ROUGH PATH STABILITY OF (SEMI-)LINEAR SPDES

PETER FRIZ AND HARALD OBERHAUSER

Abstract. We give meaning to linear and semi-linear (possibly degenerate) parabolic partial differential equations with (affine) linear rough path noise and establish stability in a rough path metric. In the case of enhanced Brownian motion (Brownian motion with its Lévy area) as rough path noise the solution coincides with the standard variational solution of the SPDE.

1. Introduction

Given a continuous, $d$-dimensional semimartingale $Z = (Z^1, \ldots, Z^d)$ consider the SPDE

$$du + L (t, x, u, Du, D^2u) dt = \sum_{k=1}^d \Lambda_k (t, x, u, Du) \circ dZ^k_t,$$  

with scalar initial data $u(0, \cdot) = u_0 (\cdot)$ on $\mathbb{R}^n$, $L$ a (semi-)linear second order operator of the form

$$L (t, x, r, p, X) = -\text{Tr} [A (t, x) \cdot X] + b (t, x) \cdot p + c (t, x, r)$$

and $\Lambda$ a collection of first order different operators $\Lambda_k (t, x, r, p)$ which are affine linear in $r, p$, that is,

$$\Lambda_k (t, x, r, p) = p \cdot \sigma_k (t, x) + r \nu_k (t, x) + g_k (t, x), \quad k = 1, \ldots, d.$$  

The contribution of this article is to give meaning to equation (1.1) when $Z (\omega)$ is replaced by a rough path $z$ (this is carried out in sections 2, 3 and 4). Our main result as stated and proven in section 4 (in section 2 we recall $\text{Lip}^\gamma$-regularity, rough paths and their metrics and the $\text{BUC}$ space of bounded, uniformly continuous real-valued functions that appear in the theorem below) is the following

Theorem 1. Let $p \geq 1$. Assume $L$ fulfills assumption \text{(4)} and the coefficients of $\Lambda = (\Lambda_1, \ldots, \Lambda_{d_1+1+2+d_3})$ fulfill assumption \text{(3)} for some $\gamma > p + 2$ (assumptions \text{(4, 3)} are given in section 4 on page 13). Let $u_0 \in \text{BUC} (\mathbb{R}^n)$ and let $z$ be a geometric $p$-rough path. Then there exists a unique $u = u^\varepsilon \in \text{BUC} ([0,T] \times \mathbb{R}^n)$ such that for any sequence $(z^\varepsilon)_\varepsilon \subset C^1 ([0,T], \mathbb{R}^d)$ such that $z^\varepsilon \to z$ in $p$-rough path sense, the viscosity solutions $(u^\varepsilon) \subset \text{BUC} ([0,T] \times \mathbb{R}^n)$ of

$$\dot{u}^\varepsilon + L (t, x, u^\varepsilon, Du^\varepsilon, D^2u^\varepsilon) = \sum_{k=1}^d \Lambda_k (t, x, u^\varepsilon, Du^\varepsilon) z^\varepsilon_k t,$$

converge locally uniformly to $u^\varepsilon$. We write formally,

$$du + L (t, x, u, Du, D^2u) dt = \Lambda (t, x, u, Du) dz_t, \quad u (0, \cdot) = u_0 (\cdot).$$

Moreover, we have the contraction property

$$\sup_{(t, x) \in \mathbb{R}^n \times (0,T]} | u^\varepsilon (t, x) - \hat{u}^\varepsilon (t, x) | \leq e^{CT} \sup_{x \in \mathbb{R}^n} | u_0 (x) - \hat{u}_0 (x) |$$

(C given by (4.2)) and continuity of the solution map $(z, u_0) \mapsto u^\varepsilon$

$$C^{0,p\text{-var}} ([0,T], G^{[p]} (\mathbb{R}^d)) \times \text{BUC} (\mathbb{R}^n) \to \text{BUC} ([0,T] \times \mathbb{R}^n).$$
The resulting theory of rough PDEs can then be used (in a “rough-pathwise” fashion) to give meaning (and then existence, uniqueness, stability, etc.) to large classes of stochastic partial differential equations which has numerous benefits as discussed in section 6. By combining well-known Wong–Zakai type results of the $L^2$-theory of SPDEs [29, 8, 45, 44] with convergence of piecewise linear approximations to “enhanced” Brownian motion (EBM) in rough path sense, e.g. [22, Chapter 13 and 14], we show that the solutions provided by above theorem when applied with EBM as rough path are in fact the usual $L^2$-solutions of the variational approach. [31, 39, 41]. This “intersection” of RPDE/SPDE theory is made precise in Section 5. However, let us emphasize that neither theory is “contained” in the other, even in the case of Brownian driving noise. An appealing feature of our RPDE approach is that it can handle degenerate situations (including pure first order SPDEs) and automatically yields continuous versions of SPDE solutions without requiring dimension-dependent regularity assumptions on the coefficients (as pointed out by Krylov [30], a disadvantage of the $L^2$ theory of SPDEs). On the other hand, our regularity assumption (in particular in the noise terms) are more stringent than what is needed to ensure existence and uniqueness in the $L^2$ theory of SPDEs. Below we sketch our approach and the outline of this article.

1.1. Robustification. In fact, it is part of folklore that the equation (1.1) can be given a pathwise meaning in the case when there is no gradient noise ($\sigma = 0$ in (1.2)).

**Classical robustification:** if $\sigma = 0$ in (1.2) and also (for simplicity of presentation only) $\nu = \nu(x)$ (i.e. no time dependence) one can take a smooth path $z$ and solve the auxiliary differential equation $\dot{\phi} = \phi \sum_j \nu_j (x) dz^j \equiv \phi \nu \cdot dz$. The solution is given by

$$\phi_t = \phi_0 \exp \left( \int_0^t \nu(x) \cdot dz \right) = \phi_0 \exp (\nu(x) \cdot z_t)$$

and induces the flow map $\phi(t, \phi_0) := \phi_0 \exp (\nu(x) \cdot z_t)$; observe that these expressions can be extended by continuity to any continuous path $z$ such as a typical realization of $Z(\cdot)$. The point is that this transform allows to transform the SPDE into a random PDE (sometimes called the Zakai equation in robust form): it suffices to introduce $v$ via the ”outer transform” $u(t, x) = \phi(t, v(t, x))$ which leads immediately to

$$v(t, x) = \exp (-\nu(x) \cdot z_t) u(t, x).$$

An elementary computation then shows that $v$ solves a linear PDE given by an affine linear operator $\theta L$ in $v, Dv, D^2v$ with coefficients that will depend on $z$ resp. $Z_t(\cdot)$,

$$dv + \theta L(t, x, v, Dv, D^2v) dt = 0.$$

Moreover, one can conclude from this representation that $u = u(z)$ is continuous with respect to the uniform metric $|z - \tilde{z}|_{[0,T]} = \sup_{r \in [0,T]} |z_r - \tilde{z}_r|$. This provides a fully pathwise “robust” approach (the extension to vector field $\nu = \nu(t, x)$ with sufficiently smooth time-dependence is easy).

**Rough path robustification:** The classical robustification does not work in presence of general gradient noise. In fact, we can not expect PDE solutions to

$$du + L(t, x, u, Du, D^2u) dt = \sum_{k=1}^{d} A_k(t, x, u, Du) dz^k,$$

(which are well-defined for smooth $z : [0, T] \rightarrow \mathbb{R}^d$) to depend continuously on $z$ in uniform topology (cf. the “twisted approximations” of section 6). Our main result is that $u = u(z)$ is continuous with respect to rough path metric.\(^1\) That is, if $(z^n) \subset C^1([0, T], \mathbb{R}^d)$ is Cauchy in rough path metric then $(u^n)$ will converge to a limit which will be seen to depend only on the (rough path) limit of $z^n$ (and not on the approximating sequence). As a consequence, it

\(^1\)Two (smooth) paths $z, \tilde{z}$ are close in rough path metric iff $z$ is close to $\tilde{z}$ AND sufficiently many iterated integrals of $z$ are close to those of $\tilde{z}$. More details are given later in this article as needed.
is meaningful to replace \( z \) above by an abstract (geometric) rough path \( z \) and the analogue of Lyons’ universal limit theorem \([37]\) holds.

1.2. Structure and outline. We shall prefer to write the right hand side of (1.1) in the equivalent form

\[
\sum_{i=1}^{d_1} (Du \cdot \sigma_i (t, x)) \circ dZ_1^i + u \sum_{j=1}^{d_2} \nu_j (t, x) \circ dZ_2^j + \sum_{k=1}^{d_3} g_k (t, x) \circ dZ_3^k
\]

where \( Z \equiv (Z^1, Z^2, Z^3) \) and \( Z^1 \) is a \( d_1 \)-dimensional, continuous semimartingale. Our approach is based on a pointwise (viscosity) interpretation of (1.1): we successively transform away the noise terms such as to transform the SPDE, ultimately, into a random PDE. The big scheme of the paper is

\[
\begin{align*}
\text{Transformation 1} & \quad \mapsto \quad \text{Transformation 1} \quad \mapsto \quad \text{Transformation 2} \\
\text{Transformation 2} & \quad \mapsto \quad \text{Transformation 3}
\end{align*}
\]

\( u \) where \( u^1 \) has the (gradient) noise driven by \( Z^1 \) removed;

\( u^1 \) where \( u^{12} \) has the remaining noise driven by \( Z^2 \) removed;

\( u^{12} \) where \( \tilde{u} \) has the remaining noise driven by \( Z^3 \) removed.

None of these transformations is new on its own. The first is an example of Kunita’s stochastic characteristics method; the second is known as robustification (also known as Doss–Sussman transform); the third amounts to change \( u^{12} \) additively by a random amount and has been used in virtually every SPDE context with additive noise. What is new is that the combined transformation can be managed and is compatible with rough path convergence; for this we have to remove all probability from the problem: In fact, we will transform an RPDE (rough PDE) solution \( u \) into a classical PDE solution \( \tilde{u} \) in which the coefficients depend on various rough flows (i.e. the solution flows to rough differential equations) and their derivatives. Stability results of rough path theory and viscosity theory, in the spirit of \([6, 5]\), then play together to yield the desired result. Upon using the canonical rough path lift of the observation process in this RPDE one has constructed a robust version of the SPDE solution of equation (1.1).

We note that the viscosity/Stratonovich approach allows us to avoid any ellipticity assumption on \( L \); we can even handle the fully degenerate first order case. In turn, we only obtain \( BUC \) (bounded, uniformly continuous) solutions. Stronger assumptions would allow to discuss all this in a classical context (i.e. \( \tilde{u} \) would be a \( C^{1,2} \) solution) and SPDE solution can then be seen to have certain spatial regularity, etc.

We should remark that the usual way to deal with (1.1), which goes back to Pardoux, Krylov, Rozovskii, and others, \([29, 8, 45, 44]\), is to find solutions in a suitable functional analytic setting; e.g. such that solutions evolve in suitable Sobolev spaces. The equivalence of this solution concept with the RPDE approach as presented in sections 2 to 4 is then discussed in section 5. Interestingly, there has been no success until now (despite the advances by Deya–Gubinelli–Tindel \([23, 12]\) and Teichmann \([13]\)) to include (1.1) in a setting of abstract rough evolution equations on infinite-dimensional spaces.

Acknowledgement 2. The first author would like to thank the organizers and participants of the Filtering Workshop in June 2010, part of the Newton Institute’s SPDE program, where parts of this work was first presented. The second author would like to thank the organizers and participants of the Rough path and SPDE workshop, part of the Newton Institute’s SPDE program and expresses his gratitude for a Newton Institute Junior membership grant. Partial support from the European Unions Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement nr. 258237 is gratefully acknowledged. HO is indebted to G. Barles, P. Souganidis and the participants of the C.I.M.E. meeting on HJB-equations in 2011 for helpful conversations. Both authors would like to thank the referees for their valuable comments.

Transformation 2 and 3 could actually be performed in 1 step; however, the separation leads to a simpler analytic tractability of the transformed equations.
2. Background on viscosity theory and rough paths

Let us recall some basic ideas of (second order) viscosity theory \[9, 15\] and rough path theory \[37, 38\]. As for viscosity theory, consider a real-valued function \( u = u(t, x) \) with \( t \in [0, T], x \in \mathbb{R}^n \) and assume \( u \in C^2 \) is a classical subsolution,

\[
\partial_t u + F(t, x, u, Du, D^2u) \leq 0,
\]

where \( F \) is a (continuous) function, degenerate elliptic in the sense that

\[
F(t, x, r, p, A + B) \leq F(t, x, r, p, A)
\]

whenever \( B \geq 0 \) in the sense of symmetric matrices (cf. \[9\]). The idea is to consider a (smooth) test function \( \varphi \) and look at a local maxima \((\hat{t}, \hat{x})\) of \( u - \varphi \). Basic calculus implies that

\[
Du(\hat{t}, \hat{x}) = D\varphi(\hat{t}, \hat{x}), D^2u(\hat{t}, \hat{x}) \leq D\varphi(\hat{t}, \hat{x})
\]

and, from degenerate ellipticity,

\[
\partial_t \varphi + F(\hat{t}, \hat{x}, u, D\varphi, D^2\varphi) \leq 0.
\]

This suggests to define a viscosity supersolution (at the point \((\hat{t}, \hat{x})\)) to \( \partial_t + F = 0 \) as a continuous function \( u \) with the property that \((2.1)\) holds for any test function. Similarly, viscosity subsolutions are defined by reversing inequality in \((2.1)\); viscosity solutions are both super- and subsolutions. A different point of view is to note that \( u(t, x) \leq u(\hat{t}, \hat{x}) - \varphi(\hat{t}, \hat{x}) + \varphi(t, x) \) for \((t, x)\) near \((\hat{t}, \hat{x})\). A simple Taylor expansion then implies

\[
(2.2) \quad u(t, x) \leq u(\hat{t}, \hat{x}) + a(\hat{t} - t) + p \cdot (x - \hat{x}) + \frac{1}{2} (x - \hat{x})^T \cdot (x - \hat{x}) + o(||x - \hat{x}||^2 + ||\hat{t} - t||)
\]

as \( ||x - \hat{x}||^2 + ||\hat{t} - t|| \to 0 \) with \( a = \partial_t \varphi(\hat{t}, \hat{x}), p = D\varphi(\hat{t}, \hat{x}), X = D^2\varphi(\hat{t}, \hat{x}) \). Moreover, if \((2.2)\) holds for some \((a, p, X)\) and \( u \) is differentiable, then \( a = \partial_t u(\hat{t}, \hat{x}), p = Du(\hat{t}, \hat{x}), X \leq D^2u(\hat{t}, \hat{x}) \), hence by degenerate ellipticity

\[
\partial_t \varphi + F(\hat{t}, \hat{x}, u, D\varphi, D^2\varphi) \leq 0.
\]

Pushing this idea further leads to a definition of viscosity solutions based on a generalized notion of “\((\partial_t u, Du, D^2u)\)” for non-differentiable \( u \), the so-called parabolic semijets, and it is a simple exercise to show that both definitions are equivalent. The resulting theory (existence, uniqueness, stability, ...) is without doubt one of the most important recent developments in the field of partial differential equations. As a typical result, the initial value problem \((\partial_t + F)u = 0, u(0, \cdot) = u_0 \in BUC(\mathbb{R}^n)\) has a unique solution in \( BUC([0, T] \times \mathbb{R}^n) \) provided \( F = F(t, x, u, Du, D^2u) \) is continuous, degenerate elliptic, proper (i.e. increasing in the \( u \) variable) and satisfies a (well-known) technical condition. In fact, uniqueness follows from a stronger property known as comparison: assume \( u \) (resp. \( v \)) is a supersolution (resp. subsolution) and \( u_0 \geq v_0; \) then \( u \geq v \) on \([0, T] \times \mathbb{R}^n\). A key feature of viscosity theory is what workers in the field simply call stability properties. For instance, it is relatively straightforward to study \((\partial_t + F)u = 0 \) via a sequence of approximate problems, say \((\partial_t + F^n)u^n = 0, \) provided \( F^n \to F \) locally uniformly and some apriori information on the \( u^n \) (e.g. locally uniform convergence, or locally uniform boundedness). Note the stark contrast to the classical theory where one has to control the actual derivatives of \( u^n \).

The idea of stability is also central to rough path theory. Given a collection \((V_1, \ldots, V_d)\) of (sufficiently nice) vector fields on \( \mathbb{R}^n \) and \( z \in C^1([0, T], \mathbb{R}^d) \) one considers the (unique) solution \( y \) to the ordinary differential equation

\[
(2.3) \quad \dot{y}(t) = \sum_{i=1}^{d} V_i(y) \dot{z}^i(t), \quad y(0) = y_0 \in \mathbb{R}^n.
\]

\( ^3 \)BUC (...) denotes the space of bounded, uniformly continuous functions.

\( ^4 \)(3.14) of the User’s Guide [9].

\( ^5 \)What we have in mind here is the Barles–Perthame method of semi-relaxed limits [15].
The question is, if the output signal $y$ depends in a stable way on the driving signal $z$ (one handles time-dependent vector fields $V = V(t, y)$ by considering the $(d + 1)$-dimensional driving signal $t \mapsto (t, z_t)$). The answer, of course, depends strongly on how to measure distance between input signals. If one uses the supremums norm, so that the distance between driving signals $z, \tilde{z}$ is given by $|z - \tilde{z}|_{\infty;[0,T]} = \sup_{r \in [0,T]} |z_r - \tilde{z}_r|$, then the solution will in general not depend continuously on the input.

**Example 3.** Take $n = 1, d = 2, V = (V_1, V_2) = (\sin (\cdot), \cos (\cdot))$ and $y_0 = 0$. Obviously,

$$z^n(t) = \left(\frac{1}{n} \cos (2\pi n^2 t), \frac{1}{n} \sin (2\pi n^2 t)\right)$$

converges to $0$ in $\infty$-norm whereas the solutions to $\dot{y}^n = V(y^n) \dot{z}^n, y_0^n = 0$, do not converge to zero (the solution to the limiting equation $\dot{y} = 0$).

If $|z - \tilde{z}|_{\infty;[0,T]}$ is replaced by the (much) stronger distance

$$|z - \tilde{z}|_{1\text{-var};[0,T]} = \sup_{(t_i) \subset [0,T]} \sum |z_{t_i, t_{i+1}} - \tilde{z}_{t_i, t_{i+1}}|,$$

(using the notation $z_{s,t} := z_t - z_s$) it is elementary to see that now the solution map is continuous (in fact, locally Lipschitz); however, this continuity does not lend itself to push the meaning of (2.3): the closure of $C^1$ (or smooth) paths in variation is precisely $W^{1,1}$, the set of absolutely continuous paths (and thus still far from a typical Brownian path). Lyons’ theory of rough paths exhibits an entire cascade of $(p$-variation or $1/p$-Hölder type rough path) metrics, for each $p \in [1, \infty)$, on path-space under which such ODE solutions are continuous (and even locally Lipschitz) functions of their driving signal. For instance, the “rough path” $p$-variation distance between two smooth $\mathbb{R}^d$-valued paths $z, \tilde{z}$ is given by

$$\max_{j=1,\ldots,[p]} \left( \sup_{(t_i) \subset [0,T]} \left( \sum |z_{t_i, t_{i+1}}^{(j)} - \tilde{z}_{t_i, t_{i+1}}^{(j)}|^{1/p} \right)^p \right),$$

where $z_{s,t}^{(j)} = \int dz_{r_1} \otimes \cdots \otimes dz_{r_j}$ with integration over the $j$-dimensional simplex $\{ s < r_1 < \cdots < r_j < t \}$.

This allows to extend the very meaning of (2.3), in a unique and continuous fashion, to driving signals which live in the abstract completion of smooth $\mathbb{R}^d$-valued paths (with respect to rough path $p$-variation or a similarly defined $1/p$-Hölder metric). The space of so-called $p$-rough paths is precisely this abstract completion. In fact, this space can be realized as genuine path space, where $G^{[p]}(\mathbb{R}^d)$ is the free step-$[p]$ nilpotent group over $\mathbb{R}^d$, equipped with Carnot–Caratheodory metric; realized as a subset of $1 + \mathfrak{t}^{[p]}(\mathbb{R}^d)$ where

$$\mathfrak{t}^{[p]}(\mathbb{R}^d) = \mathbb{R}^d \oplus (\mathbb{R}^d)^{\otimes 2} \oplus \cdots \oplus (\mathbb{R}^d)^{\otimes [p]}$$

is the natural space for (up to $[p]$) iterated integrals of a smooth $\mathbb{R}^d$-valued path. For instance, almost every realization of $d$-dimensional Brownian motion $B$ **enhanced with its iterated stochastic integrals in the sense of Stratonovich**, i.e. the matrix-valued process given by

$$B^{(2)} := \left( \int_0 \mathfrak{t} \circ dB \right)_{i,j \in \{1, \ldots, d\}}$$

yields a path $B(\omega)$ in $G^2(\mathbb{R}^d)$ with finite $1/p$-Hölder (and hence finite $p$-variation) regularity, for any $p > 2$. ($B$ is known as Brownian rough path.) We remark that $B^{(2)} = \frac{1}{2} B \otimes B + A$ where the anti-symmetric part of the matrix, $A := \text{Anti}(B^{(2)})$, is known as Lévy’s stochastic area; in other words $B(\omega)$ is determined by $(B, A)$, i.e. Brownian motion **enhanced with Lévy’s area**. A similar construction work when $B$ is replaced by a generic multi-dimensional continuous semimartingales; see [22, Chapter 14] and the references therein.

---

8In the strict terminology of rough path theory: geometric $p$-rough paths.
3. Transformations

3.1. Inner and outer transforms. Throughout, \( F = F(t, x, r, p, X) \) is a continuous scalar-valued function on \([0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{S}(n)\), \( \mathbb{S}(n) \) denotes the space of symmetric \( n \times n \)-matrices, and \( F \) is assumed to be non-increasing in \( X \) (degenerate elliptic) and proper in the sense of \([7,11]\). Time derivatives of functions are denoted by upper dots, spatial derivatives (with respect to \( x \)) by \( D, D^2 \), etc. Further, we use \((\cdot, \cdot)\) to denote tensor contraction\(^7\), i.e.

\[
\langle p, q \rangle_{j_1, \ldots, j_n} = \sum_{i_1, \ldots, i_m} p_{i_1, \ldots, i_m} q_{j_1, \ldots, j_m}, \quad p \in (\mathbb{R}^l)^{\otimes m}, \quad q \in (\mathbb{R}^l)^{\otimes n} \otimes \left((\mathbb{R}^l)^{\otimes n}\right)^{\otimes m}.
\]

Lemma 4. [Inner Transform] Let \( \psi \) (with respect to sense of (7.1). Time derivatives of functions are denoted by upper dots, spatial derivatives \( \partial \)).

\[
\psi (t, x) \text{ is a viscosity subsolution (always assumed BUC) of}
\]

\[
\partial_t u + F(t, x, r, Du, D^2u) - \sum_{i=1}^d \langle Du \cdot \sigma_i (t, x) \rangle \dot{z}_i^t = 0; \quad u(0, .) = u_0 (.)
\]

iff \( w(t, x) := u(t, \psi(t, x)) \) is a viscosity subsolution of

\[
\partial_t w + F^\psi (t, x, w, Dw, D^2w) = 0; \quad w(0, .) = u_0 (.)
\]

where

\[
F^\psi (t, x, r, p, X) = F(t, \psi_t(x), r, \langle p, D\psi_t^{-1}|_{\psi_t(x)} \rangle, \langle X, D\psi_t^{-1}|_{\psi_t(x)} \otimes D\psi_t^{-1}|_{\psi_t(x)} \rangle + \langle p, D^2\psi_t^{-1}|_{\psi_t(x)} \rangle)
\]

and

\[
D\psi_t^{-1}|_x = \left(\frac{\partial (\psi_t^{-1}(t, x))^k}{\partial x^i}\right)_{i=1, \ldots, n} \quad \text{and} \quad D^2\psi_t^{-1}|_x = \left(\frac{\partial^2 (\psi_t^{-1}(t, x))^k}{\partial x^i x^j}\right)_{i, j=1, \ldots, n}.
\]

The same statement holds if one replaces the word subsolution by supersolution throughout.

Remark 5. The regularity assumptions on \( \sigma \) with respect to \( t \) can be obviously relaxed here. Treating time and space variable similarly will be convenient in the rough path framework where sharp results on time-dependent vector fields are hard to find in the literature (but see [3])

If we specialize from general \( F \) to a semilinear \( L : [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{S}(n) \to \mathbb{R} \) we get transformation 1 as a corollary.

Corollary 6. [Transformation 1] Let \( \psi = \psi(t, x) \) be the ODE flow of \( dy = \sigma(t, y) dz \), as above. Define \( L = L(t, x, r, p, X) \) by

\[
L = - \text{Tr}[A(t, x) \cdot X] + b(t, x) \cdot p + c(t, x, r);
\]

define also the transform

\[
L^\psi = - \text{Tr}[A^\psi(t, x) \cdot X] + b^\psi(t, x) \cdot p + c^\psi(t, x, r)
\]

where

\[
A^\psi(t, x) = \langle A(t, \psi_t(x)), D\psi_t^{-1}|_{\psi_t(x)} \otimes D\psi_t^{-1}|_{\psi_t(x)} \rangle,
\]

\[
b^\psi(t, x) \cdot p = b(t, \psi_t(x)) \cdot \langle p, D\psi_t^{-1}|_{\psi_t(x)} \rangle - \text{Tr}(A(t, \psi_t) \cdot \langle p, D^2\psi_t^{-1}|_{\psi_t(x)} \rangle),
\]

\[
c^\psi(t, x, r) = c(t, \psi_t(x), r).
\]

\(^7\)We also use \( \cdot \) to denote contraction over only index or to denote matrix multiplication.
Then \( u \) is a solution (always assumed \( \mathsf{BUC} \)) of
\[
\partial_t u + L(t, x, u, Du, D^2 u) = \sum_{i=1}^{d} (Du \cdot \sigma_i(t, x)) \frac{\partial}{\partial t} z_i^j; \quad u(0,.) = u_0(.)
\]
if and only if \( u^1(t, x) := u(t, \psi(t, x)) \) is a solution of
\[
(3.3) \quad \partial_t + L^\psi = 0; \quad u^1(0,.) = u_0(.)
\]

**Proof of Lemma 4.** Set \( y = \psi_t(x) \). When \( u \) is a classical sub-solution, it suffices to use the chain rule and definition of \( F^\psi \) to see that
\[
\dot{w}(t, x) = \dot{u}(t, y) + Du(t, y) \cdot \dot{\psi}_t(x) = \dot{u}(t, y) + Du(t, y) \cdot \sigma(y) \dot{z}_t,
\]
\[
\leq F(t, y, u(t, y), Du(t, y), D^2 u(t, y)) = F^\psi(t, x, t, y, Du(t, y), D^2 u(t, y), \dot{w}(t, x))
\]
The case when \( u \) is a viscosity sub-solution of (3.1) is not much harder: suppose that \( (\bar{t}, \bar{x}) \) is a maximum of \( w - \xi \), where \( \xi \in C^2([0, T] \times \mathbb{R}^n) \) and define \( \varphi(t, y) = \xi(t, \psi_t^{-1}(y)) \). Set \( \bar{y} = \psi_t(\bar{x}) \) so that
\[
F(\bar{t}, \bar{y}, u(\bar{t}, \bar{y}), Du(\bar{t}, \bar{y}), D^2 u(\bar{t}, \bar{y})) = F^\psi(\bar{t}, \bar{x}, w(\bar{t}, \bar{x}), D\xi(\bar{t}, \bar{x}), D^2\xi(\bar{t}, \bar{x}))
\]
Obviously, \( (\bar{t}, \bar{y}) \) is a maximum of \( u - \varphi \), and since \( u \) is a viscosity sub-solution of (3.1) we have
\[
\dot{\varphi}(\bar{t}, \bar{y}) + D\varphi(\bar{t}, \bar{y}) \sigma(\bar{t}, \bar{y}) \dot{z}(\bar{t}) \leq F(\bar{t}, \bar{y}, u(\bar{t}, \bar{y}), Du(\bar{t}, \bar{y}), D^2 u(\bar{t}, \bar{y}))
\]
On the other hand, \( \xi(t, x) = \varphi(t, \psi_t(x)) \) implies \( \dot{\xi}(\bar{t}, \bar{x}) = \dot{\varphi}(\bar{t}, \bar{y}) + D\varphi(\bar{t}, \bar{y}) \sigma(\bar{t}, \bar{y}) \dot{z}(\bar{t}) \) and putting things together we see that
\[
\dot{\xi}(\bar{t}, \bar{x}) \leq F^\psi(\bar{t}, \bar{x}, w(\bar{t}, \bar{x}), D\xi(\bar{t}, \bar{x}), D^2\xi(\bar{t}, \bar{x}))
\]
which says precisely that \( w \) is a viscosity sub-solution of (3.2). Replacing maximum by minimum and \( \leq \) by \( \geq \) in the preceding argument, we see that if \( u \) is a super-solution of (3.1), then \( w \) is a super-solution of (3.2).

Conversely, the same arguments show that if \( v \) is a viscosity sub- (resp. super-) solution for (3.2), then \( u(t, y) = w(t, \psi^{-1}(y)) \) is a sub- (resp. super-) solution for (3.1). \( \blacksquare \)

We prepare the next lemma by agreeing that for a sufficiently smooth function \( \phi = \phi(t, r, x) : [0, T] \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R} \) we shall write
\[
\dot{\phi} = \frac{\partial \phi}{\partial t}(t, r, x), \quad \phi' = \frac{\partial \phi}{\partial r}(t, r, x), \quad D\phi = \bigg( \frac{\partial \phi}{\partial x^i}(t, r, x) \bigg)_{i=1, ..., n} \quad \text{and} \quad D^2\phi = \bigg( \frac{\partial^2 \phi}{\partial x^i \partial x^j}(t, r, x) \bigg)_{i,j=1, ..., n}.
\]

**Lemma 7.** [Outer transform] Let \( \phi = \phi(t, r, x) \in C^{1,2} \) and assume that \( \forall (t, x), \ r \mapsto \phi(t, r, x) \) is an increasing diffeomorphism on the real line. Then \( u \) is a sub-solution of \( \partial_t u + F(t, x, u, Du, D^2 u) = 0, \ u(0,.) = u_0(.) \) if and only if
\[
v(t, x) = \phi^{-1}(t, u(t, x), x)
\]
is a sub-solution of \( \partial_t v + \phi F(t, x, v, Dv, D^2v) = 0, \ v(0,.) = \phi^{-1}(0, u_0(x), x) \) with
\[
(3.4) \quad \phi F(t, x, r, p, X) = \frac{\dot{\phi}}{\phi'} + \frac{1}{\phi'} F(t, x, \phi, D\phi + \phi' p) + \phi'' p \otimes p + D\phi' \otimes p + p \otimes D\phi' + D^2\phi + \phi' X
\]
where \( \phi \) and all derivatives are evaluated at \( (t, r, x) \). The same statement holds if one replaces the word sub-solution by supersolution throughout.
Proof. ($\Rightarrow$) We show the first implication, i.e. assume $u$ is a subsolution of $\partial_t u + F = 0$ and set $v(t, x) = \phi^{-1}(t, u(t, x), x)$. By definition, $(a, p, X) \in \mathcal{P}^{2+} v(s, z)$ (the parabolic superjet, cf. [3] section 8) iff

$$v(t, x) \leq v(s, z) + a(t - s) + p \cdot (x - z) + \frac{1}{2} (x - z)^T \cdot X \cdot (x - z) + o \left( |t - s| + |x - z|^2 \right)$$

as $(t, x) \to (s, z)$. Since $\phi(t, \cdot, x)$ is increasing,

$$\phi(t, v(t, x), x) \leq \phi(t, *, x)$$

with

$$* = v(s, z) + a(t - s) + p \cdot (x - z) + \frac{1}{2} (x - z)^T \cdot X \cdot (x - z) + o \left( |t - s| + |x - z|^2 \right)$$

and using a Taylor expansion on $\phi$ in all three arguments we see that the right hand side equals

$$\phi(s, v(s, z), z) + \dot{\phi}(s, v(s, z), z)(t - s) + \phi_s(v(s, z), z)a(t - s) + \phi_{ss}(v(s, z), z)p \cdot (x - z)$$

$$+ \frac{1}{2} \phi''_{ss}(v(s, z), z)(x - z)^T \cdot X \cdot (x - z) + D\phi_s(v(s, z), z)(t - s) + \frac{1}{2} (x - z)^T \cdot D^2\phi_s(v(s, z), z)(x - z)$$

$$+ (x - z)^T \cdot (D\phi')(s, v(s, z), z) \otimes p \cdot (x - z)$$

$$+ (x - z)^T \cdot p \otimes (D\phi)'(s, v(s, z), z) \cdot (x - z)$$

$$+ (x - z)^T \cdot \phi''_{ss}(v(s, z), z)p \otimes p \cdot (x - z) + o \left( |t - s| + |x - z|^2 \right)$$

as $(s, z) \to (t, x)$.

Hence,

$$\left( \phi'(s, v(s, z), z) + \dot{\phi}'(s, v(s, z), z)a, D\phi_s(v(s, z), z) + \phi_{ss}(v(s, z), z)p, \right.$$

$$\phi''_{ss}(v(s, z), z)p \otimes p + D(\phi)'(s, v(s, z), z) \otimes p + p \otimes (D\phi)'(s, v(s, z), z) + D^2\phi_s(v(s, z), z) + \phi_{ss}(v(s, z), z)X \left) \right.$$

belongs to $\mathcal{P}^{2+} u(s, z)$ and since $u$ is a subsolution this immediately shows

$$\dot{\phi}(s, v(s, z), z) + \phi'(s, v(s, z), z)a + F(s, z, \phi_s(v(s, z), z), D\phi_s(v(s, z), z) + \phi_{ss}(v(s, z), z)p,$$

$$\phi''_{ss}(v(s, z), z)p \otimes p + D(\phi)'(s, v(s, z), z) \otimes p + p \otimes (D\phi)'(s, v(s, z), z) + D^2\phi_s(v(s, z), z) + \phi_{ss}(v(s, z), z)X \right) \leq 0.$$
Using several times equalities of the type \((f \circ f^{-1})' = f'_f (f^{-1})' = id\) cancels the terms involving \(\phi, \psi\) and their derivatives and we are left with \(F\), i.e.

\[
\phi^{-1} (\phi F) = F.
\]

This finishes the proof. \(\blacksquare\)

**Corollary 8.** [Transformation 2] Assume \(\nu = (\nu_1, \ldots, \nu_d) \subset C_b^{0,2}([0, T] \times \mathbb{R}^d)\) (i.e. continuous, bounded and twice differentiable in the second variable with bounded derivatives). Assume \(\phi = \phi (t, x, r)\) is determined by the ODE

\[
\dot{\phi} = \phi \sum_{j=1}^d \nu_j (t, x) \dot{z}_j^t \equiv \phi \nu (t, x) \cdot \dot{z}_t, \quad \phi (0, x, r) = r.
\]

Define \(L = L (t, x, r, p, X)\) by

\[
L = -Tr [A (t, x) \cdot X] + b (t, x) \cdot p + c (t, x, r);
\]

define also

\[
(3.5) \quad \phi L (t, x, r, X) = -Tr [A (t, x) \cdot X] + \phi b (t, x) \cdot p + \phi c (t, x, r)
\]

where

\[
\phi b (t, x) \cdot p \equiv b (t, x) \cdot p - \frac{2}{\phi'} Tr [A (t, x) \cdot D\phi' \otimes p]
\]

\[
\phi c (t, x, r) \equiv -\frac{1}{\phi'} Tr [A (t, x) \cdot (D^2 \phi)] + \frac{1}{\phi'} b (t, x) \cdot (D\phi) + \frac{1}{\phi'} c (t, x, \phi)
\]

with \(\phi\) and all its derivatives evaluated at \((t, r, x)\). Then

\[
\partial_t w + L (t, x, w, Dw, D^2 w) - w \nu (t, x) \cdot \dot{z} (t) = 0
\]

if and only if \(v (t, x) = \phi^{-1} (t, w (t, x), x)\) satisfies

\[
\partial_t v + \phi L (t, x, v, Dv, D^2 v) = 0.
\]

**Proof.** Obviously,

\[
\phi (t, x, r) = r \exp \left( \int_0^t \sum_{j=1}^d \nu_j (s, x) \dot{z}_j^s \right).
\]

This implies that \(\phi' = \phi / r\) and \(D\phi'\) do not depend on \(r\) so that indeed \(\phi b (t, x)\) defined above has no \(r\) dependence. Also note that \(\phi'' = 0\) and \(\phi / \phi = d \cdot \dot{z} \equiv \sum_{j=1}^d d_j (t, x) \dot{z}_j^t\). It follows, for general \(F\), that

\[
\phi F (t, x, r, p, X) = r d \cdot \dot{z} + \frac{1}{\phi'} F (t, x, \phi, D\phi + \phi' p),
\]

\[
D\phi' \otimes p + p \otimes D\phi' + D^2 \phi + \phi' X)
\]

and specializing to \(F = L - w \nu \cdot \dot{z}\), of the assumed (semi-) linear form, we see that

\[
\phi L = -\frac{1}{\phi'} Tr [A (t, x) \cdot (D\phi' \otimes p + p \otimes D\phi' + D^2 \phi + \phi' X)]
\]

\[
+ \frac{1}{\phi'} b (t, x) \cdot (D\phi + \phi' p) + \frac{1}{\phi'} c (t, x, \phi)
\]

where \(\phi\) and all derivatives are evaluated at \((t, r, x)\). Observe that \(\phi L\) is again linear in \(X\) and \(p\). It now suffices to collect the corresponding terms to obtain \(3.5\). \(\blacksquare\)

We shall need another (outer) transform to remove additive noise.
Lemma 9. [Transformation 3] Let $g \in C \left( [0, T] \times \mathbb{R}^n, \mathbb{R}^d \right)$ and set $\varphi (t, x) = \int_0^t g(s, x) \, dz_s = \sum_{i=1}^d \int_0^t g_i(s, x) \, dz_s^i$. Define

\[
L (t, x, r, p, X) = - \text{Tr} [A (t, x) \cdot X] + b (t, x) \cdot p + c (t, x, r);
\]

\[
L_\sigma (t, x, r, p, X) = - \text{Tr} [A (t, x) \cdot X] + b (t, x) \cdot p + c_\sigma (t, x, r)
\]

with $c_\sigma (t, x, r) = \text{Tr} \left[ A (t, x) \cdot D^2 \varphi (t, x) \right] - b (t, x) \cdot D \varphi (t, x) + c (t, x, r - \varphi (t, x))$

Then $v$ solves

\[
\partial_t v + L (t, x, v, Dv, D^2 v) - g(t, x) \cdot \dot{z} (t) = 0
\]

if and only if $\tilde{v} (t, x) = v (t, x) + \varphi (t, x)$ solves

\[
\partial_t \tilde{v} + L_\varphi (t, x, \tilde{v}, D\tilde{v}, D^2 \tilde{v}) = 0.
\]

Proof. Left to reader. 

3.2. The full transformation. As before, let

\[
L (t, x, r, p, X) := - \text{Tr} [A (t, x) X] + b (t, x) \cdot p + c (t, x, r)
\]

where $A : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{S}^n$, $b : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $f : [0, T] \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$. Let us also define the following (linear, first order) differential operators,

\[
M_k (t, x, u, Du) = \sigma_k (t, x) \cdot Du \quad \text{for} \quad k = 1, \ldots, d_1
\]

\[
M_{d_1+k} (t, x, u, Du) = u \nu_k (t, x) \quad \text{for} \quad k = 1, \ldots, d_2
\]

\[
M_{d_1+d_2+k} (t, x, u, Du) = g_k (t, x) \quad \text{for} \quad k = 1, \ldots, d_3.
\]

The combination of transformations 1, 2 and 3 leads to the following

Proposition 10. Let $z^1 \in C^1 \left( [0, T], \mathbb{R}^{d_1} \right)$, $\sigma = (\sigma_1, \ldots, \sigma_{d_1}) \in C^2_{b} \left( [0, T] \times \mathbb{R}^n, \mathbb{R}^n \right)$ and denote the ODE flow of $dy = \sigma (t, y) \, dz$ with $\psi$, i.e.

\[
\psi (t, x) = \sigma (t, \psi (t, x)) \dot{z}_t^1, \quad \psi (0, x) = x \in \mathbb{R}^n.
\]

Further, let $z^2 \in C^1 \left( [0, T], \mathbb{R}^{d_2} \right)$ and let $\nu = (\nu_1, \ldots, \nu_{d_2})$ be a collection of $C^0_{b} \left( [0, T] \times \mathbb{R}^n, \mathbb{R} \right)$ functions and define $\phi = \phi (t, r, x)$ as solution to the linear ODE

\[
\dot{\phi} = \phi \nu (t, \psi_t (x)) \dot{z}_t^2, \quad \phi (0, r, x) = r \in \mathbb{R}.
\]

Further, let $z^3 \in C^1 \left( [0, T], \mathbb{R}^{d_3} \right)$ and define for given $g = (g_1, \ldots, g_{d_3}) \in C \left( [0, T] \times \mathbb{R}^n, \mathbb{R}^{d_3} \right)$, $\varphi (t, x)$ as the integrator

\[
\varphi (t, x) = \int_0^t \phi g^\psi (s, x) \, d\dot{z}_s^3,
\]

where $\phi g^\psi (t, x) = \frac{1}{\phi' (t, x)} g (t, \psi_t (x))$.

At last, set $z_t := (z_t^1, z_t^2, z_t^3) \in \mathbb{R}^{d_1} \oplus \mathbb{R}^{d_2} \oplus \mathbb{R}^{d_3} \cong \mathbb{R}^d$. Then $u$ is a viscosity solution of

\[
\partial_t u + L (t, x, u, Du, D^2 u) = \Lambda (t, x, u, Du) \dot{z}_t,
\]

\[
u (0, x) = u_0 (x),
\]

iff $\tilde{u} (t, x) = \phi^{-1} (t, u (t, \psi (t, x))) \cdot x + \varphi (t, x)$ is a viscosity solution of

\[
\partial_t \tilde{u} + \tilde{L} (t, x, \tilde{u}, D\tilde{u}, D^2 \tilde{u}) = 0
\]

\[
\tilde{u} (0, x) = u_0 (x)
\]

where $\tilde{L} = \left[ \phi \left( L^\psi \right) \right]_{\varphi}$ is obtained via transformations 1, 2 and 3 (in the given order).

---

\( ^8 \)Since $\phi$ is linear in $r$, there is no $r$ dependence in its derivative $\phi'$.
The proof is now finished.

Remark 11. Transformation 2 and 3 could have been performed in one step, by considering

$$\dot{\phi} = \phi \nu^\psi (t, x) \cdot \dot{z}_t^2 + g^\psi (t, x) \cdot \dot{z}_t^3, \quad \phi (t, r, x) \big|_{t=0} = r.$$  

Indeed, the usual variation of constants formula gives immediately

$$\phi (t, x) = r \exp \left( \int_0^t \nu^\psi (s, x) \, dz_s^2 \right) + \int_0^t e^\left( \int_s^t \nu^\psi (r, x) \, dz_r^2 \right) g^\psi (s, x) \cdot dz_s^3$$  

and one easily sees that transformations 2 and 3 just split above expression in two terms; with the benefit of keeping the algebra somewhat simpler (after all, we want explicit understandings of the transformed equations).

Remark 12. Related to the last remark, generic noise of the form $H (t, x, u) \, dz$ can be removed with this technique. The issue is that the transformed equations quickly falls beyond available viscosity theory (e.g. standard comparison results do no longer apply) cf. [35] [13].

Proof. We first remove the terms driven by $z^1$: to this end we apply transformation 1 with $L (t, x, r, p, X)$ replaced by $L - r \nu \cdot \dot{z}^2 - g \cdot \dot{z}^3$. The transformed solution, $u^1 (t, x) = u (t, \psi_t (x))$, satisfies the equation

$$\left( \partial_t + L^\psi \right) u^1 = - u^1 \nu (t, \psi_t (x)) \cdot \dot{z}_t^2 - g (t, \psi_t (x)) \cdot \dot{z}_t^3 = 0$$  

We then remove the terms driven by $z^2$ by applying transformation 2 with $L^\psi - \psi^\psi \cdot \dot{z}^3$. The transformed solution $u^{12} (t, x) = \phi^{-1} (t, u^1 (t, x), x)$ satisfies the equation with operator

$$\left( \partial_t + \phi \left( L^\psi - \psi^\psi \cdot \dot{z}^3 \right) \right)$$  

i.e.

$$\partial_t u^{12} + \phi (L^\psi) u^{12} - \frac{1}{\phi (c^\psi)} g^\psi \cdot \dot{z}^3 = 0.$$  

It now remains to apply transformation 3 to remove the remaining terms driven by $z^3$. The transformed solution is precisely $\tilde{u}$, as given in the statement of this proposition, and satisfies the equation

$$\left( \partial_t + \left[ \phi \left( L^\psi \right) \right] \phi \right) \tilde{u} = 0.$$  

The proof is now finished. ■

3.3. Rough transformation. We need to understand transformations 1,2,3 when $(z^1, z^2, z^3)$ becomes a rough path, say $z$. There is some “tri-diagonal” structure: (3.7) can be solved as function of $z^1$ alone;

$$d\psi_t (x) = \sigma (t, \psi_t (x)) \, dz_t^1 \text{ with } \psi_0 (x) = x.$$  

(3.8) is tantamount to

$$\phi (t, r, x) = r \exp \left[ \int_0^t \nu (s, \psi_s (x)) \, dz_s^2 \right].$$  

As for $\varphi = \varphi (t, x)$, note that

$$1/\phi' (t, r, x) = \tilde{\phi} (t, x) \equiv \exp \left[ - \int_0^t \nu (s, \psi_s (x)) \, dz_s^2 \right]$$  

so that

$$\varphi (t, x) = \int_0^t \tilde{\phi} (s, x) g (s, \psi_s (x)) \, dz_s^3.$$
Lemma 13. Let \( z \) be a geometric \( p \)-rough path; that is, an element in \( C^{0,p}\text{-var} ([0,T], G^{|p|} (\mathbb{R}^d)) \). Let \( \gamma > p \geq 1 \). Assume
\[
\begin{align*}
\sigma &= (\sigma_1, \ldots, \sigma_d_1) \subset \text{Lip}^\gamma ([0,T] \times \mathbb{R}^n, \mathbb{R}^n), \\
\nu &= (\nu_1, \ldots, \nu_d_2) \subset \text{Lip}^{\gamma - 1} ([0,T] \times \mathbb{R}^n, \mathbb{R}), \\
g &= (g_1, \ldots, g_d_3) \subset \text{Lip}^{\gamma - 1} ([0,T] \times \mathbb{R}^n, \mathbb{R}).
\end{align*}
\]
Then \( \psi, \phi \) and \( \varphi \) depend (in local uniform sense) continuously on \((z^1, z^2, z^3)\) in rough path sense. Under the stronger regularity assumption \( \gamma > p + 2 \); this also holds for the first and second derivatives (with respect to \( x \)) of \( \psi, \psi^{-1}, \phi, \tilde{\phi} \) and \( \varphi \). In particular, we can define \( \psi, \phi \) and \( \varphi \) when \((z^1, z^2, z^3)\) is replaced by a genuine geometric \( p \)-rough path \( z \) and write \( \psi^x, \phi^x, \varphi^x \) to indicate this dependence.

Proof. Given \( z \) one can build a “time-space” rough path \((t, z)\) of \((t, z^1, z^2, z^3)\) since the additionally needed iterated integrals against \( t \) are just Young integrals, cf. [22] Chapter 12. Define
\[
W_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, W_2 = \begin{pmatrix} 0 & \sigma (t, \psi) \\ \sigma (t, \psi) & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, W_4 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \rho \cdot \nu (t, \psi) \\ \rho \cdot \nu (t, \psi) & 0 \\ 0 & 0 & \tilde{\rho} \cdot \nu (t, \psi) \end{pmatrix}, W_4 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \tilde{\rho} \cdot g (t, \psi) \end{pmatrix}.
\]
The assumptions on \( \sigma, \nu \) and \( g \) guarantee that
\[
W = (W_1, \ldots, W_4) : \mathbb{R}^{1 + d_1 + 2d_2 + d_3} \to L (\mathbb{R}^{1 + d_1 + 2d_2 + d_3})
\]
is \( \text{Lip}^\gamma \) (we work with \( \mathbb{R} \) for the time coordinate instead of the closed subset \([0, T] \subset \mathbb{R} \) since by the classic Whitney–Stein extension theorem (see e.g. [42]) we can always find \( \text{Lip}^\gamma \) resp. \( \text{Lip}^{\gamma - 1} \) extensions of \( \sigma, \nu \) and \( g \)). Hence, the “full RDE” (parametrized by \( x \in \mathbb{R}^n \) and \( r \in \mathbb{R} \))
\[
d \begin{pmatrix} t \\ \psi \\ \phi \\ \varphi \end{pmatrix} = W \begin{pmatrix} t \\ \psi \\ \phi \\ \varphi \end{pmatrix} d (t, z) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \sigma (t, \psi) & 0 & 0 \\ 0 & \rho \cdot \nu (t, \psi) & 0 & 0 \\ 0 & 0 & \tilde{\rho} \cdot \nu (t, \psi) & 0 \end{pmatrix} d (t, z)
\]
has a unique global solution\(^9\) (with obvious initial condition that the flows \( \psi, \phi, \varphi \) evaluated at \( t = 0 \) are the identity maps). Further, every additional degree of Lipschitz regularity allows for one further degree of differentiability of the solution flow with corresponding stability in rough path sense, see [36, 37, 22].

4. Semirelaxed limits and rough PDEs

The goal is to understand
\[
\partial_t u + L(t, x, u, Du, D^2 u) = \sum_{i=1}^{d_1} (\sigma_i (t, x) \cdot Du) z^1_i + u \sum_{j=1}^{d_2} \nu_j (t, x) z^2_j + \sum_{k=1}^{d_3} g_k (t, x) z^3_k
\]
in the case when \((z^1, z^2, z^3)\) becomes a rough path. To this end we first give the assumptions on \( L \).

Assumption 1. Assume \( L : [0,T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{S}^n \to \mathbb{R} \) is of the form
\[
(4.1) \quad L(t, x, r, p, X) = -\text{Tr} [A(t, x)X] + b(t, x) \cdot p + c(t, x, r)
\]
\(^9\)although \( W \) fails to be bounded, the particular structure of the system where one can first solve for \( \psi \) and then construct the other quantities by rough integration makes it clear that no explosion can happen. The same situation is discussed in detail in [22] Chapter 11.
with

(1) \( A = a \cdot a^T \) for some \( a : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times n'} \)

(2) \( a : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times n'} \) and \( b : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n \) are bounded, continuous in \( t \) and Lipschitz continuous in \( x \), uniformly in \( t \in [0, T] \)

(3) \( c : [0, T] \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R} \) is continuous, bounded whenever \( r \) remains bounded, and with a lower Lipschitz bound, i.e. \( \exists C < 0 \) s.t.

\[
(4.2) \quad c(t, x, r) - c(t, x, s) \geq C(r - s) \text{ for all } r \geq s, \quad (t, x) \in [0, T] \times \mathbb{R}^n.
\]

Assumption 1 guarantees that a comparison result holds for \( \partial_t + L \); see the appendix and [15, Section V, Lemma 7.1] or [6] for details. Further we need the assumptions on the coefficients in \( \Lambda \).

**Assumption 2. Assume that**

\[
\sigma = (\sigma_1, \ldots, \sigma_{d_1}) \subset Lip^\gamma ([0, T] \times \mathbb{R}^n, \mathbb{R}^n),
\]

\[
\nu = (\nu_1, \ldots, \nu_{d_2}) \subset Lip^{\gamma - 1} ([0, T] \times \mathbb{R}^n, \mathbb{R}),
\]

\[
g = (g_1, \ldots, g_{d_3}) \subset Lip^{\gamma - 1} ([0, T] \times \mathbb{R}^n, \mathbb{R}).
\]

Let us now replace the (smooth) driving signals of the earlier sections by a \( d = (d_1 + d_2 + d_3) \)-dimensional driving signal \( z^\epsilon \) and impose convergence to a genuine geometric \( p \)-rough path \( z \), that is, in the notation of [22, Chapter 14]

\[
z \in C^{0,p\text{-var}} ([0, T], G[p] ([d_1 + d_2 + d_3])).
\]

We can now prove our main result.

**Theorem 14. Let** \( p \geq 1 \). Assume \( L \) fulfills assumption [4] and the coefficients of \( \Lambda = (\Lambda_1, \ldots, \Lambda_{d_1 + d_2 + d_3}) \) fulfill [5] for some \( \gamma > p + 2 \). Let \( u_0 \in BUC (\mathbb{R}^n) \) and \( z \in C^{0,p\text{-var}} ([0, T], G[p] (\mathbb{R}^d)) \). Then there exists a unique \( u = u^\epsilon \in BUC ([0, T] \times \mathbb{R}^n) \) such that for any sequence \((z^\epsilon)_n \subset C^1 ([0, T], \mathbb{R}^d)\) such that \( z^\epsilon \to z \) in \( p \)-rough path sense, the viscosity solutions \((u^\epsilon) \subset BUC ([0, T] \times \mathbb{R}^n)\) of

\[
\dot{u}^\epsilon + L (t, x, u^\epsilon, Du^\epsilon, D^2 u^\epsilon) = \sum_{k=1}^d \Lambda_k (t, x, u^\epsilon, Du^\epsilon) \dot{z}^k, \quad u^\epsilon (0, \cdot) = u_0 (\cdot),
\]

converge locally uniformly to \( u^\epsilon \). We write formally [11]

\[
(4.3) \quad du + L (t, x, u, Du, D^2 u) dt = \Lambda (t, x, u, Du) dz, \quad u(0, \cdot) = u_0 (\cdot).
\]

Moreover, we have the contraction property

\[
|u^\epsilon - \hat{u}^\epsilon|_{\infty; \mathbb{R}^n \times [0, T]} \leq e^{CT} |u_0 - \hat{u}_0|_{\infty; \mathbb{R}^n}
\]

(C given by (4.2)) and continuity of the solution map \((z, u_0) \mapsto u^\epsilon\) from

\[
C^{0,p\text{-var}} ([0, T], G[p] (\mathbb{R}^d)) \times BUC (\mathbb{R}^n) \to BUC ([0, T] \times \mathbb{R}^n).
\]

**Proof.** We shall write \( \psi^\epsilon, \phi^\epsilon, \varphi^\epsilon \) for the objects (solutions of rough differential equations and integrals) built upon \( z \), as discussed in the last section (lemma [13]) and also write \( \psi^\epsilon, \phi^\epsilon, \varphi^\epsilon \) when the driving signal is \( z^\epsilon \). Recall from [3,6] that \( \Lambda = (\Lambda_1, \ldots, \Lambda_d) \) is a collection of linear, first order differential operators. We use the same technique of “rough semi-relaxed limits” as in [6]: the key remark being that

\[
[ \phi^\epsilon (L \psi^\epsilon) ] \varphi^\epsilon \to [ \phi^\epsilon (L \psi^\epsilon) ] \varphi^\epsilon
\]

holds locally uniformly, as function of \((t, x, r, p, X)\). Secondly, applying the transformations, the (classical) viscosity solutions \( u^\epsilon \) can be used to define a new function \( \tilde{u}^\epsilon \) by setting

\[
\tilde{u}^\epsilon (t, x) = (\phi^\epsilon)^{-1} (t, u^\epsilon (t, \psi^\epsilon (t, x)), x) + \varphi^\epsilon (t, x);
\]

---

10 The regularity assumptions on the vector fields with respect to \( t \) could be relaxed here, cf. remark 5

11 The intrinsic meaning of this “rough” PDE is discussed in definition 15 below.
and proposition 10 in section 3 show that \( \bar{u}^\varepsilon \) is a (classical) viscosity solution of
\[
d\bar{u}^\varepsilon + \left[ L^\varepsilon (\bar{u}^\varepsilon) \right]_{\bar{u}^\varepsilon} (t, x, \bar{u}^\varepsilon, D\bar{u}^\varepsilon, D^2\bar{u}^\varepsilon) = 0.
\]
Now one uses the comparison result for parabolic viscosity solutions (as given in the appendix) to conclude that there exists a constant \( C > 0 \) such that
\[
\sup_{\varepsilon \in (0, 1)} \sup_{t \in [0, T]} |\bar{u}^\varepsilon (t, x)| < (1 + |u_0|_\infty) C T;
\]
This in turn implies (thanks to the uniform control on \( \varphi^\varepsilon, \phi^\varepsilon, \psi^\varepsilon \) as \( \varepsilon \to 0 \)) by using the rough path representations discussed in section 3.3 that \( \bar{u}^\varepsilon \) remains locally uniformly bounded (as \( \varepsilon \to 0 \)). Together with the stability of (classical) viscosity solutions (c.f. 6) the proof is finished.

The reader may wonder if \( u \) is the solution in a sense beyond the “formal” equation
\[
du + L (t, x, u, Du, D^2u) dt = \Lambda (t, x, u, Du) dz, \ u (0, \cdot) \equiv u_0 (\cdot).
\]
Inspired by the definition given by Lions–Souganidis in 35 we give

**Definition 15.** \( u \) is a solution to the rough partial differential equation (4.3) if and only if \( \bar{u} (t, x) = (\phi^\varepsilon)^{-1} (t, u (t, \psi^\varepsilon (t, x)), x) + \varphi^\varepsilon (t, x) \)
\[
d\bar{u} + \tilde{L} (t, x, \bar{u}, D\bar{u}, D^2\bar{u}) = 0, \ \bar{u} (0, \cdot) = u_0 (\cdot)
\]
in viscosity sense where
\[
\tilde{L} = \left[ L^\varepsilon (\varphi^\varepsilon) \right]_{\varphi^\varepsilon}.
\]

**Corollary 16.** Under the assumptions of Theorem 14 there exists a unique solution in \( \text{BUC} ([0, T] \times \mathbb{R}^n) \) to the RPDE (4.3).

**Proof.** Existence is clear from theorem 14. Uniqueness is inherited from uniqueness to the Cauchy problem for \( \left( \partial_t + \tilde{L} \right) = 0 \) which follows from a comparison theorem for parabolic viscosity solution (c.f. Theorem 28 in the appendix). ■

## 5. RPDEs and Variational Solutions of SPDEs

A classic approach to second order parabolic SPDEs (especially the Zakai equation from nonlinear filtering) is the so-called \( L^2 \)-theory for SPDEs, due to Pardoux, Rozovskii, Krylov et. al., cf. 31, 39, 40. Sufficient conditions for existence, uniqueness in this setting are classical (brief recalls are given below). On the other hand, our main theorem on RPDEs driven by rough paths, theorem 14 can be applied with almost every realization of Enhanced Brownian motion, \( B (\omega) \), that is Brownian motion enhanced with Lévy’s stochastic area, the standard example of a (random) rough path. The aim of this section is to show that, whenever the (not too far from optimal) assumptions of both theories are met, the resulting solutions coincide. We focus on the model case of linear SPDEs; i.e.

\[
L (t, x, r, p, X) = -\text{Tr} [A(t, x)X] + b (t, x) \cdot p + c (t, x) r + f (t, x).
\]

### 5.1. \( L^2 \) Solutions

Given is a filtered probability space \( (\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P}) \), which satisfies the usual conditions and carries a \( d \)-dimensional Brownian motion \( B \). Denote with \( H^m (\mathbb{R}^n) \) the usual Sobolev space, i.e. the subspace of \( L^p (\mathbb{R}^n) \) consisting of functions whose generalized derivatives up to order \( m \) are in \( L^p (\mathbb{R}^n) \). Equipped with the norm
\[
|f|_{H^m (\mathbb{R}^n)} = \left( \sum_{0 \leq k \leq m} |\partial_{i_1} \cdots \partial_{i_k} f|_{L^2 (\mathbb{R}^n)}^p \right)^{1/2}
\]
\( H^m (\mathbb{R}^n) \) becomes a separable Hilbert space and the variational approach makes use of the triple
\[
(H^m (\mathbb{R}^n)) \hookrightarrow L^2 (\mathbb{R}^n) \subset (L^2 (\mathbb{R}^n))^* \hookrightarrow (H^m (\mathbb{R}^n))^*. 
\]
We make the following assumptions on the coefficients of $L$ and $\Lambda$.

**Assumption 3.** For $i, j \in \{1, \ldots, n\}$, $k \in \{1, \ldots, d\}$

1. $a^{ij}, b_i, c, f$ as well as $\sigma^k_i, \nu^k_i, g$ are elements of $L^2([0, T] \times \mathbb{R}^n, \mathbb{R})$ and $\sigma^k_i, \nu^k_i, g$ have one and $a^{ij}, \sigma^k_i$ have two, bounded (uniformly in $t$) continuous derivatives in space,
2. $f, g \in L^2([0, T] \times \mathbb{R}^n)$,
3. $A = a \cdot d^2$, and $\exists \lambda > 0$ such that $\forall t \in [0, T]$ $z^T \cdot A(t, x) \cdot z \geq \lambda |z|^2$ $\forall x, z \in \mathbb{R}^n$.

A $L^2$-solution is then defined as follows

**Definition 17.** Let $u_0 \in L^2(\mathbb{R}^n)$ and assume $L$ and $\Lambda$ fulfill assumption 3. We say that a $L^2(\mathbb{R}^n)$-valued, strongly continuous and $(\mathcal{F}_t)$-adapted process $u = (u_t)_{t \in [0,T]}$ is a $L^2$-solution of

\begin{align}
(5.2) \quad \frac{du}{dt} &= L u dt + \Lambda u dB_t \\
\quad u(0, \cdot) &= u_0(\cdot)
\end{align}

if

1. we have $\mathbb{P}$-a.s. that $u_t \in H^1(\mathbb{R}^n)$ for a.e. $t \in [0, T]$ and $\mathbb{P} \left( \int_0^T |u_t|^2_{H^1(\mathbb{R}^n)} \, dt < \infty \right) = 1$,
2. $\forall \varphi \in C_0^\infty(\mathbb{R}^n)$ we have

\begin{align}
(5.3) \quad \langle u_t, \varphi \rangle - \langle u_0, \varphi \rangle = \int_0^t \langle u_r, \tilde{L}^* \varphi \rangle \, dr + \sum_{k=1}^d \int_0^t \langle u_r, \Lambda_k^* \varphi \rangle \, dB_r^k (d\lambda \otimes d\mathbb{P}) - a.s.,
\end{align}

here $\tilde{L} \varphi := L \varphi + \frac{1}{2} \sum_{k=1}^d A_k \Lambda_k \varphi$.

**Remark 18.** The difference with the standard definition, cf. [11] Chapter IV, p130, is that we additionally assume enough regularity on the coefficients for the existence of the adjoint of $\tilde{L}$ and to switch between divergence and non-divergence form.

**Theorem 19.** Under assumption 3 there exists a unique $L^2$-solution of $(5.2)$.

**Proof.** The standard variational approach as for example presented in [11] Chapter 4, Theorem 1 (see also [31] [39] [40]) guarantees the existence of an $L^2(\mathbb{R}^n)$-valued, strongly continuous in $t$, $(\mathcal{F}_t)$-adapted process $(u_t)$ which fulfills part (2) of definition 17 as well as that $\forall \varphi \in C_0^\infty(\mathbb{R}^n)$ we have $\mathbb{P}$-a.s. (using the Einstein summation convention)

\[
\langle u_t, \varphi \rangle - \langle u_0, \varphi \rangle = \int_0^t \left( -\langle \tilde{a}^{ij} \partial_j u_r, \partial_i \varphi \rangle + \langle \tilde{\sigma}^r_i \partial_i u_r + \tilde{\nu}^r_i (r) u_r + \tilde{f}_r(r), \varphi \rangle \right) \, dr \\
+ \int_0^t \langle \sigma^k_i (r) \partial_j u_r + \nu^k_i (r) u_r + g_k (r), \varphi \rangle \, dB_r^k
\]
where

\[ \tilde{a}^{ij} = a^{ij} + \frac{1}{2} \sum_{k=1}^{d} \sigma_k^i \sigma_k^j \]

\[ \tilde{b}^i = b^i + \frac{1}{2} \sum_{k=1}^{d} \left( \sigma_k^i \left( \partial_j \sigma_k^j \right) + 2 \nu_k \sigma_k^i \right) \]

\[ \tilde{c} = c + \frac{1}{2} \sum_{k=1}^{d} \left( \sigma_k^i \left( \partial_k \nu_k \right) + \nu_k^2 \right) \]

\[ \tilde{f} = f + \frac{1}{2} \sum_{k=1}^{d} \left( \sigma_k^i \partial_k g + \nu_k g_k + g \right) \]

and

\[ \tilde{b} = \tilde{b} - \left( \partial_j \tilde{a}^{ij} \right) \]

(i.e. \( \tilde{a} \) and \( \tilde{b} \) are the coefficients that appear in \( \tilde{L} \) due to the switch from Ito to Stratonovich integration and \( \tilde{b} \) results from the switch to divergence form). Now using integration by parts we can rewrite the above divergence form into the adjoint formulation as required by definition \( 17 \).

We can now prove the main result of this section which identifies the RPDE solution with the classic \( L^2 \)-solution whenever both exist.

**Proposition 20.** Assume that \( L \) and \( \Lambda \) fulfill assumption \( 3 \) as well as the assumptions of Theorem \( 14 \). If we denote with \( u^B \) the RPDE solution given in Theorem \( 14 \) driven by Enhanced Brownian motion \( B \) then \( \left( u^B_t \right)_{t \geq 0} \) is a (and hence a version of the unique) \( L^2 \)-solution.

**Proof.** Step 1. Assume additionally to assumption \( 3 \) that all coefficients appearing in \( L \) and \( \Lambda \) are in \( C_0^\infty ((0, T) \times \mathbb{R}^n) \) (in step 2 we get rid of this assumption). Define the adapted, piecewise linear approximation \( B^n \) to \( B \) as

\[ B^n_t = B^n_{t_{k-1}} + n \left( \frac{t - t_{k-1}}{T} \right) (B_{t_k} - B_{t_{k-1}}) \]

for \( t \in [t_k, t_{k+1}) \) with \( t_k = k \frac{T}{n} \). For every \( n \in \mathbb{N} \) we denote with \( u(B_n) \) the \( L^2 \)-solutions of \( 5.2 \) where \( B \) is replaced by \( B^n \) and with \( u \) the unique \( L^2 \)-solution of theorem \( 19 \). Further, denote with \( B^n = S_2 (B^n) \) the rough path lift of \( B^n \) and with \( u^{B^n} \) resp. \( u^B \) the viscosity solution for the random rough path \( B^n \) resp. \( B \) given in theorem \( 14 \). For \( \epsilon > 0 \) and \( n \in \mathbb{N} \) the regularity of the coefficients allows to identify \( u^{B_n, \epsilon} \) with \( u(B^n, \omega) \) (both are the unique, bounded, smooth solutions of a parabolic PDE with smooth coefficients which depend on \( \omega \)). The Wong–Zakai result in \( 29 \) Theorem 2.1] (all coefficients are smooth and additionally \( f(t, \cdot), g(t, \cdot) \) have compact support, hence are \( H^5 (\mathbb{R}^n) \)-valued), tells us that

\[ u(B_n) \to u(B) \in L^2 \left( [0, T], H^1 \right) \text{ a.s.} \]

where the convergence takes place in \( L^2 \left( [0, T], H^1 \right) \) and hence also in \( L^2 \left( [0, T] \times \mathbb{R}^n \right) \) (much more is true here of course). On the other hand, we know that a.s. \( B^n \) converges to \( B \) in rough path metric (see \( 21 \)) and from \( 14 \) we conclude that

\[ u^{B^n} \to u^B \text{ a.s.} \]

locally uniformly on \([0, T] \times \mathbb{R}^n\). It is an now easy matter to identify the \( L^2 \) and loc. uniform limit,

\[ u(B) = u^B \text{ a.s.} \]

(viewing \( u(B) \) a.s. \( C ([0, T], L^2 (\mathbb{R}^n)) \)-valued, this means that for a.e. \( \omega \), \( \forall t \in [0, T], u_t (B(\omega)) = u^B_t(\omega) \) as equality in \( L^2 (\mathbb{R}^n) \); in particular, \( u^{B(\omega)} \) constitutes a continuous version in \( t, x \); once more much more is true here). Further we know that \( u(B) \in L^2 \left( [0, T], H^1 (\mathbb{R}^n) \right) \) \( \mathbb{P} \)-a.s. and
hence fulfills definition \[17\]. Hence, we conclude that \((u^B_\epsilon)\) is the unique \(L^2\)-solution of \((5.2)\) (strictly speaking, a continuous function in the equivalence class).

**Step 2.** For every \(\epsilon > 0\) truncate all coefficients outside a ball of radius \(\epsilon^{-1}\) and smooth by convolution with a mollifier \(m^\epsilon\) (viz. \(m^\epsilon(t,x) = \epsilon^{-n+1} m(\frac{t}{\epsilon}, \frac{x}{\epsilon})\) where \(m\) has compact support, is non-negative and has total mass one) to arrive at the new operators \(L^\epsilon\) and \(\Lambda^\epsilon\). It is easy to see that \(L^\epsilon, \Lambda^\epsilon\) again fulfill assumption \[3\], hence Theorem \[\ref{LinearOperatorStability}\] applies and gives existence and uniqueness of the associated \(L^2\)-solution. Denote with \(u^{\epsilon,B}\) the associated RPDE solution\[17\] (with random rough path \(B\)) and note that by step 1, \(u^{\epsilon,B}\) coincides with the \(L^2\)-solution. We now claim that

\[ u^{\epsilon,B} \to u^B \text{ a.s.} \]

uniformly on compacts. To see this, note that by the construction given in theorem \[\ref{UniformContinuityAndCompactness}\] \(u^B\) is the composition of a viscosity solution \(\tilde{u}^\epsilon\) with rough path flows and \(\tilde{u}^\epsilon\) itself is a solution of a linear PDE

\[ \partial v^\epsilon_t + \mathcal{T} (t, x, v^\epsilon, Dv^\epsilon, D^2v^\epsilon) = 0; \]

the precise form of \(\mathcal{T}\) is as in Theorem \[\ref{UniformContinuityAndCompactness}\] given by the transformation via rough path flows, that is

\[ \mathcal{T} = \left[ \phi^{\epsilon,B} \left( (L^\epsilon)^{(\psi^{\epsilon,B})} \right) \right]_{\varphi^{\epsilon,B}}. \]

(\(\phi^{\epsilon,B}, \psi^{\epsilon,B}\) and \(\varphi^{\epsilon,B}\) denote the rough flows associated with the truncated and mollified vector fields). Further we claim that the truncated and mollified \(\text{Lip}\gamma\)-vector fields (appearing in \(\Lambda^\epsilon\)) converge locally uniformly with locally uniform \(\text{Lip}\gamma\) bounds: given \(V \in \text{Lip}\gamma\) denote \(V^\epsilon\) as the vector field given by truncation outside radius \(\epsilon^{-1}\) and convolution of \(V\) with \(m^\epsilon\). Of course, \(V^\epsilon\) converges locally uniformly to \(V\) and (cf. \[12, p123, p159\]) locally uniform \(\text{Lip}\gamma\) bounds are readily seen to hold true for every \(V^\epsilon, \epsilon > 0\). An interpolation argument then shows, locally, convergence in \(\text{Lip}\gamma\) for \(\gamma' < \gamma\) (we only need \(\gamma' = \gamma - 1\)). Given a geometric \(p\)-rough path \(z\) with \(p < \gamma\) it then follows from \[\cite{22}\] Corollary 10.39 (together with the remark that the \(\|\cdot\|_{\text{Lip}^{\gamma-1}}\)-norm can be replaced by the local Lipschitz norm, restricted to a big enough ball in which both RDE solutions live) that the (unique) RDE solutions (started at a fixed point) to \(dy^\epsilon = V^\epsilon(y^\epsilon') dz\) converge as \(\epsilon \to 0\) to the (unique) RDE solution of \(dy = V(y) dz\).

As in \[\cite{22}\] Theorem 11.12 and Theorem 11.13] this convergence can be raised to the level of \(C^k\)-diffeomorphisms, provided \(V\) is assumed to be \(\text{Lip}^{\gamma+k-1}\) for \(k \in \mathbb{N}\) – the case of interest to us is given by \(\gamma > 4\) and \(p \in (2, \gamma - 2)\) which results in stability of the flow seen as \(C^2\)-diffeomorphism. This shows that \(\mathcal{T} \to \left[ (L^\epsilon)^{(\psi^B)} \right]_{\varphi^B}\) as \(\epsilon \to 0\) uniformly on compacts and the stability properties of viscosity solutions guarantee (the same argument as given in theorem \[\ref{UniformContinuityAndCompactness}\]) that \(u^\epsilon \to u\), hence \(u^{\epsilon,B} \to u^B\) (loc. uniformly on \([0, T] \times \mathbb{R}^n\) a.s. From the first step it follows that \(u^{\epsilon,B}\) is the unique \(L^2\)-solution, i.e. \(\forall t \in (0, T]\)

\[ \left< u^{\epsilon,B}_t, \varphi \right> - \left< u_0, \varphi \right> = \int_0^t \left< u^{\epsilon,B}_t, (L^\epsilon)^* \varphi \right> dt + \sum_{k=1}^d \int_0^t \left< u^{B,\epsilon}_t, (\Lambda^\epsilon)^k \varphi \right> dB^k_t \]

Sending \(\epsilon \to 0\) in above equality shows that point (2) of definition \[17\] is fulfilled. Now for every \(\epsilon > 0\), classic variational arguments, see \[\cite{11}\] Chapter 4, Theorem 1, p131, show that there exists a constant \(C^\epsilon > 0\) which depends only on \(T, n, d\) and \(\sup_{t,x,i,j,k} \left( |a^{\epsilon,ij}|, |b^{\epsilon,i}|, |\sigma^{\epsilon,ij}|, |\nu_k^{\epsilon,i,j}| \right) \)

\[16\] The mollification uses values of the coefficients for \(t\) outside \([0, T]\) therefore we simply define the coefficients for \(t \in \mathbb{R} \setminus [0, T]\) by constant continuation.

\[17\] with abuse of notation we identify the operators \(L^\epsilon, L\) (and similarly \(\Lambda, \Lambda^\epsilon\)) with functions on \([0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{S}^n \to \mathbb{R}\) as required in the viscosity setting in the obvious way.
(which is finite by assumption 3) s.t.

\[
\mathbb{E} \left[ \sup_{t \in [0,T]} |u_t^B|_{L^2(\mathbb{R}^n)}^2 + \int_0^T |u_t^B|_{H^1(\mathbb{R}^n)}^2 \, dt \right] 
\leq C \epsilon \left[ |u_0|_{L^2(\mathbb{R}^n)}^2 + \mathbb{E} \int_0^T \left( |f_t^B|_{H^{-1}(\mathbb{R}^n)}^2 + \sum_{k=d_1+d_2+1}^d \left( |g_t^k|_{L^2(\mathbb{R}^n)}^2 \right) \right) \, dt \right].
\]

By the estimate

\[ |f_t^B|_{H^{-1}(\mathbb{R}^n)} \lesssim |f_t^B|_{L^2(\mathbb{R}^n)} = \left| (f(\cdot,1_{|.|<\epsilon^{-1}}) \ast m_t^B) \right|_{L^2(\mathbb{R}^n)} \leq |f(\cdot,1_{|.|<\epsilon^{-1}})|_{L^2(\mathbb{R}^n)} \leq |f|_{L^2(\mathbb{R}^n)} \]

(and similarly for \( g^B \)), the right-hand side can be uniformly bounded in \( \epsilon \), leading to the desired regularity properties of \( u^B \) (as required by point (2) in definition 17).

**Remark 21.** Classical \( L^2 \)-theory of SPDEs gives, with probability one, \( u(t,\cdot,\omega) \in L^2(\mathbb{R}^n) \) for all \( t \in [0,T] \) and then in the Sobolev space \( H^1(\mathbb{R}^n) \) for a.e. \( t \in [0,T] \). It is not clear, in general, if a continuous (in \( t,x \)) version of \( u \) exists. Under further regularity assumptions one finds that \( u(t,\cdot,\omega) \) takes values in higher Sobolev spaces \( H^l(\mathbb{R}^n) \), \( l = 1,2,... \). Since Sobolev embedding theorems are dimension-dependent (recall \( H^l(\mathbb{R}^n) \subset C(\mathbb{R}^n) \) when \( l > n/2 \)) the regularity required for a continuous version will grow with the dimension \( n \). In contrast, our approach effectively gives sufficient conditions, without dimension dependence, under which \( L^2 \)-solutions admit continuous versions. We note that such considerations also motivated Krylov’s \( L^p \)-theory [32] p185

5.2. A \( L^1_{loc} \)-solution. Theorem [4] applied with enhanced Brownian motion provides the unique RPDE viscosity solution even if

1. \( L \) is degenerate elliptic,
2. \( u_0 \in BUC(\mathbb{R}^n) \).

Under such conditions one can not hope for the existence of a \( L^2 \)-solution: the degeneracy of \( L \) does not lead to \( H^1 \)-regularity in space and the initial data \( u_0 \) does not fit into a \( L^2 \)-theory (in fact \( L^p \) for \( 1 \leq p < \infty \), e.g. by taking \( u_0 \equiv 1 \); however one could consider weighted Sobolev spaces). Hence, our new assumptions read,

**Assumption 4.** For \( i,j \in \{1,\ldots,n\}, k \in \{1,\ldots,d\} \),

1. \( a^{ij}, b_i, c, f \) as well as \( \sigma^i_k, \nu_k, g_k \) are in \( C_b([0,T] \times \mathbb{R}^n) \) and \( \sigma^i_k, \nu_k, g \) have one and \( a^{ij}, \sigma^i_k \) have two, continuous, bounded (uniformly in \( t \)) derivatives in space,
2. \( f,g \in L^2([0,T] \times \mathbb{R}^n) \),
3. \( A = a \cdot a^T \).

Motivated by above remarks we give the following definition.

**Definition 22.** By an \( L^1_{loc} \)-solution we mean a \( L^1_{loc}(\mathbb{R}^n) \)-valued strongly continuous \((\mathcal{F}_t)\)-adapted process \( u = \{u_t\}_{t \in [0,T]} \) such that \( \forall \varphi \in C^\infty_0(\mathbb{R}^n) \) we have

\[
\langle u, \varphi \rangle - \langle u_0, \varphi \rangle = \int_0^T \left< u_t, \left( \bar{L}^* \right) \varphi \right> \, dt + \sum_{k=1}^d \int_0^T (u, (\Lambda_k)^* \varphi) \, dB_t^k + (d\lambda \otimes d\mathbb{P}) - a.s.
\]

where \( \bar{L} \varphi := L \varphi + \frac{1}{2} \sum_{k=1}^d \Lambda_k \Lambda_k \varphi \).

**Remark 23.** Above definition comes of course with a caveat: \( L^1_{loc} \) is not a Banach space and the standard uniqueness results do not apply. However, note that we could have given a more restrictive definition of a weak solution by using weighted \( L^p \) or Orlicz-spaces instead of \( L^1_{loc} \). Either way, we are not aware of a uniqueness theory for degenerate SPDEs in either such a setup which seems to be a challenging question in its own right. Below we only give the existence for \( L^1_{loc} \)-weak solutions by showing that the viscosity RPDE solution is a \( L^1_{loc} \)-solution.
**Proposition 24.** Let $B$ be a $d$-dimensional Brownian motion, $u_0 \in \text{BUC}$ and assume $L,A$ fulfill the conditions of theorem [4]. Then $(u^B_t)_{t \in [0,T]}$ is a $L^1_{\text{loc}}$-solution.

**Proof.** For $\epsilon > 0$ consider the elliptic operator $L^\epsilon := L + \epsilon \sum_{i=1}^d \partial_i^2$ and truncate and mollify $u_0$ to get $u_0^\epsilon \in C^\infty$ s.t. $u_0^\epsilon \to u_0$ uniformly on compacts in $\mathbb{R}^n$. Proposition [20] shows for $\epsilon > 0$ that the RPDE solution $u^\epsilon \in \text{BUC}([0,T] \times \mathbb{R}^n)$ associated with $(L^\epsilon, A, u_0^\epsilon, B)$ is (in the equivalence class of) the unique $L^2$-solution; especially $u^\epsilon$ is a $L^1_{\text{loc}}$-solution and therefore fulfills $(d\lambda \otimes d\mathbb{P}) - \text{a.s. that}$

$$
\langle u^\epsilon, \varphi \rangle - \langle u_0^\epsilon, \varphi \rangle = \int_0^\infty \int_0^\infty \left( \int_0^t \left( \int_0^\infty \sum_{i=1}^d \partial_i^2 \varphi \right) dr + \int_0^\infty \sum_{i=1}^d \partial_i^2 \varphi \right) dr + \int_0^\infty \sum_{k=1}^d \int_0^\infty \langle u^\epsilon, \Lambda^\epsilon \varphi \rangle dB_k^r
$$

Conclude by noting that the locally uniform converge $u^\epsilon \to u$ on $[0,T] \times \mathbb{R}^n$ follows from the stability properties of standard viscosity solutions ($u^\epsilon$ is given by a transformation with RDE flows as a standard viscosity solution with an extra term including a Hessian which vanishes as $\epsilon \to 0$). 

6. **Applications to stochastic partial differential equations**

We now discuss some further applications of theorem [4] when applied to a stochastic driving signal, i.e. by taking $z$ to be a realization of a continuous semi-martingale $Z$ and its stochastic area, say $Z(\omega) = (Z, A)$; the most prominent example being Brownian motion and Lévy’s area.

**Remark 25.** [Itô versus Stratonovich] Note that similar SPDEs in Itô-form need not be, in general, well-posed. Consider the following (well-known) linear example

$$
\dot{u} = u_{xx} dt + \lambda u_{xx} dt, \quad \lambda \geq 0.
$$

A simple computation shows that $v(x,t) := u(x-B_t,t)$ solves the (deterministic) PDE $\dot{v} = (\lambda - 1/2)v_{xx}$. From elementary facts about the heat equation we recognize that, for $\lambda < 1/2$, this equation, with given initial data $v_0 = u_0$, is not well-posed. In the (Itô-) SPDE literature, starting with [39], this has led to coercivity conditions, also known as super-parabolicity assumptions, in order to guarantee well-posedness.

**Remark 26.** [Regularity of noise coefficients] Applied in the semimartingale context (finite $p$-variation for any $p > 2$) the regularity assumption of theorem [4] reads $\text{Lip}^{\delta + \epsilon}$, $\epsilon > 0$. While our arguments do not appear to leave much room for improvement we insist that working directly with Stratonovich flows (rather than rough flows) will not bring much gain: the regularity requirements are essentially the same. Itô flows, on the other hand, require one degree less in regularity. In turn, there is a potential loss of well-posedness and the resulting SPDE is not robust as a function of its driving noise (similar to classical Itô stochastic differential equations).

**Remark 27.** [Space-time regularity of SPDE solutions] Since $u(t,x)$ is the image of a (classical) PDE solution under various (inner and outer) flows of diffeomorphisms, it suffices to impose conditions on the coefficients on $L$ which guarantee that existence of nice solutions to $\partial_t + \left[ \partial_x^\alpha (L^\epsilon \partial_x^\alpha ) \right] u$. For instance, if the driving rough path $z$ has $1/p$-Hölder regularity, it is not hard to formulate conditions that guarantee that the rough PDE solutions is an element of $C^{1/p, 2+\delta}$ for suitable $\delta > 0$. Indeed, it is sheer matter of conditions-book-keeping to solve $\partial_t + \left[ \partial_x^\alpha (L^\epsilon \partial_x^\alpha ) \right] u$ under (known and sharp) conditions in Hölder spaces, cf. [33] Section 9, p. 140], with $C^{1+\delta/2, 2+\delta}$ regularity. Unwrapping the change of variables will not destroy spatial regularity (thanks to sufficient smoothness of our diffeomorphisms for fixed $t$) but will most definitely reduce time regularity to $1/p$-Hölder.

We now turn to the applications. Throughout we prefer to explain the main point rather than spelling out theorems under obvious conditions; the reader with familiarity with rough
Rough path theory will realize that formulating and proving such statements follows easily from well-known results once continuous dependence in rough path norm is established (which is done in Theorem 14).

(Approximations) Any approximation result to a Brownian motion \( B \) (or more generally, a continuous semimartingale) in rough path topology implies a corresponding (weak or strong) limit theorem for such SPDEs: it suffices that an approximation to \( B \) converges in rough path topology; as is well known (e.g. [22, Chapter 13] and the references therein) examples include piecewise linear, mollifier, and Karhunen-Loève approximations, as well as (weak) Donsker type random walk approximations [2]. The point being made, we shall not spell out more details here.

(Support results) In conjunction with known support properties of \( B \) (e.g. [34] in \( p \)-variation rough path topology or [16] for a conditional statement in Hölder rough path topology) continuity of the SPDE solution as a function of \( B \) immediately implies Stroock–Varadhan type support descriptions for such SPDEs. In the linear, Brownian noise case, approximations and support of SPDEs have been studied in great detail [28, 27, 25, 24, 26].

(Large deviation results) Another application of our continuity result is the ability to obtain large deviation estimates when \( B \) is replaced by \( \varepsilon B \) with \( \varepsilon \to 0 \); indeed, given the known large deviation behaviour of \( (\varepsilon B, \varepsilon^2 A) \) in rough path topology (e.g. [34] in \( p \)-variation and [19] in Hölder rough path topology) it suffices to recall that large deviation principles are stable under continuous maps.

(SPDEs with non-Brownian noise) Yet another benefit of our approach is the ability to deal with SPDEs with non-Brownian and even non-semimartingale noise. For instance, one can take \( z \) as (the rough path lift of) fractional Brownian motion with Hurst parameter \( 1/4 < H < 1/2 \), cf. [8] or [17], a regime which is “rougher” than Brownian and notoriously difficult to handle, or a diffusion with uniformly elliptic generator in divergence form with measurable coefficients; see [20]. Much of the above (approximations, support, large deviation) results also extend, as is clear from the respective results in the above-cited literature.

(SPDEs with higher level rough paths without extra effort) In contrast to the approach by Gubinelli-Tindel [23], no extra effort is necessary when \( z \) is a higher level rough path (the prominent example being fractional Brownian motion with \( H \in (1/4, 1/3) \)).

(Approximation of Wong-Zakai type with modified drift term) For brevity let us write \( L, \Lambda \) and \( \Lambda_k \) instead of \( L(t, x, u, Du, D^2 u) \), \( \Lambda(t, x, u, Du) \) and \( \Lambda_k(t, x, u, Du) \) in this section and consider the SPDE

\[
du + Ldt = \sum_k \Lambda_k \circ dZ^k.
\]

Equivalently, we write

\[
du + Ldt = \Lambda dZ
\]

where \( Z \) denotes the Stratonovich lift of \( (Z^1, \ldots, Z^d) \). Recall that \( \log Z \) takes values in \( \mathbb{R}^d \oplus so(d) \). Define \( \tilde{Z} \) by perturbing the Lévy area as follows

\[
\log \tilde{Z} := \log Z + (0, \Gamma_t)
\]

where \( \Gamma \in C^{1-\text{var}}([0, T], so(d)) \). Then the solution to

\[
d\tilde{u} + Ldt = \Lambda d\tilde{Z}
\]

is identified with

\[
d\tilde{u} + Ldt = \Lambda d\tilde{Z} + \sum_{i,j \in \{1, \ldots, d\}} [\Lambda_i, \Lambda_j] d\Gamma^{i,j}.
\]

The practical relevance is that one can construct approximations \( (Z^n) \) to \( Z \), each \( Z^n \) only dependent on finitely many points, which converge uniformly to \( Z \) with the “wrong” area (cf. [18]); that is,

\[
\left( Z^n, \int Z^n dZ^n \right) \to \tilde{Z}
\]
in $p$-variation rough path sense, $p \in (2, 3)$. The solutions to the resulting approximations will then converge to the solution of the “wrong” limiting equation

$$d\hat{u} + L dt = \sum_{k=1}^{d} \Lambda_{k} \circ dZ^{k} + \sum_{i,j \in \{1, \ldots, d\}} [\Lambda_{i}, \Lambda_{j}] d\Gamma^{i,j}.$$

The formal proof is easy; it suffices to analyze the equations (rough) differential equations for $(\psi, \phi, \alpha)$ in presence of area perturbation; see [18], and then identify the corresponding operators $[\phi [L_{y}^{\gamma}]]_{\alpha}$. Actually, one can pick any multi-index $\gamma = (\gamma_{1}, \ldots, \gamma_{N}) \in \{1, \ldots, d\}^{N}$ and find (uniform) approximations such as to make the higher iterated Lie brackets $\Lambda_{\gamma} = [\Lambda_{\gamma_{1}}, [\Lambda_{\gamma_{2}}, \ldots, [\Lambda_{\gamma_{N-1}}, \Lambda_{\gamma_{N}}] \ldots \ldots]$, or even any linear combination of them, appear by perturbing the higher order iterated integrals.

**SPDEs with delayed Brownian input** A interesting concrete example of the previous discussion arises when the 2-dimensional driving signal is Brownian motion with its $\varepsilon$-delay; say

$$du^{\varepsilon} + L dt = \Lambda_{1} \circ dB_{\varepsilon}^{\varepsilon} + \Lambda_{2} \circ dB.$$

where $B_{t-\varepsilon}^{\varepsilon} := B_{t-\varepsilon}$. Observe that in the classical setting this can be solved (as flow) on $[0, \varepsilon]$, then on $[\varepsilon, 2\varepsilon]$ and so on. As $\varepsilon \to 0$, $(B_{t}^{\varepsilon}, B_{t})$ converges in rough path sense to $(B_{t}, B_{t})$ with non-trivial area $-t/2$ (see [22, Chapter 14]). In other words, $u^{\varepsilon} \to u$ in probability and locally uniformly where

$$du + L dt = (\Lambda_{1} + \Lambda_{2}) \circ dB + [\Lambda_{1}, \Lambda_{2}] dt$$

(Robustness of the Zakai SPDE in nonlinear filtering) Nonlinear filtering is concerned with the estimation of the conditional law of a Markov process; to be precise consider

$$\begin{align*}
    dX_{t} & = \mu (X_{t}) dt + V (X_{t}) dB_{t} + \sigma (X_{t}) d\tilde{B}_{t} \\
    dY_{t} & = h (X_{t}) dt + d\tilde{B}_{t}
\end{align*} \tag{6.1}$$

where $B$ and $\tilde{B}$ are independent, multidimensional Brownian motions. The goal is to compute for a given real-valued function $\varphi$

$$\pi_{t} (\varphi) = \mathbb{E} [\varphi (X_{t}) | \sigma (Y_{r}, r \leq t)]$$

and from basic principles it follows that there exists a map $\phi^{\varphi}_{t} : C ([0, T], \mathbb{R}^{d_{Y}}) \to \mathbb{R}$ such that

$$\phi^{\varphi}_{t} (Y | [0, t]) = \pi_{t} (\varphi) \quad \mathbb{P} \text{- a.s.} \tag{6.2}$$

As pointed out by Clark [7], this classic formulation is not justified in practice since only discrete observations of $Y$ are available and the functional $\phi^{\varphi}_{t}$ is only defined up to nullsets on pathspace (which includes the in practice observed, bounded variation path). He showed that in the case of uncorrelated noise ($\sigma \equiv 0$ in (6.1)) there exists a unique $\phi^{\varphi}_{t} : C ([0, T], \mathbb{R}^{d_{Y}}) \to \mathbb{R}$ which is continuous in uniform norm and fulfills (6.2), thus providing a version of the conditional expectation $\pi_{t} (\varphi)$ which is robust under approximations in uniform norm of the observation $Y$.

Unfortunately in the correlated noise case this is no longer true.\textsuperscript{[18]}

In [10] it was recently shown that in this case robustness still holds in a rough path sense. Now recall that under well-known conditions [39, 41, 1], $\pi_{t}$ can be written in the form

$$\pi_{t} (\varphi) = \int_{\mathbb{R}^{d_{Y}}} \varphi (x) \frac{u_{t} (x)}{\int_{\mathbb{R}^{d_{Y}}} u_{t} (\tilde{x}) d\tilde{x}} dx \tag{6.3}$$

\textsuperscript{18} We quote Mark Davis [11]

“It must, regretfully, be pointed out that the results for correlated noise cannot, unlike those for the independence case, be extended to vector observations. This is because the corresponding operators (...) do not in general commute whereas with no noise correlation they are multiplication operators which automatically commute.”

See also the counterexample given in [10].
where \( u_t \in L^1(\mathbb{R}^n) \) a.s. and \((u_t)\) is the \(L^2\)-solution of the (dual) Zakai SPDE

\[
d u_t = G^* dt + \sum_k N_k u_t dY_t^k
\]

\[(6.4)\]

\[
= \left( G^* + \frac{1}{2} \sum_k N_k N_k \right) u_t dt + \sum_k N_k u_t \circ dY_t^k
\]

with \( G \) denoting the generator of the diffusion \( X \), \( Y \) a Brownian motion under a measure change and

\[
(N_k u) (t, x) = \sigma_k^t (t, x) \partial_t u_t (x) + h (x) \cdot u_t (x).
\]

Using Theorem (14) in combination with Proposition 20 one can now construct a solution of \( (6.4) \) which depends continuously on the observation in rough path metric. The results in [10] (where one works directly with Kallianpur–Striebel functional) suggest that one can use the representation \( 6.3 \) to establish robustness. However, to this end it is necessary to show that \( u^2_t \in L^1 \) (i.e. a rough pathwise version of the discussion in [11] Chapter 5) which we hope to discuss in the future in detail. Finally, let us note that the gradient term in the noise \( N_k u \) explains rather intuitively why in the general, correlated noise case rough path metrics are required: as pointed out above, correction terms are picked up by the brackets \([N_i, N_j] \) but if \( \sigma = 0 \) then \([N_i, N_j] = 0 \), hence no extra terms appear. In fact, solving \( (6.4) \) for the case of \( \sigma \equiv 0 \) reduces via the method of characteristics to solving an SDE with commuting vector fields which is well-known to be robust under approximations of the driving signal (i.e. the observation \( Y \)) in uniform norm.

7. Appendix: Comparison for parabolic equations

Let \( u \in \text{BUC} ([0, T] \times \mathbb{R}^n) \) be a subsolution to \( \partial_t + F \); that is,

\[
\partial_t u + F (t, x, u, Du, D^2 u) \leq 0
\]

if \( u \) is smooth and with the usual viscosity definition otherwise. Similarly, let \( v \in \text{BUC} ([0, T] \times \mathbb{R}^n) \) be a supersolution.

**Theorem 28.** Assume condition (3.14) of the User’s Guide [9], uniformly in \( t \), together with uniform continuity of \( F = F (t, x, r, p, X) \) whenever \( r, p, X \) remain bounded. Assume also a (weak form of) properness: there exists \( C \) such that

\[
F (t, x, r, p, X) - F (t, x, s, p, X) \geq C (r - s) \quad \forall r \leq s,
\]

and for all \( t \in [0, T] \) and all \( x, p, X \). Then comparison holds. That is,

\[
u (0, \cdot) - v (0, \cdot) \Rightarrow u \leq v \text{ on } [0, T] \times \mathbb{R}^n.
\]

**Proof.** It is easy to see that \( \tilde{u} = e^{-Ct} u \) is a subsolution to a problem which is proper in the usual sense; that is \( 7.1 \) holds with \( C = 0 \) which is tantamount to require that \( F \) is non-decreasing in \( r \). The standard arguments (e.g. [9] or the appendix of [6] or also [14]) then apply with minimal adaptations.

**Corollary 29.** Under the assumptions of the theorem above let \( u, v \) be two solutions, with initial data \( u_0, v_0 \) respectively. Then

\[
|u - v|_{\infty; \mathbb{R}^n \times [0, T]} \leq e^{CT} |u_0 - v_0|_{\infty; \mathbb{R}^n}
\]

with \( C \) being the constant from \( 7.1 \).

**Proof.** Use again the transformation \( \tilde{u} = e^{-Ct} u, \tilde{v} = e^{-Ct} v \). Then \( \tilde{v} + |u_0 - v_0|_{\infty; \mathbb{R}^n} \) is a supersolution of a problem to which standard comparison arguments apply; hence,

\[
\tilde{u} \leq \tilde{v} + |u_0 - v_0|_{\infty; \mathbb{R}^n}.
\]

Applying the same reasoning to \( \tilde{u} + |u_0 - v_0|_{\infty; \mathbb{R}^n} \) and finally undoing the transformation gives the result.
[30] N. V. Krylov. An analytic approach to SPDEs. In Stochastic partial differential equations: six perspectives, volume 64 of Math. Surveys Monogr., pages 185–242. Amer. Math. Soc., Providence, RI, 1999.

[31] N. V. Krylov and B. L. Rozovskiĭ. Itô equations in Banach spaces and strongly parabolic stochastic partial differential equations. Dokl. Akad. Nauk SSSR, 249(2):285–289, 1979.

[32] Nicolai Krylov. An analytical approach to SPDEs. In René Carmona and Boris Rozovskiĭ, editors, Stochastic partial differential equations: six perspectives, pages 185–242. AMS, 1999.

[33] Nicolai V. Krylov. Lectures on Elliptic and Parabolic Equations in Sobolev Spaces. Graduate Studies in Mathematics, 2008.

[34] M. Ledoux, Z. Qian, and T. Zhang. Large deviations and support theorem for diffusion processes via rough paths. Stochastic Process. Appl., 102(2):265–283, 2002.

[35] Pierre-Louis Lions and Panagiotis E. Souganidis. Fully nonlinear stochastic pde with semilinear stochastic dependence. C. R. Acad. Sci. Paris Sér. I Math., 331(8):617–624, 2000.

[36] Terry Lyons. Differential equations driven by rough signals. Rev. Mat. Iberoamericana, 14(2):215–310, 1998.

[37] Terry Lyons and Zhongmin Qian. System Control and Rough Paths. Oxford University Press, 2002. Oxford Mathematical Monographs.

[38] Terry J. Lyons, Michael Caruana, and Thierry 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, With an introduction concerning the Summer School by Jean Picard.

[39] E. Pardoux. Stochastic partial differential equations and filtering of diffusion processes. Stochastics, 3(2):127–167, 1979.

[40] Claudia Prévôt and Michael Röckner. A concise course on stochastic partial differential equations, volume 1905 of Lecture Notes in Mathematics. Springer, Berlin, 2007.

[41] B. L. Rozovskiĭ. Stochastic evolution systems, volume 35 of Mathematics and its Applications (Soviet Series). Kluwer Academic Publishers Group, Dordrecht, 1990. Linear theory and applications to nonlinear filtering, Translated from the Russian by A. Yarkho.

[42] E.M. Stein. Singular integrals and differentiability properties of functions. Princeton mathematical series. Princeton University Press, 1970.

[43] Josef Teichmann. Another approach to some rough and stochastic partial differential equations. 2009. arXiv:0908.2814.

[44] Giannmario Tessitore and Jerzy Zabczyk. Wong-Zakai approximations of stochastic evolution equations. J. Evol. Equ., 6(4):621–655, 2006.

[45] Krystyna Twardowska. Wong-Zakai approximations for stochastic differential equations. Acta Appl. Math., 43(3):317–359, 1996.

PF is affiliated with WIAS, Mohrenstrasse 39, 10117 Berlin and TU Berlin, Institut für Mathematik, Strasse des 17. Juni 136, 10623 Berlin.

HO is corresponding author and affiliated with TU Berlin, Institut für Mathematik, Strasse des 17. Juni 136, 10623 Berlin.

E-mail address: h.oberhauser@gmail.com