Available online at www.sciencedirect.com

ScienceDirect

Ann. I. H. Poincaré – AN 28 (2011) 27–46

ANNALES DE L'INSTITUT HENRI POINCARÉ ANALYSE NON LINÉAIRE

www.elsevier.com/locate/anihpc

A (rough) pathwise approach to a class of non-linear stochastic partial differential equations

Michael Caruana [a], Peter K. Friz [b,c,*], Harald Oberhauser [b]

[a] *King Fahd University of Petroleum and Minerals, Saudi Arabia*
[b] *TU Berlin, Germany*
[c] *WIAS Berlin, Germany*

Received 23 September 2009; received in revised form 4 November 2010; accepted 4 November 2010

Available online 9 November 2010

Abstract

We consider non-linear parabolic evolution equations of the form $\partial_t u = F(t, x, Du, D^2u)$, subject to noise of the form $H(x, Du) \circ dB$ where H is linear in Du and $\circ dB$ denotes the Stratonovich differential of a multi-dimensional Brownian motion. Motivated by the essentially pathwise results of [P.-L. Lions, P.E. Souganidis, Fully nonlinear stochastic partial differential equations, C. R. Acad. Sci. Paris Sér. I Math. 326 (9) (1998) 1085–1092] we propose the use of rough path analysis [T.J. Lyons, Differential equations driven by rough signals, Rev. Mat. Iberoamericana 14 (2) (1998) 215–310] in this context. Although the core arguments are entirely deterministic, a continuity theorem allows for various probabilistic applications (limit theorems, support, large deviations, ...).

Keywords: Parabolic viscosity PDEs; Stochastic PDEs; Rough path theory

1. Introduction

Let us recall some basic ideas of (second order) viscosity theory [13,15] and rough path theory [41,42]. As for viscosity theory, consider a real-valued function $u = u(x)$ with $x \in \mathbb{R}^n$ and assume $u \in C^2$ is a classical super-solution,

$$-G\big(x, u, Du, D^2u\big) \geqslant 0,$$

where G is a (continuous) function, *degenerate elliptic* in the sense that $G(x, u, p, A) \leqslant G(x, u, p, A + B)$ whenever $B \geqslant 0$ in the sense of symmetric matrices. The idea is to consider a (smooth) test function φ which touches u from below at some point $\bar{x}$. Basic calculus implies that $Du(\bar{x}) = D\varphi(\bar{x})$, $D^2u(\bar{x}) \geqslant D^2\varphi(\bar{x})$ and, from degenerate ellipticity,

$$-G\big(\bar{x}, \varphi, D\varphi, D^2\varphi\big) \geqslant 0. \tag{1.1}$$

* Corresponding author at: TU and WIAS Berlin, Germany.
E-mail addresses: friz@math.tu-berlin.de, friz@wias-berlin.de (P.K. Friz).

0294-1449/$ – see front matter
doi:10.1016/j.anihpc.2010.11.002

This suggests to define a *viscosity super-solution* (at the point $\bar{x}$) to $-G = 0$ as a (lower semi-) continuous function u with the property that (1.1) holds for any test function which touches u from below at $\bar{x}$. Similarly, *viscosity sub-solutions* are (upper semi-) continuous functions defined via test functions touching u from above and by reversing inequality in (1.1); *viscosity solutions* are both super- and sub-solutions. Observe that this definition covers (completely degenerate) first order equations as well as parabolic equations, e.g. by considering $\partial_t - F = 0$ on $\mathbb{R}^+ \times \mathbb{R}^n$ where F is degenerate elliptic. The resulting theory (existence, uniqueness, stability, ...) is without doubt one of most important recent developments in the field of partial differential equations. As a typical result,[1] one has existence and uniqueness result in the class of bounded solutions to the initial value problem $(\partial_t - F)u = 0, u(0, \cdot) = u_0 \in \mathrm{BUC}(\mathbb{R}^n)$, provided $F = F(t, x, Du, D^2u)$ is continuous, degenerate elliptic and satisfies a (well-known) technical condition (see Condition 1 below). In fact, uniqueness follows from a stronger property known as *comparison*: assume u (resp. v) is a sub-solution (resp. super-solution) and $u_0 \leqslant v_0$; then $u \leqslant v$ 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 a priori information on the u^n (e.g. locally uniform convergence, or locally uniform boundedness[2]). 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 $(V_1, \dots, V_d)$ 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

$$\dot{y}(t) = \sum_{i=1}^{d} V_i(y)\dot{z}^i(t), \quad y(0) = y_0 \in \mathbb{R}^n. \tag{1.2}$$

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

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

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

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

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

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 (1.2): 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 (1.2), in a unique and continuous fashion, to driving signals which live in the abstract

[1] BUC(...) denotes the space of bounded, uniformly continuous functions; BC(...) denotes the space of bounded, continuous functions.

[2] What we have in mind here is the *Barles–Perthame method of semi-relaxed limits*. We shall use this method in the proof of Theorem 1 and postpone precise references until then.

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[3] is precisely this abstract completion. In fact, this space can be realized as genuine path space,

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

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

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

is the natural state 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

$$B^{(2)} := \left(\int_0^{\cdot} B^i \circ dB^j\right)_{i,j\in\{1,\dots,d\}} \tag{1.3}$$

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 $B^{(2)}=\frac{1}{2}B\otimes B+A$ where $A:=\operatorname{Anti}(B^{(2)})$ 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.*

Turning to the main topic of this paper, we follow [35,36,38] in considering a real-valued function of time and space $u=u(t,x)\in \mathrm{BC}([0,T)\times\mathbb{R}^n)$ which solves the non-linear partial differential equation

$$\begin{aligned} du &= F\big(t,x,Du,D^2u\big)\,dt+\sum_{i=1}^d H_i(x,Du)\,dz^i \\ &\equiv F\big(t,x,Du,D^2u\big)\,dt+H(x,Du)\,dz \end{aligned} \tag{1.4}$$

in viscosity sense. When $z:[0,T]\to\mathbb{R}^d$ is C^1 then, subject to suitable conditions on F and H, this falls in the standard setting of viscosity theory as discussed above. This can be pushed further to $z\in W^{1,1}$ (see e.g. [35, Remark 4] and the references given there) but the case when $z=z(t)$ has only "Brownian" regularity (just below $1/2$-Hölder, say) falls dramatically outside the scope of the standard theory. The reader can find a variety of examples (drawing from fields as diverse as stochastic control theory, pathwise stochastic control, interest rate theory, front propagation and phase transition in random media, ...) in the articles [36,34] justifying the need of a theory of (non-linear) *stochastic partial differential equations* (SPDEs) in which z in (1.4) is taken as a Brownian motion.[4] In the same series of articles a satisfactory theory is established for the case of non-linear Hamiltonian with no spatial dependence, i.e. $H=H(Du)$. The contribution of this article is to deal with non-linear F and $H=H(x,Du)$, linear in Du, although we suspect that the marriage of rough path and viscosity methodology will also prove useful in further investigations on *fully non-linear* (i.e. both F and H) stochastic partial differential equations.[5] To fix ideas, we give the following example, suggested in [36] and carefully worked out in [8,9].

Example 2 *(Pathwise stochastic control).* Consider

$$dX=b(X;\alpha)\,dt+W(X;\alpha)\circ d\tilde{B}+V(X)\circ dB,$$

where b,W,V are (collections of) sufficiently nice vector fields (with b,W dependent on a suitable control $\alpha=\alpha(t)\in\mathcal{A}$, applied at time t) and $\tilde{B},B$ are multi-dimensional (independent) Brownian motions. Define[6]

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

[4] ... in which case (1.4) is understood in Stratonovich form.

[5] The use of rough path analysis in the context of non-linear SPDEs was verbally conjectured by P.L. Lions in his 2003 Courant lecture.

[6] Remark that any optimal control $\alpha(\cdot)$ here will depend on knowledge of the entire path of B. Such anticipative control problems and their link to classical stochastic control problems were discussed early on by Davis and Burnstein [14].

$$v(x,t;B)=\inf_{\alpha\in\mathcal{A}}\mathbb{E}\left[\left(g\left(X_T^{x,t}\right)+\int_t^T f\left(X_s^{x,t},\alpha_s\right)ds\right)\,\Big|\,B\right]$$

where $X^{x,t}$ denotes the solution process to the above SDE started at $X(t)=x$. Then, at least by a formal computation,

$$dv+\inf_{\alpha\in\mathcal{A}}\left[b(x,\alpha)Dv+L_a v+f(x,\alpha)\right]dt+Dv\cdot V(x)\circ dB=0$$

with terminal data $v(\cdot,T)\equiv g$, and $L_\alpha=\sum W_i^2$ in Hörmander form. Setting $u(x,t)=v(x,T-t)$ turns this into the initial value (Cauchy) problem,

$$du=\inf_{\alpha\in\mathcal{A}}\left[b(x,\alpha)Du+L_a u+f(x,\alpha)\right]dt+Du\cdot V(x)\circ dB_{T-\cdot}$$

with initial data $u(\cdot,0)\equiv g$; and hence of a form which is covered by Theorem 1 below. Indeed, $H=(H_1,H_2)$, $H_i(x,p)=p\cdot V_i(x)$, is linear in p. (Moreover, the rough driving signal in Theorem 1 is taken as $\mathbf{z}_t:=\mathbf{B}_{T-t}(\omega)$ where $\mathbf{B}(\omega)$ is a fixed Brownian rough path, run backwards in time.[7])

Returning to the general setup of (1.4), the results [35,36,38] are in fact *pathwise* and apply to any continuous path $z\in C([0,T],\mathbb{R}^d)$, this includes Brownian and even rougher sources of noise; however, the assumption was made that $H=H(Du)$ is independent of x. The rôle of x-dependence is an important one (as it arises in applications such as Example 2): the results of Lions–Souganidis imply that the map

$$z\in C^1\left([0,T],\mathbb{R}^d\right)\mapsto u(\cdot,\cdot)\in C\left([0,T],\mathbb{R}^n\right)$$

depends continuously on z *in uniform topology*; thereby giving existence/uniqueness results to

$$du=F\left(t,x,Du,D^2u\right)dt+\sum_{i=1}^d H_i(Du)\,dz^i$$

for *every* continuous path $z:[0,T]\to\mathbb{R}^d$. When the Hamiltonian depends on x, this ceases to be true; indeed, take $F\equiv 0$, $d=2$ and $H_i(x,p)=pV_i(x)$ where V_1,V_2 are the vector fields from Example 1. Solving the characteristic equations shows that u is expressed in terms of the (inverse) flow associated to $dy=V_1(y)\,dz^1+V_2(y)\,dz^2$, and we have already seen that the solution of this ODE does not depend continuously on $z=(z^1,z^2)$ in uniform topology.[8]

Of course, this type of problem can be prevented by strengthening the topology: the Lyons' theory of rough paths does exhibit an entire cascade of (p-variation or $1/p$-Hölder type rough path) metrics (for each $p\geqslant 1$) on path-space under which such ODE solutions are continuous functions of their driving signal. This suggests to extend the Lions–Souganidis theory from a pathwise to a *rough* pathwise theory. We shall do so for a rich class of fully-non-linear F and Hamiltonians $H(x,Du)$ linear in Du. This last assumption allows for a global change of coordinates which mimicks a classical trick in SPDE analysis (which, to the best of our knowledge, goes back to Tubaro [47], Kunita [32, Chapter 6] and Rozovskiĭ [46], see also Iftimie and Varsan [30]; similar techniques have also proven useful when $H=H(x,u)$ – we shall comment on this in Section 8) where a SPDE is transformed into a random PDE (i.e. one that can be solved with deterministic methods by fixing the randomness). In doing so, the interplay between rough path and viscosity methods is illustrated in a transparent way and everything boils down to combine the stability properties of viscosity solution with those of differential equations in the rough path sense. We have the following result.[9]

Theorem 1. *Let $p\geqslant 1$ and $(z^\varepsilon)\subset C^\infty([0,T],\mathbb{R}^d)$ be Cauchy in (p-variation) rough path topology with rough path limit $\mathbf{z}\in C^{0,p\text{-var}}([0,T],G^{[p]}(\mathbb{R}^d))$. Assume*

$$u_0^\varepsilon\in \mathrm{BUC}\left(\mathbb{R}^n\right)\to u_0\in\mathrm{BUC}\left(\mathbb{R}^n\right),$$

7 Alternatively, the proof of Theorem 1 is trivially modified to directly accommodate terminal data problems.

8 We shall push this remark much further in Theorem 2 below.

9 Unless otherwise stated we shall always equip the spaces BC and BUC with the topology of locally uniform convergence.

locally uniformly as $\varepsilon \to 0$. *Let* $F = F(t, x, p, X)$ *be continuous, degenerate elliptic, and assume that* $\partial_t - F$ *satisfies* $\Phi^{(3)}$*-invariant comparison* (*cf. Definition* 1 *below*). *Assume that* $V = (V_1, \ldots, V_d)$ *is a collection of* $\mathrm{Lip}^{\gamma+2}(\mathbb{R}^n; \mathbb{R}^n)$ *vector fields with* $\gamma > p$. *Assume existence of* (*necessarily unique*[10]) *viscosity solutions* $u^\varepsilon \in \mathrm{BC}([0, T) \times \mathbb{R}^n)$ *to*

$$du^\varepsilon = F\left(t, x, Du^\varepsilon, D^2u^\varepsilon\right) dt - Du^\varepsilon(t, x) \cdot V(x)\, dz^\varepsilon(t), \tag{1.5}$$

$$u^\varepsilon(0, \cdot) = u_0^\varepsilon \tag{1.6}$$

and assume that the resulting family $(u^\varepsilon\colon \varepsilon > 0)$ *is uniformly bounded.*[11] *Then*

(i) *there exists a unique* $u \in \mathrm{BC}([0, T) \times \mathbb{R}^n)$, *only dependent on* $\mathbf{z}$ *and* u_0 *but not on the particular approximating sequences, such that* $u^\varepsilon \to u$ *locally uniformly. We write* (*formally*)

$$du = F\left(t, x, Du, D^2u\right) dt - Du(t, x) \cdot V(x)\, d\mathbf{z}(t), \tag{1.7}$$

$$u(0, \cdot) = u_0, \tag{1.8}$$

and also $u = u^{\mathbf{z}}$ *when we want to indicate the dependence on* $\mathbf{z}$;

(ii) *we have the contraction property*

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

where $\hat{u}^{\mathbf{z}}$ *is defined as limit of* $\hat{u}^\eta$, *defined as in* (1.5) *with* u^ε *replaced by* $\hat{u}^\eta$ *throughout*;

(iii) *the solution map* $(\mathbf{z}, u_0) \mapsto u^{\mathbf{z}}$ *from*

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

is continuous.

Our proof actually allows for BC initial data in the above theorem, since existence of solutions to the approximate problems (1.5) is assumed. Our preference for BUC initial data comes from the fact that existence results are typically established under some assumption of uniform continuity (e.g. [12, Theorem 3.1]). Conversely, under a mild sharpening of the structural assumption which is still satisfied in all our examples the solutions constructed in the above theorem can be seen to be BUC ("bounded uniformly continuous") in time-space. When $F = F(Du, D^2u)$, as in the setting of [35,38], a spatial modulus is easy to obtain; in the above generality the comparison proof (based on doubling of the spatial variable) can be adapted to obtain a spatial modulus of continuity, uniform in time (this is implemented in [22] for instance). Curiously, a modulus in time cannot be established so directly; it is known however that a "modulus of continuity in space" implies "modulus of continuity in time" (cf. Lemma 9.1 in [3]). We shall return to such regularity questions in detail in a separate note.

The reader may wonder if $u \in \mathrm{BC}([0, T) \times \mathbb{R}^n)$ constructed in the above theorem solves a well-defined "rough" PDE, apart from the formal equation (1.7). The answer is, in essence, that u is also a solution in the sense of Lions and Souganidis [35,36,38] provided their definition is translated, mutatis mutandis, to the present rough PDE setting. While we suspect that such a point of view will be the key to a (rough) pathwise understanding of fully non-linear stochastic partial differential equations, the present situation (H linear in Du) allows for a simpler understanding, still in the spirit of Lions–Souganidis (to be specific, see [35, Theorem 2.4]). The details of this are best given after the proof of Theorem 1; we thus postpone further discussion on this to Section 7.

2. Condition for comparison

We shall always assume that $F = F(t, x, p, X)$ is continuous and degenerate elliptic. A sufficient condition[12] for comparison of (bounded) solutions to $\partial_t = F$ on $[0, T) \times \mathbb{R}^n$ is given by

---

[10] This follows from the first 5 lines in the proof of this theorem.

[11] A simple sufficient conditions is boundedness of $F(\cdot, \cdot, 0, 0)$ on $[0, T] \times \mathbb{R}^n$, and the assumption that $u_0^\varepsilon \to u_0$ uniformly, as can be seen by comparison with function of the type $(t, x) \mapsto \pm C(t + 1)$, with sufficiently large C.

[12] ... which actually implies degenerate ellipticity, cf. p. 18 in [13, (3.14)].

Condition 1. (See [13, (3.14)].) There exists a function $\theta : [0,\infty] \to [0,\infty]$ with $\theta(0+) = 0$, such that for each fixed $t \in [0,T]$,

$$F\big(t, x, \alpha(x-\tilde{x}), X\big) - F\big(t, \tilde{x}, \alpha(x-\tilde{x}), Y\big) \leqslant \theta\big(\alpha|x-\tilde{x}|^2 + |x-\tilde{x}|\big) \tag{2.1}$$

for all $\alpha > 0$, $x, \tilde{x} \in \mathbb{R}^n$ and $X, Y \in S^n$ (the space of $n \times n$ symmetric matrices) satisfy

$$-3\alpha \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \leqslant \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leqslant 3\alpha \begin{pmatrix} I & -I \\ -I & I \end{pmatrix}. \tag{2.2}$$

Furthermore, we require $F = F(t, x, p, X)$ to be uniformly continuous whenever p, X remain bounded.

Although this seems part of the folklore in viscosity theory[13] only the case when $\mathbb{R}^n$ is replaced by a bounded domain is discussed in the standard Refs. ([13, (3.14) and Section 8] or [15, Sections V.7, V.8]; in this case the very last requirement on uniform continuity can be omitted). For this reason and the reader's convenience we have included a full proof of parabolic comparison on $[0,T) \times \mathbb{R}^n$ under the above condition in Appendix A.

Remark 1 *(Stability under sup, inf etc.).* Using elementary inequalities of the type

$$\big|\sup(a) - \sup(b)\big| \leqslant \sup|a-b| \quad \text{for } a, b \in \mathbb{R},$$

one immediately sees that if $F_\gamma, F_{\gamma,\beta}$ satisfy (2.1) for γ, β in some index set with a common modulus θ, then $\inf_\gamma F_\gamma$, $\sup_\beta \inf_\gamma F_{\beta,\gamma}$ etc. again satisfy (2.1). Similar remarks apply to the uniform continuity property; provided there exists, for any $R < \infty$, a common modulus of continuity σ_R, valid whenever p, X are of norm less than R.

3. Invariant comparison

To motivate our key assumption on F we need some preliminary remarks on the transformation behaviour of

$$Du = (\partial_1 u, \ldots, \partial_n u), \qquad D^2 u = (\partial_{ij} u)_{i,j=1,\ldots,n}$$

under change of coordinates on $\mathbb{R}^n$ where $u = u(t,\cdot)$, for fixed t. Let us allow the change of coordinates to depend on t, say $v(t,\cdot) := u(t, \phi_t(\cdot))$ where $\phi_t : \mathbb{R}^n \to \mathbb{R}^n$ is a diffeomorphism. Differentiating $v(t, \phi_t^{-1}(\cdot)) = u(t,\cdot)$ twice, followed by evaluation at $\phi_t(y)$, we have, with summation over repeated indices,

$$\partial_i u\big(t, \phi_t(x)\big) = \partial_k v(t,x) \partial_i \phi_t^{-1;k}\big|_{\phi_t(x)},$$
$$\partial_{ij} u\big(t, \phi_t(x)\big) = \partial_{kl} v(t,x) \partial_i \phi_t^{-1;k}\big|_{\phi_t(x)} \partial_j \phi_t^{-1;l}\big|_{\phi_t(x)} + \partial_k v(t,x) \partial_{ij} \phi_t^{-1;k}\big|_{\phi_t(x)}.$$

We shall write this, somewhat imprecisely[14] but convenient, as

$$Du|_{\phi_t(x)} = \big\langle Dv|_x, D\phi_t^{-1}\big|_{\phi_t(x)}\big\rangle,$$
$$D^2 u\big|_{\phi_t(x)} = \big\langle D^2 v\big|_x, D\phi_t^{-1}\big|_{\phi_t(x)} \otimes D\phi_t^{-1}\big|_{\phi_t(x)}\big\rangle + \big\langle Dv|_x, D^2 \phi_t^{-1}\big|_{\phi_t(x)}\big\rangle. \tag{3.1}$$

Let us now introduce $\Phi^{(k)}$ as the class of all flows of C^k-diffeomorphisms of $\mathbb{R}^n$, $\phi = (\phi_t\colon\ t \in [0,T])$, such that $\phi_0 = \mathrm{Id}$, $\forall \phi \in \Phi^{(k)}$ and such that ϕ_t and ϕ_t^{-1} have k bounded derivatives, uniformly in $t \in [0,T]$. We say that $\phi(n) \to \phi$ in $\Phi^{(k)}$ iff for all multi-indices α with $|\alpha| \leqslant k$

$$\partial_\alpha \phi(n) \to \partial_\alpha \phi_t, \qquad \partial_\alpha \phi(n)^{-1} \to \partial_\alpha \phi_t^{-1} \quad \text{locally uniformly in } [0,T] \times \mathbb{R}^n.$$

Definition 1 *($\Phi^{(k)}$-invariant comparison; F^ϕ).* Let $k \geqslant 2$ and define $F^\phi((t, x, p, X))$ as

$$F\big(t, \phi_t(x), \big\langle p, D\phi_t^{-1}\big|_{\phi_t(x)}\big\rangle, \big\langle X, D\phi_t^{-1}\big|_{\phi_t(x)} \otimes D\phi_t^{-1}\big|_{\phi_t(x)}\big\rangle + \big\langle p, D^2\phi_t^{-1}\big|_{\phi_t(x)}\big\rangle\big). \tag{3.2}$$

---

[13] E.g. in Section 4.4. of Barles' 1997 lecture notes, http://www.phys.univ-tours.fr/~barles/Toulcours.pdf, or Section V.9 in [15].

[14] Strictly speaking, one should view $(Du, D^2u)|.$ as *second order* cotangent vector, the pull-back of $(Dv, D^2v)|_x$ under ϕ_t^{-1}.

We say that $\partial_t = F$ satisfies $\Phi^{(k)}$-invariant comparison if, for every $\phi \in \Phi^{(k)}$, comparison holds for bounded solutions of $\partial_t - F^\phi = 0$. More precisely, if u is a bounded upper semi-continuous sub- and v a bounded lower semi-continuous super-solution to this equation and $u(0,\cdot) \leqslant v(0,\cdot)$ then $u \leqslant v$ on $[0,T) \times \mathbb{R}^n$.

4. Examples

Example 3 *(F linear).* Suppose that $\sigma(t,x) : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n \times n'}$ and $b(t,x) : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ are bounded, continuous in t and Lipschitz continuous in x, uniformly in $t \in [0,T]$. If $F(t,x,p,X) = \mathrm{Tr}[\sigma(t,x)\sigma(t,x)^T X] + b(t,x) \cdot p$, then $\Phi^{(3)}$-invariant comparison holds. Although this is a special case of the following example, let us point out that F^ϕ is of the same form as F with σ, b replaced by

$$\sigma^\phi(t,x)^k_m = \sigma^i_m\big(t,\phi_t(x)\big)\partial_i \phi_t^{-1;k}\big|_{\phi_t(x)}, \quad k = 1,\dots,n;\ m = 1,\dots,n',$$
$$b^\phi(t,x)^k = \big[b^i\big(t,\phi_t(x)\big)\partial_i \phi_t^{-1;k}\big|_{\phi_t(x)}\big] + \sum_{i,j}\big(\sigma^i_m \sigma^j_m \partial_{ij}\phi_t^{-1;k}\big|_{\phi_t(y)}\big), \quad k = 1,\dots,n.$$

By defining properties of flows of diffeomorphisms, $t \mapsto \partial_i \phi_t^{-1;k}|_{\phi_t(x)}, \partial_{ij}\phi_t^{-1;k}|_{\phi_t(y)}$ is continuous and the C^3-boundedness assumption inherent in our definition of $\Phi^{(3)}$ ensures that σ^ϕ, b^ϕ are Lipschitz in x, uniformly in $t \in [0,T]$. It is then easy to see (cf. the argument of [15, Lemma 7.1]) that F^ϕ satisfies Condition 1 for every $\phi \in \Phi^{(3)}$. This implies that $\Phi^{(3)}$-invariant comparison holds for bounded solutions of $\partial_t - F^\phi = 0$.

Example 4 *(F quasi-linear).* Let

$$F(t,x,p,X) = \mathrm{Tr}\big[\sigma(t,x,p)\sigma(t,x,p)^T X\big] + b(t,x,p). \tag{4.1}$$

We assume $b = b(t,x,p) : [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is continuous, bounded and Lipschitz continuous in x and p, uniformly in $t \in [0,T]$. We also assume that $\sigma = \sigma(t,x,p) : [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n \times n'}$ is a continuous, bounded map such that

- $\sigma(t,\cdot,p)$ is Lipschitz continuous, uniformly in $(t,p) \in [0,T] \times \mathbb{R}^n$;
- there exists a constant $c_1 > 0$, such that[15]

$$\forall p,q \in \mathbb{R}^n: \quad \big|\sigma(t,x,p) - \sigma(t,x,q)\big| \leqslant c_1 \frac{|p-q|}{1+|p|+|q|} \tag{4.2}$$

for all $t \in [0,T]$ and $x \in \mathbb{R}^n$.

We show that F^ϕ satisfies Condition 1 for every $\phi \in \Phi^{(3)}$; this implies that $\Phi^{(3)}$-invariant comparison holds for $\partial_t = F$ with F given by (4.1). To see this we proceed as follows. For brevity denote

$$p = \alpha(x - \tilde{x}), \qquad J_\cdot = D\phi_t^{-1}\big|_{\phi_t(\cdot)}, \qquad H_\cdot = D^2\phi_t^{-1}\big|_{\phi_t(\cdot)},$$
$$\sigma_\cdot = \sigma\big(t,\phi_t(\cdot),\langle p, J_\cdot\rangle\big), \qquad a_\cdot = \sigma_\cdot \sigma_\cdot^T, \qquad b_\cdot = b\big(t,\phi_t(\cdot),\langle p, J_\cdot\rangle\big)$$

so that

$$\begin{aligned} F^\phi(t,x,p,X) &= \mathrm{Tr}\big[a_x\big(\langle X, J_x \otimes J_x\rangle + \langle p, H_x\rangle\big)\big] + b_x \\ &= \mathrm{Tr}\big[J_x a_x J_x^T X\big] + b_x + \mathrm{Tr}\big[a_x \langle p, H_x\rangle\big]. \end{aligned}$$

Hence

$$F^\phi(t,\tilde{x},p,Y) - F^\phi(t,x,p,X) = \underbrace{\mathrm{Tr}\big[J_{\tilde{x}} a_{\tilde{x}} J_{\tilde{x}}^T Y - J_x a_x J_x^T X\big]}_{=:\text{(i)}} + \underbrace{b_{\tilde{x}} - b_x}_{=:\text{(ii)}} + \underbrace{\mathrm{Tr}\big[a_{\tilde{x}}\langle p, H_{\tilde{x}}\rangle - a_x\langle p, H_x\rangle\big]}_{=:\text{(iii)}}.$$

[15] A condition of this type also appears also in [2].

To estimate (i) note that $J_x a_x J_x^T = J_x\sigma_x(J_x\sigma_x)^T$. The $\mathbb{R}^{2n}\times\mathbb{R}^{2n}$ matrix

$$\begin{pmatrix} (J_x\sigma_x)(J_x\sigma_x)^T & J_x\sigma_x(J_{\tilde{x}}\sigma_{\tilde{x}})^T \\ (J_{\tilde{x}}\sigma_{\tilde{x}})(J_x\sigma_x)^T & J_{\tilde{x}}\sigma_{\tilde{x}}(J_{\tilde{x}}\sigma_{\tilde{x}})^T \end{pmatrix}$$

is positive semi-definite and thus we can multiply it to both sides of the inequality

$$\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leqslant 3\alpha \begin{pmatrix} I & -I \\ -I & I \end{pmatrix}.$$

The resulting inequality is stable under evaluating the trace and so one gets

$$\begin{aligned} &\mathrm{Tr}\big[J_x\sigma_x(J_x\sigma_x)^T\cdot X - J_{\tilde{x}}\sigma_{\tilde{x}}(J_{\tilde{x}}\sigma_{\tilde{x}})^T\cdot Y\big] \\ &\quad\leqslant 3\alpha\,\mathrm{Tr}\big[(J_x\sigma_x)(J_x\sigma_x)^T - J_x\sigma_x(J_{\tilde{x}}\sigma_{\tilde{x}})^T - J_{\tilde{x}}\sigma_{\tilde{x}}(J_x\sigma_x)^T + J_{\tilde{x}}\sigma_{\tilde{x}}(J_{\tilde{x}}\sigma_{\tilde{x}})^T\big] \\ &\quad= 3\alpha\,\mathrm{Tr}\big[(J_x\sigma_x - J_{\tilde{x}}\sigma_{\tilde{x}})(J_x\sigma_x - J_{\tilde{x}}\sigma_{\tilde{x}})^T\big] \\ &\quad= 3\alpha\|J_x\sigma_x - J_{\tilde{x}}\sigma_{\tilde{x}}\|^2 \end{aligned}$$

(using that $\mathrm{Tr}[.\cdot .^T]$ defines an inner product for matrices and gives rise to the Frobenius matrix norm $\|.\|$). Hence, by the triangle inequality and Lipschitzness of the Jacobian of the flow (which follows a fortiori from the boundedness of the second order derivatives of the flow),

$$\begin{aligned} \|J_x\sigma_x - J_{\tilde{x}}\sigma_{\tilde{x}}\| &\leqslant \|J_x\sigma_x - J_x\sigma_{\tilde{x}}\| + \|J_x\sigma_{\tilde{x}} - J_{\tilde{x}}\sigma_{\tilde{x}}\| \\ &\leqslant \|J_x\|\|\sigma_x - \sigma_{\tilde{x}}\| + \|J_x - J_{\tilde{x}}\|\|\sigma_{\tilde{x}}\| \\ &\leqslant \|J_x\|\|\sigma_x - \sigma_{\tilde{x}}\| + c_2(\sigma,\phi)|x-\tilde{x}|. \end{aligned}$$

Since $\sigma(t,\cdot,q)$ is Lipschitz continuous (uniformly in $(t,q)\in[0,T]\times\mathbb{R}^n$) and $\phi_t(\cdot)$ is Lipschitz continuous (uniformly in $t\in[0,T]$), we can use our assumption (4.2) on σ, to see

$$\|\sigma_x - \sigma_{\tilde{x}}\| \leqslant (\mathrm{const})\times|x-\tilde{x}|. \tag{4.3}$$

Indeed,

$$\begin{aligned} \|\sigma_x - \sigma_{\tilde{x}}\| &= \big\|\sigma\big(t,\phi_t(x),p\cdot J_x\big) - \sigma\big(t,\phi_t(\tilde{x}),p\cdot J_{\tilde{x}}\big)\big\| \\ &\leqslant \big\|\sigma\big(t,\phi_t(x),p\cdot J_x\big) - \sigma\big(t,\phi_t(\tilde{x}),p\cdot J_x\big)\big\| + \big\|\sigma\big(t,\phi_t(\tilde{x}),p\cdot J_x\big) - \sigma\big(t,\phi_t(\tilde{x}),p\cdot J_{\tilde{x}}\big)\big\| \\ &\leqslant c_2(\sigma,\phi)|x-\tilde{x}| + c_1\frac{\alpha|x-\tilde{x}||J_x - J_{\tilde{x}}|}{1+\alpha|x-\tilde{x}|(|J_x|+|J_{\tilde{x}}|)}; \end{aligned}$$

and, noting that $\phi_t\circ\phi_t^{-1} = \mathrm{Id}$ and $\sup_{(t,x)\in[0,T]\times\mathbb{R}^n}\|D\phi_t|_x\|\leqslant c_3$ implies $\|J_x\| = \|D\phi_t^{-1}|_{\phi_t(x)}\|\geqslant 1/c_3$, we have

$$\begin{aligned} c_1\frac{\alpha|x-\tilde{x}||J_x - J_{\tilde{x}}|}{1+\alpha|x-\tilde{x}|(|J_x|+|J_{\tilde{x}}|)} &\leqslant |x-\tilde{x}|\cdot\frac{c_1\alpha|J_x - J_{\tilde{x}}|}{\alpha|x-\tilde{x}|(|J_x|+|J_{\tilde{x}}|)} \\ &\leqslant |x-\tilde{x}|\frac{c_4(\sigma,\phi)|x-\tilde{x}|}{|x-\tilde{x}|(|J_x|+|J_{\tilde{x}}|)} \\ &\leqslant c_5(\sigma,\phi)|x-\tilde{x}|. \end{aligned}$$

Putting things together we have

$$\big|(\mathrm{i})\big| \leqslant c_6(\sigma,\phi)\alpha|x-\tilde{x}|^2.$$

As for (ii), we have that,

$$\begin{aligned} |b_x - b_{\tilde{x}}| &\leqslant \big|b\big(t,\phi_t(x),\langle p,J_x\rangle\big) - b\big(t,\phi_t(\tilde{x}),\langle p,J_x\rangle\big)\big| + \big|b\big(t,\phi_t(\tilde{x}),\langle p,J_x\rangle\big) - b\big(t,\phi_t(\tilde{x}),\langle p,J_{\tilde{x}}\rangle\big)\big| \\ &\leqslant c_7(b)\big(\big|\phi_t(x)-\phi_t(\tilde{x})\big| + |p||J_{\tilde{x}} - J_x|\big) \end{aligned}$$

where $c_7(b)$ is the (uniform in $t\in[0,T]$) Lipschitz bound for $b(t,\cdot,\cdot)$. To get the required estimate we again use the regularity of the flow. Finally, for (iii),

$$\begin{aligned}\text{(iii)} &= \mathrm{Tr}\big[a_{\tilde{x}}\langle p, H_{\tilde{x}}\rangle - a_{\tilde{x}}\langle p, H_x\rangle\big] + \mathrm{Tr}\big[a_{\tilde{x}}\langle p, H_x\rangle - a_x\langle p, H_x\rangle\big]\\ &= \mathrm{Tr}\big[a_{\tilde{x}}\langle p, H_{\tilde{x}} - H_x\rangle\big] + \mathrm{Tr}\big[(a_{\tilde{x}} - a_x)\langle p, H_x\rangle\big].\end{aligned}$$

Using Cauchy–Schwartz (with inner product $\mathrm{Tr}[.\cdot .^T]$) and $p = \alpha(x - \tilde{x})$ it is clear that boundedness of H and a (i.e. $\sup_x |H_x| < \infty$ uniformly in $t \in [0, T]$ and similarly for a) and Lipschitz continuity (i.e. $|H_x - H_{\tilde{x}}| \leqslant (\text{const}) \times |x - \tilde{x}|$ uniformly in $t \in [0, T]$ and similar for a) will suffice to obtain the (desired) estimate

$$\big|\text{(iii)}\big| \leqslant c_8 \times \alpha|x - \tilde{x}|^2.$$

Only Lipschitz continuity of $a_x = \sigma_x\sigma_x^T$ requires a discussion. But this follows, thanks to boundedness of $\sup_x |\sigma_x|$, from showing Lipschitzness of $x \mapsto \sigma_x = \sigma(t, \phi_t(x), \langle p, J_x\rangle)$ uniformly in $t \in [0, T]$ which was already seen in (4.3). This shows that F^ϕ satisfies (2.1), for any $\phi \in \Phi^{(3)}$. To see that F^ϕ satisfies Condition 1 it only remains to see that $F^\phi(t, x, p, X)$ is uniformly continuous whenever p, X remain bounded. To see this first observe that the flow map $\phi_t(x)$, as function of $(t, x) \in [0, T] \times \mathbb{R}^n$, is uniformly continuous (but not bounded) while the derivatives of the (inverse) flow, given by $J.$, $H.$ above, are bounded uniformly continuous maps as functions of t, x. One now easily concludes with the fact the observations that (a) the product of BUC function is again BUC and (b) the composition of a BUC function with a UC function is again BUC.

Example 5 *(F of Hamilton–Jacobi–Bellman type).* From the above examples and Remark 1, we see that $\Phi^{(3)}$-invariant comparison holds when F is given by

$$F(t, x, p, X) = \inf_{\gamma\in\Gamma}\big\{\mathrm{Tr}\big[\sigma(t, x; \gamma)\sigma(t, x; \gamma)^T X\big] + b(t, x; \gamma)\cdot p\big\},$$

the usual non-linearity in the Hamilton–Jacobi–Bellman equation, and more generally

$$F(t, x, p, X) = \inf_{\gamma\in\Gamma}\big\{\mathrm{Tr}\big[\sigma(t, x, p; \gamma)\sigma(t, x, p; \gamma)^T\cdot X\big] + b(t, x, p; \tau)\big\}$$

whenever the conditions in Examples 3 and 4 are satisfied uniformly with respect to $\gamma \in \Gamma$.

Example 6 *(F of Isaac type).* Similarly, $\Phi^{(3)}$-invariant comparison holds for

$$F(t, x, p, X) = \sup_\beta \inf_\gamma\big\{\mathrm{Tr}\big[\sigma(t, x; \beta, \gamma)\sigma(t, x; \beta, \gamma)^T X\big] + b(t, x; \beta, \gamma)\cdot p\big\}$$

(such non-linearities arise in Isaac equation in the theory of differential games), and more generally

$$F(t, x, p, X) = \sup_\beta \inf_\gamma\big\{\mathrm{Tr}\big[\sigma(t, x, p; \beta, \gamma)\sigma(t, x, p; \beta, \gamma)^T\cdot X\big] + b(t, x, p; \beta, \gamma)\big\}$$

whenever the conditions in Examples 3 and 4 are satisfied uniformly with respect to $\beta \in \mathcal{B}$ and $\gamma \in \Gamma$, where $\mathcal{B}$ and Γ are arbitrary index sets.

5. Some lemmas

Lemma 1. *Let $z : [0, T] \to \mathbb{R}^d$ be smooth and assume that we are given C^3-bounded vector fields*[16] *$V = (V_1, \ldots, V_d)$. Then ODE*

$$dy_t = V(y_t)\,dz_t, \quad t \in [0, T]$$

has a unique solution flow (of C^3-diffeomorphisms) $\phi = \phi^z \in \Phi^{(3)}$.

Proof. Standard, e.g. Chapter 4 in [21]. □

[16] In particular, if the vector fields are Lip^γ, $\gamma > p + 2$, $p \geqslant 1$, then they are also C^3-bounded.

Proposition 1. *Let z, V and ϕ be as in Lemma 1. Then u is a viscosity sub- (resp. super-) solution*

$$\dot{u}(t,x) = F\big(t,x,Du,D^2u\big) - Du(t,x)\cdot V(x)\dot{z}(t) \tag{5.1}$$

if and only if $v(t,x) := u(t,\phi_t(x))$ is a viscosity sub- (resp. super-) solution of

$$\dot{v}(t,x) = F^{\phi}\big(t,x,Dv,D^2v\big) \tag{5.2}$$

where F^{ϕ} was defined in (3.2).

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

$$\begin{aligned}\dot{v}(t,x) &= \dot{u}(t,y) + Du(t,y)\cdot\dot{\phi}_t(x) = \dot{u}(t,y) + Du(t,y)\cdot V(y)\dot{z}_t \\ &\leqslant F\big(t,y,Du(t,y),D^2u(t,y)\big) = F^{\phi}\big(t,x,Dv(t,x),D^2v(t,x)\big).\end{aligned}$$

The case when u is a viscosity sub-solution of (5.1) is not much harder: suppose that $(\bar{t},\bar{x})$ is a maximum of $v-\xi$, where $\xi \in C^2([0,T]\times\mathbb{R}^n)$ and define $\psi \in C^2([0,T]\times\mathbb{R}^n)$ by $\psi(t,y) = \xi(t,\phi_t^{-1}(y))$. Set $\bar{y} = \phi_{\bar{t}}(\bar{x})$ so that

$$F\big(\bar{t},\bar{y},D\psi(\bar{t},\bar{y}),D^2\psi(\bar{t},\bar{y})\big) = F^{\phi}\big(\bar{t},\bar{x},D\xi(\bar{t},\bar{x}),D^2\xi(\bar{t},\bar{x})\big).$$

Obviously, $(\bar{t},\bar{y})$ is a maximum of $u-\psi$, and since u is a viscosity sub-solution of (5.1) we have

$$\dot{\psi}(\bar{t},\bar{y}) + D\psi(\bar{t},\bar{y})V(\bar{y})\dot{z}(\bar{t}) \leqslant F\big(\bar{t},\bar{y},D\psi(\bar{t},\bar{y}),D^2\psi(\bar{t},\bar{y})\big).$$

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

$$\dot{\xi}(\bar{t},\bar{x}) \leqslant F^{\phi}\big(\bar{t},\bar{x},D\xi(\bar{t},\bar{x}),D^2\xi(\bar{t},\bar{x})\big)$$

which says precisely that v is a viscosity sub-solution of (5.2). Replacing maximum by minimum and $\leqslant$ by $\geqslant$ in the preceding argument, we see that if u is a super-solution of (5.1), then v is a super-solution of (5.2).

Conversely, the same arguments show that if v is a viscosity sub- (resp. super-) solution for (5.2), then $u(t,y) = v(t,\phi^{-1}(y))$ is a sub- (resp. super-) solution for (5.1). □

6. Proof of the main result

Proof of Theorem 1. Using Lemma 1, we see that $\phi^{\varepsilon} \equiv \phi^{z^{\varepsilon}}$, the solution flow to $dy = V(y)\,dz^{\varepsilon}$, is an element of $\Phi \equiv \Phi^{(3)}$. Set $F^{\varepsilon} := F^{\phi^{\varepsilon}}$. From Proposition 1, we know that u^{ε} is a solution to

$$du^{\varepsilon} = F\big(t,y,Du^{\varepsilon},D^2u^{\varepsilon}\big)\,dt - Du^{\varepsilon}(t,y)\cdot V(y)\,dz^{\varepsilon}(t), \quad u^{\varepsilon}(0,\cdot) = u_0^{\varepsilon}$$

if and only if v^{ε} is a solution to $\partial_t - F^{\varepsilon} = 0$ with $v^{\varepsilon}(0,\cdot) = u_0^{\varepsilon}$. Let $\phi^{\mathbf{z}}$ denote the solution flow to the rough differential equation

$$dy = V(y)\,d\mathbf{z}.$$

Thanks to $\mathrm{Lip}^{\gamma+2}$-regularity of the vector fields $\phi^{\mathbf{z}} \in \Phi$, and in particular a flow of C^3-diffeomorphisms. Set $F^{\mathbf{z}} = F^{\phi^{\mathbf{z}}}$. The “universal” limit theorem [41] holds, in fact, on the level of flows of diffeomorphisms (see [40] and [21, Chapter 11] for more details) tells us that, since z^{ε} tends to $\mathbf{z}$ in rough path sense,

$$\phi^{\varepsilon} \to \phi^{\mathbf{z}} \quad \text{in } \Phi$$

so that, by continuity of F (more precisely: uniform continuity on compacts), we easily deduce that

$$F^{\varepsilon} \to F^{\mathbf{z}} \quad \text{locally uniformly.}$$

From the method of semi-relaxed limits (Lemma 6.1 and Remarks 6.2, 6.3 and 6.4 in [13], see also [15]) the pointwise (relaxed) limits

$$\bar{v} := \limsup{}^* v^\varepsilon,$$
$$\underline{v} := \liminf{}_* v^\varepsilon,$$

are USC sub- resp. LSC super-solutions to $\partial_t - F^{\mathbf{z}} = 0$. Boundedness of $\bar{v}$, $\underline{v}$ is also clear by assumption that (u^ε) is uniformly bounded. Moreover, since $v^\varepsilon(0,\cdot) = u_0^\varepsilon \to u_0$ locally uniformly as $\varepsilon \to 0$ it is not hard to see that $\bar{v}(0,\cdot) = \underline{v}(0,\cdot) = u_0$. (For the reader's convenience we have included a proof of this in Appendix A.) By assumption on Φ-invariant comparison, the equation $\partial_t - F^{\mathbf{z}} = 0$ satisfies comparison. It follows that $v := \bar{v} = \underline{v}$ is the unique (and continuous, since $\bar{v}$, $\underline{v}$ are upper resp. lower semi-continuous) solution to

$$\partial_t v = F^{\mathbf{z}} v, \quad v(0,\cdot) = u_0(\cdot)$$

(and hence that v does not depend on the approximating sequence to $\mathbf{z}$). Moreover, using a simple Dini-type argument (e.g. [13, p. 35] or [1, Lemme 4.1]) one sees that this limit must be uniform on compacts. The proof of (i) is finished by setting

$$u^{\mathbf{z}}(t,x) := v\big(t, \big(\phi_t^{\mathbf{z}}\big)^{-1}(x)\big).$$

(ii) The comparison $|u^{\mathbf{z}} - \hat{u}^{\mathbf{z}}|_{\infty;[0,T]\times\mathbb{R}^n} \leqslant |u_0 - \hat{u}_0|_{\infty;\mathbb{R}^n}$ is a simple consequence of comparison for $v, \hat{v}$ (solutions to $\partial_t v = F^{\mathbf{z}} v$). At last, to see (iii), we argue in the very same way as in (i), starting with

$$F^{\mathbf{z}_n} \to F^{\mathbf{z}} \quad \text{locally uniformly}$$

to see that $v^n \to v$ locally uniformly, i.e. uniformly on compacts. □

7. Applications to stochastic partial differential equations

Applications to SPDEs are path-by-path, i.e. by taking $\mathbf{z}$ to be a typical realization of Brownian motion and Lévy's area, $\mathbf{B}(\omega) \equiv (B, A)$, also known as enhanced Brownian motion or Brownian rough path. The continuity property (iii) of our Theorem 1 allows to identify (1.7) with $\mathbf{z} = \mathbf{B}(\omega)$ as *Stratonovich solution* to the non-linear SPDE

$$du = F\big(t, x, Du, D^2u\big)\,dt - Du \cdot V(x) \circ dB, \quad u(0,\cdot) = u_0.$$

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

$$du = F\big(t, x, Du, D^2u\big)\,dt - Du \cdot V(x)\,d\mathbf{B}, \quad u(0,\cdot) = u_0,$$

as constructed in Theorem 1. If one takes this piecewise linear approximation property as *definition* of a solution in Stratonovich sense[17] this identification is trivially settled. More interestingly, there is a number of Wong–Zakai approximation results for SPDEs, ranging from [5,48] to [28,29]. Any solution of ours that is also covered in the aforementioned references is then indeed a Stratonovich SPDE solution in the usual sense.[18] Of course, (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.) At last, we point out that our solution also constitutes a *stochastic viscosity solution* in the sense of Lions and Souganidis [35,36,38]: adapted to the present setting, and recalling the notation used in the proof of Theorem 1, this amounts to call u a (stochastic viscosity) solution if $v(t,x) := u(t, (\phi_t^{\mathbf{B}})^{-1}(x))$ satisfies the (random) PDE $\partial_t v = F^{\mathbf{B}} v$ in viscosity sense.[19] Observe that uniqueness of stochastic viscosity solutions then follows from the classical theory of viscosity solutions of fully non-linear second-order partial differential equations. (After all, our assumptions guarantee that $\partial_t - F^{\mathbf{B}}$ satisfies comparison.)

[17] … commonly done in the context of anticipating stochastic analysis, see [43,10] for instance.

[18] The same logic has been used by T. Lyons in [39] to identify rough differential equation driven by $\mathbf{B}$ as Stratonovich SDE solutions.

[19] The actual definition of Lions–Souganidis is a localized version of this and allows for noise of the form $H(x, Du) \circ dB$ with H non-linear in Du. When H is linear in Du, the standing assumption in the present paper, the global and local definition are easily seen to be equivalent.

Remark 2 *(Itô versus Stratonovich).* Note that similar **SPDEs in Itô-form** need not be, in general, well-posed. Consider the following (well-known) linear example

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

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

Remark 3 *(Regularity of V).* Applied to the Brownian context (finite p-variation for any $p > 2$) the regularity assumption of Theorem 1 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. 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 4 *(Space–time regularity of SPDE solutions).* Since $u(t,x) = v(t, \phi_t^{\mathbf{B}}(x))$ and $\phi_t^{\mathbf{B}}$ is a flow of C^3- diffeomorphisms the regularity of u is readily reduced to regularity properties of v, classical viscosity solution to $\partial_t v = F^{\mathbf{B}} v$. Unless one makes very specific assumptions on F this is a difficult problem in its own right; see the relevant remarks in [13] for instance.

Let us now give some applications, typical in the sense that they have been studied in great detail in the case of classical (Stratonovich) stochastic differential equations.

Approximations. *Any* approximation result to $\mathbf{B}$ 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. [21, Chapter 13] and the references therein) examples include piecewise linear-, mollifier- and Karhunen–Loeve approximations, as well as (weak) Donsker type random walk approximations [4]. The point being made, we shall not spell out more details here.

Twisted approximations. The following result implies en passant that there is no (classical) pathwise theory of SPDEs in presence of spatial dependence in the Hamilonian terms.

Theorem 2. *Let $V = (V_1, \ldots, V_d)$ be a collection of C^∞-bounded vector fields on $\mathbb{R}^n$ and B a d-dimensional standard Brownian motion. Then, for every $\alpha = (\alpha_1, \ldots, \alpha_N) \in \{1, \ldots, d\}^N$, $N \geqslant 2$, there exists* (*piecewise*) *smooth approximations (z^k) to B, with each z^k only dependent on $\{B(t)\colon t \in D^k\}$ where (D^k) is a sequence of dissections of* $[0,T]$ *with mesh tending to zero, such that almost surely*

$$z^k \to B \quad \textit{uniformly on } [0,T],$$

but u^k, solutions to

$$du^k = F\big(t, x, Du^k, D^2u^k\big)\,dt - Du^k(t,x)\cdot V(x)\,dz^k, \quad u^k(0,\cdot) = u_0 \in \mathrm{BUC}\big(\mathbb{R}^n\big)$$

(*with assumptions on F as formulated in Theorem* 1) *converge almost surely locally uniformly to the solution of the "wrong" differential equation*

$$du = \big[F\big(t,x,Du,D^2u\big) - Du(t,x)\cdot V_\alpha(x)\big]\,dt - Du(t,x)\cdot V(x)\circ dB$$

where V_α is the bracket-vector field given by $V_\alpha = [V_{\alpha_1}, [V_{a_2}, \ldots [V_{\alpha_{N-1}}, V_{\alpha_N}]]]$.

Proof. The rough path regularity of $\mathbf{B}(\omega)$ implies that higher iterated (Stratonovich) integrals are deterministically defined; see [39, First theorem]. Doing this up to level N yields a (rough path) $S_N(\mathbf{B})$ and we perturbe it in the highest level, linearly in the

$\left[e_{\alpha_1},\left[e_{a_2},\ldots[e_{\alpha_{N-1}},e_{\alpha_N}]\right]\right]$-direction

of $S_N(\mathbf{B})$ viewed as element in the step-N free nilpotent Lie algebra. This yields a (level-N) rough path $\tilde{\mathbf{B}}$ and we can find approximations (z^k) that converge almost surely to $\tilde{\mathbf{B}}$ in rough path topology (see [17]). One identifies standard RDEs driven by $\tilde{\mathbf{B}}$ as RDEs-with-drift (driven along the original vector fields by $d\mathbf{B}$, and along V_α by dt). The resulting identification obviously holds on the level of RDE flows and thus

$$u^{z^k}(t,x)=v\big(t,\big(\phi_t^{z^k}\big)^{-1}(x)\big)\to u^{\tilde{\mathbf{B}}}(t,x)=v\big(t,\big(\phi_t^{\tilde{\mathbf{B}}}\big)^{-1}(x)\big).$$

The flow identification then implies that

$$du=F\big(t,x,Du,D^2u\big)\,dt-Du(t,x)\cdot V(x)\,d\tilde{\mathbf{B}}$$

is equivalent to the equation with $V(x)\,d\tilde{\mathbf{B}}$ replaced by $V(x)\,d\mathbf{B}+V_\alpha(x)\,dt$. □

Remark 5. The attentive reader will have noticed that the preceding result also holds when the Stratonovich differential $\circ\, dB$ is replaced by dz for some $z\in C^1([0,T],\mathbb{R}^d)$; it can then be viewed as result on the effective behaviour of a (deterministic) non-linear parabolic equations with coefficients that exhibit highly oscillatory behaviour in time.

Support results. In conjunction with known support properties of $\mathbf{B}$ (e.g. [33] 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. 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 [27,26,24,23,25].

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. [33] in p-variation and [19] 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. [11] or [18], 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.

8. Further remarks

We have discussed a rough paths approach to stochastic partial differential equation of the form

$$du=F\big(t,x,Du,D^2u\big)\,dt+\sum_{i=1}^d H_i(x,Du)\circ dB^i$$

with fully non-linear F and Hamiltonian $H=H(x,Du)$, linear in Du. When F is (semi-) linear, uniformly elliptic, there are various results (based on *backward stochastic differential equations*) for solving SPDE with general $H_i=H_i(u,Du,x)$; see [45] for instance. Under a semi-linearity assumption ($F=\Delta u+f(t,x,Du)$) and non-linear $H_i=H_i(Du,x)$, subject to restrictive algebraic properties, a pathwise approach was carried out by Iftimie and Varsan [30].

It is worth pointing out that SPDEs of the form

$$du=F\big(t,x,Du,D^2u\big)\,dt+\sum_{i=1}^d G_i(u,x)\circ dB^i,$$

have also benefited from global (Doss–Sussmann) type transformations; see [37,6,7]. Although this suggests that the general rough path methodology of the present paper is also applicable for such SPDEs, by considering rough PDEs of the form

$$du = F\big(t, x, Du, D^2 u\big)\,dt + G(u, x)\,d\mathbf{z},$$

the matter is far from straight-forward: a transformation based on the (stochastic/rough) flow associsted to G, see [37], leads to a transformed PDE which does not fit in the standard viscosity theory and we shall return to this in future work.

Acknowledgements

M. Caruana was supported by EPSRC grant EP/E048609/1. P.K. Friz has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement No. 258237. H. Oberhauser was supported by a DOC-fellowship of the Austrian Academy of Sciences. Part of this work was undertaken while the last two authors visited the Radon Institute (Austria). The authors would like to thank Guy Barles for a helpful email exchange.

Appendix A. Comparison for parabolic equations

Recall that USC (resp. LSC) refers to upper (resp. lower) semi-continuity. Let $u \in \mathrm{USC}([0,T) \times \mathbb{R}^n)$ be a bounded sub-solution to $\partial_t - F$; that is, $\partial_t u - F(t, x, Du, D^2u) \leqslant 0$ if u is smooth and with the usual viscosity definition otherwise. Similarly, let $v \in \mathrm{LSC}([0,T) \times \mathbb{R}^n)$ be a bounded super-solution.

Theorem 3. *Assume Condition* 1. *Then comparison holds. That is,*

$$u_0 \leqslant v_0 \quad \text{on } \mathbb{R}^n \quad \Rightarrow \quad u \leqslant v \quad \text{on } [0,T) \times \mathbb{R}^n,$$

where $u_0 = u(0,\cdot) \in \mathrm{USC}(\mathbb{R}^n)$ *and* $v_0 = v(0,\cdot) \in \mathrm{LSC}(\mathbb{R}^n)$ *denote the* (*bounded*) *initial data.*

Proof. We follow the argument given in the User's Guide [13, Section 8]. Without loss of generality, we may assume that $\partial_t u - F(t, x, Du, D^2u) \leqslant -c < 0$ and that $\lim_{t\to T} u(t,x) = -\infty$ uniformly in $x \in \mathbb{R}^n$. We aim to contradict the existence of a point $(s,z) \in (0,T) \times \mathbb{R}^n$ such that

$$u(s,z) - v(s,z) = \delta > 0.$$

To this end, consider a maximum point $(\hat{t}, \hat{x}, \hat{y}) \in [0,T) \times \mathbb{R}^n \times \mathbb{R}^n$ of

$$\phi(t,x,y) = u(t,x) - v(t,y) - \frac{\alpha}{2}|x-y|^2 - \varepsilon\big(|x|^2 + |y|^2\big).$$

We first argue that, for small (resp. large) enough values of ε and α, the optimizing time parameter $\hat{t} \in [0,T)$ cannot be zero. Indeed, assuming $\hat{t} = 0$ we can estimate

$$\begin{aligned} \delta - 2\varepsilon|z|^2 &= \phi(s,z,z) \\ &\leqslant \phi(0,\hat{x},\hat{y}) \\ &= \sup_{x,y}\left[u_0(x) - v_0(y) - \frac{\alpha}{2}|x-y|^2 - \varepsilon\big(|x|^2+|y|^2\big)\right]. \end{aligned}$$

From Lemma 3.1 in the User's Guide (applied to the $\mathrm{USC}(\mathbb{R}^n)$ resp. $\mathrm{LSC}(\mathbb{R}^n)$ map given by $u_0(x) - \varepsilon|x|^2$ resp. $v_0(y) - \varepsilon|y|^2$) it follows that

$$\begin{aligned} \lim_{\alpha\to\infty} \phi(0,\hat{x},\hat{y}) &= \sup_x\big[u_0(x) - v_0(x) - 2\varepsilon|x|^2\big] \\ &\leqslant |u_0 - v_0|_{\infty;\mathbb{R}^n} \leqslant 0 \quad \text{by assumption.} \end{aligned}$$

In particular, there exists $\alpha_0 = \alpha_0(\delta)$ such that $\phi(0,\hat{x},\hat{y}) < \delta/3$ for $\alpha \geqslant \alpha_0$. If we then choose $\varepsilon \leqslant \varepsilon_0 = \varepsilon_0(\delta, z)$, determined by $2\varepsilon_0|z|^2 = \delta/3$ for instance, we are left with the contradiction

$$\delta - \delta/3 \leqslant \delta - 2\varepsilon|z|^2 \leqslant \phi(0, \hat{x}, \hat{y}) < \delta/3.$$

It follows that $\hat{t} \in (0, T)$ whenever $\varepsilon \leqslant \varepsilon_0$ and $\alpha \geqslant \alpha_0$, which we shall assume from here on. (In fact, we shall send $\varepsilon \to 0$, and then $\alpha \to \infty$, in what follows.)

Again, the plan is to arrive at a contradiction (so that we have to reject the existence of a point $(s, z) \in (0, T) \times \mathbb{R}^n$ at which $u(s, z) - v(s, z) > 0$) altogether. To this end, let us rewrite $\phi(t, x, y)$ as

$$\phi(t, x, y) = u^\varepsilon(t, x) - v^\varepsilon(t, y) - \frac{\alpha}{2}|x - y|^2$$

where $u^\varepsilon(t, x) = u(t, x) - \varepsilon|x|^2$ and $v^\varepsilon(t, y) = v(t, y) + \varepsilon|y|^2$. Since u^ε (resp. v^ε) are upper (resp. lower) semi-continuous we can apply the (parabolic) theorem of sums [13, Theorem 8.3] at $(\hat{t}, \hat{x}, \hat{y})$ to learn that there are numbers a, b and $X, Y \in S^n$ such that

$$\bigl(a, \alpha(\hat{x} - \hat{y}), X\bigr) \in \bar{\mathcal{P}}^{2,+} u^\varepsilon(\hat{t}, \hat{x}), \qquad \bigl(b, \alpha(\hat{x} - \hat{y}), Y\bigr) \in \bar{\mathcal{P}}^{2,-} v^\varepsilon(\hat{t}, \hat{y}) \tag{A.1}$$

such that $a - b = 0$ and such that one has the estimate (2.2). It is easy to see (cf. [13, Remark 2.7]) that (A.1) is equivalent to

$$\bigl(a, \alpha(\hat{x} - \hat{y}) + 2\varepsilon\hat{x}, X + 2\varepsilon I\bigr) \in \bar{\mathcal{P}}^{2,+} u(\hat{t}, \hat{x}),$$
$$\bigl(b, \alpha(\hat{x} - \hat{y}) - 2\varepsilon\hat{y}, Y - 2\varepsilon I\bigr) \in \bar{\mathcal{P}}^{2,-} v(\hat{t}, \hat{y});$$

using that $\partial_t u - F(t, x, Du, D^2u) \leqslant -c$ and $\partial_t v - F(t, x, Dv, D^2v) \geqslant 0$ (always understood in the sense of viscosity sub- resp. super-solutions) we then see that

$$a - F\bigl(\hat{t}, \hat{x}, \alpha(\hat{x} - \hat{y}) + 2\varepsilon\hat{x}, X + 2\varepsilon I\bigr) \leqslant -c,$$
$$b - F\bigl(\hat{t}, \hat{y}, \alpha(\hat{x} - \hat{y}) - 2\varepsilon\hat{y}, Y - 2\varepsilon I\bigr) \geqslant 0.$$

Using $a = b$, this implies

$$0 \leqslant c \leqslant F\bigl(\hat{t}, \hat{x}, \alpha(\hat{x} - \hat{y}) + 2\varepsilon\hat{x}, X + 2\varepsilon I\bigr) - F\bigl(\hat{t}, \hat{y}, \alpha(\hat{x} - \hat{y}) - 2\varepsilon\hat{y}, Y - 2\varepsilon I\bigr).$$

The last step consists in showing that the right-hand side converges to zero by first sending $\varepsilon \to 0$ and then $\alpha \to \infty$. (This yields the desired contradiction which ends the proof.) If ε were absent (e.g. set $\varepsilon = 0$ throughout) we would estimate

$$F\bigl(\hat{t}, \hat{x}, \alpha(\hat{x} - \hat{y}), X\bigr) - F\bigl(\hat{t}, \hat{y}, \alpha(\hat{x} - \hat{y}), Y\bigr) \leqslant \theta\bigl(\alpha|\hat{x} - \hat{y}|^2 + |\hat{x} - \hat{y}|\bigr)$$

and conclude (Lemma 3.1 in the User's Guide) that

$$\alpha|\hat{x} - \hat{y}|^2, |\hat{x} - \hat{y}| \to 0 \quad \text{as } \alpha \to \infty$$

in conjunction with continuity of θ at $0+$. The present case, $\varepsilon > 0$, is essentially reduced to the case $\varepsilon = 0$ by adding/subtracting

$$F\bigl(\hat{t}, \hat{x}, \alpha(\hat{x} - \hat{y}), X\bigr) - F\bigl(\hat{t}, \hat{y}, \alpha(\hat{x} - \hat{y}), Y\bigr).$$

It follows that $c \leqslant \text{(i)} + \text{(ii)} + \text{(ii)}$ where

$$\text{(i)} = \bigl|F\bigl(\hat{t}, \hat{x}, \alpha(\hat{x} - \hat{y}) + 2\varepsilon\hat{x}, X + 2\varepsilon I\bigr) - F\bigl(\hat{t}, \hat{x}, \alpha(\hat{x} - \hat{y}), X\bigr)\bigr|,$$
$$\text{(ii)} = \bigl|F\bigl(\hat{t}, \hat{y}, \alpha(\hat{x} - \hat{y}) - 2\varepsilon\hat{y}, Y - 2\varepsilon I\bigr) - F\bigl(\hat{t}, \hat{y}, \alpha(\hat{x} - \hat{y}), Y\bigr)\bigr|,$$
$$\text{(iii)} = \theta\bigl(\alpha|\hat{x} - \hat{y}|^2 + |\hat{x} - \hat{y}|\bigr).$$

From Lemma 2 below, we see that (a) $p = \alpha(\hat{x} - \hat{y})$ remains, for fixed α, bounded as $\varepsilon \to 0$, (b) $2\varepsilon|\hat{x}|$ and $2\varepsilon|\hat{y}|$ tend to zero as $\varepsilon \to 0$, for fixed α, and (c)

$$\limsup_{\alpha\to\infty} \limsup_{\varepsilon\to 0} \bigl(\alpha|\hat{x} - \hat{y}|^2 + |\hat{x} - \hat{y}|\bigr) = \lim_{\alpha\to\infty} \limsup_{\varepsilon\to 0} \bigl(\alpha|\hat{x} - \hat{y}|^2 + |\hat{x} - \hat{y}|\bigr) = 0.$$

We also note that (2.2) implies that any matrix norm of X, Y is bounded by a constant times α, independent of ε. Since F is assumed to be uniformly continuous whenever its gradient and Hessian argument remain in abounded set, combining all this information shows that

$$\limsup_{\varepsilon\to 0}(\mathrm{i}), \qquad \limsup_{\varepsilon\to 0}(\mathrm{ii}), \qquad \limsup_{\varepsilon\to 0}(\mathrm{iii})$$

all tend to 0 as $\alpha \to \infty$. In summary,

$$0 < c \leqslant \lim_{\alpha\to\infty} \limsup_{\varepsilon\to 0}\big[(\mathrm{i}) + (\mathrm{ii}) + (\mathrm{ii})\big] = 0$$

which is the desired contradiction. The proof is now finished. □

Lemma 2. *Let $u \in \mathrm{USC}([0,T)\times\mathbb{R}^n)$ bounded from above and $v \in \mathrm{LSC}([0,T)\times\mathbb{R}^n)$ bounded from below. Consider a maximum point $(\hat t, \hat x, \hat y) \in (0,T)\times\mathbb{R}^n\times\mathbb{R}^n$ of*

$$\phi(t,x,y) = u(t,x) - v(t,y) - \frac{\alpha}{2}|x-y|^2 - \varepsilon\big(|x|^2 + |y|^2\big),$$

where $\alpha, \varepsilon > 0$. Then

$$\limsup_{\varepsilon\to 0} \alpha(\hat x - \hat y) = C(\alpha) < \infty, \tag{A.2}$$

$$\limsup_{\alpha\to\infty} \limsup_{\varepsilon\to 0} \varepsilon\big(|\hat x|^2 + |\hat y|^2\big) = 0, \tag{A.3}$$

$$\limsup_{\alpha\to\infty} \limsup_{\varepsilon\to 0} \left(\frac{\alpha}{2}|\hat x - \hat y|^2 + |\hat x - \hat y|\right) = 0. \tag{A.4}$$

Remark 6. A similar lemma is found in [31] or (without t dependence) in Barles' book [1, Lemme 4.3]; the order in which limits are taken is important and suggests the notation

$$\limsup_{\varepsilon\ll\frac{1}{\alpha}\to 0}(\ldots) := \limsup_{\alpha\to\infty}\limsup_{\varepsilon\to 0}(\ldots), \qquad \liminf_{\varepsilon\ll\frac{1}{\alpha}\to 0}(\ldots) := \liminf_{\alpha\to\infty}\liminf_{\varepsilon\to 0}(\ldots).$$

Proof. We start with some notation, where unless otherwise stated $t \in [0,T]$ and $x, y \in \mathbb{R}^n$,

$$M_{\alpha,\varepsilon} := \sup_{t,x,y}\phi(t,x,y) = u(\hat t,\hat x) - v(\hat t,\hat y) - \frac{\alpha}{2}|\hat x - \hat y|^2 - \varepsilon\big(|\hat x|^2 + |\hat y|^2\big),$$

$$M(h) := \sup_{t,x,y:\ |x-y|\leqslant h}\big[u(t,x) - v(t,y)\big] \geqslant \sup_{t,x}\big[u(t,x) - v(t,x)\big],$$

$$M' :=\downarrow \lim_{h\to 0} M(h).$$

(As indicated, M' exists as limit of $M(h)$, non-increasing in h and bounded from below.)

Step 1. Take $t = x = y = 0$ as argument of $\phi(t,x,y)$. Since $M_{\alpha,\varepsilon} = \sup\phi$ we have

$$c = u(0,0) - v(0,0) \leqslant M_{\alpha,\varepsilon} = u(\hat t,\hat x) - v(\hat t,\hat y) - \frac{\alpha}{2}|\hat x - \hat y|^2 - \varepsilon\big(|\hat x|^2 + |\hat y|^2\big)$$

and hence, for a suitable constant C (e.g. $C^2 := \sup u + \sup(-v) + c$)

$$\frac{\alpha}{2}|\hat x - \hat y|^2 + \varepsilon\big(|\hat x|^2 + |\hat y|^2\big) \leqslant C^2$$

which implies

$$|\hat x - \hat y| \leqslant C\sqrt{2/\alpha} \tag{A.5}$$

and hence $\alpha|\hat x - \hat y| \leqslant \sqrt{2\alpha}C$ which is the first claimed estimate (A.2).

Step 2. We first argue that it is enough to show the (two) estimates

$$\limsup_{\varepsilon\ll\frac{1}{\alpha}\to 0}\big[u(\hat t,\hat x)-v(\hat t,\hat y)\big]\leqslant M'\leqslant \liminf_{\varepsilon\ll\frac{1}{\alpha}\to 0} M_{\alpha,\varepsilon}. \tag{A.6}$$

Indeed, from $\frac{\alpha}{2}|\hat x-\hat y|^2+\varepsilon(|\hat x|^2+|\hat y|^2)=u(\hat t,\hat x)-v(\hat t,\hat y)-M_{\alpha,\varepsilon}$ it readily follows that

$$\begin{aligned}\limsup_{\varepsilon\ll\frac{1}{\alpha}\to 0}\frac{\alpha}{2}|\hat x-\hat y|^2+\varepsilon\big(|\hat x|^2+|\hat y|^2\big)&\leqslant \limsup_{\varepsilon\ll\frac{1}{\alpha}\to 0}\big[u(\hat t,\hat x)-v(\hat t,\hat y)-M_{\alpha,\varepsilon}\big]\\ &=\limsup_{\varepsilon\ll\frac{1}{\alpha}\to 0}\big[u(\hat t,\hat x)-v(\hat t,\hat y)\big]-\liminf_{\varepsilon\ll\frac{1}{\alpha}\to 0} M_{\alpha,\varepsilon}\\ &\leqslant 0\quad (\text{and hence} =0).\end{aligned}$$

This already gives (A.3) and also (A.4), noting that

$$|\hat x-\hat y|=\alpha^{-1/2}\alpha^{1/2}|\hat x-\hat y|\leqslant\frac{1}{2\alpha}+\frac{\alpha}{2}|\hat x-\hat y|^2.$$

We are left to show (A.6). For the first estimate, it suffices to note that, from (A.5) and the definition of $M(h)$ applied with $h=C\sqrt{2/\alpha}$,

$$\begin{aligned}\limsup_{\varepsilon\ll\frac{1}{\alpha}\to 0}\big[u(\hat t,\hat x)-v(\hat t,\hat y)\big]&\leqslant\limsup_{\varepsilon\ll\frac{1}{\alpha}\to 0} M\bigg(\sqrt{\frac{2}{\alpha}}C\bigg)\\ &=\lim_{\alpha\to\infty} M\bigg(\sqrt{\frac{2}{\alpha}}C\bigg)=M'.\end{aligned}$$

We now turn to the second estimate in (A.6). From the very definition of M' as $\lim_{h\to 0}M(h)$, there exists a family (t_h,x_h,y_h) so that

$$|x_h-y_h|\leqslant h\quad\text{and}\quad u(t_h,x_h)-v(t_h,x_h)\to M'\quad\text{as } h\to 0. \tag{A.7}$$

For every α,ε we may take (t_h,x_h,y_h) as argument of ϕ; since $M_{\alpha,\varepsilon}=\sup\phi$ we have

$$u(t_h,x_h)-v(t_h,y_h)-\frac{\alpha}{2}h^2-\varepsilon\big(|x_h|^2+|y_h|^2\big)\leqslant M_{\alpha,\varepsilon}. \tag{A.8}$$

Take now $\varepsilon=\varepsilon(h)\to 0$ with $h\to 0$; fast enough so that $\varepsilon(|x_h|^2+|y_h|^2)\to 0$; for instance $\varepsilon(h):=h/(1+(|x_h|^2+|y_h|^2))$ would do. It follows that

$$\begin{aligned}M'&=\lim_{h\to 0}u(t_h,x_h)-v(t_h,y_h)\\ &=\liminf_{h\to 0}u(t_h,x_h)-v(t_h,y_h)-\frac{\alpha}{2}h^2-\varepsilon\big(|x_h|^2+|y_h|^2\big)\\ &\leqslant\liminf_{h\to 0}M_{\alpha,\varepsilon_h}=\liminf_{\varepsilon\to 0}M_{\alpha,\varepsilon}\quad\text{by monotonicity of } M_{\alpha,\varepsilon}\text{ in }\varepsilon.\end{aligned}$$

Since this is valid for every α, we also have

$$M'\leqslant\liminf_{\alpha\to\infty}\liminf_{\varepsilon\to 0}M_{\alpha,\varepsilon}.$$

This is precisely the second estimate in (A.6) and so the proof is finished. □

Appendix B. Initial data under semi-relaxed limits

Let $(v^\varepsilon\colon\varepsilon>0)$ denote a family of uniformly bounded viscosity solutions $v^\varepsilon(t,x)$ to

$$\partial_t v^\varepsilon-F^\varepsilon\big(t,x,Dv^\varepsilon,D^2v^\varepsilon\big)=0,\qquad v^\varepsilon(0,\cdot)=g^\varepsilon\in\mathrm{BC}\big([0,T]\times\mathbb{R}^n\big)$$

where $F^\varepsilon=F^\varepsilon(t,x,p,X)$ is a continuous function of its arguments. Assume $g^\varepsilon\to g\in\mathrm{BC}$ locally uniformly and $F^\varepsilon\to F$ locally uniformly and recall that the semi-relaxed limits are defined by

$$\bar{v}(t,x) := \limsup_{(s,y)\in[0,T]\times\mathbb{R}^n:\ s\to t,\ y\to x,\ \varepsilon\to 0} v^\varepsilon(s,y),$$
$$\underline{v}(t,x) := \liminf_{(s,y)\in[0,T]\times\mathbb{R}^n:\ s\to t,\ y\to x,\ \varepsilon\to 0} v^\varepsilon(s,y).$$

Proposition 2. *We have* $\bar{v}(0,x) = \underline{v}(0,x) = g(x)$.

Proof. We adapt the argument of Fleming and Soner [15, Section VII.5] to our setting and focus on showing $\bar{v}(0,x) = g(x)$, the other equality being similar. Trivially $\bar{v}(0,x) \geqslant g(x)$. Suppose equality does not hold. Then there exist $x_0 \in \mathbb{R}^n$ and $\delta > 0$ so that

$$\bar{v}(0,x_0) = g(x_0) + \delta.$$

We can assume without loss of generality $g(x_0) = 0$; for otherwise consider $\tilde{v}^\varepsilon(t,x) := v^\varepsilon(t,x) - g(x_0)$. Since $g^\varepsilon \to g$ uniformly near x_0 there are $\rho > 0$ and $\varepsilon_0 > 0$ such that

$$g^\varepsilon(x) < \delta/2 \quad \text{whenever } |x-x_0|^2 < \rho,\ \varepsilon < \varepsilon_0. \tag{B.1}$$

(We could take $\rho = 1$ in fact.) Define the (smooth) test-function

$$w(t,x) = \gamma t + K|x-x_0|^2$$

where $K = (\sup_{\varepsilon>0}|v^\varepsilon|_{\infty;[0,T]\times\mathbb{R}^n} + 1)/\rho$ and $\gamma \geqslant K$ will be chosen later. Now, if x is such that $g^\varepsilon(x) \geqslant \delta/2$, and if $\varepsilon < \varepsilon_0$, then (B.1) implies that we must have $|x-x_0|^2 \geqslant \rho$; it then follows that

$$w(t,x) \geqslant K|x-x_0|^2 \geqslant K\rho \geqslant v^\varepsilon(t,x) - g^\varepsilon(x_0) + \big[1 + g^\varepsilon(x_0)\big].$$

By making ε_0 smaller if necessary we can assume that $|g^\varepsilon(x_0)| < 1/2$, say, for all $\varepsilon < \varepsilon_0$ which shows that

$$w(t,x) > v^\varepsilon(t,x) - g^\varepsilon(x_0) \quad \text{whenever } g^\varepsilon(x) \geqslant \delta/2,\ \varepsilon < \varepsilon_0. \tag{B.2}$$

For $\varepsilon > 0$, choose

$$(t_\varepsilon, x_\varepsilon) \in \operatorname{argmax}\big\{v^\varepsilon(t,x) - w(t,x):\ (t,x) \in [0,T]\times\mathbb{R}^n\big\}.$$

Also the definition of $\bar{v}$ implies that there exist $\varepsilon_n \to 0$ and $(s_n, y_n) \to (0, x_0)$ such that

$$\delta = \bar{v}(0,x_0) = \lim_n v^{\varepsilon_n}(s_n, y_n).$$

Set $(t_n, x_n) = (t_{\varepsilon_n}, x_{\varepsilon_n})$. Then

$$\liminf_{n\to\infty} v^{\varepsilon_n}(t_n,x_n) - w(t_n,x_n) \geqslant v^{\varepsilon_n}(s_n,y_n) - w(s_n,y_n) = \delta. \tag{B.3}$$

We claim that $t_n \neq 0$ for sufficiently large n. Indeed, if $t_n = 0$ then $v^{\varepsilon_n}(t_n,x_n) = g^{\varepsilon_n}(x_n)$. The above inequality then yields that there is n_0 such that, for all $n \geqslant n_0$

$$g^{\varepsilon_n}(x_n) = v^{\varepsilon_n}(t_n,x_n) \geqslant \frac{\delta}{2} + w(t_n,x_n) \geqslant \frac{\delta}{2}$$

and, from (B.2), $w(t_n,x_n) > v^{\varepsilon_n}(t_n,x_n) - g^{\varepsilon_n}(x_0)$. It follows that

$$\liminf_{n\to\infty} v^{\varepsilon_n}(t_n,x_n) - w(t_n,x_n) \leqslant \liminf_{n\to\infty} g^{\varepsilon_n}(x_0) = 0$$

which contradicts (B.3). Hence, for all $n \geqslant n_0$ we have $t_n \neq 0$ and the viscosity property of v^ε gives (with all derivatives evaluated at t_n, x_n)

$$\begin{aligned} 0 &\geqslant \partial_t w(t_n,x_n) - F^{\varepsilon_n}\big(t_n, x_n, Dw, D^2 w\big) \\ &= \gamma - F^{\varepsilon_n}\big(t_n, x_n, 2K(x_n - x_0), 2KI\big). \end{aligned}$$

Note that $x_n = x_n^\gamma$ (to indicate the dependence on γ) remains bounded, uniformly in $n \in \{n_0, \ldots\}$ and $\gamma \in [K, \infty)$. Since F^{ε_n} is locally uniformly continuous it is also locally uniformly bounded. A contradiction is now obtained by taking γ large enough. □

References

[1] Guy Barles, Solutions de viscosité des équations de Hamilton–Jacobi, Springer, 2004.
[2] Guy Barles, Samuel Biton, Mariane Bourgoing, Olivier Ley, Uniqueness results for quasilinear parabolic equations through viscosity solutions' methods, Calc. Var. Partial Differential Equations 18 (2) (2003) 159–179.
[3] Guy Barles, Samuel Biton, Olivier Ley, A geometrical approach to the study of unbounded solutions of quasilinear parabolic equations, Arch. Ration. Mech. Anal. 162 (4) (2002) 287–325.
[4] Emmanuel Breuillard, Peter Friz, Martin Huesmann, From random walks to rough paths, Proc. Amer. Math. Soc. 137 (2009) 3487–3496.
[5] Zdzisław Brzeźniak, Franco Flandoli, Almost sure approximation of Wong–Zakai type for stochastic partial differential equations, Stochastic Process. Appl. 55 (2) (1995) 329–358.
[6] Rainer Buckdahn, Jin Ma, Stochastic viscosity solutions for nonlinear stochastic partial differential equations. I, Stochastic Process. Appl. 93 (2) (2001) 181–204.
[7] Rainer Buckdahn, Jin Ma, Stochastic viscosity solutions for nonlinear stochastic partial differential equations. II, Stochastic Process. Appl. 93 (2) (2001) 205–228.
[8] Rainer Buckdahn, Jin Ma, Pathwise stochastic Taylor expansions and stochastic viscosity solutions for fully nonlinear stochastic PDEs, Ann. Probab. 30 (3) (2002) 1131–1171.
[9] Rainer Buckdahn, Jin Ma, Pathwise stochastic control problems and stochastic HJB equations, SIAM J. Control Optim. 45 (6) (2007) 2224–2256 (electronic).
[10] Laure Coutin, Peter Friz, Nicolas Victoir, Good rough path sequences and applications to anticipating stochastic calculus, Ann. Probab. 35 (3) (2007) 1172–1193.
[11] Laure Coutin, Zhongmin Qian, Stochastic analysis, rough path analysis and fractional Brownian motions, Probab. Theory Related Fields 122 (1) (2002) 108–140.
[12] Michael G. Crandall, Viscosity Solutions: A Primer, Lecture Notes in Math., vol. 1660, 1995.
[13] Michael G. Crandall, Hitoshi Ishii, Pierre-Louis Lions, User's guide to viscosity solutions of second order partial differential equations, Bull. Amer. Math. Soc. (N.S.) 27 (1) (1992) 1–67.
[14] Mark H.A. Davis, Gabriel Burstein, A deterministic approach to stochastic optimal control with application to anticipative control, Stoch. Stoch. Rep. 40 (3–4) (1992) 203–256.
[15] Wendell H. Fleming, H. Mete Soner, Controlled Markov Processes and Viscosity Solutions, second edition, Stoch. Model. Appl. Probab., vol. 25, Springer, New York, 2006.
[16] P. Friz, T. Lyons, D. Stroock, Lévy's area under conditioning, Ann. Inst. H. Poincaré Probab. Statist. 42 (1) (2006) 89–101.
[17] Peter Friz, Harald Oberhauser, Rough path limits of the Wong–Zakai type with a modified drift term, J. Funct. Anal. 256 (2009) 3236–3256.
[18] Peter Friz, Nicolas Victoir, Differential equations driven by Gaussian signals, Ann. Inst. H. Poincaré Probab. Statist. 46 (2) (2010) 369–413.
[19] Peter Friz, Nicolas Victoir, Approximations of the Brownian rough path with applications to stochastic analysis, Ann. Inst. H. Poincaré Probab. Statist. 41 (4) (2005) 703–724.
[20] Peter Friz, Nicolas Victoir, On uniformly subelliptic operators and stochastic area, Probab. Theory Related Fields 142 (3–4) (2008) 475–523.
[21] Peter K. Friz, Nicolas B. Victoir, Multidimensional Stochastic Processes as Rough Paths: Theory and Applications, Cambridge Stud. Adv. Math., vol. 120, Cambridge University Press, Cambridge, 2010.
[22] Y. Giga, S. Goto, H. Ishii, M.-H. Sato, Comparison principle and convexity preserving properties for singular degenerate parabolic equations on unbounded domains, Indiana Univ. Math. J. 40 (2) (1991) 443–470.
[23] I. Gyöngy, The stability of stochastic partial differential equations and applications. I, Stoch. Stoch. Rep. 27 (2) (1989) 129–150.
[24] I. Gyöngy, The stability of stochastic partial differential equations and applications. Theorems on supports, in: Stochastic Partial Differential Equations and Applications, II, Trento, 1988, in: Lecture Notes in Math., vol. 1390, Springer, Berlin, 1989, pp. 91–118.
[25] I. Gyöngy, The stability of stochastic partial differential equations. II, Stoch. Stoch. Rep. 27 (3) (1989) 189–233.
[26] I. Gyöngy, The approximation of stochastic partial differential equations and applications in nonlinear filtering, Comput. Math. Appl. 19 (1) (1990) 47–63.
[27] István Gyöngy, On stochastic partial differential equations. Results on approximations, in: Topics in Stochastic Systems: Modelling, Estimation and Adaptive Control, in: Lecture Notes in Control and Inform. Sci., vol. 161, Springer, Berlin, 1991, pp. 116–136.
[28] István Gyöngy, György Michaletzky, On Wong–Zakai approximations with δ-martingales, in: Stochastic Analysis with Applications to Mathematical Finance, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 460 (2041) (2004) 309–324.
[29] István Gyöngy, Anton Shmatkov, Rate of convergence of Wong–Zakai approximations for stochastic partial differential equations, Appl. Math. Optim. 54 (3) (2006) 315–341.
[30] Bogdan Iftimie, Constantin Varsan, A pathwise solution for nonlinear parabolic equations with stochastic perturbations, Cent. Eur. J. Math. 1 (3) (2003) 367–381 (electronic).
[31] M. Kobylanski, Backward stochastic differential equations and partial differential equations with quadratic growth, Ann. Probab. 28 (2) (2000) 558–602.
[32] Hiroshi Kunita, Stochastic Flows and Stochastic Differential Equations, Cambridge Stud. Adv. Math., vol. 24, Cambridge University Press, Cambridge, 1997, reprint of the 1990 original.
[33] M. Ledoux, Z. Qian, T. Zhang, Large deviations and support theorem for diffusion processes via rough paths, Stochastic Process. Appl. 102 (2) (2002) 265–283.
[34] P.-L. Lions, P.E. Souganidis, Viscosity solutions of fully nonlinear stochastic partial differential equations, in: Viscosity Solutions of Differential Equations and Related Topics, Kyoto, 2001, Sūrikaisekikenkyūsho Kōkyūroku 1287 (2002) 58–65 (in Japanese).

[35] Pierre-Louis Lions, Panagiotis E. Souganidis, Fully nonlinear stochastic partial differential equations, C. R. Acad. Sci. Paris Sér. I Math. 326 (9) (1998) 1085–1092.
[36] Pierre-Louis Lions, Panagiotis E. Souganidis, Fully nonlinear stochastic partial differential equations: non-smooth equations and applications, C. R. Acad. Sci. Paris Sér. I Math. 327 (8) (1998) 735–741.
[37] Pierre-Louis Lions, Panagiotis E. Souganidis, Fully nonlinear stochastic PDE with semilinear stochastic dependence, C. R. Acad. Sci. Paris Sér. I Math. 331 (8) (2000) 617–624.
[38] Pierre-Louis Lions, Panagiotis E. Souganidis, Uniqueness of weak solutions of fully nonlinear stochastic partial differential equations, C. R. Acad. Sci. Paris Sér. I Math. 331 (10) (2000) 783–790.
[39] Terry Lyons, Differential equations driven by rough signals, Rev. Mat. Iberoamericana 14 (2) (1998) 215–310.
[40] Terry Lyons, Zhongmin Qian, Flow of diffeomorphisms induced by a geometric multiplicative functional, Probab. Theory Related Fields 112 (1) (1998) 91–119.
[41] Terry Lyons, Zhongmin Qian, System Control and Rough Paths, Oxford Math. Monogr., Oxford University Press, 2002.
[42] Terry J. Lyons, Michael Caruana, Thierry Lévy, Differential equations driven by rough paths, in: 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, in: Lecture Notes in Math., vol. 1908, Springer, Berlin, 2007.
[43] David Nualart, The Malliavin Calculus and Related Topics, Probab. Appl. (N. Y.), second edition, Springer-Verlag, Berlin, 2006.
[44] E. Pardoux, Stochastic partial differential equations and filtering of diffusion processes, Stochastics 3 (2) (1979) 127–167.
[45] Étienne Pardoux, Shi Ge. Peng, Backward doubly stochastic differential equations and systems of quasilinear SPDEs, Probab. Theory Related Fields 98 (2) (1994) 209–227.
[46] B.L. Rozovskiĭ, Evolyutsionnye stokhasticheskie sistemy, in: Lineinaya teoriya i prilozkheniya k statistike sluchainykh protsessov (Linear Theory and Applications to the Statistics of Random Processes), Nauka, Moscow, 1983.
[47] Luciano Tubaro, Some results on stochastic partial differential equations by the stochastic characteristics method, Stoch. Anal. Appl. 6 (2) (1988) 217–230.
[48] Krystyna Twardowska, An approximation theorem of Wong–Zakai type for nonlinear stochastic partial differential equations, Stoch. Anal. Appl. 13 (5) (1995) 601–626.