

Regularity Theory for Rough Partial Differential Equations and Parabolic Comparison Revisited

Joscha Diehl, Peter K. Friz and Harald Oberhauser

Abstract Partial differential equations driven by rough paths are studied. We return to the investigations of [Caruana, Friz and Oberhauser: A (rough) pathwise approach to a class of non- linear SPDEs, *Annales de l’Institut Henri Poincaré/Analyse Non Linéaire* 2011, 28, pp. 27–46], motivated by the Lions–Souganidis theory of viscosity solutions for SPDEs. We continue and complement the previous (uniqueness) results with general existence and regularity statements. Much of this is transformed to questions for deterministic parabolic partial differential equations in viscosity sense. On a technical level, we establish a refined parabolic theorem of sums which may be useful in its own right.

Keywords Rough partial differential equations · Regularity · Existence · Parabolic comparison and theorem of sums

1991 Mathematics Subject Classification 35R99 · 60H15

1 Introduction

In Ref. [5], and then [6], inspired by earlier works of Lions–Souganidis on stochastic viscosity theory [24, 25, 27, 28], rough path stability (with respect to the multi-dimensional driving differential $dz = \dot{z}dt$) was established for the parabolic Cauchy

J. Diehl

Technische Universität Berlin, Berlin, Germany

P.K. Friz (✉)

Technische Universität Berlin and WIAS Berlin, Berlin, Germany

e-mail: friz@math.tu-berlin.de

H. Oberhauser

Oxford-MAN, University of Oxford, Oxford, UK

© Springer International Publishing Switzerland 2014

D. Crisan et al. (eds.), *Stochastic Analysis and Applications 2014*,

Springer Proceedings in Mathematics & Statistics 100,

DOI 10.1007/978-3-319-11292-3_8

problem with spatial domain \mathbb{R}^n ,

$$du = F\left(t, x, Du, D^2u\right) dt - \sum_{i=1}^d Du(t, x) \cdot V_i(x) dz^i, \quad u(0, \cdot) = u_0.$$

Loosely speaking, if a family (z^ϵ) is Cauchy in rough path metric, with rough path limit \mathbf{z} (example: typical mollifications of Brownian motion will satisfy this condition with probability one; the rough path limit is identified as Brownian motion and Lévy's stochastic area; see e.g. [15]), then the resulting PDE solutions of the above problem, (u^ϵ) will converge locally uniformly to a limit which is seen to depend only on \mathbf{z} , and not on the particular sequence (z^ϵ) . In particular, this allows a (rough)pathwise and robust view on stochastic partial differential equations (SPDEs). Immediate (probabilistic) benefits of this approach include support theorems, large deviations and a variety of limit theorems for SPDEs; see [6] for a discussion. Another nice application of robustness in the driving signal is that it quickly leads to *splitting* results for such (rough, and then stochastic) PDEs; see [17]. There are various extensions to noise other than $H(x, Du) = \sum_{i=1}^d Du \cdot V_i(x)$. The “fully non-linear” case $H = H(Du)$, with non-linear dependence on Du , is quite intricate and discussed in Ref. [25], and then [26] with regard to applications; adding x -dependence i.e. $H = H(x, Du)$ is a difficult problem, cf. the forthcoming book by Lions–Souganidis and also forthcoming joint work with P. Gassiat. The case $H = H(x, u)$ is considered in Refs. [12, 27]; the later reference makes a link to backward stochastic differential equations *with rough drivers*. Even the “fully linear” case, with both F and $H = H(x, u, Du)$ linear (in u and its derivatives), is interesting as it covers the Zakai equation from filtering theory (e.g. [1]): robust dependence on \mathbf{z} (the “observation” path) is a classical problem, which goes back to the engineering literature of the late seventies; the rough path point of view has recently led to resolution of this problem; [11, 16].

Having commented on the importance of such classes of (rough) partial differential equations, let us describe the contribution of this paper. We complement the stability result of Ref. [6], which settled uniqueness, with general existence and regularity results, giving conditions for a space-time modulus of continuity (en passant, this justifies regarding solutions as elements in *BUC*-spaces, as is common in viscosity theory) and also spatial Lipschitz regularity. Our conditions are satisfied for large classes of F ; for instance, infima and/or suprema over linear weakly elliptic operators, as is typical in stochastic control resp. differential games. As for the noise term, we have focused on $H(x, Du) = Du \cdot V(x)$, but adaptations to (linear) $H = H(x, u, Du)$ or the setting of Ref. [12] are not difficult.

As a matter of fact, after reduction of the RPDE problem to a (classical) viscosity problem, we are in the need of fairly general (parabolic) comparison results on \mathbb{R}^n . Unfortunately, we failed to find this in the literature which forced us to revisit and adapt some results from viscosity theory. This is a common situation, of course, but since the necessary work seems to go well beyond a routine exercise, and may be of independent interest, we opted to include a reasonably self-contained discussion (Sect. 2) which may be read independently from the rough path/RPDE considerations of Sect. 3.

2 Parabolic Comparison Revisited

2.1 Structural Conditions on F

Let $F = F(t, x, u, p, X) : [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times S^n \rightarrow \mathbb{R}$ be continuous and degenerate elliptic i.e. non-decreasing in X . Here S^n denotes the space of symmetric $n \times n$ matrices. Assume also that there exists γ such that, uniformly in t, x, p, X ,

$$\gamma(u - v) \leq F(t, x, v, p, X) - F(t, x, u, p, X) \text{ whenever } v \leq u. \quad (1)$$

When $\gamma \geq 0$ such F s are called *proper*. Since we will be interested in parabolic problems of the form $\partial_t - F$ a suitable change of variable ($u \leftrightarrow e^{\gamma t} u$) shows that $\gamma < 0$ does not cause trouble. Assume furthermore that there exists, for all $R > 0$, a function $\theta_R : [0, \infty] \rightarrow [0, \infty]$ with $\theta_R(0+) = 0$, such that

$$\begin{aligned} & F(t, x, r, \alpha(x - \tilde{x}), X) - F(t, \tilde{x}, r, \alpha(x - \tilde{x}), Y) \\ & \leq \theta_R \left(\alpha |x - \tilde{x}|^2 + |x - \tilde{x}| \right) \end{aligned} \quad (2)$$

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

$$-3\alpha \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \leq \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq 3\alpha \begin{pmatrix} I & -I \\ -I & I \end{pmatrix}. \quad (3)$$

Under these conditions, comparison for the Cauchy–Dirichlet problem $\partial_t - F = 0$ on $(0, T) \times \Omega$, with Ω bounded, holds (User’s Guide, Chap. 8). We shall be interested in comparison for bounded (semi-)continuous, sub- and super-) solutions on $(0, T] \times \mathbb{R}^n$. In particular, the unboundedness of \mathbb{R}^n leads us to the following additional assumption: assume $F = F(t, x, u, p, X)$ is *uniformly continuous* (UC) whenever u, p, X remain bounded; i.e.

$$\forall R > 0 : F|_{[0, T] \times \mathbb{R}^n \times [-R, R] \times B_R \times M_R} \text{ is uniformly continuous} \quad (4)$$

where B_R, M_R denote (open) balls of radius R in \mathbb{R}^n, S^n respectively.¹ Remark that these structural conditions are satisfied when $F = F(u, p, X)$ is proper (no t, x dependence).

2.2 Statement of Theorems

We write BC, BUC, BUSC, BLSC for bounded continuous, bounded uniformly continuous and bounded upper- resp. lower semi-continuous functions.

¹ Using any of the equivalent norms on S^n .

Theorem 1 (Comparison and spatial regularity) *Assume F satisfies the assumptions of Sect. 2.1. Consider $u \in \text{BUSC}([0, T] \times \mathbb{R}^n)$, $v \in \text{BLSC}([0, T] \times \mathbb{R}^n)$, extended to $[0, T] \times \mathbb{R}^n$ via their semi-continuous envelopes; i.e.²*

$$u(T, x) = \limsup_{\substack{(t, y) \in [0, T] \times \mathbb{R}^n: \\ t \rightarrow T, y \rightarrow x}} u(t, y), \quad v(T, x) = \liminf_{\substack{(t, y) \in [0, T] \times \mathbb{R}^n: \\ t \rightarrow T, y \rightarrow x}} v(t, y). \quad (5)$$

Assume that, in the sense of parabolic viscosity sub- and super-solutions³

$$\partial_t u - F(t, x, u, Du, D^2 u) \leq 0 \leq \partial_t v - F(t, x, v, Dv, D^2 v) \quad \text{on } (0, T) \times \mathbb{R}^n. \quad (6)$$

Then the following statements hold true.

(i) *The validity of (6) extends to $Q := (0, T] \times \mathbb{R}^n$.*

(ii) *If $u_0 := u(0, \cdot)$, $v_0 := v(0, \cdot) \in \text{BUC}(\mathbb{R}^n)$ and $u_0 \leq v_0$ on \mathbb{R}^n one has the “key” estimate*

$$u(t, x) - v(t, y) \leq \inf_{\alpha} \left[\frac{\alpha}{2} |x - y|^2 + l(\alpha) \right],$$

valid for all $(t, x, y) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n$, where $l(\alpha)$ tends to 0 as $\alpha \uparrow \infty$, uniformly in $t \in [0, T]$.

Remark 1 Since $\tilde{u}(t, x) = e^{-\gamma t} u(t, x)$ [resp. $\tilde{v}(t, x) = e^{-\gamma t} v(t, x)$] is a sub- [resp. super-]solution to $(\partial_t - \tilde{F})\tilde{u} + \gamma\tilde{u} = 0$ with

$$\tilde{F}(t, x, p, X) = e^{-\gamma t} F(t, x, e^{\gamma t} \tilde{u}, e^{\gamma t} D\tilde{u}, e^{\gamma t} D^2 \tilde{u})$$

we can always reduce to the case that $\gamma > 0$. In particular, we shall give the proof under this assumption.

Remark 2 The key estimate immediately implies comparison (take $x = y$)

$$u \leq v \text{ on } [0, T] \times \mathbb{R}^n.$$

By a 2ϵ argument, it also yields a spatial modulus for any solution u ; uniform in $t \in [0, T]$. Indeed, for fixed $t \leq T$ pick α large enough so that $l(\alpha) < \epsilon/2$; for any $x, y : |x - y|$ small enough (only depending on α and hence ϵ) we have $u(t, x) - u(t, y) < \epsilon$. By switching the roles of x and y , if necessary, we see $|u(t, x) - u(t, y)| < \epsilon$.

² If one assumes that given $u \in \text{BUSC}([0, T] \times \mathbb{R}^n)$, $v \in \text{BLSC}([0, T] \times \mathbb{R}^n)$ satisfy (6) on $(0, T] \times \mathbb{R}^n$ then it already follows that (5) holds true. This follows from the so-called Accessibility Theorem [7].

³ As is well-known, the precise meaning of (6) is expressed (equivalently) in terms of “touching” test-functions or in term of sub- and super-jets. We shall switch between these points of view without further comments.

Theorem 2 (Spatial Lipschitz regularity) *Assume F satisfies the assumptions of Sect. 2.1 with the strengthening that the modulus θ_R is linear, i.e. $\theta_R(x) = \theta_R x$ for a constant $\theta_R > 0$. Let $u \in BC([0, T] \times \mathbb{R}^n)$ be a solution to*

$$\partial_t u - F(t, x, u, Du, D^2 u) = 0 \text{ on } (0, T] \times \mathbb{R}^n. \quad (7)$$

and assume $u(0, \cdot)$ to be Lipschitz with Lipschitz constant L_{u_0} . Then, for all $t \in (0, T]$, $u(t, \cdot)$ is Lipschitz uniformly in $t \in [0, T]$, with Lipschitz constant $e^{\tilde{\gamma}t}/2 + 2L_{u_0}^2 + \theta_R^2/\gamma$ where $\tilde{\gamma} := 2(\theta_R + 1)$.

Theorem 3 (Time-space regularity) *Assume F satisfies the assumptions of Sect. 2.1, with the strengthening*

$$\forall R > 0 : F|_{[0, T] \times \mathbb{R}^n \times [-R, R] \times B_R \times M_R} \text{ is bounded, uniformly continuous.} \quad (8)$$

(i) Let $u \in BC([0, T] \times \mathbb{R}^n)$ be a viscosity solution to $\partial_t - F = 0$ on $(0, T] \times \mathbb{R}^n$ with initial data $u_0 = u(0, \cdot) \in BUC(\mathbb{R}^n)$. Then

$$u = u(t, x) \in BUC([0, T] \times \mathbb{R}^n).$$

(ii) If, in addition, θ_R is linear and if F has also linear growth in the Hessian, i.e. there exists an $M > 0$ such that

$$|F(t, x, r, p, X)| \leq M(1 + |X|),$$

for all $t \in [0, T]$, $x \in \mathbb{R}^n$, $r \in \mathbb{R}$, $p \in \mathbb{R}^n$ and $X, Y \in S^n$, and $u_0 = u(0, \cdot) \in BC \cap \text{Lip}(\mathbb{R}^n)$ then u is $1/2$ -Hölder in time (uniformly in space) and Lipschitz in space (uniformly in $t \in [0, T]$).

Theorem 4 (Existence) *Assume F satisfies the assumptions of Sect. 2.1, with the strengthening (8), as above. Let $u_0 \in BUC(\mathbb{R}^n)$. Then there exists a unique bounded viscosity solution to the initial value problem*

$$\begin{aligned} \partial_t u - F(t, x, u, Du, D^2 u) &= 0 \text{ on } (0, T] \times \mathbb{R}^n, \\ u(0, \cdot) &= u_0 \end{aligned}$$

and $u = u(t, x) \in BUC([0, T] \times \mathbb{R}^n)$.

2.3 Comments on the Existing Literature

The above ensemble of results gives, under natural conditions, a fairly complete picture of the (model) case of bounded solutions to parabolic problems on \mathbb{R}^n . While it is not harder to think of further generalizations (e.g. unbounded solutions,

discontinuous F , Dirichlet problem on unbounded domains ...), the setting here is appropriate for the study of stochastic viscosity⁴ and rough partial differential equations. Moreover, and this is the *raison d'être* of the present section, one is hard-pressed to find these results in a similar form in the literature.

Let us be specific and point to the closest we are aware of: results in the spirit of Theorem 1, part (ii) and 3, part (i), are found in Ref. [19]. As for Theorem 1, part (ii), related results are found in Ref. [7, 18, 22] (and the references therein). Regularity results relating to Theorem 3, part (ii), are found in Ref. [20]. (The works [20, 21] are based on a useful continuous dependence result whose proof is a variant of the comparison principle proof⁵ If one restricts to the first order case, so that $F = F(t, x, u, Du)$, the theory is less involved and there is a fairly complete text book literature [2, 3]. For instance, the natural analogue of condition 4 in a first order setting appear in Ref. [2], p. 49, as condition (H11) and p. 136, as condition (H23).

On a technical level, the deepest result of this section is an extension of the *Theorem of Sums* (TOS), cf. Sect. 2.9, and it does not appear that one can get uniform continuity in space-time easily without it. To elaborate on this point, recall that almost every modern treatise of second order comparison relies in one way or another on the TOS, also known as *Crandall-Ishii Lemma* [9]. A parabolic version of the TOS on $(0, T) \times \mathbb{R}^n$ then underlies most second order (parabolic) comparison results; such as those in [10, Chap. 8] or [14, Chap. 5]. As is well-known, its application requires a barrier at time T ; e.g. replace a subsolution u by $u^\gamma := u - \gamma/(T - t)$ or so, followed by $\gamma \downarrow 0$ in the end. The downside is that an initially bounded function u is turned into an unbounded function u^γ ; consequently various localizations of the non-linearity are necessary to deal with the resulting unboundedness. (An example of the resulting complication was also seen in Ref. [12].) But then, establishing a spatial modulus of solutions with the (standard) form of the parabolic Theorem of Sums leads to an (a priori) dependence of the spatial modulus in time; establishing the (desired) uniformity in $t \in [0, T]$, cf. (iii) above, would then entail, if possible at all, a painstaking checking of uniformity in γ for all double limits in the technical Lemma 6 below. All these difficulties, as will be seen, are avoided by our extension of the (parabolic) TOS on $Q := (0, T] \times \mathbb{R}^n$.⁶ Since, in fact, the TOS is often used as a “black-box” theorem, in particular in the stochastic control literature such as [14, Chap. 5]), there seems to be every reason to use it in its strongest possible form; in this sense our extension seems to be of independent interest. Concluding this short subsection, it appears that Theorems 1–4, which are otherwise proved by more or less well-known techniques, have not been obtained previously in this form for lack of a suitable TOS, valid at terminal time.

⁴ All results of Lions–Souganidis on stochastic viscosity theory are stated in BUC-spaces with spatial domain \mathbb{R}^n .

⁵ These ideas may prove useful in establishing rates of convergence for equations driven by (z^c) , convergent in rough path sense.

⁶ The point here, of course, is to handle appropriately terminal time T which is a well-documented subtlety in (parabolic) viscosity theory; some mistakes in the early literature were corrected in Ref. [7].

2.4 Proof of Theorem 1: Parabolic Comparison, Part (i)

Proof Assume $u \in \text{BUSC}([0, T] \times \mathbb{R}^n)$ solves $\partial_t u - F(t, x, u, Du, Du) \leq 0$ with “properness” $\gamma \geq 0$; with initial data $u(0, \cdot)$ on $(0, T) \times \mathbb{R}^n$. By assumption, u is extended to $\text{BUSC}([0, T] \times \mathbb{R}^n)$ by setting

$$u(T, x) = \lim_{t \uparrow T, y \rightarrow x} \sup u(t, y).$$

Assume $u - \phi$ has a (strict) max at (T, \bar{x}) , relative to $[0, T] \times \mathbb{R}^n$. (The test function ϕ is defined in an open neighbourhood of $(T, \bar{x}) \in [0, T] \times \mathbb{R}^n$.) In the remainder of this section we establish

$$\partial_t \phi(T, \bar{x}) - F\left(T, \bar{x}, u(T, \bar{x}), D\phi(T, \bar{x}), D^2\phi(T, \bar{x})\right) \leq 0.$$

To this end, start by taking $(t^n, x^n) \in (0, T) \times \mathbb{R}^n$ s.t. $(t^n, x^n) \rightarrow (T, \bar{x})$ and $u(t^n, x^n) \rightarrow u(T, \bar{x})$. Set $\alpha_n := T - t^n \downarrow 0$. Then take

$$(t_n, x_n) \in \arg \max \left(u - \phi - \frac{\alpha_n^2}{T - t} \right) \equiv \arg \max \psi_n.$$

over $[0, T] \times \mathbb{R}^n$. In order to guarantee that the sequence $(t_n, x_n) \in [0, T] \times \mathbb{R}^n$ remains in a compact, say $[T/2, T] \times \bar{B}_1(\bar{x})$, we make the assumption (without loss of generality) that $\phi(T, \bar{x}) = 0$ and $\phi(t, x) > 3|u|_\infty$ for $(t, x) \notin [T/2, T] \times \bar{B}_1(\bar{x})$; this implies $(t_n, x_n) \in [T/2, T] \times \bar{B}_1(\bar{x})$ for n large enough, as desired. By compactness, $(t_n, x_n) \rightarrow (\tilde{t}, \tilde{x})$ at least along a subsequence $n(k)$. We shall run through the other sequence (t^n, x^n) along the same subsequence and relabel both to keep the same notation. Note $\psi_n(t_n, x_n)$ is non-decreasing and bounded, hence $\psi_n(t_n, x_n) \rightarrow l$. Since $\psi_n(t_n, x_n) \leq (u - \phi)(t_n, x_n)$ it follows (using USC of $u - \phi$) that $l \leq (u - \phi)(\tilde{t}, \tilde{x})$. On the other hand,

$$\psi_n(t_n, x_n) \geq \psi_n(t^n, x^n) = (u - \phi)(t^n, x^n) - \underbrace{\frac{\alpha_n^2}{T - t^n}}_{=\alpha_n}$$

and hence $l \geq (u - \phi)(T, \bar{x})$. Since (T, \bar{x}) was a strict maximum point for $u - \phi$ conclude that $(\tilde{t}, \tilde{x}) = (T, \bar{x})$ is the common limit of the sequences (t^n, x^n) , (t_n, x_n) . Now we note that

$$(u - \phi)(t_n, x_n) \geq \psi_n(t_n, x_n) \geq (u - \phi)(t^n, x^n) - \alpha_n$$

which implies that $u(t_n, x_n) \geq u(t^n, x^n) + o(1)$ where $o(1) \rightarrow 0$ as $n \rightarrow \infty$. By definition of a subsolution,

$$\partial_t \phi(t_n, x_n) - F\left(t_n, x_n, u(t_n, x_n), D\phi(t_n, x_n), D^2\phi(t_n, x_n)\right) \leq 0$$

and hence, using properness of F (more specifically, (1) applied with $u = u(t_n, x_n)$ and $v = u(t^n, x^n) + o(1)$),

$$\begin{aligned} -F(u(t_n, x_n)) &\geq -F(u(t^n, x^n) + o(1)) + \gamma(u(t_n, x_n) - (u(t^n, x^n) + o(1))) \\ &\geq -F(u(t^n, x^n)) + o(1), \end{aligned}$$

also using uniform continuity of F as function of u over compacts, we obtain

$$\partial_t \phi(t_n, x_n) - F(t_n, x_n, u(t^n, x^n), D\phi(t_n, x_n), D^2\phi(t_n, x_n)) \leq o(1).$$

Sending $n \rightarrow \infty$ yields (use continuity of ϕ and F)

$$\partial_t \phi(T, \bar{x}) - F(T, \bar{x}, u(T, \bar{x}), D\phi(T, \bar{x}), D^2\phi(T, \bar{x})) \leq 0,$$

as desired. \square

2.5 Proof of Theorem 1: Parabolic Comparison, Part (ii)

Proof By assumption, $u(t, x) - v(t, y)$ is bounded on $[0, T] \times \mathbb{R}^n \times \mathbb{R}^n$. Let $(\hat{t}, \hat{x}, \hat{y})$ be a maximum point of

$$\phi(t, x, y) := u(t, x) - v(t, y) - \frac{\alpha}{2} |x - y|^2 - \varepsilon (|x|^2 + |y|^2) \quad (9)$$

over $[0, T] \times \mathbb{R}^n \times \mathbb{R}^n$ where $\alpha > 0$ and $\varepsilon > 0$; such a maximum exists since $\phi \in \text{USC}([0, T] \times \mathbb{R}^n \times \mathbb{R}^n)$ and $\phi \rightarrow -\infty$ as $|x|, |y| \rightarrow \infty$. (The presence $\varepsilon > 0$ amounts to a barrier at ∞ in space). The plan is to show a “key estimate” of the form

$$u(t, x) - v(t, y) \leq \inf_{\alpha} \left[\frac{\alpha}{2} |x - y|^2 + l(\alpha) \right], \quad (10)$$

valid on $[0, T] \times \mathbb{R}^n \times \mathbb{R}^n$, where $l(\alpha)$ tends to 0 as $\alpha \uparrow \infty$. Thanks to the very definition of $(\hat{t}, \hat{x}, \hat{y})$ as $\arg \max$ of $\phi = \phi(t, x, y)$, we obtain the estimate

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

Note that $(\hat{t}, \hat{x}, \hat{y})$ depends on α, ε . We shall consider the cases $\hat{t} = 0$ and $\hat{t} \in (0, T]$ separately. In the first case $\hat{t} = 0$ we have

$$\phi(0, \hat{x}, \hat{y}) = \sup_{x, y} \left[u_0(x) - v_0(y) - \frac{\alpha}{2} |x - y|^2 - \varepsilon (|x|^2 + |y|^2) \right] =: A_{\alpha, \varepsilon}$$

and Lemma 5 below asserts that $A_{\alpha,\varepsilon} \rightarrow \sup_x [u_0(x) - v_0(x)] \leq 0$ as $(\varepsilon, \alpha) \rightarrow (0, \infty)$. The second case is $\hat{t} \in (0, T]$ and we will show

$$\phi(\hat{t}, \hat{x}, \hat{y}) \leq B_{\alpha,\varepsilon} \text{ where } \left(\limsup_{\varepsilon \rightarrow 0} B_{\alpha,\varepsilon} \right) \rightarrow 0 \text{ as } \alpha \rightarrow \infty; \quad (11)$$

it is here that we will use Theorem of Sums and viscosity properties. (Since

$$\phi(\hat{t}, \hat{x}, \hat{y}) \leq u(\hat{t}, \hat{x}) - v(\hat{t}, \hat{y})$$

we can and will use the fact that it is enough to consider the case $u(\hat{t}, \hat{x}) - v(\hat{t}, \hat{y}) \geq 0$.) Leaving the details of this to below, let us quickly complete the argument: our discussion of the two cases above gives $\phi(\hat{t}, \hat{x}, \hat{y}) \leq A_{\alpha,\varepsilon} \vee B_{\alpha,\varepsilon}$ and hence

$$u(t, x) - v(t, y) \leq \frac{\alpha}{2} |x - y|^2 + \varepsilon (|x|^2 + |y|^2) + A_{\alpha,\varepsilon} \vee B_{\alpha,\varepsilon};$$

we emphasize that this estimate is valid for all $t, x, y \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n$ and $\alpha, \varepsilon > 0$. Take now $\limsup_{\varepsilon \rightarrow 0}$ on the right hand side, then optimize over $\alpha > 0$, to obtain the key estimate

$$u(t, x) - v(t, y) \leq \inf_{\alpha} \left\{ \frac{\alpha}{2} |x - y|^2 + l(\alpha) \right\}$$

where we may take $l(\alpha) := \limsup_{\varepsilon \rightarrow 0} A_{\alpha,\varepsilon} \vee \limsup_{\varepsilon \rightarrow 0} B_{\alpha,\varepsilon}$, noting that $l(\alpha)$ indeed tends to 0 as $\alpha \rightarrow \infty$. It remains to prove the estimate (11). To this end, rewrite ϕ 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 as given in the appendix at $(\hat{t}, \hat{x}, \hat{y}) \in (0, T] \times \mathbb{R}^n \times \mathbb{R}^n$ to learn that there are numbers a, b and $X, Y \in S^n$ such that

$$(a, \alpha(\hat{x} - \hat{y}), X) \in \bar{\mathcal{P}}_Q^{2,+} u^\varepsilon(\hat{t}, \hat{x}), \quad (b, \alpha(\hat{x} - \hat{y}), Y) \in \bar{\mathcal{P}}_Q^{2,-} v^\varepsilon(\hat{t}, \hat{y}) \quad (12)$$

such that $a - b \geq 0$ (equality if $\hat{t} \in (0, T)$, although this does not matter), and such that one has the two-sided matrix estimate (3). It is easy to see (cf. [10, Remark 2.7]) that (12) is equivalent to

$$\begin{aligned} (a, \alpha(\hat{x} - \hat{y}) + 2\varepsilon\hat{x}, X + 2\varepsilon I) &\in \bar{\mathcal{P}}_Q^{2,+} u(\hat{t}, \hat{x}), \\ (b, \alpha(\hat{x} - \hat{y}) - 2\varepsilon\hat{y}, Y - 2\varepsilon I) &\in \bar{\mathcal{P}}_Q^{2,-} v(\hat{t}, \hat{y}). \end{aligned}$$

Using the viscosity sub- and super-solution properties (and part (i) in the case that $\hat{t} = T$) we then see that

$$\begin{aligned} a - F(\hat{t}, \hat{x}, u(\hat{t}, \hat{x}), \alpha(\hat{x} - \hat{y}) + 2\varepsilon\hat{x}, X + 2\varepsilon I) &\leq 0, \\ b - F(\hat{t}, \hat{y}, v(\hat{t}, \hat{y}), \alpha(\hat{x} - \hat{y}) - 2\varepsilon\hat{y}, Y - 2\varepsilon I) &\geq 0. \end{aligned}$$

Note that (using $a - b \geq 0$)

$$\begin{aligned} &F(\hat{t}, \hat{y}, v(\hat{t}, \hat{y}), \alpha(\hat{x} - \hat{y}) - 2\varepsilon\hat{y}, Y - 2\varepsilon I) \\ &\quad - F(\hat{t}, \hat{x}, u(\hat{t}, \hat{x}), \alpha(\hat{x} - \hat{y}) + 2\varepsilon\hat{x}, X + 2\varepsilon I) \leq 0 \end{aligned} \quad (13)$$

Trivially, (recall it is enough to consider the case $u(\hat{t}, \hat{x}) \geq v(\hat{t}, \hat{y})$)

$$\begin{aligned} \gamma\phi(\hat{t}, \hat{x}, \hat{y}) &\leq \gamma(u(\hat{t}, \hat{x}) - v(\hat{t}, \hat{y})) \\ &\leq F(\hat{t}, \hat{y}, v(\hat{t}, \hat{y}), \alpha(\hat{x} - \hat{y}) - 2\varepsilon\hat{y}, Y - 2\varepsilon I) \\ &\quad - F(\hat{t}, \hat{y}, u(\hat{t}, \hat{x}), \alpha(\hat{x} - \hat{y}) - 2\varepsilon\hat{y}, Y - 2\varepsilon I) \\ &\leq F(\hat{t}, \hat{x}, u(\hat{t}, \hat{x}), \alpha(\hat{x} - \hat{y}) + 2\varepsilon\hat{x}, X + 2\varepsilon I) \\ &\quad - F(\hat{t}, \hat{y}, u(\hat{t}, \hat{x}), \alpha(\hat{x} - \hat{y}) - 2\varepsilon\hat{y}, Y - 2\varepsilon I) \end{aligned}$$

where we used (13) in the last estimate. If ε were absent (e.g. set $\varepsilon = 0$ throughout) we would estimate, with $R := |u|_\infty \vee |v|_\infty$,

$$\begin{aligned} &F(\hat{t}, \hat{x}, u(\hat{t}, \hat{x}), \alpha(\hat{x} - \hat{y}), X) - F(\hat{t}, \hat{y}, u(\hat{t}, \hat{x}), \alpha(\hat{x} - \hat{y}), Y) \\ &\leq \theta_R \left(\alpha |\hat{x} - \hat{y}|^2 + |\hat{x} - \hat{y}| \right) =: B_\alpha \end{aligned}$$

and since $\alpha |\hat{x} - \hat{y}|^2 + |\hat{x} - \hat{y}| \leq 2\alpha |\hat{x} - \hat{y}|^2 + 1/\alpha \rightarrow 0$ as $\alpha \rightarrow \infty$, thanks to [10, Lemma 3.1], we see that $B_\alpha \rightarrow 0$ with $\alpha \rightarrow \infty$, which is enough to conclude. The present case where $\varepsilon > 0$ is essentially reduced to the case $\varepsilon = 0$ by adding/subtracting

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

but we need some refined properties of $(\hat{t}, \hat{x}, \hat{y})$ as collected in Lemma 6: (a) $p = \alpha(\hat{x} - \hat{y})$ remains, for fixed α , bounded as $\varepsilon \rightarrow 0$, (b) $2\varepsilon|\hat{x}|$ and $2\varepsilon|\hat{y}|$ tend to zero as $\varepsilon \rightarrow 0$ for fixed (large enough) α ; this follows from the fact, that for α large enough we must have $\limsup_{\varepsilon \rightarrow 0} \varepsilon|\hat{x}|^2 = c_\alpha < \infty$ (after all, c_α tends to zero with $\alpha \rightarrow \infty$) and by rewriting $\limsup_{\varepsilon \rightarrow 0} \varepsilon|\hat{x}| \leq \sqrt{c_\alpha} \limsup_{\varepsilon \rightarrow 0} \sqrt{\varepsilon} = 0$, (c) that $\limsup_{\varepsilon \rightarrow 0} \left(\frac{\alpha}{2} |\hat{x} - \hat{y}|^2 + |\hat{x} - \hat{y}| \right) \rightarrow 0$ as $\alpha \rightarrow \infty$. We also note that (3) implies (d): any matrix norm of X, Y is bounded by a constant times α , independent of ε .

We can now return to the estimate of ϕ and clearly have

$$\phi(\hat{t}, \hat{x}, \hat{y}) \leq \frac{1}{\gamma} [(i) + (ii) + (iii)] =: B_{\alpha, \varepsilon}$$

where

$$\begin{aligned} (i) &= |F(\hat{t}, \hat{x}, u(\hat{t}, \hat{x}), \alpha(\hat{x} - \hat{y}) + 2\varepsilon\hat{x}, X + 2\varepsilon I) - F(\hat{t}, \hat{x}, u(\hat{t}, \hat{x}), \alpha(\hat{x} - \hat{y}), X)| \\ (ii) &= |F(\hat{t}, \hat{y}, u(\hat{t}, \hat{x}), \alpha(\hat{x} - \hat{y}) - 2\varepsilon\hat{y}, Y - 2\varepsilon I) - F(\hat{t}, \hat{y}, u(\hat{t}, \hat{x}), \alpha(\hat{x} - \hat{y}), Y)| \\ (iii) &= \theta_R \left(\alpha |\hat{x} - \hat{y}|^2 + |\hat{x} - \hat{y}| \right). \end{aligned}$$

From (a), (d) above the gradient and Hessian argument in F as seen in (i), (ii), i.e.

$$\alpha(\hat{x} - \hat{y}) \pm 2\varepsilon\hat{x} \text{ and } X + 2\varepsilon I, Y - 2\varepsilon I,$$

remain in a bounded set, for fixed α , uniformly as $\varepsilon \rightarrow 0$. From (b) above and the assumed uniform continuity properties of F (i.e. (4)), it then follows that for fixed (large enough) α

$$(i), (ii) \rightarrow 0 \text{ as } \varepsilon \rightarrow 0.$$

On the other hand, continuity of θ_R at 0+ together with (c) above shows that also $(iii) \rightarrow 0$ as $\varepsilon \ll \frac{1}{\alpha} \rightarrow 0$. We conclude that

$$B_{\alpha, \varepsilon} \rightarrow 0 \text{ as } \varepsilon \ll \frac{1}{\alpha} \rightarrow 0,$$

which implies (11), as desired. The proof is now finished. \square

Lemma 5 Assume $u_0, v_0 \in \text{BUC}(\mathbb{R}^n)$. Then

$$\begin{aligned} &\sup_{x, y} \left[u_0(x) - v_0(y) - \frac{\alpha}{2} |x - y|^2 - \varepsilon (|x|^2 + |y|^2) \right] \\ &\rightarrow \sup_x [u_0(x) - v_0(x)] \text{ as } (\varepsilon, \frac{1}{\alpha}) \rightarrow (0, 0). \end{aligned}$$

Proof (Similar to [2, Lemme 2.9] but we include full details for the reader's convenience.)

Without loss of generality $M := \sup_x [u_0(x) - v_0(x)] > 0$; for otherwise replace u_0 by $u_0 + 2|M|$. Write $M_{\alpha, \varepsilon}$ for the achieved maximum (at \hat{x}, \hat{y} , say) of the left-hand-side. Obviously, $u_0(x) - v_0(x) - 2\varepsilon|x|^2 \leq M_{\alpha, \varepsilon}$ for any x and so

$$M \leq \lim_{\substack{\varepsilon \rightarrow 0 \\ \alpha \rightarrow \infty}} \inf M_{\alpha, \varepsilon}.$$

It follows that we can and will consider ε (α) small (large) enough so that $M_{\alpha,\varepsilon} > 0$. On the other hand, $|u_0|, |v_0| \leq R < \infty$ and so

$$0 \leq M_{\alpha,\varepsilon} \leq 2R - \frac{\alpha}{2} |\hat{x} - \hat{y}|^2 - \varepsilon (|\hat{x}|^2 + |\hat{y}|^2)$$

from which we deduce $\frac{\alpha}{2} |\hat{x} - \hat{y}|^2 \leq 2R$, or $|\hat{x} - \hat{y}| \leq \sqrt{4R/\alpha}$. By omitting the (positive) penalty terms, we can also estimate

$$\begin{aligned} M_{\alpha,\varepsilon} &\leq u_0(\hat{x}) - v_0(\hat{y}) \\ &\leq u_0(\hat{x}) - v_0(\hat{x}) + \sigma_{v_0}(\sqrt{4R/\alpha}) \\ &\leq M + \sigma_{v_0}(\sqrt{4R/\alpha}) \end{aligned}$$

where σ_{v_0} denotes the modulus of continuity of v_0 . It follows that

$$\lim_{\substack{\varepsilon \rightarrow 0 \\ \alpha \rightarrow \infty}} \sup M_{\alpha,\varepsilon} \leq M$$

which shows that the $\lim M_{\alpha,\varepsilon}$ (as $\varepsilon \rightarrow 0, \alpha \rightarrow \infty$) exists and is equal to M . \square

Lemma 6 *Let $u \in \text{BUSC}([0, T] \times \mathbb{R}^n)$ and $v \in \text{BLSC}([0, T] \times \mathbb{R}^n)$. Consider a maximum point $(\hat{t}, \hat{x}, \hat{y}) \in (0, T] \times \mathbb{R}^n \times \mathbb{R}^n$ of*

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

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

$$\lim_{\varepsilon \rightarrow 0} \sup \alpha (\hat{x} - \hat{y}) = C(\alpha) < \infty, \quad (14)$$

$$\lim_{\alpha \rightarrow \infty} \sup \lim_{\varepsilon \rightarrow 0} \sup \varepsilon (|\hat{x}|^2 + |\hat{y}|^2) = 0, \quad (15)$$

$$\lim_{\alpha \rightarrow \infty} \sup \lim_{\varepsilon \rightarrow 0} \sup \left(\frac{\alpha}{2} |\hat{x} - \hat{y}|^2 + |\hat{x} - \hat{y}| \right) = 0. \quad (16)$$

The order in which limits are taken is important and suggests the notation

$$\lim_{\varepsilon < \frac{1}{\alpha} \rightarrow 0} \sup (...) := \lim_{\alpha \rightarrow \infty} \sup \lim_{\varepsilon \rightarrow 0} \sup (...), \quad \lim_{\varepsilon < \frac{1}{\alpha} \rightarrow 0} \inf (...) := \lim_{\alpha \rightarrow \infty} \inf \lim_{\varepsilon \rightarrow 0} \inf (...).$$

Proof A similar Lemma (without t dependence) is found in Barles' book [2, Lemme 4.3]. Again, we include details for the reader's convenience.

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

$$\begin{aligned} M_{\alpha, \varepsilon} &:= \sup_{t, x, y} \phi(t, x, y) = u(\hat{t}, \hat{x}) - v(\hat{t}, \hat{y}) - \frac{\alpha}{2} |\hat{x} - \hat{y}|^2 - \varepsilon \left(|\hat{x}|^2 + |\hat{y}|^2 \right); \\ M(h) &:= \sup_{t, x, y: |x-y| \leq h} [u(t, x) - v(t, y)] \geq \sup_{t, x} [u(t, x) - v(t, x)] \\ M' &:= \downarrow \lim_{h \rightarrow 0} M(h), \end{aligned}$$

where 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

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

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

$$\frac{\alpha}{2} |\hat{x} - \hat{y}|^2 + \varepsilon \left(|\hat{x}|^2 + |\hat{y}|^2 \right) \leq C^2$$

which implies

$$|\hat{x} - \hat{y}| \leq C \sqrt{2/\alpha} \quad (17)$$

and hence $\alpha |\hat{x} - \hat{y}| \leq \sqrt{2\alpha} C$ which is the first claimed estimate (14).

Step 2: We first argue that it is enough to show the two estimates

$$\limsup_{\varepsilon \ll \frac{1}{\alpha} \rightarrow 0} [u(\hat{t}, \hat{x}) - v(\hat{t}, \hat{y})] \leq M' \leq \liminf_{\varepsilon \ll \frac{1}{\alpha} \rightarrow 0} M_{\alpha, \varepsilon}. \quad (18)$$

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

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

This, together with (17), already gives (15) and also (16).

We are left to show (18). For the first estimate, it suffices to note that, from (17) and the definition of $M(h)$ applied with $h = C\sqrt{2/\alpha}$,

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

We now turn to the second estimate in (18). From the very definition of M' as $\lim_{h \rightarrow 0} M(h)$, there exists a family (t_h, x_h, y_h) so that

$$|x_h - y_h| \leq h \text{ and } u(t_h, x_h) - v(t_h, y_h) \rightarrow M' \text{ as } h \rightarrow 0 \quad (19)$$

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

$$u(t_h, x_h) - v(t_h, y_h) - \frac{\alpha}{2}h^2 - \varepsilon(|x_h|^2 + |y_h|^2) \leq M_{\alpha, \varepsilon} \quad (20)$$

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

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

Since this is valid for every α , we also have

$$M' \leq \liminf_{\alpha \rightarrow \infty} \liminf_{\varepsilon \rightarrow 0} M_{\alpha, \varepsilon}.$$

This is precisely the second estimate in (18) and so the proof is finished. \square

2.6 Proof of Theorem 2: Lipschitz Regularity in Space

Remark 3 The argument will be similar to the regularity proof in [20]. Like there, it can be adapted to show Hölder continuity in x for Hölder initial data and a corresponding condition on the modulus θ_R .

Proof We modify the proof of Theorem 1 (ii) by adding a specific t -dependence to the test function. This idea can for example also be found in Ref. [20]. Since the proof is similar to the proof of Theorem 1 (ii) we will omit redundant arguments.

Let $(\hat{t}, \hat{x}, \hat{y})$ be a maximum point of

$$\phi(t, x, y) := u(t, x) - u(t, y) - \frac{\alpha}{2} e^{\tilde{\gamma}t} |x - y|^2 - \varepsilon (|x|^2 + |y|^2) \quad (21)$$

over $[0, T] \times \mathbb{R}^n \times \mathbb{R}^n$ where $\alpha > 0$ and $\varepsilon > 0$. The plan is to show the estimate

$$u(t, x) - u(t, y) \leq \inf_{\alpha} \left[e^{\tilde{\gamma}t} \frac{\alpha}{2} |x - y|^2 + (2L_{u_0}^2 + \frac{\theta_R^2}{\gamma}) \frac{1}{\alpha} \right], \quad (22)$$

valid on $[0, T] \times \mathbb{R}^n \times \mathbb{R}^n$. It then immediately follows, that

$$u(t, x) - u(t, y) \leq (e^{\tilde{\gamma}t} \frac{1}{2} + 2L_{u_0}^2 + \frac{\theta_R^2}{\gamma}) |x - y|.$$

Thanks to the very definition of $(\hat{t}, \hat{x}, \hat{y})$ as $\arg \max$ of $\phi(t, x, y) = u(t, x) - u(t, y) - \frac{\alpha}{2} |x - y|^2 - \varepsilon (|x|^2 + |y|^2)$, we obtain the estimate

$$u(t, x) - u(t, y) \leq \frac{\alpha}{2} e^{\tilde{\gamma}t} |x - y|^2 + \varepsilon (|x|^2 + |y|^2) + \phi(\hat{t}, \hat{x}, \hat{y}).$$

Note that $(\hat{t}, \hat{x}, \hat{y})$ depends on α, ε . We shall consider the cases $\hat{t} = 0$ and $\hat{t} \in (0, T]$ separately. In the first case $\hat{t} = 0$ we have

$$\phi(0, \hat{x}, \hat{y}) = \sup_{x, y} \left[u_0(x) - u_0(y) - \frac{\alpha}{2} |x - y|^2 - \varepsilon (|x|^2 + |y|^2) \right] =: A_{\alpha, \varepsilon}$$

and Lemma 7 below asserts that

$$\limsup_{\varepsilon \rightarrow 0} A_{\alpha, \varepsilon} \leq 2L_{u_0}^2 \frac{1}{\alpha}. \quad (23)$$

The second case is $\hat{t} \in (0, T]$ and we will show

$$\phi(\hat{t}, \hat{x}, \hat{y}) \leq B_{\alpha, \varepsilon} \text{ where } \left(\limsup_{\varepsilon \rightarrow 0} B_{\alpha, \varepsilon} \right) \leq \frac{\theta_R^2}{\gamma} \frac{1}{\alpha}; \quad (24)$$

it is here that we will use Theorem of Sums and viscosity properties. From (23) and (24) we can then deduce (22) just as in the proof of Theorem (1) (ii).

It remains to prove the estimate (24). To this end, as before, we can deduce the existence of

$$\begin{aligned} (a, \alpha e^{\tilde{\gamma}\hat{t}} (\hat{x} - \hat{y}) + 2\varepsilon\hat{x}, X + 2\varepsilon I) &\in \bar{\mathcal{P}}_Q^{2,+} u(\hat{t}, \hat{x}), \\ (b, \alpha e^{\tilde{\gamma}\hat{t}} (\hat{x} - \hat{y}) - 2\varepsilon\hat{y}, Y - 2\varepsilon I) &\in \bar{\mathcal{P}}_Q^{2,-} u(\hat{t}, \hat{y}), \end{aligned}$$

with the difference, that in this case $a - b \geq \frac{\alpha}{2} \bar{\gamma} e^{\bar{\gamma} \hat{t}} |\hat{x} - \hat{y}|^2$ (since the time derivative of the test function does not vanish) and such that one has the two-sided matrix estimate

$$-3\alpha e^{\bar{\gamma} \hat{t}} \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \leq \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq 3\alpha e^{\bar{\gamma} \hat{t}} \begin{pmatrix} I & -I \\ -I & I \end{pmatrix}. \quad (25)$$

Using the viscosity sub- and super-solution properties we then see that

$$\begin{aligned} & F(\hat{t}, \hat{y}, u(\hat{t}, \hat{y}), \alpha e^{\bar{\gamma} \hat{t}} (\hat{x} - \hat{y}) - 2\varepsilon \hat{y}, Y - 2\varepsilon I) \\ & - F(\hat{t}, \hat{x}, u(\hat{t}, \hat{x}), \alpha e^{\bar{\gamma} \hat{t}} (\hat{x} - \hat{y}) + 2\varepsilon \hat{x}, X + 2\varepsilon I) \leq -\frac{\alpha}{2} \bar{\gamma} e^{\bar{\gamma} \hat{t}} |\hat{x} - \hat{y}|^2. \end{aligned} \quad (26)$$

Trivially, (recall it is enough to consider the case $u(\hat{t}, \hat{x}) \geq u(\hat{t}, \hat{y})$)

$$\begin{aligned} \gamma \phi(\hat{t}, \hat{x}, \hat{y}) & \leq \gamma (u(\hat{t}, \hat{x}) - u(\hat{t}, \hat{y})) \\ & \leq F(\hat{t}, \hat{y}, u(\hat{t}, \hat{y}), \alpha e^{\bar{\gamma} \hat{t}} (\hat{x} - \hat{y}) - 2\varepsilon \hat{y}, Y - 2\varepsilon I) \\ & \quad - F(\hat{t}, \hat{y}, u(\hat{t}, \hat{x}), \alpha e^{\bar{\gamma} \hat{t}} (\hat{x} - \hat{y}) - 2\varepsilon \hat{y}, Y - 2\varepsilon I) \\ & \leq F(\hat{t}, \hat{x}, u(\hat{t}, \hat{x}), \alpha e^{\bar{\gamma} \hat{t}} (\hat{x} - \hat{y}) + 2\varepsilon \hat{x}, X + 2\varepsilon I) \\ & \quad - F(\hat{t}, \hat{y}, u(\hat{t}, \hat{x}), \alpha e^{\bar{\gamma} \hat{t}} (\hat{x} - \hat{y}) - 2\varepsilon \hat{y}, Y - 2\varepsilon I) - \frac{\alpha}{2} \bar{\gamma} e^{\bar{\gamma} \hat{t}} |\hat{x} - \hat{y}|^2 \end{aligned}$$

where we used (26) in the last estimate. So

$$\phi(\hat{t}, \hat{x}, \hat{y}) \leq \frac{1}{\gamma} [(i) + (ii) + (iii)] =: B_{\alpha, \varepsilon}$$

where (remember the choice of $\bar{\gamma} = 2(\theta_R + 1)$)

$$\begin{aligned} (i) &= |F(\hat{t}, \hat{x}, u(\hat{t}, \hat{x}), \alpha (\hat{x} - \hat{y}) + 2\varepsilon \hat{x}, X + 2\varepsilon I) - F(\hat{t}, \hat{x}, u(\hat{t}, \hat{x}), \alpha (\hat{x} - \hat{y}), X)| \\ (ii) &= |F(\hat{t}, \hat{y}, u(\hat{t}, \hat{x}), \alpha (\hat{x} - \hat{y}) - 2\varepsilon \hat{y}, Y - 2\varepsilon I) - F(\hat{t}, \hat{y}, u(\hat{t}, \hat{x}), \alpha (\hat{x} - \hat{y}), Y)| \\ (iii) &= \theta_R (\alpha |\hat{x} - \hat{y}|^2 + |\hat{x} - \hat{y}|) - \frac{\alpha}{2} \bar{\gamma} e^{\bar{\gamma} \hat{t}} |\hat{x} - \hat{y}|^2 \\ &= \theta_R |\hat{x} - \hat{y}| + \alpha \left(\theta_R - \frac{1}{2} \bar{\gamma} e^{\bar{\gamma} \hat{t}} \right) |\hat{x} - \hat{y}|^2 \\ &\leq \theta_R |\hat{x} - \hat{y}| - \alpha |\hat{x} - \hat{y}|^2 \\ &\leq \theta_R^2 \frac{1}{\alpha}. \end{aligned}$$

The last inequality follows from $\sup_r [c_1 r - c_2 r^2] \leq c_1^2 c_2^{-1}$, for $c_1, c_2 > 0$. As before we deduce for fixed (large enough) α that (i), (ii) $\rightarrow 0$ as $\varepsilon \rightarrow 0$. Hence

$$\limsup_{\varepsilon \rightarrow 0} B_{\alpha, \varepsilon} \leq \frac{\theta_R^2}{\gamma} \frac{1}{\alpha}. \quad \square$$

Lemma 7 Assume u_0 is bounded and Lipschitz with Lipschitz constant L_{u_0} . Then

$$\limsup_{\varepsilon \rightarrow 0} \sup_{x, y} \left[u_0(x) - u_0(y) - \frac{\alpha}{2} |x - y|^2 - \varepsilon (|x|^2 + |y|^2) \right] \leq 2L_{u_0}^2 \frac{1}{\alpha}.$$

Proof Write $M_{\alpha, \varepsilon}$ for the achieved maximum (at \hat{x}, \hat{y} , say) of the left-hand-side (for fixed α and ε). By comparing to evaluation at $x = y = 0$ we have

$$0 \leq u_0(\hat{x}) - u_0(\hat{y}) - \frac{\alpha}{2} |\hat{x} - \hat{y}|^2 - \varepsilon (|\hat{x}|^2 + |\hat{y}|^2)$$

from which we deduce $\frac{\alpha}{2} |\hat{x} - \hat{y}|^2 - L_{u_0} |\hat{x} - \hat{y}| \leq 0$. Hence $|\hat{x} - \hat{y}| \leq 2 \frac{L_{u_0}}{\alpha}$. By omitting the (positive) penalty terms, we can also estimate

$$M_{\alpha, \varepsilon} \leq u_0(\hat{x}) - u_0(\hat{y}) \leq L_{u_0} |\hat{x} - \hat{y}| \leq 2 \frac{L_{u_0}^2}{\alpha}.$$

It follows that $\limsup_{\varepsilon \rightarrow 0} M_{\alpha, \varepsilon} \leq 2 \frac{L_{u_0}^2}{\alpha}$. □

2.7 Proof of Theorem 3: Regularity in Time-Space

Remark 4 The linear growth condition in (ii) is especially satisfied for linear problems. Even in this case the Hölder regularity cannot be improved, as maybe seen in the standard heat equation started with initial data $u_0(x) = \max\{|x|, 1\}$.

Proof (i): We adapt the argument from [4, Lemma 9.1]; see also [3]. From Theorem 1, there exists a spatial modulus m for $u(t, \cdot)$, uniform over $t \in [0, T]$. Given $0 \leq t_0 < t \leq T$ and $x_0, x \in \mathbb{R}^n$ we now estimate, using the triangle inequality,

$$|u(t, x) - u(t_0, x_0)| \leq m(|x_0 - x|) + |u(t, x_0) - u(t_0, x_0)|.$$

We shall show that $|u(t, x_0) - u(t_0, x_0)|$ goes to zero as $t \downarrow t_0$, uniformly in $x_0 \in \mathbb{R}^n$ and $t_0 \in [0, T]$. We will show a little more. Fix $x_0 \in \mathbb{R}^n$ and $R \in (0, \infty)$; for instance $R = 1$ would do (and there is no need to track dependence in R). We claim that for every $\eta > 0$ one can find constants $C = C(\eta)$, $K = K(\eta)$, not dependent on x_0 and t_0 , such that, for all $x \in B_{R/2}(x_0)$ and $y \in B_R(x_0)$ and all $t \in [t_0, T]$

$$u(t, y) - u(t_0, x) \leq \eta + C|y - x|^2 + K(t - t_0) \quad (27)$$

and

$$u(t, y) - u(t_0, x) \geq -\eta - C|y - x|^2 - K(t - t_0). \quad (28)$$

(Choosing $x = y = x_0$ in these estimates shows that $|u(t, x_0) - u(t_0, x_0)| \leq \inf \{\eta + K(\eta)(t - t_0) : \eta > 0\}$ which immediately gives the desired uniform continuity in time, uniformly in x_0 .) We only prove (27), (28) being proved in an analogous way. In the sequel, x is fixed in $B_{R/2}(x_0)$. Rewrite (27) as

$$u - \chi \leq 0 \text{ on } [t_0, T] \times B_R(x_0)$$

where $\chi(t, y) := u(t_0, x) + \eta + C|y - x|^2 + K(t - t_0)$. We shall see below we can find C , the choice of which only depends on η (and in a harmless way on $|u|_{\infty; [0, T] \times \mathbb{R}^n}$, R and $m(\cdot)$ but not on K and not on x_0, t_0), such that $u - \chi \leq 0$ on the parabolic boundary of $[t_0, T] \times B_R(x_0)$. The extension to the interior is then based on the maximum principle. More precisely, we can choose K depending on η (and again in a harmless way $|u|_{\infty; [0, T] \times \mathbb{R}^n}$, R and $m(\cdot)$) such that χ is a (smooth) strict supersolution of $\partial_t - F$ on $(t_0, T) \times B_R(x_0)$;

$$K - F(t, y, \chi(t, y), 2C(y - x), 2CI) > 0 \text{ on } (t_0, T) \times B_R(x_0).$$

Indeed, by properness we have

$$K - F(t, y, \chi(t, y), 2C(y - x), 2CI) > K - F(t, y, -|u|_{\infty}, 2C(y - x), 2CI);$$

noting $|y - x| \leq 2R$ so that $p := 2C(y - x)$, $X := 2CI$ remain in a bounded set whose size may depend on η through C , it then follows by our structural assumption on the non-linearity⁷ that we can pick $K = K(\eta)$ large enough such as to achieve the claimed strict inequality. (Note that this choice of K is uniformly in t_0 provided we can find C with the correct dependences.) Since, on the other hand, u is a viscosity solution (hence subsolution), it follows from the very definition of a subsolution that

$$K - F(\hat{t}, \hat{y}, \chi(t, y), 2C(\hat{y} - x), 2CI) \leq 0$$

whenever $(\hat{t}, \hat{y}) \in (t_0, T] \times B_R(x_0)$ is a maximum point of $u - \chi$. (Note that $\hat{t} = T$ is possible here, we then rely on part (i) of Theorem 1.) This contradiction shows that maximum points of $u - \chi$ over $[t_0, T] \times \bar{B}_R(x_0)$ are necessarily achieved on the parabolic boundary

$$(t, y) \in [t_0, T] \times \partial B_R(x_0) \cup \{t_0\} \times \bar{B}_R(x_0).$$

The remainder of the proof is thus concerned with showing that $u - \chi \leq 0$ on this parabolic boundary. Consider first the case that $t \in [t_0, T]$ and $|y - x_0| = R$. Since $x \in B_{R/2}(x_0)$ we must have $|y - x| \geq R/2$ and it thus suffices to take $C \geq 8|u|_{\infty; [0, T] \times \mathbb{R}^n}/R^2$ to ensure that

$$u(t_0, y) \leq u(t_0, x) + \eta + C|y - x|^2 + K(t - t_0) \quad (29)$$

⁷ ... notably boundedness of $F(\cdot, \cdot, y, p, X)$ when y, p, X remain in a bounded set ...

for all $t \in [t_0, T]$ and $y \in B_R(x_0)$, and any $\eta, K \geq 0$. The second case to be considered is $t = t_0$ and $y \in \bar{B}_R(x_0)$. We want to see that for every η there exists $C = C(\eta)$ such that

$$u(t_0, y) \leq u(t_0, x) + \eta + C|y - x|^2 \text{ for all } y \in \bar{B}_R(x_0); \quad (30)$$

but this follows immediately from the fact (cf. Theorem 1) that $u(t_0, \cdot)$ has a spatial modulus m . Indeed: If there were $\eta > 0$ such that for all C there are points y_C so that $u(t_0, y) > u(t_0, x) + \eta + C|y - x|^2$, then $|y_C - x|^2 \leq 2|u|_{\infty; [0, T] \times \mathbb{R}^n} / C \rightarrow 0$ with $C \rightarrow \infty$ and a contradiction to

$$m(|y_C - x|) \geq u(t_0, y) - u(t_0, x) \geq \eta > 0.$$

is obtained as soon as C is chosen large enough and this choice depends only on $\eta, |u|_{\infty; [0, T] \times \mathbb{R}^n}$ and m . Since all these quantities are independent of t_0 , so is our choice of C .

(ii): By Theorem 2 we get for every $t \in [0, T]$ that $u(t, \cdot)$ is Lipschitz continuous with the same Lipschitz constant, say L .

For the regularity in time, let η be given. Choose $C = \frac{L^2}{4\eta}$, then (30) holds. If we then fix R large enough, (29) is also fulfilled.

We can now choose K explicitly. Indeed, with $K \geq M(1 + 2C)$, χ is a supersolution. Using this in (27), we get

$$u(t, x) - u(t_0, x) \leq \eta + M\left(1 + \frac{L^2}{4\eta}\right)t,$$

for all $t \in [t_0, T]$. Optimization with respect to η leads to

$$u(t, x) - u(t_0, x) \leq \tilde{L}t^{1/2},$$

where \tilde{L} depends only on M and L . □

2.8 Proof of Theorem 4: Existence

At last, we discuss existence via Perron's Method; the only difficulty in the proof is to produce subsolutions and supersolutions.

Proof Step 1: Assume u_0 is Lipschitz continuous with Lipschitz constant L . Define for $z \in \mathbb{R}^n, \varepsilon > 0$

$$\psi_{\varepsilon, z}(x) := u_0(z) - L\left(|x - z|^2 + \varepsilon\right)^{1/2}.$$

We will show that there exists $A_\varepsilon \leq 0$ (non-positive, yet to be chosen) such that

$$u_{\varepsilon, z}(t, x) := A_\varepsilon t + \psi_{\varepsilon, z}(x)$$

is a (classical) subsolution of $\partial_t - F = 0$. To this end we first note that $Du_{\varepsilon,z} = D\psi_{\varepsilon,z}$ and $D^2u_{\varepsilon,z} = D^2\psi_{\varepsilon,z}$ are bounded by LC_ε where C is a constant dependent on ε . We also note that (for any non-positive choice of A_ε)

$$u_{\varepsilon,z}(t, x) \leq u_{\varepsilon,z}(0, x) = \psi_{\varepsilon,z}(x) \leq u_0(z) - L|x - z| \leq u_0(x),$$

thanks to L -Lipschitzness of u_0 . Since $F = F(t, x, u, p, X)$ is assumed to be proper, and thus in particular anti-monotone in u , we have

$$\begin{aligned} & \partial_t u_{\varepsilon,z} - F(t, x, u_{\varepsilon,z}, Du_{\varepsilon,z}, D^2u_{\varepsilon,z}) \\ &= A_\varepsilon - F(t, x, u_{\varepsilon,z}, D\psi_{\varepsilon,z}, D^2\psi_{\varepsilon,z}) \\ &\leq A_\varepsilon - F(t, x, |u_0|_\infty, D\psi_{\varepsilon,z}, D^2\psi_{\varepsilon,z}). \end{aligned}$$

Since $|u_0|_\infty < \infty$ and $|D\psi_{\varepsilon,z}|, |D^2\psi_{\varepsilon,z}| \leq LC_\varepsilon$ we can use the assumed boundedness of F over sets where u, p, X remain bounded. In particular, we can pick A_ε negative, large enough, such that

$$\partial_t u_{\varepsilon,z} - F(t, x, u_{\varepsilon,z}, Du_{\varepsilon,z}, D^2u_{\varepsilon,z}) \leq \dots \leq 0.$$

We now define the sup of all these subsolutions,

$$\hat{u}(t, x) := \sup_{\varepsilon \in (0,1], z \in \mathbb{R}^n} u_{\varepsilon,z}(t, x) \leq u_0(x) \leq |u_0|_\infty < \infty.$$

Then

$$-\infty < A_1 T - L \leq A_1 t - L = u_{1,x}(t, x) \leq \hat{u}(t, x) \leq u_0(x) \leq |u_0|_\infty < \infty,$$

so that \hat{u} is bounded. We moreover have

$$\hat{u}(0, x) = \sup_{\varepsilon \in (0,1], z \in \mathbb{R}^n} \psi_{\varepsilon,z}(x) = \sup_{\varepsilon \in (0,1]} u_0(x) - L\varepsilon^{1/2} = u_0(x).$$

The upper semicontinuous envelope $\underline{u}(t, x) := \hat{u}^*$ is then (cf. Proposition 8.2 in [8] for instance) also a bounded subsolution to $\partial_t - F = 0$.

Step 2: We show that $\hat{u}(t, x)$ is continuous at $t = 0$; this implies that

$$\underline{u}(0, x) := \hat{u}(0, x) = u_0(x)$$

and thus yields a sub-solution with the correct initial data. Let $(t^n, x^n) \rightarrow (0, x)$. First we show lower semicontinuity, i.e.

$$\liminf_{n \rightarrow \infty} \hat{u}(t^n, x^n) \geq \hat{u}(0, x).$$

Let $\delta > 0$. Choose $\tilde{\varepsilon}, \tilde{z}$ such that

$$u_{\tilde{\varepsilon}, \tilde{z}}(0, x) \geq \hat{u}(0, x) - \delta.$$

Let M be a bound for $|Du_{\tilde{\varepsilon}, \tilde{z}}|$ (and hence for $|D\psi_{\tilde{\varepsilon}, \tilde{z}}|$). Choose N such that for $n \geq N$

$$|t^n|, |x^n - x| \leq \min \left\{ \frac{\delta}{A_{\tilde{\varepsilon}}}, \frac{\delta}{M} \right\}.$$

Then

$$\begin{aligned} \hat{u}(t^n, x^n) &\geq u_{\tilde{\varepsilon}, \tilde{z}}(t^n, x^n) \\ &= u_{\tilde{\varepsilon}, \tilde{z}}(t^n, x^n) - u_{\tilde{\varepsilon}, \tilde{z}}(0, x) + u_{\tilde{\varepsilon}, \tilde{z}}(0, x) \\ &= A_{\tilde{\varepsilon}} t^n + \psi_{\tilde{\varepsilon}, \tilde{z}}(x^n) - \psi_{\tilde{\varepsilon}, \tilde{z}}(x) + u_{\tilde{\varepsilon}, \tilde{z}}(0, x) \\ &\geq \hat{u}(0, x) - 3\delta, \end{aligned}$$

which proves the lower semicontinuity.

For upper semicontinuity, notice that

$$\begin{aligned} u_{\varepsilon, z}(s, y) &= A_{\varepsilon} s + \psi_{\varepsilon, z}(y) \\ &\leq A_{\varepsilon} s + u_0(y) \\ &\leq u_0(y), \end{aligned}$$

where we have used that $A_{\varepsilon} \leq 0$ and that $\psi_{\varepsilon, z}(y) \leq u_0(y)$, as shown above. Hence, $\hat{u}(s, y) \leq u_0(y)$, and then for $(t^n, x^n) \rightarrow (0, x)$, we have

$$\limsup_n \hat{u}(t^n, x^n) \leq \limsup_n u_0(x^n) = u_0(x) = \hat{u}(0, x).$$

Hence \hat{u} is also upper semicontinuous at $(0, x)$ and hence continuous at $(0, x)$.

Step 3: Similarly, one constructs a bounded super-solution with correct (bounded, Lipschitz) initial data u_0 . Perron's method then applies and yields a bounded viscosity solution to $\partial_t - F = 0$ with bounded, Lipschitz initial data.

Step 4: Let now $u_0 \in \text{BUC}(\mathbb{R}^n)$ and u_0^n be a sequence of bounded Lipschitz functions such that $|u_0^n - u_0|_{\infty} \rightarrow 0$. By the previous step there exists a bounded solution u^n to $\partial_t - F = 0$ with initial data $u^n(0, \cdot) = u_0^n$. (It is also unique by comparison.) Since F is proper ($\gamma \geq 0$), the solutions form a contraction in the sense

$$|u^n - u^m|_{\infty; [0, T] \times \mathbb{R}^n} \leq |u_0^n - u_0^m|_{\infty; \mathbb{R}^n}$$

(This follows immediately from comparison and properness.). Hence u^n is Cauchy in supremum norm and converges to a continuous bounded function $u : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$. By Lemma 6.1 in the User's Guide we then have that u is a bounded solution

to $\partial_t - F = 0$ with $\text{BUC}(\mathbb{R}^n)$ initial data. By comparison, it is the unique (bounded) solution with this initial data. At last, Corollary 3 shows that the solution is BUC in time space. \square

2.9 Parabolic Theorem of Sums Revisited

We start some recalls on parabolic jets. If $u : (0, T) \times \mathbb{R}^n \rightarrow \mathbb{R}$ its parabolic semijet $\mathcal{P}^{2,+}u$ is defined by $(b, p, X) \in \mathbb{R} \times \mathbb{R}^n \times \mathcal{S}^n$ lies in $\mathcal{P}^{2,+}u(s, z)$ if $(s, z) \in (0, T) \times \mathbb{R}^n$ and

$$\begin{aligned} u(t, x) &\leq u(s, z) + b(t - s) + \langle p, x - z \rangle \\ &\quad + \frac{1}{2} \langle X(x - z), x - z \rangle + o(|t - s| + |x - z|^2) \end{aligned}$$

as $(0, T) \times \mathbb{R}^n \ni (t, x) \rightarrow (s, z)$. Consider now $u : Q \rightarrow \mathbb{R}$ where $Q = (0, T] \times \mathbb{R}^n$. The parabolic semijet relative to Q , write $\mathcal{P}_Q^{2,+}u$, as used in Ref. [19] for instance, is defined by $(b, p, X) \in \mathbb{R} \times \mathbb{R}^n \times \mathcal{S}^n$ lies in $\mathcal{P}_Q^{2,+}u(s, z)$ if $(s, z) \in (0, T) \times \mathbb{R}^n$ and

$$\begin{aligned} u(t, x) &\leq u(s, z) + b(t - s) + \langle p, x - z \rangle \\ &\quad + \frac{1}{2} \langle X(x - z), x - z \rangle + o(|t - s| + |x - z|^2) \end{aligned}$$

as $Q \ni (t, x) \rightarrow (s, z)$. Note that $\mathcal{P}_Q^{2,+}u(s, z) = \mathcal{P}^{2,+}u(s, z)$ for $(s, z) \in (0, T) \times \mathbb{R}^n$. Note also the special behaviour of the semijet at time T in the sense that

$$(b, p, X) \in \mathcal{P}_Q^{2,+}u(T, z) \implies \forall b' \leq b : (b', p, X) \in \mathcal{P}_Q^{2,+}u(T, z). \quad (31)$$

Closures of these jets are defined in the usual way; e.g.

$$(b, p, X) \in \bar{\mathcal{P}}_Q^{2,+}u(T, z)$$

iff $\exists (t_n, z_n; b_n, p_n, X_n) \in Q \times \mathbb{R} \times \mathbb{R}^n \times \mathcal{S}^n : (b_n, p_n, X_n) \in \bar{\mathcal{P}}_Q^{2,+}u(t_n, z_n)$ and

$$(t_n, z_n; u(t_n, z_n); b_n, p_n, X_n) \rightarrow (T, z; u(T, z); b, p, X).$$

Recall the classical parabolic Theorem of Sums.

Theorem 8 [9, Theorem 7] *Let $u_1, u_2 \in \text{USC}((0, T) \times \mathbb{R}^n)$ and $w \in \text{USC}((0, T) \times \mathbb{R}^{2n})$ be given by*

$$w(t, x) = u_1(t, x_1) + u_2(t, x_2)$$

Suppose that $s \in (0, T)$, $z = (z_1, z_2) \in \mathbb{R}^{2n}$, $b \in \mathbb{R}$, $p = (p_1, p_2) \in \mathbb{R}^{2n}$, $A \in \mathcal{S}^{2n}$ with

$$(b, p, A) \in \mathcal{P}^{2,+}w(s, z). \quad (32)$$

Assume moreover that there is an $r > 0$ such that for every $M > 0$ there is a C such that for $i = 1, 2$

$$b_i \leq C \text{ whenever } (b_i, q_i, X_i) \in \mathcal{P}^{2,+}w(t, x_i), \quad (33)$$

$$|x_i - z_i| + |s - t| < r \text{ and } |u_i(t, x_i)| + |q_i| + \|X_i\| \leq M.$$

Then for each $\varepsilon > 0$ there exists $(b_i, X_i) \in \mathbb{R} \times \mathcal{S}^n$ such that

$$(b_i, p_i, X_i) \in \bar{\mathcal{P}}^{2,+}u_i(s, z_i)$$

and

$$-\left(\frac{1}{\varepsilon} + \|A\|\right)I \leq \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \leq A + \varepsilon A^2 \text{ and } b_1 + b_2 = b. \quad (34)$$

The proof of the above Theorem is usually reduced (cf. Lemma 8 in [9]) to the case $b = 0$, $z = 0$, $p = 0$ and $v_1(s, 0) = v_2(s, 0) = 0$, where (in order to avoid confusion) we write v_i instead of u_i . Condition (32) translates than to

$$v_1(t, x_1) + v_2(t, x_2) - \frac{1}{2} \langle Ax, x \rangle \leq 0 \text{ for all } (t, x) \in (0, T) \times \mathbb{R}^{2n}; \quad (35)$$

this also means that the left-hand-side as a function of (t, x_1, x_2) has a global maximum at $(s, 0, 0)$. The assertion of the (reduced) Theorem is then the existence of $(b_i, X_i) \in \mathbb{R} \times \mathcal{S}^n$ such that $(b_i, 0, X_i) \in \bar{\mathcal{P}}^{2,+}v_i(s, 0)$ for $i = 1, 2$ and (34) holds with $b = 0$.

We now give the main result of this subsection.

Theorem 9 Assume that u_i has a finite extension to $(0, T] \times \mathbb{R}^n$, $i = 1, 2$, via its semi-continuous envelopes, that is,

$$u_i(T, x) = \limsup_{\substack{(t,y) \in (0,T) \times \mathbb{R}^n: \\ t \uparrow T, y \rightarrow x}} u_i(t, y) < \infty.$$

Then the above Theorem remains valid at $s = T$ if

$$\mathcal{P}^{2,+}w(s, z) \text{ and } \bar{\mathcal{P}}^{2,+}u_i(s, z_i)$$

is replaced by

$$\mathcal{P}_Q^{2,+}w(T, z) \text{ and } \bar{\mathcal{P}}_Q^{2,+}u_i(T, z_i)$$

and the final equality in (34) is replaced by

$$b_1 + b_2 \geq b. \quad (36)$$

Remark 5 If we knew (but we don't!) that the final conclusion is $(b_i, p_i, X_i) \in \bar{\mathcal{P}}_Q^{2,+} u(T, z_i)$, rather than just being an element in the closure $\bar{\mathcal{P}}_Q^{2,+} u(T, z_i)$, then we could trivially diminish the b_i 's such as to have $b_1 + b_2 = b$; cf. (31).

Remark 6 If one knows that w is left-accessible in the sense of [7] it should be possible to simplify the proof by avoiding doubling in t (and thus the penalty parameter m) and by using the classical parabolic (instead of elliptic) Theorem of Sums on $(0, T)$ in step 1 of the proof below. Left-accessibility is actually given in our application since then $u_1, -u_2$ are viscosity sub-/supersolutions. In that case condition (33) is also immediately satisfied, see e.g. Lemma V.6.1 in [14].

Proof Step 1: We focus on the reduced setting (and thus write v_i instead of u_i) and (following the proof of Lemma 8 in [9]) redefine $v_i(t_i, x_i)$ as $-\infty$ when $|x_i| > 1$ or $t_i \notin [T/2, T]$. We can also assume that (35) is strict if $t < s = T$ or $x \neq 0$. For the rest of the proof, we shall abbreviate $(t_1, t_2), (x_1, x_2)$ etc. by (t, x) . With this notation in mind we set

$$w(t, x) = v_1(t_1, x_1) + v_2(t_2, x_2) - \frac{1}{2} \langle Ax, x \rangle.$$

By the extension via semi-continuous envelopes, there exist a sequence $(t^n, x^n) \in (0, T)^2 \times (\mathbb{R}^n)^2$, such that

$$(t^n, x^n) \equiv (t^{1,n}, t^{2,n}, x^{1,n}, x^{2,n}) \rightarrow (T, T, 0, 0)$$

and

$$w(t^{1,n}, t^{2,n}, x^{1,n}, x^{2,n}) \rightarrow w(T, T, 0, 0).$$

We now consider w with a penalty term for $t_1 \neq t_2$ and a barrier at time T for both t_1 and t_2 .

$$\psi_{m,n}(t, x) = w(t, x) - \left\{ \frac{m}{2} |t_1 - t_2|^2 + \sum_{i=1}^2 \left(T - t^{i,n} \right)^2 / (T - t_i) \right\},$$

indexed by $(m, n) \in \mathbb{N}^2$, say. By assumption w has a maximum at $(T, T, 0, 0)$ which we may assume to be strict (otherwise subtract suitable forth order terms ...). Define now

$$(\hat{t}, \hat{x}) \in \arg \max \psi_{m,n} \text{ over } [T - r, T]^2 \times \bar{B}_r(0)^2$$

where $r = T/2$ (for instance). When we want to emphasize dependence on m, n we write $(\hat{t}_{m,n}, \hat{x}_{m,n})$. We shall see below (Step 2) that there exists increasing sequences $m = m(k), n = n(k)$ so that

$$(\hat{t}, \hat{x})|_{m=m(k), n=n(k)} \rightarrow (T, T, 0, 0). \quad (37)$$

Using the (elliptic) Theorem of Sums in the form of [9, Theorem 1] we find that there are

$$(b_i, p_i, X_i) \in \bar{\mathcal{P}}^{2,+}_{v_i}(\hat{t}_i, \hat{x}_i)$$

(where $\hat{t}_i \rightarrow T, \hat{x}_i \rightarrow 0$ as $k \rightarrow \infty$) such that the first part of (34) holds and

$$A \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}, \quad b_i = m(t_i - t_{3-i}) + (T - t^{i,\varepsilon})^2 / (T - t_i)^2.$$

for $i = 1, 2$. Note that

$$b_1 + b_2 = m(t_1 - t_2) + m(t_2 - t_1) + (\text{positive terms}) \geq 0;$$

since each b_i is bounded above by the assumptions and the estimates on the X_i it follows that the b_i lie in precompact sets. Upon passing to the limit $k \rightarrow \infty$ we obtain points

$$(b_i, p_i, X_i) \in \bar{\mathcal{P}}^{2,+}_{v_i}(T, 0), \quad i = 1, 2;$$

with $b_1 + b_2 \geq 0$.

Step 2: We still have to establish (37). We first remark that for arbitrary (strictly) increasing sequences $m(k), n(k)$, compactness implies that

$$\{(\hat{t}_{m(k),n(k)}, \hat{x}_{m(k),n(k)}) : k \geq 1\} \in [T - r, T]^2 \times \bar{B}_r(0)^2$$

has limit points. Note also $\hat{t}_1, \hat{t}_2 \in [T - r, T)$ thanks to the barrier at time T . The key technical ingredient for the remained of the argument is and we postpone details of these to Step 3 below:

$$\begin{aligned} w(\hat{t}, \hat{x}) - \psi_{m,n}(\hat{t}, \hat{x}) &= \left\{ \frac{m}{2} |\hat{t}_1 - \hat{t}_2|^2 + \sum_{i=1}^2 (T - t^{i,n})^2 / (T - \hat{t}_i) \right\} \\ &\rightarrow 0 \text{ as } \frac{1}{n} \ll \frac{1}{m} \rightarrow 0. \end{aligned} \quad (38)$$

In particular, for every $k > 0$ there exists $m(k)$ such that for all $m \geq m(k)$

$$\limsup_{n \rightarrow \infty} \{ \dots \} < \frac{1}{k}.$$

By making $m(k)$ larger if necessary we may assume that $m(k)$ is (strictly) increasing in k . Furthermore there exists $n(m(k), k) = n(k)$ such that for all $n \geq n(k) : \{\dots\} < 2/k$. Again, we may make $n(k)$ larger if necessary so that $n(k)$ is strictly increasing. Recall $t^{1,n(k)} - t^{2,n(k)} \rightarrow T - T = 0$ as $k \rightarrow \infty$. For reasons that will become apparent further below, we actually want the stronger statement that

$$\frac{m(k)}{2} \left| t^{1,n(k)} - t^{2,n(k)} \right|^2 \rightarrow 0 \text{ as } k \rightarrow \infty \quad (39)$$

which we can achieve by modifying $n(k)$ such as to run to ∞ even faster. Note that the so-constructed $m = m(k)$, $n = n(k)$ has the property

$$\left[w(\hat{t}, \hat{x}) - \psi_{m,n}(\hat{t}, \hat{x}) \right] |_{m=m(k), n=n(k)} = \{\dots\} |_{m=m(k), n=n(k)} \rightarrow 0 \text{ as } k \rightarrow \infty. \quad (40)$$

By switching to a subsequence (k_l) if necessary we may also assume (after relabeling) that

$$(\hat{t}_{m(k),n(k)}, \hat{x}_{m(k),n(k)}) \rightarrow (\tilde{t}, \tilde{x}) \in [T - r, T]^2 \times \bar{B}_r(0)^2 \text{ as } k \rightarrow \infty.$$

In the sequel we think of (\hat{t}, \hat{x}) as this sequence indexed by k . We have

$$\begin{aligned} w(\tilde{t}, \tilde{x}) &\geq \limsup_{k \rightarrow \infty} w(\hat{t}, \hat{x}) |_{m=m(k), n=n(k)} \text{ by upper-semi-continuity} \\ &= \limsup_{k \rightarrow \infty} \psi_{m,n}(\hat{t}, \hat{x}) |_{m=m(k), n=n(k)} \text{ thanks to (40).} \end{aligned} \quad (41)$$

On the other hand, thanks to the particular form of our time- T barrier,

$$\begin{aligned} \psi_{m,n}(\hat{t}, \hat{x}) &\geq \psi_{m,n}(t^n, x^n) \\ &= w(t^n, x^n) - \left\{ \frac{m}{2} \left| t^{1,n} - t^{2,n} \right|^2 + \sum_{i=1}^2 \left(T - t^{i,n} \right) \right\}. \end{aligned}$$

Take now $m = m(k)$, $n = n(k)$ as constructed above. Then

$$\begin{aligned} \psi_{m,n}(\hat{t}, \hat{x}) |_{m=m(k), n=n(k)} &\geq w(t^{n(k)}, x^{n(k)}) \\ &\quad - \left\{ \frac{m(k)}{2} \left| t^{1,n(k)} - t^{2,n(k)} \right|^2 + \sum_{i=1}^2 \left(T - t^{i,n(k)} \right) \right\} \end{aligned}$$

The first term in the curly bracket goes to zero (with $k \rightarrow \infty$) thanks to (39), the other term goes to zero since $t^{i,n} \rightarrow T$ with $n \rightarrow \infty$, and hence also along $n(k)$. On the other hand (recall $x^{i,n} \rightarrow 0$)

$$w(t^{n(k)}, x^{n(k)}) \rightarrow v_1(T, 0) + v_2(T, 0) - \frac{1}{2} \langle A0, 0 \rangle = 0 \text{ as } k \rightarrow \infty.$$

(In the reduced setting $v_1(T, 0) = v_2(T, 0) = 0$.) It follows that

$$\lim_{k \rightarrow \infty} \inf \psi_{m,n}(\hat{t}, \hat{x})|_{m=m(k), n=n(k)} = 0.$$

Together with (41) we see that $w(\tilde{t}, \tilde{x}) \geq 0$. But $w(T, T, 0, 0) = 0$ was a strict maximum in $[T - r, T]^2 \times \bar{B}_r(0)^2$ and so we must have $(\tilde{t}, \tilde{x}) = (T, T, 0, 0)$.

Step 3: Set

$$M(h) = \sup_{\substack{(t,x) \in [T-r, T]^2 \times \bar{B}_r(0)^2 \\ |t_1 - t_2| < h}} w(t_1, t_2, x_1, x_2) \text{ and } M' = \lim_{h \rightarrow 0} M(h)$$

It is enough to show

$$\limsup_{\frac{1}{n} < < \frac{1}{m} \rightarrow 0} w(\hat{t}, \hat{x}) \leq M' \leq \liminf_{\frac{1}{n} < < \frac{1}{m} \rightarrow 0} \psi_{m,n}(\hat{t}, \hat{x}). \quad (42)$$

since the claimed

$$\begin{aligned} w(\hat{t}, \hat{x}) - \psi_{m,n}(\hat{t}, \hat{x}) &= \left\{ \frac{m}{2} |\hat{t}_1 - \hat{t}_2|^2 + \sum_{i=1}^2 \left(T - t^{i,n} \right)^2 / (T - \hat{t}_i) \right\} \\ &\rightarrow 0 \text{ as } \frac{1}{n} < < \frac{1}{m} \rightarrow 0. \end{aligned}$$

follows from

$$\begin{aligned} \limsup_{\frac{1}{n} < < \frac{1}{m} \rightarrow 0} \{ \dots \} &\leq \limsup_{\frac{1}{n} < < \frac{1}{m} \rightarrow 0} w(\hat{t}, \hat{x}) - \liminf_{\frac{1}{n} < < \frac{1}{m} \rightarrow 0} \psi_{m,n}(\hat{t}, \hat{x}) \\ &\leq 0 \text{ (and hence } = 0). \end{aligned}$$

Note that $w(\hat{t}, \hat{x})$ is bounded on $[T - r, T]^2 \times \bar{B}_r(0)^2$ so that

$$|\hat{t}_1 - \hat{t}_2|^2 = O(1/m) \implies w(\hat{t}, \hat{x}) \leq M(\text{const}/\sqrt{m}).$$

On the other hand, from the very definition of M' as $\lim_{h \rightarrow 0} M(h)$, there exists a family (t_h, x_h) so that

$$|t_{1,h} - t_{2,h}| \leq h \text{ and } w(t_h, x_h) \rightarrow M' \text{ as } h \rightarrow 0 \quad (43)$$

For every m, n we may take (t_h, x_h) as argument of $\psi_{m,n}$ (which itself has a maximum at \hat{t}, \hat{x}); hence

$$w(t_h, x_h) - \frac{m}{2} h^2 - \sum_{i=1}^2 \left(T - t^{i,n} \right)^2 / (T - t_{i,h}) \leq \psi_{m,n}(\hat{t}, \hat{x}). \quad (44)$$

Take now a sequence $n = n(h)$, fast enough increasing as $h \searrow$ such that $(T - t^{i,n})^2 / (T - t_{i,h}) \rightarrow 0$ with $h \rightarrow 0$. It follows that

$$\begin{aligned} M' &= \lim_{h \rightarrow 0} w(t_h, x_h) \\ &= \liminf_{h \rightarrow 0} \left(w(t_h, x_h) - \frac{m}{2} h^2 - \sum_{i=1}^2 (T - t^{i,n(h)})^2 / (T - t_{i,h}) \right) \\ &\leq \liminf_{h \rightarrow 0} \psi_{m,n(h)}(\hat{t}, \hat{x}) = \liminf_{n \rightarrow \infty} \psi_{m,n}(\hat{t}, \hat{x}) \text{ by monotonicity of } \sup \psi_{m,n} \text{ in } n. \end{aligned}$$

(In the last equality we used that $t^{i,n} \uparrow T$; this shows that $\sup \psi_{m,n}$ is indeed monotone in n .) The proof is now finished. \square

3 RPDEs: Existence, Uniqueness and Regularity

We start with a brief recall on *rough path theory*; [15, 29–31]. 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

$$\dot{y}(t) = \sum_{i=1}^d V_i(y) \dot{z}^i(t), \quad y(0) = y_0 \in \mathbb{R}^n. \quad (45)$$

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. If $|z - \tilde{z}|_{\infty; [0, T]}$ is replaced by the (much) stronger distance $(z_{s,t} := z_t - z_s)$

$$|z - \tilde{z}|_{1\text{-H\"ol}; [0, T]} = \sup_{0 \leq s < t \leq T} \frac{|z_{s,t} - \tilde{z}_{s,t}|}{|t - s|},$$

it is elementary to see that now the solution map is continuous (in fact, locally Lipschitz); however, this continuity does not lend itself to push the meaning of (45): the closure of smooth paths in this metric yields precisely C^1 . Lyons' theory of rough paths exhibits an entire cascade of α -H\"older type rough path (or, as a variation on the scheme: $(1/\alpha)$ -variation) metrics, for each $\alpha \in (0, 1]$ on path-space under which such ODE solutions are continuous (and even locally Lipschitz) functions of their driving signal. For instance, the “rough path” α -H\"older distance between two smooth \mathbb{R}^d -valued paths z, \tilde{z} is given by

$$\max_{j=1, \dots, [1/\alpha]} \sup_{0 \leq s < t \leq T} \sum \frac{|z_{s,t}^{(j)} - \tilde{z}_{s,t}^{(j)}|}{|t - s|^{j\alpha}}$$

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 (45), 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 α -Hölder metric). The space of so-called α -Hölder rough paths⁸ is precisely this abstract completion. In fact, this space can be realized as genuine path space,

$$C^{0,\alpha\text{-Höl}}([0, T], G^{[1/\alpha]}(\mathbb{R}^d))$$

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

$$\mathfrak{t}^{[1/\alpha]}(\mathbb{R}^d) = \mathbb{R}^d \oplus (\mathbb{R}^d)^{\otimes 2} \oplus \cdots \oplus (\mathbb{R}^d)^{\otimes [1/\alpha]}$$

is the natural state space for (up to $[1/\alpha]$) 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

$$B^{(2)} := \left(\int_0^\cdot B^i \circ dB^j \right)_{i,j \in \{1, \dots, d\}} \quad (46)$$

yields a path $\mathbf{B}(\omega)$ in $G^2(\mathbb{R}^d)$ with finite α -Hölder, for any $\alpha < 1/2$. (\mathbf{B} is known as *Brownian rough path*.) We remark that $B^{(2)} = \frac{1}{2}B \otimes B + A$ where $A := \text{Anti}(B^{(2)})$ 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 section, we follow [25, 26, 28] in considering a real-valued function of time and space $u = u(t, x) \in \text{BUC}([0, T] \times \mathbb{R}^n)$ which solves the nonlinear partial differential equation

$$\begin{aligned} du &= F(t, x, Du, D^2u) dt + \sum_{i=1}^d H_i(x, Du) dz^i \\ &\equiv F(t, x, Du, D^2u) dt + H(x, Du) dz \end{aligned} \quad (47)$$

in viscosity sense. When $z : [0, T] \rightarrow \mathbb{R}^d$ 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. [25, Remark 4] and the references given there) but, as was pointed out by various authors, the case when $z = z(t)$ has

⁸ In the strict terminology of rough path theory: geometric α -Hölder rough paths.

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 [24, 26] justifying the need of a theory of (non-linear) *stochastic partial differential equations* (SPDEs) in which z in (47) is taken as a Brownian motion.⁹ In the same series of articles a satisfactory theory is established for the case of non-linear Hamiltonian with no spatial dependence, i.e. $H = H(Du)$. Our present discussion follows [6] in that we consider non-linear F and $H = H(x, Du)$, linear in Du . The following condition may be considered as a global version of the corresponding definition put forward by Lions–Souganidis.

Definition 1 Let ϕ denote the solution flow to the RDE $dy = V(y) dz(t)$. (As is well known, this yields a C^3 -flow of diffeomorphisms provided $V = (V_1, \dots, V_d)$ is a collection of $\text{Lip}^{\gamma+2}(\mathbb{R}^n; \mathbb{R}^n)$ vector fields with $\gamma > 1/\alpha$ and if

$$z \in C^{0, \alpha-H\ddot{o}l}([0, T], G^{[1/\alpha]}(\mathbb{R}^d)).$$

A continuous function u is called a *rough viscosity solution* to

$$du = F(t, x, Du, D^2u) dt - Du(t, x) \cdot V(x) dz(t)$$

if $v(t, \cdot) := u(t, \phi_t(\cdot))$ is a viscosity solution to

$$\partial_t v - \tilde{F}(t, x, Dv, D^2v) = 0$$

where $\tilde{F}(t, x, p, X)$ is given by

$$\begin{aligned} F\left(t, \phi_t(x), \left\langle p, D\phi_t^{-1}|_{\phi_t(x)} \right\rangle, \left\langle X, D\phi_t^{-1}|_{\phi_t(x)} \otimes D\phi_t^{-1}|_{\phi_t(x)} \right\rangle \right. \\ \left. + \left\langle p, D^2\phi_t^{-1}|_{\phi_t(x)} \right\rangle\right). \end{aligned} \quad (48)$$

It should be noted that in the case when z arises from a smooth \mathbb{R}^d -valued path z , the definition is consistent with the interpretation

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

this is verified in Ref. [6], for instance. The following result adds existence and regularity to the main result of Ref. [6].

⁹ ... in which case (47) is understood in Stratonovich form.

Theorem 10 Let $\alpha \in (0, 1]$ and $(z^\varepsilon) \subset C^\infty([0, T], \mathbb{R}^d)$ be Cauchy in $(\alpha\text{-H\"older})$ rough path topology with rough path limit $\mathbf{z} \in C^{0, \alpha\text{-H\"ol}}([0, T], G^{[1/\alpha]}(\mathbb{R}^d))$. Assume

$$u_0^\varepsilon \in \text{BUC}(\mathbb{R}^n) \rightarrow u_0 \in \text{BUC}(\mathbb{R}^n),$$

uniformly as $\varepsilon \rightarrow 0$. Let $F = F(t, x, p, X)$ be continuous, degenerate elliptic, and assume that $\partial_t - \tilde{F}$ where \tilde{F} is given in (48) satisfies the assumptions of Sect. 2.1 with assumption (2.1) strengthened to ¹⁰

$$\forall R > 0 : \tilde{F}|_{[0, T] \times \mathbb{R}^n \times [-R, R] \times B_R \times M_R} \text{ is bounded, uniformly continuous.}$$

Assume that $V = (V_1, \dots, V_d)$ is a collection of $\text{Lip}^{\gamma+2}(\mathbb{R}^n; \mathbb{R}^n)$ vector fields with $\gamma > 1/\alpha$. Then

(i) **Existence, uniqueness:** there exists unique BUC-solutions to the approximate problems

$$du^\varepsilon = F\left(t, x, Du^\varepsilon, D^2u^\varepsilon\right) dt - Du^\varepsilon(t, x) \cdot V(x) dz^\varepsilon(t), \quad (49)$$

$$u^\varepsilon(0, \cdot) = u_0^\varepsilon \quad (50)$$

and the locally uniform limit $u = \lim_{\varepsilon \rightarrow 0} u^\varepsilon$ exists, and is the unique BUC $([0, T] \times \mathbb{R}^n)$ rough viscosity solution (Definition 1) to

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

$$u(0, \cdot) = u_0 \in \text{BUC}(\mathbb{R}^n).$$

In particular, $u = u^\mathbf{z}$ only depends on \mathbf{z} and u_0 but not on the particular approximating sequence $\{z^\varepsilon\}$.

(ii) **Robustness:** The solution map $(\mathbf{z}, u_0) \mapsto u^\mathbf{z}$ from

$$C^{0, \alpha\text{-H\"ol}}([0, T], G^{[p]}(\mathbb{R}^d)) \times \text{BUC}(\mathbb{R}^n) \rightarrow \text{BUC}([0, T] \times \mathbb{R}^n)$$

is continuous.

(iii) **Regularity:** Make the additional assumption that \tilde{F} has modulus θ_R which is actually linear, and that \tilde{F} has linear growth in the Hessian. Then, if $u_0 \in \text{Lip}^1(\mathbb{R}^n)$, i.e. bounded and Lipschitz continuous, then u is Lipschitz in space and $\min(\alpha, 1/2)$ -H\"older in time.

Proof By assumption, $F(\cdot, \phi, (\cdot) \cdot, 0, 0) = \tilde{F}(\cdot, \cdot, 0, 0)$ is bounded on $[0, T] \times \mathbb{R}^n$, and the assumption that $u_0^\varepsilon \rightarrow u_0$ uniformly, as can be seen by comparison with function of the type $(t, x) \mapsto \pm C(t+1)$, with sufficiently large C , independent of

¹⁰ This may be expressed in terms of F ; in particular F then satisfies $\Phi^{(3)}$ -invariant comparison as introduced in Ref. [6]; there it was also checked that these structural assumptions are satisfied by many examples; notably those arising from pathwise stochastic control theory.

ε , that the family

$$\{u^\varepsilon\}$$

is locally uniformly bounded in ε . We also note that our structural assumptions imply both existence and comparison (hence uniqueness) of the approximate problems. Since RDE solution flows (as C^3 -flows of diffeomorphisms) depend continuously on the rough driving signals [15, Theorem 11.12, 11.13], the corresponding $\tilde{F}(\varepsilon)$ given by (48), with ϕ give as the solution flow to $dy = V(y) dz^\varepsilon$, converge locally uniformly to \tilde{F} based on the RDE solution flow to $dy = V(y) dz$. Thanks to the Barles-Perthame procedure (which we apply to the transformed equation, cf (iii) below) we see that

$$u = \lim_{\varepsilon \rightarrow 0} u^\varepsilon$$

exists locally uniformly and is a rough viscosity solution in the sense of Definition 1. Since solutions to $\partial_t - \tilde{F}$ with BUC initial data, are also BUC in space time, we see that $u \in \text{BUC}([0, T] \times \mathbb{R}^n)$. This settles (i).

The argument for (ii) is identical, one just considers approximate RDE problems $dy = V(y) dz^\varepsilon$ where each z^ε may be a genuine rough paths (rather than a smooth, approximating path).

(iii) Since $v(t, \cdot) := u(t, \phi_t(\cdot))$ is a viscosity solution to

$$\partial_t v - \tilde{F}(t, x, Dv, D^2v) = 0$$

it follows from the results in Sect. 2 that v is Lipschitz in space and $(1/2)$ -Hölder in time. Since the time-space regularity

$$(x, t) \mapsto \phi_t(\cdot) \text{ resp. } \phi_t^{-1}(\cdot)$$

is well-understood [15], namely α -Hölder in time and Lipschitz (and more) in space, the claimed regularity of the rough viscosity solution u follows. \square

4 Examples

Let us verify in some detail that Theorem 10 applies to two concrete rough resp. stochastic partial differential equations of interest.

4.1 Stochastic HJB Equation

Following [25] consider the following stochastic Hamilton–Jacobi–Bellman equation

$$du = \inf_{\alpha \in A} L_\alpha(t, x, Du, D^2u) dt + Du(t, x) \cdot V_i(x) \circ dB.$$

The case of constant vector fields $V = (V_1, \dots, V_d)$ was treated in the aforementioned reference; one then has a truly pathwise theory, i.e. for every continuous path $B : [0, T] \rightarrow \mathbb{R}^d$. Moreover, the solution is robust in the sense that it is the “Wong–Zakai” limit of approximate problems where B is replaced by (piecewise) smooth B^ε , uniformly convergent to B as $\varepsilon \rightarrow 0$. This was extended with the aid of rough path theory to non-constant vector fields in Ref. [6]. It was assumed that all approximate problems have a unique (viscosity) solution, convergence then takes places if (B^ε) is Cauchy in rough path metric. Although existence is not an issue here (one could use stochastic control theory and then the stability result of [6]), Theorem 10 applies on purely analytical grounds and also gives provides information about the regularity of the solution. To be a little more specific, one assumes

$$L_\alpha(t, x, p, X) = \text{Tr} \left[\sigma_\alpha(t, x) \sigma_\alpha(t, x)^T X \right] + b_\alpha(t, x) \cdot p,$$

where $\sigma_\alpha(t, x) : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n'}$ and $b_\alpha(t, x) : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ are bounded, continuous, and Lipschitz continuous in x , uniformly in $t \in [0, T]$; with all these properties valid uniformly in $\alpha \in A$. Assume also that V_1, \dots, V_d are in $\text{Lip}^5(\mathbb{R}^n)$, say, reflecting the regularity need for a α -Hölder driving rough path, any $\alpha < 1/2$, such as Brownian motion enhanced to a rough path. A rigorous application of Theorem 10 is easily justified, along the lines of [6] example 3 and 5 on pages 33, 35 respectively. (The only structural aspect of \tilde{F} which was not verified explicitly in that paper was boundedness of \tilde{F} , but this an easy consequence of the fact that the auxiliary (rough) flows have bounded derivatives.)

4.2 Zakai Equation

As is well-known, the filtering problem for non-linear diffusion leads to the following stochastic partial differential equation. To be precise, consider

$$dX_t = b(X_t) dt + \sigma(X_t) dB_t + V(X_t) d\tilde{B}_t \quad (51)$$

$$dY_t = h(X_t) dt + d\tilde{B}_t$$

where B and \tilde{B} are independent, multidimensional Brownian motions. Note that the generator of X is of the form

$$\mathcal{L}u = \frac{1}{2} \text{Tr} \left[\sigma(x) \sigma(x)^T D^2 u \right] + \frac{1}{2} \text{Tr} \left[V(x) V(x)^T D^2 u \right] + b(x) \cdot Du$$

Define also the first order operators, $V = (V_1, \dots, V_d)$

$$\mathcal{N}_k u = V_k(x) \cdot Du + h(x) u, \quad k = 1, \dots, d$$

and note that the (formal) adjoint $\mathcal{N}_k^* u$ is of the form $\mathcal{N}_k^* u = H(x, u, Du)$ with H linear in u and Du . The goal is to compute for a given real-valued function φ

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

and from basic principles it follows that there exists a map $\phi_t^\varphi : C([0, T], \mathbb{R}^{d_Y}) \rightarrow \mathbb{R}$ such that

$$\phi_t^\varphi(Y|_{[0,t]}) = \pi_t(\varphi) \quad \mathbb{P} - a.s. \quad (52)$$

As pointed out by M. Clark in the late seventies, there is a problem here since a basic (practical!) requirement of a filter is robustness, say in the form of continuous dependence of the observation path. He then showed that in the case of uncorrelated noise ($\sigma \equiv 0$ in (51)) (which corresponds to the case $V = 0$ above) there exists a unique $\bar{\phi}_t^\varphi : C([0, T], \mathbb{R}^n) \rightarrow \mathbb{R}$ which is continuous in uniform norm and fulfills (52), thus providing a version of the conditional expectation $\pi_t(\varphi)$ which is robust under approximations in uniform norm of the observation Y . Unfortunately in the correlated noise case this is no longer true! In Ref. [11] it was recently shown that in this case robustness still holds in a rough path sense. Now recall that under well-known conditions (e.g. [1]), π_t can be written in the form

$$\pi_t(\varphi) = \int_{\mathbb{R}^{d_X}} \varphi(x) \frac{u_t(x)}{\int u_t(\tilde{x}) d\tilde{x}} dx \quad (53)$$

where $u_t \in L^1(\mathbb{R}^n)$ a.s. and (u_t) is the L^2 -solution of the (dual) Zakai SPDE¹¹

$$\begin{aligned} du_t &= \mathcal{L}^* u_t dt + \sum_k \mathcal{N}_k^* u_t dY_t^k \\ &= \left(\mathcal{L}^* - \frac{1}{2} \sum_k \mathcal{N}_k^* \mathcal{N}_k^* \right) u_t dt + \sum_k \mathcal{N}_k^* u_t \circ dY_t^k \\ &\equiv F(x, u, Du, D^2 u) dt + H(x, u, Du) \circ dY_t^i. \end{aligned}$$

Note that \mathcal{L}^* has the “stochastic” parabolicity property which here means that F is degenerate elliptic. The resulting stochastic PDE does not quite fall in the class considered in Theorem 10, for H has additional dependence on u , but the method still works. A more complicated transformation, detailed in Ref. [16], allows to given (rough)pathwise meaning to the above equation. Indeed, one can show that $v(t, x) = \psi^{-1}(t, u(t, \phi(t, x)), x)$ satisfies a parabolic equation of the form

$$\partial_t v - \tilde{F}(x, v, Dv, D^2 v) = 0$$

¹¹ Consistency of L^2 -theory with RPDE theory has been established in Ref. [16].

with initial data $v(0, \cdot) = u(0, \cdot)$. Moreover, with this definition u will depend continuously on the driving signal (i.e. the observation path Y) in rough path topology; thereby solving the robustness problem for the (un-normalized) conditional density.

In fact, \tilde{F} turns out to be linear in v , Dv , D^2v and there is an explicit, if complicated, expression. In particular then, if b , σ are bounded and Lipschitz, $V \in \text{Lip}^{\gamma+2}$ and $h \in \text{Lip}^{\gamma+1}$ for some $\gamma > 2$, \tilde{F} is seen to meet the assumption of Sect. 2, necessary to arrive at the same conclusions as those stated in Theorem 10. An appealing feature here is that one can immediately handle degenerate situations (including the case when F degenerates to a purely first order operator) and also that one gets automatically continuous versions of solutions to the Zakai equation, without requiring dimension-dependent regularity assumptions on the coefficients (as pointed out by Krylov [23], a disadvantage of the L^2 theory of SPDEs). On the other hand, our regularity assumption (in particular in the noise terms) are more stringent¹² than what is needed to ensure existence and uniqueness in the L^2 theory of SPDEs.

Acknowledgments J. Diehl is supported by DFG Priority Program 1324; P. Friz and H. Oberhauser received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement nr. 258237. P. Friz would like to thank G. Barles for a helpful email exchange, E. Jakobsen for pointing us to [20, 21], M. Soner and T. Souganidis for discussions during the Oberwolfach meetings on *Stochastic Analysis in Finance and Insurance and Rough Paths and PDEs*.

References

1. A. Bain, D. Crisan, *Fundamentals of Stochastic Filtering* (Springer, New York, 2009)
2. G. Barles, *Solutions de viscosité des équations de Hamilton-Jacobi* (Springer, Berlin, 2004)
3. M. Bardi, I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations* (Birkhaeuser, Boston, 1997)
4. G. Barles, S. Biton, O. Ley, A geometrical approach to the study of unbounded solutions of quasilinear parabolic equations. Arch. Ration. Mech. Anal. **162**(4), 287–325 (2002)
5. M. Caruana, P. Friz, Partial differential equations driven by rough paths. J. Differ. Equ. **247**(1), 140–173 (2009)
6. M. Caruana, P. Friz, H. Oberhauser, A (rough) pathwise approach to a class of nonlinear SPDEs, Annales de l'Institut Henri Poincaré/Analyse non linéaire vol. 28, pp. 27–46 (ISSN: 0294-1449) (2011)
7. Y.-G. Chen, Y. Giga, S. Goto, Remarks on viscosity solutions for evolution equations. Proc. Japan Acad. Ser. A Math. Sci. **67**(10), 323–328 (1991)
8. M.G. Crandall, *Viscosity Solutions: A Primer*, Lecture Notes in Mathematics (Springer, New York, 1996)
9. M.G. Crandall, H. Ishii, The maximum principle for semicontinuous functions. Differ. Integral Equ. **3**(6), 1001–1014 (1990)
10. M.G. Crandall, H. Ishii, P.-L. Lions, User's guide to viscosity solutions of second order partial differential equations. Bull. Amer. Math. Soc. (N.S.) **27**(1), 1–67 (1992)
11. D. Crisan, J. Diehl, P. Friz, Harald Oberhauser Robust filtering: correlated noise and multidimensional observation, Ann Appl Probab. **23**(5), 2139–2160 (2013). doi:[10.1214/12-AAP896](https://doi.org/10.1214/12-AAP896)

¹² Above regularity assumptions can be weakened with the “joint lift” approach carried out in Ref. [13].

12. J. Diehl, F. Peter, Backward stochastic differential equations with rough drivers. *The Annals of Probability* 2012, Vol. 40, (4) 1715–1758, Publisher: Institute of Mathematical Statistics
13. J. Diehl, H. Oberhauser, S. Riedel, A Levy-area between Brownian motion and rough paths with applications to robust non-linear filtering and RPDEs. *Stoch Process Appl.* **125**(1), 161–181 January 2015. <http://dx.doi.org/10.1016/j.spa.2014.08.005>
14. H.W. Fleming, H.M. Soner, *Controlled Markov Processes and Viscosity Solutions, Volume 25 of Stochastic Modelling and Applied Probability*, 2nd edn. (Springer, New York, 2006)
15. P.K. Friz, N.B. Victoir, *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*, *Cambridge Studies in Advanced Mathematics* (Cambridge University Press, Cambridge, 2010)
16. P. Friz, H. Oberhauser, Rough path stability of (semi-)linear SPDEs. *Probab Theory Relat Fields.* **158**(1-2), 401–434 February 2014
17. P. Friz, H. Oberhauser, On the splitting-up method for rough (partial) differential equations, *J. Differ. Equ.*, Vol. 251 (2) pp. 316–338 ISSN: 0022–0396 (2011)
18. Y. Giga, *Surface Evolution Equations: A Level Set Approach* (Birkhauser, Basel, 2006)
19. 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), 443–470 (1991)
20. E.R. Jakobsen, K.H. Karlsen, Continuous dependence estimates for viscosity solutions of fully nonlinear degenerate parabolic equations. *J. Differ. Equ.* **183**(2), 497–525 (2002). (ISSN: 0022-0396)
21. E.R. Jakobsen, K.H. Karlsen, Continuous dependence estimates for viscosity solutions of integro-PDEs. *J. Differ. Equ.* **212**(2), 278–318 (2002)
22. Petri Juutinen, On the definition of viscosity solutions for parabolic equations. *Proc. Amer. Math. Soc.* **129**, 2907–2911 (2001)
23. N.V. Krylov, An analytic approach to SPDEs, in *Stochastic partial differential equations: six perspectives of Math. Surveys Monogr.*, Amer. Math. Soc., Providence, vol. 64 pp. 185–242 (1999)
24. P.-L. Lions, P.E. Souganidis, Viscosity solutions of fully nonlinear stochastic partial differential equations. *Sūrikaiseikikenkyūsho Kōkyūroku*, (1287):58–65. Viscosity solutions of differential equations and related topics (Japanese) (Kyoto, 2001) (2002)
25. P.-L. Lions, P.E. Souganidis, Fully nonlinear stochastic partial differential equations. *C. R. Acad. Sci. Paris Sér. I Math.* **326**(9), 1085–1092 (1998)
26. P.-L. Lions, P.E. Souganidis, Fully nonlinear stochastic partial differential equations: non-smooth equations and applications. *C. R. Acad. Sci. Paris Sér. I Math.* **327**(8), 735–741 (1998)
27. P.-L. Lions, P.E. Souganidis, Fully nonlinear stochastic pde with semilinear stochastic dependence. *C. R. Acad. Sci. Paris Sér. I Math.* **331**(8), 617–624 (2000)
28. P.-L. Lions, P.E. Souganidis, Uniqueness of weak solutions of fully nonlinear stochastic partial differential equations. *C. R. Acad. Sci. Paris Sér. I Math.* **331**(10), 783–790 (2000)
29. T. Lyons, Differential equations driven by rough signals. *Rev. Mat. Iberoam.* **14**(2), 215–310 (1998)
30. T. Lyons, Z. Qian, *System Control and Rough Paths*, *Oxford Mathematical Monographs* (Oxford University Press, Oxford, 2002)
31. T.J. Lyons, M. Caruana, T. Lévy, Differential equations driven by rough paths, volume 1908 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007. Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004, With an introduction concerning the Summer School by Jean Picard