathbb{R}^d$, $\eta:[0,T]\to\mathbb{R}^d$ are continuous mappings, on which more assumptions will be presented later.

Assume for the moment that η is actually a smooth path. Then (67) is connected to the PDE

$$\begin{aligned} &\partial_t u(t,x)+\frac{1}{2}\operatorname{Tr}[\sigma(t,x)\sigma(t,x)^T D^2u(t,x)]+\langle b(t,x),Du(t,x)\rangle\\ &\quad +f(t,x,u(t,x),Du(t,x)\sigma(t,x))+H(x,u(t,x))\dot\eta_t=0, \quad t\in[0,T),x\in\mathbb{R}^n,\\ &u(T,x)=g(x), \quad x\in\mathbb{R}^n. \end{aligned} \tag{68}$$

We will make this connection explicit after introducing the following adaption (and strengthening) of previous assumptions:

(MA1) There exists a constant $C_\sigma>0$ such that for $(t,x)\in[0,T]\times\mathbb{R}^n$

$$\begin{aligned} |\sigma(t,x)| &\le C_\sigma,\\ |\partial_{x_i}\sigma(t,x)| &\le C_\sigma, \quad i=1,\dots,n. \end{aligned}$$

(MA2) There exists a constant $C_b>0$ such that for $(t,x)\in[0,T]\times\mathbb{R}^n$

$$\begin{aligned} |b(t,x)| &\le C_b,\\ |\partial_x b(t,x)| &\le C_b. \end{aligned}$$

(MF1) There exists a constant $C_{1,f}>0$ such that for $(t,x,y,z)\in[0,T]\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m$

$$\begin{aligned} |f(t,x,y,z)| &\le C_{1,f},\\ |\partial_z f(t,x,y,z)| &\le C_{1,f}. \end{aligned}$$

(MF2) There exists a constant $C_{2,f} > 0$ such that such that for $(t,x,y,z) \in [0,T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m$

$$\partial_y f(t,x,y,z) \leq C_{2,f}.$$

(MF3) There exists a constant $C_{3,f} > 0$ such that such that for $(t,x,y,z) \in [0,T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m$

$$\partial_x f(t,x,y,z) \leq C_{3,f} + C_{3,f}|z|^2,$$

and f is uniformly continuous in x, uniformly in (t,y,z).

(MG1) g is bounded and uniformly continuous.

We again consider for a smooth (or rough) path η the flow (parametrized by x)

$$\phi(t,x,y) = y + \int_t^T \sum_{k=1}^d H_k(x, \phi(r,x,y)) d\eta^k(r). \tag{69}$$

Proposition 41. *Assume (MA1), (MA2), (MF1), (MF2), (MF3), (MG1) and let H be Lipschitz on $\mathbb{R}^n \times \mathbb{R}$. Let η be a smooth path. Then there exists a unique viscosity solution* [23] *to* (68) *in $BUC([0,T],\mathbb{R}^n)$.*

It is given by

$$u(t,x) := Y_t^{t,x},$$

where for every $(t_0,x_0) \in [0,T] \times \mathbb{R}^n$ the process $(Y^{t_0,x_0}, Z^{t_0,x_0})$ is the solution to (67).

Proof. The fact that u is a bounded, uniformly continuous viscosity solution follows from Proposition 2.5 and Theorem 3.4 in [2]. Uniqueness of a bounded viscosity solution to (68) follows from Theorem 3.6. Note, that since $\partial y f$ is bounded, we can choose $h_\varepsilon \equiv T$ in the statement of the theorem and hence get uniqueness on the entire interval $(0,T]$. Every BUC function on $(0,T] \times \mathbb{R}^n$ has a unique extension to $[0,T] \times \mathbb{R}^n$. Hence u is unique in $BUC([0,T],\mathbb{R}^n)$. □

Remark 42. *It is also possible to show existence of a (unique) solution to* (68) *by purely deterministic methods, see e.g. Theorem 2 in [28].*

Let now $\eta^n, n = 1,2,\ldots$, be smooth paths in $\mathbb{R}^d$. Let $\gamma > p \geq 1$ and assume $\eta^n \to \eta^0$ in p-variation, for a $\eta^0 \in C^{0,p-\mathrm{var}}([0,T], G^{[p]}(\mathbb{R}^d))$. Assume (MA1), (MA2), (MF1), (MF2), (MF3), (MG1) and $(H_{p,\gamma})$, so that especially Theorem 36 holds true. It follows that the corresponding u^n (as given in Proposition 41) converge *pointwise* to some function u^0, i.e.

$$u^n(t,x) \to u^0(t,x) \qquad t \in [0,T], x \in \mathbb{R}^n.$$

Again, the limiting function u^0 does not depend on the approximating sequence, but only on the limiting rough path η^0. We could hence define this u^0 to be the solution solution to (68). But it is not straightforward, via this approach, to show uniform convergence on compacta as well as continuity of the solution map. We hence work directly on the PDEs, as in [10] and [32]. First we get the respective versions of Lemma 37 and Lemma 40.

[23] For an introduction to the theory of viscosity solutions we refer the reader to [18].

Lemma 43. *Assume (MA1), (MA2), (MF1), (MF2), (MG1) and let $H(x,\cdot) = (H_1(x,\cdot),\dots,H_d(x,\cdot))$ be a collection of Lipschitz vector fields on $\mathbb{R}$. Let a smooth path η be given. Let u be the unique viscosity solution to* (68)*.*

Then $v(t,x) := \phi^{-1}(t,x,u(t,x))$ is a viscosity solution to

$$\begin{aligned}
&\partial_t v(t,x) + \frac{1}{2}\operatorname{Tr}[\sigma(t,x)\sigma(t,x)^T D^2 v(t,x)] + \langle b(t,x), Dv(t,x)\rangle \\
&\qquad + \tilde f(t,x,v(t,x),Dv(t,x)\sigma(t,x)) = 0, \quad t\in[0,T), x\in\mathbb{R}^n,\\
&v(T,x) = g(x), \quad x\in\mathbb{R}^n,
\end{aligned}$$

where (in what follows the ϕ will always be evaluated at $(t,x,\tilde y)$)

$$\begin{aligned}
\tilde f(t,x,\tilde y,\tilde z) = \frac{1}{\partial_y\phi}\Big\{ & f\left(t,x,\phi,\partial_y\phi\tilde z + \partial_x\phi\sigma(t,x)\right) + \langle\partial_x\phi, b(t,x)\rangle + \frac{1}{2}\operatorname{Tr}\left[\partial_{xx}\phi\sigma(t,x)\sigma(t,x)^T\right]\\
&+ \langle \tilde z, (\partial_{xy}\phi\sigma(t,x))^T\rangle + \frac{1}{2}\partial_{yy}\phi|\tilde z|^2\Big\}.
\end{aligned}$$

Proof. This is an application of Lemma 5 in [32]. □

Lemma 44. *Let $p\ge 1$, $\eta\in C^{0,p-\mathrm{var}}([0,T],G^{[p]}(\mathbb{R}^d))$ and $\gamma>p$. Assume (MA1), (MA2), (MF1), (MF2), (MF3), (MG1) and ($H_{p,\gamma}$). Let ϕ be the flow corresponding to equation* (69) *(solved as a rough differential equation). Then*

$$\begin{aligned}
\tilde f(t,x,\tilde y,\tilde z) = \frac{1}{\partial_y\phi}\Big\{ & f\left(t,x,\phi,\partial_y\phi\tilde z + \partial_x\phi\sigma(t,x)\right) + \langle\partial_x\phi, b(t,x)\rangle + \frac{1}{2}\operatorname{Tr}\left[\partial_{xx}\phi\sigma(t,x)\sigma(t,x)^T\right]\\
&+ \langle \tilde z, (\partial_{xy}\phi\sigma(t,x))^T\rangle + \frac{1}{2}\partial_{yy}\phi|\tilde z|^2\Big\}
\end{aligned}$$

satisfies:

- *There exists a constant $\tilde C_{1,f}>0$ depending only on C_σ, C_b, $C_{1,f}$, C_H and $||\eta||_{p-\mathrm{var};[0,T]}$ such that for $(t,x,\tilde y,\tilde z)\in[0,T]\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m$*

$$\begin{aligned}
|\tilde f(t,x,\tilde y,\tilde z)| &\le \tilde C_{1,f} + \tilde C_{1,f}|\tilde z|^2,\\
|\partial_{\tilde z}\tilde f(t,x,\tilde y,\tilde z)| &\le \tilde C_{1,f} + \tilde C_{1,f}|\tilde z|.
\end{aligned}$$

- *There exists a constant $\tilde C_{\mathrm{unif}}>0$ that only depends on C_σ, C_b, $C_{2,f}$, C_H and $||\eta||_{p-\mathrm{var};[0,T]}$ such that for every $\varepsilon>0$ there exists an $h_\varepsilon>0$ that only depends on C_σ, C_b, C_H and $||\eta||_{p-\mathrm{var};[0,T]}$ such that for $(t,x,\tilde y,\tilde z)\in[T-h_\varepsilon,T]\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m$*

$$\partial_{\tilde y}\tilde f(t,x,\tilde y,\tilde z) \le \tilde C_{\mathrm{unif}} + \varepsilon|\tilde z|^2.$$

- *There exists a $\tilde C_{3,f}>0$ that only depends on C_σ, C_b, $C_{2,f}$, $C_{3,f}$, C_H and $||\eta||_{p-\mathrm{var};[0,T]}$ such that for $(t,x,\tilde y,\tilde z)\in[0,T]\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m$*

$$\partial_x\tilde f(t,x,\tilde y,\tilde z) \le \tilde C_{3,f} + \tilde C_{3,f}|\tilde z|^2.$$

Proof. The first three inequalities follow as in Lemma 40. Now for $i \le n$ we have

$$
\begin{aligned}
&\partial_{x_i} \tilde{f}(t,x,\tilde{y},\tilde{z}) \\
&= -\partial_{x_i y}\phi \frac{1}{\partial_y \phi} \tilde{f}(t,x,\tilde{y},\tilde{z}) \\
&\quad + \frac{1}{\partial_y\phi}\Big[\partial_y f(t,x,\phi,\partial_y\phi\tilde{z} + \partial_x\phi\sigma(t,x))\partial_{x_i}\phi \\
&\qquad + \partial_z f(t,x,\phi,\partial_y\phi\tilde{z} + \partial_x\phi\sigma(t,x))\left(\partial_{x_i y}\phi\tilde{z} + \partial_{x_i x}\phi\sigma(t,x) + \partial_x\phi\partial_{x_i}\sigma(t,x)\right)^T \\
&\qquad + \langle \partial_{x_i x}\phi, b(t,x)\rangle + \langle \partial_x\phi, \partial_{x_i} b(t,x)\rangle \\
&\qquad + \frac{1}{2}\operatorname{Tr}\left[\partial_{x_i xx}\phi\sigma(t,x)\sigma(t,x)^T\right] + \frac{1}{2}\operatorname{Tr}\left[\partial_{xx}\phi\partial_{x_i}\sigma(t,x)\sigma(t,x)^T\right] + \frac{1}{2}\operatorname{Tr}\left[\partial_{xx}\phi\sigma(t,x)\partial_{x_i}\sigma(t,x)^T\right] \\
&\qquad + \langle \tilde{z}, (\partial_{x_i xy}\phi\sigma(t,x))^T\rangle + \langle \tilde{z}, (\partial_{xy}\phi\partial_{x_i}\sigma(t,x))^T\rangle + \frac{1}{2}\partial_{x_i yy}\phi|\tilde{z}|^2\Big].
\end{aligned}
$$

So

$$
\begin{aligned}
&|\partial_{x_i}\tilde{f}(t,x,\tilde{y},\tilde{z})| \\
&\le |\partial_{x_i y}\phi|\left|\frac{1}{\partial_y\phi}\right||\tilde{f}(t,x,\tilde{y},\tilde{z})| \\
&\quad + \left|\frac{1}{\partial_y\phi}\right|\Big[|\partial_y f(t,x,\phi,\partial_y\phi\tilde{z} + \partial_x\phi\sigma(t,x))||\partial_{x_i}\phi| \\
&\qquad + |\partial_z f(t,x,\phi,\partial_y\phi\tilde{z} + \partial_x\phi\sigma(t,x))|\left(|\partial_{x_i y}\phi||\tilde{z}| + |\partial_{x_i x}\phi||\sigma(t,x)| + |\partial_x\phi||\partial_{x_i}\sigma(t,x)|\right) \\
&\qquad + |\langle\partial_{x_i x}\phi||b(t,x)| + |\partial_x\phi||\partial_{x_i} b(t,x)| \\
&\qquad + \frac{1}{2}|\partial_{x_i xx}\phi||\sigma(t,x)|^2 + |\partial_{xx}\phi||\partial_{x_i}\sigma(t,x)||\sigma(t,x)| \\
&\qquad + |\tilde{z}||\partial_{x_i xy}\phi||\sigma(t,x)| + |\tilde{z}||\partial_{xy}\phi||\partial_{x_i}\sigma(t,x)| + \frac{1}{2}|\partial_{x_i yy}\phi||\tilde{z}|^2\Big]| \\
&\le \tilde{C}_{3,f} + \tilde{C}_{3,f}|\tilde{z}|^2
\end{aligned}
$$

with a constant $\tilde{C}_{3,f}$ only depending on C_σ, C_b, $C_{1,f}$, $C_{2,f}$, $C_{3,f}$, C_H and $||\eta||_{p-\text{var};[0,T]}$. Here we have used the first inequality of the statement to bound $\tilde{f}$, (MF1), (MF2), (MF3) to bound f and its y and z derivatives and Lemma 3.4 to bound the flow and its derivatives.

Summing over i then yields the desired result. □

Theorem 45. *Let $\gamma > p \ge 1$ and let $\eta^n, n = 1, 2, \ldots$ be smooth paths in $\mathbb{R}^d$. Assume*

$$\eta^n \to \eta$$

in p-variation, for a rough path $\eta \in C^{0,p-\text{var}}([0,T], G^{[p]}(\mathbb{R}^d))$. Assume (MA1), (MA2), (MF1), (MF2), (MF3), (MG1) and ($H_{p,\gamma}$). Let $u^n \in BUC([0,T]\times\mathbb{R}^n)$ be the unique solution to (68) *with driving path η^n (Proposition 41). Then there exists $u \in BC([0,T]\times\mathbb{R}^n)$, only dependent on η but not on the approximating sequence η^n, such that*

$$u^n \to u \qquad \textit{locally uniformly.}$$

We write (formally)

$$
\begin{aligned}
&du + \left[\frac{1}{2}\operatorname{Tr}[\sigma(t,x)\sigma(t,x)^T D^2 u(t,x)] + \langle b(t,x), Du(t,x)\rangle + f(t,x,u(t,x),Du(t,x)\sigma(t,x))\right]dt \\
&\qquad + H(x,u(t,x))d\eta(t) = 0, \quad t \in [0,T), x \in \mathbb{R}^n, \qquad (70) \\
&u(T,x) = g(x), \quad x \in \mathbb{R}^n.
\end{aligned}
$$

Furthermore, the solution map

$$C^{0,p-\text{var}}([0,T],G^{[p]}(\mathbb{R}^d))\times BUC(\mathbb{R}^n)\to BC([0,T]\times\mathbb{R}^n),$$
$$(\eta,g)\mapsto u$$

is continuous.

At last we have the stochastic representation

$$u(t,x)=Y_t^{t,x},$$

where $Y^{t,x}$ is (the Y-component of) the solution to the BSDE (67).

Remark 46. *Equations like* (70) *have been considered in [32]. The setting there is more general in the sense that the vector field H in front of the rough path is allowed to also depend on the gradient. On the other hand, their f is independent of the gradient and H is linear.*

For the proof we apply the same ideas as in the proof of Theorem 1 in [10]. We however mimic our analysis of the BSDE case (Theorem 36) and proceed on small intervals; a similar approach was carried out in Lions-Souganidis [48].

Remark 47. *We suspect the solution to actually lie in $BUC([0,T]\times\mathbb{R}^n)$. Showing this would involve adapting the comparison theorem 3.6 to directly yield a modulus of continuity for solutions, as it has been done in [28] under different assumptions on the coefficients.*

Proof. For the sake of unified notation, we write u^0 for the (rough PDE) solution u; similarly, we write η^0 for the rough path η.

1. **Existence**

Let $\phi^n, n\geq 0$ be the (ODE, for $n\geq 1$ and RDE, when $n=0$) solution flow (parametrized by x)

$$\phi^n(t,x,y)=y+\int_t^T H(x,\phi^n(r,x,y))d\eta^n(r).$$

Then, by Lemma 43, for $n\geq 1$, u^n is a solution to (68) if and only if $v^n(t,x):=(\phi^n)^{-1}(t,x,u^n(t,x))$ is a solution to

$$\begin{aligned}&\partial_t v^n(t,x)+\frac{1}{2}\operatorname{Tr}[\sigma(t,x)\sigma(t,x)^T D^2 v^n(t,x)]+\langle b(t,x),Dv^n(t,x)\rangle\\&\quad+\tilde f^n(t,x,v^n(t,x),Dv^n(t,x)\sigma(t,x))=0,\quad t\in[0,T),x\in\mathbb{R}^n,\\&v^n(T,x)=g(x),\quad x\in\mathbb{R}^n,\end{aligned}\tag{71}$$

where

$$\tilde f^n(t,x,\tilde y,\tilde z)=\frac{1}{\partial_y\phi^n}\Big\{f\left(t,x,\phi^n,\partial_y\phi^n\tilde z+\partial_x\phi^n\sigma(t,x)\right)+\langle\partial_x\phi^n,b(t,x)\rangle+\frac{1}{2}\operatorname{Tr}\left[\partial_{xx}\phi^n\sigma(t,x)\sigma(t,x)^T\right]$$
$$+\langle\tilde z,(\partial_{xy}\phi^n\sigma(t,x))^T\rangle+\frac{1}{2}\partial_{yy}\phi^n|\tilde z|^2\Big\}.$$

In the proof of Theorem 36 we have already seen that $\tilde f^n\to\tilde f^0$, locally uniformly. From the method of semi-relaxed limits (Lemma 6.1, Remark 6.2-6.4 in [18]), the pointwise (relaxed) limits

$$\bar v^0:=\limsup^* v^n,\quad \underline{v}^0:=\liminf_* v^n,$$

are viscosity (sub resp. super) solutions to the (transformed) PDE (71) for $n = 0$. Here we have used the fact, that $\bar{v}^0$ and $\underline{v}^0$ are indeed finite, say bounded in norm by $M > 0$. This follows from the Feynman-Kac representation (Proposition 41) for each u^n, in combination with bounds (uniform in (t_0, x_0) and n) on the corresponding BSDEs (Corollary 2.2 in [43]). (Although not completely obvious, such uniform bounds can also be obtained without BSDE arguments; one would need to exploit comparison for (68), and then (71), clearly valid when $n \geq 1$, with rough path estimates for RDE solutions which will serve as sub- and super-solutions without spatial structure.) Note also, that the semi-relaxed limiting procedure preserves the terminal value (see e.g. Proposition 5.1 in [31] or Section 10 in [10]).

By Lemma 44 the function $\tilde{f}^0$ satisfies the conditions of Theorem 3.6. Hence the PDE (71) for $n = 0$ satisfies comparison on $[T-h, T]$ for h sufficiently small (as long as $T - h > 0$), and h only depends on M and the constants $\tilde{C}_{\text{unif}}$, $\tilde{C}_{1,f}$ and $\tilde{C}_{2,f}$ for $\tilde{f}^0$ given by Lemma 44. So $v^0(t,x) := \bar{v}^0(t,x) = \underline{v}^0(t,x), t \in [T-h, T]$ is the unique (and continuous, since $\bar{v}, \underline{v}$ are respectively upper resp. lower semi-continuous) solution to (71) with $n = 0$ on $[T-h, T]$. Moreover, using a Dini-type argument (Remark 6.4 in [18]), one sees that this limit must be uniform on compact sets. Undoing the transformation, we see that $u^n \to u^0$ locally uniformly on $[T-h, T]$, where $u^0(t,x) := \phi^0(t, x, v^0(t,x)),\ t \in [T-h, T]$.

We proceed to the next subinterval. We use the same argument as above, we just work with a different transformation. For $n \geq 0$ let $\phi^{n,T-h}$ be the solution flow started at time $T - h$, i.e.

$$\phi^{n,T-h}(t,x,y) = y + \int_t^{T-h} H(x, \phi^{n,T-h}(r,x,y)) d\eta^n(r).$$

Then, for $n \geq 1$, $u^n|_{[0,T-h]}$ is a solution to

$$\begin{aligned} &\partial_t u^n(t,x) + \frac{1}{2} \operatorname{Tr}[\sigma(t,x)\sigma(t,x)^T D^2 u^n(t,x)] + \langle b(t,x), Du^n(t,x) \rangle \\ &\quad + f(t,x,u^n(t,x), Du^n(t,x)\sigma(t,x)) + H(x, u^n(t,x))\dot{\eta}_r = 0, \quad t \in [0, T-h], x \in \mathbb{R}^n, \\ &u(T-h, x) = \phi^n(T-h, x, v^n(T-h, x)), \quad x \in \mathbb{R}^n. \end{aligned}$$

if and only if $v^{n,T-h}(t,x) := (\phi^{n,T-h})^{-1}(t,x,u^n(t,x))$ is a solution to

$$\begin{aligned} &\partial_t v^{n,T-h}(t,x) + \frac{1}{2} \operatorname{Tr}[\sigma(t,x)\sigma(t,x)^T D^2 v^{n,T-h}(t,x)] + \langle b(t,x), Dv^{n,T-h}(t,x) \rangle \\ &\quad + \tilde{f}^{n,T-h}(t,x,v^{n,T-h}(t,x), \sigma(t,x) Dv^{n,T-h}(t,x)) = 0, \quad t \in (0, T-h), x \in \mathbb{R}^n, \\ &v^{n,T-h}(T,x) = \phi^n(T-h, x, v^n(T-h, x)), \quad x \in \mathbb{R}^n, \end{aligned}$$

where of course $\tilde{f}^{n,T-h}$ is defined as $\tilde{f}^n$ was, with ϕ^n replaced by $\phi^{n,T-h}$.

Now we have already shown that the terminal values of these PDEs converge, e.g.

$$\phi^n(T-h, \cdot, v^n(T-h, \cdot)) \to \phi(T-h, \cdot, v(T-h, \cdot)), \text{ locally uniformly.}$$

As before, one also shows that $\tilde{f}^{n,T-h} \to \tilde{f}^{0,T-h}$, locally uniformly. By Theorem 3.6 we again get comparison, now on $[T-2h, T-h]$, and hence again via the method of semi-relaxed limits we arrive at [24]

$$v^{n,T-h} \to v^{0,T-h} \quad \text{locally uniformly on } [T-2h, T-h] \times \mathbb{R}^n.$$

[24] Remark 6.3 in [18] does not take into account converging terminal values. But the result is immediate: the relaxed limit is a sub resp. super solution by Lemma 6.3 and by Proposition 2 in [10] their terminal value is exactly the limit of the given converging terminal values.

Hence $u^n \to u^0$ locally uniformly on $[T-2h, T-h]$, where $u^0(t,x) = \phi^{0,T-h}(t,x,v^{0,T-h}(t,x))$. Iterating this argument up to time 0 we get

$$u^n \to u^0 \quad \text{locally uniformly on } [0,T] \times \mathbb{R}^n,$$

where u^0 is defined on intervals of length h as above. [25]

2. **Uniqueness, Continuity of solution map**

Uniqueness of the limit and continuity of the solution map now follow by the same arguments as in the proof of Theorem 36, adapted to the PDE setting.

3. **Stochastic representation**

Let $\eta^n \to \eta^0$ as above. Denote by u^n the solution to the corresponding PDE (rough PDE for $n = 0$). Denote by $Y^{n,t,x}$ the solution to the BSDE (67) (BSDE with rough driver for $n = 0$) corresponding to the path η^n.

Then, using the result from step 1, the stochastic representation in the case of a smooth path from Proposition 41 and the convergence of the BSDEs from Theorem 36, we get

$$u^0(t,x) = \lim_{n\to\infty} u^n(t,x) = \lim_{n\to\infty} Y_t^{n,t,x} = Y_t^{0,t,x}.$$

□

3.4 Connection to BDSDEs

Let $\Omega^1 = C([0,T],\mathbb{R}^d)$, $\Omega^2 = C([0,T],\mathbb{R}^m)$, with the respective Wiener measures $\mathbb{P}^1$, $\mathbb{P}^2$ on them. Let $\Omega = \Omega^1 \times \Omega^2$, with the product measure $\mathbb{P} := \mathbb{P}^1 \otimes \mathbb{P}^2$. For $(\omega^1,\omega^2) \in \Omega$ let $B(\omega^1,\omega^2) = \omega^1$ be the coordinate mapping with respect to the first component. Analogously $W(\omega^1,\omega^2) = \omega^2$ is the coordinate mapping with respect to the second component. In particular, B is a d-dimensional Brownian motion and W is an independent m-dimensional Brownian motion.

Define $\mathcal{F}_t := \mathcal{F}^B_{t,T} \vee \mathcal{F}^W_{0,t}$, where $\mathcal{F}^B_{t,T} := \sigma(B_r : r \in [t,T]), \mathcal{F}^W_{0,t} := \sigma(W_r : r \in [0,t])$. Note that $\mathcal{F}$ is not a filtration, since it is neither increasing nor decreasing. In this setting, Pardoux and Peng [55] considered backward doubly stochastic differential equations (BDSDEs). An $\mathcal{F}$-adapted process (Y,Z) is called a solution to the BDSDE

$$Y_t = \xi + \int_t^T f(r,Y_r,Z_r)dr + \int_t^T H(X_r,Y_r) \circ dB_r - \int_t^T Z_r dW_r, \tag{72}$$

if $\mathbb{E}[\sup_{t\le T} |Y_t|^2] < \infty$, $\mathbb{E}[\int_0^T |Z_r|^2 dr] < \infty$ and (Y,Z) satisfies $\mathbb{P}$-a.s. (72) for $t \le T$. Here X is again the semimartingale

$$X_t = x + \int_0^t \sigma_r dW_r + \int_0^t b_r dr.$$

[25] The attentive reader will observe that convergence at $t = 0$ is not immediate, since Theorem 3.6 was not formulated to give comparison at $t = 0$. But we can argue by extending the coefficients as well as the (rough) paths η^n for $t \in [-1,0]$ as

$$\sigma(t,x) := \sigma(0,x), \qquad b(t,x) := b(0,x), \qquad f(t,x,y,z) := f(0,x,y,z), \qquad \eta^n_t := \eta^n_0$$

and considering the PDEs on the interval $[-1,T]$.

Under appropriate (essentially Lipschitz) conditions on f and H they were able to show existence and uniqueness of a solution. [26] The connection to BSDEs with rough driver is given by the following

Theorem 48. *Let $p \in (2,3), \gamma > p$. Let $\xi \in L^\infty(\mathcal{F}_T)$. Let f be a random function satisfying (F1) and (F2). Moreover, assume (A1), (A2), (F1), (F2) and $(H_{p,\gamma})$.*

Then by Theorem 1.1 in [55] there exists a unique solution (Y,Z) to the BDSDE

$$Y_t = \xi + \int_t^T f(r, Y_r, Z_r)dr + \int_t^T H(X_r, Y_r) \circ dB_r - \int_t^T Z_r dW_r.$$

Let $\mathbf{B}_t = \exp(B_t + A_t)$ be the Enhanced Brownian motion (over B) [27], especially $\mathbf{B} \in C_0^{0,p-\mathrm{var}}([0,T], G^2(\mathbb{R}^d))$ $\mathbb{P}^1$ a.s.. By setting $\mathbf{B} = 0$ on a null set, we get $\mathbf{B} \in C_0^{0,p-\mathrm{var}}([0,T], G^2(\mathbb{R}^d))$. By Theorem 36 we can, for every $\omega^1 \in \Omega^1$, construct the solution to the BSDE with rough driver

$$\begin{aligned} Y^{rp}(\omega^1, \cdot)_t = \xi(\cdot) &+ \int_t^T f(r, Y_r^{rp}, Z_r^{rp})dr + \int_t^T H(X_r, Y^{rp}(\omega^1, \cdot))d\mathbf{B}_r(\omega^1) \\ &- \int_t^T Z^{rp}(\omega^1, \cdot)dW_r(\cdot), \quad t \in [0,T]. \end{aligned}$$

We then have for $\mathbb{P}^1 - a.e.\ \omega^1$ that $\mathbb{P}^2 - a.s.$

$$Y_t(\omega^1, \cdot) = Y_t^{rp}(\omega^1, \cdot), \quad t \le T$$

and

$$Z_t(\omega^1, \cdot) = Z_t^{rp}(\omega^1, \cdot), \quad dt \otimes \mathbb{P}^2 a.s..$$

Proof. As in the proof of Theorem 36, in the BDSDE setting, one can transform the integral belonging to the Brownian motion B away. In [7] it was shown, that if we let ϕ be the stochastic (Stratonovich) flow

$$\phi(\omega^1; t, x, y) = y + \int_t^T H(x, \phi(\omega^1; r, x, y)) \circ dB_r(\omega^1),$$

then with $\tilde{Y}_t := \phi^{-1}(t, X_t, Y_t), \tilde{Z}_t := \frac{1}{\partial_y \phi(t, X_t, Y_t)} Z_t$ we have $\mathbb{P}$-a.s.

$$\tilde{Y}_t(\omega^1, \omega^2) = \xi(\omega^2) + \int_t^T \tilde{f}(\omega^1, \omega^2; r, X_r, \tilde{Y}_r(\omega^1, \omega^2), \tilde{Z}_r(\omega^1, \omega^2))dr - \int_t^T \tilde{Z}_r(\omega^1, \omega^2)dW_r(\omega^2), \quad t \le T. \tag{73}$$

Here

$$\begin{aligned} \tilde{f}(\omega^1, \omega^2; t, x, \tilde{y}, \tilde{z}) := \frac{1}{\partial_y \phi} \Big\{ & f\left(\omega^2; t, \phi, \partial_y \phi \tilde{z} + \partial_x \phi \sigma_t\right) + \langle \partial_x \phi, b_t \rangle + \frac{1}{2} \operatorname{Tr}\left[\partial_{xx} \phi \sigma_t \sigma_t^T\right] \\ & + \langle \tilde{z}, (\partial_{xy} \phi \sigma_t)^T \rangle + \frac{1}{2} \partial_{yy} \phi |\tilde{z}|^2 \Big\}, \end{aligned}$$

[26] Pardoux and Peng considered equations, where the Stratonovich integral was actually a backward integral. But if H is smooth enough, the formulations are equivalent. See also Section 4 in [7].

[27] $\mathbf{B}$ is precisely d-dimensional Brownian motion enhanced with its iterated integrals in Stratonovich sense; it is in $1-1$ correspondence with Brownian motion enhanced with Lévy's area; exp denotes the exponential map from the Lie algebra $\mathbb{R}^d \oplus so(d)$ to the group, realized inside the truncated tensor algebra. See e.g. section 13 in [33] for more details.

where ϕ and its derivatives are always evaluated at $(\omega^1; x, \tilde{y})$. Especially, by a Fubini type theorem (e.g. Theorem 3.4.1 in [5]), there exists Ω_0^1 with $\mathbb{P}^1(\Omega_0^1) = 1$ such that for $\omega^1 \in \Omega_0^1$ equation (73) holds true $\mathbb{P}^2$ *a.s.*.

On the other hand we can construct ω^1-wise the rough flow

$$\phi^{rp}(\omega^1; t, x, y) = y + \int_t^T H(x, \phi^{rp}(\omega^1; r, x, y)) d\mathbf{B}_r(\omega^1).$$

Assume for the moment that we have global comparison, so that we can solve the transformed BSDE uniquely, i.e. for every $\omega^1 \in \Omega^1$, we have $\mathbb{P}^2$ *a.s.*

$$\begin{aligned} \tilde{Y}_t^{rp}(\omega^1, \omega^2) = \xi(\omega^2) + &\int_t^T \tilde{f}^{rp}(\omega^1; r, \tilde{Y}_r^{rp}(\omega^1, \omega^2), \tilde{Z}_r^{rp}(\omega^1, \omega^2)) dr \\ &- \int_t^T \tilde{Z}_r^{rp}(\omega^1, \omega^2) dW_r(\omega^2), \quad t \leq T, \end{aligned}$$

where

$$\begin{aligned} \tilde{f}^{rp}(\omega^1, \omega^2; t, x, \tilde{y}, \tilde{z}) := \frac{1}{\partial_y \phi^{rp}} \Big\{ &f\left(\omega^2; t, \phi^{rp}, \partial_y \phi^{rp} \tilde{z} + \partial_x \phi^{rp} \sigma_t\right) + \langle \partial_x \phi^{rp}, b_t \rangle + \frac{1}{2} \operatorname{Tr}\left[\partial_{xx} \phi^{rp} \sigma_t \sigma_t^T\right] \\ &+ \langle \tilde{z}, (\partial_{xy} \phi^{rp} \sigma_t)^T \rangle + \frac{1}{2} \partial_{yy} \phi^{rp} |\tilde{z}|^2 \Big\}, \end{aligned}$$

where ϕ and its derivatives are always evaluated at $(\omega^1; x, \tilde{y})$. It is a classical rough path result, that there exists Ω_1^1 with $\mathbb{P}^1(\Omega_1^1) = 1$ such that for $\omega^1 \in \Omega_1^1$, we have

$$\phi^{rp}(\omega^1; \cdot, \cdot, \cdot) = \phi(\omega^1; \cdot, \cdot, \cdot).$$

Combining above results we have for $\omega^1 \in \Omega_0^1 \cap \Omega_1^1$ that $(\tilde{Y}_t(\omega^1, \cdot), \tilde{Z}_t(\omega^1, \cdot))$ and $(\tilde{Y}_t^{rp}(\omega^1, \cdot), \tilde{Z}_t^{rp}(\omega^1, \cdot))$ satisfy the same BSDE. Hence, we have by uniqueness

$$\tilde{Y}_t(\omega^1, \cdot) = \tilde{Y}_t^{rp}(\omega^1, \cdot), t \leq T, \quad \mathbb{P}^2 \ a.s.$$

and

$$\tilde{Z}_t(\omega^1, \cdot) = \tilde{Z}_t^{rp}(\omega^1, \cdot), \quad dt \otimes \mathbb{P}^2 a.s..$$

By reversing the transformation we get the desired result for Y and Z.

Now, since comparison does *not* necessarily hold globally, we must argue differently. Define $A^k := \{\omega^1 \in \Omega^1 : ||\mathbf{B}(\omega^1)||_{p-\text{var}} \leq k\}$. Then on A^k we have for an $h = h(k) > 0$ comparison on $[T - h, T]$, and we argue on subsequent intervals as above. Now, since $\mathbb{P}(\bigcup_k A^k) = 1$, we get the desired result. □

3.5 Technical results

3.5.1 Comparison for BSDEs

Definition 3.1. Let $\xi \in L^\infty(\mathcal{F}_T)$, W an m-dimensional Brownian motion and f a predictable function on $\Omega \times \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}^m$.

We call an adapted process (Y, Z, C) a *supersolution to the BSDE with data* (ξ, f) if $Y \in H^\infty_{[0,T]}$, $Z \in H^2_{[0,T]}$, C is an adapted right continuous increasing process and

$$Y_t = \xi + \int_t^T f(r, Y_r, Z_r) dr - \int_t^T Z_r dW_r + \int_t^T dC_r, \qquad t \leq T.$$

We call (Y, Z, C) a *subsolution to the BSDE with data* (ξ, f) if $(Y, Z, -C)$ is a supersolution.

The following statement as well as its proof are based on Theorem 2.6 in [43].

Theorem 3.2. *There exists a (universal) strictly positive function* $\delta : \mathbb{R}_+^2 \to (0,\infty)$ *such that the following statement is true.*

Let $(Y^{(1)}, Z^{(1)}, C^{(1)})$ *be a supersolution to the BSDE with data* $(\xi^{(1)}, f^{(1)})$. *Let* $(Y^{(2)}, Z^{(2)}, C^{(2)})$ *be a subsolution to the BSDE with data* $(\xi^{(2)}, f^{(2)})$. *Let* $M \in \mathbb{R}_+$ *be a bound for* $Y^{(1)}$ *and* $Y^{(2)}$, *i.e.*

$$||Y^{(1)}||_\infty, ||Y^{(2)}||_\infty \leq M.$$

Assume that $\mathbb{P}$*-a.s.*

$$f^{(1)}(t, Y_t^{(1)}, Z_t^{(1)}) \leq f^{(2)}(t, Y_t^{(1)}, Z_t^{(1)}), \quad \forall t \in [0,T],$$
$$\xi^{(1)} \leq \xi^{(2)}.$$

Assume that there exist constants $C > 0, L > 0, K > 0$ *such that for* $(t,y,z) \in [0,T] \times [-M,M] \times \mathbb{R}^m$

$$|f^{(2)}(t,y,z)| \leq L + C|z|^2 \quad \mathbb{P}-a.s.,$$
$$|\partial_z f^{(2)}(t,y,z)| \leq K + C|z| \quad \mathbb{P}-a.s.$$

Assume that there exists a constant $N > 0$ *such that for* $(t,y,z) \in [0,T] \times [-M,M] \times \mathbb{R}^m$

$$\partial_y f^{(2)}(t,y,z) \leq N + \delta(C,M)|z|^2 \quad \mathbb{P}-a.s. \tag{74}$$

Then $\mathbb{P}$*-a.s.*

$$Y_t^{(1)} \leq Y_t^{(2)}, \quad 0 \leq t \leq T. \tag{75}$$

Remark 3.3. *We note that, as in Theorem 2.6 of [43], the assumptions could be weakened by replacing the constants* L, K, N *with deterministic functions* $l_t \in L^1(0,T), k_t \in L^2(0,T)$ *and* $n_t \in L^1(0,T)$.

In our application of Theorem 3.2 in the proof of Theorem 36, condition (74) *is not satisfied on* $[0,T]$. *But we are able to choose* $h > 0$ *small enough, such that it is satisfied on* $[T-h,T]$. *Comparison* (75) *then holds on* $[T-h,T]$.

Proof. Let $\lambda > 0, B > 1$ be constants, to be specified later on. We begin by constructing several functions, whose properties we will rely on later in the proof. Define

$$\gamma(\tilde{y}) := \gamma_{\lambda,B}(\tilde{y}) := \frac{1}{\lambda} \log\left(\frac{e^{\lambda B \tilde{y}} + 1}{B}\right) - M, \quad \tilde{y} \in \mathbb{R}.$$

Then

$$\gamma^{-1}(y) = \frac{1}{\lambda B} \log\left(B e^{\lambda(y+M)} - 1\right), \quad \gamma'(\tilde{y}) = B \frac{1}{1 + e^{-\lambda B \tilde{y}}}.$$

Denote $g(y) := e^{-\lambda(y+M)}$, then $0 < g \leq 1$, on $[-M,M]$. Define

$$w(y) := \gamma'(\gamma^{-1}(y)) = B - g(y).$$

Then

$$w'(y) = \lambda g(y), \quad w''(y) = -\lambda^2 g(y),$$
$$\frac{w'(y)}{w(y)} = \frac{\lambda g(y)}{B - g(y)}, \quad \frac{w''(y)}{w(y)} = \frac{-\lambda^2 g(y)}{B - g(y)}.$$

In particular $w > 0$ on $[-M, M]$.

Define $\alpha(y) := \gamma^{-1}(y)$. Then, since $(Y^{(1)}, Z^{(1)}, C^{(1)})$ is a supersolution to the BSDE with data $(\xi^{(1)}, f^{(1)})$, Itô formula gives

$$\alpha(Y_t^{(1)}) = \alpha(Y_0^{(1)}) - \int_0^t \alpha'(Y_r^{(1)}) f^{(1)}(r, Y_r^{(1)}, Z_r^{(1)}) dr + \int_0^t \alpha'(Y_r^{(1)}) Z_r^{(1)} dW_r$$
$$- \int_0^t \alpha'(Y_r^{(1)}) dC_r + \int_0^t \alpha''(Y_r^{(1)}) |Z_r^{(1)}|^2 dr.$$

Define

$$\tilde{Y^{(1)}} := \alpha(Y^{(1)}), \quad \tilde{Z^{(1)}} := \frac{Z^{(1)}}{\gamma'(\tilde{Y}^{(1)})} = \frac{Z^{(1)}}{w(Y^{(1)})}.$$

and

$$F^{(1)}(t, \tilde{y}, \tilde{z}) := \frac{1}{\gamma'(\tilde{y})} \left[f^{(1)}(t, \gamma(\tilde{y}), \gamma'(\tilde{y})\tilde{z}) + \frac{1}{2}\gamma''(\tilde{y})|\tilde{z}|^2 \right].$$

Since $\alpha' > 0$ we have that $(\tilde{Y}^{(1)}, \tilde{Z}^{(1)}, \int_0^\cdot \alpha'(Y_r^{(1)}) dC_r^{(1)})$ is a supersolution to the BSDE with data $(\alpha(\xi^{(1)}), F^{(1)})$. Analogously we have that $(\tilde{Y}^{(2)}, \tilde{Z}^{(2)}, \int_0^\cdot \alpha'(Y_r^{(2)}) dC_r^{(2)})$ is a subsolution to the BSDE with data $(\alpha(\xi^{(2)}), F^{(2)})$. Since α is increasing, it is now enough to verify that $\tilde{Y}^{(1)} \leq \tilde{Y}^{(2)}$.

For that we will verify, that $F^{(2)}$ satisfies the conditions of Proposition 2.9 in [43]. Especially we will show, that there exist constants $A, G > 0$ such that

$$\partial_y f^{(2)}(t, y, z) + A|\partial_z f^{(2)}(t, y, z)|^2 \leq G, \quad \forall (t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^m. \tag{76}$$

For simplicity denote $F := F^{(2)}, f := f^{(2)}$. Denote $y = \gamma(\tilde{y}),\ z = \gamma'(\tilde{y})\tilde{z} = w(y)\tilde{z}$. For convenience w and its derivatives will always be evaluated at y. Then

$$\partial_{\tilde{z}} F(t, \tilde{y}, \tilde{z}) = \partial_z f(t, y, z) + z\frac{w'}{w},$$
$$\partial_{\tilde{y}} F(t, \tilde{y}, \tilde{z}) = \frac{1}{w}\left[\frac{1}{2}w''|z|^2 + w'\left(\partial_z f(t, y, z) z - f(t, y, z)\right)\right] + \partial_y f(t, y, z).$$

Hence

$$\partial_{\tilde{y}} F(t, \tilde{y}, \tilde{z}) \leq \frac{1}{w}\left[\frac{1}{2}w''|z|^2 + w'\big(|z|[K + C|z|] + L + C|z|^2\big)\right] + \partial_y f(t, y, z)$$

and

$$|\partial_{\tilde{z}} F(t, \tilde{y}, \tilde{z})|^2 \leq \left[K + C|z| + \frac{w'}{w}|z|\right]^2.$$

So, for $A > 0$

$$(\partial_{\tilde{y}} F + A|\partial_{\tilde{z}} F|^2)(t,\tilde{y},\tilde{z}) \leq |z|^2 \left[\frac{1}{2}\frac{w''}{w} + \frac{w'}{w}2C + A(C + \frac{w'}{w})^2\right] + K|z|\left[\frac{w'}{w} + 2A\left(C + \frac{w'}{w}\right)\right] + \frac{w'}{w}L + \partial_y f(t,y,z) + AK^2.$$

Note, that for the second term we have

$$\begin{aligned} K|z|\left[\frac{w'}{w} + 2A\left(C + \frac{w'}{w}\right)\right] &\leq K|z|\left[(1+2A)\left(C + \frac{w'}{w}\right)\right] \\ &\leq A\left(C + \frac{w'}{w}\right)^2 |z|^2 + \frac{(1+2A)^2}{A}K^2. \end{aligned}$$

Hence

$$\begin{aligned} (\partial_{\tilde{y}} F + A|\partial_{\tilde{z}} F|^2)(t,\tilde{y},\tilde{z}) &\leq |z|^2 \left[\frac{1}{2}\frac{w''}{w} + \frac{w'}{w}2C + 2A(C + \frac{w'}{w})^2\right] \\ &\quad + \frac{w'}{w}L + \partial_y f(t,y,z) + \left(A + \frac{(1+2A)^2}{A}\right)K^2. \end{aligned}$$

Now

$$\begin{aligned} \frac{1}{2}\frac{w''}{w} + \frac{w'}{w}2C + 2A(C + \frac{w'}{w})^2 &= \frac{1}{2}\frac{w''}{w} + \frac{w'}{w}2C + 2AC^2 + 4AC\frac{w'}{w} + 2A(\frac{w'}{w})^2 \\ &= -\frac{\lambda^2}{2}\frac{g(y)}{B - g(y)} + 2C(1+2A)\frac{\lambda g(y)}{B - g(y)} + 2AC^2 + 2A\frac{\lambda^2 g(y)^2}{(B - g(y))^2} \\ &= \frac{g(y)}{(B - g(y))^2}\Big[-\frac{\lambda^2}{2}(B - g(y)) + 2C(1+2A)\lambda(B - g(y)) \\ &\qquad + 2A\lambda^2 g(y)\Big] + 2AC^2 \\ &= \frac{g(y)}{(B - g(y))^2}\left[\frac{\lambda^2}{2}((1+4A)g(y) - B) + 2C(1+2A)\lambda(B - g(y))\right] + 2AC^2. \end{aligned}$$

For all $A < 1$ we hence have

$$\frac{1}{2}\frac{w''}{w} + \frac{w'}{w}2C + 2A(C + \frac{w'}{w})^2 \leq \frac{g(y)}{(B - g(y))^2}\left[\frac{\lambda^2}{2}(5g(y) - B) + 2C3\lambda(B - g(y))\right] + 2AC^2.$$

Now, choose $B = 6$. Hence $5g(y) - B \leq -1,\ y \in [-M, M]$. Then choose $\lambda = \lambda(C)$ sufficiently large such that the term in square brackets is strictly negative, say smaller then -1 for all $y \in [-M, M]$. This is possible since it is a polynomial in λ and the leading power has a negative coefficient. Then for $y \in [-M, M]$

$$\begin{aligned} \frac{g(y)}{(6 - g(y))^2}\left[\frac{\lambda^2}{2}(5g(y) - 6) + 2C3\lambda(6 - g(y))\right] &\leq -\frac{g(y)}{(6 - g(y))^2} \\ &\leq -\frac{1}{36}e^{-\lambda 2M} =: -2\delta, \end{aligned}$$

where δ depends only M and λ and hence only on M and C, i.e.

$$\delta = \delta(C, M) = \frac{1}{72}e^{-\lambda(C)2M}.$$

Now choose $A \in (0,1)$ small enough such that $2AC^2 < \delta$. If then for some $N > 0$ we have

$$\partial_y f(t,y,z) \leq N + \delta(C,M)|z|^2,$$

it follows that

$$\begin{aligned}(\partial_{\tilde{y}}F + A|\partial_{\tilde{z}}F|^2)(t,\tilde{y},\tilde{z}) &\leq \frac{w'}{w}L + N + \left(A + \frac{(1+2A)^2}{A}\right)K^2 \\ &\leq \frac{\lambda}{B-1}L + N + \left(A + \frac{(1+2A)^2}{A}\right)K^2 \\ &=: G.\end{aligned}$$

So we have shown (76) and comparison then follows from Proposition 2.9 in [43]. □

3.5.2 Properties of rough flows

Consider the solution flow ϕ to

$$\phi(t,x,y) = y + \int_t^T H(x,\phi(r,x,y))d\eta_r, \tag{77}$$

where H and η will be specified in a moment. We need to control

$$\partial_y\phi - 1, \partial_x\phi, \partial_{xx}\phi, \partial_{xy}\phi, \partial_{yy}\phi, \partial_{yyy}\phi, \partial_{xyy}\phi, \partial_{xxy}\phi$$

over a small interval $[T-h,T]$. Note that each of the above expressions is 0 when evaluated at $t = T$.

Lemma 3.4. *Let $p \geq 1$, $\eta \in C^{0,p-\mathrm{var}}([0,T], G^{[p]}(\mathbb{R}^d))$ and $\gamma > p$. Assume that $H_i = H_i(x,y)$ has joint regularity of the form*

$$\sup_{i=1,\ldots,d} |H_i(\cdot,\cdot)|_{\mathrm{Lip}^{\gamma+2}(R^{n+1})} \leq c_1$$

and

$$\|\zeta\|_{p\text{-}var;[0,T]} \leq c_2.$$

Then, the solution to (77) induces a flow of C^3 diffeomorphisms, parametrized by $x \in \mathbb{R}^n$, and there exists a positive $L = L(c_1,c_2,T)$ so that, uniformly over $x \in R^n, y \in R$ and $t \in [0,T]$

$$\max\left\{\partial_x\phi, \partial_y\phi, \frac{1}{\partial_y\phi}, \partial_{xx}\phi, \partial_{xy}\phi, \partial_{yy}\phi, \partial_{yyy}\phi, \partial_{xyy}\phi, \partial_{xxy}\phi\right\} < L.$$

Moreover, for every $\varepsilon > 0$ there exists a positive $\delta = \delta(\varepsilon, c_1, c_2)$ so that, uniformly over $x \in R^n, y \in R$ and $t \in [T-\delta, T]$

$$\max\{\partial_x\phi, \partial_y\phi - 1, \partial_{xx}\phi, \partial_{xy}\phi, \partial_{yy}\phi, \partial_{yyy}\phi, \partial_{xyy}\phi, \partial_{xxy}\phi\} < \varepsilon.$$

Proof. Consider the extended RDE

$$\begin{aligned} d\xi &= 0 \\ -d\phi &= H(\xi,\phi)\,d\zeta \end{aligned}$$

with terminal data $(\xi_T, \phi_T) = (x, y)$. The assumption on (H_i) implies that (ξ, ϕ) evolves according to a rough differential equation with $\mathrm{Lip}^{\gamma+2}$-vector fields. In this case, the ensemble

$$\hat{\phi} = (\xi, \phi, \partial_x \phi, \partial_y \phi, \partial_{xx}\phi, \partial_{xy}\phi, \partial_{yy}\phi, \partial_{yyy}\phi, \partial_{xyy}\phi, \partial_{xxy}\phi)$$

can be seen to be the (unique[28], non-explosive) solution to an RDE along $\mathrm{Lip}^{\gamma-1}_{loc}$ vector fields. Thanks to non-explosivity we can, for fixed terminal data

$$\hat{\phi}_T = (x, y, 0, 1, 0, 0, 0, 0, 0, 0)\,,$$

localize the problem and assume without loss of generality that the above ensemble is driven along $\mathrm{Lip}^{\gamma-1}$ vector fields. Since we want estimates that are *uniform* in x, y we make another key observations: there is no loss of generality in taking $(x, y) = (0, 0)$ provided H is replaced by $H_{x,y} = H(x + \cdot, y + \cdot)$. This also shifts the derivatives (evaluated at some (x, y)) to derivatives evaluated at $(0, 0)$. As announced, we can now safely localize, and assume that the vector fields required for $\hat{\phi}$, obtain by taking formal (x, y) derivatives in

$$\begin{aligned} d\xi &= 0 \\ -d\phi &= H(\xi, \phi)\, d\zeta, \end{aligned}$$

are globally $\mathrm{Lip}^{\gamma-1}$. A basic estimate (Theorem 10.14 in [33]) for RDE solutions implies that for some $C = C(p, \gamma)$

$$\left|\hat{\phi}_t - \hat{\phi}_T\right| \leq \left|\hat{\phi}\right|_{p\text{-var};[t,T]} = C \times \varphi_p\left(|H_{x,y}|_{\mathrm{Lip}^{\gamma+2}} \|\zeta\|_{p\text{-var};[T-h,T]}\right),$$

where $\varphi_p(x) = \max(x, x^p)$.At last, we note that $|H_{x,y}|_{\mathrm{Lip}^{\gamma+2}} = |H|_{\mathrm{Lip}^{\gamma+2}}$ thanks to invariance of such Lip norms under translation. The proof is then easily finished. □

Lemma 3.5. *Assume the setting of the previous lemma. Assume that $\eta^n, n \geq 1$ is a sequence of p rough paths that converge to a rough path η^0 in p-variation.*

Then locally uniformly on $[0, T] \times \mathbb{R}^n \times \mathbb{R}$

$$(\phi^n, \frac{1}{\partial_y \phi^n}, \partial_y \phi^n, \partial_{yy}\phi^n, \partial_x \phi^n, \partial_{xx}\phi^n, \partial_{yx}\phi^n) \to (\phi^0, \frac{1}{\partial_y \phi^0}, \partial_y \phi^0, \partial_{yy}\phi^0, \partial_x \phi^0, \partial_{xx}\phi^0, \partial_{yx}\phi^0)$$

Proof. Using enlargment of the state space as in the proof of Lemma 3.4 we can apply the same reasoning as in Theorem 11.14 and Theorem 11.15 in [33] to get the desired result. □

3.5.3 Comparison for PDEs

We consider the equation

$$-\partial_t u - \frac{1}{2}\mathrm{Tr}[\sigma(t,x)\sigma(t,x)^T D^2 u] - \langle b(t,x), Du\rangle - f(t, x, u, Du\sigma(t,x)) = 0, \quad (t,x) \in (0,T) \times \mathbb{R}^n, \tag{78}$$

where $f : [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$, $b : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times m}$, are a continuous functions.

The following statement as well as its proof are a modification of Theorem 3.2 in [43]. (The statement is not in its most general form, but adjusted to what we need in the main text.)

[28] This is actual a subtle point since uniqueness in general requires $\mathrm{Lip}^{\gamma}_{loc}$-regularity. The point is that the RDEs obtained by differentiating the flow have a special structure so that for the final level of derivatives only rough integration is need; as is well known, for this it suffices to have $\mathrm{Lip}^{\gamma-1}_{loc}$ regularity. Chapter 11 in [33] contains a detailed discussion of this.

Theorem 3.6. *Assume that there exists a constant $C_b > 0$ such that for $(t,x),(t,y) \in [0,T] \times \mathbb{R}^n$*

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq C_b|x-y|,$$
$$|b(t,x)| \leq C_b.$$

Assume that there exists a constant $C_\sigma > 0$ such that for $(t,x),(t,y) \in [0,T] \times \mathbb{R}^n$

$$|\sigma(t,x) - \sigma(t,y)| \leq C_\sigma|x-y|,$$
$$|\sigma(t,x)| \leq C_\sigma.$$

Assume that there exists a constant $C_{1,f} > 0$ such that for $(t,x,y,z) \in [0,T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m$

$$|f(t,x,y,z)| \leq C_{1,f}(1+|z|^2),$$
$$|\partial_z f(t,x,y,z)| \leq C_{1,f}(1+|z|).$$

Assume that there exists a constant $C_{2,f}$ such that for every $\varepsilon > 0$ there exists an $h_\varepsilon \in (0,T]$ such that for $(t,x,y,z) \in [T-h_\varepsilon, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m$ we have

$$\partial_y f(t,x,y,z) \leq C_{2,f} + \varepsilon|z|^2. \tag{79}$$

Assume that there exists a constant $C_{3,f} > 0$ such that for $(t,x,y,z) \in [0,T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m$

$$|\partial_x f(t,x,y,z)| \leq C_{3,f}(1+|z|^2).$$

Let u, v be a bounded semicontinuous sub-(resp. super-)solution to (78) *on $(0,T) \times \mathbb{R}^n$, with $u(T,\cdot) \leq v(T,\cdot)$. Then there exists an $\varepsilon^* = \varepsilon^*(||u||_\infty \vee ||v||_\infty, C, C_{2,f}) > 0$ such that for $(t,x) \in (T-h_{\varepsilon^*}, T] \times \mathbb{R}^n$ we have*

$$u(t,x) \leq v(t,x).$$

Proof. 1. Reduction

We will first transform the PDE into a PDE with coefficient that satisfies a certain structure condition (equation (24) in [43]). Set $M := \max\{||u||_\infty, ||v||_\infty\} + 1$.

Let $\lambda > 0, A > 1, K > 0$ be constants to be chosen later. We begin by constructing several functions, whose properties we will rely on later in the proof. Define

$$\varphi(\tilde{y}) := \frac{1}{\lambda} \ln\left(\frac{e^{\lambda A \tilde{y}} + 1}{A}\right) : \mathbb{R} \to (-\frac{\ln(A)}{\lambda}, \infty).$$

We will have to choose $A \geq e^{\lambda 2 M e^{Kt}}$, since we shall need later on that $\{e^{Kt}(y-M) : y \in [-M,M]\}$ is contained in the range of φ. Then

$$\varphi'(\tilde{y}) = A\frac{1}{1+e^{-\lambda A \tilde{y}}}, \quad \varphi^{-1}(y) = \frac{1}{\lambda A} \ln\left(Ae^{\lambda y} - 1\right).$$

Define $r(y) := \varphi^{-1}(e^{Kt}(y-M))$, its inverse $s(\tilde{y}) := \varphi(\tilde{y})e^{-Kt} + M$ and $g(y) := e^{-\lambda e^{Kt}(y-M)}$: $[-M,M] \to [1, e^{\lambda 2 M e^{Kt}}]$. Then $g'(y) = -\lambda e^{Kt} g(y)$. Define

$$w(y) := e^{-Kt}\varphi'(r(y)) = \partial_{\tilde{y}} s|_{\tilde{y}=r(y)} = e^{-Kt}\left[A - e^{-\lambda e^{Kt}(y-M)}\right] = e^{-Kt}\left[A - g(y)\right],$$

which is non-negative for $A \geq e^{\lambda 2 M e^{Kt}}$. Then

$$w'(y) = \lambda g(y), \quad w''(y) = -e^{Kt}\lambda^2 g(y).$$

Let now $u(t,x)$ be a solution to (78). Let $\tilde{u}(t,x) := r(u(t,x))$. Then $u(t,x) = s(\tilde{u}(t,x))$, and hence

$$\begin{aligned} \partial_{x_i} u(t,x) &= \varphi'(\tilde{u}(t,x))e^{-Kt}\partial_{x_i}\tilde{u}(t,x), \\ \partial_{x_j x_i} u(t,x) &= \varphi''(\tilde{u}(t,x))e^{-Kt}\partial_{x_j}\tilde{u}(t,x)\partial_{x_i}\tilde{u}(t,x) + \varphi'(\tilde{u}(t,x))e^{-Kt}\partial_{x_j x_i}\tilde{u}(t,x), \end{aligned}$$

i.e.

$$\begin{aligned} Du(t,x) &= \varphi'(\tilde{u}(t,x))e^{-Kt}D\tilde{u}(t,x), \\ D^2 u(t,x) &= \varphi''(\tilde{u}(t,x))e^{-Kt}D\tilde{u}(t,x) \otimes D\tilde{u}(t,x) + \varphi'(\tilde{u}(t,x))e^{-Kt}D^2\tilde{u}(t,x). \end{aligned}$$

Hence

$$\begin{aligned} \partial_t \tilde{u}(t,x) &= \frac{1}{\varphi'(\tilde{u}(t,x))}\left[Ke^{Kt}(u(t,x)-M) + e^{Kt}\partial_t u(t,x)\right] \\ &= \frac{1}{\varphi'(\tilde{u}(t,x))}Ke^{Kt}(u(t,x)-M) \\ &\quad - \frac{1}{\varphi'(\tilde{u}(t,x))}e^{Kt}\left[\frac{1}{2}\operatorname{Tr}[\sigma(t,x)\sigma(t,x)^T D^2 u(t,x)] + \langle b(t,x), Du(t,x)\rangle\right] \\ &\quad - \frac{1}{\varphi'(\tilde{u}(t,x))}e^{Kt}f(t,x,u(t,x),Du(t,x)\sigma(t,x)) \\ &= K\frac{\varphi(\tilde{u}(t,x))}{\varphi'(\tilde{u}(t,x))} \\ &\quad - \frac{1}{2}\operatorname{Tr}[\sigma(t,x)\sigma(t,x)^T D^2\tilde{u}(t,x)] - \frac{\varphi''(\tilde{u}(t,x))}{\varphi'(\tilde{u}(t,x))}\frac{1}{2}\operatorname{Tr}[\sigma(t,x)\sigma(t,x)^T D\tilde{u}(t,x)\otimes D\tilde{u}(t,x)] \\ &\quad - \langle b(t,x), D\tilde{u}(t,x)\rangle - \frac{1}{\varphi'(\tilde{u}(t,x))}e^{Kt}f(t,x,s(\tilde{u}(t,x)),\varphi'(\tilde{u}(t,x))e^{-Kt}D\tilde{u}(t,x)\sigma(t,x)). \end{aligned}$$

Analogously, by resorting to test functions, one shows, that if u (resp. v) is a viscosity sub-(resp. super-) solution to (78), then $\tilde{u}(t,x) := r(u(t,x))$ (resp. $\tilde{v}(t,x) := r(v(t,x))$) is a viscosity sub-(resp. super-) solution to

$$\begin{aligned} &-\partial_t\tilde{u}(t,x) - \frac{1}{2}\operatorname{Tr}[\sigma(t,x)\sigma(t,x)^T D^2\tilde{u}(t,x)] - \langle(b(t,x), D\tilde{u}(t,x)\rangle \\ &\qquad - \tilde{f}(t,x,\tilde{u}(t,x),D\tilde{u}(t,x)\sigma(t,x)) = 0,\ t\in(0,T), x\in\mathbb{R}^n, \end{aligned} \tag{80}$$

where, denoting from now on $y = s(\tilde{y})$, $z = w(y)\tilde{z}$,

$$\begin{aligned} \tilde{f}(t,x,\tilde{y},\tilde{z}) &= -K\frac{\varphi(\tilde{y})}{\varphi'(\tilde{y})} + \frac{\varphi''(\tilde{y})}{\varphi'(\tilde{y})}\frac{1}{2}|\tilde{z}|^2 \\ &\quad + \frac{1}{\varphi'(\tilde{y})}e^{Kt}f(t,x,s(\tilde{y}),\varphi'(\tilde{y})e^{-Kt}\tilde{z}) \\ &= -K\frac{y-M}{w(y)} + w'(y)\frac{1}{2}|\tilde{z}|^2 + \frac{1}{w(y)}f(t,x,y,w(y)\tilde{z}). \end{aligned}$$

We also obviously have $\tilde{u}(T,\cdot) \leq \tilde{v}(T,\cdot)$.

We will bound the $\tilde{y}$-derivative of $\tilde{f}$, while at the same time choosing the constants K, λ, A. First

$$\begin{aligned}
\partial_{\tilde{y}}\tilde{f}(t,x,\tilde{y},\tilde{z}) &= -K(1-(y-M)\frac{w'(y)}{w(y)})) + \frac{1}{2}\frac{w''(y)}{w(y)}|z|^2 \\
&\quad - \frac{w'(y)}{w(y)}f(t,x,y,z) + \partial_y f(t,x,y,z) \\
&\quad + \frac{w'(y)}{w(y)}\partial_z f(t,x,y,z)z \\
&\leq -K(1-(y-M)\frac{w'(y)}{w(y)})) + \frac{1}{2}\frac{w''(y)}{w(y)}|z|^2 \\
&\quad + \frac{w'(y)}{w(y)}C_{1,f}(1+|z|^2) + \partial_y f(t,x,y,z) \\
&\quad + \frac{w'(y)}{w(y)}C_{1,f}(1+|z|)|z| \\
&\leq \frac{|z|^2}{w(y)}\left(\frac{1}{2}w''(y) + C_{1,f}w'(y) + C_{1,f}w'(y)\right) \\
&\quad - K(1-(y-M)\frac{w'(y)}{w(y)}) + \partial_y f(t,x,y,z) + C_{1,f}\frac{w'(y)}{w(y)} + C_{1,f}\frac{w'(y)}{w(y)}|z|.
\end{aligned}$$

Now using

$$C_{1,f}\frac{w'(y)}{w(y)}|z| \leq \frac{|z|^2}{w(y)}w'(y) + \frac{w'(y)}{w(y)}C_{1,f}^2,$$

we get

$$\begin{aligned}
\partial_{\tilde{y}}\tilde{f}(t,x,\tilde{y},\tilde{z}) \leq \frac{|z|^2}{w(y)}&\left(\frac{1}{2}w''(y) + (2C_{1,f}+1)w'(y)\right) \\
&- K + \partial_y f(t,x,y,z) + \frac{w'(y)}{w(y)}\left(C_{1,f} + K(y-M) + C_{1,f}^2\right).
\end{aligned} \tag{81}$$

Note that

$$C_{1,f} + K(y-M) + C_{1,f}^2 \leq C_{1,f} - K + C_{1,f}^2, \quad y \in [-(M-1), M-1].$$

Hence we can choose $K_0 = K_0(C_{1,f})$ sufficiently large, such that

$$C_{1,f} + K_0(y-M) + C_{1,f}^2 \leq -1, \quad y \in [-(M-1), M-1].$$

Then we have that for all choices of $K_0 > K$, and all choices $\lambda > 0$ that the last term in (81),

$$\frac{w'(y)}{w(y)}\left(C_{1,f} + K(y-M) + C_{1,f}^2\right) = \frac{e^{-\lambda e^{Kt}(y-M)}}{A - e^{-\lambda e^{Kt}(y-M)}}\lambda e^{Kt}\left(C_{1,f} + K(y-M) + C_{1,f}^2\right),$$

is negative for $y \in [-(M-1), M-1]$ as long as $A > e^{\lambda 2Me^{Kt}}$. We now fix $K = K(C_{1,f}, C_{2,f}) = \max\{K_0(C_{1,f}), C_{2,f}\} + 1$. Then

$$\begin{aligned}
\frac{1}{2}w''(y) + (2C_{1,f}+1)w'(y) &= -\frac{1}{2}e^{Kt}\lambda^2 g(y) + \lambda(2C_{1,f}+1)g(y) \\
&\leq -\frac{1}{2}\lambda^2 g(y) + \lambda(2C_{1,f}+1)g(y) \\
&= g(y)\lambda\left[(2C_{1,f}+1) - \frac{1}{2}\lambda\right].
\end{aligned}$$

So, if we choose $\lambda = \lambda(C_{1,f}) = 4C_{1,f} + 4$, we have

$$\frac{1}{2}w''(y) + (2C_{1,f} + 1)w'(y) \leq g(y)(4C_{1,f} + 4)(-1) \leq -(4C_{1,f} + 4) \leq -1.$$

We now fix $A = A(\lambda(C_{1,f}), M, K(C_{1,f}, C_{2,f})) = A(M, C_{1,f}, C_{2,f}) = e^{\lambda 2Me^{KT}} + 1$. Then for the first term in (81)

$$\begin{aligned}
\frac{|z|^2}{w(y)}\left(\frac{1}{2}w''(y) + (2C_{1,f} + 1)w'(y)\right) &= \frac{|z|^2}{e^{\lambda 2Me^{KT}} + 1 - e^{-\lambda e^{Kt}(y-M)}}e^{Kt}\left(\frac{1}{2}w''(y) + (2C_{1,f} + 1)w'(y)\right) \\
&\leq -\frac{|z|^2}{e^{\lambda 2Me^{KT}} + 1 - e^{-\lambda e^{Kt}(y-M)}}e^{Kt} \\
&\leq -\frac{|z|^2}{e^{\lambda 2Me^{KT}}}e^{Kt} \\
&< -\delta|z|^2 < 0,
\end{aligned}$$

with

$$\delta = \delta(\lambda(C_{1,f}), K(C_{1,f}, C_{2,f}), M) = \delta(M, C_{1,f}, C_{2,f}) = \frac{e^{Kt}}{e^{\lambda 2Me^{KT}} + 1} > 0.$$

We now set $\varepsilon^* = \varepsilon^*(M, C_{2,f}, C_{2,f}) := \frac{\delta}{2}$. Then on $[T - h_{\varepsilon^*}, T]$ we have

$$\partial_y f(t, x, y, z) \leq C_{2,f} + \frac{\delta}{2}|z|^2,$$

and hence we get that on $[T - h_{\varepsilon^*}, T]$ (remember that $K \geq C_{2,f} + 1$)

$$\begin{aligned}
\partial_{\tilde{y}}\tilde{f}(t, x, \tilde{y}, \tilde{z}) \leq{}& \frac{|z|^2}{w(y)}\left(\frac{1}{2}w''(y) + (2C_{1,f} + 1)w'(y)\right) \\
&- K + \partial_y f(t, x, y, z) + \frac{w'(y)}{w(y)}\left(C_{1,f} + K(y - M) + C_{1,f}^2\right) \\
\leq{}& -\delta|z|^2 - K + C_{2,f} + \frac{\delta}{2}|z|^2 \\
\leq{}& -\delta|z|^2 + \frac{\delta}{2}|z|^2 - 1 \\
={}& -\frac{\delta}{2}|z|^2 - 1 \\
={}& -\frac{\delta}{2}|w(y)|^2|\tilde{z}|^2 - 1 \\
\leq{}& -\tilde{K}(1 + |\tilde{z}|^2), \qquad \text{for } y \in [-(M-1), M-1].
\end{aligned}$$

for some $\tilde{K} > 0$.

Moreover by the definition of $\tilde{f}$ and the assumptions on f it is straightforward to bound the other partial derivatives of $\tilde{f}$. So in total we get $\tilde{K}, \tilde{C} > 0$ such that for $t \in [T - h_{\varepsilon^*}, T], y \in [-(M-1), M-1], \tilde{z} \in \mathbb{R}^m$

$$\begin{aligned}
\partial_{\tilde{y}}\tilde{f}(t, x, \tilde{y}, \tilde{z}) &\leq -\tilde{K}(1 + |\tilde{z}|^2), \\
|\partial_x \tilde{f}(t, x, \tilde{y}, \tilde{z})| &\leq \tilde{C}(1 + |\tilde{z}|^2), \\
|\partial_{\tilde{z}}\tilde{f}(t, x, \tilde{y}, \tilde{z})| &\leq \tilde{C}(1 + |\tilde{z}|).
\end{aligned} \tag{82}$$

Let $\underline{M} := r(-(M-1))$, $\bar{M} := r(M-1)$. Then, since u, v take values in $[-(M-1), M-1]$, $\tilde{u}, \tilde{v}$ take values in $[\underline{M}, \bar{M}]$. We can then define

$$\tilde{\tilde{f}}(t,x,\tilde{y},\tilde{z}) := \begin{cases} \tilde{f}(t,x,\underline{M},\tilde{z}) - \tilde{K}(1+|\tilde{z}|^2)(\tilde{y}-\underline{M}) & ,\tilde{y} < \underline{M}, \\ \tilde{f}(t,x,\tilde{y},\tilde{z}) & ,\tilde{y} \in [\underline{M},\bar{M}], \\ \tilde{f}(t,x,\bar{M},\tilde{z}) - \tilde{K}(1+|\tilde{z}|^2)(\tilde{y}-\bar{M}) & ,\bar{M} < \tilde{y}. \end{cases}$$

This function $\tilde{\tilde{f}}$ then satisfies [29] for some $\tilde{K}, \tilde{C} > 0$ and for all $t \in [T - h_{\varepsilon^*}, T], \tilde{y} \in \mathbb{R}, \tilde{z} \in \mathbb{R}^m$

$$\begin{aligned} \partial_{\tilde{y}} \tilde{\tilde{f}}(t,x,\tilde{y},\tilde{z}) &\leq -\tilde{K}(1+|\tilde{z}|^2), \\ |\partial_x \tilde{\tilde{f}}(t,x,\tilde{y},\tilde{z})| &\leq \tilde{C}(1+|\tilde{z}|^2), \\ |\partial_{\tilde{z}} \tilde{\tilde{f}}(t,x,\tilde{y},\tilde{z})| &\leq \tilde{C}(1+|\tilde{z}|+|\tilde{y}||\tilde{z}|). \end{aligned} \tag{83}$$

and $\tilde{u}, \tilde{v}$ are also sub-(resp. super-) solution to (80) with $\tilde{f}$ replaced by $\tilde{\tilde{f}}$. [30] We can hence assume the validity of (83) for $\tilde{f}$.

2. Comparison under structure condition

Let $\tilde{u}$, $\tilde{v}$ be a semicontinuous sub-(resp. super-)solution to

$$\begin{aligned} &-\partial_t \tilde{u}(t,x) - \frac{1}{2}\operatorname{Tr}[\sigma(t,x)\sigma(t,x)^T D^2\tilde{u}(t,x)] - \langle (b(t,x), D\tilde{u}(t,x)\rangle \\ &\quad - \tilde{f}(t,x,\tilde{u}(t,x), D\tilde{u}(t,x)\sigma(t,x)) = 0, \quad t \in (0,T), x \in \mathbb{R}^n \end{aligned}$$

where $\tilde{f}$ satisfies (83) for $t \in [0,T], \tilde{y} \in \mathbb{R}, \tilde{z} \in \mathbb{R}^m$. Let $\tilde{u}$ be bounded above and $\tilde{v}$ be bounded. Assume $\tilde{u}(T,\cdot) \leq \tilde{v}(T,\cdot)$. We show $\tilde{u} \leq \tilde{v}$ on $(0,T] \times \mathbb{R}^n$.

First of all, note that $\tilde{u}_\gamma(t,x) := \tilde{u}(t,x) - \frac{\gamma}{t}$ is also a subsolution. Since $\tilde{u} \leq \tilde{v}$ follows from $\tilde{u}_\gamma \leq \tilde{v}$ in the limit $\gamma \to 0$, it suffices to prove comparison under the additional assumption

$$\lim_{t\to 0} \tilde{u}(t,x) = -\infty, \qquad \text{uniformly on } \mathbb{R}^n.$$

Define

$$L := \sup_{x\in\mathbb{R}^n, t\in(0,T]} [\tilde{u}(t,x) - \tilde{v}(t,x)]$$

and also

$$\begin{aligned} L(h) &:= \sup_{|x-x'|\leq h, t\in(0,T]} \left[\tilde{u}(t,x) - \tilde{v}(t,x')\right], \\ L' &:= \lim_{h\to 0} L(h). \end{aligned}$$

[29] Note that $\tilde{\tilde{f}}$ is not necessarily continuously differentiable in $\tilde{y}$ anymore, but, as was noted on page 48, we can directly work with functions that are only (locally) Lipschitz and bound the corresponding Lipschitz constants.

[30] The reason one wants bounds globally in y is that is that the proof involves $\tilde{u}_\gamma = \tilde{u} - \gamma/t$ which is unbounded.

In fact, it is possible to carry out the comparison proof without penalizing $t = 0$. It suffices to use the slightly more general version of the parabolic theorem of sums in Theorem 50. Following this approach also leads to comparison at $t = 0$, if we take into consideration the remarks in [14] on the accessibility of a subsolution.

One has of course $L \leq L'$. We will show $L' \leq 0$. Consider

$$\psi_{\varepsilon,\eta}(t,x,x') := \tilde{u}(t,x) - \tilde{v}(t,x') - \frac{|x-x'|^2}{\varepsilon^2} - \eta(|x|^2 + |x'|^2).$$

Let $L_{\varepsilon,\eta}$ be the maximum of $\psi_{\varepsilon,\eta}$ and $(\hat{t},\hat{x},\hat{x'}) = (\hat{t}_{\varepsilon,\eta},\hat{x}_{\varepsilon,\eta},\hat{x'}_{\varepsilon,\eta}) \in (0,T] \times \mathbb{R}^n$ a maximizing point, which exists by the assumptions on $\tilde{u}$ and $\tilde{v}$.

We argue by contradiction. Hence assume $\tilde{u}(s,z) - \tilde{v}(s,z) > \delta$ for some (s,z). Then also $L' > \delta$. We first argue, that for small enough values of ε,η the optimizing time parameter $\hat{t}$ cannot be T. Indeed, assuming $\hat{t} = T$ we can estimate

$$\begin{aligned}
\delta - 2\eta|z|^2 &= \psi_{\varepsilon,\eta}(s,z,z) \\
&\leq \psi_{\varepsilon,\eta}(T,\hat{x},\hat{x}') \\
&= \sup_{x,x'} \left[u(T,x) - v(T,x') - \frac{|x-x'|^2}{\varepsilon^2} - \eta(|x|^2 + |x'|^2) \right].
\end{aligned}$$

Now by Theorem 3.1 in [18], applied to $u(T,x) - \eta|x|^2$ and $v(T,x) + \eta|x'|^2$, we get

$$\lim_{\varepsilon\to 0} \psi_{\varepsilon,\eta}(T,\hat{x},\hat{x}') = \sup_x \left[u(T,x) - v(T,x) - 2\eta|x|^2 \right] \leq \sup_x \left[u(T,x) - v(T,x) \right] \leq 0.$$

It follows that for ε,η small enough, $\hat{t} \neq T$. Also, by assumption $\tilde{u}(s,z) - \tilde{v}(s,z) > \delta$; hence we have for η small enough, that $\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}') \geq L_{\varepsilon,\eta} \geq \delta > 0$. We assume to be in this scenario from now on.

By applying the parabolic Theorem on Sums (e.g. Theorem 8.3 in [18]), we get

$$\begin{aligned}
(b,\hat{p},\hat{X}) &\in \bar{\mathcal{P}}^{2,+}\tilde{u}(\hat{t},\hat{x}), \\
(b',\hat{p}',\hat{X}') &\in \bar{\mathcal{P}}^{2,+}\tilde{v}(\hat{t},\hat{x}'),
\end{aligned}$$

such that $b - b' = 0$, $\hat{p} = 2\frac{\hat{x}-\hat{x}'}{\varepsilon^2} + 2\eta\hat{x}$, $\hat{p}' = 2\frac{\hat{x}-\hat{x}'}{\varepsilon^2} - 2\eta\hat{x}'$ and

$$\begin{pmatrix} \hat{X} & 0 \\ 0 & -\hat{X}' \end{pmatrix} \leq \frac{4}{\varepsilon^2} \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + 4\eta \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}. \tag{84}$$

Indeed, defining $\varphi(t,x,x') := \frac{|x-x'|^2}{\varepsilon^2} + \eta(|x|^2 + |x'|^2)$ we have

$$A := D^2\varphi(t,x,x') = \frac{2}{\varepsilon^2} \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + 2\eta \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

Then

$$A^2 = \left(\frac{8}{\varepsilon^4} + 8\frac{\eta}{\varepsilon^2} \right) \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + 4\eta^2 \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

By the Theorem on Sums, for every $a > 0$ there exist said elements of the jets, such that

$$\begin{pmatrix} \hat{X} & 0 \\ 0 & -\hat{X}' \end{pmatrix} \leq A + aA^2.$$

Hence we can choose a so small such that (84) holds.

By the viscosity property,

$$\begin{aligned}0 \le \operatorname{Tr}[\sigma\sigma^T(\hat{t},\hat{x})\hat{X}] - \operatorname{Tr}[\sigma\sigma^T(\hat{t},\hat{x}')\hat{X}'] + \langle b(\hat{t},\hat{x}),\hat{p}\rangle - \langle b(\hat{t},\hat{x}'),\hat{p}'\rangle \\ + \tilde{f}(\hat{t},\hat{x},\tilde{u}(\hat{t},\hat{x}),\hat{p}\sigma(\hat{t},\hat{x})) - \tilde{f}(\hat{t},\hat{x}',\tilde{v}(\hat{t},\hat{x}'),\hat{p}'\sigma(\hat{t},\hat{x}')) \\ = (i) + (ii) + (iii).\end{aligned}$$

Where

$$\begin{aligned}(i) &:= \operatorname{Tr}[\sigma\sigma^T(\hat{t},\hat{x})\hat{X}] - \operatorname{Tr}[\sigma\sigma^T(\hat{t},\hat{x}')\hat{X}'],\\ (ii) &:= \langle b(\hat{t},\hat{x}),\hat{p}\rangle - \langle b(\hat{t},\hat{x}'),\hat{p}'\rangle,\\ (iii) &:= \tilde{f}(\hat{t},\hat{x},\tilde{u}(\hat{t},\hat{x}),\hat{p}\sigma(\hat{t},\hat{x})) - \tilde{f}(\hat{t},\hat{x}',\tilde{v}(\hat{t},\hat{x}'),\hat{p}'\sigma(\hat{t},\hat{x}')).\end{aligned}$$

Multiplying (84) with $\begin{pmatrix}\sigma(\hat{t},\hat{x})\\ \sigma(\hat{t},\hat{x}')\end{pmatrix}$ from the right side, with $\begin{pmatrix}\sigma(\hat{t},\hat{x})\\ \sigma(\hat{t},\hat{x}')\end{pmatrix}^T$ from the left and then taking the trace, we get

$$\begin{aligned}(i) &= \operatorname{Tr}[\sigma\sigma^T(\hat{t},\hat{x})\hat{X}] - \operatorname{Tr}[\sigma\sigma^T(\hat{t},\hat{x}')\hat{X}']\\ &\le \frac{4}{\varepsilon^2}||\sigma(\hat{t},\hat{x}) - \sigma(\hat{t},\hat{x}')||_2^2 + 4\eta\left(||\sigma(\hat{t},\hat{x})||_2^2 + ||\sigma(\hat{t},\hat{x}')||_2^2\right)\\ &\le C_\sigma\frac{4}{\varepsilon^2}|\hat{x}-\hat{x}'|^2 + 8\eta C_\sigma^2.\end{aligned}$$

Moreover

$$\begin{aligned}(ii) &= \langle b(\hat{t},\hat{x}'),\hat{p}'\rangle - \langle b(\hat{t},\hat{x}),\hat{p}\rangle\\ &\le |b(\hat{t},\hat{x}') - b(\hat{t},\hat{x})||2\frac{\hat{x}-\hat{x}'}{\varepsilon^2}| + |b(\hat{t},\hat{x})||2\eta\hat{x}' + 2\eta\hat{x}|\\ &\le C_b|\hat{x}-\hat{x}'||2\frac{\hat{x}-\hat{x}'}{\varepsilon^2}| + C_b 2\eta(|\hat{x}'| + |\hat{x}|)\\ &= C_b 2\frac{|\hat{x}-\hat{x}'|^2}{\varepsilon^2} + C_b 2\eta(|\hat{x}'| + |\hat{x}|).\end{aligned}$$

We have

$$\begin{aligned}(iii) &= \tilde{f}(\hat{t},\hat{x},\tilde{u}(\hat{t},\hat{x}),\hat{p}\sigma(\hat{t},\hat{x})) - \tilde{f}(\hat{t},\hat{x}',\tilde{v}(\hat{t},\hat{x}'),\hat{p}'\sigma(\hat{t},\hat{x}'))\\ &= \int_0^1 [\partial_x\tilde{f}((*))(\hat{x}-\hat{x}') + \partial_{\tilde{y}}\tilde{f}((*))(\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}')) + \partial_{\tilde{z}}\tilde{f}((*))(\hat{p}\sigma(\hat{t},\hat{x}) - \hat{p}'\sigma(\hat{t},\hat{x}'))]d\lambda,\end{aligned}$$

where

$$(*) := (\hat{t}, \lambda\hat{x} + (1-\lambda)\hat{x}', \lambda\tilde{u}(\hat{t},\hat{x}) + (1-\lambda)\tilde{v}(\hat{t},\hat{x}'), \lambda\hat{p}\sigma(\hat{t},\hat{x}) + (1-\lambda)\hat{p}'\sigma(\hat{t},\hat{x}')).$$

We know $|\tilde{v}(\hat{t},\hat{x}')| \le ||\tilde{v}||_\infty < \infty$ and by the upper boundedness of $\tilde{u}$ and by the definition of the maximizier we get $\infty > C \ge \tilde{u}(\hat{t},\hat{x}) \ge \tilde{u}(T,0) - ||\tilde{v}||_\infty > \infty$. Hence $\lambda\tilde{u}(\hat{t},\hat{x}) + (1-\lambda)\tilde{v}(\hat{t},\hat{x}')$ is always bounded and we can assume that actually

$$|\partial_{\tilde{z}}\tilde{f}(t,x,\tilde{y},\tilde{z})| \le \tilde{C}(1+|\tilde{z}|).$$

Remember moreover that we assume η small enough, such that $\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}') \ge L_{\varepsilon,\eta} \ge \delta > 0$. Especially we have $|\hat{x}-\hat{x}'| \le \varepsilon\sqrt{\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}')}$. Hence we can estimate (let

$$(**) := \lambda \hat{p}\sigma(\hat{t},\hat{x}) + (1-\lambda)\hat{p}'\sigma(\hat{t},\hat{x}'))$$

$$\begin{aligned}
(iii) \le & \int_0^1 [\tilde{C}(1+|(**)|^2)|\hat{x}-\hat{x}'| - \tilde{K}(1+|(**)|^2)(\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}')) \\
& \quad + \tilde{C}(1+|(**)|)|\hat{p}\sigma(\hat{t},\hat{x}) - \hat{p}'\sigma(\hat{t},\hat{x}')|]d\lambda \\
\le & \int_0^1 [\tilde{C}(1+|(**)|^2)\varepsilon\sqrt{\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}')} - \tilde{K}(1+|(**)|^2)(\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}')) \\
& \quad + \tilde{C}\vartheta(1+|(**)|)^2(\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}')) + \frac{1}{\vartheta(\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}'))}|\hat{p}\sigma(\hat{t},\hat{x}) - \hat{p}'\sigma(\hat{t},\hat{x}')|^2]d\lambda \\
\le & \int_0^1 [\tilde{C}(1+|(**)|^2)\varepsilon\sqrt{\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}')} - \tilde{K}(1+|(**)|^2)(\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}')) \\
& \quad + \tilde{C}\vartheta 2(1+|(**)|^2)(\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}')) + \frac{1}{\vartheta(\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}'))}|\hat{p}\sigma(\hat{t},\hat{x}) - \hat{p}'\sigma(\hat{t},\hat{x}')|^2]d\lambda.
\end{aligned}$$

Choose $\vartheta = \frac{\tilde{C}}{6\tilde{K}}$. We then have for $\varepsilon^2 < \frac{\tilde{C}\sqrt{\delta}}{3\tilde{K}}$

$$\begin{aligned}
(iii) \le & -\frac{\tilde{K}}{3}\left(\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}')\right) + |\hat{p}\sigma(\hat{t},\hat{x}) - \hat{p}'\sigma(\hat{t},\hat{x}')|^2 \frac{1}{\vartheta\left(\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}')\right)} \\
\le & -\frac{\tilde{K}}{3}\left(\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}')\right) + (|\hat{p}|C_\sigma|\hat{x}-\hat{x}'| + |\hat{p}-\hat{p}'|C_\sigma)^2 \frac{1}{\vartheta\left(\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}')\right)} \\
\le & -\frac{\tilde{K}}{3}\left(\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}')\right) + (2\frac{|\hat{x}-\hat{x}'|^2}{\varepsilon^2} + 2C_\sigma\eta|\hat{x}||\hat{x}-\hat{x}'| + C_\sigma|2\eta\hat{x} + 2\eta\hat{x}'|)^2 \frac{1}{\vartheta\left(\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x}')\right)}.
\end{aligned}$$

Now, Lemma 3.5 in [43] (see also Lemma 2 in [10]) yields

$$\begin{aligned}
&\liminf_{\varepsilon\to 0}\liminf_{\eta\to 0}\left[\tilde{u}(\hat{t},\hat{x}) - \tilde{v}(\hat{t},\hat{x'})\right] = L', \\
&\limsup_{\varepsilon\to 0}\limsup_{\eta\to 0}\frac{|\hat{x}-\hat{x}'|}{\varepsilon} = 0, \\
&\limsup_{\varepsilon\to 0}\limsup_{\eta\to 0}\eta(|\hat{x}|^2 + |\hat{x}'|^2) = 0.
\end{aligned}$$

Hence we get $\limsup_{\varepsilon\to 0}\limsup_{\eta\to 0}(i) \le 0$, $\limsup_{\varepsilon\to 0}\limsup_{\eta\to 0}(ii) \le 0$, and $\limsup_{\varepsilon\to 0}\limsup_{\eta\to 0}(iii) \le -\frac{\tilde{K}}{3}L'$. Combining, we arrive at

$$0 \le -\frac{\tilde{K}}{3}L',$$

which is the desired contradiction. □

In the proof above we need a version of the classical theorem of sums, that does not seem to be available in the literature. Let us recall the usual statement first.

Theorem 49 ([19, Thm 7]). *Let $u_1, u_2 \in \mathrm{USC}((0,T)\times\mathbb{R}^n)$ and $w \in \mathrm{USC}\left((0,T)\times\mathbb{R}^{2n}\right)$ be given by*

$$w(t,x) = u_1(t,x_1) + u_2(t,x_2)$$

Suppose that $s \in (0,T)$, $z = (z_1,z_2) \in \mathbb{R}^{2n}$, $b \in \mathbb{R}$, $p = (p_1,p_2) \in \mathbb{R}^{2n}$, $A \in \mathcal{S}^{2n}$ with

$$(b,p,A) \in \mathcal{P}^{2,+}w(s,z). \tag{85}$$

Assume moreover that there is an $r>0$ such that for every $M>0$ there is a C such that for $i=1,2$

$$\begin{aligned} b_i &\leq C \text{ whenever } (b_i, q_i, X_i) \in \mathcal{P}^{2,+} w(t, x_i), \\ |x_i - z_i| + |s-t| &< r \text{ and } |u_i(t, x_i)| + |q_i| + \|X_i\| \leq M. \end{aligned} \tag{86}$$

Then for each $\varepsilon>0$ there exists $(b_i,\, X_i) \in \mathbb{R} \times \mathcal{S}^n$ such that

$$(b_i, p_i, X_i) \in \bar{\mathcal{P}}^{2,+} u(s, z_i)$$

and

$$-\left(\frac{1}{\varepsilon} + \|A\|\right) I \leq \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \leq A + \varepsilon A^2 \text{ and } b_1 + b_2 = b. \tag{87}$$

The proof of the above theorem is reduced (cf. Lemma 8 in [19]) to the case $b=0, z=0, p=0$ and $v_1(s,0) = v_2(s,0) = 0$, where (in order to avoid confusion) we write v_i instead of u_i. Condition (85) translates than to

$$v_1(t, x_1) + v_2(t, x_2) - \frac{1}{2}\langle Ax, x\rangle \leq 0 \text{ for all } (t,x) \in (0,T) \times \mathbb{R}^{2n}; \tag{88}$$

this also means that the left-hand-side as a function of (t, x_1, x_2) has a global maximum at $(s,0,0)$. The assertion of the (reduced) theorem is then the existence of $(b_i,\, X_i) \in \mathbb{R} \times \mathcal{S}^n$ such that $(b_i, 0, X_i) \in \bar{\mathcal{P}}^{2,+} v_i(s,0)$ for $i=1,2$ and (87) holds with $b=0$.

Theorem 50. *Assume that u_i has a finite extension to $(0,T] \times \mathbb{R}^n$, $i=1,2$, via its semi-continuous envelopes, that is,*

$$u_i(T,x) = \limsup_{\substack{(t,y) \in (0,T) \times \mathbb{R}^n: \\ t \uparrow T, y \to x}} u_i(t,y) < \infty.$$

Then the above theorem remains valid at $s=T$ if

$$\mathcal{P}^{2,+} w(s,z) \text{ and } \bar{\mathcal{P}}^{2,+} u(s, z_i)$$

is replaced by

$$\mathcal{P}_Q^{2,+} w(T,z) \text{ and } \bar{\mathcal{P}}_Q^{2,+} u(T, z_i)$$

and the final equality in (87) is replaced by

$$b_1 + b_2 \geq b. \tag{89}$$

Remark 51. *If we knew (but we don't!) that the final conclusion is $(b_i, p_i, X_i) \in \mathcal{P}^{2,+} u(T, z_i)$, rather than just being an element in the closure $\bar{\mathcal{P}}_Q^{2,+} u(T, z_i)$, then we could trivially diminish the b_i's such as to have $b_1 + b_2 = b$.*

Proof. **Step 1:** We focus on the reduced setting (and thus write v_i instead of u_i) and (following the proof of Lemma 8 in [19]) redefine $v_i(t_i, x_i)$ as $-\infty$ when $|x_i| > 1$ or $t_i \notin [T/2, T]$. We can also assume that (88) is strict if $t < s = T$ or $x \neq 0$. For the rest of the proof, we shall abbreviate $(t_1, t_2), (x_1, x_2)$ etc by (t, x). With this notation in mind we set

$$w(t,x) = v_1(t_1, x_1) + v_2(t_2, x_2) - \frac{1}{2}\langle Ax, x\rangle.$$

By the extension via semi-continuous envelopes, there exist a sequence $(t^n, x^n) \in (0,T)^2 \times (\mathbb{R}^n)^2$, such that

$$(t^n, x^n) \equiv \left(t^{1,n}, t^{2,n}, x^{1,n}, x^{2,n}\right) \to (T,T,0,0)\,.$$

We now consider w with a penalty term for $t_1 \neq t_2$ and a barrier at time T for both t_1 and t_2.

$$\psi_{m,n}(t,x) = w(t,x) - \left\{ \frac{m}{2} |t_1 - t_2|^2 + \sum_{i=1}^{2} \left(T - t^{i,n}\right)^2 / (T - t_i) \right\},$$

indexed by $(m,n) \in \mathbb{N}^2$, say. By assumption w has a maximum at $(T,T,0,0)$ which we may assume to be strict (otherwise subtract suitable forth order terms ...). Define now

$$\left(\hat{t}, \hat{x}\right) \in \arg\max \psi_{m,n} \text{ over } [T-r,T]^2 \times \bar{B}_r(0)^2$$

where $r = T/2$ (for instance). When we want to emphasize dependence on m,n we write $\left(\hat{t}_{m,n}, \hat{x}_{m,n}\right)$. We shall see below (Step 2) that there exists increasing sequences $m = m(k)\,, n = n(k)$ so that

$$\left(\hat{t}, \hat{x}\right)|_{m=m(k), n=n(k)} \to (T,T,0,0)\,. \tag{90}$$

Using the (elliptic) theorem of sums in the form of [19, Theorem 1] we find that there are

$$(b_i, p_i, X_i) \in \bar{\mathcal{P}}^{2,+} v_i\left(\hat{t}_i, \hat{x}_i\right)$$

(where $\hat{t}_i \to T, \hat{x}_i \to 0$ as $k \to \infty$) such that the first part of (87) holds and

$$A \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}, \; b_i = m(t_i - t_{3-i}) + \left(T - t^{i,\varepsilon}\right)^2 / (T - t_i)^2\,.$$

for $i = 1,2$. Note that

$$b_1 + b_2 = m(t_1 - t_2) + m(t_2 - t_1) + (\text{positive terms}) \geq 0;$$

since each b_i is bounded above by the assumptions and the estimates on the X_i it follows that the b_i lie in precompact sets. Upon passing to the limit $k \to \infty$ we obtain points

$$(b_i, p_i, X_i) \in \bar{\mathcal{P}}^{2,+} v_i(T,0)\,, \;\; i = 1,2;$$

with $b_1 + b_2 \geq 0$.

Step 2: We still have to establish (90). We first remark that for arbitrary (strictly) increasing sequences $m(k)\,, n(k)$, compactness implies that

$$\left\{\left(\hat{t}_{m(k),n(k)}, \hat{x}_{m(k),n(k)}\right) : k \geq 1\right\} \in [T-r,T]^2 \times \bar{B}_r(0)^2$$

has limit points. Note also $\hat{t}_1, \hat{t}_2 \in [T-r, T)$ thanks to the barrier at time T. The key technical ingredient for the remained of the argument is and we postpone details of these to Step 3 below:

$$w\left(\hat{t}, \hat{x}\right) - \psi_{m,n}\left(\hat{t}, \hat{x}\right) = \left\{ \frac{m}{2} \left|\hat{t}_1 - \hat{t}_2\right|^2 + \sum_{i=1}^{2} \left(T - t^{i,n}\right)^2 / \left(T - \hat{t}_i\right) \right\} \to 0 \text{ as } \frac{1}{n} << \frac{1}{m} \to 0. \tag{91}$$

In particular, for every $k > 0$ there exists $m(k)$ such that for all $m \geq m(k)$

$$\limsup_{n\to\infty} \{...\} < \frac{1}{k}.$$

By making $m(k)$ larger if necessary we may assume that $m(k)$ is (strictly) increasing in k. Furthermore there exists $n(m(k), k) = n(k)$ such that for all $n \geq n(k) : \{...\} < 2/k$. Again, we may make $n(k)$ larger if necessary so that $n(k)$ is strictly increasing. Recall $t^{1,n(k)} - t^{2,n(k)} \to T - T = 0$ as $k \to \infty$. For reasons that will become apparent further below, we actually want the stronger statement that

$$\frac{m(k)}{2} \left| t^{1,n(k)} - t^{2,n(k)} \right|^2 \to 0 \text{ as } k \to \infty \tag{92}$$

which we can achieve by modifying $n(k)$ such as to run to ∞ even faster. Note that the so-constructed $m = m(k), n = n(k)$ has the property

$$\left[w\left(\hat{t}, \hat{x}\right) - \psi_{m,n}\left(\hat{t}, \hat{x}\right) \right] |_{m=m(k),n=(k)} = \{...\} |_{m=m(k),n=(k)} \to 0 \text{ as } k \to \infty. \tag{93}$$

By switching to a subsequence (k_l) if necessary we may also assume (after relabeling) that

$$\left(\hat{t}_{m(k),n(k)}, \hat{x}_{m(k),n(k)}\right) \to \left(\tilde{t}, \tilde{x}\right) \in [T - r, T]^2 \times \bar{B}_r(0)^2 \text{ as } k \to \infty.$$

In the sequel we think of $\left(\hat{t}, \hat{x}\right)$ as this sequence indexed by k. We have

$$\begin{aligned} w\left(\tilde{t}, \tilde{x}\right) &\geq \limsup_{k\to\infty} w\left(\hat{t}, \hat{x}\right) |_{m=m(k),n=(k)} \quad \text{by upper-semi-continuity} \\ &= \limsup_{k\to\infty} \psi_{m,n}\left(\hat{t}, \hat{x}\right) |_{m=m(k),n=(k)} \text{ thanks to (93).} \end{aligned} \tag{94}$$

On the other hand, thanks to the particular form of our time-T barrier,

$$\begin{aligned} & \psi_{m,n}\left(\hat{t}, \hat{x}\right) \\ \geq\ & \psi_{m,n}\left(t^n, x^n\right) \\ =\ & w\left(t^n, x^n\right) - \left\{ \frac{m}{2} \left| t^{1,n} - t^{2,n} \right|^2 + \sum_{i=1}^{2} \left(T - t^{i,n}\right) \right\}. \end{aligned}$$

Take now $m = m(k), n = n(k)$ as constructed above. Then

$$\begin{aligned} & \psi_{m,n}\left(\hat{t}, \hat{x}\right) |_{m=m(k),n=(k)} \\ \geq\ & w\left(t^{n(k)}, x^{n(k)}\right) \\ & - \left\{ \frac{m(k)}{2} \left| t^{1,n(k)} - t^{2,n(k)} \right|^2 + \sum_{i=1}^{2} \left(T - t^{i,n(k)}\right) \right\} \end{aligned}$$

The first term in the curly bracket goes to zero (with $k \to \infty$) thanks to (92), the other term goes to zero since $t^{i,n} \to T$ with $n \to \infty$, and hence also along $n(k)$. On the other hand (recall $x^{i,n} \to 0$)

$$w\left(t^{n(k)}, x^{n(k)}\right) \to v_1(T, 0) + v_2(T, 0) - \frac{1}{2} \langle A0, 0 \rangle = 0 \text{ as } k \to \infty.$$

(In the reduced setting $v_1(T, 0) = v_2(T, 0) = 0$.) It follows that

$$\liminf_{k\to\infty} \psi_{m,n}\left(\hat{t}, \hat{x}\right) |_{m=m(k),n=(k)} = 0.$$

Together with (94) we see that $w\left(\tilde{t}, \tilde{x}\right) \geq 0$. But $w(T, T, 0, 0) = 0$ was a strict maximum in $[T - r, T]^2 \times \bar{B}_r(0)^2$ and so we must have $\left(\tilde{t}, \tilde{x}\right) = (T, T, 0, 0)$.

Step 3: Set

$$M(h)=\sup_{\substack{(t,x)\in[T-r,T)^2\times\bar{B}_r(0)^2\\|t_1-t_2|<h}} w(t_1,t_2,x_1,x_2) \text{ and } M'=\lim_{h\to 0}M(h)$$

It is enough to show

$$\limsup_{\frac{1}{n}<<\frac{1}{m}\to 0} w\left(\hat{t},\hat{x}\right)\le M'\le \liminf_{\frac{1}{n}<<\frac{1}{m}\to 0}\psi_{m,n}\left(\hat{t},\hat{x}\right). \tag{95}$$

since the claimed

$$w\left(\hat{t},\hat{x}\right)-\psi_{m,n}\left(\hat{t},\hat{x}\right)=\left\{\frac{m}{2}\left|\hat{t}_1-\hat{t}_2\right|^2+\sum_{i=1}^{2}\left(T-t^{i,n}\right)^2/\left(T-\hat{t}_i\right)\right\}\to 0 \text{ as } \frac{1}{n}<<\frac{1}{m}\to 0.$$

follows from

$$\begin{aligned}\limsup_{\frac{1}{n}<<\frac{1}{m}\to 0}\{...\} &\le \limsup_{\frac{1}{n}<<\frac{1}{m}\to 0} w\left(\hat{t},\hat{x}\right)-\liminf_{\frac{1}{n}<<\frac{1}{m}\to 0}\psi_{m,n}\left(\hat{t},\hat{x}\right)\\ &\le 0 \text{ (and hence } =0).\end{aligned}$$

Note that $w\left(\hat{t},\hat{x}\right)$ is bounded on $[T-r,T]^2\times\bar{B}_r(0)^2$ so that

$$\left|\hat{t}_1-\hat{t}_2\right|^2=O(1/m)\implies w\left(\hat{t},\hat{x}\right)\le M\left(\text{const}/\sqrt{m}\right).$$

On the other hand, from the very definition of M' as $\lim_{h\to 0}M(h)$, there exists a family (t_h,x_h) so that

$$|t_{1,h}-t_{2,h}|\le h \text{ and } w(t_h,x_h)\to M' \text{ as } h\to 0 \tag{96}$$

For every m,n we may take (t_h,x_h) as argument of $\psi_{m,n}$ (which itself has a maximum at $\hat{t},\hat{x}$); hence

$$w(t_h,x_h)-\frac{m}{2}h^2-\sum_{i=1}^{2}\left(T-t^{i,n}\right)^2/(T-t_{i,h})\le\psi_{m,n}\left(\hat{t},\hat{x}\right). \tag{97}$$

Take now a sequence $n=n(h)$, fast enough increasing as $h\searrow$ such that $\left(T-t^{i,n}\right)^2/(T-t_{i,h})\to 0$ with $h\to 0$. It follows that

$$\begin{aligned}M' &= \lim_{h\to 0}w(t_h,x_h)\\ &= \liminf_{h\to 0}\left(w(t_h,x_h)-\frac{m}{2}h^2-\sum_{i=1}^{2}\left(T-t^{i,n(h)}\right)^2/(T-t_{i,h})\right)\\ &\le \liminf_{h\to 0}\psi_{m,n(h)}\left(\hat{t},\hat{x}\right)=\liminf_{n\to\infty}\psi_{m,n}\left(\hat{t},\hat{x}\right) \text{ by monotonicity of } \sup\psi_{m,n} \text{ in } n.\end{aligned}$$

(In the last equality we used that $t^{i,n}\uparrow T$; this shows that $\sup\psi_{m,n}$ is indeed monotone in n.) The proof is now finished. □

References

[1] G.B. Arous and F. Castell. Flow decomposition and large deviations. *Journal of functional analysis*, 140(1):23–67, 1996.

[2] G. Barles, R. Buckdahn, and E. Pardoux. Backward stochastic differential equations and integral-partial differential equations. *Stochastics An International Journal of Probability and Stochastic Processes*, 60(1):57–83, 1997.

[3] F. Baudoin and M. Hairer. A version of hörmanders theorem for the fractional brownian motion. *Probability theory and related fields*, 139(3):373–395, 2007.

[4] J.M. Bismut. Conjugate convex functions in optimal stochastic control. *J. Math. Anal. Appl*, 44(1973):384–404, 1973.

[5] V. I. Bogachev. *Measure theory. Vol. I.* Springer-Verlag, Berlin, 2007.

[6] Leo Breiman. *Probability.* Classics in Applied Mathematics. SIAM, Philadelphia, PA, 1992.

[7] R. Buckdahn and J. Ma. Stochastic viscosity solutions for nonlinear stochastic partial differential equations. I. *Stochastic Process. Appl.*, 93(2):181–204, 2001.

[8] R. Buckdahn and J. Ma. Pathwise stochastic control problems and stochastic HJB equations. *SIAM J. Control Optim.*, 45(6):2224–2256 (electronic), 2007.

[9] M. Caruana and P. K. Friz. Partial differential equations driven by rough paths. *J. Differential Equations*, 247(1):140–173, 2009.

[10] Michael Caruana, Peter K. Friz, and Harald Oberhauser. A (rough) pathwise approach to a class of non-linear stochastic partial differential equations. *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*, 28(1):27 – 46, 2011.

[11] T. Cass and P. Friz. Densities for rough differential equations under hörmander's condition. *Annals of mathematics*, 171(3):2115–2141, 2010.

[12] T. Cass, C. Litterer, and T. Lyons. Integrability estimates for gaussian rough differential equations. *Arxiv preprint arXiv:1104.1813*, 2011.

[13] T. Cass and T. Lyons. Evolving communities with individual preferences. *preprint*, 2010.

[14] Y.G. Chen, Y. Giga, and S. Goto. Remarks on viscosity solutions for evolution equations. *Proceedings of the Japan Academy. Series A Mathematical sciences*, 67(10):323–328, 1991.

[15] J. M. C. Clark. The design of robust approximations to the stochastic differential equations of nonlinear filtering. In *Communication systems and random process theory (Proc. 2nd NATO Advanced Study Inst., Darlington, 1977)*, pages 721–734. NATO Advanced Study Inst. Ser., Ser. E: Appl. Sci., No. 25. Sijthoff & Noordhoff, Alphen aan den Rijn, 1978.

[16] J. M. C. Clark and D. Crisan. On a robust version of the integral representation formula of nonlinear filtering. *Probab. Theory Related Fields*, 133(1):43–56, 2005.

[17] L. Coutin and Z. Qian. Stochastic analysis, rough path analysis and fractional brownian motions. *Probability theory and related fields*, 122(1):108–140, 2002.

[18] M. G. Crandall, H Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.

[19] M.G. Crandall and H. Ishii. The maximum principle for semicontinuous functions. *Differential Integral Equations*, 3(6):1001–1014, 1990.

[20] D. Crisan and B. Rozovskiĭ (eds). *The Oxford handbook of nonlinear filtering.* Oxford: Oxford University Press. xiv, 1063 p. , 2011.

[21] AM Davie. Differential equations driven by rough paths: An approach via discrete approximation. *Applied Mathematics Research eXpress*, 2007(2007), 2007.

[22] M. H. A. Davis. On a multiplicative functional transformation arising in nonlinear filtering theory. *Z. Wahrsch. Verw. Gebiete*, 54(2):125–139, 1980.

[23] M. H. A. Davis. A pathwise solution of the equations of nonlinear filtering. *Teor. Veroyatnost. i Primenen.*, 27(1):160–167, 1982.

[24] M. H. A. Davis and M. P. Spathopoulos. Pathwise nonlinear filtering for nondegenerate diffusions with noise correlation. *SIAM J. Control Optim.*, 25(2):260–278, 1987.

[25] M.H.A. Davis. Pathwise nonlinear filtering. Stochastic systems: the mathematics of filtering and identification and applications, Proc. NATO Adv. Study Inst., Les Arcs/Savoie/ France 1980, 505-528 (1981)., 1981.

[26] M.H.A. Davis. Pathwise nonlinear filtering with correlated noise. Crisan, Dan (ed.) and Rozovskiĭ, Boris (ed.), The Oxford handbook of nonlinear filtering. Oxford: Oxford University Press. 403-424 (2011)., 2011.

[27] A. Deya, M. Gubinelli, and S. Tindel. Non-linear rough heat equations. *Arxiv preprint arXiv:0911.0618*, 2009.

[28] J. Diehl, P.K. Friz, and H. Oberhauser. Parabolic Comparison Revisited and Applications. *Preprint*, 2011.

[29] H. Doss. Liens entre équations différentielles stochastiques et ordinaires. *Ann. Inst. H. Poincaré Sect. B (N.S.)*, 13(2):99–125, 1977.

[30] N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Math. Finance*, 7(1):1–71, 1997.

[31] W.H. Fleming and H.M. Soner. *Controlled Markov processes and viscosity solutions.* Springer Verlag, 2006.

[32] P. K. Friz and H. Oberhauser. Rough path stability of SPDEs arising in non-linear filtering. *Arxiv preprint arXiv:1005.1781*, 2010.

[33] P. K. Friz and N. B. Victoir. *Multidimensional stochastic processes as rough paths*, volume 120 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 2010. Theory and applications.

[34] P.K. Friz. Continuity of the Ito-map for Hoelder rough paths with applications to the support theorem in Hoelder norm. *Probability and Partial Differential Equations in Modern Applied Mathematics, IMA Volumes in Mathematics and its Applications, Vol 140 (2005) 117-135*, 2005.

[35] JG Gaines and TJ Lyons. Variable step size control in the numerical solution of stochastic differential equations. *SIAM Journal on Applied Mathematics*, pages 1455–1484, 1997.

[36] M. Gubinelli and S. Tindel. Rough evolution equations. *Arxiv preprint arXiv:0803.0552v1*, 2008.

[37] I. Gyöngy and N. Krylov. Existence of strong solutions for itô's stochastic equations via approximations. *Probability theory and related fields*, 105(2):143–158, 1996.

[38] I. Gyongy and T. Martínez. On stochastic differential equations with locally unbounded drift. *Czechoslovak Mathematical Journal*, 51(4):763–783, 2001.

[39] M. Hairer and N.S. Pillai. Ergodicity of hypoelliptic sdes driven by fractional brownian motion. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 47, pages 601–628. Institut Henri Poincaré, 2011.

[40] BM Hambly and TJ Lyons. Stochastic area for brownian motion on the sierpinski gasket. *The Annals of Probability*, 26(1):132–148, 1998.

[41] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, pages 13–30, 1963.

[42] O. Kallenberg. *Foundations of modern probability.* Springer Verlag, 2002.

[43] M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Probab.*, 28(2):558–602, 2000.

[44] T.G. Kurtz and P. Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *The Annals of Probability*, 19(3):1035–1070, 1991.

[45] H. J. Kushner. A robust discrete state approximation to the optimal nonlinear filter for a diffusion. *Stochastics*, 3(2):75–83, 1979.

[46] M. Ledoux, Z. Qian, and T. Zhang. Large deviations and support theorem for diffusion processes via rough paths. *Stochastic processes and their applications*, 102(2):265–283, 2002.

[47] G. Liang, T. J. Lyons, and Z. Qian. Backward stochastic dynamics on a filtered probability space. *Arxiv preprint arXiv:0904.0377*, 2009.

[48] P.-L. Lions and P. E. Souganidis. Fully nonlinear stochastic pde with semilinear stochastic dependence. *C. R. Acad. Sci. Paris Sér. I Math.*, 331(8):617–624, 2000.

[49] T. J. Lyons. Differential equations driven by rough signals. I. An extension of an inequality of L. C. Young. *Math. Res. Lett.*, 1(4):451–464, 1994.

[50] T. J. Lyons, M. Caruana, and T. Lévy. *Differential equations driven by rough paths*, volume 1908 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007. Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004.

[51] T. J. Lyons and Z. Qian. *System Control and Rough Paths.* Oxford University Press, 2002. Oxford Mathematical Monographs.

[52] T.J. Lyons. Differential equations driven by rough signals. *Revista Matemática Iberoamericana*, 14(2):215–310, 1998.

[53] Y. Mishura. *Stochastic calculus for fractional Brownian motion and related processes.* Number 1929. Springer, 2008.

[54] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. *Stochastic partial differential equations and their applications*, pages 200–217, 1992.

[55] É. Pardoux and S. Peng. Backward doubly stochastic differential equations and systems of quasilinear SPDEs. *Probability Theory and Related Fields*, 98(2):209–227, 1994.

[56] É. Pardoux and S. G. Peng. Adapted solution of a backward stochastic differential equation. *Systems Control Lett.*, 14(1):55–61, 1990.

[57] E. Pardoux and SG Peng. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61, 1990.

[58] H. Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61 of *Stochastic Modelling and Applied Probability.* Springer-Verlag, Berlin, 2009.

[59] P.E. Protter. *Stochastic integration and differential equations.* Springer Verlag, 2004.

[60] H. J. Sussmann. On the gap between deterministic and stochastic ordinary differential equations. *Ann. Probability*, 6(1):19–41, 1978.

[61] J. Teichmann. Another approach to some rough and stochastic partial differential esquations. *Arxiv preprint arXiv:0908.2814v1*, 2009.

[62] L.C. Young. An inequality of the hölder type, connected with stieltjes integration. *Acta Mathematica*, 67(1):251–282, 1936.

[63] M. Zaehle. Forward integrals and stochastic differential equations. *Progress in Probability*, pages 293–302, 2002.