Eikonal equations and pathwise solutions to fully non-linear SPDEs

Peter K Friz, Paul Gassiat, Pierre Louis Lions, Panagiotis E Souganidis

Eikonal equations and pathwise solutions to fully non-linear SPDEs

Peter K. Friz[1,*], Paul Gassiat[2,**], Pierre-Louis Lions[3] and Panagiotis E. Souganidis[4,***]

November 3, 2016

Abstract

We study the existence and uniqueness of the stochastic viscosity solutions of fully nonlinear, possibly degenerate, second order stochastic pde with quadratic Hamiltonians associated to a Riemannian geometry. The results are new and extend the class of equations studied so far by the last two authors.

Key words. Fully non-linear stochastic partial differential equations; eikonal equations, pathwise stability, rough paths.
AMS Class. Numbers. 35R99, 60H15

1 Introduction

The theory of stochastic viscosity solutions, including existence, uniqueness and stability, developed by two of the authors (Lions and Souganidis [4, 5, 6, 7, 8, 9]) is concerned with pathwise solutions to fully nonlinear, possibly degenerate, second order stochastic pde, which, in full generality, have the form

$$\begin{cases} du = F(D^2u, Du, u, x, t)dt + \sum_{i=1}^{d} H_i\left(Du, u, x\right) d\xi^i \ \text{ in } \ \mathbb{R}^N \times (0,T], \\ u = u_0 \text{ on } \mathbb{R}^N \times \{0\}; \end{cases} \tag{1}$$

here F is degenerate elliptic and $\xi = (\xi^1, \cdots, \xi^d)$ is a continuous path. A particular example is a d-dimensional Brownian motion, in which case (1) should be interpreted in the Stratonovich sense. Typically, $u \in \mathrm{BUC}(\mathbb{R}^N \times [0,T])$, the space of bounded uniformly continuous real-valued functions on $\mathbb{R}^N \times [0,T]$.

For the convenience of the reader we present a quick general overview of the theory: The Lions–Souganidis theory applies to rather general paths when $H = H(p)$ and, as established in [6, 9], there is a very precise trade off between the regularity of the paths and H. When $H = H(p,x)$ and $d = 1$, the results of [9] deal with general continuous, including Brownian paths, and the theory requires certain global structural conditions on H involving higher order (up to three) derivatives in x and p. Under similar conditions, Lions and Souganidis [10] have also established the wellposedness of (1) for $d > 1$ and Brownian paths. For completeness we note that, when ξ is smooth, for example C^1, (1) falls

within the scope of the classical Crandall-Lions viscosity theory – see, for example, Crandall, Ishii and Lions [2].
The aforementioned conditions are used to control the length of the interval of existence of smooth solutions of the so-called doubled equation

$$dw = (H(D_x w, x) - H(-D_y w, y))d\xi \ \text{ in } \ \mathbb{R}^N \times (t_0 - h_*, t_0 + h_*) \tag{2}$$

with initial datum

$$w(x, y, t_0) = \lambda |x - y|^2 \tag{3}$$

as $\lambda \to \infty$ and uniformly for $|x - y|$ appropriately bounded.
It was, however, conjectured in [9] that, given a Hamiltonian H, it may be possible to find initial data other than $\lambda|x - y|^2$ for the doubled equation, which are better adapted to H, thus avoiding some of the growth conditions. As a matter of fact this was illustrated by an example when $N = 1$.
In this note we follow up on the remark above about the structural conditions on H and identify a better suited initial data for (2) for the special class of quadratic Hamiltonians of the form

$$H(p, x) := (g^{-1}(x)p, p) = \sum_{i,j=1}^{N} g^{i,j}(x)\, p_i p_j, \tag{4}$$

which are associated to a Riemannian geometry in $\mathbb{R}^N$ and do not satisfy the conditions mentioned earlier, where

$$g = (g_{i,j})_{1 \le i,j \le N} \in C^2(\mathbb{R}^N; \mathcal{S}^N) \tag{5}$$

is positive definite, that is there exists $C > 0$ such that, for all $w \in \mathbb{R}^N$,

$$\frac{1}{C}|w|^2 \le \sum_{i,j} g_{i,j}(x)\, w^i w^j \le C|w|^2. \tag{6}$$

It follows from (4) and (6) that g is invertible and $g^{-1} = (g^{i,j})_{1 \le i,j \le N} \in C^2(\mathbb{R}^N; \mathcal{S}^N)$ is also positive definite; here $\mathcal{S}^N$ is the space of $N \times N$-symmetric matrices and (p, q) denotes the usual inner product of the vectors p, $q \in \mathbb{R}^N$. When dealing with (1) it is necessary to strengthen (5) and we assume that

$$g, g^{-1} \in C_b^2(\mathbb{R}^N; \mathcal{S}^N), \tag{7}$$

where $C_b^2(\mathbb{R}^N; \mathcal{S}^N)$ is the set of functions bounded in $C^2(\mathbb{R}^N; \mathcal{S}^N)$. Note that in this case (6) is implied trivially.
The distance $d_g(x, y)$ with respect to g of two points $x, y \in \mathbb{R}^N$ is given by

$$d_g(x, y) := \inf\left\{ \int_0^1 \frac{1}{2}(g(\gamma_t)\dot{\gamma}_t, \dot{\gamma}_t)^{1/2} dt : \gamma \in C^1([0,1], \mathbb{R}^N), \gamma_0 = x, \gamma_1 = y \right\},$$

and their associated "energy" is

$$e_g(x, y) := d_g^2(x, y) = \inf\left\{ \int_0^1 \frac{1}{4}(g(\gamma_t)\dot{\gamma}_t, \dot{\gamma}_t) dt : \gamma \in C^1([0,1], \mathbb{R}^N), \gamma_0 = x, \gamma_1 = y \right\}. \tag{8}$$

Note that, if $g = I$ the identity $N \times N$ matrix in $\mathbb{R}^N$, then $d_I(x, y) = \frac{1}{2}|x - y|$, the usual Euclidean distance, and $e_I(x, y) = \frac{1}{4}|x - y|^2$; more generally, (6) implies, with $C = c^2$ and for all $x, y \in \mathbb{R}^N$,

$$\frac{1}{2c}|x - y| \le d_g(x, y) \le \frac{1}{2}c|x - y|.$$

In addition, we assume that

$$\text{there exists } \Upsilon > 0 \text{ such that } e_g \in C^1(\{(x,y) \in \mathbb{R}^N \times \mathbb{R}^N : d_g(x,y) < \Upsilon\}); \tag{9}$$

in the language of differential geometry (9) is the same as to say that the manifold $(\mathbb{R}^N, g)$ has strictly positive injectivity radius. We remark that (7) is sufficient for (9) (see, for example, Proposition 4.3), though (far) from necessary.

We continue with some terminology and notation that we will need in the paper. We write I_N for the identity matrix in $\mathbb{R}^N$. A modulus is a nondecreasing, subadditive function $\omega : [0,\infty) \to [0,\infty)$ such that $\lim_{r\to 0} \omega(r) = \omega(0) = 0$. We write $u \in \mathrm{UC_g}\left(\mathbb{R}^N\right)$ if $|u(x) - u(y)| \leq \omega\left(d_g(x,y)\right)$ for some modulus ω, and, given $u \in \mathrm{UC_g}\left(\mathbb{R}^N\right)$, we denote by ω_u its modulus. When u is also bounded, we write $u \in \mathrm{BUC_g}\left(\mathbb{R}^N\right)$ and may take its modulus bounded. We denote by USC (resp. LSC) the set of upper- (resp. lower) semicontinuous functions in $\mathbb{R}^N$, and BUSC (resp. BLSC) is the set of bounded functions in USC (resp. LSC). For a bounded continuous function $u : \mathbb{R}^k \to \mathbb{R}$, for some $k \in \mathbb{N}$, and $A \subset \mathbb{R}^k$, $\|u\|_{\infty,A} := \sup_A |u|$. If $a, b \in \mathbb{R}$, then $a \wedge b := \min(a,b)$, $a_+ := \max(a,0)$ and $a_- := \max(-a,0)$. Given a modulus ω and $\lambda > 0$, we use the function $\theta : (0,\infty) \to (0,\infty)$ defined by

$$\theta(\omega,\lambda) := \sup_{r\geq 0}\{\omega(r) - \lambda r^2/2\}; \tag{10}$$

and observe that, in view of the assumed properties of the modulus,

$$\lim_{\lambda\to\infty} \theta(\omega;\lambda) = 0. \tag{11}$$

Finally, for $k \in \mathbb{N}$, $C_0^k([0,T];\mathbb{R}) := \{\zeta \in C^k([0,T];\mathbb{R}) : \zeta_0 = 0\}$ and, of for any two $\zeta, \xi \in C_0([0,T];\mathbb{R})$, we set

$$\Delta_T^+ := \max_{s\leq T}(\xi_s - \zeta_s) \geq 0 \quad \text{and} \quad \Delta_T^- := \max_{s\leq T}\{-(\xi_s - \zeta_s)\} \geq 0. \tag{12}$$

We review next the approach taken in [4, 5, 6, 7, 8, 9] to define solutions to (1). The key idea is to show that the solutions of the initial value problems with smooth paths, which approximate locally uniformly the given continuous one, form a Cauchy family in $BUC(\mathbb{R}^N \times [0,T])$ for all $T > 0$, and thus converge to a limit which is independent of the regularization. This limit is considered as the solution to (1). It follows that the solution operator for (1) is the extension in the class of continuous paths of the solution operator for smooth paths. Then [4, 5, 6, 7, 8, 9] introduced an intrinsic definition for a solution, called stochastic viscosity solution, which is satisfied by the uniform limit. Moreover, it was shown that the stochastic viscosity solutions satisfy a comparison principle and, hence, are intrinsically unique and can be constructed by the classical Perron's method (see [9] and [13] for the complete argument). The assumptions on the Hamiltonians mentioned above were used in these references to obtain both the Cauchy property and the intrinsic uniqueness.

To prove the Cauchy property the aforementioned references consider the solutions to (1) corresponding to two different smooth paths ζ_1 and ζ_2 and establish an upper bound for the sup-norm of their difference. The classical viscosity theory provides immediately such a bound, which, however, depends on the L^1-norm of $\dot\zeta_1 - \dot\zeta_2$. Such a bound is, of course, not useful since it blows up, as the paths approximate the given continuous path ξ. The novelty of the Lions-Souganidis theory is that it is possible to obtain far better control of the difference of the solutions based on the sup-norm of $\zeta_1 - \zeta_2$ at the expense of some structural assumptions on H. In the special case of (1) with $F = 0$ and H

independent of x, a sharp estimate was obtained in [9]. It was also remarked there that such bound cannot be expected to hold for spatially dependent Hamiltonians without additional restrictions.

In this note we take advantage of the very particular quadratic structure of H and obtain a local in time bound on the difference of two solutions with smooth paths. That the bound is local is due to the need to deal with smooth solutions of the Hamilton-Jacobi part of the equation. Quadratic Hamiltonians do not satisfy the assumptions in [9]. Hence, the results here extend the class of (1) for which there exists a well posed solution. The bound obtained is also used to give an estimate for the solutions to (1),(4) corresponding to different merely continuous paths as well as a modulus of continuity.

Next we present the results and begin with the comparison of solutions with smooth and different paths. Since the assumptions on the metric g are slightly stronger in the presence of the second order term in (1), we state two theorems. The first is for the first-order problem

$$\begin{cases} du - (g^{-1}(x)Du, Du)d\xi = 0 \text{ in } \mathbb{R}^N \times (0,T], \\ u = u_0 \text{ on } \mathbb{R}^N \times \{0\}, \end{cases} \tag{13}$$

and the second for (1) with H given by (4). Then we discuss the extension property and the comparison for general paths.

We first assume that we have smooth driving signals and estimate the difference of solutions. Since we are working with "classical" viscosity solutions, we write u_t and $\dot{\xi}_t$ in place of of du and $d\xi_t$.

Theorem 1.1. *Assume* (5), (6) *and* (9) *and let* $\xi, \zeta \in C_0^1([0,T];\mathbb{R})$ *and* $u_0, v_0 \in \mathrm{BUC_g}(\mathbb{R}^N)$. *If* $u \in \mathrm{BUSC}(\mathbb{R}^N \times [0,T])$ *and* $v \in \mathrm{BLSC}(\mathbb{R}^N \times [0,T])$ *are respectively viscosity sub- and super-solutions to*

$$u_t - (g^{-1}(x)Du, Du)\dot{\xi} \leq 0 \text{ in } \mathbb{R}^N \times (0,T] \quad u(\cdot,0) \leq u_0 \text{ on } \mathbb{R}^N,$$

and

$$v_t - (g^{-1}(x)Dv, Dv)\dot{\zeta} \geq 0 \text{ in } \mathbb{R}^N \times (0,T] \quad v(\cdot,0) \geq v_0 \text{ on } \mathbb{R}^N,$$

then, if

$$\Delta_T^+ + \Delta_T^- < \frac{1}{2\left(\|u_0\|_{\infty;\mathbb{R}^N} + \|v_0\|_{\infty;\mathbb{R}^N}\right)}\Upsilon^2, \tag{14}$$

$$\sup_{\mathbb{R}^N \times [0,T]} (u - v) \leq \sup_{\mathbb{R}^N} (u_0 - v_0) + \theta(\omega_{u_0} \wedge \omega_{v_0}, \frac{1}{\Delta_T^+}). \tag{15}$$

We consider now the second-order fully nonlinear equation (1) with quadratic Hamiltonians, that is the initial value problem

$$\begin{cases} du = F\left(D^2u, Du, u, x, t\right)dt + (g^{-1}(x)Du, Du)d\xi \quad \text{ in } \mathbb{R}^N \times 0, T], \\ u(\cdot,0) = u_0 \in \mathrm{BUC}(\mathbb{R}^N), \end{cases} \tag{16}$$

and introduce assumptions on F in order to have a result similar to Theorem 1.1.
In order to be able to have some checkable structural conditions on F, we find it necessary to replace (5) and (9) by the stronger conditions (7) and

$$\text{there exists } \Upsilon > 0 \text{ such that } D^2 d_g^2 \text{ is bounded on } \{(x,y) : d_g(x,y) < \Upsilon\}. \tag{17}$$

As far as $F \in \ C(\mathcal{S}^N \times \mathbb{R}^N \times \mathbb{R}^N \times [0,T]; \mathbb{R})$ is concerned we assume that it is degenerate elliptic, that is for all $X, Y \in \mathcal{S}^N$ and $(p, r, x, t) \in \mathbb{R}^N \times \mathbb{R}^N \times [0,T]$,

$$F(X, p, r, x, t) \leq F(Y, p, r, x, t) \ \text{ if } \ X \leq Y; \tag{18}$$

Lipschitz continuous in r, that is

$$\text{there exists } L > 0 \text{ such that } |F(X, p, r, x, t) - F(X, p, s, x, t)| \leq L|s - r|; \tag{19}$$

bounded in (x, t), in the sense that

$$\sup_{\mathbb{R}^N \times [0,T]} |F(0, 0, 0, \cdot, \cdot)| < \infty; \tag{20}$$

and uniformly continuous for bounded (X, p, r), that is, for any $R > 0$,

$$F \text{ is uniformly continuous on } M_R \times B_R \times [-R, R] \times \mathbb{R}^N \times [0,T], \tag{21}$$

where M_R and B_R are respectively the balls of radius R in $\mathcal{S}^N$ and $\mathbb{R}^N$.
Similarly to the classical theory of viscosity solutions, it is also necessary to assume something more about the joint continuity of F in X, p, x, namely that

$$\begin{cases} \text{for each } R > 0 \text{ there exists a modulus } \omega_{F,R} \text{ such that, for all } \alpha, \varepsilon > 0 \text{ and uniformly on} \\ \ t \in [0,T] \text{ and } r \in [-R, R], \\ F(X, \alpha D_x d_g^2(x,y), r, x, t) - F(Y, -\alpha D_y d_g^2(x,y), r, y, t) \leq \omega_{F,R}(\alpha d_g^2(x,y) + d_g(x,y) + \varepsilon), \\ \text{whenever } d_g(x,y) < \Upsilon \text{ and } X, Y \in \mathcal{S}^N \text{ are such that, for } A = D^2_{(x,y)} d_g^2(x,y), \\ -(\alpha^2 \varepsilon^{-1} + \|A\|) \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \leq \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq \alpha A + \varepsilon A^2. \end{cases} \tag{22}$$

Note that in the deterministic theory the above assumption is stated using the Euclidean distance. Here it is convenient to use d_g and as a result we find it necessary to strengthen the assumptions on the metric g.
To simplify the arguments below, instead of (19), we will assume that F monotone in r, that is

$$\text{there exists } \rho > 0 \text{ such that } \ F(X, p, r, x, t) - F(X, p, s, x, t) \geq \rho(s - r) \text{ whenever } s \geq r; \tag{23}$$

this is, of course, not a restriction since we can always consider the change $u(x,t) = e^{(L+\rho)t} v(x,t)$, which yields an equation for v with a new F satisfying (23) and path ξ' such that $\dot{\xi}'_t = e^{(L+\rho)t} \dot{\xi}_t$.
To state the result we introduce some additional notation. For $\gamma > 0$, we write

$$\tilde{\theta}(\omega; \gamma) := \sup_{r \geq 0} (\omega(r) - \frac{\gamma}{2} r), \tag{24}$$

and, for $\xi, \zeta \in C^1([0,\infty); \mathbb{R})$,

$$\Delta_T^{\gamma,+} := \sup_{t \leq T} \int_0^t e^{\gamma s} (\dot{\xi}_s - \dot{\zeta}_s) ds \ \text{ and } \ \Delta_T^{\gamma,-} := \sup_{t \leq T} [-\int_0^t e^{\gamma s} (\dot{\xi}_s - \dot{\zeta}_s) ds.] \tag{25}$$

Finally, for bounded $u_0, v_0 : \mathbb{R}^N \to \mathbb{R}$, let

$$K := \frac{2}{\rho} [\sup_{(x,t) \in \mathbb{R}^N \times [0,T]} F(0,0,0,x,t) + \|u_0\|_{\infty;\mathbb{R}^N} + \|v_0\|_{\infty;\mathbb{R}^N}].$$

We have:

Theorem 1.2. *Assume* (7), (17), (18), (19), (20), (21) *and* (22), *and let* $\xi, \zeta \in C_0^1([0,T];\mathbb{R})$ $u_0, v_0 \in \mathrm{BUC}(\mathbb{R}^N)$ *and* $T > 0$. *If* $u \in \mathrm{BUSC}_g(\mathbb{R}^N \times [0,T])$ *and* $v \in \mathrm{BLSC}(\mathbb{R}^N \times [0,T])$ *are respectively viscosity sub- and super-solutions of*

$$u_t - F(D^2u, Du, u, x, t) - (g^{-1}(x)Du, Du)\dot{\xi} \leq 0 \quad \text{in } \mathbb{R}^N \times (0,T] \qquad u(0,\cdot) \leq u_0 \quad \text{on } \mathbb{R}^N, \tag{26}$$

and

$$v_t - F(D^2v, Dv, v, x, t) - (g^{-1}(x)Dv, Dv)\dot{\zeta} \geq 0 \quad \text{in } \mathbb{R}^N \times (0,T] \qquad v(0,\cdot) \geq u_0 \quad \text{on } \mathbb{R}^N, \tag{27}$$

then, if $\Gamma := \{\gamma > 0 : \Delta_T^{\gamma,+} + \Delta_T^{\gamma,-} < \frac{\Upsilon^2}{4K} \text{ and } \Delta_T^{\gamma,-} < 1\}$,

$$\begin{cases} \sup_{\mathbb{R}^N \times [0,T]} (u - v) \leq \sup_{\mathbb{R}^N} (u_0 - v_0)_+ \\ + \inf_{\gamma \in \Gamma} [\theta(\omega_{u_0} \wedge \omega_{v_0}, \frac{1}{\Delta_T^{\gamma,+}}) + \frac{1}{\rho}\tilde{\theta}(\omega_{F,K}; \gamma) + \frac{1}{\rho}\omega_{F,K}(2(K(\Delta_T^{\gamma,+} + \Delta_T^{\gamma,-}))^{1/2}]. \end{cases} \tag{28}$$

Under their respective assumptions, Theorem 1.1 and Theorem 1.2 imply that, for paths $\xi \in C^1([0,\infty);\mathbb{R})$ and $g \in \mathrm{BUC}_g(\mathbb{R}^N)$, the initial value problems (13) and (16) have well-defined solution operators

$$\mathcal{S} : (u_0, \xi) \mapsto u \equiv \mathcal{S}^{\xi}[u_0].$$

The main interest in the estimates (15) and (28) is that they provide a unique continuous extension of this solution operator to all $\xi \in C([0,\infty);\mathbb{R})$. Since the proof is a simple reformulation of (15) and (28), we omit it.

Theorem 1.3. *Under the assumptions of Theorem 1.1 and Theorem 1.2, the solution operator* $\mathcal{S} : BUC(\mathbb{R}^N) \times C^1([0,\infty);\mathbb{R}) \to BUC(\mathbb{R}^N \times [0,T])$ *admits a unique continuous extension to* $\bar{\mathcal{S}} : BUC(\mathbb{R}^N) \times C([0,\infty);\mathbb{R}) \to BUC(\mathbb{R}^N \times [0,T])$. *In addition, for each* $T > 0$, *there exists a nondecreasing* $\Phi : [0,\infty) \to [0,\infty]$, *depending only on* T *and the moduli and* sup*-norms of* $u_0, v_0 \in \mathrm{BUC_g}$, *such that* $\lim_{r \to 0} \Phi(r) = \Phi(0) = 0$, *and, for all* $\xi, \zeta \in C([0,T];\mathbb{R})$,

$$\left\| \mathcal{S}^{\xi}[u_0] - \mathcal{S}^{\zeta}[v_0] \right\|_{\infty;\mathbb{R}^N \times [0,T]} \leq \|u_0 - v_0\|_{\infty;\mathbb{R}^N} + \Phi\left(\|\xi - \zeta\|_{\infty;[0,T]} \right). \tag{29}$$

We also remark that for both problems the proofs yield a, uniform in $t \in [0,T]$ and $\|\xi - \zeta\|_{\infty;[0,T]}$, estimate for $u(x,t) - v(y,t)$. Applied to the solutions of (13) and (16), this yields a (spatial) modulus of continuity which depends only on the initial datum, g and F but not ξ. This allows to see (as in [3] and [4, 5, 6, 7, 8, 9]) that $\mathcal{S}$ and then $\bar{\mathcal{S}}$ indeed takes values in $BUC(\mathbb{R}^N \times [0,T])$.

An example of F that satisfies the assumptions of Theorem 1.2 is the Hamilton-Jacobi-Isaacs operator

$$F(M,p,r,x,t) = \inf_{\alpha} \sup_{\beta} \left\{ \mathrm{tr}\left(\sigma_{\alpha\beta}\sigma_{\alpha\beta}^T(p,x)\, M \right) + b_{\alpha\beta}(p,x) - c_{\alpha\beta}(x) r \right\}, \tag{30}$$

with

$$\sigma, b, c \text{ bounded uniformly in } \alpha, \beta \tag{31}$$

such that, for some modulus ω and constant $C > 0$ and uniformly in α, β,

$$|\sigma_{\alpha\beta}(p,x) - \sigma_{\alpha\beta}(q,y)| \leq C(|x-y| + \frac{|p-q|}{|p|+|q|}), \tag{32}$$

and

$$|b_{\alpha\beta}(p,x) - b_{\alpha\beta}(q,y)| \leq \omega((1+|p|+|q|)|x-y| + |p-q|), \qquad |c_{\alpha\beta}(x) - c_{\alpha\beta}(y)| \leq \omega(|x-y|). \tag{33}$$

The paper is organized as follows. In the next section we prove Theorem 1.1. Section 3 is about the proof of Theorem 1.2. In the last section we state and prove a result showing that (7) implies (17) and verify that (30) satisfies the assumptions of Theorem 1.2.

2 The first order case: The proof of Theorem 1.1

We begin by recalling without proof the basic properties of the Riemannian energy e_g which we need in this paper. For more discussion we refer to, for example, [12] and the references therein.

Proposition 2.1. *Assume* (5), (6) *and* (9). *The Riemannian energy* e_g *defined by* (8) *is (locally) absolutely continuous, almost everywhere differentiable and satisfies the Eikonal equations*

$$(g^{-1}(y)D_y e_g, D_y e_g) = (g^{-1}(x)D_x e_g, D_x e_g) = e_g(x,y), \tag{34}$$

on a subset E *of* $\mathbb{R}^N \times \mathbb{R}^N$ *of full measure. Moreover,* $\{(x,y) \in \mathbb{R}^N \times \mathbb{R}^N : d_g(x,y) < \Upsilon\} \subset E$.

The next lemma, which is based on (34) and the properties of g, is about an observation which plays a vital role in the proofs.
To this end, for $x, y \in \mathbb{R}^N$, $\lambda > 0$ and $\xi, \zeta \in \mathrm{C}_0^1([0,T])$, we set

$$\Phi^\lambda(x,y,t) := \frac{\lambda e_g(x,y)}{1 - \lambda(\xi_t - \zeta_t)}. \tag{35}$$

Lemma 2.2. *Assume* (5), (6) *and* (9) *and choose* $\lambda < 1/\Delta_T^+$. *Then*

$$\frac{\lambda e_g}{1 + \lambda \Delta_T^-} \le \Phi^\lambda \le \frac{\lambda e_g}{1 - \lambda \Delta_T^+} \quad \text{on} \quad \mathbb{R}^N \times \mathbb{R}^N \times [0,T]. \tag{36}$$

In addition, in the set $\{(x,y) \in \mathbb{R}^N \times \mathbb{R}^N : d_g(x,y) < \Upsilon\}$, Φ^λ *is a classical solution of*

$$w_t = (g^{-1}(x)D_x w, D_x w)\dot{\xi} - (g^{-1}(y)D_y w, D_y w)\dot{\zeta}. \tag{37}$$

Proof. The first inequality is immediate from the definition (25) of $\Delta_T^\pm$. To prove (37), we observe that, in view of Proposition 2.1, we have

$$\begin{aligned}
\Phi_t^\lambda &= \frac{\lambda^2 e_g(x,y)}{(1 - \lambda(\xi_t - \zeta_t))^2}(\dot{\xi}_t - \dot{\zeta}_t), \\
(g^{-1}(x)D_x\Phi, D_x\Phi) &= \frac{\lambda^2}{(1 - \lambda(\xi_t - \zeta_t))^2}(g^{-1}(x)D_x e_g, D_x e_g) \\
&= \frac{\lambda^2 e_g(x,y)}{(1 - \lambda(\xi_t - \zeta_t))^2} \\
&= \frac{\lambda^2}{(1 - \lambda(\xi_t - \zeta_t))^2}(g^{-1}(y)D_y e_g, D_y e_g) \ = (g^{-1}(y)D_y\Phi, D_y\Phi),
\end{aligned}$$

and, hence, whenever $d_g(x,y) < \Upsilon$, the claim follows. □

The proof of Theorem 1.1 follows the standard procedure of doubling variables. The key idea introduced in [5] is to use special solutions of the Hamiltonian part of the equation as test functions in all the comparison type-arguments, instead of the typical $\lambda|x-y|^2$ used in the "deterministic" viscosity theory. As already pointed out earlier, in the case of general Hamiltonians, the construction of the test functions in [5] is tedious and requires structural conditions on H. The special form of the problem at hand, however, yields easily such tests functions, which are provided by Lemma 2.2.

Proof of Theorem 1.1. To prove (15) it suffices to show that, for all λ in a left-neighborhood of $(\Delta_T^+)^{-1}$, that is for $\lambda \in ((\Delta_T^+)^{-1} - \epsilon, (\Delta_T^+)^{-1})$ for some $\epsilon > 0$, and $x, y \in \mathbb{R}^N$ and $t \in [0, T]$,

$$\begin{aligned} u(x,t) - v(y,t) &\leq \Phi^\lambda(x,y,t) + \sup_{x',y' \in \mathbb{R}^N} \left(u_0(x') - v_0(y') - \lambda e_g(x',y')\right) \\ &\leq \Phi^\lambda(x,y,t) + \sup_{\mathbb{R}^N}(u_0 - v_0) + \sup_{x',y' \in \mathbb{R}^N} \left(v_0(x') - v_0(y') - \lambda e_g(x',y')\right). \end{aligned} \tag{38}$$

Indeed taking $x = y$ in (38) we find

$$\begin{aligned} u(x,t) - v(x,t) &\leq \sup_{\mathbb{R}^N}(u_0 - v_0) + \sup_{x',y' \in \mathbb{R}^N} \left(v_0(x') - v_0(y') - \lambda d^2(x',y')/2\right)_+ \\ &\leq \sup_{\mathbb{R}^N}(u_0 - v_0) + \sup_{r \geq 0} \left(\omega_{v_0}(r) - \lambda r^2/2\right)_+ \\ &= \sup_{\mathbb{R}^N}(u_0 - v_0) + \theta\left(\omega_{v_0}, \lambda\right), \end{aligned}$$

and we conclude letting $\lambda \to (\Delta_T^+)^{-1}$.
We begin with the observation that, since constants are solutions of (13),

$$u \ \leq \ \|u_0\|_{\infty;\mathbb{R}^N} \text{ and } \ -v \ \leq \ \|v_0\|_{\infty;\mathbb{R}^N}. \tag{39}$$

Next we fix $\delta, \alpha > 0$ and $0 < \lambda < (\Delta_T^+)^{-1}$ and consider the map

$$(x,y,t) \to u(x,t) - v(y,t) - \Phi^\lambda(x,y,t) - \delta\left(|x|^2 + |y|^2\right) - \alpha t,$$

which, in view of (39), achieves its maximum at some $(\hat{x}, \hat{y}, \hat{t}) \in \mathbb{R}^N \times \mathbb{R}^N \times [0,T]$ –note that below to keep the notation simple we omit the dependence of $(\hat{x}, \hat{y}, \hat{t})$ on λ, δ, α.
Let

$$\begin{aligned} M_{\lambda,\alpha,\delta} &:= \max_{\mathbb{R}^N \times \mathbb{R}^N \times [0,T]} u(x,t) - v(y,t) - \Phi^\lambda(x,y,t) - \delta\left(|x|^2 + |y|^2\right) - \alpha t \\ &= u(\hat{x},\hat{t}) - v(\hat{y},\hat{t}) - \Phi^\lambda(\hat{x},\hat{y},\hat{t}) - \delta\left(|\hat{x}|^2 + |\hat{y}|^2\right) - \alpha\hat{t}. \end{aligned}$$

The lemma below summarizes a number of important properties of $(\hat{x}, \hat{y}, \hat{t})$. Since the arguments in the proof are classical in the theory of viscosity solutions, see for example [1], [2], we omit the details.

Lemma 2.3. *Suppose that the assumptions of Theorem 1.1 hold. Then:*

$$\begin{cases} \textit{(i) for any fixed } \lambda, \alpha > 0, \ \lim_{\delta \to 0} \delta(|\hat{x}|^2 + |\hat{y}|^2) = 0, \\ \textit{(ii)} \ \ e_g(\hat{x},\hat{y}) \leq 2(1/\lambda + \Delta_T^-)(\|u\|_\infty + \|v\|_\infty), \\ \textit{(iii) if } d_g(\hat{x},\hat{y}) \leq \Upsilon, \ \textit{then} \\ (g^{-1}(\hat{x}) D_x \Phi^\lambda(\hat{x},\hat{y},\hat{t}), D_x \Phi^\lambda(\hat{x},\hat{y},\hat{t})) + \\ (g^{-1}(\hat{y}) D_y \Phi^\lambda(\hat{x},\hat{y},\hat{t}), D_y \Phi^\lambda(\hat{x},\hat{y},\hat{t})) \leq 2\lambda(1 - \lambda\Delta_T^+)^{-1}(\|u\|_\infty + \|v\|_\infty), \\ \textit{(iv)} \ \lim_{\delta \to 0} M_{\lambda,\alpha,\delta} = M_{\lambda,\alpha,0}. \end{cases} \tag{40}$$

Next we argue that, for any λ in a sufficiently small left-neighborhood of $(\Delta_T^+)^{-1}$, we have $d_g(\hat{x},\hat{y}) < \Upsilon$, which yields that the eikonal equation for e are valid at these points.
In view of the bound on $d_g^2(\hat{x},\hat{y}) = e_g(\hat{x},\hat{y})$ that follows from part (ii) of Lemma 2.3, it suffices to choose λ so that

$$2(1/\lambda + \Delta_T^-)\left(\|u\|_\infty + \|v\|_\infty\right) < \Upsilon^2.$$

Taking into account that we also need $\Delta_T^+ < 1/\lambda$, we are led to the condition

$$\Delta_T^+ + \Delta_T^- < \frac{1}{\lambda} + \Delta_T^- \leq \frac{1}{2\left(\|u\|_\infty + \|v\|_\infty\right)}\Upsilon^2;$$

and finding such λ is possible in view of (14).
If $\hat{t} \in (0,T]$, we use the inequalities satisfied by u and v in the viscosity sense, noting that to simplify the notation we omit the explicit dependence of derivatives of Φ on $(\hat{x},\hat{y},\hat{t})$, and we find, in view of Lemma 2.2 and the Cauchy-Schwarz's inequality,

$$\begin{aligned}
0 &\geq \Phi_t^\lambda + \alpha - (g^{-1}(\hat{x})(D_x\Phi^\lambda + 2\delta\hat{x}), (D_x\Phi^\lambda + 2\delta\hat{x}))\dot{\xi}_{\hat{t}} + (g^{-1}(\hat{y})(D_y\Phi^\lambda - 2\delta\hat{y}), (D_y\Phi^\lambda - 2\delta\hat{y}))\dot{\zeta}_{\hat{t}} \\
&\geq \alpha - \|\dot{\xi}\|_{\infty;[0,T]}\left(2\delta(g^{-1}(\hat{x})D_x\Phi^\lambda, D_x\Phi^\lambda)^{1/2}(g^{-1}(\hat{x})\hat{x},\hat{x})^{1/2} + \delta^2(g^{-1}(\hat{x})\hat{x},\hat{x})\right) \\
&\qquad - \|\dot{\zeta}\|_{\infty;[0,T]}\left(2\delta(g^{-1}(\hat{y})D_x\Phi^\lambda, D_y\Phi^\lambda)^{1/2}(g^{-1}(\hat{y})\hat{y},\hat{y})^{1/2} + \delta^2(g^{-1}(\hat{y})\hat{y},\hat{y})\right).
\end{aligned}$$

Using again Lemma 2.3 (i)-(iii), we can now let $\delta \to 0$ to obtain $\alpha \leq 0$, which is a contradiction.
It follows that, for all δ small enough, we must have $\hat{t} = 0$ and, hence,

$$M_{\lambda,\alpha,\delta} \leq (u_0(\hat{x}) - v_0(\hat{y}) - \lambda e(\hat{x},\hat{y})) \leq \sup_{\mathbb{R}^N}(u_0 - v_0) + \theta\left(\omega_{u_0} \wedge \omega_{v_0}, \lambda\right).$$

Letting first $\delta \to 0$ and then $\alpha \to 0$, concludes the proof of (38). □

3 The second-order case: The proof of Theorem 1.2

Since the proof of Theorem 1.2 is in many places very similar to that of Theorem 1.1, we omit arguments that follow along straightforward modifications.
In the next lemma we introduce the modified test functions, which here will depend on an additional parameter γ corresponding to a time exponential. Since its proof is similar to the one of Lemma 2.2, we omit it.

Lemma 3.1. *Fix $T, \lambda > 0, \gamma \geq 0$, $\xi, \zeta \in C^1([0,T];\mathbb{R}^N)$ with $\xi_0 = \zeta_0 = 0$ and assume that $\lambda\Delta_T^{\gamma,+} < 1$. Then*

$$\Phi^{\lambda,\gamma}(x,y,t) := \frac{\lambda e^{\gamma t}}{1 - \lambda\int_0^t e^{\gamma s}(\dot{\xi}_s - \dot{\zeta}_s)ds} e_g(x,y)$$

is a classical solution, in $\left\{(x,y) \in \mathbb{R}^N \times \mathbb{R}^N : d(x,y) < \Upsilon\right\} \times [0,T]$, of

$$w_t - \gamma w - (g^{-1}(x)D_x w, D_x w)\dot{\xi} + (g^{-1}(y)D_y w, D_y w)\dot{\zeta} = 0.$$

Next we specify the range of λ's we will use. We set

$$\bar{\lambda} := (\Delta_T^{\gamma,+})^{-1} \text{ and } \underline{\lambda} := \frac{4K}{\Upsilon - 4K\Delta_T^{\gamma,-}}, \tag{41}$$

and observe that, in view of our assumptions, we have $\bar{\lambda} > \underline{\lambda}$. We say that λ is admissible for fixed γ and α, if $\lambda \in (\underline{\lambda}, \bar{\lambda})$.
Also note that, if u, v, u_0, v_0, ξ, ζ and F are as in the statement of Theorem 1.2, then

$$\sup_{\mathbb{R}^N \times [0,T]} (u - v) \leq K. \tag{42}$$

For fixed $\delta > 0$ and λ admissible we consider the map

$$(x, y, t) \to u(x, t) - v(y, t) - \Phi^{\lambda,\gamma}(x, y, t) - \delta \left(|x|^2 + |y|^2 \right),$$

which, in view of (39), achieves its maximum at some $(\hat{x}, \hat{y}, \hat{t}) \in \mathbb{R}^N \times \mathbb{R}^N \times [0, T]$ –as before to keep the notation simple we omit the dependence of $(\hat{x}, \hat{y}, \hat{t})$ on λ, δ.
Let

$$\begin{aligned} M_{\lambda,\gamma,\delta} &:= \max_{(x,y,t) \in \mathbb{R}^N \times \mathbb{R}^N \times [0,T]} u(x, t) - v(y, t) - \Phi^{\lambda,\gamma}(x, y, t) - \delta(|x|^2 + |y|^2) \\ &= u(\hat{x}, \hat{t}) - v(\hat{y}, \hat{t}) - \Phi^{\lambda,\gamma}(\hat{x}, \hat{y}, \hat{t}) - \delta(|\hat{x}|^2 + |\hat{y}|^2). \end{aligned} \tag{43}$$

The following claim is the analogue of Lemma 2.3. As before when writing Φ and its derivatives we omit their arguments.

Lemma 3.2. *Under the assumptions of Theorem 1.2 and for λ admissible we have:*

$$\begin{cases} \textit{(i)}\ \lim_{\delta \to 0} \delta(|\hat{x}|^2 + |\hat{y}|^2) = 0, \quad \textit{(ii)}\ e_g(\hat{x}, \hat{y}) \leq 2K(\frac{1}{\lambda} + \Delta_T^{\gamma,-}), \quad \textit{(iii)}\ |D_x \Phi^{\lambda,\gamma}|^2 \leq \frac{2\lambda e^{\gamma T}}{1 - \lambda \Delta_T^{\gamma,+}} K \\ \textit{and} \quad \textit{(iv)}\ \lim_{\delta \to 0} M_{\lambda,\gamma,\delta} = M_{\lambda,\gamma,0}. \end{cases}$$

Proof of Theorem 1.2. If, for some sequence $\delta \to 0$, $\hat{t} = 0$, then

$$M_{\lambda,\gamma,0} = \lim_{\delta \to 0} M_{\lambda,\gamma,\delta} \leq u_0(\hat{x}) - v_0(\hat{y}) - \Phi^{\lambda,\gamma}(\hat{x}, \hat{y}, 0) \leq \| (u_0 - v_0)_+ \|_\infty + \theta(\omega_{u_0} \wedge \omega_{v_0}, \lambda). \tag{44}$$

We now treat the case where $\hat{t} \in (0, T]$ for all δ small enough.
Since, in view of Lemma 3.2(ii) and the assumptions (recalling that λ is admissible), the test-function $\Phi^{\lambda,\gamma}$ is smooth at $(\hat{x}, \hat{y}, \hat{t})$, it follows from the theory of viscosity solutions (see, for example, [2]) that

$$\begin{aligned} 0 \geq \Phi_t^{\lambda,\gamma} - F(X + 2\delta I, D_x \Phi^{\lambda,\gamma} + 2\delta \hat{x}, u(\hat{x}, \hat{t}), \hat{x}, \hat{t}) - (g^{-1}(\hat{x})(D_x \Phi^{\lambda,\gamma} + 2\delta \hat{x}), D_x \Phi^{\lambda,\gamma} + 2\delta \hat{x}) \dot{\xi}_{\hat{t}} \\ + F(Y - 2\delta I, -D_y \Phi^{\lambda,\gamma} - 2\delta \hat{y}, v(\hat{y}, \hat{t}), \hat{y}, \hat{t}) + (g^{-1}(\hat{y})(D_y \Phi^{\lambda,\gamma} + 2\delta \hat{y}), D_y \Phi^{\lambda,\gamma} + 2\delta \hat{y}) \dot{\zeta}_{\hat{t}}, \end{aligned} \tag{45}$$

where $X, Y \in \mathcal{S}^N$ are such that for a given $\varepsilon > 0$,

$$-\left(\frac{\hat{\alpha}^2}{\varepsilon} + \hat{\alpha} |D^2 e_g(\hat{x}, \hat{y})| \right) \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \leq \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq \hat{\alpha} D^2 e_g(\hat{x}, \hat{y}) + \varepsilon (D^2 e_g(\hat{x}, \hat{y}))^2 \tag{46}$$

and

$$\hat{\alpha} := \frac{\lambda e^{\gamma \hat{t}}}{1 - \lambda \int_0^{\hat{t}} e^{\gamma s} (\dot{\xi}_s - \dot{\zeta}_s) ds} = \frac{\Phi^{\lambda,\gamma}(\hat{x}, \hat{y}, \hat{t})}{e_g(\hat{x}, \hat{y})}. \tag{47}$$

Then, as in the usual proof of the comparison of viscosity solutions, combining (45) and (23), we get that

$$\rho(u(\hat{x}, \hat{t}) - v(\hat{y}, \hat{t}))_+ \leq (a) + (b) + (c) + (d), \tag{48}$$

where

$$(a) := -F(X, D_x\Phi^{\lambda,\gamma}, u(\hat{x},\hat{t}), \hat{x}, \hat{t}) + F(X + 2\delta, D_x\Phi^{\lambda,\gamma} + 2\delta\hat{x}, u(\hat{x},\hat{t}), \hat{x}, \hat{t}), \tag{49}$$

$$(b) := F(Y, -D_y\Phi^{\lambda,\gamma}, v(\hat{y},\hat{t}), \hat{x}, \hat{t}) - F(Y + 2\delta I, -D_y\Phi^{\lambda,\gamma} - 2\delta\hat{y}, v(\hat{y},\hat{t}), \hat{x}, \hat{t}), \tag{50}$$

$$(c) := \begin{cases} \Phi_t^{\lambda,\gamma} + \gamma\Phi^{\lambda,\gamma} + (g^{-1}(\hat{x})(D_x\Phi^{\lambda,\gamma} + 2\delta\hat{x}), D_x\Phi^{\lambda,\gamma} + 2\delta\hat{x})\dot{\xi}_{\hat{t}} \\ -(g^{-1}(\hat{y})(D_y\Phi^{\lambda,\gamma} - 2\delta\hat{y}), D_y\Phi^{\lambda,\gamma} - 2\delta\hat{y})\dot{\zeta}_{\hat{t}}, \end{cases} \tag{51}$$

and

$$(d) := -\gamma\Phi^{\lambda,\gamma} + F(X, D_x\Phi^{\lambda,\gamma}, u(\hat{x},\hat{t}), \hat{x}, \hat{t}) - F(Y, -D_y\Phi^{\lambda,\gamma}, v(\hat{y},\hat{t}), \hat{y}, \hat{t}). \tag{52}$$

Since (17) and (46) imply that X and Y stay bounded, in view of (21), we get $\limsup_{\delta\to 0}((a)+(b)) = 0$.

Moreover, the quadratic form of the equation satisfied by $\Phi^{\lambda,\gamma} = \hat{\alpha}e_g$ gives

$$(c) \le C\delta|D_{\hat{x}}\Phi^{\lambda,\gamma}|(|\hat{x}| + |\hat{y}|)\left(\|\dot{\xi}\|_{\infty;[0,T]} + \|\dot{\zeta}\|_{\infty;[0,T]}\right),$$

and using Lemma 3.2 (i),(iii) we find $\lim_{\delta\to 0}(c) = 0$.
For the last term, note that Lemma 3.2 (ii) and (22) yield, always at the point $(\hat{x}, \hat{y}, \hat{t})$,

$$\begin{aligned} (d) &= -\gamma\hat{\alpha}e_g + F(X, \hat{\alpha}D_x e_g, u(\hat{x},\hat{t}), \hat{x}, \hat{t}) - F(Y, -\hat{\alpha}D_y e_g, v(\hat{y},\hat{t}), \hat{y}, \hat{t}) \\ &\le -\frac{\gamma\hat{\alpha}}{2}d^2(\hat{x},\hat{y}) + \omega_{F,K}\left(d(\hat{x},\hat{y}) + \hat{\alpha}d^2(\hat{x},\hat{y}) + \varepsilon\right) \\ &\le -\frac{\gamma\hat{\alpha}}{2}d^2(\hat{x},\hat{y}) + \omega_{F,K}(\hat{\alpha}d^2(\hat{x},\hat{y})) + \omega_{F,K}(d(\hat{x},\hat{y}) + \omega_{F,K}(\varepsilon) \\ &\le \tilde{\theta}(\omega_{F,K}, \gamma) + \omega_{F,K}(2(K(\frac{1}{\lambda} + \Delta^{\gamma;-})^{1/2}) + \omega_{F,K}(\varepsilon). \end{aligned}$$

Combining the last four estimates and (44) and letting $\varepsilon \to 0$ we find that, for all $\lambda \in (\underline{\lambda}, \bar{\lambda})$

$$\begin{aligned} u(x,t) - v(x,t) &\le M_{\lambda,\gamma,0} = \lim_{\delta\to 0} M_{\lambda,\gamma,\delta} \\ &\le \left\|(u_0 - v_0)_+\right\|_\infty + \theta(\omega_{u_0} \wedge \omega_{v_0}, \lambda) + \frac{1}{\rho}\tilde{\theta}(\omega_{F,K}, \gamma) + \frac{1}{\rho}\omega_{F,K}(2(K(\frac{1}{\lambda} + \Delta_T^{\gamma;-})^{1/2}). \end{aligned}$$

Letting $\lambda \to \bar{\lambda}$ and using the continuity of θ in the last argument, we finally obtain that, for all $\gamma \in \Gamma$,

$$u - v \le \left\|(u_0 - v_0)_+\right\|_\infty + \theta\left(\omega_{u_0} \wedge \omega_{v_0}, (\Delta_T^{\gamma;+})^{-1}\right) + \frac{1}{\rho}\tilde{\theta}(\omega_{F,K}; \gamma) + \frac{1}{\rho}\omega_{F,K}(2(K(\frac{1}{\lambda} + \Delta_T^{\gamma;-})^{1/2}).$$

□

4 The properties of the geodesic energy and the assumptions of Theorem 1.2

In this section we prove that C_b^2-bounds on g and g^{-1} imply (17) and verify that the F in (30), if (32) and (33) hold, satisfies the assumptions of Theorem 1.2.
We begin with the former.

Proposition 4.1. *Assume* (7). *Then there exists* $\Upsilon > 0$ *such that, in the set* $\{(x,y) : d_g(x,y) < \Upsilon\}$, e_g *is twice continuously differentiable and* (17) *is satisfied with bounds depending only on appropriate norms of* g, g^{-1}.

Proof. We begin by recalling some basic facts concerning geodesics and distances.
For each fixed point x, there is a unique geodesic with starting velocity $v = \dot{\gamma}(0)$ given by $\gamma_t = X_t$, where $(X,P)_{t\geq 0}$ is the solution to the characteristic equations

$$\begin{cases} \dot{X}_s = 2g^{-1}(X_s)P_s \quad & X_0 = x, \\ \dot{P}_s = -(Dg^{-1}(X_s)P_s, P_s) \quad & P_0 = p = \frac{1}{2}g(x)v. \end{cases} \tag{53}$$

Equivalently $(\gamma)_{t\geq 0}$ satisfies the second order system of ode

$$\ddot{\gamma}_t + \Gamma^k_{ij}(\gamma_t)\,\dot{\gamma}^i_t\dot{\gamma}^j_t = 0 \quad \gamma_0 = x,\ \dot{\gamma}_0 = v \tag{54}$$

with

$$\Gamma^k_{ij} := g^{k\ell}\left(\partial_i g_{\ell j} + \partial_j g_{\ell i} - \partial_\ell g_{ij}\right). \tag{55}$$

It is easy to see that (53) has a global solution $(X,P)_{t\geq 0}$, since, in view of (6) and (7) as well as the invariance of the flow, we have, for $t \geq 0$,

$$|P_t| \approx (g^{-1}(X_t)P_t, P_t) = (g^{-1}(X_0)P_0, P_0) \approx |p|^2. \tag{56}$$

As a consequence, the projected end-point map $E_x(p) := X_1(x,p)$ is well-defined for any p.
We note that the energy along a geodesic γ emerging from $\gamma_0 = x$ has a simple expression in terms of $p = P_0$ or $v = \dot{\gamma}_0$. Indeed, invariance of the Hamilonian $H(x,p) = (g^{-1}(x)p, p)$ under the flow yields

$$|\dot{\gamma}_0|^2_g := (g(x)\,v, v) = 4\,H(x,p) = 4\int_0^1 H(X_t,P_t)dt = \int_0^1 (g(\gamma_t)\,\dot{\gamma}_t, \dot{\gamma}_t)dt. \tag{57}$$

It is a basic fact that distance minimizing curves (geodesics) are also energy minimizing. Indeed, given x, y, (53) and equivalently (54), are the first-order optimality necessary conditions for these minimization problems. Hence, in view of (57),

$$\begin{aligned} e_g(x,y) &= \inf\left\{\frac{1}{4}|\dot{\gamma}_0|^2_g : \gamma \text{ satisfies (54) with } \gamma_0 = x \text{ and } \gamma_1 = y\right\} \\ &= \inf\left\{H(X_0,P_0) : (X,P) \text{ satisfies (53) with } X_0 = x \text{ and } X_1 = y\right\}. \end{aligned}$$

A standard compactness argument implies the existence of at least one geodesic connecting two given points x, y. In general, however, more than one geodesic from x to y may exist, each determined by its initial velocity $\dot{\gamma}_0 = v$ or equivalently $P_0 = p = \frac{1}{2}g(x)v$, upon departure from x.
It turns out that, for y close to x, there exists exactly one geodesic. Indeed, if $g^{-1} \in C^2$, it is clear from (53) that $E_x = (p \mapsto X_1(x,p))$ has C^1-dependence in p. Since $D_pE_x(p)$ is non-degenerate in a neighborhood of $p = 0$, it follows from the inverse function theorem that, for y close enough to x, one can solve $y = E_x(p)$ uniquely for $p = E_x^{-1}(y)$ with C^1-dependence in y. Hence

$$e_g(x,y) = H(x,p) = H(x, E_x^{-1}(y)). \tag{58}$$

The gradient $D_x e_g(x,y)$ points in the direction of maximal increase of $x \mapsto e(x,y)$. Since $E_x^{-1}(y) = p$ is precisely the co-velocity of X at $X_0 = x$ and X is the geodesic from x to $X_1 = y$, it follows that

$$D_x e_g(x,y) = -E_x^{-1} y = -p. \tag{59}$$

This easily implies that, for points x, y close enough, the energy has continuous second derivatives. Indeed, existence of continuous mixed derivatives $D^2_{xy} e_g$ follows immediately from the C^1-regularity of $E_x^{-1} = E_x^{-1}(y)$. Concerning $D^2_{xx} e_g$ (and by symmetry $D^2_{yy} e_g$) we set $p = p_0$ above and note that by exchanging the roles of x, y, we have $D_y e_g(x,y) = -E_y^{-1}(x) = P_1(x, p_0) =: p_1$ and so, from (53),

$$D_y e_g(x,y) + D_x e_g(x,y) = p_1 - p_0 = -\int_0^1 H_x(X_s, P_s)\, ds. \tag{60}$$

The existence and continuity of $D^2_{xx} e_g$ is then clear, since the right-hand side above has C^1-dependence in x as is immediate from (53) and $g^{-1} \in C^2$.
Since all the above considerations have been so far local, it is necessary to address the (global) question of regularity in a strip around the diagonal $\{x = y\}$. For this we need to control $\alpha_1 := D_p E_x(p) = D_p X_1(x,p)$. We do this by considering the tangent flow

$$(\alpha_t, \beta_t) := (D_p X_t(x,p), D_p P_t(x,p)),$$

which solves the matrix-valued linear ode

$$\begin{cases} \dot\alpha_s = 2Dg^{-1}(X_s)P_s\alpha_s + 2g^{-1}(X_s)\beta_s, \qquad \alpha_0 = 0, \\ \dot\beta_s = -(D^2 g^{-1}(X_s)P_s, P_s)\alpha_s - 2\, Dg^{-1}(X_s)P_s\beta_s, \qquad \beta_0 = I. \end{cases} \tag{61}$$

We now argue that, uniformly in x,

$$D_p E_x(p) = \alpha_1 \approx 2\, g^{-1}(x).$$

It follows that $D_p E_x$ is non-degenerate, again uniformly in x. Indeed if $\alpha_0 = 0$, whenever p is small, $X_\cdot \approx x, \beta_\cdot \approx I$ and we have

$$\begin{pmatrix} \dot\alpha \\ \dot\beta \end{pmatrix} = \begin{pmatrix} \text{(small)} & 2g^{-1}(X_\cdot) \\ \text{(small)} & \text{(small)} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

Next we prove the above claim. First, it follows from (56) that if p is small, then $P_t(x,p)$ stays small, over, for example, a unit time interval $[0,1]$. Moreover, since $\dot X = 2\, g^{-1}(X)P$, the boundedness of g^{-1} yields that the path $X_\cdot(x,p)$ also stays close to $X_0 = x$, again uniformly on $[0,1]$. Furthermore, the C^1- and C^2-bounds on g^{-1} yield that the matrices $Dg^{-1}(X_s)P_s$ and $(D^2 g^{-1}(X_s)P_s, P_s)$ are small along (X_s, P_s), while $2g^{-1}(X_s)$ is plainly bounded. This implies that $\alpha_\cdot$ and $\beta_\cdot$ will stay bounded. Then $\dot\beta_\cdot$ will be small, and, hence, $\beta_\cdot$ will be close to $\beta_0 = I$, uniformly on $[0,1]$. In turn, $\dot\alpha_s$ is the sum of a small term plus $2\ g^{-1}(X_s)\beta_s \approx 2g^{-1}(x)$. In other words,

$$\dot\alpha_s = 2g^{-1}(x) + \Theta_s(x,p),$$

where

$$\lim_{\delta\to 0} \sup_{|p|\le\delta} \sup_{s\in[0,1]} \sup_{x\in\mathbb{R}^N} |\Theta_s(x,p)| = 0.$$

Since $D_pE_x(p) = \alpha_1 = \int_0^1 \dot{\alpha}_s ds$, it follows that, there exists some $\delta > 0$, which can be taken proportional to M^{-4} where $M = 1 + \|g\|_{C^2} + \|g^{-1}\|_{C^2}$, such that

$$|D_pE_x(p) - 2g^{-1}(x)| \leq \|g\|_\infty^{-1} \text{ for all } x, p \in \mathbb{R}^N, |p| \leq \delta.$$

Note that the choice of the constant $\|g\|_\infty^{-1} > 0$ on the right-hand side guarantees that $D_pE_x(p)$ remains non-degenerate, uniformly in x.

It follows by the inverse function theorem that $p \mapsto E_x(p)$ is a diffeomorphism from B_δ onto a neighbourhood of x. We claim that this neighbourhood contains a ball of radius $\Upsilon > 0$, which may be taken independent of x.

For this we observe that, with $p = E_x^{-1}(y)$, (58) yields

$$d_g(x, y) = \sqrt{e_g(x, y)} = \sqrt{H(x, p)} = \sqrt{(g^{-1}(x)p, p)}.$$

Hence it suffices to choose $\Upsilon > 0$ small enough so that $(g^{-1}(x)p, p) \leq \Upsilon^2$ implies $|p| \leq \delta$, an obvious choice being $\Upsilon = \delta/c$, where c^2 is the ellipticity constant of g.

At last, we note that (59), in conjunction with the just obtained quantitative bounds, implies that D^2e_g is bounded on $\left\{(x, y) \in \mathbb{R}^N \times \mathbb{R}^N : d_g(x, y) < \Upsilon\right\}$. Indeed, with $p = E_x^{-1}(y)$, we have $-D^2_{xy}e_g(x, y) = D_yE_x^{-1}(y) = (D_pE_x(p))^{-1} \approx \frac{1}{2}g(x)$ which readily leads to bounds of the second mixed derivatives, uniformly over x, y of distance at most Υ. Similar uniform bounds for $D^2_{xx}e_g$ (and then $D^2_{yy}e_g$) are obtained by differentiating (60) with respect to x and estimating the resulting right-hand side. □

The comparison proofs in the viscosity theory typically employ quadratic penalty function $\phi(x, y) = \frac{1}{2}|x - y|^2$ and make use of (trivial) identities such as

$$D_x\phi + D_y\phi = (x - y) + (y - x) = 0$$

and

$$\begin{pmatrix} D^2_{xx}\phi & D^2_{xy}\phi \\ D^2_{yx}\phi & D^2_{yy}\phi \end{pmatrix} = \begin{pmatrix} I & -I \\ -I & I \end{pmatrix}, \quad (p,\, q) \begin{pmatrix} D^2_{xx}\phi & D^2_{xy}\phi \\ D^2_{yx}\phi & D^2_{yy}\phi \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} \leq |p - q|^2 .$$

To see what one can expect in more general settings, consider first the case of g obtained from the Euclidean metric, written in different coordinates, say $x = \Psi^{-1}(\tilde{x})$, in which case, we have

$$e_g(x, y) = |\Psi(x) - \Psi(y)|^2 .$$

If Ψ is bounded in C^2, it is immediate that

$$\begin{aligned} D_xe + D_ye &= (\Psi(x) - \Psi(y))D\Psi(x) + (\Psi(y) - \Psi(x))D\Psi(y) \\ &= (\Psi(x) - \Psi(y))(D\Psi(y) - D\Psi(y)), \end{aligned}$$

and, hence,

$$|D_xe + D_ye| \lesssim |x - y|^2 \tag{62}$$

and, similarly,

$$(p,\, q) \begin{pmatrix} D^2_{xx}e & D^2_{xy}e \\ D^2_{yx}e & D^2_{yy}e \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} \lesssim |p - q|^2 + |x - y|^2 .$$

Unfortunately no such arguments work in the case of general Riemannian metric, since, in general, there is no change of variables of the form $|\Psi(x) - \Psi(y)|$ that reduces $d_g(x, y)$ to a Euclidean distance.

The next two propositions provide estimates that can be used in the comparison proofs in place of the exact identities above.

Proposition 4.2. *Assume* (7). *Then there exists* $\Upsilon > 0$ *such that whenever* $d_g(x,y) < \Upsilon$,

$$|D_x e_g + D_y e_g| \leq L\,|x-y|^2 \tag{63}$$

with a constant L *that depends only on the* C^1*-bounds of* g, g^{-1}.

Proof. As pointed out in the proof of Proposition 4.1, for all $(x,y) : d_g(x,y) < \Upsilon$,

$$D_y e_g\,(x,y) + D_x e_g\,(x,y) = -\int_0^1 (Dg^{-1}(X_s)P_s, P_s)ds.$$

Using that $g^{-1} \in C^1$ and the fact that g is uniformly comparable to the Euclidean metric we get

$$\begin{aligned} |D_y e\,(x,y) + D_x e\,(x,y)| &\leq \|g^{-1}\|_{C^1} \int_0^1 |P_s|^2\,ds \\ &\leq \|g^{-1}\|_{C^1}\|g\|_\infty \int_0^1 (g^{-1}\,(X_s)\,P_s, P_s)ds \end{aligned}$$

and, hence, thanks to invariance of the Hamiltonian under the flow and (58),

$$|D_y e\,(x,y) + D_x e\,(x,y)| \lesssim \int_0^1 H(X_s,P_s)ds = H\,(X_0,P_0) = e_g\,(x,y)\,.$$

Using again that g is uniformly comparable to the Euclidean metric the proof is finished.

□

The following claim applies in particular under condition (7), which, in view of Propostion 4.1, implies C^2-regularity of the energy near the diagonal. The proof is based on an argument in a forthcoming paper by the last two authors [11].

Proposition 4.3. *Assume there exists* $\Upsilon > 0$ *such that, in the set* $\{(x,y) : d_g(x,y) < \Upsilon\}$, e_g *is twice continuously differentiable. Then, whenever* $d_g(x,y) < \Upsilon$,

$$\begin{pmatrix} D^2_{xx}e_g & D^2_{xy}e_g \\ D^2_{yx}e_g & D^2_{yy}e_g \end{pmatrix} \leq L \begin{pmatrix} I_N & -I_N \\ -I_N & I_N \end{pmatrix} + L\,|x-y|^2\,I_{2N}$$

with a constant L *that only depends on the* C^2*-bounds of* g, g^{-1}.

Proof. In view of the assumed C^2-regularity of the energy, we find that, as $\varepsilon \to 0$,

$$\begin{cases} \varepsilon^2 \left((p,q) \begin{pmatrix} D^2_{xx}e_g & D^2_{xy}e_g \\ D^2_{yx}e_g & D^2_{yy}e_g \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} \right) \\ = e_g\,(x+\varepsilon p, y+\varepsilon q) + e_g\,(x-\varepsilon p, y-\varepsilon w) - 2e_g\,(x,y) \end{cases}$$

We estimate the second-order difference on the right-hand side, keeping $v = \varepsilon p, w = \varepsilon q$ fixed.
Let $(\gamma_t)_{t\in[0,1]}$ be a geodesic connecting x, y, parametrized at constant speed (in the metric g) so that

$$|\dot\gamma_t| \leq c\,(g\,(\gamma_t)\,\dot\gamma_t, \dot\gamma_t)^{1/2} = 2cd_g\,(x,y) \leq C\,|x-y|\,, \tag{64}$$

and, for $\Delta := w - v$, define the paths

$$\gamma_t^{\pm} := \gamma_t \pm (v + t\Delta)$$

which connect $x \pm v$ to $y \pm w$. Then

$$\begin{aligned} & e\left(x+v, y+w\right)+e\left(x-v, y-w\right)-2e\left(x, y\right) \\ \leq & \int_0^1 \frac{1}{4}\left(g\left(\gamma_t^+\right)\dot{\gamma}_t^+, \dot{\gamma}_t^+\right) dt+\int_0^1 \frac{1}{4}\left(g\left(\gamma_t^-\right)\dot{\gamma}_t^-, \dot{\gamma}_t^-\right) dt-2\int_0^1 \frac{1}{4}\left(g\left(\gamma_t\right)\dot{\gamma}_t, \dot{\gamma}_t\right) dt \\ = & \frac{1}{4}\int_0^1 \left[\left(g\left(\gamma_t^+\right)\left(\dot{\gamma}_t+\Delta\right), \dot{\gamma}_t+\Delta\right)+\left(g\left(\gamma_t^-\right)\left(\dot{\gamma}_t-\Delta\right), \dot{\gamma}_t-\Delta\right) dt-2\left(g\left(\gamma_t\right)\dot{\gamma}_t, \dot{\gamma}_t\right)\right] dt \end{aligned}$$

Using the C^2-regularity of g, writing $\delta = v + t\Delta$, and noting that that $|\delta| \leq |v| + |w|$, we find

$$g\left(\gamma_t^{\pm}\right) = g\left(\gamma_t\right) \pm \left(Dg\left(\gamma_t\right), \delta\right) + \frac{1}{2}\left(D^2 g\left(\gamma_t\right)\delta, \delta\right) + o\left(|v|^2 + |w|^2\right).$$

Collecting terms (in $g, Dg, D^2 g$) then leads to

$$e\left(x+v, y+w\right)+e\left(x-v, y-w\right)-2e\left(x, y\right) \leq \frac{1}{4}\int_0^1 \left[(i)+(ii)+(iii)+(E)\right] dt$$

where

$$\begin{cases} (i) := 2\left(g\left(\gamma_t\right)\Delta, \Delta\right) \\ (ii) := 4\left(\left(Dg\left(\gamma_t\right), \delta\right)\dot{\gamma}_t, \Delta\right) \\ (iii) := \left(\left(D^2 g\left(\gamma_t\right)\delta, \delta\right)\dot{\gamma}_t, \dot{\gamma}_t\right)+\left(\left(D^2 g\left(\gamma_t\right)\delta, \delta\right)\Delta, \Delta\right) \end{cases}$$

It is immediate that,

$$\begin{cases} |(i)| \leq 2\left\|g\right\|_\infty |w-v|^2 \\ |(ii)| \leq 4C\left\|Dg\right\|_\infty \left(|v|+|w|\right)|x-y|\,|w-v| \\ |iii| \leq \left\|D^2 g\right\|_\infty \left(|v|^2+|w|^2\right)\left(C^2|x-y|^2+|w-v|^2\right). \end{cases}$$

Moreover, expanding g, as $v, w \to 0$, we find

$$(E) = (|\dot{\gamma}_t|^2 + |\Delta|^2) o\left(|v|^2+|w|^2\right) = (C^2|x-y|^2 + |w-v|^2) o\left(|v|^2+|w|^2\right).$$

With $v = \varepsilon p, w = \varepsilon q$ all terms above are of order $O\left(\varepsilon^2\right)$, with the exception of the second term contributing to (iii) and the error term (E) which are actually $o\left(\varepsilon^2\right)$, and hence negligible as $\varepsilon \to 0$. Indeed

$$\begin{aligned} & e(x+\varepsilon p, y+\varepsilon q)+e(x-\varepsilon p, y-\varepsilon w)-2e(x, y) \\ \leq & \frac{\varepsilon^2}{4}[2\|g\|_\infty |p-q|^2 + 4C\|Dg\|_\infty (|p|+|q|)|x-y||p-q| \\ & +\|D^2 g\|_\infty (|p|^2+|q|^2)(C^2|x-y|^2+O(\varepsilon^2)) + (|x-y|^2+O(\varepsilon^2))o(1)]. \end{aligned}$$

Using again the Cauchy- Schwarz inequality to handle the middle term in the estimate, we find, for some $K > 0$ that only depends on g,

$$\begin{aligned} & (p, q)\begin{pmatrix} D^2_{xx} e_g & D^2_{xy} e_g \\ D^2_{yx} e_g & D^2_{yy} e_g \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix} \\ \leq & \left[\frac{1}{2}\left\|g\right\|_\infty |p-q|^2 + C\left\|Dg\right\|_\infty \left(|p|+|q|\right)|x-y|\,|p-q| + \frac{C^2}{4}\left\|D^2 g\right\|_\infty \left(|p|^2+|q|^2\right)|x-y|^2\right] \\ \leq & K|p-q|^2 + K|x-y|^2\left(|p|^2+|q^2|\right). \end{aligned}$$

□

We conclude by checking that the assumptions of Theorem 1.2 are satisfied for F of Hamilton-Jacobi-Isaacs type. Note that condition (7) is valid, as demanded by that theorem.

Proposition 4.4. *Let F be given by* (30)*, satisfying* (32) *and* (33)*. Then F satisfies* (18)-(22)*.*

Proof. Since it is clear that all assumptions are stable under taking sup and inf, we will only treat the quasilinear case

$$F(M,p,r,x) = \mathrm{Tr}(a(p,x)M) + b(p,x) - c(x)r,$$

and we concentrate on (22), since the others are obvious.
With t, x, y, r, α, X, Y taken as in the statement of the assumption, we have

$$\begin{aligned} & F(X, \alpha D_x d_g^2(x,y), r, x, t) - F(Y, -\alpha D_y d_g^2(x,y), r, x, t) \\ = \;& b(p,x) - b(q,y) + \mathrm{Tr}\left(\sigma\sigma^T(p,x)X - \sigma\sigma^T(q,y)Y\right) + (c(x)-c(y))r, \end{aligned}$$

where for simplicity we write $p := \alpha D_x d_g^2(x,y)$ and $q := -\alpha D_y d_g^2(x,y)$.
Noting that Proposition 4.2 yields

$$|p-q| \leq \alpha K |x-y|^2,$$

and using the eikonal equation for $e_g = d_g^2$, for some $C > 0$ we find

$$\alpha C^{-1}|x-y| \leq |p| + |q| \leq \alpha C |x-y|.$$

The assumptions on b also give

$$\begin{aligned} b(p,x) - b(q,y) &\leq \omega\left((1+|p|+|q|)\,|x-y| + |p-q|\right) \\ &\leq \omega\left((1+C\alpha|x-y|)\,|x-y| + K\alpha|x-y|^2\right). \end{aligned}$$

For the second order term, we use Proposition 4.3 and get

$$\begin{aligned} \mathrm{Tr}\left(\sigma\sigma^T(p,x)X - \sigma\sigma^T(q,y)Y\right) &= (\sigma(p,x),\, \sigma(q,y)) \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \begin{pmatrix} \sigma(p,x) \\ \sigma(q,y) \end{pmatrix} \\ &\leq \alpha\,(\sigma(p,x),\, \sigma(q,y)) \begin{pmatrix} D^2_{xx}e & D^2_{xy}e \\ D^2_{yx}e & D^2_{yy}e \end{pmatrix} \begin{pmatrix} \sigma(p,x) \\ \sigma(q,y) \end{pmatrix} \\ &\quad + \varepsilon \left\| \begin{pmatrix} D^2_{xx}e & D^2_{xy}e \\ D^2_{yx}e & D^2_{yy}e \end{pmatrix} \begin{pmatrix} \sigma(p,x) \\ \sigma(q,y) \end{pmatrix} \right\|^2 \\ &\leq K\alpha(|\sigma(p,x) - \sigma(q,y)|^2 + |x-y|^2) + K\varepsilon \\ &\leq K'\alpha\left(|x-y|^2 + \frac{|p-q|^2}{(|p|+|q|)^2}\right) + K'\varepsilon \\ &\leq K''\left(\alpha|x-y|^2 + \varepsilon\right). \end{aligned}$$

□

References

[1] G. Barles, *Solutions de viscosité des équations de Hamilton-Jacobi*. Springer, 2004.

[2] M. Crandall, H. Ishii and P.-L. Lions (1992) : "User's Guide to Viscosity Solutions of Second Order Partial Differential Equations", *Bull. Amer. Math. Soc.*, **27**, 1-67.

[3] J. Diehl, P. Friz, H. Oberhauser. Regularity theory for RPDEs and parabolic comparison revisited. Stochastic Analysis and Applications (Ed. Dan Crisan et al), Springer Proceedings in Mathematics & Statistics, Vol. 100, 2014

[4] P.-L. Lions and P. E. Souganidis. Viscosity solutions of fully nonlinear stochastic partial differential equations. Sūrikaisekikenkyūsho Kōkyūroku, (1287):58–65, 2002. Viscosity solutions of differential equations and related topics (Japanese) (Kyoto, 2001).

[5] P.-L. Lions and P. E. Souganidis. Fully nonlinear stochastic partial differential equations. C. R. Acad. Sci. Paris Sér. I Math., 326(9):1085–1092, 1998.

[6] P.-L. Lions and 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.

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

[8] P.-L. Lions and 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.

[9] P.-L. Lions and P. E. Souganidis. Stochastic viscosity solutions. Book, in preparation.

[10] P.-L. Lions and P. E. Souganidis. Stochastic viscosity solutions of spatially dependent Hamilton-Jacobi equations with multiple paths. In preparation.

[11] P.-L. Lions and P. E. Souganidis. In preparation.

[12] C. Mantegazza and A.C. Mennucci. Hamilton-Jacobi Equations and Distance Functions on Riemannian Manifolds Applied Math. and Optim., vol 47, (2002)

[13] B. Seeger. Perron's method for stochastic viscosity solutions. Preprint, arXiv:1605.01108

([1]) Technische Universität Berlin and Weierstraß–Institut für Angewandte Analysis und Stochastik
Straße des 17. Juni 136
10623 Berlin
email: friz@math.tu-berlin.de

([2]) CEREMADE, Université de Paris-Dauphine,
Place du Maréchal-de-Lattre-de-Tassigny
75775 Paris cedex 16, France
email: gassiat@ceremade.dauphine.fr

([3]) Collège de France and CEREMADE, Université de Paris-Dauphine,
1, Place Marcellin Berthelot
75005 Paris Cedex 5, France
email: lions@ceremade.dauphine.fr

([4]) Department of Mathematics
University of Chicago
Chicago, IL 60637, USA
email: souganidis@math.uchicago.edu

(*) Partially supported by the European Research Council through StG-258237 / CoG-683166 and DFG research unit FOR2402.

(**) Affiliated to TU Berlin while this work was commenced and partially supported by the European Research Council through StG-258237.

(***) Partially supported by the National Science Foundation grant DMS-1266383.