Rough path stability of SPDEs arising in non-linear filtering

Peter Friz
TU Berlin and WIAS Berlin

Harald Oberhauser
TU Berlin

May 11, 2010

Abstract

We prove a longstanding conjecture [Lyons, T. J.; Differential equations driven by rough signals. Rev. Mat. Iberoamericana 14 (1998), no. 2, 215–310] concerning the applicability of rough path analysis for stochastic partial differential equations arising from the theory of non-linear filtering.

1 Introduction

"*Stochastic filtering is concerned with the estimation of the conditional law of a Markov process, given observations of some function of it. The normal formulation (due to Zakai) looks at the case where the process is of diffusion type and splits into a first part (known as the signal) and a second part, known as the observation process with values in a vector space, and whose martingale part has stationary increments independent of the signal. In this case, Zakai showed that it was possible to completely describe the conditional density of the signal given knowledge of the observation process. In fact, the density evolves according to an SPDE of parabolic type. It is a commutative equation, and so the relationship between the observation process and the conditional density is a relatively stable one. On the other hand, it is really rather rare that real filtering problems present themselves with the noise in the observation process being independent of the signal. And the transformation involved in making it so involves the solution of a generic SDE which will not commute. It follows that to do robust and stable filtering it is important to measure the area process as well as the values of the observation process.*"

The preceding paragraph is taken from (the motivating introduction of) Lyons' seminal 1998 paper [26] on rough paths, a theory that proposes to understand paths together with their areas in order to guarantee robustness of differential equations driven by such paths. In other words, we have the conjecture that rough path theory is relevant for non-linear filtering via the stability analysis of Zakai-type SPDEs; a conjecture supported in particular by the works of I. Gyöngy [21, 20] who obtained explicit results of how "correct" approximations to the observation process, but wrong approximation of its area process, lead to convergence to the "wrong" solution. A quantitative form of this conjecture is the assertion that solutions to the filtering problem (i.e. the afore-mentioned SPDEs) depend continuously on the observation process *seen as a rough path*. The proof of this is the content of this paper. More precisely, we consider the SPDE[1]

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

with scalar initial data $u(0,\cdot) = u_0(\cdot)$, assumed to be BUC (bounded and uniformly continuous) on $\mathbb{R}^n$, and give meaning and rough path stability for the solution to the above equations as a function of $\mathbf{Y}$, the (canonical semi-martingale) enhancement of the d-dimensional observation process Y. Here L is a linear second order operator of the form

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

Actually, our approach allows to deal with L semi-linear in the sense that f above may depend non-linearly on r with no extra effort. The first order operators $\Lambda_k = \Lambda_k(t, x, r, p)$ are affine linear in r, p. That is,

$$\Lambda_k(t, x, r, p) = \left(p \cdot \sigma_k(t, x)\right) + r\, \nu_k(t, x) + g_k(t, x).$$

We shall prefer to write the rhs of (1) in the equivalent form

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

where $Y \equiv (Y^1, Y^2, Y^3)$ is a $(d_1 + d_2 + d_3)$-dimensional semimartingale. Our approach, however, is based on a pointwise (viscosity) interpretation of (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

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

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 v 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 is seen to be 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 $\tilde{v}$ 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, first introduced for a class non-linear problems [4], then play together to yield the desired result. Upon using the canonical rough path lift of the observation process in this RPDE, a truly (rough)pathwise point of view, one has constructed a robust version of the SPDE solution in equation. We note that the Stratonovich form of our SPDE allows us to *avoid any ellipticity assumption on L*; we can

[1] Using similar notation as in [22].

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

even handle the fully degenerate first order case. In turn, we only obtain BUC solutions. Stronger assumptions would allow to discuss all this in a classical context (i.e. $\tilde{v}$ would be a $C^{1,2}$ solution); SPDE solution can then be seen to have certain spatial regularity.

In the presence of gradient (rough) noise only this approach was implemented in [3]. We should remark that the usual way to deal with (1), which goes back to Krylov, Rozovskii, Pardoux and others, is to find solutions in a suitable functional analytic setting; e.g. such that solutions evolve in suitable Sobolev spaces. Interestingly, there has been no success until now (despite the advances by Gubinelli–Tindel [14] and Teichmann [30]) to include (1) in a setting of abstract *rough* evolution equations on infinite-dimensional spaces.

Acknowledgement 1 *Part of this work was carried out at the Newton Institute in Cambridge during the SPDE program in Spring 2010. HO would like to express his gratitude for a Junior membership grant.*

2 Recalls on viscosity theory and rough paths

Let us recall some basic ideas of (second order) viscosity theory [6, 7] and rough path theory [27, 28]. 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 $F(t,x,r,p,A+B) \leq F(t,x,r,p,A)$ whenever $B \geq 0$ in the sense of symmetric matrices. The idea is to consider a (smooth) test function φ and look at a local maxima $\left(\hat{t}, \hat{x}\right)$ of $u - \varphi$. Basic calculus implies that $Du\left(\hat{t}, \hat{x}\right) = D\varphi\left(\hat{t}, \hat{x}\right), D^2 u\left(\hat{t}, \hat{x}\right) \leq D\varphi\left(\hat{t}, \hat{x}\right)$ and, from degenerate ellipticity,

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

This suggests to define a *viscosity supersolution* (at the point $\left(\hat{x}, \hat{t}\right)$) to $\partial_t + F = 0$ as a continuous function u with the property that (2) holds for any test function. Similarly, *viscosity subsolutions* are defined by reversing inequality in (2); *viscosity solutions* are both super- and subsolutions. A different point of view is to note that $u(t,x) \leq u\left(\hat{t}, \hat{x}\right) - \varphi\left(\hat{t}, \hat{x}\right) + \varphi(t,x)$ for (t,x) near $\left(\hat{t}, \hat{x}\right)$. A simple Taylor expansion then implies

$$u(t,x) \leq u\left(\hat{t}, \hat{x}\right) + a\left(t - \hat{t}\right) + p \cdot (x - \hat{x}) + \frac{1}{2}(x - \hat{x})^T \cdot X \cdot (x - \hat{x}) + o\left(|\hat{x} - x|^2 + \left|\hat{t} - t\right|\right) \tag{3}$$

as $|\hat{x} - x|^2 + \left|\hat{t} - t\right| \to 0$ with $a = \partial_t \varphi\left(\hat{t}, \hat{x}\right), p = D\varphi\left(\hat{t}, \hat{x}\right), X = D^2\varphi\left(\hat{t}, \hat{x}\right)$. Moreover, if (3) holds for some (a, p, X) and u is differentiable, then $a = \partial_t u\left(\hat{t}, \hat{x}\right), p = Du\left(\hat{t}, \hat{x}\right), X \leq D^2 u\left(\hat{t}, \hat{x}\right)$, hence by degenerate ellipticity

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

Pushing this idea further leads to a definition of viscosity solutions based on a generalized notion of "$\left(\partial_t u, Du, D^2 u\right)$" for nondifferentiable 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 $u_0 \geq v_0$; 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 straight-forward 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

$$\dot{y}(t) = \sum_{i=1}^{d} V_i(y) \dot{z}^i(t), \quad y(0) = y_0 \in \mathbb{R}^n. \tag{4}$$

The question is, if the output signal y depends in a stable way on the driving signal z. The answer, of course, depends strongly on how to measure distance between input signals. If one uses the ∞ norm, so that the distance between driving signals $z, \tilde{z}$ is given by $|z - \tilde{z}|_{\infty;[0,T]}$, 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,*

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

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

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

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 (4): 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

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

[3] BUC $(\ldots)$ denotes the space of bounded, uniformly continuous functions.
[4] (3.14) of the User's Guide [6].
[5] What we have in mind here is the *Barles–Perthame method of semi-relaxed limits* [7].

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 (4), 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,

$$C^{0,p\text{-var}}\left([0,T], G^{[p]}\left(\mathbb{R}^d\right)\right) \text{ resp. } C^{0,1/p\text{-Höl}}\left([0,T], G^{[p]}\left(\mathbb{R}^d\right)\right)$$

where $G^{[p]}\left(\mathbb{R}^d\right)$ is the free step-$[p]$ nilpotent group over $\mathbb{R}^d$, equipped with Carnot–Caratheodory metric; realized as a subset of $1 + \mathfrak{t}^{[p]}\left(\mathbb{R}^d\right)$ 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 statespace 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

$$B^{(2)} := \left(\int_0^{\cdot} B^i \circ dB^j\right)_{i,j\in\{1,\ldots,d\}} \tag{5}$$

yields a path $\mathbf{B}(\omega)$ in $G^2\left(\mathbb{R}^d\right)$ 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 $B^{(2)} = \frac{1}{2} B \otimes B + A$ where $A := \text{Anti}\left(B^{(2)}\right)$ 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 work when B is replaced by a generic multi-dimensional continuous semimartingales; see [13, 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 non-increasing in X (degenerate elliptic). 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. $\langle p, q \rangle_{j_1,\ldots,j_n} \equiv \sum_{i_1,\ldots,i_m} p_{i_1,\ldots,i_m} q^{i_1,\ldots,i_m}_{j_1,\ldots,j_n}$, $p \in \left(\mathbb{R}^l\right)^{\otimes m}$, $q \in \left(\mathbb{R}^l\right)^{\otimes n} \otimes \left(\left(\mathbb{R}^l\right)'\right)^{\otimes m}$.

Lemma 3 (Inner Transform) *Let $z \in C^1\left([0,T], \mathbb{R}^d\right)$, $\sigma = (\sigma_1, \ldots, \sigma_d) \subset C_b^2(\mathbb{R}^n, \mathbb{R}^n)$ and $\psi = \psi(t,x)$ the ODE flow of $dy = \sigma(t,y)\,dz$, i.e.*

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

[6] In the strict terminology of rough path theory: geometric p-rough paths.
[7] We also use $\cdot$ to denote contraction over only index or to denote matrix multiplication.

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

$$\partial_t u + F\left(t, x, r, Du, D^2 u\right) = \sum_{i=1}^{d} \left(Du \cdot \sigma_i\left(t, x\right)\right) \dot{z}_t^i; \;\; u\left(0, .\right) = u_0\left(.\right) \tag{6}$$

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

$$\partial_t w + F^{\psi}\left(t, x, w, Dw, D^2 w\right) = 0; \;\; w\left(0, .\right) = u_0\left(.\right) \tag{7}$$

where

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

and

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

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

Corollary 4 (Transformation 1) *Let $\psi = \psi\left(t, x\right)$ be the ODE flow of $dy = \sigma\left(t, y\right) dz$, as above. Define $L = L\left(t, x, r, p, X\right)$ by*

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

define also the transform

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

where

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

Then u is a solution (always assumed BUC) of

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

if and only if $w\left(t, x\right) := u\left(t, \psi\left(t, x\right)\right)$ is a solution of

$$\partial_t w + L^{\psi}\left(t, x, w, Dw, D^2 w\right) = 0; \;\; w\left(0, .\right) = u_0\left(.\right) \tag{8}$$

Proof. Set $y = \psi_t\left(x\right)$. When u is a classical sub-solution, it suffices to use the chain-rule and definition of F^{ψ} to see that

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

The case when u is a viscosity sub-solution of (6) 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 $\varphi(t,y) = \xi\left(t,\psi_t^{-1}(y)\right)$. Set $\bar{y}=\psi_{\bar{t}}(\bar{x})$ so that

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

Obviously, $(\bar{t},\bar{y})$ is a maximum of $u-\varphi$, and since u is a viscosity sub-solution of (6) we have

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

On the other hand, $\xi(t,x)=\varphi(t,\psi_t(x))$ implies $\dot{\xi}(\bar{t},\bar{x}) = \dot{\varphi}(\bar{t},\bar{y}) + D\varphi(\bar{t},\bar{y})V(\bar{y})\dot{z}(\bar{t})$ and putting things together we see that

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

which says precisely that w is a viscosity sub-solution of (7). Replacing maximum by minimum and $\le$ by $\ge$ in the preceding argument, we see that if u is a super-solution of (6), then w is a super-solution of (7).
Conversely, the same arguments show that if v is a viscosity sub- (resp. super-) solution for (7), then $u(t,y) = w\left(t,\psi_t^{-1}(y)\right)$ is a sub- (resp. super-) solution for (6). ■

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

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

Lemma 5 (Outer transform) *Let $\phi=\phi(t,r,x)\in C^{1,2,2}$ and strictly increasing in r. (It follows that $\phi(t,.,x)$ is an increasing diffeomorphism on the real line). Then u is a subsolution of $\partial_t u + F\left(t,x,r,Du,D^2u\right)=0$, $u(0,.)=u_0(.)$ if and only if*

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

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

$$\begin{aligned}
{}^{\phi}F(t,x,r,p,X) &= \frac{\dot{\phi}}{\phi'} + \frac{1}{\phi'}F\big(t,x,\phi,D\phi+\phi' p, \\
&\quad \phi'' p\otimes p + D\phi'\otimes p + p\otimes D\phi' + D^2\phi + \phi' X\big)
\end{aligned} \tag{9}$$

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, $(a,p,X)\in\mathcal{P}^{2;+}v(s,z)$ iff

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

Since $\phi(t,.,x)$ is increasing,

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

and using a Taylor expansion on ϕ in all three arguments we see that the rhs equals

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

Hence,

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

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

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

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 (9) for some F. Set $u(t,x) := \phi(t,v(t,x),x)$. By above argument we know that v is a subsolution of ${}^{\phi^{-1}}\left({}^{\phi}F\right)(t,x,r,p,X)$. For brevity write $\psi(t,.,x) = \phi^{-1}(t,.,x)$. Then

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

Using several times equalities of the type $\left(f\circ f^{-1}\right)'=f'_{f^{-1}}\left(f^{-1}\right)'=id$ 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 6 (Transformation 2) *Assume* $\nu=(\nu_1,\dots,\nu_d)\subset C_b^{0,2}([0,T]\times\mathbb{R}^n)$. *Assume* $\phi=\phi(t,x,r)$ *is determined by the ODE*

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

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

$$\begin{aligned}
{}^{\phi}L(t,x,r,X) &= -Tr[A(t,x)\cdot X]+{}^{\phi}b(t,x)\cdot p+{}^{\phi}c(t,x,r)\;\;\textit{where} \qquad (10)\\
{}^{\phi}b(t,x)\cdot p &\equiv b(t,x)\cdot p-\frac{2}{\phi'}Tr\left[A(t,x)\cdot D\phi'\otimes p\right]\\
{}^{\phi}c(t,x,r) &\equiv -\frac{1}{\phi'}Tr\left[A(t,x)\cdot(D^2\phi)\right]+\frac{1}{\phi'}b(t,x)\cdot(D\phi)+\frac{1}{\phi'}c(t,x,\phi)
\end{aligned}$$

where ϕ *and all its derivatives are evaluated at* (t,r,x). *Then*

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

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

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

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

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

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

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

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 (10). ■

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

Lemma 7 (Transformation 3) *Set* $\alpha(t,x) = \int_0^t g(s,x) \cdot \dot{z}_s$. *Define*

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

Then v *solves*

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

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

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

Proof. Left to reader. ■

3.2 The full transformation

As before, let

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

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,

$$\begin{aligned} M_k(t,x,u,Du) &= Du\cdot\sigma_k(t,x) \text{ for } k=1,\dots,d_1, \\ M_{d_1+k}(t,x,u,Du) &= u\,\nu_k(t,x) \text{ for } k=1,\dots,d_2, \\ M_{d_1+d_2+k}(t,x,u,Du) &= g_k(t,x) \text{ for } k=1,\dots,d_3. \end{aligned} \tag{11}$$

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

Proposition 8 *Let* $z^1 \in C^1\left([0,T],\mathbb{R}^{d_1}\right)$, $\sigma = (\sigma_1,\dots,\sigma_{d_1}) \subset C_b^{0,2}([0,T]\times\mathbb{R}^n,\mathbb{R}^n)$ *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*

$$\dot{\psi}(t,x) = \sigma(t,\psi(t,x))\,\dot{z}_t^1,\ \psi(0,x) = x \in \mathbb{R}^n. \tag{12}$$

Further, let $z^2 \in C^1\left([0,T],\mathbb{R}^{d_2}\right)$ *and let* $\nu = (\nu_1,\dots,\nu_{d_2})$ *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*

$$\dot{\phi} = \phi\,\underbrace{\nu(t,\psi_t(x))}_{\equiv \nu^{\psi}(t,x)}\dot{z}_t^2,\ \ \phi(0,r,x) = r \in \mathbb{R}. \tag{13}$$

Further, let $z^3 \in C^1\left([0,T],\mathbb{R}^{d_3}\right)$ *and define* $\alpha(t,x)$ *as the integral*[8]

$$\alpha(t,x) = \int_0^t {}^{\phi}\left(g^{\psi}\right)(s,x)\cdot \dot{z}_s^3, \tag{14}$$

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

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

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

$$\begin{aligned}\partial_t u+L\left(t,x,u,Du,D^2u\right)&=\Lambda(t,x,u,Du)\,\dot{z}_t,\\ u(0,x)&=u_0(x),\end{aligned}\qquad(15)$$

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

$$\begin{aligned}\partial_t\tilde{v}+\tilde{L}\left(t,x,\tilde{v},D\tilde{v},D^2\tilde{v}\right)&=0\\ \tilde{v}(0,x)&=u_0(x)\end{aligned}\qquad(16)$$

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

Remark 9 *Transformation 2 and 3 could have been performed in one step, by considering*

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

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 separate the effects of this; with the benefit of keeping the algebra simpler.

Remark 10 *Related to the last remark, generic noise of the form* $H(x,t,u)\,dz$ *can be removed with this technique. The issue is that the transformed equations quickly falls beyond available theory, cf. [25].*

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 so-transformed solution, $w(t,x)=u\left(t,\psi_t(x)\right)$, satisfies the equation

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

We then remove the terms driven by z^2 by applying transformation 2 with $L^{\psi}-g^{\psi}\cdot\dot{z}^3$. The so-transformed solution $v(t,x)=\phi^{-1}\left(t,w(t,x),x\right)$ satisfies the equation with operator $\left(\partial_t+{}^{\phi}\left(L^{\psi}-g^{\psi}\cdot\dot{z}^3\right)\right)$; i.e.

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

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

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

The proof is now finished. ∎

3.3 Rough transformation

We need to understand transformations 1,2,3 when $\left(z^1, z^2, z^3\right)$ becomes a rough path, say $\mathbf{z}$. It comes as no surprise that there will be some dependency between the flows ψ, ϕ and α since we solve the "stochastic characteristics" of a system of the form $\partial_t w = \sum_i (\sigma_i \cdot Dw) z^{1;i} + w \sum_j \nu_j z^{2;j} + \sum_k g_k \cdot z^{3;k}$. In fact there is a "triadiagonal" structure: (12) can be solved as function of z^1 alone;

$$d\psi_t(x) = \sigma(t, \psi_t(x))\, dz_t^1 \text{ with } \psi_0(x) = x. \tag{17}$$

(13) is then tantamount to

$$\phi(t, r, x) = r \exp\left[\int_0^t \nu(s, \psi_s(x))\, dz_s^2\right]. \tag{18}$$

As for $\alpha = \alpha(t, x)$, 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

$$\alpha(t, x) = \int_0^t \tilde{\phi}(s, x)\, g(s, \psi_s(x))\, dz_s^3. \tag{19}$$

Lemma 11 *Let $\mathbf{z}$ be a geometric p-rough path; that is, an element in $C^{0,p\text{-}var}\left([0,T], G^{[p]}\left(\mathbb{R}^{d_1+d_2+d_3}\right)\right)$. Let $\gamma > p \geq 1$. Assume*

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

Then ψ, ϕ and α depend (in local uniform sense) continuously on $\left(z^1, z^2, z^3\right)$ 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 α. In particular, we can define ψ, ϕ and α when $\left(z^1, z^2, z^3\right)$ is replaced by a genuine geometric p-rough path $\mathbf{z}$ and write $\psi^{\mathbf{z}}, \phi^{\mathbf{z}}, \alpha^{\mathbf{z}}$ to indicate this dependence.

Proof. Given $\mathbf{z}$ one can build a canonical rough path $(\mathbf{t}, \mathbf{z})$ of the "time-space path" $\left(t, z^1, z^2, z^3\right)$, a special case of Young pairings of rough paths, [13, Chapter 14]. Define the following vectorfields on $\mathbb{R}^{1+n+3}$,

$$V_1(y) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, V_2(y) = \begin{pmatrix} 0 \\ \sigma(\psi, t) \\ 0 \\ 0 \\ 0 \end{pmatrix}, V_3(y) = \begin{pmatrix} 0 \\ 0 \\ \phi.\nu(t, \psi) \\ -\tilde{\phi}.\nu(t, \psi) \\ 0 \end{pmatrix}, V_4(y) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \tilde{\phi}.g(t, \psi) \end{pmatrix}.$$

where $y = \left(t, \psi, \phi, \tilde{\phi}, \alpha\right)^T \in \mathbb{R}^{1+n+3}$. The assumptions on σ, ν and g guarantee that $V = (V_1, \ldots, V_4) : \mathbb{R}^{1+n+3} \to L\left(\mathbb{R}^{1+d_1+d_2+d_3}, \mathbb{R}^{1+n+3}\right)$ is a collection of $Lip_{loc}^{\gamma-1}\left(\mathbb{R}^{1+n+3}\right)$ vector fields.

Hence, the "full" RDE

$$dy_t = V(y_t)\, d(\mathbf{t}, \mathbf{z}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \sigma(t,\psi) & 0 & 0 \\ 0 & 0 & \phi.\nu(t,\psi) & 0 \\ 0 & 0 & -\tilde{\phi}.\nu(t,\psi) & 0 \\ 0 & 0 & 0 & \tilde{\phi}.g(t,\psi) \end{pmatrix} d(\mathbf{t}, \mathbf{z}) \tag{20}$$

with initial condition $\begin{pmatrix} 0, & x, & r, & 1, & 0 \end{pmatrix}^T \in \mathbb{R}^{1+n+3}$ at time 0 has either a solution on $[0,T]$ or a solution on $[0,\tau] \subset [0,T]$ which explodes at a time $\tau < T$. Exploiting the "tridiagonal" structure described above we can show that the claimed $(\gamma-1)$-Lipschitz regularity is enough for uniqueness and exclude explosion (for details on rough integration and rough flows, cf. [13],[28],[27]) : First note that the assumption $\sigma \in Lip^\gamma$ immediately guarantees a unique solution ψ on $[0,T]$ of

$$d\psi = \sigma(t,\psi)\, d\mathbf{z}. \tag{21}$$

One can then construct the rough path lift of (t, ψ, z^2), again using Young pairings and the fact that (21) provides us with iterated integrals of ψ. Rough integration of this Young-pairing against the 1-form $\varphi = (\varphi_1, \ldots, \varphi_{d_2})$, a collection of $Lip_{loc}^{\gamma-1}(\mathbb{R}\times [0,T] \times \mathbb{R}^n, \mathbb{R})$ vector fields,

$$\varphi(\phi, t, \psi) := \phi.\nu(t,\psi),$$

shows that ϕ is uniquely defined on $[0,T]$ under the weaker $(\gamma-1)$ regularity of ν. Similarly, one sees that $\tilde{\phi}$ and α can be expressed as rough integrals against $Lip_{loc}^{\gamma-1}$ 1-forms which then implies the claimed uniquess of the solution of the RDE (20). Finally, note that every additional degree of Lipschitz regularity allows for one further degree of differentiability of the solution flow with corresponding stability in rough path sense. ■

Remark 12 *The regularity assumptions of σ, ν, g with respect to t can be relaxed; see the forthcoming work by Caruana and Gurko.*

Remark 13 *One could perform transformations 1,2,3 in different order, leading at this point to different dependencies of the rough flows. A trivial consequence of theorem 14 is that the order of the transformations does not matter for the definition of the rough PDE and we found the current sequence to be the algebraically most tractable one.*

4 Semirelaxed limits and rough PDEs

The goal is to understand

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

in the case when (z^1, z^2, z^3) becomes a rough path. To this end we first need quantitative assumptions on L; we shall assume that

$$L(t,x,r,p,X) = -\mathrm{Tr}\left[A(t,x)^T X\right] + b(t,x) \cdot p + f(t,x,r) \tag{22}$$

where $A(t,x) = \sigma(t,x)\sigma^T(t,x)$ and[9] $\sigma : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n\times n'}$, $b(t,x) : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ are bounded, continuous in t and Lipschitz continuous in x, uniformly in $t \in [0,T]$. We also assume that $f : [0,T] \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ is continuous, bounded whenever r remains bounded, and with a lower Lipschitz bound, i.e. $\exists C < 0$ s.t.

$$f(t,x,r) - f(t,x,s) \geq C(r-s) \text{ for all } r \geq s,\ (t,x) \in [0,T] \times \mathbb{R}^n. \tag{23}$$

This will guarantee that comparison holds for $\partial_t + L$; see the appendix for some details. Let us now replace the (smooth) driving signals of the earlier sections by a $(d_1+d_2+d_3)$-dimensional driving signal z^ε and *impose* convergence to a genuine geometric p-rough path $\mathbf{z}$. That is, in the notation of [13, 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 shall write $\psi^{\mathbf{z}}, \phi^{\mathbf{z}}, \alpha^{\mathbf{z}}$ for the objects objects (solution of rough differential equations and integrals) built upon $\mathbf{z}$, as discussed in the last section (lemma 11). We also write $\psi^\varepsilon, \phi^\varepsilon, \alpha^\varepsilon$ when the driving signal is z^ε. Recall from (11) that $\Lambda = (\Lambda_1, \ldots, \Lambda_d)$ is a collection of linear, first order differential operators. We have

Theorem 14 *Let $p \geq 1$. Let L be of the above form (22) with assumptions on A, b, f as detailed above. Assume furthermore that the coefficients of $\Lambda = (\Lambda_1, \ldots, \Lambda_{d_1+d_2+d_3})$ have Lip$^\gamma$-regularity for $\gamma > p+2$. Let $u_0 \in BUC$ and $\mathbf{z}$ a geometric p-rough path. Then there exists a unique $u = u^{\mathbf{z}} \in BUC([0,T] \times \mathbb{R}^n)$, write formally*[10]*,*

$$du + L\left(t,x,u,Du,D^2u\right)dt = \Lambda(t,x,u,Du)\,d\mathbf{z},\ u(0,\cdot) \equiv u_0 \tag{24}$$

such that for any smooth sequence $z^\varepsilon \to \mathbf{z}$ in p-rough path sense, u^ε, viscosity solutions (always assumed BUC) of the classical PDE problems

$$\dot{u}^\varepsilon + L\left(t,x,u^\varepsilon,Du^\varepsilon,D^2u^\varepsilon\right) = \Lambda(t,x,u^\varepsilon,Du^\varepsilon)\,\dot{z}^\varepsilon,\ u(0,\cdot) \equiv u_0,$$

converge locally uniformly to $u^{\mathbf{z}}$. Moreover, we have the contraction property

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

(C given by 23) and continuity of the solution map $(\mathbf{z},u_0) \mapsto u^{\mathbf{z}}$ from

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

Proof. We use the same technique of "rough semi-relaxed limits" as in [4]: the key remark being that

$$\tilde{L}^\varepsilon := \left[{}^{\phi^\varepsilon}\left(L^{\psi^\varepsilon}\right)\right]_{\alpha^\varepsilon} \to \left[{}^{\phi^{\mathbf{z}}}\left(L^{\psi^{\mathbf{z}}}\right)\right]_{\alpha^{\mathbf{z}}}$$

holds uniformly on compact sets (as function of (t,x,r,p,X)). Secondly, the viscosity solutions u^ε can be transformed to

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

[9] this matrix-valued σ has to be distinguished from the $\mathbb{R}^n$-valued noise vectorfields $\sigma_1, \ldots, \sigma_{d_1}$.
[10] The intrinsic meaning of this "rough" PDE is discussed in definition 15 below.

$\tilde{v}^{\varepsilon}$ being a solution of the "random" PDE

$$\partial_t \tilde{v}^{\varepsilon} + \tilde{L}^{\varepsilon}\left(t, x, \tilde{v}^{\varepsilon}, D\tilde{v}^{\varepsilon}, D^2\tilde{v}^{\varepsilon}\right) = 0,\ \ \tilde{v}^{\varepsilon}(0,\cdot) = u_0(\cdot).$$

It is a simple matter, using comparison for instance, to see that for suitable C,

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

this in turn implies (thanks to the uniform control on $\alpha^{\varepsilon}, \phi^{\varepsilon}, \psi^{\varepsilon}$ as $\varepsilon \to 0$, using the rough path representations discussed in section 3.3) that $u^{\varepsilon}(t,x)$ remains locally uniform bounded (as $\varepsilon \to 0$). We can then define

$$\begin{aligned} \underline{\tilde{v}}(t,x) &:= \liminf{}_* \tilde{v}^{\varepsilon}(t,x), \\ \overline{\tilde{v}}(t,x) &:= \limsup{}_* \tilde{v}^{\varepsilon}(t,x). \end{aligned}$$

The stability result for viscosity solutions (cf. [6]) guarantees that $\underline{\tilde{v}}$, resp. $\overline{\tilde{v}}$, are super-, resp. sub-, solutions of

$$\partial_t + \tilde{L} = 0,\ \text{with}\ \tilde{L} := \left[{}^{\phi^{\mathbf{z}}}\left(L^{\psi^{\mathbf{z}}}\right)\right]_{\alpha^{\mathbf{z}}}. \tag{25}$$

Hence, comparison implies $\underline{\tilde{v}} \geq \overline{\tilde{v}}$ and since by definition $\underline{\tilde{v}} \leq \overline{\tilde{v}}$ we have shown that $\tilde{v}(t,x) := \underline{\tilde{v}}(t,x) = \overline{\tilde{v}}(t,x)$ is a solution of (25). Unwrapping the transformation, that is, setting

$$u^{\mathbf{z}}(t,x) := \phi^{\mathbf{z}}\left(t, \tilde{v}\left(t, \left(\psi^{\mathbf{z}}\right)^{-1}(t,x)\right) - \alpha^{\mathbf{z}}\left(t, \left(\psi^{\mathbf{z}}\right)^{-1}(t,x)\right)\right),$$

finishes the proof. ■

The reader may wonder if u is the solution in a sense beyond the "formal" equation

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

Inspired by the definition given by Lions–Souganidis in [25] we give

Definition 15 *u is a solution to the above **rough partial differential equation** if and only if*

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

in viscosity sense where

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

Corollary 16 *There exists a unique solution to the rough PDE.*

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 comparison. ■

5 Applications to stochastic partial differential equations

In this section we discuss concrete applications to SPDEs; they can be applied to the Zakai SPDE whenever appropiate, the general interest, however, goes beyond the filtering problem. The typical applications to SPDEs are path-by-path, i.e. by taking $\mathbf{z}$ to be a realization of a continuous semi-martingale Y and its stochastic area, say $\mathbf{Y}(\omega) = (Y, A)$; the most prominent example being Brownian motion and Lévy's area. The continuity property of our main theorem allows to identify 24 with $\mathbf{z} = \mathbf{Y}(\omega)$ as *Stratonovich solution* to the SPDE

$$du + L\left(t, x, u, Du, D^2u\right) dt = \Lambda(t, x, u, Du) \circ dY, \;\; u(0, \cdot) = u_0.$$

Indeed, under the stated assumptions the Wong-Zakai approximations, in which Y is replaced by its piecewise linear approximation, based on some mesh $\left\{0, \frac{T}{n}, \frac{2T}{n} \dots, T\right\}$, the approximate solution will converge (locally uniformly on $[0, T] \times \mathbb{R}^n$ and in probability, say) to the solution of

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

as constructed in theorem 14. In view of well-known Wong-Zakai approximation results for SPDEs, ranging from [2, 31] to [20, 21], the rough PDE solution is then identified as Stratonovich solution. (At least for L uniformly elliptic: the (Stratonovich) integral interpretations can break down in degenerate situations; as example, consider non-differentiable initial data u_0 and the (one-dimensional) random transport equation $du = u_x \circ dB$ with explicit "Stratonovich" solution $u_0(x + B_t)$. A similar situation occurs for the classical transport equation $\dot{u} = u_x$, of course.)

Remark 17 (Itô versus Stratonovich) *Note that similar* ***SPDEs in Itô-form*** *need not be, in general, well-posed. Consider the following (well-known) linear example*

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

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 [29], this has led to coercivity conditions, also known as super-parabolicity assumptions, in order to guarantee well-posedness.*

Remark 18 (Regularity of noise coefficients) *Applied in the semimartingale context (finite p-variation for any* $p > 2$*) the regularity assumption of theorem 14 reads* $Lip^{4+\varepsilon}$, $\varepsilon > 0$. *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. 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 19 (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* $\partial_t + \left[{}^{\phi^{\mathbf{z}}}\left(L^{\psi^{\mathbf{z}}}\right)\right]_{\alpha^{\mathbf{z}}}$. *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 has* $C^{1/p, 2+\delta}$ *for suitable* $\delta > 0$. *Indeed,* *it is sheer matter of conditions-book-keeping to solve* $\partial_t + \left[{}^{\phi^{\mathbf{z}}}\left(L^{\psi^{\mathbf{z}}}\right)\right]_{\alpha^{\mathbf{z}}}$ *under (known and sharp) conditions in Hölder spaces, cf. [23, 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 routine theorems under obvious conditions; the reader with the slightest familiarity with rough path theory will realize that formulating such statements is a mechanical exercise.

(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. [13, Chapter 13] and the references therein) examples include piecewise linear -, mollifier - and Karhunen-Loeve approximations, as well as (weak) Donsker type random walk approximations [1]. 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. [24] in p-variation rough path topology or [8] 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. Let us note that, to the best of our knowledge, results of this type are new for such non-linear SPDEs. In the linear case, approximations and support of SPDEs have been studied in great detail [19, 18, 16, 15, 17].

(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 deiviation behaviour of $\left(\varepsilon B, \varepsilon^2 A\right)$ in rough path topology (e.g. [24] in p-variation and [11] in Hölder rough path topology) it suffices to recall that large deviation principles are stable under continuous maps. Again, large deviation estimates for non-linear SPDEs in the small noise limit appear to be new and may be hard to obtain without rough paths theory.

(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. [5] or [10], 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 [12]. 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 [14], 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

$$du + Ldt = \sum_k \Lambda_k \circ dY^k.$$

Equivalently, we write

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

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

$$\log \tilde{\mathbf{Y}} := \log \mathbf{Y} + (0, \Gamma_t)$$

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

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

is identified with

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

The practical relevance is that one can construct approximations (Y^n) to Y, each Y^n only dependent on finitely many points, which converge uniformly to Y with the "wrong" area (cf. [9]); that is,

$$\left(Y^n, \int Y^n dY^n\right) \to \tilde{\mathbf{Y}}$$

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

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

The formal proof is easy; it suffices to analyze the equations (rough) differential equations for (ψ, ϕ, α) in presence of area perturbation; see [9], and then identify the corresponding operators $\left[{}^{\phi}\left[L^{\psi}\right]\right]_{\alpha}$. Actually, one can pick any multiindex $\gamma = (\gamma_1, \dots, \gamma_N) \in \{1,\dots,d\}^N$ and find (uniform) approximations such as to make the higher iterated Lie brackets $\Lambda_\gamma = \left[\Lambda_{\gamma_1}, \dots, \left[\Lambda_{\gamma_{N-1}}, \Lambda_{\gamma_N}\right] \dots\right]$, 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 [13, Chapter 14]). In other words, $u^{\varepsilon} \to u$ in probability and locally uniformly where

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

6 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 20 *Assume condition (3.14) of the User's Guide [6], 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*

$$F(t,x,r,p,X) - F(t,x,s,p,X) \ge C(r-s) \quad \forall r \le s, \tag{26}$$

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 (26) holds with $C = 0$ which is tantamount to require that F is non-decreasing in r. The standard arguments (e.g. [6] or appendix of [4]) then apply with minimal adaptations. ■

Corollary 21 *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 (26)*.*

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. ■

References

[1] Emmanuel Breuillard, Peter Friz, and Martin Huesmann. From random walks to rough paths. *Proc. AMS*, pages 3487–3496, 2009.

[2] Zdzisław Brzeźniak and Franco Flandoli. Almost sure approximation of Wong-Zakai type for stochastic partial differential equations. *Stochastic Process. Appl.*, 55(2):329–358, 1995.

[3] Michael Caruana and Peter Friz. Partial differential equations driven by rough paths. *Journal of Differential Equations*, 247(1):140 – 173, 2009.

[4] Michael Caruana, Peter Friz, and Harald Oberhauser. A (rough) pathwise approach to fully non-linear stochastic partial differential equations. *Annales de l'IHP (C), Non Linear Analysis (accepted 2010)*, 2009.

[5] Laure Coutin and Zhongmin Qian. Stochastic analysis, rough path analysis and fractional Brownian motions. *Probab. Theory Related Fields*, 122(1):108–140, 2002.

[6] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.

[7] Wendell H. Fleming and H. Mete Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stochastic Modelling and Applied Probability*. Springer, New York, second edition, 2006.

[8] P. Friz, T. Lyons, and D. Stroock. Lévy's area under conditioning. *Ann. Inst. H. Poincaré Probab. Statist.*, 42(1):89–101, 2006.

[9] Peter Friz and Harald Oberhauser. Rough path limits of the Wong-Zakai type with a modified drift term. *J. Funct. Anal.*, 256:3236–3256, 2009.

[10] Peter Friz and Nicolas Victoir. Differential equations driven by Gaussian signals. *Ann. de l'IHP.* accepted (2009).

[11] Peter Friz and Nicolas Victoir. Approximations of the Brownian rough path with applications to stochastic analysis. *Ann. Inst. H. Poincaré Probab. Statist.*, 41(4):703–724, 2005.

[12] Peter Friz and Nicolas Victoir. On uniformly subelliptic operators and stochastic area. *Probab. Theory Related Fields*, 142(3-4):475–523, 2008.

[13] Peter K. Friz and Nicolas B. Victoir. *Multidimensional stochastic processes as rough paths: theory and applications.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2010.

[14] M. Gubinelli and S. Tindel. Rough evolution equations. *Ann. Prob. 38, pp. 1-75*, 2010.

[15] I. Gyöngy. The stability of stochastic partial differential equations and applications. I. *Stochastics Stochastics Rep.*, 27(2):129–150, 1989.

[16] I. Gyöngy. 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.*, pages 91–118. Springer, Berlin, 1989.

[17] I. Gyöngy. The stability of stochastic partial differential equations. II. *Stochastics Stochastics Rep.*, 27(3):189–233, 1989.

[18] I. Gyöngy. The approximation of stochastic partial differential equations and applications in nonlinear filtering. *Comput. Math. Appl.*, 19(1):47–63, 1990.

[19] István Gyöngy. 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.*, pages 116–136. Springer, Berlin, 1991.

[20] István Gyöngy and György Michaletzky. On Wong-Zakai approximations with δ-martingales. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 460(2041):309–324, 2004. Stochastic analysis with applications to mathematical finance.

[21] István Gyöngy and Anton Shmatkov. Rate of convergence of Wong-Zakai approximations for stochastic partial differential equations. *Appl. Math. Optim.*, 54(3):315–341, 2006.

[22] 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.

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

[24] 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.

[25] 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.

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

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

[28] 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.

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

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

[31] Krystyna Twardowska. An approximation theorem of Wong-Zakai type for nonlinear stochastic partial differential equations. *Stochastic Anal. Appl.*, 13(5):601–626, 1995.