Probab. Theory Relat. Fields (2014) 158:401–434
DOI 10.1007/s00440-013-0483-2

Rough path stability of (semi-)linear SPDEs

Peter Friz · Harald Oberhauser

Received: 21 April 2011 / Revised: 3 September 2012 / Published online: 8 February 2013

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.

Keywords Rough paths · Viscosity solutions · SPDEs · Zakai equation

Mathematics Subject Classification 60H15 · 35K58 · 60G35

1 Introduction

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

$$du + L(t, x, u, Du, D^2u)dt = \sum_{k=1}^{d} \Lambda_k(t, x, u, Du) \circ dZ_t^k, \tag{1.1}$$

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

P. Friz
WIAS, Mohrenstrasse 39, 10117 Berlin, Germany

P. Friz · H. Oberhauser (✉)
Institut für Mathematik, TU Berlin, Strasse des 17. Juni 136, 10623 Berlin, Germany
e-mail: h.oberhauser@gmail.com

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

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

$$\Lambda_k(t,x,r,p) = p\cdot\sigma_k(t,x) + r\,\nu_k(t,x) + g_k(t,x), \quad k=1,\dots,d. \tag{1.2}$$

The contribution of this article is to give meaning to Eq. (1.1) when $Z(\omega)$ is replaced by a rough path $\mathbf{z}$ (this is carried out in Sects. 2, 3 and 4). Our main result as stated and proven in Sect. 4 (in Sect. 2 we recall Lip^γ-regularity, rough paths and their metrics and the BUC space of continuous, 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 1 and the coefficients of $\Lambda = (\Lambda_1,\dots,\Lambda_{d_1+d_2+d_3})$ fulfill Assumption 2 for some $\gamma > p+2$ (Assumption 1 and 2 are given in Sect. 4). Let $u_0 \in BUC(\mathbb{R}^n)$ and let $\mathbf{z}$ be a geometric p-rough path. Then there exists a unique $u = u^{\mathbf{z}} \in BUC([0,T]\times\mathbb{R}^n)$ such that for any sequence $(z^\epsilon)_\epsilon \subset C^1([0,T],\mathbb{R}^d)$ such that $z^\varepsilon \to \mathbf{z}$ in p-rough path sense, the viscosity solutions $(u^\varepsilon) \subset BUC([0,T]\times\mathbb{R}^n)$ of*

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

converge locally uniformly on $[0,T]\times\mathbb{R}^n$ to $u^{\mathbf{z}}$. We write formally,

$$du + L(t,x,u,Du,D^2u)dt = \Lambda(t,x,u,Du)\,d\mathbf{z}_t, \quad u(0,\cdot) = u_0(\cdot).$$

Moreover, we have the contraction property

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

(C given by (4.2)) and continuity of the solution map

$$C^{0,p\text{-}var}\left([0,T],G^{[p]}\left(\mathbb{R}^d\right)\right)\times BUC\left(\mathbb{R}^n\right) \to BUC\left([0,T]\times\mathbb{R}^n\right)$$
$$(\mathbf{z},u_0) \mapsto u^{\mathbf{z}}.$$

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 Sect. 6. By combining well-known Wong–Zakai type results of the L^2-theory of SPDEs [3,29,44,45] 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 Sect. 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 on the coefficients (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 Eq. (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 $Z_\cdot(\omega)$. 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 ${}^{\phi}L$ in v, Dv, $D^2 v$ with coefficients that will depend on z resp. $Z_t(\omega)$,

$$dv + {}^{\phi}L\left(t, x, v, Dv, D^2 v\right) dt = 0.$$

Moreover, one can conclude from this representation that $u = u(z)$ is continuous with respect to the uniform metric $|z - \tilde{z}|_{\infty;[0,T]} = \sup_{r \in [0,T]} |z_r - \tilde{z}_r|$. 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

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

(which are well-defined for smooth $z : [0, T] \to \mathbb{R}^d$) to depend continuously on z in uniform topology (cf. the "twisted approximations" of Sect. 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 is meaningful to replace z above by an abstract (geometric) rough path $\mathbf{z}$ 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

$$\sum_{i=1}^{d_1} (Du \cdot \sigma_i (t, x)) \circ dZ_t^{1;i} + u \sum_{j=1}^{d_2} \nu_j (t, x) \circ dZ^{2;j} + \sum_{k=1}^{d_3} g_k (t, x) \circ dZ^{3;k}$$

where $Z \equiv (Z^1, Z^2, Z^3)$ and Z^i is a d_i-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

$$u \overset{\text{Transformation 1}}{\mapsto} u^1 \text{ where } u^1 \text{ has the (gradient) noise driven by } Z^1 \text{ removed;}$$
$$u^1 \overset{\text{Transformation 2}}{\mapsto} u^{12} \text{ where } u^{12} \text{ has the remaining noise driven by } Z^2 \text{ removed;}$$
$$u^{12} \overset{\text{Transformation 3}}{\mapsto} \tilde{u} \text{ where } \tilde{u} \text{ has the remaining noise driven by } Z^3 \text{ removed.}$$

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 know 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.[2] 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 [5,6], 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 Eq. (1.1). We note that the viscosity/Stratonovich approach allows us to *avoid any ellipticity*

[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.

[2] Transformation 2 and 3 could actually be performed in 1 step; however, the separation leads to a simpler analytic tractability of the transformed equations.

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 the 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, [3,29,44,45], 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 Sects. 2–4 is then discussed in Sect. 5. Interestingly, there has been no success until now (despite the advances by Deya–Gubinelli–Tindel [12,23] and Teichmann [43]) to include (1.1) in a setting of abstract *rough* evolution equations on infinite-dimensional spaces.

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,

$$\partial_t u + F\left(t, x, u, Du, D^2 u\right) \leq 0,$$

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

$$F\left(t, x, r, p, A + B\right) \leq F\left(t, x, r, p, A\right)$$

whenever $B \geq 0$ in the sense of symmetric matrices (cf. [9]). The idea is to consider a (smooth) test function φ and look at a local maxima $(\hat{t}, \hat{x})$ of $u - \varphi$. Basic calculus implies that $Du(\hat{t},\hat{x}) = D\varphi(\hat{t},\hat{x})$, $D^2 u(\hat{t},\hat{x}) \leq D\varphi(\hat{t},\hat{x})$ and, from degenerate ellipticity,

$$\partial_t \varphi + F\left(\hat{t}, \hat{x}, u, D\varphi, D^2\varphi\right) \leq 0. \tag{2.1}$$

This suggests to define a *viscosity supersolution* (at the point $(\hat{x}, \hat{t})$) 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

$$\begin{aligned} u\left(t,x\right) \leq u\left(\hat{t},\hat{x}\right) + a\left(t-\hat{t}\right) + p\cdot\left(x-\hat{x}\right) + \frac{1}{2}\left(x-\hat{x}\right)^T \cdot X \cdot \left(x-\hat{x}\right) \\ + o\left(\left|\hat{x}-x\right|^2 + \left|\hat{t}-t\right|\right) \end{aligned} \tag{2.2}$$

as $|\hat{x}-x|^2+|\hat{t}-t| \to 0$ 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 $a = \partial_t u(\hat{t}, \hat{x})$, $p = Du(\hat{t}, \hat{x})$, $X \le D^2 u(\hat{t}, \hat{x})$, hence by degenerate ellipticity

$$\partial_t\varphi + F\left(\hat{t}, \hat{x}, u, p, X\right) \le 0.$$

Pushing this idea further leads to a definition of viscosity solutions based on a generalized notion of "$(\partial_t u, Du, D^2 u)$" 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,[3] the initial value problem $(\partial_t + F)u = 0$, $u(0, \cdot) = u_0 \in \mathrm{BUC}(\mathbb{R}^n)$ has a unique solution in $\mathrm{BUC}([0, T] \times \mathbb{R}^n)$ 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.[4] In fact, uniqueness follows from a stronger property known as *comparison:* assume u (resp. v) is a supersolution (resp. subsolution) and $u_0 \ge v_0$; then $u \ge v$ 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).[5] 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

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

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 $|z-\tilde{z}|_{;[0,T]} = \sup_{r\in[0,T]} |z_r - \tilde{z}_r|$, then the solution will in general *not* depend continuously on the input.

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

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

[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].

converges to 0 in ∞-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

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

(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

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

where $z^{(j)}_{s,t} = \int dz_{r_1} \otimes \cdots \otimes dz_{r_j}$ 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[6] 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

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

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

$$B^{(2)} := \left(\int_0^{\cdot} B^i \circ dB^j \right)_{i,j\in\{1,\dots,d\}} \tag{2.4}$$

yields a path $\mathbf{B}(\omega)$ in $G^2(\mathbb{R}^d)$ with finite $1/p$-Hölder (and hence finite p-variation) regularity, for any $p > 2$. ($\mathbf{B}$ is known as *Brownian rough path.*) We remark that

[6] In the strict terminology of rough path theory: geometric p-rough paths.

$B^{(2)} = \frac{1}{2} B \otimes B + A$ where the anti-symmetric part of the matrix, $A :=$Anti$(B^{(2)})$, is known as *Lévy's stochastic area;* in other words $\mathbf{B}(\omega)$ is determined by (B, A), i.e. Brownian motion *enhanced with Lévy's area.* A similar construction works when B is replaced by a generic multi-dimensional continuous semimartingale; see [22, Chapter 14] and the references therein.

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 (7.1). Time derivatives of functions are denoted by upper dots, spatial derivatives (with respect to x) by D, D^2, etc. Further, we use, $\langle ., . \rangle$ to denote tensor contraction,[7] i.e. $\langle p, q \rangle_{j_1,\ldots,j_n} \equiv \sum_{i_1,\ldots,i_m} p_{i_1,\ldots,i_m} q^{i_1,\ldots,i_m}_{j_1,\ldots,j_n}$, $p \in (\mathbb{R}^l)^{\otimes m}$, $q \in (\mathbb{R}^l)^{\otimes n} \otimes ((\mathbb{R}^l)')^{\otimes m}$.

Lemma 3 (Inner transform) *Let* $z \in C^1([0, T], \mathbb{R}^d), \sigma = (\sigma_1, \ldots, \sigma_d) \subset C_b^2([0, T] \times \mathbb{R}^n, \mathbb{R}^n)$ *(the space of continuous and twice differentiable, bounded functions with bounded derivatives) and* $\psi = \psi(t, x)$ *the ODE flow of* $dy = \sigma(y) dz$*, i.e.*

$$\dot{\psi}(t, x) = \sum_{i=1}^{d} \sigma_i(t, \psi(t, x)) \dot{z}_t^i, \quad \dot{\psi}(0, x) = x \in \mathbb{R}^n.$$

Then u *is a viscosity subsolution (always assumed* BUC*) of*

$$\partial_t u + F\left(t, x, r, Du, D^2 u\right) - \sum_{i=1}^{d} (Du \cdot \sigma_i(t, x)) \dot{z}_t^i = 0; \quad u(0, .) = u_0(.) \tag{3.1}$$

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

$$\partial_t w + F^{\psi}\left(t, x, w, Dw, D^2 w\right) = 0; \quad w(0, .) = u_0(.) \tag{3.2}$$

where

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

[7] We also use $\cdot$ to denote contraction over only index or to denote matrix multiplication.

and

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

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

Remark 4 The regularity assumptions on σ 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 [4]).

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 5 (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*

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

define also the transform

$$L^\psi = -Tr\left[A^\psi(t,x)\cdot X\right] + b^\psi(t,x)\cdot p + c^\psi(t,x,r)$$

where

$$\begin{aligned} A^\psi(t,x) &= \left\langle A(t,\psi_t(x)), D\psi_t^{-1}|_{\psi_t(x)} \otimes D\psi_t^{-1}|_{\psi_t(x)}\right\rangle, \\ b^\psi(t,x)\cdot p &= b(t,\psi_t(x))\cdot\left\langle p, D\psi_t^{-1}|_{\psi_t(x)}\right\rangle - Tr\left(A(t,\psi_t)\cdot\left\langle p, D^2\psi_t^{-1}|_{\psi_t(x)}\right\rangle\right), \\ c^\psi(t,x,r) &= c(t,\psi_t(x),r). \end{aligned}$$

Then u is a solution (always assumed BUC) of

$$\partial_t u + L\left(t,x,u,Du,D^2u\right) = \sum_{i=1}^d \left(Du\cdot\sigma_i(t,x)\right)\dot{z}_t^i; \quad u(0,.) = u_0(.)$$

if and only if $u^1(t,x) := u(t,\psi(t,x))$ is a solution of

$$\partial_t + L^\psi = 0; \quad u^1(0,.) = u_0(.). \tag{3.3}$$

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^ψ to see that

$$\begin{aligned}\dot{w}(t,x) &= \dot{u}(t,y) + Du(t,y)\cdot\dot{\psi}_t(x) = \dot{u}(t,y) + Du(t,y)\cdot\sigma(y)\dot{z}_t \\ &\leq F\Big(t,y,u(t,y),Du(t,y),D^2u(t,y)\Big) \\ &= F^{\psi}\Big(t,x,w(t,x),Dw(t,x),D^2w(t,x)\Big).\end{aligned}$$

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 $\varphi \in C^2((0,T)\times\mathbb{R}^n)$ by $\varphi(t,y) = \xi(t,\psi_t^{-1}(y))$. Set $\bar{y} = \psi_{\bar{t}}(\bar{x})$ so that

$$F\left(\bar{t},\bar{y},u\left(\bar{t},\bar{y}\right),D\varphi\left(\bar{t},\bar{y}\right),D^2\varphi\left(\bar{t},\bar{y}\right)\right) = F^{\psi}\left(\bar{t},\bar{x},w\left(\bar{t},\bar{x}\right),D\xi\left(\bar{t},\bar{x}\right),D^2\xi\left(\bar{t},\bar{x}\right)\right).$$

Obviously, $(\bar{t},\bar{y})$ is a maximum of $u-\varphi$, and since u is a viscosity sub-solution of (3.1) we have

$$\dot{\varphi}\left(\bar{t},\bar{y}\right) + D\varphi\left(\bar{t},\bar{y}\right)\sigma\left(\bar{t},\bar{y}\right)\dot{z}\left(\bar{t}\right) \leq F\Big(\bar{t},\bar{y},u\left(\bar{t},\bar{y}\right),D\varphi\left(\bar{t},\bar{y}\right),D^2\varphi\left(\bar{t},\bar{y}\right)\Big).$$

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

$$\dot{\xi}\left(\bar{t},\bar{x}\right) \leq F^{\psi}\Big(\bar{t},\bar{x},w\left(\bar{t},\bar{x}\right),D\xi\left(\bar{t},\bar{x}\right),D^2\xi\left(\bar{t},\bar{x}\right)\Big)$$

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). □

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

$$\begin{aligned}\dot{\phi} &= \frac{\partial\phi(t,r,x)}{\partial t}, \phi' = \frac{\partial\phi(t,r,x)}{\partial r}, \\ D\phi &= \left(\frac{\partial\phi(t,r,x)}{\partial x^i}\right)_{i=1,\ldots,n} \quad \text{and} \quad D^2\phi = \left(\frac{\partial^2\phi(t,r,x)}{\partial x^i\partial x^j}\right)_{i,j=1,\ldots,n}.\end{aligned}$$

Lemma 6 [Outer transform] *Let $\phi = \phi(t,r,x) \in C^{1,2,2}$ and assume that $\forall(t,x), r \mapsto \phi(t,r,x)$ is an increasing diffeomorphism on the real line. Then u is a subsolution of $\partial_t u + F(t,x,u,Du,D^2u) = 0$, $u(0,.) = u_0(.)$ if and only if*

$$v(t,x) = \phi^{-1}(t,u(t,x),x)$$

is a subsolution of $\partial_t v +^{\phi} F(t, x, v, Dv, D^2 v) = 0$, $v(0, .) = \phi^{-1}(0, u_0(x), x)$ *with*

$$ {}^{\phi}F(t, x, r, p, X) = \frac{\dot{\phi}}{\phi'} + \frac{1}{\phi'} F\big(t, x, \phi, D\phi + \phi' p, \\ \phi'' p \otimes p + D\phi' \otimes p + p \otimes D\phi' + D^2\phi + \phi' X\big) \tag{3.4} $$

where ϕ *and all derivatives are evaluated at* (t, r, x). *The same statement holds if one replaces the word subsolution by supersolution throughout.*

Proof ($\Longrightarrow$) 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. [9, Section 8]) iff

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

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

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

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 ϕ in all three arguments we see that the right hand side equals

$$ \phi(s, v(s, z), z) + \dot{\phi}_{s,v(s,z),z}(t - s) + \phi'_{s,v(s,z),z} a(t - s) + \phi'_{s,v(s,z),z} p \cdot (x - z) \\ + \frac{1}{2}\phi'_{s,v(s,z),z}(x - z)^T \cdot X \cdot (x - z) + D\phi_{s,v(s,z),z} \cdot (x - z) \\ + \frac{1}{2}(x - z)^T \cdot D^2\phi_{s,v(s,z),z} \cdot (x - z) \\ + (x - z)^T \cdot \left(D\left(\phi'\right)\right)_{s,v(s,z),z} \otimes p \cdot (x - z) \\ + (x - z)^T \cdot p \otimes (D\phi)'_{s,v(s,z),z} \cdot (x - z) \\ + (x - z)^T \cdot \phi''_{s,v(s,z),z} p \otimes p \cdot (x - z) + o\left(|t - s| + |x - z|^2\right) \text{ as } (s, z) \to (t, x) $$

Hence,

$$ \Big(\dot{\phi}_{s,v(s,z),z} + \phi'_{s,v(s,z),z} a, D\phi_{s,v(s,z),z} + \phi'_{s,v(s,z),z} p, \phi''_{s,v(s,z),z} p \otimes p + D\left(\phi'\right)_{s,v(s,z),z} \\ \otimes p + p \otimes (D\phi)'_{s,v(s,z),z} + D^2\phi_{s,v(s,z),z} + \phi'_{s,v(s,z),z} X\Big) $$

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

$$\dot{\phi}_{s,v(s,z),z} + \phi'_{(s,v(s,z),z)}a + F\Big(s, z, \phi_{(s,v(s,z),z)}, D\phi_{s,v(s,z),z} + \phi'_{s,v(s,z),z}p,$$
$$\phi''_{s,v(s,z),z}p\otimes p + D\left(\phi'\right)_{s,v(s,z),z}\otimes p + p\otimes (D\phi)'_{s,v(s,z),z} + D^2\phi_{s,v(s,z),z} + \phi'_{s,v(s,z),z}X\Big) \le 0.$$

Dividing by ϕ' (>0) shows that v is a subsolution of $\partial_t v + F^\phi = 0$.

($\Longleftarrow$) Assume v is a subsolution of $\partial_t v +^\phi F = 0,^\phi F$ defined as in (3.4 for some F. Set $u(t,x) := \phi(t, v(t,x), x)$. By above argument we know that v is a subsolution of ${}^{\phi^{-1}}({}^\phi F)(t,x,r,p,X)$. For brevity write $\psi(t,.,x) = \phi^{-1}(t,.,x)$. Then

$$\begin{aligned}
&{}^{\phi^{-1}}\left({}^\phi F\right)(t,x,r,p,X)\\
&= \frac{\psi_{t,r,x}}{\psi'_{t,r,x}} + \frac{1}{\psi'_{t,r,x}}\,{}^\phi F\Big(t,x,\psi_{(t,r,x)}, D\psi_{t,r,x} + \psi'_{t,r,x}p,\\
&\quad \psi''_{t,r,x}p\otimes p + D\left(\psi'\right)_{t,r,x}\otimes p + p\otimes (D\psi)'_{t,r,x} + D^2\psi_{t,r,x} + \psi'_{t,r,x}X\Big)\\
&= \frac{\psi_{t,r,x}}{\psi'_{t,r,x}} + \frac{1}{\psi'_{t,r,x}}\left[\frac{\dot{\phi}_{t,\psi_{t,r,x},x}}{\phi'_{t,\psi_{t,r,x},x}} + \frac{1}{\phi'_{t,\psi_{t,r,x},x}}F\Big(t,x,\phi\left(t,\psi_{t,r,x},x\right),\right.\\
&\quad D\phi_{t,\psi_{t,r,x},x} + \phi'_{t,\psi_{t,r,x},x}\left\{D\psi_{t,r,x} + \psi'_{t,r,x}p\right\},\\
&\quad \phi''_{t,\psi_{t,r,x},x}p\otimes p + D\left(\phi'\right)_{t,\psi_{t,r,x},x}\otimes p + p\otimes (D\phi)'_{t,\psi_{t,r,x},x} + D^2\phi_{t,\psi_{t,r,x},x}\\
&\quad \left.+\phi'_{t,\psi_{t,r,x},x}\left\{\psi''_{t,r,x}p\otimes p + D\left(\psi'\right)_{t,r,x}\otimes p + p\otimes (D\psi)'_{t,r,x} + D^2\psi_{t,r,x} + \psi'_{t,r,x}X\right\}\Big)\right].
\end{aligned}$$

Using several times equalities of the type $(f\circ f^{-1})' = f'_{f^{-1}}(f^{-1})' = id$ cancels the terms involving ϕ, ψ and their derivatives and we are left with F, i.e.

$${}^{\phi^{-1}}\left({}^\phi F\right) = F.$$

This finishes the proof. □

Corollary 7 (Transformation 2) *Assume* $\nu = (\nu_1,\ldots,\nu_d) \subset C_b^{0,2}([0,T]\times\mathbb{R}^n)$ *(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*

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

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

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

define also

$${}^\phi L(t,x,r,X) = -Tr\left[A(t,x)\cdot X\right] + {}^\phi b(t,x)\cdot p + {}^\phi c(t,x,r) \tag{3.5}$$

where

$$
\begin{aligned}
{}^{\phi}b\,(t,x)\cdot p &\equiv b\,(t,x)\cdot p-\frac{2}{\phi'}Tr\left[A\,(t,x)\cdot D\phi'\otimes p\right]\\
{}^{\phi}c\,(t,x,r) &\equiv -\frac{1}{\phi'}Tr\left[A\,(t,x)\cdot (D^2\phi)\right]+\frac{1}{\phi'}b\,(t,x)\cdot(D\phi)+\frac{1}{\phi'}c\,(t,x,\phi)
\end{aligned}
$$

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

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

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

$$
\partial_t v+{}^{\phi}L\left(t,x,v,Dv,D^2v\right)=0.
$$

Proof Obviously,

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

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 $\dot{\phi}/\phi=d\cdot\dot{z}\equiv\sum_{j=1}^d d_j(t,x)\dot{z}_t^j$. It follows, for general F, that

$$
\begin{aligned}
{}^{\phi}F\,(t,x,r,p,X)=r\,d\cdot\dot{z}+\frac{1}{\phi'}F\,\big(t,x,\phi,D\phi+\phi'p,\\
D\phi'\otimes p+p\otimes D\phi'+D^2\phi+\phi'X\big)
\end{aligned}
$$

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

$$
\begin{aligned}
{}^{\phi}L=-\frac{1}{\phi'}\mathrm{Tr}\left[A\,(t,x)\cdot(D\phi'\otimes p+p\otimes D\phi'+D^2\phi+\phi'X)\right]\\
+\frac{1}{\phi'}b\,(t,x)\cdot\left(D\phi+\phi'p\right)+\frac{1}{\phi'}c\,(t,x,\phi)
\end{aligned}
$$

where ϕ 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). □

We shall need another (outer)transform to remove additive noise.

Lemma 8 (Transformation 3) *Let $g \in C([0,T] \times \mathbb{R}^n, \mathbb{R}^d)$ 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*

$$\begin{aligned} L(t,x,r,p,X) &= -Tr[A(t,x)\cdot X] + b(t,x)\cdot p + c(t,x,r);\\ L_\varphi(t,x,r,p,X) &= -Tr[A(t,x)\cdot X] + b(t,x)\cdot p + c_\varphi(t,x,r)\\ \text{with } c_\varphi(t,x,r) &= Tr\left[A(t,x)\cdot D^2\varphi(t,x)\right] - b(t,x)\cdot D\varphi(t,x)\\ &\quad + c(t,x,r-\varphi(t,x)) \end{aligned}$$

Then v solves

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

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

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

Proof Left to reader. □

3.2 The full transformation

As before, let

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

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

$$\begin{aligned} M_k(t,x,u,Du) &= \sigma_k(t,x)\cdot Du \quad \text{for } k=1,\ldots,d_1\\ M_{d_1+k}(t,x,u,Du) &= u\,\nu_k(t,x) \quad \text{for } k=1,\ldots,d_2\\ M_{d_1+d_2+k}(t,x,u,Du) &= g_k(t,x) \quad \text{for } k=1,\ldots,d_3. \end{aligned} \tag{3.6}$$

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

Proposition 9 *Let $z^1 \in C^1([0,T],\mathbb{R}^{d_1})$, $\sigma = (\sigma_1,\ldots,\sigma_{d_1}) \subset C_b^2([0,T]\times\mathbb{R}^n,\mathbb{R}^n)$ and denote the ODE flow of $dy = \sigma(t,y)dz$ with ψ, i.e. $\psi : [0,T]\times\mathbb{R}^n \to \mathbb{R}^n$ satisfies*

$$\dot{\psi}(t,x) = \sigma(t,\psi(t,x))\,\dot{z}_t^1, \quad \psi(0,x) = x \in \mathbb{R}^n. \tag{3.7}$$

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

$$\dot{\phi} = \phi\,\underbrace{\nu(t,\psi_t(x))}_{\equiv d^\psi(t,x)}\dot{z}_t^2, \quad \phi(0,r,x) = r \in \mathbb{R}. \tag{3.8}$$

Further, let $z^3 \in C^1([0,T],\mathbb{R}^{d_3})$ *and define for given* $g = (g_1,\ldots,g_{d_3}) \in C([0,T]\times\mathbb{R}^n,\mathbb{R}^{d_3})$, $\varphi(t,x)$ *as the integral*[8]

$$\varphi(t,x) = \int_0^t {}^{\phi}g^{\psi}(s,x)\,d\dot{z}_s^3, \tag{3.9}$$

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

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*

$$\partial_t u + L\left(t,x,u,Du,D^2u\right) = \Lambda(t,x,u,Du)\,\dot{z}_t, \tag{3.10}$$

$$u(0,x) = u_0(x), \tag{3.11}$$

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

$$\partial_t\tilde{u} + \tilde{L}\left(t,x,\tilde{u},D\tilde{u},D^2\tilde{u}\right) = 0 \tag{3.12}$$

$$\tilde{u}(0,x) = u_0(x) \tag{3.13}$$

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

Remark 10 Transformation 2 and 3 could have been performed in one step, by considering

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

Indeed, the usual variation of constants formula gives immediately

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

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 11 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. [13,35].

8 Since ϕ is linear in r, there is no r dependence in its derivative ϕ'.

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

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

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

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

i.e.

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

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

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

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 $\mathbf{z}$. There is some "tri-diagonal" structure: (3.7) can be solved as function of z^1 alone;

$$d\psi_t\left(x\right) = \sigma\left(t, \psi_t\left(x\right)\right) dz_t^1 \quad \text{with} \quad \psi_0\left(x\right) = x. \tag{3.14}$$

(3.8) is tantamount to

$$\phi\left(t, r, x\right) = r \exp\left[\int_0^t \nu\left(s, \psi_s\left(x\right)\right) dz_s^2\right]. \tag{3.15}$$

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

$$\varphi(t,x) = \int_0^t \tilde{\phi}(s,x)\,g(s,\psi_s(x))\,dz_s^3. \tag{3.16}$$

Lemma 12 *Let* $\mathbf{z}$ *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*

$$\begin{aligned}
\sigma &= \left(\sigma_1,\ldots,\sigma_{d_1}\right) \subset Lip^{\gamma}\left([0,T]\times\mathbb{R}^n,\mathbb{R}^n\right),\\
\nu &= \left(\nu_1,\ldots,\nu_{d_2}\right) \subset Lip^{\gamma-1}\left([0,T]\times\mathbb{R}^n,\mathbb{R}\right),\\
g &= \left(g_1,\ldots,g_{d_3}\right) \subset Lip^{\gamma-1}\left([0,T]\times\mathbb{R}^n,\mathbb{R}\right).
\end{aligned}$$

Then ψ,ϕ *and* φ *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* φ*. In particular, we can define* ψ,ϕ *and* φ *when* (z^1,z^2,z^3) *is replaced by a genuine geometric* p*-rough path* $\mathbf{z}$ *and write* $\psi^{\mathbf{z}},\phi^{\mathbf{z}},\varphi^{\mathbf{z}}$ *to indicate this dependence.*

Proof Given $\mathbf{z}$ one can build a "time-space" rough path $(\mathbf{t},\mathbf{z})$ 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

$$W_1 = \begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix},\quad W_2 = \begin{pmatrix}0\\\sigma(t,\psi)\\0\\0\\0\end{pmatrix},\quad W_4 = \begin{pmatrix}0\\0\\r.\phi.\nu(t,\psi)\\-\tilde{\phi}.\nu(t,\psi)\\0\end{pmatrix},\quad W_4 = \begin{pmatrix}0\\0\\0\\0\\\tilde{\phi}.g(t,\psi)\end{pmatrix}.$$

The assumptions on σ, ν and g guarantee that

$$W = (W_1,\ldots,W_4) : \mathbb{R}^{1+d_1+2d_2+d_3} \to L\left(\mathbb{R}^{1+d},\mathbb{R}^{1+d_1+2d_2+d_3}\right)$$

is Lip^{γ} (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 Lip^{γ} resp. $Lip^{\gamma-1}$ extensions of σ, ν and g). Hence, the "full RDE" (parametrized by $x\in\mathbb{R}^n$ and $r\in\mathbb{R}$) reads as

$$d\begin{pmatrix} t \\ \psi \\ \phi \\ \tilde{\phi} \\ \varphi \end{pmatrix} = W\begin{pmatrix} t \\ \psi \\ \phi \\ \tilde{\phi} \\ \varphi \end{pmatrix} d(\mathbf{t},\mathbf{z}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \sigma(t,\psi) & 0 & 0 \\ 0 & 0 & \phi.\nu(t,\psi) & 0 \\ 0 & 0 & -\tilde{\phi}.\nu(t,\psi) & 0 \\ 0 & 0 & 0 & \tilde{\phi}g(t,\psi) \end{pmatrix} d(\mathbf{t},\mathbf{z})$$

and has a unique global solution[9] (with obvious initial condition that the flows ψ,ϕ,φ 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 [22,36,37]. □

4 Semirelaxed limits and rough PDEs

The goal is to understand

$$\partial_t u + L\left(t,x,u,Du,D^2u\right) = \sum_{i=1}^{d_1}\left(\sigma_i(t,x)\cdot Du\right)\dot{z}_t^{1;i} + u\sum_{j=1}^{d_2}\nu_j(t,x)\dot{z}_t^{2;j} + \sum_{k=1}^{d_3} g_k(t,x)\dot{z}_t^{3;k}$$

in the case when $\left(z^1,z^2,z^3\right)$ 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

$$L(t,x,r,p,X) = -\mathrm{Tr}\left[A(t,x)X\right] + b(t,x)\cdot p + c(t,x,r) \tag{4.1}$$

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.

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

[9] Although W fails to be bounded, the particular structure of the system where one can first solve for ψ and then construct the other quantities by rough integration makes it clear that no explosion can happen. The same situtation is discussed in detail in [22, Chapter 11].

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 Λ.

Assumption 2 Assume that[10]

$$\begin{aligned} \sigma &= \left(\sigma_1, \ldots, \sigma_{d_1}\right) \subset Lip^{\gamma}\left([0,T] \times \mathbb{R}^n, \mathbb{R}^n\right), \\ \nu &= \left(\nu_1, \ldots, \nu_{d_2}\right) \subset Lip^{\gamma-1}\left([0,T] \times \mathbb{R}^n, \mathbb{R}\right), \\ g &= \left(g_1, \ldots, g_{d_3}\right) \subset Lip^{\gamma-1}\left([0,T] \times \mathbb{R}^n, \mathbb{R}\right). \end{aligned}$$

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^{ε} and *impose* convergence to a genuine geometric p-rough path $\mathbf{z}$, that is, in the notation of [22, Chapter 14]

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

We can now prove our main result.

Theorem 13 *Let $p \geq 1$. Assume L fulfills Assumption 1 and the coefficients of $\Lambda = (\Lambda_1, \ldots, \Lambda_{d_1+d_2+d_3})$ fulfill Assumption 2 for some $\gamma > p+2$. Let $u_0 \in BUC(\mathbb{R}^n)$ and $\mathbf{z} \in C^{0,p\text{-}var}([0,T], G^{[p]}(\mathbb{R}^d))$. Then there exists a unique $u = u^{\mathbf{z}} \in BUC([0,T] \times \mathbb{R}^n)$ such that for any sequence $(z^{\epsilon})_{\epsilon} \subset C^1([0,T], \mathbb{R}^d)$ such that $z^{\varepsilon} \to \mathbf{z}$ in p-rough path sense, the viscosity solutions $(u^{\varepsilon}) \subset BUC([0,T] \times \mathbb{R}^n)$ of*

$$\dot{u}^{\varepsilon} + L\left(t, x, u^{\varepsilon}, Du^{\varepsilon}, D^2 u^{\varepsilon}\right) = \sum_{k=1}^{d} \Lambda_k\left(t, x, u^{\varepsilon}, Du^{\varepsilon}\right) \dot{z}^{k;\varepsilon} \quad u^{\epsilon}(0, \cdot) = u_0(.),$$

converge locally uniformly to $u^{\mathbf{z}}$. We write formally,[11]

$$du + L\left(t, x, u, Du, D^2 u\right) dt = \Lambda(t, x, u, Du)\, d\mathbf{z} \quad u(0, \cdot) = u_0(.) \tag{4.3}$$

Moreover, we have the contraction property

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

(C given by (4.2)) and continuity of the solution map

$$\begin{aligned} C^{0,p\text{-}var}\left([0,T], G^{[p]}\left(\mathbb{R}^d\right)\right) \times BUC\left(\mathbb{R}^n\right) &\to BUC\left([0,T] \times \mathbb{R}^n\right) \\ (\mathbf{z}, u_0) &\mapsto u^{\mathbf{z}}. \end{aligned}$$

[10] The regularity assumptions on the vector fields with respect to t could be relaxed here, cf. Remark 4.

[11] The intrinsic meaning of this “rough” PDE is discussed in Definition 14 below.

Proof We shall write $\psi^{\mathbf{z}}, \phi^{\mathbf{z}}, \varphi^{\mathbf{z}}$ for the objects (solutions of rough differential equations and integrals) built upon $\mathbf{z}$, as discussed in the last section (Lemma 12) and also write $\psi^{\varepsilon}, \phi^{\varepsilon}, \varphi^{\varepsilon}$ when the driving signal is z^{ε}. 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

$$\left[{}^{\phi^{\varepsilon}}\left(L^{\psi^{\varepsilon}}\right)\right]_{\varphi^{\varepsilon}} \to \left[{}^{\phi^{\mathbf{z}}}\left(L^{\psi^{\mathbf{z}}}\right)\right]_{\varphi^{\mathbf{z}}}$$

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

$$\tilde{u}^{\varepsilon}(t,x) = \left(\phi^{\varepsilon}\right)^{-1}\left(t, u^{\varepsilon}\left(t, \psi^{\varepsilon}(t,x)\right), x\right) + \varphi^{\varepsilon}(t,x);$$

and Proposition 9 in Sect. 3 show that $\tilde{u}^{\varepsilon}$ is a (classical) viscosity solution of

$$d\tilde{u}^{\varepsilon} + \left[{}^{\phi^{\varepsilon}}\left(L^{\psi^{\varepsilon}}\right)\right]_{\varphi^{\varepsilon}}\left(t, x, \tilde{u}^{\varepsilon}, D\tilde{u}^{\varepsilon}, D^2\tilde{u}^{\varepsilon}\right) = 0.$$

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

$$\sup_{\substack{\varepsilon \in (0,1] \\ t \in [0,T] \\ x \in \mathbb{R}^n}} \left|\tilde{u}^{\varepsilon}(t,x)\right| < \left(1 + |u_0|_{\infty}\right) e^{CT};$$

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 Sect. 3.3 that $\tilde{u}^{\varepsilon}$ remains locally uniform 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

$$du + L\left(t, x, u, Du, D^2u\right)dt = \Lambda(t, x, u, Du)\, d\mathbf{z} \quad u(0,\cdot) \equiv u_0(\cdot).$$

Inspired by the definition given by Lions–Souganidis in [35] we give

Definition 14 u is a solution to the **rough partial differential equation** (4.3) if and only if $\tilde{u}(t,x) = (\phi^{\mathbf{z}})^{-1}(t, u(t, \psi^{\mathbf{z}}(t,x)), x) + \varphi^{\mathbf{z}}(t,x)$

$$d\tilde{u} + \tilde{L}\left(t, x, \tilde{u}, D\tilde{u}, D^2\tilde{u}\right) = 0, \quad \tilde{u}(0,\cdot) = u_0(\cdot)$$

in viscosity sense where

$$\tilde{L} = \left[{}^{\phi^{\mathbf{z}}}\left(L^{\psi^{\mathbf{z}}}\right)\right]_{\varphi^{\mathbf{z}}}.$$

Corollary 15 *Under the assumptions of Theorem 13 there exists a unique solution in* $BUC([0,T]\times\mathbb{R}^n)$ *to the RPDE* (4.3).

Proof Existence is clear from Theorem 13. Uniqueness is inherited from uniqueness to the Cauchy problem for $(\partial_t+\tilde{L})=0$ which follows from a comparison theorem for parabolic viscosity solution (c.f. Theorem 27 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 13, can be applied with almost every realization of Enhanced Brownian motion, $\boldsymbol{B}(\omega)$, 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.

$$L(t,x,r,p,X)=-\mathrm{Tr}\left[A(t,x)X\right]+b(t,x)\cdot p+c(t,x)r+f(t,x). \quad (5.1)$$

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

$$|f|_{H^m(\mathbb{R}^n)}=\left(\sum_{0\le k\le m}\left|\partial_{i_1},\ldots,\partial_{i_k}f\right|^p_{L^2(\mathbb{R}^n)}\right)^{1/2}$$

$H^m(\mathbb{R}^n)$ becomes a separable Hilbert space and the variational approach makes use of the triple

$$\left(H^m\left(\mathbb{R}^n\right)\right)\hookrightarrow L^2\left(\mathbb{R}^n\right)\backsimeq\left(L^2\left(\mathbb{R}^n\right)\right)^{\star}\hookrightarrow\left(H^m\left(\mathbb{R}^n\right)\right)^{\star}.$$

We make the following assumptions on the coefficients of L and Λ.

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

(1) a^{ij}, b_i, c, f as well as σ_k^i, ν_k^i, g are elements of[12] $C_b([0,T]\times\mathbb{R}^n,\mathbb{R})$ and σ_k^i, ν_k^i, g have one and a^{ij}, σ_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) $A = a\cdot a^T$, and $\exists\lambda > 0$ such that $\forall t\in[0,T]$

$$z^T \cdot A(t,x)\cdot z \geq \lambda\,|z|^2 \quad \forall x, z \in \mathbb{R}^n.$$

A L^2-solution is then defined as follows

Definition 16 Let $u_0 \in L^2(\mathbb{R}^n)$ and assume L and Λ 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

$$\begin{aligned} du &= Ludt + \Lambda u \circ dB_t \\ u(0,.) &= u_0(.) \end{aligned} \tag{5.2}$$

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 $\mathbb{P}(\int_0^T |u_r|^2_{H^1(\mathbb{R}^n)}dr < \infty) = 1$,
(2) $\forall\varphi\in C_0^\infty(\mathbb{R}^n)$ we have[13]

$$\langle u_., \varphi\rangle - \langle u_0,\varphi\rangle = \int_0^{.} \left\langle u_r, \widetilde{L}^\star\varphi\right\rangle dr + \sum_{k=1}^d \int_0^{.} \left\langle u_r, \Lambda_k^\star\varphi\right\rangle dB_r^k \quad (d\lambda\otimes d\mathbb{P}) - a.s., \tag{5.3}$$

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

Remark 17 The difference with the standard definition, cf. [41, Chapter IV, p130], is that we additionally assume enough regularity on the coefficients for the existence of the adjoint of $\widetilde{L}$ and to switch between divergence and non-divergence form.[14]

Theorem 18 *Under Assumption 3 there exists a unique L^2-solution of (5.2).*

Proof The standard variational approach as for example presented in [41, 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

[12] C_b denotes the bounded continuous functions and C_0 the subset of compactly supported functions.

[13] $\widetilde{L}^\star$ and $\Lambda^\star$ denote the formal adjoint operators of $\widetilde{L}$ and Λ, the stochastic integral is understood in the Itô sense and $\langle .,.\rangle$ denotes the scalar product on $L^2(\mathbb{R}^n)$.

[14] In the classic variational approach this can be avoided by working throughout with L in divergence form (resulting in no smoothness requirement on the coefficients in space instead of the existence of one derivative; in fact, except for the free terms, only boundedness and measurability of coefficients in combination with superparabolicity is sufficient, cf. [41, Chapter IV]).

16 as well as that $\forall \varphi \in C_0^\infty(\mathbb{R}^n)$ we have $\mathbb{P}$-a.s (using the Einstein summation convention)[15]

$$\langle u_t, \varphi \rangle - \langle u_0, \varphi \rangle = \int_0^t \left(-\left\langle \tilde{a}^{ij} \partial_j u_r, \partial_i \varphi \right\rangle + \left\langle \overline{b}_r^i \partial_i u_r + \tilde{c}_r u_r + \tilde{f}_r, \varphi \right\rangle \right) dr$$
$$+ \int_0^t \left\langle \sigma_k^j(r)\, \partial_j u_r + \nu_k(r)\, u_r + g_k(r), \varphi \right\rangle dB_r^k$$

where

$$\tilde{a}^{ij} = a^{ij} + \frac{1}{2}\sum_{k=1}^{d} \sigma_k^i \sigma_k^j$$
$$\tilde{b}^i = b^i + \frac{1}{2}\sum_{k=1}^{d} \left(\sigma_k^j \left(\partial_j \sigma_k^i \right) + 2\nu_k \sigma_k^i \right)$$
$$\tilde{c} = c + \frac{1}{2}\sum_{k=1}^{d} \left(\sigma_k^i \left(\partial_i \nu_k \right) + \nu_k^2 \right)$$
$$\tilde{f} = f + \frac{1}{2}\sum_{k=1}^{d} \left(\sigma_k^i \partial_i g + \nu_k g_k + g \right)$$

and

$$\overline{b}^i = \tilde{b}^i - \left(\partial_j \tilde{a}^{ij} \right)$$

(i.e. $\tilde{a}$ and $\tilde{b}$ are the coefficients that appear in $\widetilde{L}$ due to the switch from Itô to Stratonovich integration and $\overline{b}$ results from the switch to divergence form). Now using integration by parts we can rewrite the above divergence form into the adjoint formulation (5.3) as required by Definition 16. □

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 19 *Assume that L and Λ fulfill Assumption 3 as well as the assumptions of Theorem 13. If we denote with $u^{\boldsymbol{B}}$ the RPDE solution given in Theorem 13 driven by Enhanced Brownian motion $\boldsymbol{B}$ then $(u_t^{\boldsymbol{B}})_{t\geq 0}$ 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 Λ 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

[15] For $i, j \in \{1, \ldots, n\}$ and $k \in \{1, \ldots, d_1, \ldots, d_1 + d_2, \ldots, d_1 + d_2 + d_3\}$ and setting $\sigma_k = 0$ for $k > d_1$, $\nu_k = 0$ for $k \leq d_1$ or $k > d_1 + d_2$ and $g_k = 0$ for $k \leq d_1 + d_2$.

$$B_t^n = B_{t_{k-1}} + n\frac{(t-t_{k-1})}{T}\left(B_{t_k} - B_{t_{k-1}}\right)$$

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 18. Further, denote with $\boldsymbol{B}^n = S_2(B^n)$ the rough path lift of B^n and with $u^{\boldsymbol{B}_n}$ resp. $u^{\boldsymbol{B}}$ the viscosity solution for the random rough path $\boldsymbol{B}^n$ resp. $\boldsymbol{B}$ given in theorem 13. For $\epsilon > 0$ and $n \in \mathbb{N}$ the regularity of the coefficients allows to identify $u^{\boldsymbol{B}_n(\omega)}$ with $u(B^n(\omega))$ (both are the unique, bounded, smooth solutions of a parabolic PDE with smooth coefficients which depend on ω). The Wong–Zakai result in [29, Theorem 2.1] (all coefficients are smooth and additionally $f(t,.)$, $g(t,.)$ 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 $L^2([0,T], H^1)$ and hence also in $L^2([0,T]\times\mathbb{R}^n)$ (*much* more is true here of course). On the other hand, we know that a.s. $\boldsymbol{B}^n$ converges to $\boldsymbol{B}$ in rough path metric (see [21]) and from 13 we conclude that

$$u^{\boldsymbol{B}_n} \to u^{\mathbf{B}} \text{ a.s.}$$

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

$$u(B) = u^{\mathbf{B}} \text{ a.s}$$

(viewing $u(B)$ a.s. $C([0,T], L^2(\mathbb{R}^n))$-valued, this means that for a.e. $\omega, \forall t \in [0,T], u_t(B(\omega)) = u_t^{\mathbf{B}(\omega)}$ as equality in $L^2(\mathbb{R}^n)$; in particular, $u^{\mathbf{B}(\omega)}$ constitutes a continuous version in t, x; once more much more is true here). Further we know that $u(B) \in L^2([0,T], H^1(\mathbb{R}^n))$ $\mathbb{P}$-a.s. and hence fulfills Definition 16. Hence, we conclude that $(u_t^{\mathbf{B}})$ 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 ϵ^{-1} and smooth by convolution with a mollifier m^ϵ (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)[16] to arrive at the new operators L^ϵ and Λ^ϵ. It is easy to see that $L^\epsilon, \Lambda^\epsilon$ again fulfill Assumption 3, hence Theorem 18 applies and gives existence and uniqueness of the associated L^2-solution. Denote with $u^{\epsilon;\boldsymbol{B}}$ the associated RPDE solution[17] (with random rough path $\boldsymbol{B}$) and note that by step 1, $u^{\epsilon;\boldsymbol{B}}$ coincides with the L^2-solution. We now claim that

$$u^{\epsilon,\boldsymbol{B}} \to u^{\boldsymbol{B}} \quad a.s.$$

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

[17] With abuse of notation we identify the operators L^ϵ, 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.

uniformly on compacts. To see this, note that by the construction given in Theorem 13, $u^{\boldsymbol{B}}$ 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

$$\partial v_t^{\epsilon} + \overline{L}^{\epsilon}\left(t, x, v^{\epsilon}, Dv^{\epsilon}, D^2 v^{\epsilon}\right) = 0;$$

the precise form of $\overline{L}^{\epsilon}$ is as in Theorem 13 given by the transformation via rough path flows, that is

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

(where $\phi^{\boldsymbol{B},\epsilon}$, $\psi^{\boldsymbol{B},\epsilon}$ and $\varphi^{\boldsymbol{B},\epsilon}$ denote the rough flows associated with the truncated and mollified vector fields). Further we claim that the truncated and mollified Lip^{γ}-vector fields (appearing in Λ^{ϵ}) converge locally uniformly with locally uniform Lip^{γ} bounds: given $V \in Lip^{\gamma}$ denote V^{ϵ} as the vector field given by truncation outside radius ϵ^{-1} and convolution of V with m^{ϵ}. Of course, V^{ϵ} converges locally uniformly to V and (cf. [42, pp. 123, 159]) locally uniform Lip^{γ} bounds are readily seen to hold true for every $V^{\epsilon}, \epsilon > 0$. An interpolation argument then shows, locally, convergence in $Lip^{\gamma'}$ for $\gamma' < \gamma$ (we only need $\gamma' = \gamma - 1$). Given a geometric p-rough path z with $p < \gamma$ it then follows from [22, Corollary 10.39] (together with the remark that the $|.|_{Lip^{\gamma-1}}$-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 [22, Theorems 11.12 and 11.13] this convergence can be raised to the level of C^k–diffeomorphisms, provided V is assumed to be $Lip^{\gamma+k-1}$ 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 $\overline{L}^{\epsilon} \to [\phi^{\boldsymbol{B}}(L^{\psi^{\boldsymbol{B}}})]_{\varphi^{\boldsymbol{B}}}$ as $\epsilon \to 0$ uniformly on compacts and the stability properties of viscosity solutions guarantee (the same argument as given in Theorem 13) that $v^{\epsilon} \to v$, hence $u^{\boldsymbol{B},\epsilon} \to u^{\boldsymbol{B}}$ (loc. uniformly on $[0, T) \times \mathbb{R}^n$) a.s. From the first step it follows that $u^{\boldsymbol{B},\epsilon}$ is the unique L^2-solution, i.e. $\forall t \in [0, T]$

$$\left\langle u_t^{\boldsymbol{B},\epsilon}, \varphi\right\rangle - \langle u_0, \varphi\rangle = \int_0^t \left\langle u_t^{\boldsymbol{B},\epsilon}, \left(\widetilde{L}^{\epsilon}\right)^{\star} \varphi\right\rangle dr + \sum_{k=1}^{d} \int_0^t \left\langle u_r^{\boldsymbol{B},\epsilon}, \left(\Lambda_k^{\epsilon}\right)^{\star} \varphi\right\rangle dB_r^k$$

Sending $\epsilon \to 0$ in above equality shows that point (2) of Definition 16 is fulfilled. Now for every $\epsilon > 0$, classic variational arguments, see [41, Chapter 4, Theorem 1, p131], show that there exists a constant $C^{\epsilon} > 0$ which depends only on T, n, d and $\sup_{t,x,i,j,k}(|\tilde{a}^{\epsilon;ij}|, |b^{\epsilon;i}|, |\sigma_k^{\epsilon;i}|, |\nu_k^{\epsilon;i}|)$ (which is finite by Assumption 3) s.t.

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|u_r^{\epsilon,\boldsymbol{B}}\right|^2_{L^2(\mathbb{R}^n)}+\int_0^T\left|u_r^{\epsilon,\boldsymbol{B}}\right|^2_{H^1(\mathbb{R}^n)}dr\right]$$
$$\leq C^{\epsilon}\cdot\left[|u_0|^2_{L^2(\mathbb{R}^n)}+\mathbb{E}\int_0^T\left(\left|f_r^{\epsilon}\right|^2_{H^{-1}(\mathbb{R}^n)}+\sum_{k=d_1+d_2+1}^{d=d_1+d_2+d_3}\left|\left(g_r^{\epsilon}\right)^k\right|^2_{L^2(\mathbb{R}^n)}\right)dr\right].$$

By the estimate

$$\left|f_r^{\epsilon}\right|_{H^{-1}(\mathbb{R}^n)}\lesssim\left|f_r^{\epsilon}\right|_{L^2(\mathbb{R}^n)}=\left|\left(f(.)\,1_{|.|<\epsilon^{-1}}\right)\star m^{\epsilon}\right|_{L^2(\mathbb{R}^n)}\leq\left|f(.)\,1_{|.|<\epsilon^{-1}}\right|_{L^2(\mathbb{R}^n)}$$
$$\leq|f|_{L^2(\mathbb{R}^n)}$$

(and similarly for g^{ϵ}), the right-hand side can be uniformly bounded in ϵ, leading to the desired regularity properties of $u^{\boldsymbol{B}}$ (as required by point (2) in Definition 16). □

Remark 20 Classical L^2-theory of SPDEs gives, with probability one, $u(t,.,\omega)\in L^2(\mathbb{R}^n)$ 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 $u(t,.,\omega)$ takes values in higher Sobolev spaces $H^l(\mathbb{R}^n), l=1,2\ldots$. 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, p. 185].

5.2 A L^1_{loc}-solution

Theorem 13 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 σ_k^i,ν_k,g_k are in $C_b([0,T]\times\mathbb{R}^n)$ and σ_k^i,ν_k^i,g have one and a^{ij},σ_k^i 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 21 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 $\forall\varphi \in C_0^\infty(\mathbb{R}^n)$ we have

$$\langle u_., \varphi\rangle - \langle u_0, \varphi\rangle = \int_0^{\cdot} \left\langle u_r, \left(\widetilde{L}\right)^\star \varphi\right\rangle dr + \sum_{k=1}^{d} \int_0^{\cdot} \left(u, (\Lambda_k)^\star \varphi\right)_r dB_r^k \quad (d\lambda \otimes d\mathbb{P}) - a.s.$$

where $\widetilde{L}\varphi := L\varphi + \frac{1}{2}\sum_{k=1}^d \Lambda_k\Lambda_k\varphi$.

Remark 22 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 23 *Let B be a d-dimensional Brownian motion, $u_0 \in BUC$ and assume L, Λ fulfill the conditions of Theorem 13. Then $(u_t^{\boldsymbol{B}})_{t\in[0,T]}$ is a L^1_{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 $u_0^\epsilon \in C_c^\infty$ s.t. $u_0^\epsilon \to u_0$ uniformly on compacts in $\mathbb{R}^n$. Proposition 19 shows for $\epsilon > 0$ that the RPDE solution $u^\epsilon \in BUC([0,T]\times\mathbb{R}^n)$ associated with $(L^\epsilon, \Lambda^\epsilon, u_0^\epsilon, \boldsymbol{B})$ is (in the equivalence class of) the unique L^2-solution; especially u^ϵ is a L^1_{loc}-solution and therefore fulfills $(d\lambda\otimes d\mathbb{P}) - a.s.$ that

$$\left\langle u^\epsilon_., \varphi\right\rangle - \left\langle u^\epsilon_0, \varphi\right\rangle = \int_0^{\cdot} \left\langle u^\epsilon_r, \left(\widetilde{L^\epsilon}\right)^\star \varphi\right\rangle dr + \epsilon\int_0^{\cdot} \left\langle u^\epsilon_r, \sum_{i=1}^d \partial_i^2\varphi\right\rangle dr + \sum_{k=1}^d \int_0^{\cdot} \left\langle u^\epsilon_r, \Lambda_i^\star\varphi\right\rangle dB_r^k$$

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^ϵ 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 13 when applied to a stochastic driving signal, i.e. by taking $\mathbf{z}$ to be a realization of a continuous semi-martingale Z and its stochastic area, say $\mathbf{Z}(\omega) = (\mathbf{Z}, \mathbf{A})$; the most prominent example being Brownian motion and Lévy's area.

Remark 24 (*Itô versus Stratonovich*) Note that similar **SPDEs in Itô-form** need not be, in general, well-posed. Consider the following (well-known) linear example

$$du = u_x dB + \lambda u_{xx} dt, \quad \lambda \geq 0.$$

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 25 (*Regularity of noise coefficients*) Applied in the semimartingale context (finite p-variation for any $p > 2$) the regularity assumption of Theorem 13 reads $\mathrm{Lip}^{4+\varepsilon}$, $\varepsilon > 0$. 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. It 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 26 (*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 $\partial_t + [\phi^{\mathbf{z}}(L^{\psi^{\mathbf{z}}})]_{\alpha^{\mathbf{z}}}$. For instance, if the driving rough path $\mathbf{z}$ 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 $\partial_t + [\phi^{\mathbf{z}}(L^{\psi^{\mathbf{z}}})]_{\alpha^{\mathbf{z}}}$ 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 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 13).

(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-Loeve 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 $\mathbf{B}$ (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 $\mathbf{B}$ 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 [24–28].

(Large deviation results) Another application of our continuity result is the ability to obtain large deviation estimates when B is replaced by εB 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 $\mathbf{z}$ 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 $\mathbf{z}$ 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, Λ and Λ_k instead of $L(t, x, u, Du, D^2u)$, $\Lambda(t, x, u, Du)$ 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

$$du + Ldt = \Lambda d\mathbf{Z}$$

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

$$\log \tilde{\mathbf{Z}} := \log \mathbf{Z} + (\mathbf{0}, \Gamma_{\mathbf{t}})$$

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

$$d\tilde{u} + Ldt = \Lambda d\tilde{\mathbf{Z}}$$

is identified with

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

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,

$$\left(Z^n, \int Z^n dZ^n\right) \to \tilde{\mathbf{Z}}$$

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

$$d\tilde{u} + Ldt = \sum_{k=1}^{d} \Lambda_k \circ dZ^k + \sum_{i,j\in\{1,\ldots,d\}} \left[\Lambda_i, \Lambda_j\right] d\Gamma^{i,j}.$$

The formal proof is easy; it suffices to analyze the equations (rough) differential equations for (ψ, ϕ, α) in presence of area perturbation; see [18], and then identify the corresponding operators $[^{\phi}[L^{\psi}]]_{\alpha}$. 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 $\Lambda_{\gamma} = [\Lambda_{\gamma_1}, \ldots, [\Lambda_{\gamma_{N-1}}, \Lambda_{\gamma_N}] \ldots]$, 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 ε-delay; say

$$du^{\varepsilon} + Ldt = \Lambda_1 \circ dB^{\varepsilon}_{\cdot-\varepsilon} + \Lambda_2 \circ dB.$$

where $B^{\varepsilon}_{t-\varepsilon} := B_{t-\varepsilon}$. 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^{ε}_t, 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

$$du + Ldt = (\Lambda_1 + \Lambda_2) \circ dB + [\Lambda_1, \Lambda_2]\, dt$$

(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

$$\begin{aligned} dX_t &= \mu(X_t)\, dt + V(X_t)\, dB_t + \sigma(X_t)\, d\tilde{B}_t \\ dY_t &= h(X_t)\, dt + d\tilde{B}_t \end{aligned} \tag{6.1}$$

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

$$\pi_t(\varphi) = \mathbb{E}\left[\varphi(X_t) \,|\sigma(Y_r, r \le t)\right]$$

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

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

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 ϕ^{φ}_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 $\overline{\phi}_t^{\varphi} : C([0,T], \mathbb{R}^n) \to \mathbb{R}$ which is continuous in uniform norm and fulfills (6.2), thus providing a version of the conditional expecation $\pi_t(\varphi)$ which is robust under approximations in uniform norm of the observation Y. Unfortunately in the correlated noise case this is no longer true![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 [1,39,41], π_t can be written in the form

$$\pi_t(\varphi) = \int_{\mathbb{R}^{d_X}} \varphi(x) \frac{u_t(x)}{\int u_t(\tilde{x})\, d\tilde{x}} dx \tag{6.3}$$

where $u_t \in L^1(\mathbb{R}^n)$ a.s. and (u_t) is the L^2-solution of the (dual) Zakai SPDE

$$\begin{aligned} du_t &= G^\star dt + \sum_k N_k u_t dY_t^k \\ &= \left(G^\star + \frac{1}{2} \sum_k N_k N_k \right) u_t dt + \sum_k N_k u_t \circ dY_t^k \end{aligned} \tag{6.4}$$

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

$$(N_k u)(t,x) = \sigma_k^i(x)\, \partial_i u_t(x) + h(x)\,.u_t(x)\,. \tag{6.5}$$

Using Theorem (13) in combination with Proposition 19 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 $u_t^z \in L^1$ (i.e. a rough pathwise version of the discussion in [41, 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.

[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].

Acknowledgments 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.

Appendix: comparison for parabolic equations

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

$$\partial_t u + F\left(t, x, u, Du, D^2 u\right) \le 0$$

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

Theorem 27 *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*

$$F(t,x,r,p,X) - F(t,x,s,p,X) \ge C(r-s) \quad \forall r \le s, \tag{7.1}$$

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

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

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 28 *Under the assumptions of the theorem above let u, v be two solutions, with initial data u_0, v_0 respectively. Then*

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

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,

$$\tilde{u} \le \tilde{v} + |u_0 - v_0|_{\infty;\mathbb{R}^n}.$$

Applying the same reasoning to $\tilde{u} + |u_0 - v_0|_{\infty;\mathbb{R}^n}$ and finally undoing the transformation gives the result. □

References

1. Bain, A., Crişan, D.: Fundamentals of Stochastic Filtering. Applications of Mathematics. Springer, Berlin (2008)
2. Breuillard, E., Friz, P.K., Huesmann, M.: From random walks to rough paths. Proc. AMS, pp. 3487–3496 (2009)
3. Brzeźniak, Z., Flandoli, F.: Almost sure approximation of Wong-Zakai type for stochastic partial differential equations. Stoch. Process. Appl. **55**(2), 329–358 (1995)
4. Caruana, M., Friz, P.K., L Gyurko.: Rough Differential Equations With Time-Dependent Vector Fields. Working paper (2012)
5. Caruana, M., Friz, P.K.: Partial differential equations driven by rough paths. J. Differ. Eq. **247**(1), 140–173 (2009)
6. Caruana, M., Friz, P.K., Oberhauser, H.: A (rough) pathwise approach to a class of non-linear stochastic partial differential equations. Ann. Inst. H. Poincaré Anal. Non Linéaire **28**(1), 27–46 (2011)
7. Clark, J.M.C.: The design of robust approximations to stochastic differential equations in nonlinear filtering. In: Skwirzynsky, J.K. (ed.) Communication Systems in Random Processes Theory. Sijthoff, Nordhoff (1978)
8. Coutin, L., Qian, Z.: Stochastic analysis, rough path analysis and fractional Brownian motions. Probab. Theory Related Fields **122**(1), 108–140 (2002)
9. Crandall, M.G., Ishii, H., Lions, P.-L.: User's guide to viscosity solutions of second order partial differential equations. Bull. Am. Math. Soc. (N.S.) **27**(1), 1–67 (1992)
10. Crisan, D., Diehl, J., Friz, P.K., Oberhauser, H.: Robust filtering: multdimensional noise and multidimensional observation. Ann. Appl. Probab. (in press) (2012)
11. Davis, M.: Pathwise nonlinear filtering with noise correlation. In: Crisan, D., Rozovskiĭ, B. (eds.) Oxford Handbook of Nonlinear Filtering. Oxford University Press, Oxford (2010)
12. Deya, A., Gubinelli, M., Tindel, S.: Non-linear rough heat equations. Probab. Theory Related Fields, pp. 1–51 (2011)
13. Diehl, J., Friz, P.K.: Backward stochastic differential equations with rough drivers. Ann. Probab. **40**, 1715–1758 (2012)
14. Diehl, J., Friz, P., Oberhauser, H.: Parabolic comparison revisited and applications. Submitted (2011)
15. Fleming, W.H., Mete Soner, H.: Controlled Markov Processes and Viscosity Solutions, volume 25 of Stochastic Modelling and Applied Probability, second edn., Springer, New York (2006)
16. Friz, P.K., Lyons, T., Stroock, D.: Lévy's area under conditioning. Ann. Inst. H. Poincaré Probab. Statist. **42**(1), 89–101 (2006)
17. Friz, P.K., Victoir, N.: Differential equations driven by gaussian signals. In: Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, vol. 46, pp. 369–413. Institut Henri Poincaré (2010)
18. Friz, P.K., Oberhauser, H.: Rough path limits of the Wong-Zakai type with a modified drift term. J. Funct. Anal. **256**, 3236–3256 (2009)
19. Friz, P.K., Victoir, N.: Approximations of the Brownian rough path with applications to stochastic analysis. Ann. Inst. H. Poincaré Probab. Statist. **41**(4), 703–724 (2005)
20. Friz, P.K., Victoir, N.: On uniformly subelliptic operators and stochastic area. Probab. Theory Related Fields **142**(3–4), 475–523 (2008)
21. Friz, P.K.: Continuity of the Itô-map for Hölder rough paths with applications to the support theorem in Hölder norm. In: Probability and partial differential equations in modern applied mathematics, vol. 140 of IMA Vol. Math. Appl., pp. 117–135. Springer, New York (2005)
22. Friz, P.K., Victoir, N.B.: Multidimensional stochastic processes as rough paths: theory and applications. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge (2010)
23. Gubinelli, M., Tindel, S.: Rough evolution equations. Ann. Prob. **38**, 1–75 (2010)
24. Gyöngy, I.: The stability of stochastic partial differential equations and applications. I. Stoch. Stoch. Rep. **27**(2), 129–150 (1989)
25. Gyöngy, I.: The stability of stochastic partial differential equations and applications. Theorems on supports. In: Stochastic partial differential equations and applications, II (Trento, 1988), volume 1390 of Lecture Notes in Math., pp. 91–118. Springer, Berlin (1989)
26. Gyöngy, I.: The stability of stochastic partial differential equations. II. Stoch. Stoch. Rep. **27**(3), 189–233 (1989)
27. Gyöngy, I.: The approximation of stochastic partial differential equations and applications in nonlinear filtering. Comput. Math. Appl. **19**(1), 47–63 (1990)

28. Gyöngy, I.: On stochastic partial differential equations. Results on approximations. In: Topics in stochastic systems: modelling, estimation and adaptive control, volume 161 of Lecture Notes in Control and Inform. Sci., pp. 116–136. Springer, Berlin (1991)
29. Gyöngy, I., Shmatkov, A.: Rate of convergence of Wong-Zakai approximations for stochastic partial differential equations. Appl. Math. Optim. **54**(3), 315–341 (2006)
30. Krylov, N.V.: An analytic approach to SPDEs. In: Stochastic partial differential equations: six perspectives, volume 64 of Math. Surveys Monogr., pp. 185–242. Am. Math. Soc., Providence, RI (1999)
31. Krylov, N.V., Rozovskiĭ, B.L.: Itô equations in Banach spaces and strongly parabolic stochastic partial differential equations. Dokl. Akad. Nauk SSSR **249**(2), 285–289 (1979)
32. Krylov, N.V.: An analytical approach to SPDEs. In: Carmona, R., Rozovskiĭ, B. (eds), Stochastic Partial Differential Equations: Six Perspectives, pp. 185–242. AMS (1999)
33. Krylov, N.V.: Lectures on Elliptic and Parabolic Equations in Sobolev Spaces, Graduate Studies in Mathematics (2008)
34. Ledoux, M., Qian, Z., Zhang, T.: Large deviations and support theorem for diffusion processes via rough paths. Stoch. Process. Appl. **102**(2), 265–283 (2002)
35. Lions, P.L., Souganidis, P.E.: Fully nonlinear stochastic pde with semilinear stochastic dependence. C. R. Acad. Sci. Paris Sér. I Math. **331**(8), 617–624 (2000)
36. Lyons, T.J.: Differential equations driven by rough signals. Rev. Mat. Iberoamericana **14**(2), 215–310 (1998)
37. Lyons, T.J., Qian, Z.: System Control and Rough Paths. Oxford Mathematical Monographs. Oxford University Press, Oxford (2002)
38. Lyons, T.J., Caruana, M., Lévy, T.: Differential equations driven by rough paths, volume 1908 of Lecture Notes in Mathematics (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). Springer, Berlin (2007)
39. Pardoux, E.: Stochastic partial differential equations and filtering of diffusion processes. Stochastics **3**(2), 127–167 (1979)
40. Prévôt, C., Röckner, M.: A concise course on stochastic partial differential equations, volume 1905 of Lecture Notes in Mathematics. Springer, Berlin (2007)
41. Rozovskiĭ, B.L.: Stochastic evolution systems, volume 35 of Mathematics and its Applications (Soviet Series). Kluwer Academic Publishers Group, Dordrecht: Linear theory and applications to nonlinear filtering. Translated from the Russian by A, Yarkho (1990)
42. Stein, E.M.: Singular integrals and differentiability properties of functions. Princeton mathematical series. Princeton University Press, Princeton (1970)
43. Teichmann, J.: Another approach to some rough and stochastic partial, differential equations. arXiv:0908.2814 (2009)
44. Tessitore, G., Zabczyk, J.: Wong-Zakai approximations of stochastic evolution equations. J. Evol. Equ. **6**(4), 621–655 (2006)
45. Twardowska, K.: Wong-Zakai approximations for stochastic differential equations. Acta Appl. Math. **43**(3), 317–359 (1996)