J. Math. Pures Appl. 210 (2026) 103872

Contents lists available at ScienceDirect

Journal de Mathématiques Pures et Appliquées

journal homepage: www.elsevier.com/locate/matpur

A comparison principle based on couplings of partial integro-differential operators ☆

Serena Della Corte [a], Fabian Fuchs [b,d], Richard C. Kraaij [a,*], Max Nendel [c]

[a] *Delft Institute of Applied Mathematics, Delft University of Technology, Mekelweg 4, Delft, the Netherlands*
[b] *Center for Mathematical Economics, Bielefeld University, Universitätsstraße 25, Bielefeld, Germany*
[c] *Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue W, Waterloo, Canada*
[d] *Department of AI, Data and Decision Sciences, Luiss Guido Carli, Viale Romania 32, Roma, Italy*

A R T I C L E I N F O

Article history:
Received 31 October 2024
Available online 9 February 2026

MSC:
primary 35J60, 35D40, 45K05
secondary 49L25, 49Q22

Keywords:
Comparison principle
Viscosity solution
Hamilton–Jacobi–Bellman–Isaacs equation
Coupling of operators
Lyapunov function
Mixed topology

A B S T R A C T

This paper is concerned with a comparison principle for viscosity solutions to Hamilton–Jacobi (HJ), –Bellman (HJB), and –Isaacs (HJI) equations for general classes of partial integro-differential operators. Our approach contributes to the literature in three ways: (1) We cast the Crandall–Ishii Lemma into a test function framework to tackle a wide class of second-order integro-differential operators in the spirit of the classical doubling of variables method. (2) We provide a unified approach to estimate the difference of Hamiltonians by adapting the probabilistic notion of couplings to an analytic setting. (3) We strengthen the sup-norm contractivity resulting from the comparison principle to one that encodes continuity in the strict topology. We apply our theory to a variety of examples, in particular, to second-order differential operators and, more generally, generators of spatially inhomogeneous Lévy processes.

R É S U M É

Cet article porte sur un principe de comparaison pour les solutions de viscosité d'équations de Hamilton–Jacobi (HJ), –Bellman (HJB) et –Isaacs (HJI), pour des classes générales d'opérateurs intégro-différentiels partiels. Notre approche contribue à la littérature des trois manières suivantes : (1) Nous reformulons le lemme de Crandall–Ishii dans un cadre de fonctions tests afin de traiter une large classe d'opérateurs intégro-différentiels du second ordre, dans l'esprit de la méthode classique du doublement des variables. (2) Nous proposons une approche unifiée afin d'estimer la différence des Hamiltoniens en adaptant la notion probabiliste de couplages à un cadre analytique. (3) Nous renforçons la contraction en norme supremum issue du principe de comparaison en une contraction encodant la continuité pour la topologie stricte. Nous appliquons notre théorie à une variété

☆ This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1283/2 2021 – 317210226 and by The Netherlands Organisation for Scientific Research (NWO), grant number 613.009.148.

* Corresponding author.
E-mail addresses: s.dellacorte@tudelft.nl (S. Della Corte), fabian.fuchs@uni-bielefeld.de (F. Fuchs), r.c.kraaij@tudelft.nl (R.C. Kraaij), mnendel@uwaterloo.ca (M. Nendel).

https://doi.org/10.1016/j.matpur.2026.103872

d'exemples, en particulier aux opérateurs différentiels du second ordre et, plus généralement, aux générateurs de processus de Lévy spatialement inhomogènes.

1. Introduction

In this work, we provide a new perspective on comparison principles for viscosity solutions to the Hamilton–Jacobi equation

$$f - \lambda H f = h, \qquad \lambda > 0,\, h \in C_b(\mathbb{R}^q), \tag{1.1}$$

for Hamiltonians H of the type

$$\begin{aligned} Hf(x) = \langle b(x), \nabla f(x)\rangle + \frac{1}{2}\operatorname{Tr}\left(\Sigma\Sigma^T(x) D^2 f(x)\right) \\ + \int \left[f(x+\mathbf{z}) - f(x) - \chi_{B_1(0)}(\mathbf{z})\langle \mathbf{z}, \nabla f(x)\rangle\right]\mu_x(\mathrm{d}\mathbf{z}) + \mathcal{H}(\nabla f(x)) \end{aligned} \tag{1.2}$$

and, more generally, for those in Bellman and Isaacs form

$$\begin{aligned} Hf(x) &= \sup_{\theta\in\Theta}\{H_\theta f(x) - \mathcal{I}(x,\theta)\} \quad \text{and} \\ Hf(x) &= \sup_{\theta_1\in\Theta_1}\inf_{\theta_2\in\Theta_2}\{H_{\theta_1,\theta_2} f(x) - \mathcal{I}(x,\theta_1,\theta_2)\}, \end{aligned}$$

with H_θ and H_{θ_1,θ_2} as in (1.2) but with θ and (θ_1, θ_2) dependent coefficients, respectively, and an appropriate cost functional $\mathcal{I}$.

Motivated by convex Hamiltonians, for which no unique classical or weak solutions exist in general, [19] introduced the notion of viscosity solutions. The seminal works [15,17,29,30,37] explore this framework for first-order equations.

Most modern comparison proofs for operators containing second-order terms are based on results of [34,35]. Using then recent advances for generalized differentials, [16] provided what is nowadays known as the Crandall–Ishii Lemma. An overview over uniqueness results for viscosity solutions to degenerate elliptic equations is given in the User's Guide [18].

The treatment of non-local operators was initially motivated by problems in optimal control theory; see [2,5,41] for early examples with non-local operators. The works [7,32,33] give a non-local version of the Crandall–Ishii Lemma by adapting the original procedure in [18], see also [3,4,25,26,31]. We also refer to [22] for an overview of the Hilbertian setting, [10] for comparison principles for convex monotone semigroups on spaces of continuous functions, to [20] for the classical well-posedness of convex Cauchy problems on L^p, to [28] for a comparison principle in the framework of G-Lévy processes, and to [9] for a comparison principle for HJB equations on the set of probability measures.

Our approach and the main results of this paper, Theorem 3.12, Corollary 3.13, and Theorem 5.10, contribute to the literature in the following three ways:

(1) We reinterpret the classical doubling-of-variables method in the context of second-order equations by casting the Crandall–Ishii Lemma into a test function framework. This adaptation allows us to effectively handle non-local integral operators, such as generators of Lévy processes, in the same framework as second-order operators, paving the way for stability results.

(2) We translate the key estimate on the difference of Hamiltonians in terms of an adaptation of the probabilistic notion of couplings, providing a unified approach that applies to both continuous and discrete operators. We point out that [17] also discusses a coupling point of view, but only for first order operators.
(3) We strengthen the typical comparison principle using Lyapunov functionals from a sup-norm contractivity result to what we call the *strict comparison principle*, cf. Definition 2.4, which encodes continuity in the strict, also called mixed, topology, cf. [11,40].

The rest of the paper is organized as follows: Section 2 introduces the notation and basic definitions. Section 3 introduces the framework, states the necessary assumptions, and formalizes the main result. In Section 4, we show how to apply our framework to operators of the form (1.2). Section 5 contains the construction of the required optimizing points and test functions. Finally, Section 6 contains the proof of the main theorem.

1.1. Review of classical comparison principles

In this section, we revisit the standard proof techniques for comparison principles starting from first-order equations, highlighting the difficulties that arise in the second-order case, which require additional technicalities, and explaining how we incorporate the necessary tools into our framework.

To this end, consider the equation

$$f - \lambda H f = h.$$

Let u and v be a viscosity sub- and supersolution for the above equation with h replaced by h_1 and h_2, respectively. The goal of comparison principles is then to show that

$$\sup_{x\in\mathbb{R}^q} u(x) - v(x) \leq \sup_{x\in\mathbb{R}^q} h_1(x) - h_2(x).$$

In the case where H is a first-order differential operator, one typically performs a penalized doubling-of-variables

$$\sup_{x\in\mathbb{R}^q} u(x) - v(x) \leq \sup_{x,x'\in\mathbb{R}^q} u(x) - v(x') - \frac{\alpha}{2} d^2(x,x'),$$

where d^2 is the squared Euclidean distance. This procedure has two immediate consequences. Firstly, assuming for simplicity that one can work with optimizers (x_α, x'_α) of the doubled equation, one can find regular test functions, mainly in terms of d^2, for the use in the definition of the sub- and supersolution property.

The above procedure then leads to the estimate

$$\begin{aligned}\sup_{x\in\mathbb{R}^q} u(x) - v(x) &\leq u(x_\alpha) - v(x'_\alpha) \\ &\leq h_1(x_\alpha) - h_2(x'_\alpha) + \lambda\left[H\left(\frac{\alpha}{2}d^2(\cdot,x'_\alpha)\right)(x_\alpha) - H\left(-\frac{\alpha}{2}d^2(x_\alpha,\cdot)\right)(x'_\alpha)\right].\end{aligned}$$

Secondly, the optimizers (x_α, x'_α) satisfy

$$\alpha d^2(x_\alpha, x'_\alpha) \to 0, \tag{1.3}$$

as $\alpha \to \infty$. That is, sending $\alpha \to \infty$ forces the optimizing points to be close, thus essentially reducing the problem to an estimate on the difference of the Hamiltonians H evaluated in the squared metric d^2. Thus, the comparison principle holds if one can show that

$$\liminf_{\alpha\to\infty} H\left(\frac{\alpha}{2}d^2(\cdot,x'_\alpha)\right)(x_\alpha) - H\left(-\frac{\alpha}{2}d^2(x_\alpha,\cdot)\right)(x'_\alpha) \le 0. \tag{1.4}$$

In practice, estimate (1.4) then translates into explicit conditions on H. Typical conditions are coercivity and Lipschitzianity. When H is, for example, of the form

$$Hf(x) = \langle b(x), \nabla f(x)\rangle + \frac{1}{2}|\nabla f(x)|^2$$

with a one-sided Lipschitz drift term b, the estimate (1.4) translates into

$$\begin{aligned}
&H\left(\frac{\alpha}{2}d^2(\cdot,x'_\alpha)\right)(x_\alpha) - H\left(-\frac{\alpha}{2}d^2(x_\alpha,\cdot)\right)(x'_\alpha) \\
&= \left[\langle b(x_\alpha), \alpha(x_\alpha - x'_\alpha)\rangle + \frac{\alpha^2}{2}d^2(x_\alpha,x'_\alpha)\right] - \left[\langle b(x'_\alpha), \alpha(x_\alpha - x'_\alpha)\rangle + \frac{\alpha^2}{2}d^2(x_\alpha,x'_\alpha)\right] \\
&\le \langle b(x_\alpha) - b(x'_\alpha), \alpha(x_\alpha - x'_\alpha)\rangle \le \alpha L_b |x_\alpha - x'_\alpha|^2,
\end{aligned}$$

which, by (1.3), tends to 0 as $\alpha \to \infty$.

For second-order operators, however, the same strategy fails. Consider, for example, $\left(\frac{1}{2}\text{ times}\right)$ the Laplacian $Hf(x) = \frac{1}{2}\Delta f(x) = \frac{1}{2}\operatorname{Tr}\left(D^2 f(x)\right)$. The estimate (1.4) yields

$$H\left(\frac{\alpha}{2}d^2(\cdot,x'_\alpha)\right)(x_\alpha) - H\left(-\frac{\alpha}{2}d^2(x_\alpha,\cdot)\right)(x'_\alpha) = 2\alpha,$$

which diverges as $\alpha \to \infty$.

The works [34,35] use the key insight that, while typical first-order comparison proofs explore sequences of optimizers of the doubled equation separately (fix x'_α and vary x for the subsolution part and vice versa), for second-order equations, one needs to treat the two sequences jointly. This insight was later formalized in [16] and the User's Guide [18, Theorem 3.2], now known as the Crandall–Ishii Lemma.

For the basic second-order differential operator $Hf(x) = \frac{1}{2}\operatorname{Tr}\left(D^2 f(x)\right)$, given $X_\alpha = D^2u(x_\alpha)$ and $Y_\alpha = D^2 v(x'_\alpha)$, or appropriate generalizations thereof, the Crandall–Ishii Lemma supplies the estimate

$$\begin{pmatrix} X_\alpha & 0 \\ 0 & -Y_\alpha \end{pmatrix} \le 3\alpha \begin{pmatrix} \mathbb{1} & -\mathbb{1} \\ -\mathbb{1} & \mathbb{1} \end{pmatrix}.$$

Conjugating the matrices with

$$C := \frac{1}{\sqrt{2}}\begin{pmatrix} \mathbb{1} & \mathbb{1} \\ \mathbb{1} & \mathbb{1} \end{pmatrix}, \tag{1.5}$$

i.e., essentially using C to couple the subsolution and supersolution inequalities, we arrive at the desired estimate

$$\frac{1}{2}\operatorname{Tr}(X_\alpha) - \frac{1}{2}\operatorname{Tr}(Y_\alpha) = \frac{1}{4}\operatorname{Tr}\begin{pmatrix} X_\alpha - Y_\alpha & X_\alpha - Y_\alpha \\ X_\alpha - Y_\alpha & X_\alpha - Y_\alpha \end{pmatrix} \le 0. \tag{1.6}$$

Extensions of this originally local argument for equations with non-local terms can be found, e.g., in [7,25,32].

1.2. Our framework

To be able to treat second-order and non-local integral operators in the same way, we reinterpret the Crandall–Ishii Lemma in a test function framework together with a coupling approach. This reformulation allows us, on one hand, to handle a variety of operators jointly and, on the other hand, to obtain a stronger version of the typical comparison principle: the strict comparison principle. Our approach is organized around the following three main components.

1.2.1. Test function framework

Examining the proof of the Crandall–Ishii Lemma, one can interpret the procedure as the construction of two smooth test functions $\phi_\alpha, \psi_\alpha \in C^\infty(\mathbb{R}^q)$ that are squeezed between u and v on one hand, and $\frac{\alpha}{2}d^2$ on the other. To be more precise, we find $\phi_\alpha, \psi_\alpha \in C^\infty(\mathbb{R}^q)$ such that for slightly perturbed optimizing points

$$u(x_\alpha) - \phi_\alpha(x_\alpha) = \sup_{x\in\mathbb{R}^q}\{u(x) - \phi_\alpha(x)\}, \qquad v(x'_\alpha) - \psi_\alpha(x'_\alpha) = \inf_{x'\in\mathbb{R}^q}\{v(x') - \psi_\alpha(x')\},$$

and

$$\phi_\alpha(x_\alpha) - \psi_\alpha(x'_\alpha) - \frac{\alpha}{2}d^2(x_\alpha, x'_\alpha) = \sup_{x,x'\in\mathbb{R}^q}\left\{u(x) - v(x') - \frac{\alpha}{2}d^2(x,x')\right\}. \tag{1.7}$$

The construction of such test functions, as given in Proposition 5.10, is independent of the specific Hamiltonian. This allows us to decouple the technical construction of the test functions from operator-specific properties.

Analogous to Section 1.1, using the sub- and supersolution properties, comparison now follows from the estimate

$$\liminf_{\alpha\to\infty} H\phi_\alpha(x_\alpha) - H\psi_\alpha(x'_\alpha) \le 0.$$

For the Laplacian $Hf(x) = \frac{1}{2}\operatorname{Tr}\left(D^2 f(x)\right)$, this translates to

$$H\phi_\alpha(x_\alpha) - H\psi_\alpha(x'_\alpha) = \frac{1}{2}\operatorname{Tr}(D^2\phi_\alpha(x_\alpha)) - \frac{1}{2}\operatorname{Tr}(D^2\psi_\alpha(x'_\alpha)). \tag{1.8}$$

At this point, in proofs using the Crandall–Ishii Lemma, the estimate (1.6) is performed by the conjugation with the matrix C, cf. equation (1.5). We formalize this step by adapting the probabilistic notion of couplings, cf. [12,36,44], and identify the choice of the matrix C with the *synchronous coupling* (also called *co-monotone coupling*).

1.2.2. Synchronous coupling

We illustrate the idea of synchronous couplings in this context using the following example: Given two Brownian motions starting in x and x', one can construct a coupling of the two by considering

$$(X(t),\, X'(t)) = (x + B(t),\, x' + B(t)), \tag{1.9}$$

where $B(t)$ is a standard Brownian motion. The generator of the coupled process (1.9) is given by

$$\begin{aligned}\widehat{H}g(x,x') &:= \frac{1}{2}\left(\partial_x + \partial_{x'}\right)^2 g(x,x') = \frac{1}{2}\operatorname{Tr}\left(\begin{pmatrix}\mathbb{1} & \mathbb{1}\\ \mathbb{1} & \mathbb{1}\end{pmatrix} D^2 g(x,x')\right)\\ &= \frac{1}{2}\operatorname{Tr}\left(CD^2 g(x,x')C^T\right),\end{aligned}$$

where we recover the matrix C in (1.5). Note that $\widehat{H}$ is a coupling of H with itself in the following sense: For any $f_1, f_2 \in C_b(\mathbb{R}^q)$ and $(f_1 \oplus f_2)(x, x') := f_1(x) + f_2(x')$, we have

$$\widehat{H}(f_1 \oplus f_2)(x, x') = Hf_1(x) + Hf_2(x').$$

Using the coupling $\widehat{H}$, we can now rewrite (1.8) as

$$H\phi_\alpha(x_\alpha) - H\psi_\alpha(x'_\alpha) = \widehat{H}\left(\phi_\alpha \oplus -\psi_\alpha\right)(x_\alpha, x'_\alpha) \leq \widehat{H}\left(\frac{\alpha}{2}d^2\right)(x_\alpha, x'_\alpha) = 0, \tag{1.10}$$

where the first equality follows from the definition of a coupling, the inequality is based on the positive maximum principle with the optimizers from equation (1.7), and the final equality is due to the fact that the synchronous coupling controls distance growth. We formalize this structure as controlled growth couplings, cf. Definition 3.3.

The same strategy can be used to treat a discretized version of the Brownian motion by considering the generator $Hf(x) = \frac{1}{2}\left[f(x+1) - f(x)\right] + \frac{1}{2}\left[f(x-1) - f(x)\right]$ of a random walk: We synchronously couple the random walk with itself using the operator

$$\widehat{H}f(x,y) = \frac{1}{2}\left[f(x+1, y+1) - f(x,y)\right] + \frac{1}{2}\left[f(x-1, y-1) - f(x,y)\right].$$

The argument in (1.10) then works for the random walk exactly as it did for the Brownian motion. In the examples in Section 4, we make these heuristic arguments rigorous, cf. Propositions 4.4 and 4.12.

In fact, the coupling approach is a key step in the proof of the main theorem, see equation (6.17).

1.2.3. Strict comparison principle

Making the above aspects of our framework explicit allows for a modification of the doubling-of-variables technique that yields a stronger comparison result in the following sense.

Recall that for a subsolution u and a supersolution v to the equation $f - \lambda Hf = h$ with h replaced by h_1 for u and by h_2 for v, the goal of the comparison principle is to show

$$\sup_{x \in \mathbb{R}^q} u(x) - v(x) \leq \sup_{x \in \mathbb{R}^q} h_1(x) - h_2(x).$$

The comparison principle, once established, thus implies sup-norm contractivity of the solution map $R(\lambda)\colon C_b(\mathbb{R}^q) \to C_b(\mathbb{R}^q)$, where $R(\lambda)h$ is the unique viscosity solution to the Hamilton–Jacobi equation (1.1).

It is well-known from examples, cf. [6,13,24,46], that the map $R(\lambda)h$ takes the form of an exponentially discounted Markovian control problem. If the dynamics admit a Lyapunov function V, having compact sublevel sets and satisfying $HV \leq c$, the controlled Markov processes satisfy tightness properties. More precisely, if the controlled process starts in a compact set K, one can find, for any time horizon $T > 0$ and $\varepsilon > 0$, a compact set $\widehat{K} \supseteq K$, given in terms of the sublevel sets of V such that, with probability $1 - \varepsilon$, the process remains in $\widehat{K}$ up to time T. Rewriting this in terms of an estimate on the solution map $R(\lambda)$, we then find

$$\sup_{x \in K} R(\lambda)h_1(x) - R(\lambda)h_2(x) \leq \varepsilon \left\|h_1 - h_2\right\| + \sup_{x \in \widehat{K}} h_1(x) - h_2(x). \tag{1.11}$$

Estimates of this type are indeed characterized by the strict topology, as was first established for linear functionals in [40, Theorem 5.1] and for convex, monotone functionals in [38, Corollary 2.10]. Note that in this paper, we do not investigate the convexity of the map $h \mapsto R(\lambda)h$, but point out that, given a convex operator H, convexity of $R(\lambda)h$ is to be expected by performing a comparison principle in terms of three variables using, e.g., variants of the three-dimensional theorem in [18, Theorem 3.2], see also the domination principle in [26, Theorem 2.22 and Corollary 2.26].

Building upon the notion of Lyapunov functions, we show that we can directly establish a variant of (1.11) for a subsolution u and supersolution v. Given its motivation, we will call this estimate the *strict comparison principle*, see Definition 2.4 and the main result, Theorem 3.12.

2. Preliminaries

2.1. Notation

Throughout the paper, let $q \in \mathbb{N}$. We write $C(\mathbb{R}^q)$ for the set of all real-valued continuous functions on $\mathbb{R}^q$, where $\mathbb{R}^q$ is endowed with the topology induced by the Euclidean distance d on $\mathbb{R}^q$. Throughout, the terminology *differentiable* and *derivative* refers to the classical concepts with respect to the Euclidian distance d on $\mathbb{R}^q$.

Let $C_b(\mathbb{R}^q)$ be the set of bounded continuous functions. For $k \in \mathbb{N}$, let $C^k(\mathbb{R}^q)$ denote the space of all real-valued functions on $\mathbb{R}^q$ that are k-times continuously differentiable. Let $C_b^k(\mathbb{R}^q)$ the set of all functions in $C^k(\mathbb{R}^q)$ with bounded derivatives up to order k. We denote the space of all smooth functions that are constant outside of a compact set by $C_c^\infty(\mathbb{R}^q)$. We write $C_u(\mathbb{R}^q)$ and $C_l(\mathbb{R}^q)$ for the set of continuous functions on $\mathbb{R}^q$ that are uniformly bounded from above and below, respectively. Moreover, we write

$$
\begin{aligned}
C_+(\mathbb{R}^q) &:= \{f \in C(\mathbb{R}^q) \,|\, f \text{ has compact sub-level sets}\}, \\
C_-(\mathbb{R}^q) &:= \{f \in C(\mathbb{R}^q) \,|\, f \text{ has compact super-level sets}\}, \\
C_c(\mathbb{R}^q) &:= \{f \in C(\mathbb{R}^q) \,|\, f \text{ is constant outside of a compact set}\}.
\end{aligned}
$$

We furthermore define the following intersections: $C_c^2(\mathbb{R}^q) = C_c(\mathbb{R}^q) \cap C^2(\mathbb{R}^q)$,

$$
C_+^2(\mathbb{R}^q) := C_+(\mathbb{R}^q) \cap C^2(\mathbb{R}^q), \qquad C_-^2(\mathbb{R}^q) := C_-(\mathbb{R}^q) \cap C^2(\mathbb{R}^q).
$$

For $a, b \in \mathbb{R}$, we write $a \vee b := \max\{a, b\}$ and $a \wedge b := \min\{a, b\}$. We denote the supremum norm by $\|\cdot\|$, that is

$$
\|f\| = \sup_{x \in \mathbb{R}^q} |f(x)|,
$$

for $f \in C_b(\mathbb{R}^q)$, while, for $u \in C(\mathbb{R}^q)$, we use the notation

$$
\lceil u \rceil := \sup_{x \in \mathbb{R}^q} u(x), \qquad \lfloor u \rfloor := \inf_{x \in \mathbb{R}^q} u(x)
$$

for a supremum or infimum over the entire space and

$$
\lceil u \rceil_C := \sup_{x \in C} u(x), \qquad \lfloor u \rfloor_C := \inf_{x \in C} u(x)
$$

for a supremum or infimum over a subset $C \subseteq \mathbb{R}^q$.

We say that a function $\omega\colon [0, \infty) \to [0, \infty)$ is a *modulus of continuity*, if ω is upper semi-continuous with $\omega(0) = 0$. We say that a function $f \in C(\mathbb{R}^q)$ *admits a modulus of continuity*, if, for every compact $K \subseteq \mathbb{R}^q$, there exists a modulus of continuity $\omega_K\colon [0, \infty) \to [0, \infty)$ such that, for all $x, y \in K$, we have

$$
|f(x) - f(y)| \le \omega_K(d(x, y)).
$$

A function $\phi\colon \mathbb{R}^q \to \mathbb{R}$ is called *semi-convex* with constant $\kappa \in \mathbb{R}$ if, for any $x_0 \in \mathbb{R}^q$, the map

$$x \mapsto \phi(x) + \frac{\kappa}{2} d^2(x, x_0)$$

is convex. Moreover, ϕ is called *semi-concave* with constant $\kappa \in \mathbb{R}$ if $-\phi$ is semi-convex with constant $-\kappa$.

We say that a function $f \in C(\mathbb{R}^q, \mathbb{R}^q)$ is *one-sided Lipschitz* if, for all $x, y \in \mathbb{R}^q$ and some constant $C \in \mathbb{R}$, we have

$$\langle x - y, f(x) - f(y) \rangle \leq C d^2(x, y).$$

For any $z \in \mathbb{R}^q$, let $s_z : \mathbb{R}^q \to \mathbb{R}^q$ be the *shift map*

$$s_z(x) = x - z.$$

For any $z_1, z_2 \in \mathbb{R}^q$, let

$$d_{z_1,z_2}(x, y) \coloneqq d\left(s_{z_1}(x), s_{z_2}(y)\right).$$

Let $f_1, f_2 \in C(\mathbb{R}^q)$. Then, we define the *direct sum* $f_1 \oplus f_2, f_1 \ominus f_2 \in C(\mathbb{R}^q \times \mathbb{R}^q)$ as

$$(f_1 \oplus f_2)(x_1, x_2) \coloneqq f_1(x_1) + f_2(x_2) \quad \text{and} \quad (f_1 \ominus f_2)(x_1, x_2) \coloneqq f_1(x_1) - f_2(x_2)$$

for all $x_1, x_2 \in \mathbb{R}^q$. For two sets of functions $F_1, F_2 \subseteq C(\mathbb{R}^q)$, we define

$$F_1 \oplus F_2 \coloneqq \{f_1 \oplus f_2 \,|\, f_1 \in F_1, f_2 \in F_2\} \quad \text{and} \quad F_1 \ominus F_2 \coloneqq \{f_1 \ominus f_2 \,|\, f_1 \in F_1, f_2 \in F_2\}.$$

2.2. Operators

We consider operators $H \subseteq C(\mathbb{R}^q) \times C(\mathbb{R}^q)$, where we identify H by its graph. As usual, the *domain* of H is given by

$$\mathcal{D}(H) \coloneqq \{f \in C(\mathbb{R}^q) \,|\, \exists\, g \in C(\mathbb{R}^q)\colon (f, g) \in H\}.$$

Let $H_1, H_2 \subseteq C(\mathbb{R}^q) \times C(\mathbb{R}^q)$. We define

$$H_1 + H_2 \coloneqq \{(f, g_1 + g_2) \,|\, (f, g_1) \in H_1, (f, g_2) \in H_2\},$$

which is an operator with domain

$$\mathcal{D}(H_1 + H_2) \coloneqq \mathcal{D}(H_1) \cap \mathcal{D}(H_2).$$

We say that H is *linear on its domain* if, for any $f, g \in \mathcal{D}(H)$ and $a \in \mathbb{R}$ such that $af + g \in \mathcal{D}(H)$, we have

$$H\left(af + g\right) = aHf + Hg.$$

We will prove the comparison principle for the equation in terms of H by relating it to two equations in terms of two restrictions of H. To do so, we will need to be able to construct test functions in the domain of H from functions in the domain of the restrictions. In particular, we will need the following notion.

Definition 2.1 *(Sequential denseness).* Let $\mathcal{D} \subseteq C_b(\mathbb{R}^q)$, $\mathcal{D}_+ \subseteq C_+(\mathbb{R}^q)$, and $\mathcal{D}_- \subseteq C_-(\mathbb{R}^q)$.

- We say that $\mathcal{D}$ is *upward sequentially dense* in $\mathcal{D}_+$ if, for any $f_\dagger \in \mathcal{D}_+$ and constant $a \in \mathbb{R}$, there exists a function $f_{\dagger,a} \in \mathcal{D}$ such that

$$\begin{cases} f_{\dagger,a}(x) = f_\dagger(x) & \text{if } f_\dagger(x) \le a, \\ a < f_{\dagger,a}(x) \le f_\dagger(x) & \text{if } f_\dagger(x) > a. \end{cases}$$

- We say that $\mathcal{D}$ is *downward sequentially dense* in $\mathcal{D}_-$ if, for any $f_\ddagger \in \mathcal{D}_-$ and constant $a \in \mathbb{R}$, there exists a function $f_{\ddagger,a} \in \mathcal{D}$ such that

$$\begin{cases} f_{\ddagger,a}(x) = f_\dagger(x) & \text{if } f_\ddagger(x) \ge a, \\ a > f_{\ddagger,a}(x) \ge f_\ddagger(x) & \text{if } f_\ddagger(x) < a. \end{cases}$$

2.3. Viscosity solutions

For $\lambda > 0$, consider $h_1 \in C_l(\mathbb{R}^q)$ and $h_2 \in C_u(\mathbb{R}^q)$ and two operators $H_1 \subseteq C_l(\mathbb{R}^q) \times C(\mathbb{R}^q)$ and $H_2 \subseteq C_u(\mathbb{R}^q) \times C(\mathbb{R}^q)$. We study the pair of equations

$$f - \lambda H_1 f \le h_1, \tag{2.1}$$

$$f - \lambda H_2 f \ge h_2. \tag{2.2}$$

The notion of viscosity solution is built upon the maximum principle.

Definition 2.2 *(Maximum principle).* We say that an operator $H \subseteq C(\mathbb{R}^q) \times C(\mathbb{R}^q)$ satisfies the *maximum principle* if, for all $f_1, f_2 \in \mathcal{D}(H)$ and $x_0 \in \mathbb{R}^q$ with

$$f_1(x_0) - f_2(x_0) = \sup_{x \in \mathbb{R}^q} \{f_1(x) - f_2(x)\},$$

we have

$$Hf_1(x_0) \le Hf_2(x_0)$$

and, analogously, for all $f_1, f_2 \in \mathcal{D}(H)$ and $x_0 \in \mathbb{R}^q$ with

$$f_1(x_0) - f_2(x_0) = \inf_{x \in \mathbb{R}^q} \{f_1(x) - f_2(x)\},$$

we have

$$Hf_1(x_0) \ge Hf_2(x_0).$$

Observe that every operator $H \subseteq C(\mathbb{R}^q) \times C(\mathbb{R}^q)$ that satisfies the maximum principle is single-valued, i.e., for all $f \in \mathcal{D}(H)$,

$$\#\{g \in C(\mathbb{R}^q) \mid (f, g) \in H\} = 1.$$

Definition 2.3 *(Viscosity sub- and supersolutions).* Let $H_1 \subseteq C_l(\mathbb{R}^q) \times C(\mathbb{R}^q)$ and $H_2 \subseteq C_u(\mathbb{R}^q) \times C(\mathbb{R}^q)$ be two operators with domains $\mathcal{D}(H_1)$ and $\mathcal{D}(H_2)$, respectively. Moreover, let $\lambda > 0$, $h_1 \in C_l(\mathbb{R}^q)$, and $h_2 \in C_u(\mathbb{R}^q)$.

(a) A bounded, upper semicontinuous function $u\colon \mathbb{R}^q \to \mathbb{R}$ is called a *(viscosity) subsolution* to (2.1) if, for all $(f, g) \in H_1$, there exists a sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^q$ such that

$$\lim_{n\to\infty} u(x_n) - f(x_n) = \sup_{x\in\mathbb{R}^q} u(x) - f(x),$$

$$\limsup_{n\to\infty} u(x_n) - \lambda g(x_n) - h_1(x_n) \le 0.$$

(b) A bounded, lower semicontinuous function $v\colon \mathbb{R}^q \to \mathbb{R}$ is called a *(viscosity) supersolution* to (2.2) if, for all $(f, g) \in H_2$, there exists a sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^q$ such that

$$\lim_{n\to\infty} v(x_n) - f(x_n) = \inf_{x\in\mathbb{R}^q} v(x) - f(x),$$

$$\liminf_{n\to\infty} v(x_n) - \lambda g(x_n) - h_2(x_n) \ge 0.$$

If $H_1 = H_2$ and $h_1 = h_2$, a function $u \in C_b(\mathbb{R}^q)$ is called a *(viscosity) solution* to the pair of equations (2.1) and (2.2) if it is both a subsolution to (2.1) and a supersolution to (2.2).

Working with test functions that have compact sub- or superlevel sets respectively, an approximating sequence can be replaced by an optimizing point in the definition of sub- or supersolution, see Appendix B.

Associated with the definition of viscosity solutions, we introduce the comparison principle, which, for $h_1 = h_2$, implies uniqueness in the viscosity sense for solutions of the Hamilton–Jacobi equation $f - \lambda H f = h$. We additionally introduce a new, stronger notion: the strict comparison principle. This name is inspired by the observation that the comparison principle implies contractivity in the sup-norm of the solution map. The strict comparison principle implies continuity in terms of the weaker, strict topology, see e.g. [40].

Definition 2.4. We say that the equations (2.1) and (2.2) satisfy

(a) the *comparison principle* if, for any subsolution u to (2.1) and any supersolution v to (2.2), we have

$$\sup_{x\in\mathbb{R}^q} u(x) - v(x) \le \sup_{x\in\mathbb{R}^q} h_1(x) - h_2(x).$$

(b) the *strict comparison principle* if, for any subsolution u to (2.1), any supersolution v to (2.2), any compact set $K \subseteq \mathbb{R}^q$ and $\varepsilon > 0$, there exist a compact set $\widehat{K} = \widehat{K}(K, \varepsilon, \|u\|, \|v\|)$ and constant $C = C(u, v, K, h_1, h_2, \lambda)$ such that we have

$$\sup_{x\in K} u(x) - v(x) \le \varepsilon C + \sup_{x\in\widehat{K}} h_1(x) - h_2(x).$$

Observe that the strict comparison principle implies the comparison principle. Indeed, by the strict comparison principle, for all $x_0 \in \mathbb{R}^q$ and $\varepsilon > 0$, there exists a constant C, independent of ε, and a compact set $\widehat{K} \subseteq \mathbb{R}^q$ such that

$$u(x_0) - v(x_0) \le \varepsilon C + \sup_{x\in\widehat{K}} h_1(x) - h_2(x) \le \varepsilon C + \sup_{x\in\mathbb{R}^q} h_1(x) - h_2(x).$$

Letting $\varepsilon \downarrow 0$, we find that $u(x_0) - v(x_0) \le \sup_{x\in\mathbb{R}^q} h_1(x) - h_2(x)$. Taking the supremum over all $x_0 \in \mathbb{R}^q$, the comparison principle follows.

3. Setup and main result

In this section, we state our main result, Theorem 3.12, that contains the strict comparison principle for operators of the type

$$\mathbb{H}f(x) = \sup_{\theta_1\in\Theta_1}\inf_{\theta_2\in\Theta_2} \mathbb{H}_{\theta_1,\theta_2} f(x) - \mathcal{I}(x,\theta_1,\theta_2) \tag{3.1}$$

in Hamilton–Jacobi–Isaacs (HJI) form.

In Section 3.1, we introduce the key concepts of our framework and appropriate penalization functions to perform localizations and the Jensen perturbation. Theorem 3.12 is stated in Section 3.2. Further technical assumptions that are needed for the main result are given in Section 3.3.

3.1. Core concepts of the framework

In this subsection, we give the main definitions underlying our framework. We consider operators of the form

$$H = A + B,$$

where A is a stochastic part, which we can couple in the sense of the Definition 3.3 below, and B is a deterministic part, which we require to be a *convex semi-monotone* operator in the sense of Definition 3.6 below. Later, H will play the role of $\mathbb{H}_{\theta_1,\theta_2}$ in equation (3.1).

Definition 3.1 *(Coupling).* Let $A \subseteq C(\mathbb{R}^q)\times C(\mathbb{R}^q)$ and $\widehat{A} \subseteq C(\mathbb{R}^{2q})\times C(\mathbb{R}^{2q})$ be linear on their respective domains. We say $\widehat{A}$ is a *coupling of* A if $\mathcal{D}(A)\oplus\mathcal{D}(A)\subseteq\mathcal{D}(\widehat{A})$ and, for any $f_1,f_2\in\mathcal{D}(A)$, we have

$$\widehat{A}\,(f_1\oplus f_2) = Af_1 + Af_2.$$

Definition 3.2 *(Controlled growth).* Let $\widehat{A}\subseteq C(\mathbb{R}^{2q})\times C(\mathbb{R}^{2q})$. We say that $\widehat{A}$ has *controlled growth* if, for any $\alpha>1$ and $z,z'\in\mathbb{R}^q$, we have $\frac{\alpha}{2}d^2_{z,z'}\in\mathcal{D}(\widehat{A})$. In addition, for any compact set $K\subseteq\mathbb{R}^q$, there exists a modulus of continuity $\omega_{\widehat{A},K}\colon[0,\infty)\to[0,\infty)$ and $x,x',y,y'\in K$ such that

$$\begin{aligned}\widehat{A}\left(\frac{\alpha}{2}d^2_{x-y,\,x'-y'}\right)(x,x') \leq& \omega_{\widehat{A},K}\Big(\alpha\left(d(x,y)+d(y,y')+d(y',x')\right)^2 \\ &+ \left(d(x,y)+d(y,y')+d(y',x')\right)\Big).\end{aligned}$$

Definition 3.3 *(Controlled growth coupling).* Let $A \subseteq C(\mathbb{R}^q)\times C(\mathbb{R}^q)$ and $\widehat{A} \subseteq C(\mathbb{R}^{2q})\times C(\mathbb{R}^{2q})$ be linear on their respective domains. We say $\widehat{A}$ is a *controlled growth coupling of* A if the following properties are satisfied:

(a) $\widehat{A}$ satisfies the *maximum principle*, cf. Definition 2.2.
(b) $\widehat{A}$ is a *coupling* of A, cf. Definition 3.1.
(c) $\widehat{A}$ has *controlled growth*, cf. Definition 3.2.

Definition 3.4 *(Local first-order operator).* We say that $B\subseteq C(\mathbb{R}^q)\times C(\mathbb{R}^q)$ is a *local first-order* operator if there exists a continuous map $\mathcal{B}:\mathbb{R}^q\times\mathbb{R}^q\to\mathbb{R}$ such that, for any $f\in\mathcal{D}(B)$, we have $Bf(x)=\mathcal{B}(x,\nabla f(x))$.

Definition 3.5 *(Local semi-monotonicity).* Let $B \subseteq C(\mathbb{R}^q) \times C(\mathbb{R}^q)$ be local first-order for some $\mathcal{B}$, cf. Definition 3.4. We say that B is *locally semi-monotone* if, for any compact set $K \subseteq \mathbb{R}^q$, there exists a modulus of continuity $\omega_{\mathcal{B},K} \colon [0,\infty) \to [0,\infty)$ such that, for all $x, y \in K$ and $\alpha > 1$,

$$\mathcal{B}(x, \alpha(x - x')) - \mathcal{B}(y, \alpha(x - x')) \leq \omega_{\mathcal{B},K}\left(\alpha d^2(x, x') + d(x, x')\right).$$

Definition 3.6 *(Convex semi-monotone operator).* We say that $B \subseteq C(\mathbb{R}^q) \times C(\mathbb{R}^q)$ is a *convex semi-monotone operator* if the following properties are satisfied:

(a) $B \subseteq C(\mathbb{R}^q) \times C(\mathbb{R}^q)$ is *locally semi-monotone* for some $\mathcal{B}$, cf. Definition 3.5.
(b) For all $x \in \mathbb{R}^q$, the map $p \mapsto \mathcal{B}(x, p)$ is *convex*.

Following typical comparison principle proofs, we will be perturbing the optimization problem

$$\sup_{x \in \mathbb{R}^q} u(x) - v(x),$$

using a variant of the doubling-of-variables procedure to ensure that we can use the properties of sub- and supersolutions. Our perturbations consist of two components:

- We need a function V that allows us to work with compact sets, see Definition 3.7.
- We perform a variant of the Jensen perturbation to construct optimizers, in which the second derivative exists, see Definition 3.10. This construction is based on a linear perturbation $\zeta_{p,z}$ and a localization function ξ_z.

Definition 3.7 *(Containment function).* We call $V \colon \mathbb{R}^q \to [0, \infty)$ a *containment function* if

(a) $\inf_{y \in \mathbb{R}^q} V(y) = 0$,
(b) V is semi-concave with semi-concavity constant κ_V,
(c) for every $c \geq 0$ the set $\{y \,|\, V(y) \leq c\}$ is compact.

Typical examples include, e.g., approximately $V(x) \approx |x|$ or quadratic $V(x) \approx x^2$. In Section 4, we will work with the following containment function.

Example 3.8 *(Containment function).* Consider

$$V(x) = \log\left(1 + \frac{1}{2}x^2\right).$$

Note, that V is semi-concave with constant $\kappa_V = 1$.

Later, we will use containment functions that behave like a Lyapunov function for the Hamiltonian H, which motivates the use of a slowly growing function V. More precisely, we will consider the following definition.

Definition 3.9 *(Lyapunov function).* Let $V : \mathbb{R}^q \to [0, \infty)$ be a containment function and let $H \subseteq C(\mathbb{R}^q) \times C(\mathbb{R}^q)$ be an operator. We say that V is a *Lyapunov function* for H, if $V \in D(H)$ and

$$\sup_{x \in \mathbb{R}^q} HV(x) < \infty.$$

The next definition introduces the functions, with which we can perform the Jensen perturbation procedure, cf. [18, Lemma A.3], to produce points, at which second derivatives exist. The variant used here creates a unique, global optimizer from a local one using ξ and then shifts it slightly with ζ. The two sets of perturbations are based on the prototypical examples of lines, i.e.,

$$\zeta_{z,p}(x) = \langle p, x - z \rangle \,, \tag{3.2}$$

and parabolas, i.e.,

$$\xi_z(x) = \frac{1}{2} d^2(x, z), \tag{3.3}$$

both centered at some $z \in \mathbb{R}^q$. We give them as a pair to capture the idea that quadratic growth dominates linear growth, cf. Definition 3.10 (d) below, which is used in our variant of Jensen's Lemma, cf. Proposition 5.4.

Definition 3.10. We call collections of maps $\{\zeta_{z,p}\}_{z\in\mathbb{R}^q, p\in\mathbb{R}^q} \subseteq C(\mathbb{R}^q)$ and $\{\xi_z\}_{z\in\mathbb{R}^q} \subseteq C^1(\mathbb{R}^q)$ with $\zeta_{z,p} \colon \mathbb{R}^q \to \mathbb{R}$ and $\xi_z \colon \mathbb{R}^q \to \mathbb{R}$ sets of *first* and *second order point penalizations*, respectively, if there exist constants $R > 0$ and $\kappa_\xi > 0$ such that for all $z \in \mathbb{R}^q$:

(a) $\zeta_{z,p}$ is linear in terms of p around z:

$$\zeta_{z,p}(y) = \langle p, y - z \rangle \,,$$

if $y \in B_R(z)$.

(b) The map ξ_z is semi-concave with constant κ_ξ.

(c) The map ξ_z is a penalization away from z:

$$\xi_z(z) = 0, \qquad \xi_z(y) > 0, \qquad \text{if } y \neq z.$$

(d) We have

$$\inf_{|p|\leq 1} \inf_{y \notin B_R(z)} \xi_z(y) + \zeta_{z,p}(y) > 0.$$

For any given $z_0, z_1 \in \mathbb{R}^q$ and $p \in \mathbb{R}^q$, we consider the maps

$$\Xi^0(y) = \Xi^0_{z_0,p}(y) := \xi_{z_0}(y) + \zeta_{z_0,p}(y),$$
$$\Xi(y) = \Xi_{z_0,p,z_1}(y) := \xi_{z_0}(y) + \zeta_{z_0,p}(y) + \xi_{z_1}(y).$$

In our example Section 4, we will work with the following choices for ζ and ξ. The first collection follows (3.2) and (3.3) and is a usual choice in the literature. The second collection is based on a cut-off of the first collection. We do this to control the action of non-local integral operators on the perturbations.

Example 3.11 *(Jensen penalization functions).* Consider the following two collections of penalization functions:

Collection 1 The base penalizations are

$$\zeta_{z,p}(x) = \langle p, x - z \rangle \,,$$
$$\xi_z(x) = \frac{1}{2} d^2(x, z).$$

Collection 2 Let $R'' > R' > R > 2$. Let $\overline{\ell} : [0, \infty) \to [0, \infty)$ be a smooth function satisfying $\overline{l}(r) = 1$ for $r < R'$ and $\overline{l}(r) = 0$ for $x > R''$. Let

$$\overline{\xi}_z(x) = (1 - \overline{\ell}(d(x,z)))(R''+1)^2 + \overline{\ell}(d(x,z))\frac{1}{2}d^2(x,z),$$

$$\overline{\zeta}_{p,z}(x) = \overline{\ell}(d(x,z)) \langle p, x - z \rangle .$$

3.2. Comparison principle

The following theorem is the main result of the paper.

Theorem 3.12 *(Strict comparison principle). Consider a containment function V and penalization functions $\{\zeta_{z,p}\}_{z\in\mathbb{R}^q, p\in\mathbb{R}^q}$ and $\{\zeta_z\}_{z\in\mathbb{R}^q}$. Let $\mathbb{H} \subseteq C(\mathbb{R}^q) \times C(\mathbb{R}^q)$ be given by*

$$\mathbb{H}f(x) = \sup_{\theta_1\in\Theta_1} \inf_{\theta_2\in\Theta_2} \{\mathbb{H}_{\theta_1,\theta_2}f(x) - \mathcal{I}(x,\theta_1,\theta_2)\},$$

$$\mathbb{H}_{\theta_1,\theta_2}f(x) = \mathbb{A}_{\theta_1,\theta_2}f(x) + \mathbb{B}_{\theta_1,\theta_2}f(x),$$

with compact, metric spaces Θ_1 and Θ_2, satisfying the technical Assumptions 3.15 and 3.16 below, and

(a) *For all $\theta_1 \in \Theta_1, \theta_2 \in \Theta_2$, $\mathbb{A}_{\theta_1,\theta_2}$ is linear on its domain and admits a controlled growth coupling $\widehat{\mathbb{A}}_{\theta_1,\theta_2}$ as in Definition 3.3 with a modulus uniform in θ_1 and θ_2.*
(b) *For all $\theta_1 \in \Theta_1, \theta_2 \in \Theta_2$, $\mathbb{B}_{\theta_1,\theta_2}$ is a convex semi-monotone operator as in Definition 3.6 with a modulus uniform in θ_1 and θ_2.*
(c) *The cost functional $\mathcal{I}\colon \mathbb{R}^q \times \Theta_1 \times \Theta_2 \to (-\infty, \infty]$ is lower semi-continuous in (x, θ_1, θ_2), upper semi-continuous in θ_2 for fixed (x, θ_1), and admits a modulus of continuity in x uniformly in (θ_1, θ_2).*
(d) *The collection of operators $\{\mathbb{H}_{\theta_1,\theta_2}\}_{\theta_1\in\Theta_1,\theta_2\in\Theta_2}$ satisfies Isaacs' condition, i.e., for all $f \in \bigcap_{\theta_1\in\Theta_1,\theta_2\in\Theta_2} \mathcal{D}(\mathbb{H}_{\theta_1,\theta_2})$,*

$$\sup_{\theta_1\in\Theta_1} \inf_{\theta_2\in\Theta_2} \{\mathbb{H}_{\theta_1,\theta_2}f(x) - \mathcal{I}(x,\theta_1,\theta_2)\} = \inf_{\theta_2\in\Theta_2} \sup_{\theta_1\in\Theta_1} \{\mathbb{H}_{\theta_1,\theta_2}f(x) - \mathcal{I}(x,\theta_1,\theta_2)\}.$$

(e) *We have $V \in \mathcal{D}(\mathbb{H})$ and*

$$c_V := \sup_{x\in\mathbb{R}^q} \sup_{\theta_1\in\Theta_1} \sup_{\theta_2\in\Theta_2} \mathbb{H}_{\theta_1,\theta_2}V(x) - \mathcal{I}(x,\theta_1,\theta_2) < \infty. \tag{3.4}$$

In particular, V is a Lyapunov function *for $\mathbb{H}$ in the sense of Definition 3.9.*

Let $H \subseteq \{(f,g) \in \mathbb{H} \,|\, f \in C_b(\mathbb{R}^q)\}$ and consider

$$f - \lambda Hf = h \tag{3.5}$$

for $\lambda > 0$ and $h \in C_b(\mathbb{R}^q)$. Let u and v be a sub- and supersolution to (3.5) with h_1 and h_2 instead of h, respectively. Then, for any compact set $K \subseteq \mathbb{R}^q$ and $\varepsilon \in (0,1)$, we have

$$\sup_{x\in K} u(x) - v(x) \leq \varepsilon C_\varepsilon + \sup_{x\in\widehat{K}_\varepsilon} h_1(x) - h_2(x), \tag{3.6}$$

where $\widehat{K}_\varepsilon := \widehat{K}_\varepsilon(K,u,v)$ and $C_\varepsilon := C_\varepsilon(K,u,v,h_1,h_2)$ are given by

$$\widehat{K}_\varepsilon := \left\{ z \in \mathbb{R}^q \,\middle|\, V(z) \leq \frac{\|u\| + \|v\|}{\varepsilon} + \lceil V \rceil_K \right\},$$
$$C_\varepsilon := \frac{2}{1-\varepsilon^2} \left(\lceil V \rceil_K + \lambda c_V \right) + \frac{1}{1-\varepsilon} \|h_1\| + \frac{1}{1-\varepsilon} \|h_2\| - \left\lfloor \frac{1}{1-\varepsilon} u - \frac{1}{1+\varepsilon} v \right\rfloor_K .$$

In particular, the strict comparison principle holds for (3.5).

The proof of Theorem 3.12 is carried out in Sections 5 and 6. In Section 5, we construct test functions and optimizers based on a variant of the Jensen perturbation result. This result is independent of the operator $\mathbb{H}$. Based on this construction, we proceed in Section 6 to verify the strict comparison principle based on properties of $\mathbb{H}$.

Examples, in which we apply the results, can be found in Section 4.

Corollary 3.13.

(a) *The strict comparison principle for HJB equations follows from Theorem 3.12 by taking Θ_2 to be a singleton.*
(b) *The strict comparison principle for HJ equations follows from Theorem 3.12 by taking both Θ_1 and Θ_2 to be singletons.*

Remark 3.14. If H is the generator of a Markov process, the comparison principle implies uniqueness of the martingale problem using [14, Theorem 3.7]. We also refer to [8,21,43] for details on the martingale problem.

3.3. Regularity and compatibility assumptions

In this section, we state the technical assumptions necessary for the proof the main theorem.

As we have a choice for the domain of our operator and only need functions with compact sub- and super-level sets, we need to ensure that the domains of the restrictions are regular enough to perform our analysis. In particular, the action of the operator on test functions and their combinations with perturbations need to be well-defined. Furthermore, we require that the domains are large enough to allow for approximations in the sense of Definition 2.1.

Considering the classical case $H = \frac{1}{2}\Delta$ with $\mathcal{D}(H) = C_c^\infty(\mathbb{R}^q)$, a natural extension that includes our prototypical functions

$$V(x) = \log\left(1 + \frac{1}{2}x^2\right), \quad \zeta_{z,p}(x) = \langle p, x - z \rangle, \quad \xi_z(x) = \frac{1}{2} d^2(x,z), \tag{3.7}$$

would be $\mathbb{H} = \frac{1}{2}\Delta$ with $\mathcal{D}(\mathbb{H}) = C^2(\mathbb{R}^q)$. Starting from general $\mathbb{H}$, this motivates the following assumption.

Assumption 3.15 *(Regularity of $\mathbb{H}$).* Let $\mathbb{H} \subseteq C(\mathbb{R}^q) \times C(\mathbb{R}^q)$ be an operator with the following three restrictions

$$H \subseteq \{(f,g) \in \mathbb{H} \,|\, f \in C_b(\mathbb{R}^q)\},$$
$$H_+ \subseteq \{(f,g) \in \mathbb{H} \,|\, f \in C_+(\mathbb{R}^q)\},$$
$$H_- \subseteq \{(f,g) \in \mathbb{H} \,|\, f \in C_-(\mathbb{R}^q)\},$$

satisfying

(a) $\mathbb{H}$ satisfies the maximum principle,

(b) $\mathcal{D}(H)$ is linear and $C_c^\infty(\mathbb{R}^q) \subseteq \mathcal{D}(H) \subseteq C_b(\mathbb{R}^q)$,
(c) $\mathcal{D}(H)$ is upward sequentially dense in $\mathcal{D}(H_+)$, as in Definition 2.1,
(d) $\mathcal{D}(H)$ is downward sequentially dense in $\mathcal{D}(H_-)$ as in Definition 2.1,
(e) $\mathcal{D}(H_+)$ is convex,
(f) for any $f \in \mathcal{D}(H)$ and $g \in \mathcal{D}(H_+)$ and $\delta \in (0,1)$, we have

$$(1-\delta)f + \delta g \in \mathcal{D}(H_+), \qquad (1+\delta)f - \delta g \in \mathcal{D}(H_-).$$

In our main theorem, Theorem 3.12, we assume that the Hamiltonian $\mathbb{H}$ has an Isaacs-type structure

$$\mathbb{H}f(x) = \sup_{\theta_1\in\Theta_1} \inf_{\theta_2\in\Theta_2} \left\{ \mathbb{A}_{\theta_1,\theta_2} f(x) + \mathbb{B}_{\theta_1,\theta_2} f(x) - \mathcal{I}(x,\theta_1,\theta_2) \right\}.$$

To ensure it is well-behaved, we need the collections of operators $\{\mathbb{A}_{\theta_1,\theta_2}\}_{\theta_1\in\Theta_1,\theta_2\in\Theta_2}$ and $\{\mathbb{B}_{\theta_1,\theta_2}\}_{\theta_1\in\Theta_1,\theta_2\in\Theta_2}$ themselves to be well-behaved as functions of (θ_1,θ_2) and when acting on the families of penalization functions, introduced in Section 3.1, in the following sense: The collections of operators $\{\mathbb{A}_{\theta_1,\theta_2}\}_{\theta_1\in\Theta_1,\theta_2\in\Theta_2}$ and $\{\mathbb{B}_{\theta_1,\theta_2}\}_{\theta_1\in\Theta_1,\theta_2\in\Theta_2}$ are continuous in (θ_1,θ_2) and the state variable x. The domains of the operators contain the (shifted) perturbations and the Lyapunov function. The operators act continuously with respect to all variables on the perturbations and the Lyapunov function composed with the shift maps $\Xi_{z_0,p,z_1} \circ s_z$ and $V \circ s_z$, respectively.

Returning to our basic example $\mathbb{H} = \frac{1}{2}\Delta$ with $\mathcal{D}(\mathbb{H}) = C^2(\mathbb{R}^q)$ and $V, \zeta_{z,p}, \xi_z$ as in (3.7), this is trivially satisfied. For a more general operator of the form

$$\mathbb{H}f(x) = b(x)\nabla f(x) + \frac{1}{2}\sigma^2(x)\Delta f(x)$$

with domain $\mathcal{D}(\mathbb{H}) = C^2(\mathbb{R}^q)$, the following assumption translates to continuity requirements on the coefficients σ and b.

Assumption 3.16 *(Compatibility of $\mathbb{A}_{\theta_1,\theta_2}$ and $\mathbb{B}_{\theta_1,\theta_2}$).* Let Θ_1 and Θ_2 be compact, metric spaces. For $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$, let $\mathbb{A}_{\theta_1,\theta_2}, \mathbb{B}_{\theta_1,\theta_2} \subseteq C(\mathbb{R}^q) \times C(\mathbb{R}^q)$. Consider a containment function V as in Definition 3.7 and penalization functions $\{\zeta_{z,p}\}_{z\in\mathbb{R}^q,p\in\mathbb{R}^q}$ and $\{\xi_z\}_{z\in\mathbb{R}^q}$ as in Definition 3.10.

(a) Let the collection $\{\mathbb{A}_{\theta_1,\theta_2}\}_{\theta_1\in\Theta_1,\theta_2\in\Theta_2}$ be *compatible* with V, $\{\zeta_{z,p}\}_{z\in\mathbb{R}^q,p\in\mathbb{R}^q}$, and $\{\xi_z\}_{z\in\mathbb{R}^q}$, i.e.,
   (1) we have

$$V \circ s_z \in \mathcal{D}(\mathbb{A}_{\theta_1,\theta_2}), \qquad \Xi_{z_0,p,z_1} \circ s_z \in \mathcal{D}(\mathbb{A}_{\theta_1,\theta_2})$$

   for any $\theta_1 \in \Theta_1$, $\theta_2 \in \Theta_2$, and $z \in \overline{B_1(0)}$,
   (2) the maps

$$(\theta_1,\theta_2,x,z_0,p,z_1,z) \mapsto \mathbb{A}_{\theta_1,\theta_2}\left(\Xi_{z_0,p,z_1} \circ s_z\right)(x),$$
$$(\theta_1,\theta_2,x,z) \mapsto \mathbb{A}_{\theta_1,\theta_2}\left(V \circ s_z\right)(x)$$

   are continuous,
   (3) the map

$$(\theta_1,\theta_2) \mapsto \mathbb{A}_{\theta_1,\theta_2} f(x)$$

   is continuous for any $x \in \mathbb{R}^q$ and $f \in \bigcap_{\theta_1\in\Theta_1,\theta_2\in\Theta_2} \mathcal{D}(\mathbb{A}_{\theta_1,\theta_2})$.

(b) Let the collection $\{\mathbb{B}_{\theta_1,\theta_2}\}_{\theta_1\in\Theta_1,\theta_2\in\Theta_2}$ be *compatible* with V, $\{\zeta_{z,p}\}_{z\in\mathbb{R}^q,p\in\mathbb{R}^q}$, and $\{\xi_z\}_{z\in\mathbb{R}^q}$, i.e.,

(1) we have

$$V\circ s_z\in\mathcal{D}(\mathbb{B}_{\theta_1,\theta_2}),\qquad \Xi_{z_0,p,z_1}\circ s_z\in\mathcal{D}(\mathbb{B}_{\theta_1,\theta_2})$$

for any $\theta_1\in\Theta_1$, $\theta_2\in\Theta_2$, and $z\in\overline{B_1(0)}$,

(2) the maps

$$(\theta_1,\theta_2,x,z_0,p,z_1)\mapsto\mathbb{B}_{\theta_1,\theta_2}\Xi_{z_0,p,z_1}(x),$$
$$(\theta_1,\theta_2,x)\mapsto\mathbb{B}_{\theta_1,\theta_2}V(x)$$

are continuous,

(3) the map

$$(\theta_1,\theta_2)\mapsto\mathbb{B}_{\theta_1,\theta_2}f(x)$$

is continuous for any $x\in\mathbb{R}^q$ and $f\in\bigcap_{\theta_1\in\Theta_1,\theta_2\in\Theta_2}\mathcal{D}(\mathbb{B}_{\theta_1,\theta_2})$.

4. Application to partial integro-differential operators

In this section, we discuss the application of our framework to partial integro-differential operators of the type

$$\mathbb{H}f(x)=\langle b(x),\nabla f(x)\rangle+\frac{1}{2}\operatorname{Tr}\left(\Sigma\Sigma^T(x)D^2f(x)\right)$$
$$+\int\left[f(x+\mathbf{z})-f(x)-\chi_{B_1(0)}(\mathbf{z})\left\langle\mathbf{z},\nabla f(x)\right\rangle\right]\mu_x(\mathrm{d}\mathbf{z})+\mathcal{H}(\nabla f(x))\qquad(4.1)$$

and suprema thereof. For simplicity, we restrict our attention to the case without an additional infimum. We split $\mathbb{H}$ into $\mathbb{A}+\mathbb{B}$ with

$$\mathbb{A}=\frac{1}{2}\operatorname{Tr}\left(\Sigma\Sigma^T(x)D^2f(x)\right)+\int\left[f(x+\mathbf{z})-f(x)-\chi_{B_1(0)}(\mathbf{z})\left\langle\mathbf{z},\nabla f(x)\right\rangle\right]\mu_x(\mathrm{d}\mathbf{z}),$$

and

$$\mathbb{B}=\langle b(x),\nabla f(x)\rangle+\mathcal{H}(\nabla f(x))$$

and specify conditions under which

- we can construct a controlled growth coupling for $\mathbb{A}$,
- we can establish local semi-monotonicity of $\mathbb{B}$,
- we can verify that $\mathbb{A}$ and $\mathbb{B}$ are compatible with the containment function $V(x)=\log\left(1+\frac{1}{2}x^2\right)$ as in Example 3.8 and Jensen penalizations $\{\zeta_{z,p}\}_{z\in\mathbb{R}^q,p\in\mathbb{R}^q}$ and $\{\xi_z\}_{z\in\mathbb{R}^q}$ as in Example 3.11,
- the containment function V is a Lyapunov function for $\mathbb{H}$.

Since we specialize to the Bellman case, property (e) in Theorem 3.12 is equivalent to V being a Lyapunov function.

As all considered functions in Examples 3.8 and 3.11 and s_z are smooth, part (1) of the compatibility assumptions for $\mathbb{A}$ and $\mathbb{B}$, cf. Assumptions 3.16 (a) and (b), hold for every part of $\mathbb{H}$ except the integral term immediately.

To simplify the verification of our conditions, we have the following two observations.

Remark 4.1. Let $\mathbb{A}_1, \mathbb{A}_2 \subseteq C(\mathbb{R}^q) \times C(\mathbb{R}^q)$ be linear on their respective domains and compatible with V, $\{\zeta_{z,p}\}_{z\in\mathbb{R}^q, p\in\mathbb{R}^q}$, and $\{\xi_z\}_{z\in\mathbb{R}^q}$ and with associated controlled growth couplings $\widehat{\mathbb{A}}_1, \widehat{\mathbb{A}}_2 \subseteq C(\mathbb{R}^{2q}) \times C(\mathbb{R}^{2q})$. Then, the operator $\mathbb{A} := \mathbb{A}_1 + \mathbb{A}_2$ is linear on its domain and compatible with V, $\{\zeta_{z,p}\}_{z\in\mathbb{R}^q, p\in\mathbb{R}^q}$, and $\{\xi_z\}_{z\in\mathbb{R}^q}$ and with associated controlled growth coupling $\widehat{\mathbb{A}} := \widehat{\mathbb{A}}_1 + \widehat{\mathbb{A}}_2$.

Remark 4.2. Let $\mathbb{B}_1, \mathbb{B}_2 \subseteq C(\mathbb{R}^q) \times C(\mathbb{R}^q)$ be compatible with V, $\{\zeta_{z,p}\}_{z\in\mathbb{R}^q, p\in\mathbb{R}^q}$, and $\{\xi_z\}_{z\in\mathbb{R}^q}$ and convex semi-monotone operators. Then $\mathbb{B} := \mathbb{B}_1 + \mathbb{B}_2$ is compatible with V, $\{\zeta_{z,p}\}_{z\in\mathbb{R}^q, p\in\mathbb{R}^q}$, and $\{\xi_z\}_{z\in\mathbb{R}^q}$ and convex semi-monotone operator.

The rest of this section is organized as follows:

- In Section 4.1, we consider drift terms and convex first-order Hamiltonians;
- In Section 4.2, we consider diffusion operators;
- In Section 4.3, we consider integral operators;
- In Section 4.4, we consider an illustrative example from optimal control.

4.1. Drift terms and convex first-order Hamiltonians

In this section, we consider the deterministic part of the operator (4.1).

Proposition 4.3. *Suppose that $\mathbb{B}$ is given by*

$$\mathbb{B}f(x) = \langle b(x), \nabla f(x)\rangle + \mathcal{H}(\nabla f(x))$$

with the drift term $x \mapsto b(x)$ locally, one-sided Lipschitz with constant $L_{b,K}$ and $\|b(x)\| \le \frac{c_b}{2}(1 + \|x\|)$ for some constant $c_b > 0$, and $p \mapsto \mathcal{H}(p)$ continuous and convex.
Then, $\mathbb{B}$ is compatible with both collections of Examples 3.8 and 3.11, cf. Assumption 3.16 (b), and convex semi-monotone. Furthermore, $V = \log(1 + \frac{x^2}{2})$ is a Lyapunov function for $\mathbb{B}$, cf. Definition 3.9.

Proof. Convex semi-monotonicity: Clearly, $\mathbb{B}$ is locally first-order with $\mathbb{B}f(x) = \langle b(x), \nabla f(x)\rangle + \mathcal{H}(\nabla f(x)) = \mathcal{B}(x, \nabla f(x))$. Additionally, for any compact set $K \subseteq \mathbb{R}^q$, $\alpha > 0$, and $x, x' \in K$, we have

$$\begin{aligned}\mathcal{B}(x, \alpha(x-x')) - \mathcal{B}(y, \alpha(x-x')) &= \langle b(x), \alpha(x-x')\rangle + \mathcal{H}(\alpha(x-x')) \\ &\quad - \langle b(x'), \alpha(x-x')\rangle - \mathcal{H}(\alpha(x-x')) \\ &= \langle b(x) - b(x'), \alpha(x-x')\rangle \\ &\quad + \mathcal{H}(\alpha(x-x')) - \mathcal{H}(\alpha(x-x')) \\ &\le \alpha L_{b,K} d^2(x,x'),\end{aligned} \tag{4.2}$$

establishing semi-monotonicity. As convexity of $p \mapsto \mathcal{B}(x,p)$ is immediate, we conclude that $\mathbb{B}$ is convex semi-monotone.

Lyapunov control: Using that $V(x) = \log\left(1 + \frac{x^2}{2}\right)$, $\nabla V(x) = \frac{2x}{2+|x|^2}$ is bounded as a function of x, b has linear growth, and that $\mathcal{H}$ is continuous, we find that

$$\sup_{x\in\mathbb{R}^q} \mathbb{B}V(x) = \sup_{x\in\mathbb{R}^q} \left\langle b(x), \frac{2x}{2+|x|^2} \right\rangle + \mathcal{H}\left(\frac{2x}{2+|x|^2}\right) < \infty.$$

Compatibility: We show the compatibility of $\mathbb{B}$, cf. Assumption 3.16 (b), by evaluation of the perturbation and containment function in the operator.
Using $\xi_z(x) = \frac{1}{2}d^2(x,z)$ and $\zeta_{z,p}(x) = \langle p, x-z\rangle$, we find for $z_0, z_1, z \in \mathbb{R}^q$ and $p \in B_1(0)$

$$\begin{aligned}\mathbb{B}(\Xi_{z_0,p,z_1} \circ s_z)(x) &= \langle b(x), (x-z-z_0) + p + (x-z-z_1)\rangle \\ &\quad + \mathcal{H}\left((x-z-z_0) + p + (x-z-z_1)\right),\end{aligned}$$

which is continuous in (x, z_0, p, z_1, z) as b and $\mathcal{H}$ are continuous. For $V(x) = \log\left(1 + \frac{1}{2}x^2\right)$ and $z \in \mathbb{R}^q$, we find

$$\mathbb{B}(V\circ s_z)(x) = \left\langle b(x), \frac{2(x-z)}{2+|x-z|^2} \right\rangle + \mathcal{H}\left(\frac{2(x-z)}{2+|x-z|^2}\right),$$

which is continuous in (x, z) as b and $\mathcal{H}$ are continuous. Thus, $\mathbb{B}$ is compatible. □

4.2. Diffusion operators

In this section, we focus on diffusion operators of the form

$$\mathbb{A}f(x) = \frac{1}{2}\operatorname{Tr}\left(\Sigma(x)\Sigma^T(x)D^2 f(x)\right),$$

where $\Sigma(x)$ is a positive semi-definite matrix for each fixed $x \in \mathbb{R}^q$.

Our main goal is to construct a controlled growth coupling for the operator $\mathbb{A}$. To illustrate the idea behind our approach, consider the simpler case of $\left(\frac{1}{2}\text{ times}\right)$ the Laplacian operator

$$\mathbb{A}_0 f(x) = \frac{1}{2}\operatorname{Tr}(D^2 f(x)),$$

which is the infinitesimal generator of a Brownian motion. The well-known synchronous coupling of two Brownian motions started from x and x', respectively, is given by $(X(t), X'(t)) = (x + B(t), x' + B(t))$ with $B(t)$ a standard Brownian motion, having generator

$$\widehat{\mathbb{A}}_0 g(x,x') = \frac{1}{2}\left(\partial_x + \partial_{x'}\right)^2 g(x,x'),$$

which satisfies $\widehat{\mathbb{A}}_0 d^2 = 0$. Aiming to generalize this, we rewrite

$$\widehat{\mathbb{A}}_0 g(x,x') = \operatorname{Tr}\left(CC^T D^2 g(x,x')\right) \quad \text{with} \quad C = \frac{1}{\sqrt{2}}\begin{pmatrix}\mathbb{1} & \mathbb{1}\\ \mathbb{1} & \mathbb{1}\end{pmatrix}.$$

In general we obtain the following result.

Proposition 4.4. *Suppose that $\mathbb{A}$ is given by*

$$\mathbb{A}f(x) = \frac{1}{2}\operatorname{Tr}\left(\Sigma(x)\Sigma^T(x)D^2 f(x)\right)$$

with $\Sigma(x)$ positive semi-definite for all $x \in \mathbb{R}^q$, $x \mapsto \Sigma(x)$ locally Lipschitz with constant $L_{\Sigma,K}$ and $\|\Sigma(x)\| \leq \frac{c_\Sigma}{2}(1 + \|x\|)$ for some constant $c_\Sigma > 0$. Consider

$$\widehat{\mathbb{A}}f(x,x') := \operatorname{Tr}\left(\widehat{\Sigma}^2(x,x')D^2 f(x,x')\right),$$

where

$$\widehat{\Sigma}^2(x,x') := \begin{pmatrix} \Sigma(x)\Sigma^T(x) & \Sigma(x')\Sigma^T(x) \\ \Sigma(x)\Sigma^T(x') & \Sigma(x')\Sigma^T(x') \end{pmatrix}.$$

Then, $\mathbb{A}$ is compatible, cf. Assumption 3.16 (a), linear on its domain, admitting the controlled growth coupling $\widehat{\mathbb{A}}$. Furthermore, $V = \log(1 + \frac{x^2}{2})$ is a Lyapunov function for $\mathbb{A}$, cf. Definition 3.9.

For the proof we make use of the following auxiliary lemma.

Lemma 4.5. *For each $x \in \mathbb{R}^q$, let $B(x)$ be a positive semi-definite matrix and consider*

$$\mathbb{A}f(x) = \frac{1}{2}\operatorname{Tr}\left(B(x)D^2 f(x)\right).$$

For any $x, x' \in \mathbb{R}^q$, let $\widehat{B}(x,x')$ be a positive semi-definite matrix having block-structure

$$\widehat{B}(x,x') = \begin{pmatrix} B(x) & B(x,x') \\ B(x,x')^T & B(x') \end{pmatrix}.$$

Define

$$\widehat{\mathbb{A}}f(x,x') := \frac{1}{2}\operatorname{Tr}\left(\widehat{B}(x,x')D^2 f(x,x')\right).$$

Then, $\widehat{\mathbb{A}}$ is a coupling of $\mathbb{A}$.

Proof. We have

$$\begin{aligned} \widehat{\mathbb{A}}(f_1 \oplus f_2)(x,x') &= \frac{1}{2}\operatorname{Tr}\left(\widehat{B}(x,x')D^2(f_1 \oplus f_2)(x,x')\right) \\ &= \frac{1}{2}\operatorname{Tr}\left(B(x)D^2 f(x)\right) + \frac{1}{2}\operatorname{Tr}\left(B(x')D^2 f(x')\right) \\ &= \mathbb{A}f_1(x) + \mathbb{A}f_2(x') \end{aligned}$$

and $\widehat{\mathbb{A}}$ satisfies the maximum principle by definition. □

Proof of Proposition 4.4. Controlled growth coupling: By Lemma 4.5, $\widehat{\mathbb{A}}$ is a coupling for $\mathbb{A}$. We thus verify that $\widehat{\mathbb{A}}$ has controlled growth. Consider $\alpha > 1$, $K \subseteq \mathbb{R}^q$ a compact set, and $x, x', y, y' \in K$. Then,

$$\begin{aligned} \widehat{\mathbb{A}}\left(\frac{\alpha}{2}d^2_{x-y,x'-y'}\right)(x,x') &= \frac{1}{2}\operatorname{Tr}\left(\widehat{\Sigma}^2(x,x')D^2\left(\frac{\alpha}{2}d^2_{x-y,x'-y'}\right)(x,x')\right) \\ &= \frac{1}{2}\operatorname{Tr}\left(\widehat{\Sigma}^2(x,x')\left(\alpha\begin{pmatrix} \mathbb{1} & -\mathbb{1} \\ -\mathbb{1} & \mathbb{1} \end{pmatrix}\right)(x,x')\right) \\ &= \frac{\alpha}{2}\operatorname{Tr}((\Sigma^T(x) - \Sigma^T(x'))(\Sigma(x) - \Sigma(x'))) \\ &\leq \alpha L^2_{\Sigma,K} d^2(x,x'), \end{aligned} \tag{4.3}$$

establishing controlled growth.

Lyapunov control: Using $V(x) = \log(1 + \frac{x^2}{2})$ and the fact that Σ has linear growth, we find that

$$\sup_{x\in\mathbb{R}^q} \mathbb{A}V(x) = \sup_{x\in\mathbb{R}^q} \frac{1}{2} \operatorname{Tr}\left(\Sigma(x)\Sigma^T(x)D^2V(x)\right) < \infty. \tag{4.4}$$

Compatibility: Using $\xi_z(x) = \frac{1}{2}d^2(x,z)$, $\zeta_{z,p}(x) = \langle p, x-z\rangle$ and $V(x) = \log(1+\frac{x^2}{2})$, we find for $z_0, z_1, z \in \mathbb{R}^q$ and $p \in B_1(0)$

$$\begin{aligned}\mathbb{A}(\Xi \circ s_z)(x) &= 2\operatorname{Tr}(\Sigma(x)\Sigma^T(x)),\\ \mathbb{A}(V \circ s_z)(x) &= \frac{1}{2}\operatorname{Tr}\left(\Sigma(x)\Sigma^T(x)D^2(V\circ s_z)(x)\right),\end{aligned}$$

which, by an analogous calculation as in equation (4.4), is continuous in (x, z_0, p, z_1, z) and (x, z). Consequently, $\mathbb{A}$ is compatible. □

4.3. Integral operators

In this section, we cover examples of spatially inhomogeneous Lévy processes that have generators of the type

$$\mathbb{A}f(x) = \int \left[f(x+\mathbf{z}) - f(x) - \chi_{B_1(0)}(\mathbf{z})\langle \mathbf{z}, \nabla f(x)\rangle\right]\mu_x(\mathrm{d}\mathbf{z}), \tag{4.5}$$

where $\chi_{B_1(0)}(\mathbf{z}) = l(|\mathbf{z}|)$ for some smooth non-decreasing function l satisfying $l = 1$ on a neighborhood of 0 and $l(r) = 0$ for $r \geq 1$.

We next specify the space from which we can take our jump measures μ_x. For this, we need to control the mass close to 0 as for large values of $\mathbf{z}$. The following function controls both:

$$W(\mathbf{z}) := \chi_{B_1(0)}(\mathbf{z})|\mathbf{z}|^2 + (1 - \chi_{B_1(0)}(\mathbf{z}))\log\left(1 + |\mathbf{z}|^2\right).$$

We take the family of jump measures $\{\mu_x\}_{x\in\mathbb{R}^q}$ from the set of equivalence classes $\mathcal{M}_W(\mathbb{R}^q) := \mathcal{M}(\mathbb{R}^q)/\sim$ with

$$\mathcal{M}(\mathbb{R}^q) := \left\{\mu \in M(\mathbb{R}^q) \,\middle|\, \int W(\mathbf{z})\,\mu(\mathrm{d}\mathbf{z}) < \infty\right\},$$

where $M(\mathbb{R}^q)$ is the set of all Borel measures on $\mathbb{R}^q$ and where

$$\mu \sim \nu \qquad \text{if and only if} \qquad \mu|_{\mathbb{R}^q\setminus\{0\}} = \nu|_{\mathbb{R}^q\setminus\{0\}}.$$

We topologize the set $\mathcal{M}_W(\mathbb{R}^q)$ by the weak topology σ_W induced by the pairings

$$\mu \mapsto \int g(\mathbf{z})\mu(\mathrm{d}\mathbf{z}) \qquad \text{for all } g \in C_W, \tag{4.6}$$

where

$$C_W := \left\{g \in C(\mathbb{R}^q) \,\middle|\, g(0) = 0, \text{ and } \sup_{\mathbf{z}\neq 0}\frac{|g(\mathbf{z})|}{W(\mathbf{z})} < \infty.\right\}.$$

Below, we construct controlled growth couplings for operators of the type (4.5). To clarify the concepts, we consider the example of an uncompensated process, i.e., having an operator of the type

$$\mathbb{A}f(x) = \int f(x+\mathbf{z}) - f(x)\mu_x(\mathrm{d}\mathbf{z}).$$

Controlled growth couplings for this type of operator are of the form

$$\widehat{\mathbb{A}}f(x,x') = \int \left[f(x+\mathbf{z}_1, y+\mathbf{z}_2) - f(x,x')\right]\pi_{x,x'}(\mathrm{d}\mathbf{z}_1, \mathbf{z}_2), \tag{4.7}$$

where $\pi_{x,x'}$ couples μ_x and $\mu_{x'}$. In the following example, we illustrate the need of being able to couple jumps synchronously.

Example 4.6 *(Random walk).* Consider the simple random walk on $\mathbb{R}$ making jumps of size 1, i.e., $\mu_x = \mu = \delta_{-1} + \delta_1$ leading to the operator

$$\mathbb{A}f(x) = \left[f(x-1) + f(x+1) - 2f(x)\right].$$

Well known couplings include walks with simultaneous jumps but independent directions, fully independent jumps, and synchronous jumps. The corresponding generators are given as in (4.7) with jump measures

$$\begin{aligned}
\pi^1 &:= \mu \otimes \mu, \\
\pi^2 &:= \delta_{(-1,0)} + \delta_{(1,0)} + \delta_{(0,-1)} + \delta_{(0,1)}, \\
\pi^3 &:= \delta_{(-1,-1)} + \delta_{(1,1)},
\end{aligned}$$

respectively. This leads to the operators

$$\begin{aligned}
\widehat{\mathbb{A}}^1 f(x,x') &= f(x-1,x'-1) + f(x-1,x'+1) \\
&\qquad + f(x+1,x'-1) + f(x+1,x'+1) - 4f(x,x'), \\
\widehat{\mathbb{A}}^2 f(x,x') &= f(x-1,x') + f(x+1,x') - 2f(x,x') \\
&\qquad + f(x,x'-1) + f(x,x'+1) - 2f(x,x'), \\
\widehat{\mathbb{A}}^3 f(x,x') &= f(x-1,x'-1) + f(x+1,x'+1) - 2f(x,x').
\end{aligned}$$

Only for the final example, we see that $\widehat{\mathbb{A}}^3 d^2 \le 0$, pointing at the necessity of the alignment of jumps.

Note that the third coupling above has different total mass, and we thus work outside the realm of the typical notion of couplings of probability measures. A second feature of coupling jump measures, not present in the example above, is that we can make one process jump, whereas the other does not.

We formalize this in the following definition.

Definition 4.7. Let $\mu, \nu \in \mathcal{M}_W(\mathbb{R}^q)$. We say that $\pi \in M(\mathbb{R}^q \times \mathbb{R}^q)$ is an *extended coupling of* μ *and* ν, if

$$\begin{aligned}
\pi\left((A \setminus \{0\}) \times \mathbb{R}^q\right) &= \mu(A \setminus \{0\}) \qquad \text{for all } A \in \mathcal{B}(\mathbb{R}^q), \\
\pi\left(\mathbb{R}^q \times (B \setminus \{0\})\right) &= \nu(B \setminus \{0\}) \qquad \text{for all } B \in \mathcal{B}(\mathbb{R}^q).
\end{aligned}$$

Remark 4.8. A variant of this coupling was introduced in [23]. The mass can be moved to the boundary of a domain. In our context, this boundary is the point 0.

Definition 4.9. Let $x \mapsto \mu_x$ be a map from $\mathbb{R}^q$ into $\mathcal{M}(\mathbb{R}^q)$. Let $(x,x') \mapsto \pi_{x,x'}$ be a map from $\mathbb{R}^{2q}$ into $M(\mathbb{R}^q \times \mathbb{R}^q)$.

(a) We say that $(x, x') \mapsto \pi_{x,x'}$ is an extended coupling of $x \mapsto \mu_x$, if, for all $x, x' \in \mathbb{R}^q$, we have that $\pi_{x,x'}$ is an extended coupling of μ_x and μ'_x.
(b) We say that $(x, x') \mapsto \pi_{x,x'}$ is locally Lipschitz, if, for any compact set $K \subseteq \mathbb{R}^q$, there exists a constant $L_{\pi,K}$ such that, for $x, x' \in K$, we have

$$\int d^2(\mathbf{z}_1, \mathbf{z}_2)\pi_{x,x'}(\mathrm{d}\mathbf{z}_1, \mathrm{d}\mathbf{z}_2) \leq L_{\pi,K} d^2(x, x').$$

Remark 4.10. Note that conditions (12), (34), and (35) in [7] for μ and j correspond to our choice of $\mathcal{M}_W(\mathbb{R}^q)$ and locally Lipschitz extended coupling $\pi_{x,x'}$.

Remark 4.11. Let $\eta\colon \mathbb{R} \to \mathbb{R}$ be any locally Lipschitz map with local Lipschitz constants $L_{\eta,K}$. Set $\mu_x := \delta_{\eta(x)} \mathbb{1}_{\eta(x)\neq 0}(x)$ and $\pi_{x,x'} = \delta_{(\eta(x),\eta(x'))}$. Then, $(x, x') \mapsto \pi_{x,x'}$ is a locally Lipschitz coupling of $x \mapsto \mu_x$ with $L_{\pi,K} = L_{\eta,K}$.

The main proposition of this subsection below aims to show that integral operators of the form (4.5) can be treated analogous to the other examples above. We work with the second collection of penalization functions, cf. Definition 3.11, to avoid integrability issues.

Proposition 4.12. *Consider*

$$\mathbb{A}f(x) = \int \left[f(x + \mathbf{z}) - f(x) - \chi_{B_1(0)} \langle \mathbf{z}, \nabla f(x)\rangle\right] \mu_x(\mathrm{d}\mathbf{z}).$$

Suppose there exists a σ_W-continuous map $x \mapsto \mu_x$ in $\mathcal{M}_W(\mathbb{R}^q)$, cf. (4.6), and that there exists a locally Lipschitz extended coupling $(x, x') \mapsto \pi_{x,x'}$ of $x \mapsto \mu_x$ with Lipschitz constant $L_{\pi,K}$ and, for $\widehat{\chi}(\mathbf{z}_1, \mathbf{z}_2) := \chi_{B_1(0)}(\mathbf{z}_1)\chi_{B_1(0)}(\mathbf{z}_2)$, set

$$\begin{aligned}\widehat{\mathbb{A}}g(x, x') := \int \Big[&g(x + \mathbf{z}_1, x' + \mathbf{z}_2) - g(x, x') \\ &- \widehat{\chi}(\mathbf{z}_1, \mathbf{z}_2) \left\langle (\mathbf{z}_1, \mathbf{z}_2)^T, \nabla g(x, x')\right\rangle \Big]\pi_{x,x'}(\mathrm{d}\mathbf{z}_1, \mathrm{d}\mathbf{z}_2).\end{aligned}$$

Assume furthermore that

$$\sup_{x\in\mathbb{R}^q} \int \log\left(1 + \frac{\frac{1}{2}|\mathbf{z}|^2 + \langle x, \mathbf{z}\rangle}{1 + \frac{1}{2}|x|^2}\right) \mu_x(\mathrm{d}\mathbf{z}) < \infty.$$

Then, $\mathbb{A}$ is compatible, cf. Definition 3.16 (a), and linear on its domain, admitting the controlled growth coupling $\widehat{\mathbb{A}}$. Furthermore, $V = \log(1 + \frac{x^2}{2})$ is a Lyapunov function for $\mathbb{A}$, cf. Definition 3.9.

The proof of Proposition 4.12 is based on the following two auxiliary lemmas. In the first, we obtain bounds on the integrand of our operator acting on the Lyapunov function V. In the second, we compute the integrand of the Lévy-type operator acting on the shifted squared metric. We prove these two lemmas following the proof of Proposition 4.12.

Lemma 4.13. *Fix $x, z \in \mathbb{R}^q$.*

(a) *For $\mathbf{z} \in \mathbb{R}^q$, we have*

$$-\log\left(1 + \frac{1}{2}|x - z|^2\right) \leq V \circ s_z(x + \mathbf{z}) - V \circ s_z(x)$$

$$\leq \log\left(1+\frac{\frac{1}{2}|\mathbf{z}|^2+\langle x-z,\mathbf{z}\rangle}{1+\frac{1}{2}|x-z|^2}\right)\leq \log\left(1+|\mathbf{z}|^2\right).$$

(b) *For* $\mathbf{z}\in B_1(0)$*, we have*

$$|V\circ s_z(x+\mathbf{z})-V\circ s_z(x)-\langle \mathbf{z},\nabla(V\circ s_z)(x)\rangle|\leq \frac{1}{2}|\mathbf{z}|^2.$$

Lemma 4.14. *We have*

$$\begin{aligned}&\frac{1}{2}d^2_{x-y,x'-y'}(x+\mathbf{z}_1,x'+\mathbf{z}_2)-\frac{1}{2}d^2_{x-y,x'-y'}(x,x')\\&\quad-\widehat{\chi}(\mathbf{z}_1,\mathbf{z}_2)\left\langle\begin{pmatrix}\mathbf{z}_1\\ \mathbf{z}_2\end{pmatrix},\nabla\left(\frac{1}{2}d^2_{x-y,x'-y'}\right)(x,x')\right\rangle\\&\leq\left(1-\frac{1}{2}\widehat{\chi}(\mathbf{z}_1,\mathbf{z}_2)\right)d^2(\mathbf{z}_1,\mathbf{z}_2)+(1-\widehat{\chi}(\mathbf{z}_1,\mathbf{z}_2))\frac{1}{2}d^2(y,y').\end{aligned}$$

Proof of Proposition 4.12. Controlled growth coupling: As $(x,x')\mapsto\pi_{x,x'}$ is a locally Lipschitz extended coupling of $x\mapsto\mu_x$, cf. Definition 4.9, we have that $\widehat{\mathbb{A}}$ is a coupling. Thus, we need to verify the controlled growth property of $\widehat{\mathbb{A}}$.

Let $x,x',y,y'\in K$ for $K\subseteq\mathbb{R}^q$ a compact set. Using Lemma 4.14, we then have

$$\begin{aligned}\widehat{\mathbb{A}}\left(\frac{\alpha}{2}d^2_{x-y,x'-y'}\right)(x,x')&\leq\frac{\alpha}{2}\int\left(1-\frac{1}{2}\widehat{\chi}(\mathbf{z}_1,\mathbf{z}_2)\right)d^2(\mathbf{z}_1,\mathbf{z}_2)\pi_{x,x'}(\mathrm{d}\mathbf{z}_1,\mathrm{d}\mathbf{z}_2)\\&\quad+\frac{\alpha}{2}\int(1-\widehat{\chi}(\mathbf{z}_1,\mathbf{z}_2))\frac{1}{2}d^2(y,y')\pi_{x,x'}(\mathrm{d}\mathbf{z}_1,\mathrm{d}\mathbf{z}_2)\\&\leq\frac{\alpha}{2}L_{\pi,K}d^2(x,x')+\frac{\alpha}{4}c'_\pi d^2(y,y'),\end{aligned}$$

where the second inequality is due to the local Lipschitz property of the map $(x,x')\mapsto\pi_{x,x'}$ and $c'_\pi>0$ exists since, for every $x,x'\in\mathbb{R}^q$, $\pi_{x,x'}\in M(\mathbb{R}^q\times\mathbb{R}^q)$. As such, $\mathbb{A}$ admits the controlled growth coupling $\widehat{\mathbb{A}}$.

Lyapunov control: Using Lemma 4.13, we find

$$\sup_{x\in\mathbb{R}^q}\mathbb{A}V(x)\leq\sup_{x\in\mathbb{R}^q}\int(1-\chi_{B_1(0)})\log(1+|\mathbf{z}|^2)+\chi_{B_1(0)}|\mathbf{z}|^2\mu_x(\mathrm{d}\mathbf{z})<\infty.$$

Compatibility: We start by establishing the continuity of $(x,z)\mapsto\mathbb{A}(V\circ s_z)(x)$. Let (x_n,z_n) converge to (x,z). We aim to apply Lemma Appendix A.1 with $\mathcal{X}=\mathbb{R}^q\setminus\{0\}$, $\nu_n=\mu_{x_n}$, and

$$\begin{aligned}\phi_n(\mathbf{z})&:=V\circ s_{z_n}(x_n+\mathbf{z})-V\circ s_{z_n}(x_n)-\chi_{B_1(0)}\langle\mathbf{z},\nabla(V\circ s_{z_n})(x_n)\rangle,\\\phi_\infty(\mathbf{z})&:=V\circ s_z(x+\mathbf{z})-V\circ s_z(x)-\chi_{B_1(0)}\langle\mathbf{z},\nabla(V\circ s_z)(x)\rangle.\end{aligned}$$

As ϕ_n is continuous, it remains to show that $\sup_{n\in\mathbb{N}}\sup_{\mathbf{z}\neq0}\frac{|\phi_n(\mathbf{z})|}{W(\mathbf{z})}<\infty$. By Lemma 4.13, we can estimate

$$|\phi_n(\mathbf{z})|\leq\chi_{B_1(0)}\frac{1}{2}|\mathbf{z}|^2+(1-\chi_{B_1(0)})\max\left\{-\log\left(1+\frac{1}{2}|x_n-z_n|^2\right),\log\left(1+|\mathbf{z}|^2\right)\right\}.$$

Since (x_n,z_n) is convergent, hence bounded, we obtain the desired estimate. Continuity of $(x,z)\mapsto\mathbb{A}(V\circ s_z)(x)$ now follows by Lemma Appendix A.1.

Using the particular form of Ξ_{z_0,p,z_1}, cf. Example 3.11, one readily verifies that the map $(x,z_0,p,z_1,z)\mapsto\mathbb{A}\left(\Xi_{z_0,p,z_1}\circ s_z\right)(x)$ is continuous with an analogous argumentation. □

Proof of Lemma 4.13. Let $y = x - z$, then we can write

$$\begin{aligned} &V \circ s_z(x + \mathbf{z}) - V \circ s_z(x) \\ &\quad = \log\left(1 + \frac{1}{2}(y + \mathbf{z})^2\right) - \log\left(1 + \frac{1}{2}|y|^2\right) = \log\left(1 + \frac{\frac{1}{2}|\mathbf{z}|^2 + \langle y, \mathbf{z}\rangle}{1 + \frac{1}{2}|y|^2}\right). \end{aligned}$$

Applying Young's inequality to $\langle y, \mathbf{z}\rangle$ leads to the upper bound

$$V \circ s_z(x + \mathbf{z}) - V \circ s_z(x) \leq \log\left(1 + \frac{|\mathbf{z}|^2 + \frac{1}{2}|y|^2}{1 + \frac{1}{2}|y|^2}\right) = \log\left(2 + \frac{|\mathbf{z}|^2 - 1}{1 + \frac{1}{2}|y|^2}\right) \leq \log\left(1 + |\mathbf{z}|^2\right).$$

Using that the first term is positive, we obtain the lower bound

$$V \circ s_z(x + \mathbf{z}) - V \circ s_z(x) \geq \log\left(1 - \frac{\frac{1}{2}|y|^2}{1 + \frac{1}{2}|y|^2}\right) = -\log\left(1 + \frac{1}{2}|y|^2\right).$$

This establishes (a). For the proof of (b), we apply Taylor's Theorem to obtain

$$\begin{aligned} |V \circ s_z(x + \mathbf{z}) - V \circ s_z(x) - \langle \nabla(V \circ s_z)(x), \mathbf{z}\rangle| &\leq \frac{1}{2}|\mathbf{z}|^2 \sup_{\mathbf{z} \in B_1(0)} \sup_{i,j} \left|\nabla^2_{i,j} V(y + \mathbf{z})\right| \\ &\leq \frac{1}{2}|\mathbf{z}|^2, \end{aligned}$$

which follows by a direct inspection of

$$\nabla^2_{i,j} V(x) = \frac{2\delta_{i,j}\left(1 + \frac{1}{2}|x|^2\right) - 2x_i x_j}{(1 + \frac{1}{2}|x|^2)^2}. \quad \square$$

Proof of Lemma 4.14. Evaluating the shift maps, calculating the gradient of the squared Euclidean distance, and expanding the squares leads to

$$\begin{aligned} &\frac{1}{2} d^2_{x-y,x'-y'}(x + \mathbf{z}_1, x' + \mathbf{z}_2) - \frac{1}{2} d^2_{x-y,x'-y'}(x, x') \\ &\qquad - \widehat{\chi}(\mathbf{z}_1, \mathbf{z}_2) \left\langle \begin{pmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{pmatrix}, \nabla\left(\frac{1}{2} d^2_{x-y,x'-y'}\right)(x, x') \right\rangle \\ &= \frac{1}{2} d^2(y + \mathbf{z}_1, y' + \mathbf{z}_2) - \frac{1}{2} d^2(y, y') - \widehat{\chi}(\mathbf{z}_1, \mathbf{z}_2) \langle y - y', \mathbf{z}_1 - \mathbf{z}_2\rangle \\ &= \frac{1}{2} d^2(\mathbf{z}_1, \mathbf{z}_2) + \langle y - y', \mathbf{z}_1 - \mathbf{z}_2\rangle - \widehat{\chi}(\mathbf{z}_1, \mathbf{z}_2) \langle y - y', \mathbf{z}_1 - \mathbf{z}_2\rangle \\ &\leq \left(1 - \frac{1}{2}\widehat{\chi}(\mathbf{z}_1, \mathbf{z}_2)\right) d^2(\mathbf{z}_1, \mathbf{z}_2) + (1 - \widehat{\chi}(\mathbf{z}_1, \mathbf{z}_2)) \frac{1}{2} d^2(y, y'), \end{aligned}$$

where in the second equality we use properties of the Euclidean distance d and the final line is due to Young's inequality. $\square$

4.4. Stochastic optimal control

In this section, we showcase how the previous examples can be bootstrapped to the setting of optimal control. We focus on a simplified setting that aims to illustrate the methodology that can be used to extend the above arguments to more complex examples.

For some compact set Θ and a continuous cost functional $\mathcal{I}\colon \Theta \to \mathbb{R}$, we consider the operator

$$\begin{aligned}\mathbb{H}f &= \sup_{\theta\in\Theta}\left\{A_\theta f(x) + B_\theta f(x) - \mathcal{I}(\theta)\right\}\\ &= \sup_{\theta\in\Theta}\left\{\frac{1}{2}\operatorname{Tr}\left(\Sigma(x,\theta)\Sigma^T(x,\theta)D^2 f(x)\right) + \langle b(x,\theta), \nabla f(x)\rangle - \mathcal{I}(\theta)\right\}.\end{aligned}$$

Such operators arise naturally in the context of (stochastic) optimal control, cf. [24,39,46]. We obtain the following result.

Proposition 4.15. *Suppose $\mathbb{H}$ is given by*

$$\mathbb{H}f(x) = \sup_{\theta\in\Theta}\left\{\frac{1}{2}\operatorname{Tr}\left(\Sigma(x,\theta)\Sigma^T(x,\theta)D^2 f(x)\right) + \langle b(x,\theta), \nabla f(x)\rangle - \mathcal{I}(\theta)\right\},$$

where we assume that

- *Θ is compact and $\mathcal{I}\colon \Theta \to \mathbb{R}$ is continuous,*
- *$(x,\theta) \mapsto b(x,\theta)$ is continuous and Lipschitz in x uniformly in θ with constant $L_{b,K}$,*
- *$\Sigma(x,\theta)$ is positive semi-definite for all $x \in \mathbb{R}^q$ and $\theta \in \Theta$, and the map $(x,\theta) \mapsto \Sigma(x,\theta)$ continuous and Lipschitz in x uniformly in θ with constant $L_{\Sigma,K}$,*
- *there are constants $c_b, c_\Sigma > 0$ such that $\|b(x,\theta)\| \le \frac{c_b}{2}(1+\|x\|)$ and $\|\Sigma(x,\theta)\| \le \frac{c_\Sigma}{2}(1+\|x\|)$.*

Then, $\{\langle b(x,\theta), \nabla f(x)\rangle\}_{\theta\in\Theta}$ and $\left\{\frac{1}{2}\operatorname{Tr}\left(\Sigma(x,\theta)\Sigma^T(x,\theta)D^2 f(x)\right)\right\}_{\theta\in\Theta}$ are compatible, $\{\langle b(x,\theta), \nabla f(x)\rangle\}_{\theta\in\Theta}$ is convex semi-monotone, $\left\{\frac{1}{2}\operatorname{Tr}\left(\Sigma(x,\theta)\Sigma^T(x,\theta)D^2 f(x)\right)\right\}_{\theta\in\Theta}$ is linear on its domain and admits a controlled growth coupling. Furthermore, $V = \log(1+\frac{x^2}{2})$ is a Lyapunov function for $\mathbb{H}$, cf. Definition 3.9.

Proof. Convex semi-monotonicity and controlled growth coupling: Let $\theta \in \Theta$. Then, by Proposition 4.3, $\mathcal{B}_\theta(x, \nabla f(x)) = \langle b(x,\theta), \nabla f(x)\rangle$ is locally first-order and, for any compact set $K \subseteq \mathbb{R}^q$, $\alpha > 0$, and $x, x' \in K$, proceeding as in equation (4.2) and using the compactness of Θ and continuity of b, we obtain the estimate

$$\mathcal{B}_\theta(x, \alpha(x-x')) - \mathcal{B}_\theta(y, \alpha(x-x')) \le \alpha L_{b,K} d^2(x,x') \quad \text{for all } \theta \in \Theta.$$

Furthermore, by Lemma 4.5, we can find a family of couplings $\{\widehat{\mathbb{A}}_\theta\}_{\theta\in\Theta}$, which, analogously to Proposition 4.4, has controlled growth so that, using the compactness of Θ and continuity of Σ, for any compact set $K \subseteq \mathbb{R}^q$, $\alpha > 0$, and $x, x' \in K$, we obtain an estimate similar to equation (4.3) uniformly in θ, i.e.

$$\widehat{\mathbb{A}}_\theta\left(\frac{\alpha}{2} d^2_{x-y,x'-y'}\right)(x,x') \le \alpha L^2_{\Sigma,K} d^2(x,x') \quad \text{for all } \theta \in \Theta.$$

Lyapunov control: As $\mathcal{I}$ is continuous and Θ is compact, $c_{\mathcal{I}} \coloneqq \inf_{\theta\in\Theta}\mathcal{I}(\theta) > -\infty$. Thus,

$$\begin{aligned}\sup_{x\in\mathbb{R}^q}\mathbb{H}f(x) &= \sup_{x\in\mathbb{R}^q}\sup_{\theta\in\Theta}\left\{\frac{1}{2}\operatorname{Tr}\left(\Sigma(x,\theta)\Sigma^T(x,\theta)D^2 f(x)\right) + \langle b(x,\theta), \nabla f(x)\rangle - \mathcal{I}(\theta)\right\}\\ &\le \sup_{x\in\mathbb{R}^q}\sup_{\theta\in\Theta}\left\{\frac{1}{2}\operatorname{Tr}\left(\Sigma(x,\theta)\Sigma^T(x,\theta)D^2 f(x)\right) + \langle b(x,\theta), \nabla f(x)\rangle\right\} - c_{\mathcal{I}}.\end{aligned}$$

Now, Lyapunov control follows analogously to Propositions 4.3 and 4.4.

Compatibility: As the coefficients $(x,\theta) \mapsto b(x,\theta)$ and $(x,\theta) \mapsto \Sigma(x,\theta)$ are continuous, analogous arguments as in Propositions 4.3 and 4.4 show the compatibility of $\{\langle b(x,\theta), \nabla f(x)\rangle\}_{\theta\in\Theta}$ and $\left\{\frac{1}{2}\operatorname{Tr}\left(\Sigma(x,\theta)\Sigma^T(x,\theta)D^2 f(x)\right)\right\}_{\theta\in\Theta}$, respectively. □

5. Construction of optimizers and test functions

In classical proofs of comparison principles, the approach to estimate $\sup u - v$ for a subsolution u and supersolution v is variable doubling or quadruplication, cf. [6, Theorem 3.1] or [18, introduction of Section 3]: For $\alpha > 1$,

$$\sup_{x\in\mathbb{R}^q} u(x) - v(x) \le \sup_{x,x'\in\mathbb{R}^q} u(x) - v(x') - \frac{\alpha}{2} d^2(x,x'). \tag{5.1}$$

Letting $\alpha \to \infty$ forces optimizing points of the right-hand side together, if they exist. In addition, by varying either of the two components, one obtains basic test functions in terms of $\frac{\alpha}{2}d^2$ for the use in the definition of the sub- and supersolution properties of u and v.

To ensure that optimizers in (5.1) exist, for small $\varepsilon > 0$, we consider the following problem that includes the containment function V and upper bounds $\sup u - v$ up to a term of order ε instead:

$$\begin{aligned}\sup_{x\in\mathbb{R}^q} \frac{1}{1-\varepsilon} u(x) - \frac{1}{1+\varepsilon} v(x)& \\ \le \sup_{x,x'\in\mathbb{R}^q} \frac{1}{1-\varepsilon} u(x) - \frac{1}{1+\varepsilon} v(x') - \frac{\alpha}{2} d^2(x,x') - \frac{\varepsilon}{1-\varepsilon} V(x) - \frac{\varepsilon}{1+\varepsilon} V(x').&\end{aligned} \tag{5.2}$$

The particular form of the factors $1-\varepsilon$ and $1+\varepsilon$ is motivated by convexity based arguments, which will show up in the proofs of Proposition 6.3 and Theorem 3.12 below.

The procedure in (5.2) would be sufficient for a standard, first-order Hamilton–Jacobi equation. The test functions produced by this procedure, however, will not be sufficient to treat second-order or integral operators. This problem was considered in [7] and [18]. We will follow their approach by considering a quadruplication-of-variables, which we also phrase in terms of sup- and inf-convolutions, cf. Definition 5.1. We then perform a Jensen-type perturbation.

As we aim to unify proofs for both integral and differential operators, we revisit the full proof and state our result in terms of test functions.

In Proposition 5.2 and Theorem 5.10 below, which can be considered to be an extended two-variable variant of the Crandall–Ishii construction [18, Theorem 3.2], we start out by considering the optimization (5.2) in terms of the sup- and inf-convolution of u and v, respectively, effectively leading to a quadruplication problem, see (5.5) below.

We then perform the Jensen perturbation, see (5.6). The rest of the proposition deals with various properties of the optimizers in relation to u and v.

In Theorem 5.10, we carry out an additional layer of smoothing operations to obtain C^∞ test functions. Consequently, we can move away from the notion of solutions in terms of sub- and superjets, which is of paramount importance to effectively treat diffusive and jump-type processes in a common framework.

For readability, we express suprema and infima using $\lceil\cdot\rceil$ and $\lfloor\cdot\rfloor$, respectively, as defined in Section 2.1. Furthermore, we define the sup- and inf-convolutions as follows.

Definition 5.1 *(*sup- *and* inf*-convolution).* Let $u\colon \mathbb{R}^q \to \mathbb{R}$ be bounded and upper semi-continuous and $v\colon \mathbb{R}^q \to \mathbb{R}$ be bounded and lower semi-continuous. For $\alpha > 1$, we define the sup*-convolution* $P^\alpha[u]$ *of* u as

$$P^\alpha[u](y) := \sup_{x\in\mathbb{R}^q} \left\{ u(x) - \frac{\alpha}{2} d^2(x,y) \right\} = \left\lceil u - \frac{\alpha}{2} d^2(\cdot,y) \right\rceil. \tag{5.3}$$

Analogously, we define the inf*-convolution* $P_\alpha[v]$ *of* v as

$$P_\alpha[v](y) := \inf_{x\in\mathbb{R}^q} \left\{ v(x) + \frac{\alpha}{2} d^2(x,y) \right\} = \left\lfloor u + \frac{\alpha}{2} d^2(\cdot,y) \right\rfloor. \tag{5.4}$$

5.1. Construction of optimizers

The main result of this section is the construction of optimizers in Proposition 5.2 below. After this, we state two auxiliary results: In Proposition 5.3, we collect various useful properties of the sup- and inf-convolutions. The results of Proposition 5.4 allow us to perturb a semi-convex function with a unique extreme point such that we get a new extreme point close by, in which the second derivative of the function exists. The result is a variant of the well-known perturbation result by Jensen, see e.g. [18, Lemma A.3].

Proposition 5.2 *(Construction of optimizers). Let u be bounded and upper semi-continuous, v be bounded and lower semi-continuous, V be a containment function as in Definition 3.7, and $\{\zeta_{z,p}\}_{z\in\mathbb{R}^q,p\in\mathbb{R}^q} \subseteq C(\mathbb{R}^q)$ and $\{\xi_z\}_{z\in\mathbb{R}^q} \subseteq C^1(\mathbb{R}^q)$ be collections of point penalizations as in Definition 3.10. Fix $\varepsilon \in (0,1)$ and $\varphi \in (0,1]$.*

Then, there exist compact sets $K_{\varepsilon,0} \subseteq K_\varepsilon \subseteq \mathbb{R}^q$ and, for any $\alpha > 1$, three pairs of variables $(y_{\alpha,0}, y'_{\alpha,0})$, (y_α, y'_α), (x_α, x'_α) in $\mathbb{R}^{2q}$ and $p_\alpha, p'_\alpha \in B_{1/\alpha}(0)$ such that the following four sets of properties hold.

Properties of $y_{\alpha,0}, y'_{\alpha,0}$

The variables $y_{\alpha,0}, y'_{\alpha,0}$ optimize $\lceil \Lambda_\alpha \rceil$, where

$$\begin{aligned}\Lambda_\alpha(y,y') :=& \frac{1}{1-\varepsilon}P^\alpha[u](y) - \frac{1}{1+\varepsilon}P_\alpha[v](y') - \frac{\alpha}{2}d^2(y,y')\\ &- \frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y) - \frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y')\end{aligned} \tag{5.5}$$

and satisfy the following property

(a) $y_{\alpha,0}, y'_{\alpha,0} \in K_{\varepsilon,0}$.

Properties of y_α, y'_α and p_α, p'_α

The pair y_α, y'_α optimizes

$$\left\lceil \Lambda_\alpha - \frac{\varepsilon}{1-\varepsilon}\varphi\Xi_1^0 - \frac{\varepsilon}{1+\varepsilon}\varphi\Xi_2^0 \right\rceil \tag{5.6}$$

and uniquely optimizes

$$\left\lceil \Lambda_\alpha - \frac{\varepsilon}{1-\varepsilon}\varphi\Xi_1 - \frac{\varepsilon}{1+\varepsilon}\varphi\Xi_2 \right\rceil, \tag{5.7}$$

where Λ_α is as in (5.5) and

$$\Xi_1^0(y) := \Xi^0_{y_{\alpha,0},p_\alpha}(y), \qquad \Xi_2^0(y') := \Xi^0_{y'_{\alpha,0},p'_\alpha}(y'),$$

$$\Xi_1(y) := \Xi_{y_{\alpha,0},p_\alpha,y_\alpha}(y), \qquad \Xi_2(y') := \Xi_{y'_{\alpha,0},p'_\alpha,y'_\alpha}(y')$$

are as in Definition 3.10. Moreover, the optimizers y_α, y'_α of (5.6) and (5.7) satisfy

(b) *We have*

$$d(y_\alpha, y_{\alpha,0}) \leq \frac{1}{\alpha}, \qquad d(y'_\alpha, y'_{\alpha,0}) \leq \frac{1}{\alpha}.$$

(c) *The second derivatives of* $P^\alpha[u]$ *and* $P_\alpha[v]$ *in* y_α *and* y'_α *exist, respectively.*

Properties of x_α, x'_α

The variables x_α, x'_α *optimize*

$$P^\alpha[u](y_\alpha) = u(x_\alpha) - \frac{\alpha}{2} d^2(x_\alpha, y_\alpha),$$
$$P_\alpha[v](y'_\alpha) = v(x'_\alpha) + \frac{\alpha}{2} d^2(x'_\alpha, y'_\alpha),$$

and satisfy

(d) x_α *and* x'_α *are the unique optimizers in the definition of* $P^\alpha[u](y_\alpha)$ *and* $P_\alpha[v](y'_\alpha)$*, respectively.*
(e) *We have that*

$$u(x_\alpha) - P^\alpha[u] \circ s_{x_\alpha - y_\alpha}(x_\alpha) = \lceil u - P^\alpha[u] \circ s_{x_\alpha - y_\alpha} \rceil,$$
$$v(x'_\alpha) - P_\alpha[v] \circ s_{x'_\alpha - y'_\alpha}(x'_\alpha) = \lfloor v - P_\alpha[v] \circ s_{x'_\alpha - y'_\alpha} \rfloor.$$

Behavior as $\alpha \to \infty$

(f) *We have* $\lim_{\alpha\to\infty} \alpha d^2(y_{\alpha,0}, y'_{\alpha,0}) = 0$.
(g) *We have*

$$\lim_{\alpha\to\infty} \alpha \left(d\left(x_\alpha, y_\alpha\right) + d\left(y_\alpha, y'_\alpha\right) + d\left(y'_\alpha, x'_\alpha\right)\right)^2 = 0.$$

(h) $x_\alpha, y_\alpha, y'_\alpha, x'_\alpha \in K_\varepsilon$.

In addition, the following estimate on $u - v$ *holds: For any compact set* $K \subseteq \mathbb{R}^q$*, there is a compact set* $\widehat{K} = \widehat{K}(K, \varepsilon, u, v)$ *given by*

$$\widehat{K} := \left\{ z \in \mathbb{R}^q \,\middle|\, V(z) \le \frac{\|u\| + \|v\|}{\varepsilon} + \lceil V \rceil_K \right\},$$

so that

(i) *For any compact set* $K \subseteq \mathbb{R}^q$,

$$\lceil u - v \rceil_K \le \frac{1}{1-\varepsilon} u(x_\alpha) - \frac{1}{1+\varepsilon} v(x'_\alpha) + \varepsilon\left(c_{\varepsilon,\varphi} + o(1)\right),$$

where

$$c_{\varepsilon,\varphi} := \frac{2}{1-\varepsilon^2}(1-\varphi)\lceil V \rceil_K - \left\lfloor \frac{1}{1-\varepsilon} u - \frac{1}{1+\varepsilon} v \right\rfloor_K,$$

and $o(1)$ *is in terms of* $\alpha \to \infty$ *for fixed* ε *and* φ.
(j) *Any limit point of the sequence* $(x_\alpha, y_\alpha, y_{\alpha,0}, y'_{\alpha,0}, y'_\alpha, x'_\alpha)$ *as* $\alpha \to \infty$ *is of the form* (z, z, z, z, z, z) *with* $z \in \widehat{K}$.

Fig. 1 visualizes the relation between the different optimizing points.

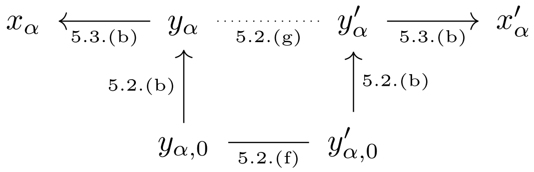

Fig. 1. Relation between the optimizing points with a note which parts of the propositions give us distance control.

Proof of Proposition 5.2. Proof of (a): As u and v are bounded, by Lemma 5.3 (a), the same holds for $\|P^\alpha[u]\|$ and $\|P_\alpha[v]\|$. Using that V has compact sublevel sets, cf. Definition 3.7, the existence of optimizers $(y_{\alpha,0}, y'_{\alpha,0})$ for $\lceil \Lambda_\alpha \rceil$ follows.

The definitions of Λ_α and the convolutions $P^\alpha[u]$ and $P_\alpha[v]$ imply that

$$\frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_{\alpha,0}) + \frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_{\alpha,0}) \leq \frac{1}{1-\varepsilon}\lceil u \rceil - \frac{1}{1+\varepsilon}\lfloor v \rfloor - \lceil \Lambda_\alpha \rceil . \tag{5.8}$$

Comparing the optimizers for $\Lambda_\alpha(y, y')$ to, e.g., the suboptimial choice $(y, y') = (\hat{y}, \hat{y})$ satisfying $V(\hat{y}) = 0$, we find

$$\frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_{\alpha,0}) + \frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_{\alpha,0}) \leq \frac{2}{1-\varepsilon}\|u\| + \frac{2}{1+\varepsilon}\|v\| .$$

From this estimate, we deduce that $(y_{\alpha,0}, y'_{\alpha,0}) \in K_{\varepsilon,0} \times K_{\varepsilon,0}$ with

$$K_{\varepsilon,0} := \left\{ y \in \mathbb{R}^q \,\middle|\, V(y) \leq \varepsilon^{-1} C_\varepsilon (\|u\| + \|v\|) \right\}$$

for some constant $C_\varepsilon > 0$ satisfying $\lim_{\varepsilon \downarrow 0} C_\varepsilon = \frac{2}{1-\varphi}$, establishing (a).

Proof of (b) and (c): For the proof of these two statements, we first move from $\lceil \Lambda_\alpha \rceil$ to its perturbed version (5.6). To do so, we use Proposition 5.4. Note, that the function $(y, y') \mapsto \Lambda_\alpha(y, y')$ of (5.5) over which we optimize in $\lceil \Lambda_\alpha \rceil$ is semi-convex with semi-convexity constant

$$\kappa = \left(\frac{2}{1-\varepsilon^2} + \frac{1}{2} \right) \alpha + \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\kappa_V > 1$$

for $\alpha > 1$. In addition, it is bounded from above and has optimizers $(y_{\alpha,0}, y'_{\alpha,0})$. We can thus apply Proposition 5.4 with

$$\eta = \frac{1}{\alpha}, \qquad \varepsilon_1 = \frac{\varepsilon}{1-\varepsilon}\varphi, \qquad \varepsilon_2 = \frac{\varepsilon}{1+\varepsilon}\varphi. \tag{5.9}$$

Consequently, it follows that there exist $p_\alpha, p'_\alpha \in B_{1/\alpha}(0)$ such that y_α, y'_α are optimizers of

$$\left\lceil \widehat{\Lambda}_\alpha \right\rceil = \widehat{\Lambda}_\alpha(y_\alpha, y'_\alpha),$$

where

$$\widehat{\Lambda}_\alpha(y, y') := \Lambda_\alpha(y, y') - \frac{\varepsilon}{1-\varepsilon}\varphi \Xi_1^0(y) - \frac{\varepsilon}{1+\varepsilon}\varphi \Xi_2^0(y') \tag{5.10}$$

with Ξ_1^0 and Ξ_2^0 as defined above. This establishes (5.6). An additional penalization around (y_α, y'_α) gives (5.7). A secondary outcome of Proposition 5.4 is that the second derivative of $\widehat{\Lambda}_\alpha$ in the optimizing point (y_α, y'_α) exists, establishing (c). Furthermore, the optimizers satisfy

$$d(y_\alpha, y_{\alpha,0}) < \eta, \qquad d(y'_\alpha, y'_{\alpha,0}) < \eta,$$

which, together with (5.9), yields

$$\max\left\{d(y_\alpha,y_{\alpha,0}),d(y'_\alpha,y'_{\alpha,0})\right\}\le\frac{1}{\alpha},$$

establishing (b).

Proof of (d): This follows immediately from Lemma 5.3 (e).

Proof of (e): We only establish

$$u(x_\alpha)-P^\alpha[u]\circ s_{x_\alpha-y_\alpha}(x_\alpha)=\lceil u-P^\alpha[u]\circ s_{x_\alpha-y_\alpha}\rceil,$$

as the second equation follows similarly. Note that by definition of $P^\alpha[u]$, we have

$$P^\alpha[u]\circ s_{x_\alpha-y_\alpha}(x)\ge u(x)-\frac{\alpha}{2}d^2\big(x,s_{x_\alpha-y_\alpha}(x)\big).$$

On the other hand, by (d), we have

$$P^\alpha[u]\circ s_{x_\alpha-y_\alpha}(x_\alpha)=P^\alpha[u](y_\alpha)=u(x_\alpha)-\frac{\alpha}{2}d^2\left(x_\alpha,y_\alpha\right).$$

Combining the two statements yields that, for any $x\in\mathbb{R}^q$,

$$\begin{aligned}u(x_\alpha)&-P^\alpha[u]\circ s_{x_\alpha-y_\alpha}(x_\alpha)\\&=\frac{\alpha}{2}d^2\left(x_\alpha,y_\alpha\right)+P^\alpha[u]\circ s_{x_\alpha-y_\alpha}(x)-P^\alpha[u]\circ s_{x_\alpha-y_\alpha}(x)\\&\ge u(x)-P^\alpha[u]\circ s_{x_\alpha-y_\alpha}(x)+\frac{\alpha}{2}\left(d^2\left(x_\alpha,y_\alpha\right)-d^2\left(x,s_{x_\alpha-y_\alpha}(x)\right)\right)\\&=u(x)-P^\alpha[u]\circ s_{x_\alpha-y_\alpha}(x)\end{aligned}$$

as the shift map preserves distances. This establishes (e).

For the proof of the final five properties, we consider the limit $\alpha\to\infty$.

Proof of (f): Consider $\lceil\Lambda_\alpha\rceil$:

$$\begin{aligned}\lceil\Lambda_\alpha\rceil=\frac{1}{1-\varepsilon}P^\alpha[u](y_{\alpha,0})-\frac{1}{1+\varepsilon}P_\alpha[v](y'_{\alpha,0})-\frac{\alpha}{2}d^2(y_{\alpha,0},y'_{\alpha,0})\\-\frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_{\alpha,0})-\frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_{\alpha,0}).\end{aligned}$$

Note, that $\lceil\Lambda_\alpha\rceil$ is decreasing in α, since $-\frac{\alpha}{2}d^2(y_{\alpha,0},y'_{\alpha,0})$, $P^\alpha[u]$, and $-P_\alpha[v]$ are decreasing in α by Lemma 5.3 (c). Note in addition that, by evaluating Λ_α in the particular choice $(y,y')=(\hat{y},\hat{y})$ as above, we have, by Lemma 5.3 (a), that

$$\lceil\Lambda_\alpha\rceil\ge\frac{1}{1-\varepsilon}P^\alpha[u](\hat{y})-\frac{1}{1+\varepsilon}P_\alpha[v](\hat{y})\ge\frac{1}{1-\varepsilon}u(\hat{y})-\frac{1}{1+\varepsilon}v(\hat{y}),$$

which is lower bounded uniformly in α. It follows that the limit $\lim_{\alpha\to\infty}\lceil\Lambda_\alpha\rceil$ exists. For any $\alpha>1$, we find

$$\begin{aligned}\lceil\Lambda_{\alpha/2}\rceil&\ge\frac{1}{1-\varepsilon}P^{\alpha/2}[u](y_{\alpha,0})-\frac{1}{1+\varepsilon}P_{\alpha/2}[v](y'_{\alpha,0})-\frac{\alpha}{4}d^2(y_{\alpha,0},y'_{\alpha,0})\\&\quad-\frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_{\alpha,0})-\frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_{\alpha,0})\\&\ge\frac{1}{1-\varepsilon}P^\alpha[u](y_{\alpha,0})-\frac{1}{1+\varepsilon}P_\alpha[v](y'_{\alpha,0})-\frac{\alpha}{4}d^2(y_{\alpha,0},y'_{\alpha,0})\end{aligned}$$

$$
\begin{aligned}
&-\frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_{\alpha,0})-\frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_{\alpha,0})\\
&\geq \lceil\Lambda_\alpha\rceil+\frac{\alpha}{4}d^2(y_{\alpha,0},y'_{\alpha,0}),
\end{aligned}
\tag{5.11}
$$

which implies that $\lim_{\alpha\to\infty}\alpha d^2(y_{\alpha,0},y'_{\alpha,0})=0$, as $\lceil\Lambda_\alpha\rceil$ and $\lceil\Lambda_{\alpha/2}\rceil$ converge to the same limit, establishing (f).

Proof of (g): We follow the same approach as in (5.11) but now expanding $P^\alpha[u](y_\alpha)$ and $P_\alpha[v](y'_\alpha)$ to obtain an optimization problem in terms of four variables.

$$
\begin{aligned}
\lceil\Lambda_{\alpha/2}\rceil &\geq \frac{1}{1-\varepsilon}P^{\alpha/2}[u](y_\alpha)-\frac{1}{1+\varepsilon}P_{\alpha/2}[v](y'_\alpha)-\frac{\alpha}{4}d^2(y_\alpha,y'_\alpha)\\
&\quad-\frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_\alpha)-\frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_\alpha)\\
&\geq \frac{1}{1-\varepsilon}u(x_\alpha)-\frac{1}{1+\varepsilon}v(x'_\alpha)-\frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_\alpha)-\frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_\alpha)\\
&\quad-\frac{\alpha}{4}\left(\frac{1}{1-\varepsilon}d^2(x_\alpha,y_\alpha)+d^2(y_\alpha,y'_\alpha)+\frac{1}{1+\varepsilon}d^2(y'_\alpha,x'_\alpha)\right)\\
&=\left\lceil\widehat{\Lambda}_\alpha\right\rceil+\frac{\alpha}{4}\left(\frac{1}{1-\varepsilon}d^2(x_\alpha,y_\alpha)+d^2(y_\alpha,y'_\alpha)+\frac{1}{1+\varepsilon}d^2(y'_\alpha,x'_\alpha)\right)\\
&\quad+\frac{\varepsilon}{1-\varepsilon}\varphi\Xi_1^0(y_\alpha)+\frac{\varepsilon}{1+\varepsilon}\varphi\Xi_2^0(y'_\alpha)
\end{aligned}
$$

by (5.10). It follows that

$$
\begin{aligned}
&\frac{\alpha}{4}\left(\frac{1}{1-\varepsilon}d^2(x_\alpha,y_\alpha)+d^2(y_\alpha,y'_\alpha)+\frac{1}{1+\varepsilon}d^2(y'_\alpha,x'_\alpha)\right)\\
&\quad\leq\lceil\Lambda_{\alpha/2}\rceil-\left\lceil\widehat{\Lambda}_\alpha\right\rceil-\frac{\varepsilon}{1-\varepsilon}\varphi\Xi_1^0(y_\alpha)-\frac{\varepsilon}{1+\varepsilon}\varphi\Xi_2^0(y'_\alpha).
\end{aligned}
$$

By (f), we obtain

$$
\lim_{\alpha\to\infty}\lceil\Lambda_\alpha\rceil=\lim_{\alpha\to\infty}\left\lceil\widehat{\Lambda}_\alpha\right\rceil
$$

and

$$
\lim_{\alpha\to\infty}\frac{\varepsilon}{1-\varepsilon}\varphi\Xi_1^0(y_\alpha)+\frac{\varepsilon}{1+\varepsilon}\varphi\Xi_2^0(y'_\alpha)=0.
$$

Consequently, we have that

$$
\lim_{\alpha\to\infty}\alpha\left(d^2(x_\alpha,y_\alpha)+d^2(y_\alpha,y'_\alpha)+d^2(y'_\alpha,x'_\alpha)\right)=0.
$$

From this, (g) follows using Young's inequality.

Proof of (h): Parts (a), (b), (f), and (g) imply (h) by considering a bounded blow-up K_ε of $K_{\varepsilon,0}$.

Proof of (i): First, note that Corollary 5.5 and the definition of η in (5.9) yield

$$
0\leq-\frac{\varepsilon}{1-\varepsilon}\varphi\Xi_1^0(y_\alpha)-\frac{\varepsilon}{1+\varepsilon}\varphi\Xi_2^0(y'_\alpha)\leq\varphi\frac{2\varepsilon}{1-\varepsilon^2}\frac{1}{\alpha}
\tag{5.12}
$$

and

$$
\lceil\Lambda_\alpha\rceil\leq\left\lceil\widehat{\Lambda}_\alpha\right\rceil=\widehat{\Lambda}_\alpha(y_\alpha,y'_\alpha)\leq\lceil\Lambda_\alpha\rceil+\varphi\frac{2\varepsilon}{1-\varepsilon^2}\frac{1}{\alpha}.
\tag{5.13}
$$

Let $K \subseteq \mathbb{R}^q$ be compact. We then obtain

$$\begin{aligned}
\lceil u-v \rceil_K &= \sup_{x\in K} u(x)-v(x) \\
&\leq \sup_{x\in K}\left\{u(x)-v(x)-\frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\left(V(x)-\lceil V\rceil_K\right)\right\} \\
&\leq \sup_{x\in K}\left\{\frac{1}{1-\varepsilon}u(x)-\frac{1}{1+\varepsilon}v(x)-\frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)V(x)\right\} \\
&\qquad + \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\lceil V\rceil_K - \varepsilon\left\lfloor \frac{1}{1-\varepsilon}u-\frac{1}{1+\varepsilon}v\right\rfloor_K \\
&\leq \sup_{x\in \mathbb{R}^q}\left\{\frac{1}{1-\varepsilon}u(x)-\frac{1}{1+\varepsilon}v(x)-\frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)V(x)\right\} \\
&\qquad + \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\lceil V\rceil_K - \varepsilon\left\lfloor \frac{1}{1-\varepsilon}u-\frac{1}{1+\varepsilon}v\right\rfloor_K \\
&\leq \lceil \Lambda_\alpha \rceil + \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\lceil V\rceil_K - \varepsilon\left\lfloor \frac{1}{1-\varepsilon}u-\frac{1}{1+\varepsilon}v\right\rfloor_K .
\end{aligned} \tag{5.14}$$

Combining this estimate with the first inequality of (5.13), dropping non-positive terms, and then (5.12), leads to

$$\begin{aligned}
\lceil u-v \rceil_K &\leq \widehat{\Lambda}_\alpha(y_\alpha, y'_\alpha) \\
&\leq \frac{1}{1-\varepsilon}u(x_\alpha) - \frac{1}{1+\varepsilon}v(x'_\alpha) \\
&\qquad + \varepsilon\left(\varphi\frac{2}{1-\varepsilon^2}\frac{1}{\alpha} + \frac{2}{1-\varepsilon^2}(1-\varphi)\lceil V\rceil_K - \left\lfloor \frac{1}{1-\varepsilon}u-\frac{1}{1+\varepsilon}v\right\rfloor_K\right),
\end{aligned}$$

which proves (i).

Proof of (j): We start by proving that any limiting point of

$$(x_\alpha, y_\alpha, y_{\alpha,0}, y'_{\alpha,0}, y'_\alpha, x'_\alpha)$$

as $\alpha \to \infty$ is of the form (z,z,z,z,z,z). We only prove $\lim_{\alpha\to\infty} d(x_\alpha, y_\alpha) = 0$, as the other limits follow analogously.

By (h), we find that, along subsequences, $(x_\alpha, y_\alpha) \to (x_0, y_0)$. Assume by contradiction that $x_0 \neq y_0$. Then, since αd^2 is increasing, we get that, for all $\alpha_0 > 1$,

$$\liminf_{\alpha\to\infty} \alpha d^2(x_\alpha, y_\alpha) \geq \alpha_0 d^2(x_0, y_0).$$

We can conclude that $\alpha d^2(x_\alpha, y_\alpha) \to \infty$, contradicting (g).

We proceed to prove that any limiting point z lies in $\widehat{K}$. Similar to (5.8), but now also using (5.12) and the first inequality of (5.13), we find

$$\begin{aligned}
&\frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_\alpha) + \frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_\alpha) \\
&\quad\leq \frac{1}{1-\varepsilon}\lceil u\rceil - \frac{1}{1+\varepsilon}\lfloor v\rfloor - \frac{\varepsilon}{1-\varepsilon}\varphi\Xi^0_1(y_\alpha) - \frac{\varepsilon}{1+\varepsilon}\varphi\Xi^0_2(y'_\alpha) - \left\lceil \widehat{\Lambda}_\alpha\right\rceil \\
&\quad\leq \frac{1}{1-\varepsilon}\lceil u\rceil - \frac{1}{1+\varepsilon}\lfloor v\rfloor + \varphi\frac{2\varepsilon}{1-\varepsilon^2}\frac{1}{\alpha} - \lceil \Lambda_\alpha\rceil .
\end{aligned}$$

Combining this with the upper bound on $-\lceil\Lambda_\alpha\rceil$ obtained from (5.14) leads to

$$\begin{aligned}\frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_\alpha) &+ \frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_\alpha)\\ &\leq \frac{1}{1-\varepsilon}\lceil u\rceil + \frac{1}{1+\varepsilon}\lfloor v\rfloor + \varphi\frac{2\varepsilon}{1-\varepsilon^2}\frac{1}{\alpha}\\ &\quad - \lceil u-v\rceil_K + \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\lceil V\rceil_K - \varepsilon\left\lfloor\frac{1}{1-\varepsilon}u - \frac{1}{1+\varepsilon}v\right\rfloor_K.\end{aligned}$$

This, in turn, yields

$$\begin{aligned}\frac{\varepsilon}{1-\varepsilon}(1-\varphi)\left(V(y_\alpha)-\lceil V\rceil_K\right) &+ \frac{\varepsilon}{1+\varepsilon}(1-\varphi)\left(V(y'_\alpha)-\lceil V\rceil_K\right)\\ &\leq \frac{1}{1-\varepsilon}\lceil u\rceil + \frac{1}{1+\varepsilon}\lfloor v\rfloor + \varphi\frac{2\varepsilon}{1-\varepsilon^2}\frac{1}{\alpha}\\ &\quad - \lceil u-v\rceil_K - \varepsilon\left\lfloor\frac{1}{1-\varepsilon}u - \frac{1}{1+\varepsilon}v\right\rfloor_K\\ &\leq 2\left(\|u\|+\|v\|\right) + \varphi\frac{\varepsilon}{1-\varepsilon^2}\frac{1}{\alpha}.\end{aligned}$$

By (g) and (h), the sequences $x_\alpha, y_\alpha, y'_\alpha, x'_\alpha$ have limit points $z\in K_\varepsilon$ as $\alpha\to\infty$. In combination with (b), we conclude that, for any such limiting point z,

$$\frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\left(V(z)-\lceil V\rceil_K\right)\leq 2\left(\|u\|+\|v\|\right),$$

establishing (j). □

We proceed by proving the auxiliary results below.

Proposition 5.3. *Let $u\colon\mathbb{R}^q\to\mathbb{R}$ be bounded and upper semi-continuous and $v\colon\mathbb{R}^q\to\mathbb{R}$ be bounded and lower semi-continuous. For $\alpha>1$, let $P^\alpha[u]$ and $P_\alpha[v]$ be the sup- and inf-convolution of u and v, respectively. Then,*

(a) *we have $\|P^\alpha[u]\|\leq\|u\|$ and $\|P_\alpha[v]\|\leq\|v\|$.*
(b) *for any $x,y\in\mathbb{R}^q$ such that*

$$P^\alpha[u](y)=u(x)-\frac{\alpha}{2}d^2(x,y),$$

we have $\frac{\alpha}{2}d^2(x,y)\leq u(x)-u(y)$. Similarly, for any $x,y\in\mathbb{R}^q$ with

$$P_\alpha[v](y)=v(x)+\frac{\alpha}{2}d^2(x,y),$$

we have $\frac{\alpha}{2}d^2(x,y)\leq v(y)-v(x)$.
(c) *$P^\alpha[u]$ and $-P_\alpha[v]$ are decreasing in α.*
(d) *$P^\alpha[u]$ and $-P_\alpha[v]$ are semi-convex with semi-convexity constant α. As a consequence, both are locally Lipschitz continuous.*
(e) *if $P^\alpha[u]$ is differentiable at y_0, then there exists a unique optimizer x_0 in (5.3) such that*

$$P^\alpha[u](y_0)=u(x_0)-\frac{\alpha}{2}d^2(x_0,y_0)$$

and $DP^\alpha[u](y_0) = \alpha(x_0 - y_0)$. *Similarly, if* $P_\alpha[v]$ *is differentiable at* y_0, *then there is a unique optimizer* x_0 *in* (5.4) *such that*

$$P_\alpha[v](y_0) = v(x_0) + \frac{\alpha}{2} d^2(x_0, y_0)$$

and $DP_\alpha[v](y_0) = -\alpha(x_0 - y_0)$.

Proof. For the proof of (a), note that, for any $x, y \in \mathbb{R}^q$, we have

$$u(x) - \frac{\alpha}{2} d^2(x, y) \le u(x).$$

This implies that

$$\lceil P^\alpha[u] \rceil = \left\lceil u - \frac{\alpha}{2} d^2 \right\rceil \le \lceil u \rceil . \tag{5.15}$$

On the other hand, we have

$$u(y) \le \left\lceil u - \frac{\alpha}{2} d^2(\cdot, y) \right\rceil = P^\alpha[u](y).$$

It follows that

$$\lfloor u \rfloor \le \lfloor P^\alpha[u] \rfloor . \tag{5.16}$$

Now, (a) follows by (5.15) and (5.16). Part (b) is equivalent to

$$P_\alpha[u] \le u \le P^\alpha[u] \quad \text{on } \mathbb{R}^q,$$

which is immediately clear from the definitions of sup- and inf-convolutions. Part (c) follows similarly from the definitions. For the proof of (d), let $y_0 \in \mathbb{R}^q$. Then, since d is the Euclidean distance, we find

$$P^\alpha[u](y) + \frac{\alpha}{2} d^2(y, y_0) = \left\lceil u + \alpha \langle y - y_0, \cdot - y_0 \rangle - \frac{\alpha}{2} d^2(\cdot, y_0) \right\rceil ,$$

where the right-hand side is convex as it is a supremum over affine functions. By [13, Proposition 2.1.5 and Theorem 2.1.7], the claim follows. Lastly, (e) follows from [13, Theorem 3.4.4] by noting that the sets over which can be optimized are compact due to the boundedness of u and v. □

Proposition 5.4. *Fix* $\eta > 0$. *Let* $\phi\colon \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}$ *be bounded from above and semi-convex with convexity constant* $\kappa \ge 1$. *Suppose that* (x_0, x_0') *is an optimizer of*

$$\phi(x_0, x_0') = \lceil \phi \rceil .$$

Let $R > 0$, $\{\zeta_{z,p}\}_{z \in \mathbb{R}^q, p \in \mathbb{R}^q} \subseteq C(\mathbb{R}^q)$ *and* $\{\xi_z\}_{z \in \mathbb{R}^q} \subseteq C^1(\mathbb{R}^q)$ *and semi-concavity constant* κ_ξ *be as in Definition 3.10.*

Fix $\varepsilon_1, \varepsilon_2 > 0$ *such that* $1 - (\varepsilon_1 + \varepsilon_2)\kappa_\xi > 0$. *Furthermore, define for* $p = (p_1, p_2) \in \mathbb{R}^q \times \mathbb{R}^q$ *the perturbed functions*

$$\phi_p(x, x') := \phi(x, x') - \varepsilon_1 \Big(\xi_{x_0}(x) + \zeta_{x_0, p_1}(x) \Big) - \varepsilon_2 \Big(\xi_{x_0'}(x') + \zeta_{x_0', p_2}(x') \Big). \tag{5.17}$$

Then, there exist $p_1, p_2 \in B_\eta(0)$, *and a pair* $(x_1, x_1') \in B_\eta(x_0) \times B_\eta(x_0')$ *globally maximizing* ϕ_p *such that the second derivative of* ϕ_p *in* (x_1, x_1') *exists.*

Corollary 5.5. *For $\eta > 0$, p and (x_1, x_1') as in Proposition 5.4, we have*

$$0 \le -\varepsilon_1\Big(\xi_{x_0}(x_1) + \zeta_{x_0,p_1}(x_1)\Big) - \varepsilon_2\Big(\xi_{x_0'}(x_1') + \zeta_{x_0',p_2}(x_1')\Big) \le \varepsilon_1\eta + \varepsilon_2\eta, \tag{5.18}$$

and

$$\lceil\phi\rceil \le \phi_{p,\varepsilon}(x_1, y_1) \le \lceil\phi\rceil + \varepsilon_1\eta + \varepsilon_2\eta. \tag{5.19}$$

The proof of the Proposition 5.4 is partly based on results from set-valued analysis. To facilitate the proof, we first introduce the necessary auxiliary definitions and results.

Definition 5.6. A set-valued function $\Gamma : A \rightrightarrows B$ is called *upper hemi-continuous at $a \in A$*, if, for all open neighbourhoods $V \subseteq B$ of $\Gamma(a)$ (meaning that $\Gamma(a) \subseteq V$), there exists a neighborhood U of a such that, for all $x \in U$, we have $\Gamma(x) \subseteq V$.

If A, B are metric, this can be equivalently formulated in terms of sequences: A set-valued map $\Gamma : A \rightrightarrows B$, which takes closed values, is upper hemi-continuous at a, if, for any sequence $a_n \to a$ and $b_n \in \Gamma(a_n)$ satisfying $b_n \to b$, we have $b \in \Gamma(a)$.

We say that Γ is upper hemi-continuous, if it is upper hemi-continuous at all points.

Lemma 5.7. *Let K be a compact, metric space and let Ξ be a metric space.*

For any $\xi \in \Xi$, let $\phi_\xi \in C(K)$ and suppose that the map $\xi \mapsto \phi_\xi$ is continuous from Ξ to $C(K)$, endowed with the supremum norm on K. Then the set-valued map $\mathrm{Opt}\colon \Xi \rightrightarrows K$ defined by

$$\mathrm{Opt}(\xi) := \{x \in K \,|\, \phi_\xi \textit{ has a maximum at } x\}$$

is upper hemi-continuous.

Proof. The result follows immediately from Berge's Maximum Theorem [1, Theorem 17.31] with $\xi \mapsto \mathrm{image}_{\phi_\xi}(K)$ being the relevant set-valued map. □

Remark 5.8. In the proof below, we will make use of the notion of a lim sup of sets. For a sequence of sets $(A_n)_{n\in\mathbb{N}}$ denote

$$\limsup_{n\to\infty} A_n = \bigcap_{n\in\mathbb{N}} \bigcup_{m\ge n} A_m$$

to be interpreted as $x \in \limsup_{n\to\infty} A_n$ if and only if there are infinitely many $n \in \mathbb{N}$ such that $x \in A_n$.

The following proof is a variant of the proof of [18, Lemma A.3] and [13, Theorem 2.3.3].

Proof of Proposition 5.4. For notational convenience, we will write $w = (x, x')$ and $w_0 = (x_0, x_0')$. Let $R > 0$ and $\{\zeta_{z,p}\}_{z\in\mathbb{R}^q, p\in\mathbb{R}^q} \subseteq C(\mathbb{R}^q)$ and $\{\xi_z\}_{z\in\mathbb{R}^q} \subseteq C^1(\mathbb{R}^q)$ be two collections of point penalizations as in Definition 3.10. Without loss of generality, we can assume that $R \ge \eta$.

We start out by making z_0 the unique optimizer by replacing ϕ with

$$\widehat{\phi}(w) = \phi(w) - \varepsilon_1\xi_{x_0}(x) - \varepsilon_2\xi_{x_0'}(x').$$

Note that as $1 - (\varepsilon_1 + \varepsilon_2)\kappa_\xi > 0$ the map $\widehat{\phi}$ is semi-convex and bounded from above with a unique optimizer w_0.

Our next step is to locally, linearly perturb $\widehat{\phi}$ to obtain ϕ_p as in equation (5.17). This procedure produces a new optimizer close to w_0, in which the second derivative of the perturbed function ϕ_p exists.

To further facilitate the analysis of optimizers, we smoothen out ϕ. To that end, let $C_\delta \colon C_b(\mathbb{R}^q) \to C_b^2(\mathbb{R}^q)$ be a mollifier with $\sup_{\delta>0} \|C_\delta f\| < \infty$ and $C_\delta f \to f$ uniformly on compacts as $\delta \downarrow 0$. Define

$$\phi_{p,\delta}(w) := (C_\delta \phi)(w) - \varepsilon_1 \Big(\xi_{x_0}(x) + \zeta_{x_0,p_1}(x) \Big) - \varepsilon_2 \Big(\xi_{x_0'}(x') + \zeta_{x_0',p_2}(x') \Big),$$

where we will read $C_0 = \mathbb{1}$ such that $\phi_{p,0} = \phi_p$ and $\phi_{0,0} = \widehat{\phi}$.

We next study the optimizers for the map $(p, \delta) \mapsto \phi_{p,\delta}$ on $\Xi = (B_1(0) \times B_1(0)) \times [0, 1]$ using Berge's Maximum Theorem with $K = \overline{B_R(w_0)}$. Set

$$\mathrm{Opt}(p, \delta) := \left\{ w \in \overline{B_R(w_0)} \,\middle|\, \phi_{p,\delta} \text{ has a local maximum at } w \in \overline{B_R(w_0)} \right\}.$$

First, note that the local nature of the problem can be removed due to the fact that the perturbations all vanish in w_0, whereas they add up to something negative outside the ball $\overline{B_R(w_0)}$ by Definition 3.10 (d), implying that

$$\mathrm{Opt}(p, \delta) = \left\{ w \in \overline{B_R(w_0)} \,\middle|\, \phi_{p,\delta} \text{ has a global maximum at } w \right\}. \tag{5.20}$$

Applying Lemma 5.7 to $(p, \delta) \mapsto \phi_{p,\delta}$ on $\Xi = (B_\eta(0) \times B_\eta(0)) \times [0, 1]$ with $K = \overline{B_R(w_0)}$, we find that the set-valued map $\mathrm{Opt} \colon \Xi \rightrightarrows K \subseteq \mathbb{R}^q \times \mathbb{R}^q$, as defined above, is upper hemi-continuous in the variables (p, δ). We can thus find a closed set U with 0 in its interior satisfying

$$U \subseteq B_\eta(0) \times B_\eta(0) \tag{5.21}$$

and $\delta_0 > 0$ such that, if $p = (p_1, p_2) \in U$ and $\delta < \delta_0$, then

$$\mathrm{Opt}(p, \delta) \subseteq \mathrm{Opt}(0, 0) \oplus B_\eta(0) = B_\eta(w_0), \tag{5.22}$$

as the unique optimizer of $\widehat{\phi}$ is w_0.

We next aim to show that the set of such optimizers has positive Lebesgue measure m. Recall that $\kappa := 1 - (\varepsilon_1 + \varepsilon_2)\kappa_\xi > 0$ is the semi-convexity constant of $\widehat{\phi}$. In particular, we will proceed to show the following steps.

Step 1: For any $\delta \in (0, \delta_0)$, we have $m(\mathrm{Opt}(U, \delta)) \geq |\kappa|^{-2d} m(U) > 0$.
Step 2: We take the limit $\delta \downarrow 0$ to obtain $m(\mathrm{Opt}(U, 0)) \geq |\kappa|^{-2d} m(U) > 0$.

Step 1. By definition, all perturbations are at least once continuously differentiable on $\overline{B_R(w_0)}$. It follows that for $p \in U$, $\delta \in (0, \delta_0)$ and $w \in \mathrm{Opt}(p, \delta)$ we have that $D(C_\delta\phi)(w) = p$. This, in turn, implies that, for fixed $\delta \in (0, \delta_0)$,

$$U \subseteq (\mathrm{Opt}(\cdot, \delta))^{-1} (\mathrm{Opt}(U, \delta)) \subseteq D(C_\delta\phi)(\mathrm{Opt}(U, \delta)). \tag{5.23}$$

We next argue towards a lower bound on the measure of $\mathrm{Opt}(U, \delta)$ for $\delta \in (0, \delta_0)$. We exclude $\delta = 0$ here, due to the possible non-smoothness of ϕ. As the convolution operator is taking averages, the semi-convexity of ϕ carries over to $C_\delta\phi$, which yields

$$-\kappa I_{2d} \leq D^2(C_\delta\phi)(w) \tag{5.24}$$

for all $w \in \mathbb{R}^{2q}$. On the other hand, if $w \in \mathrm{Opt}(U, \delta)$, we know that there is some $p \in U$ such that w maximizes $\phi_{p,\delta}$, implying that $D^2(C_\delta\phi)(w) \leq 0$. Applying (5.23), the chain rule, and (5.24), we thus obtain that, for any $\delta \in (0, \delta_0)$,

$$\begin{aligned} m(U) &\leq m(D(C_\delta\phi)(\mathrm{Opt}(U,\delta))) \\ &= \int\limits_{\mathrm{Opt}(U,\delta)} \left|\det D^2(C_\delta\phi)(w)\right| \mathrm{d}w \leq m(\mathrm{Opt}(U,\delta))|\kappa|^{2q} \end{aligned}$$

leading to the lower bound

$$0 < |\kappa|^{-2q} m(U) \leq m(\mathrm{Opt}(U,\delta)), \tag{5.25}$$

as U has non-empty interior, establishing the claim of Step 1.
Step 2. Next, we transfer our bound to $m(\mathrm{Opt}(U, 0))$. We first establish that

$$\limsup_{\delta \downarrow 0} \mathrm{Opt}(U,\delta) \subseteq \mathrm{Opt}(U,0), \tag{5.26}$$

see Remark 5.8 for the definition of the lim sup of sets. To that end, we pick an element $w \in \limsup_{\delta\downarrow 0} \mathrm{Opt}(U,\delta)$. By definition we can find a sequence $\delta_n \downarrow 0$ such that $w \in \mathrm{Opt}(U, \delta_n)$ for all $n \in \mathbb{N}$. Then, there are $p_n \in U$ such that w is an optimizer for $(\phi_{p_n,\delta_n})_{n\in\mathbb{N}}$. By the closedness of U and (5.21), U is compact, and we can therefore extract a subsequence from $(p_n)_{n\in\mathbb{N}}$ that converges to some $p_0 \in U$. By upper semi-continuity of the map $(p, \delta) \mapsto \phi_{p,\delta}$, see Lemma 5.7, we find that w maximizes ϕ_{p_0}, or in other words, $w \in \mathrm{Opt}(U, 0)$.

Thus, by (5.26), it suffices to lower bound the measure of $\limsup_{\delta\downarrow 0} \mathrm{Opt}(U,\delta)$. As a first step, note that (5.25) leads to

$$m\left(\bigcup_{\delta' \leq \delta} \mathrm{Opt}(U,\delta')\right) \geq |\kappa|^{-2q} m(U)$$

for any $\delta \in (0, \delta_0)$. Consequently, as

$$\limsup_{\delta\downarrow 0} \mathrm{Opt}(U,\delta) = \bigcap_{\delta \in (0,\delta_0)} \bigcup_{\delta' \leq \delta} \mathrm{Opt}(U,\delta'),$$

by continuity from above of the Lebesgue measure m, we find that

$$m\left(\limsup_{\delta\downarrow 0} \mathrm{Opt}(U,\delta)\right) = \lim_{\delta\downarrow 0} m\left(\bigcup_{\delta' \leq \delta} \mathrm{Opt}(U,\delta')\right) \geq |\kappa|^{-2q} m(U).$$

By (5.26), we conclude that

$$m(\mathrm{Opt}(U,0)) \geq |\kappa|^{-2q} m(U) > 0,$$

establishing the claim of Step 2.

We proceed by verifying that we can now find $p \in U$ with an optimizer z_1 in $B_\eta(z_0)$ such that the second derivative of ϕ_p in (x_1, x_1') exists.

First of all, recall that, by (5.22), we have

$$\mathrm{Opt}(U,0) \subseteq B_\eta(z_0).$$

Furthermore, by Alexandrov's theorem [13, Theorem 2.3.1], the set of points in $B_\eta(z_0)$, where the second derivative of ϕ_p exists, has full measure. As the measure of $\mathrm{Opt}(U,0)$ is positive, it follows that there exist $z_1 \in B_\eta(z_0)$ and $p \in U$ such that the second derivative of ϕ_p in z_1 exists and ϕ_p has a local maximum at z_1 in $\overline{B_R(z_0)}$. Finally, recall from (5.20) that the local optimizer is in fact a global optimizer. This establishes the claim. □

Proof of Corollary 5.5. We stay in the context of Proposition 5.4 and proceed with the proof of (5.19). By construction, we have

$$\lceil \phi \rceil = \phi(x_0, x_0') = \phi_{p,\varepsilon}(x_0, x_0') \leq \lceil \phi_{p,\varepsilon} \rceil = \phi_{p,\varepsilon}(x_1, x_1'). \tag{5.27}$$

This implies the lower bound of (5.18). Note that by the properties of point penalizations $\xi_{x_0}, \xi_{x_0'}, \zeta_{x_0,p_1}$, and ζ_{x_0',p_2} we have

$$\begin{aligned} -\varepsilon_1\Big(\xi_{x_0}(x_1) + \zeta_{x_0,p_1}(x_1)\Big) - \varepsilon_2\Big(\xi_{x_0'}(x_1') + \zeta_{x_0',p_2}(x_1')\Big) &\leq \varepsilon_1|p_1|d(x_0,x_1) + \varepsilon_2|p_2|d(x_0',x_1') \\ &\leq \varepsilon_1\eta + \varepsilon_2\eta, \end{aligned}$$

leading to the upper bound of (5.18). Consequently,

$$\begin{aligned} \phi_{p,\varepsilon}(x_1, x_1') &\leq \phi(x_1, x_1') + \varepsilon_1\eta + \varepsilon_2\eta \\ &\leq \lceil \phi \rceil + \varepsilon_1\eta + \varepsilon_2\eta. \end{aligned} \tag{5.28}$$

Combining (5.27) and (5.28), finally yields (5.19). □

5.2. Test function construction

The next proposition builds upon Proposition 5.2 to build a suitable collection of test functions for the use in the proof of the comparison principle. The sup- and inf-convolution $P^\alpha[u]$ and $P_\alpha[v]$ are not guaranteed to be smooth. However, their second derivatives exist in the relevant optimizing points.

Using the difference between Ξ_1^0 and Ξ_2^0 on one hand and Ξ_1 and Ξ_2 on the other, we are able to squeeze in a globally C^∞ function on the basis of Lemma 5.11 below, that can be used to replace $P^\alpha[u]$ and $P_\alpha[v]$. As an effect, we will approximate

$$\begin{aligned} \widehat{f}_\dagger &\approx P^\alpha[u], & f_\dagger &\approx P^\alpha[u] \circ s_{x_\alpha - y_\alpha}, \\ \widehat{f}_\ddagger &\approx P_\alpha[v], & f_\ddagger &\approx P_\alpha[v] \circ s_{x_\alpha' - y_\alpha'}, \end{aligned}$$

which will be made rigorously in next proposition for fixed ε and α. We start by first constructing $\widehat{f}_1 \in C_c^\infty(\mathbb{R}^q)$, which, by re-arrangement, satisfies

$$\widehat{f}_1(y) \approx \frac{1}{1-\varepsilon} P^\alpha[u](y) - \frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y) - \frac{\varepsilon}{1-\varepsilon}\varphi\Xi_1(y)$$

and is constant outside of a compact set. As V has compact sublevel sets and other terms on the right-hand side are bounded from above, it suffices to first perform a smooth approximation and cut off the result. For the cut-off procedure, we use functions Ω_M^+ and Ω_M^-.

Definition 5.9 *(Cut-off functions).* Let $M > 0$. We call a smooth increasing function $\Omega_M^+ : \mathbb{R} \to \mathbb{R}$ a *upper cut-off function at* M, if

$$\Omega_M^+(r) = \begin{cases} r & \text{if } r \leq M, \\ M+1 & \text{if } r \geq M+2. \end{cases}$$

We call Ω_M^- a *lower cut-off function* at M if $\Omega_M^-(r) = -\Omega_{-M}^+(-r)$.

Theorem 5.10 *(Test function construction). Consider the setting of Proposition 5.2. Fix $\varepsilon \in (0,1)$, $\varphi \in (0,1]$, and $\alpha > 1$. Then, there are functions $f_1, f_2, \widehat{f}_1, \widehat{f}_2 \in C_c^\infty(\mathbb{R}^q)$ such that*

$$f_1 = \widehat{f}_1 \circ s_{x_\alpha - y_\alpha}, \qquad f_2 = \widehat{f}_2 \circ s_{x'_\alpha - y'_\alpha}$$

and

$$\widehat{f}_\dagger := (1-\varepsilon)\widehat{f}_1 + \varepsilon(1-\varphi)V + \varepsilon\varphi\Xi_1, \qquad f_\dagger = \widehat{f}_\dagger \circ s_{x_\alpha - y_\alpha},$$
$$\widehat{f}_\ddagger := (1+\varepsilon)\widehat{f}_2 - \varepsilon(1-\varphi)V - \varepsilon\varphi\Xi_2, \qquad f_\ddagger = \widehat{f}_\ddagger \circ s_{x'_\alpha - y'_\alpha},$$

satisfying the following properties:

For $\widehat{f}_1, \widehat{f}_2$ and f_1, f_2, we have

(a) *The pair (y_α, y'_α) is the unique optimizing pair of*

$$\widehat{f}_1(y_\alpha) - \widehat{f}_2(y'_\alpha) - \frac{\alpha}{2}d^2(y_\alpha, y'_\alpha) = \left\lceil \widehat{f}_1 - \widehat{f}_2 - \frac{\alpha}{2}d^2 \right\rceil$$

and the pair (x_α, x'_α) is the unique optimizing pair of

$$f_1(x_\alpha) - f_2(x'_\alpha) - \frac{\alpha}{2}d^2_{x_\alpha - y_\alpha,\, x'_\alpha - y'_\alpha}(x_\alpha, x'_\alpha) = \left\lceil f_1 - f_2 - \frac{\alpha}{2}d^2_{x_\alpha - y_\alpha,\, x'_\alpha - y'_\alpha} \right\rceil.$$

For $\widehat{f}_\dagger, \widehat{f}_\ddagger$, and $f_\dagger, f_\ddagger$ we have

(b) *We have*

$$P^\alpha[u](y) \leq \widehat{f}_\dagger(y),$$
$$P_\alpha[v](y') \geq \widehat{f}_\ddagger(y')$$

with equality in y_α and y'_α, respectively.

(c) *We have that x_α, x'_α are the unique points such that*

$$u(x_\alpha) - f_\dagger(x_\alpha) = \lceil u - f_\dagger \rceil,$$
$$v(x'_\alpha) - f_\ddagger(x'_\alpha) = \lceil v - f_\ddagger \rceil.$$

(d) *We have*

$$D\widehat{f}_\dagger(y_\alpha) = Df_\dagger(x_\alpha) = \alpha(y_\alpha - x_\alpha),$$
$$D^2\widehat{f}_\dagger(y_\alpha) = D^2 f_\dagger(x_\alpha),$$
$$D\widehat{f}_\ddagger(y'_\alpha) = Df_\ddagger(x'_\alpha) = \alpha(x'_\alpha - y'_\alpha),$$
$$D^2\widehat{f}_\ddagger(y'_\alpha) = D^2 f_\ddagger(x'_\alpha).$$

Proof of Proposition 5.10. In this proof, we work in the context of Proposition 5.2 and will, correspondingly, follow its notation. We show the construction procedure for the test function f_1 used in the subsolution case only, as f_2 is constructed analogously. Let

$$\Pi_1^0(y) := \frac{1}{1-\varepsilon} P^\alpha[u](y) - \frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y) - \frac{\varepsilon}{1-\varepsilon}\varphi \Xi_1^0(y),$$
$$\Pi_1(y) := \frac{1}{1-\varepsilon} P^\alpha[u](y) - \frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y) - \frac{\varepsilon}{1-\varepsilon}\varphi \Xi_1(y).$$

Note that we have $\Pi_1(y_\alpha) = \Pi_1^0(y_\alpha)$ and $\Pi_1(y) < \Pi_1^0(y)$ for all $y \in \mathbb{R}^q \setminus \{y_\alpha\}$. By Lemma 5.11, we find a function $\mathfrak{f}_1 \in C^\infty(\mathbb{R}^q)$ such that

$$\Pi_1(y) < \mathfrak{f}_1(y) < \Pi_1^0(y), \quad \text{if } y \neq y_\alpha.$$

The function $\mathfrak{f}_2$ is constructed analogously. By construction of $\mathfrak{f}_1$, $\mathfrak{f}_2$ and (5.6), (y_α, y'_α) is the unique optimizer of $\lceil \mathfrak{f}_1 - \mathfrak{f}_2 - \frac{\alpha}{2} d^2 \rceil$.

As our test functions need to be constant outside a compact set, we need to cut them off in an appropriate manner. However, we need to preserve their properties in the optimizer (y_α, y'_α). Taking these conditions into account, ensures that the cut-off procedure does not create new optimizers.

The above considerations lead to the cut-off procedure $\widehat{f_1} := \Omega^-_{M_1} \circ \mathfrak{f}_1$ and $\widehat{f_2} := \Omega^+_{M_2} \circ \mathfrak{f}_2$, with $\Omega^-_{M_1}$, $\Omega^+_{M_2}$ as in Definition 5.9 and the following choice of M_1 and M_2. First pick $m_1, m_2 \in \mathbb{R}$ such that the level sets

$$\{y \in \mathbb{R}^q \,|\, \mathfrak{f}_1(y) \geq m_1\}, \qquad \{y' \in \mathbb{R}^q \,|\, \mathfrak{f}_2(y') \leq m_1\}$$

are compact, then set

$$M_1 := \min\left\{m_1,\, \mathfrak{f}_1(y_\alpha) - (\mathfrak{f}_2(y'_\alpha) - \lfloor \mathfrak{f}_2 \rfloor) - \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha)\right\},$$
$$M_2 := \max\left\{m_2,\, \mathfrak{f}_2(y'_\alpha) + (\lceil \mathfrak{f}_1 \rceil - \mathfrak{f}_1(y_\alpha)) + \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha)\right\}.$$

Using M_1 and M_2 as defined above, we find that (y_α, y'_α) is the unique optimizer of $\left\lceil \widehat{f_1} - \widehat{f_2} - \frac{\alpha}{2} d^2 \right\rceil$. To see this, denote

$$A_1 := \{y \in \mathbb{R}^q \,|\, \mathfrak{f}_1(y) \geq M_1\} \quad \text{and} \quad A_2 := \{y' \in \mathbb{R}^q \,|\, \mathfrak{f}_2(y') \leq M_2\}.$$

Thus, for $i \in \{1, 2\}$, we find $\widehat{f_i} = \mathfrak{f}_i$ on A_i, whereas

$$\widehat{f_1}(y) < \mathfrak{f}_1(y_\alpha) - (\mathfrak{f}_2(y'_\alpha) - \lfloor \mathfrak{f}_2 \rfloor) - \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha),$$
$$\widehat{f_2}(y') > \mathfrak{f}_2(y'_\alpha) + (\lceil \mathfrak{f}_1 \rceil - \mathfrak{f}_1(y_\alpha)) + \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha)$$

if $y \notin A_1$ or $y' \notin A_2$, respectively.

As $\widehat{f_1} = \mathfrak{f}_1$ on A_1 and $\widehat{f_2} = \mathfrak{f}_2$ on A_2, it suffices to show that

$$\widehat{f_1}(y) - \widehat{f_2}(y') - \frac{\alpha}{2} d^2(y, y') < \mathfrak{f}_1(y_\alpha) - \mathfrak{f}_2(y'_\alpha) - \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha)$$

if $y \in A_1^c$ or $y \in A_2^c$. For the proof of this bound, we consider the following three separate cases.

Case $y \in A_1^c$ and $y' \in A_2$: We have

$$\begin{aligned}\widehat{f}_1(y) - \widehat{f}_2(y') - \frac{\alpha}{2}d^2(y,y') &\leq \widehat{f}_1(y) - \widehat{f}_2(y') \\ &< \mathfrak{f}_1(y_\alpha) - (\mathfrak{f}_2(y'_\alpha) - \lfloor \mathfrak{f}_2 \rfloor) - \frac{\alpha}{2}d^2(y_\alpha, y'_\alpha) - \mathfrak{f}_2(y') \\ &= \mathfrak{f}_1(y_\alpha) - \mathfrak{f}_2(y'_\alpha) - \frac{\alpha}{2}d^2(y_\alpha, y'_\alpha) - (\mathfrak{f}_2(y') - \lfloor \mathfrak{f}_2 \rfloor) \\ &\leq \mathfrak{f}_1(y_\alpha) - \mathfrak{f}_2(y'_\alpha) - \frac{\alpha}{2}d^2(y_\alpha, y'_\alpha).\end{aligned}$$

Case $y \in A_1$ and $y' \in A_2^c$: Follows analogously to the case $y \in A_1^c$ and $y' \in A_2$.
Case $y \in A_1^c$ and $y' \in A_2^c$: We have

$$\begin{aligned}&\widehat{f}_1(y) - \widehat{f}_2(y') - \frac{\alpha}{2}d^2(y,y') \\ &\quad\leq \widehat{f}_1(y) - \widehat{f}_2(y') \\ &\quad< \mathfrak{f}_1(y_\alpha) - (\mathfrak{f}_2(y'_\alpha) - \lfloor \mathfrak{f}_2 \rfloor) - \frac{\alpha}{2}d^2(y_\alpha, y'_\alpha) \\ &\qquad\qquad - \Big(\mathfrak{f}_2(y'_\alpha) + (\lceil \mathfrak{f}_1 \rceil - \mathfrak{f}_1(y_\alpha)) + \frac{\alpha}{2}d^2(y_\alpha, y'_\alpha)\Big) \\ &\quad\leq \mathfrak{f}_1(y_\alpha) - \mathfrak{f}_2(y'_\alpha) - 2\frac{\alpha}{2}d^2(y_\alpha, y'_\alpha) - (\mathfrak{f}_2(y'_\alpha) - \lfloor \mathfrak{f}_2 \rfloor) - (\lceil \mathfrak{f}_1 \rceil - \mathfrak{f}_1(y_\alpha)) \\ &\quad\leq \mathfrak{f}_1(y_\alpha) - \mathfrak{f}_2(y'_\alpha) - \frac{\alpha}{2}d^2(y_\alpha, y'_\alpha).\end{aligned}$$

We conclude that the pair (y_α, y'_α) is also the unique optimizer of $\Big\lceil \widehat{f}_1 - \widehat{f}_2 - \frac{\alpha}{2}d^2 \Big\rceil$. Applying the shift maps $s_{x_\alpha - y_\alpha}$ and $s_{x'_\alpha - y'_\alpha}$, respectively, we find that (x_α, x'_α) uniquely optimize $\Big\lceil \mathfrak{f}_1 \circ s_{x_\alpha - y_\alpha} - \mathfrak{f}_2 \circ s_{x'_\alpha - y'_\alpha} - \frac{\alpha}{2}d^2_{x_\alpha - y_\alpha, x'_\alpha - y'_\alpha}\Big\rceil$. Additionally, as $M_1 \geq m_1$ and $M_2 \leq m_2$, we have $\widehat{f}_1, \widehat{f}_2 \in C_c^\infty(\mathbb{R}^q)$, establishing (a).

We next prove (b). As $r \leq \Omega^-_{M_1}(r)$,

$$\frac{1}{1-\varepsilon}P^\alpha[u](y) - \frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y) - \frac{\varepsilon}{1-\varepsilon}\varphi\Xi_1(y) = \Pi_1(y) \leq \Omega^-_{M_1} \circ \Pi_1(y) \leq \widehat{f}_1(y),$$

which, after rearrangement of terms, implies (b).

We proceed with the proof of (c). By (b) and Proposition 5.2 (e),

$$\begin{aligned}f_\dagger(x) - f_\dagger(x_\alpha) &= \widehat{f}_\dagger \circ s_{x_\alpha - y_\alpha}(x) - \widehat{f}_\dagger \circ s_{x_\alpha - y_\alpha}(x_\alpha) \\ &\geq (P^\alpha[u] \circ s_{x_\alpha - y_\alpha})(x) - (P^\alpha[u] \circ s_{x_\alpha - y_\alpha})(x_\alpha) \\ &\geq \Big(u(x) - \frac{\alpha}{2}d^2(x, s_{x_\alpha - y_\alpha}(x))\Big) - \Big(u(x_\alpha) - \frac{\alpha}{2}d^2(x_\alpha, s_{x_\alpha - y_\alpha}(x_\alpha))\Big) \\ &= u(x) - u(x_\alpha)\end{aligned}$$

with equality uniquely realized at x_α, establishing (c).

We conclude with the proof of (d). First of all, note that the equality of first and second order derivatives for $f_\dagger$ and $\widehat{f}_\dagger$ as well as for $f_\ddagger$ and $\widehat{f}_\ddagger$ follows by the chain rule.

The expressions for $D\widehat{f}_\dagger(y_\alpha)$ and $D\widehat{f}_\ddagger(y'_\alpha)$ follow from (b) and Proposition 5.2 (c) and (d). □

Lemma 5.11. *Let Π_1, Π_1^0, Π_2 and Π_2^0 be defined as in the proof of Proposition 5.10.*

Then, there exist $f_1, f_2 \in C^\infty(\mathbb{R}^q)$ such that, for all $y \in \mathbb{R}^q$,

$$\begin{aligned}\Pi_1(y) \leq & f_1(y) \leq \Pi_1^0(y),\\ \Pi_2(y) \geq & f_2(y) \geq \Pi_2^0(y)\end{aligned}$$

with equality only in y_α and y'_α, respectively.

Proof. In this proof, we only consider the case

$$\Pi_1(y) \leq f_1(y) \leq \Pi_1^0(y),$$

for $y \in \mathbb{R}^q$ with equality only in y_α, since the other statement follows analogously.

Our goal is to find f_1 by first constructing a function that is squeezed between Π_1 and Π_1^0, using the Whitney Extension Theorem [27, Theorem 2.3.6], and then modifying it to obtain f_1.

Recall that, by construction, we have that

$$\Pi_1(y) < \Pi_1^0(y) \qquad \text{for } y \in \mathbb{R}^q \setminus \{y_\alpha\}$$

and

$$\Pi_1(y_\alpha) = \Pi_1^0(y_\alpha), \quad D\Pi_1(y_\alpha) = D\Pi_1^0(y_\alpha), \quad D^2\Pi_1(y_\alpha) < D^2\Pi_1^0(y_\alpha).$$

We apply the Whitney Extension Theorem to $\frac{1}{2}(\Pi_1 + \Pi_1^0)$ on the closed set $A = \{y_\alpha\}$, yielding a function $\psi_1 \in C^2(\mathbb{R}^q)$ such that $\Pi_1 \leq \psi_1 \leq \Pi_1^0$ on $B_{2\delta}(y_\alpha)$ for some $\delta > 0$ with equality only in y_α. Inspecting the construction of ψ_1 in the proof of [45, Theorem II], we find that $\psi_1 \in C^\infty(\mathbb{R}^q)$.

Next, we modify ψ_1 such that the resulting function stays between Π_1 and Π_1^0 on all of $\mathbb{R}^q$. As smooth functions are dense in the set of continuous functions, we can find a function $\psi_2 \in C^\infty(\mathbb{R}^q)$ such that $\Pi_1 < \psi_1 < \Pi_1^0$ on $\mathbb{R}^q \setminus B_\delta(y_\alpha)$.

Then, defining

$$f_1(y) = \ell(y)\psi_1(y) + (1 - \ell(y))\psi_2(y),$$

where ℓ is a smooth function that is 1 on $B_\delta(y_\alpha)$ and 0 outside of $B_{2\delta}(y_\alpha)$, for example ℓ as defined as point (3) on [42, p. 33]. This concludes the proof. □

6. Proof of the strict comparison principle

In this section, we prove Theorem 3.12. The proof is based on a variant of the variable quadruplication procedure on the basis of

$$\begin{aligned}\sup_{x\in\mathbb{R}^q} & \frac{1}{1-\varepsilon}u(x) - \frac{1}{1+\varepsilon}v(x)\\ & \leq \sup_{x,y,y'x'\in\mathbb{R}^q} \frac{1}{1-\varepsilon}u(x) - \frac{1}{1+\varepsilon}v(x') - \frac{\alpha}{2(1-\varepsilon)}d^2(x,y) - \frac{\alpha}{2}d^2(y,y')\\ & \qquad - \frac{\alpha}{2(1+\varepsilon)}d^2(y',x') - \frac{\varepsilon}{1+\varepsilon}V(x) - \frac{\varepsilon}{1+\varepsilon}V(x'),\end{aligned}$$

which we have formalized in terms of test functions $f_\dagger, f_\ddagger$ in Proposition 5.2 and Theorem 5.10.

In a first step, we relate sub- and supersolutions for the Hamilton–Jacobi equation for H to those for H_+ and H_-; this will be carried out in Lemma 6.1. A second step is to show that $f_\dagger \in \mathcal{D}(H_+)$ and $f_\ddagger \in \mathcal{D}(H_-)$; this will be carried out in Lemma 6.2.

After establishing these technical points, we proceed to frame the comparison principle in terms of an estimate on the difference

$$\frac{H_+ f_\dagger}{1-\varepsilon} - \frac{H_- f_\ddagger}{1+\varepsilon}. \tag{6.1}$$

This reduction will be carried out in Proposition 6.3, the statement of which is more involved than typical in the literature, but leads to the improved *strict comparison principle*. Its formulation and proof hinges on the use of V as a Lyapunov function.

The statements of Lemmas 6.1, 6.2, and Proposition 6.3 can be found in Section 6.1, their proofs in Section 6.2.

We finish in Section 6.3 by estimating (6.1) in two steps leading to our final result. We first establish in Lemma 6.4 that the pre-factors $(1-\varepsilon)^{-1}$ and $(1+\varepsilon)^{-1}$ work well with the combinations of functions that define $f_\dagger, f_\ddagger$ in Theorem 5.10. We conclude the section with the proof of Theorem 3.12, where we use this split, the coupling assumption on $\mathbb{A}$, the semi-monotonicity of $\mathbb{B}$, modulus of continuity control on $\mathcal{I}$ and, again, that V is a Lyapunov function to arrive at our final result.

6.1. Comparison in terms of estimating the difference of Hamiltonians

We start with connecting the notion of sub- and supersolutions for H to those for H_+ and H_-, respectively.

Lemma 6.1. *Let H and $\mathbb{H}$ satisfy Assumption 3.15. Then, for any $h \in C_b(\mathbb{R}^q)$ and $\lambda > 0$, we have the following:*

(a) *Any viscosity subsolution of $f - \lambda H f = h$ is also a viscosity subsolution of $f - \lambda H_+ f = h$.*
(b) *Any viscosity supersolution of $f - \lambda H f = h$ is also a viscosity supersolution of $f - \lambda H_- f = h$.*

The proof follows in Section 6.2 below. In the next lemma we show that the test functions that we constructed in the previous section are in the domain of H_+ and H_-.

Lemma 6.2. *Let $\mathbb{H}$ be an operator satisfying Assumptions 3.15 and 3.16. Let $\widehat{f}_\dagger, f_\dagger$ and $\widehat{f}_\ddagger, f_\ddagger$ be as in Theorem 5.10.*

Then, $\widehat{f}_\dagger, f_\dagger \in \mathcal{D}(H_+)$ and $\widehat{f}_\ddagger, f_\ddagger \in \mathcal{D}(H_-)$.

The proof of the lemma is outlined in Section 6.2 below. We next state our key proposition, which relates the strict comparison principle to an estimate on the difference of Hamiltonians.

Proposition 6.3. *Let $\mathbb{H} \subseteq C(\mathbb{R}^q) \times C(\mathbb{R}^q)$ satisfy Assumptions 3.15 and 3.16. Let $h_1, h_2 \in C_b(\mathbb{R}^q)$, and $\lambda > 0$. Consider the equations*

$$f - \lambda H_+ f \leq h_1, \tag{6.2}$$

$$f - \lambda H_- f \geq h_2. \tag{6.3}$$

Let u and v by viscosity sub- and supersolutions to (6.2) and (6.3), respectively. For each $\varepsilon \in (0,1)$, $\varphi \in (0,1]$ and $\alpha > 1$, consider the construction of optimizers x_α, x'_α and test functions $f_\dagger, f_\ddagger$ as in Proposition 5.2 and Theorem 5.10.

Suppose there exists a map $\varepsilon \mapsto C^0_\varepsilon$, and for any $\varepsilon \in (0,1)$ a non-negative map $\varphi \mapsto C_{\varepsilon,\varphi}$ satisfying $\limsup_{\varepsilon\downarrow 0} C^0_\varepsilon < \infty$ and $\lim_{\varphi\downarrow 0} C_{\varepsilon,\varphi} = 0$ such that

$$\liminf_{\alpha\to\infty} \frac{H_+ f_\dagger(x_\alpha)}{1-\varepsilon} - \frac{H_- f_\ddagger(x'_\alpha)}{1+\varepsilon} \le \varepsilon\left(C^0_\varepsilon + C_{\varepsilon,\varphi}\right). \tag{6.4}$$

Then, for any compact set $K \subseteq \mathbb{R}^q$ and $\varepsilon \in (0,1)$,

$$\sup_{x\in K} u(x) - v(x) \le \varepsilon C_\varepsilon + \sup_{x\in \widehat{K}} h_1(x) - h_2(x),$$

where $\widehat{K}_\varepsilon := \widehat{K}_\varepsilon(K,u,v)$ and $C_\varepsilon := C_\varepsilon(K,u,v,h_1,h_2)$ are given by

$$\widehat{K}_\varepsilon := \left\{ z \in \mathbb{R}^q \,\middle|\, V(z) \le \frac{\|u\| + \|v\|}{\varepsilon} + \lceil V \rceil_K \right\},$$
$$C_\varepsilon := \lambda C^0_\varepsilon + \frac{2}{1-\varepsilon^2}\lceil V\rceil_K + \frac{1}{1-\varepsilon}\|h_1\| + \frac{1}{1-\varepsilon}\|h_2\| - \left\lfloor \frac{1}{1-\varepsilon}u - \frac{1}{1+\varepsilon}v \right\rfloor_K.$$

In particular, the strict comparison principle holds for (6.2) and (6.3).

6.2. Proof of Lemmas 6.1, 6.2, and Proposition 6.3

Proof of Lemma 6.1. We only prove the first statement as the second one follows analogously. Let u be a subsolution to $f - \lambda H f = h$ and let $(f,g) \in H_+$. Our claim thus follows if there exists x_0 satisfying

$$u(x_0) - f(x_0) = \lceil u - f \rceil, \tag{6.5}$$
$$u(x_0) - \lambda g(x_0) \le h(x_0). \tag{6.6}$$

As u is upper semi-continuous and bounded, and f has compact sublevel sets, the existence of x_0 satisfying (6.5) is immediate. We thus proceed with (6.6) using the sequential upward denseness of $\mathcal{D}(H)$ in $\mathcal{D}(H_+)$, cf. Assumption 3.15 (c). Set

$$a := f(x_0) + \lceil u\rceil - u(x_0), \qquad A := \{x \,|\, f(x) \le a\}.$$

We can thus find $(f_a, g_a) \in H$ with f_a satisfying

$$\begin{cases} f_a(x) = f(x) & \text{if } x \in A, \\ a < f_a(x) \le f(x) & \text{if } x \notin A. \end{cases}$$

We first establish that

$$u(x_0) - f_a(x_0) = \lceil u - f_a \rceil. \tag{6.7}$$

Using (6.5) and that $f = f_a$ on A, (6.7) follows by verifying that

$$u(x) - f_a(x) < u(x_0) - f(x_0), \qquad x \in A^c,$$

which follows from the definition of a:

$$\begin{aligned} u(x) - f(x) &< u(x) - a \\ &= u(x) - (f(x_0) + \lceil u \rceil - u(x_0)) \\ &= u(x_0) - f(x_0) - (\lceil u \rceil - u(x)) \\ &\leq u(x_0) - f(x_0). \end{aligned}$$

Thus, by (6.7), we can use the subsolution inequality for (f_a, g_a) in the point x_0. We obtain:

$$u(x_0) - \lambda g_a(x_0) \leq h(x_0). \tag{6.8}$$

Recalling that $f_a(x_0) = f(x_0)$ and $f_a \leq f$, we have

$$f_a(x_0) - f(x_0) = \lceil f_a - f \rceil .$$

Using the positive maximum principle for $\mathbb{H}$, cf. Assumption 3.15 (a), thus yields

$$g_a(x_0) \leq g(x_0). \tag{6.9}$$

Combining (6.8) and (6.9), leads to

$$u(x_0) - \lambda g(x_0) \leq u(x_0) - \lambda g_a(x_0) \leq h(x_0),$$

establishing (6.6) and consequently that u is a subsolution to $f - \lambda H_+ f = h$. □

Proof of Lemma 6.2. As $f_1, f_2, \widehat{f}_1, \widehat{f}_2 \in C_c^\infty(\mathbb{R}^q)$, it follows by Assumption 3.15 (b) that $f_1, f_2, \widehat{f}_1, \widehat{f}_2 \in \mathcal{D}(H)$. By compatibility, cf. Assumption 3.16, we have $V \circ s_z, \Xi \circ s_z \in \mathcal{D}(\mathbb{H})$. By Assumption 3.15 (e) and the fact that V has compact sublevel sets, cf. Definition 3.7, we thus have $(1-\varphi)V \circ s_z + \varphi \Xi \circ s_z \in \mathcal{D}(H_+)$. Consequently, $\widehat{f}_\dagger, f_\dagger \in \mathcal{D}(H_+)$ and $\widehat{f}_\ddagger, f_\ddagger \in \mathcal{D}(H_-)$ by Assumption 3.15 (f). □

Proof of Proposition 6.3. Let u be a subsolution to $f - \lambda H_+ f \leq h_1$ and v a supersolution to $f - \lambda H_- f \geq h_2$. Consider the constructions in Proposition 5.2 and Theorem 5.10 for the subsolution u, supersolution v, $\varepsilon \in (0,1)$, and $\varphi \in (0,1]$.

By Lemma 6.2, we have $f_\dagger \in \mathcal{D}(H_+)$ and $f_\ddagger \in \mathcal{D}(H_-)$ and, by Proposition 5.10 (c), we find that (x_α, x'_α) are the unique optimizers in

$$\begin{aligned} u(x_\alpha) - f_\dagger(x_\alpha) &= \lceil u - f_\dagger \rceil , \\ v(x'_\alpha) - f_\ddagger(x'_\alpha) &= \lfloor v - f_\ddagger \rfloor , \end{aligned}$$

which, by the sub- and supersolution properties for H_+ and H_-, respectively, and Lemma Appendix B.1, implies that

$$\begin{aligned} u(x_\alpha) - \lambda H_+ f_\dagger(x_\alpha) &\leq h_1(x_\alpha), \\ v(x'_\alpha) - \lambda H_- f_\ddagger(x'_\alpha) &\geq h_2(x'_\alpha). \end{aligned} \tag{6.10}$$

By Proposition 5.2 (i), we find

$$\lceil u - v \rceil_K \leq \frac{1}{1-\varepsilon} u(x_\alpha) - \frac{1}{1+\varepsilon} v(x'_\alpha) + \varepsilon \left(c_{\varepsilon,\varphi} + o(1) \right),$$

where

$$c_{\varepsilon,\varphi} := \frac{2}{1-\varepsilon^2}(1-\varphi)\lceil V\rceil_K - \left\lfloor \frac{1}{1-\varepsilon}u - \frac{1}{1+\varepsilon}v \right\rfloor_K, \tag{6.11}$$

and $o(1)$ is in terms of $\alpha \to \infty$. Using (6.10), we estimate

$$\begin{aligned}
\lceil u-v\rceil_K &\leq \frac{1}{1-\varepsilon}u(x_\alpha) - \frac{1}{1+\varepsilon}v(x'_\alpha) + \varepsilon\left(c_{\varepsilon,\varphi} + o(1)\right) \\
&\leq \frac{1}{1-\varepsilon}h_1(x_\alpha) - \frac{1}{1+\varepsilon}h_2(x'_\alpha) + \lambda\left[\frac{H_+ f_\dagger(x_\alpha)}{1-\varepsilon} - \frac{H_- f_\ddagger(x'_\alpha)}{1+\varepsilon}\right] + \varepsilon\left(c_{\varepsilon,\varphi} + o(1)\right) \\
&\leq h_1(x_\alpha) - h_2(x'_\alpha) + \lambda\left[\frac{H_+ f_\dagger(x_\alpha)}{1-\varepsilon} - \frac{H_- f_\ddagger(x'_\alpha)}{1+\varepsilon}\right] \\
&\qquad + \frac{\varepsilon}{1-\varepsilon}\|h_1\| + \frac{\varepsilon}{1+\varepsilon}\|h_2\| + \varepsilon\left(c_{\varepsilon,\varphi} + o(1)\right).
\end{aligned}$$

We next expand $c_{\varepsilon,\varphi}$ from (6.11). Furthermore, taking $\liminf_{\alpha\to\infty}$ on the right-hand side, using Proposition 5.2 (j) to treat the difference $h_1 - h_2$, and (6.4) to treat the difference of Hamiltonians, we find

$$\begin{aligned}
\lceil u-v\rceil_K \leq \lceil h_1 - h_2\rceil_{\widehat{K}} + \lambda\left(\varepsilon C_0 + C_{\varepsilon,\varphi}\right) + \frac{\varepsilon}{1-\varepsilon}\|h_1\| + \frac{\varepsilon}{1+\varepsilon}\|h_2\| \\
+ \varepsilon\left(\frac{2}{1-\varepsilon^2}(1-\varphi)\lceil V\rceil_K - \left\lfloor \frac{1}{1-\varepsilon}u - \frac{1}{1+\varepsilon}v\right\rfloor_K\right).
\end{aligned}$$

As $\varphi \in (0,1]$ was arbitrary, we can take the limit for $\varphi \downarrow 0$, which leads to

$$\begin{aligned}
\lceil u-v\rceil_K &\leq \lceil h_1 - h_2\rceil_{\widehat{K}} \\
&\quad + \varepsilon\left(\lambda C^0_\varepsilon + \frac{2}{1-\varepsilon^2}\lceil V\rceil_K + \frac{1}{1-\varepsilon}\|h_1\| + \frac{1}{1+\varepsilon}\|h_2\| - \left\lfloor \frac{1}{1-\varepsilon}u - \frac{1}{1+\varepsilon}v\right\rfloor_K\right),
\end{aligned}$$

establishing the claim. □

6.3. Proof of Theorem 3.12

We start with an auxiliary lemma that provides a detailed decomposition of the operators $\mathbb{A}$ and $\mathbb{B}$ evaluated in the test functions.

Lemma 6.4. *Let $\mathbb{A}$ and $\mathbb{B}$ satisfy Assumption 3.15 and Assumption 3.16 (a) and (b), respectively. Fix $z_0, z_1 \in \mathbb{R}^q$ and $p \in \mathbb{R}^q$. Let $\Xi = \Xi_{z_0,p,z_1}$ as in Definition 3.10, V as in Definition 3.9, and, for $\widehat{f} \in C_c^\infty(\mathbb{R}^q)$, $\varepsilon \in (0,1)$, and $\varphi \in (0,1]$, set*

$$\begin{aligned}
\widehat{f}_\dagger &:= (1-\varepsilon)\widehat{f} + \varepsilon(1-\varphi)V + \varepsilon\varphi\Xi, \\
\widehat{f}_\ddagger &:= (1+\varepsilon)\widehat{f} - \varepsilon(1-\varphi)V - \varepsilon\varphi\Xi.
\end{aligned}$$

For $z \in \mathbb{R}^q$, set $f_\dagger = \widehat{f}_\dagger \circ s_z$, and $f_\ddagger = \widehat{f}_\ddagger \circ s_z$. Then, the following statements hold:

(a) *$f_\dagger \in \mathcal{D}(A_+)$ and $f_\ddagger \in \mathcal{D}(A_-)$. Suppose furthermore that $\mathbb{A}$ is linear on its domain, then*

$$\begin{aligned}
\frac{A_+ f_\dagger}{1-\varepsilon} &= A(\widehat{f}\circ s_z) + \frac{\varepsilon}{1-\varepsilon}(1-\varphi)A_+\left(V\circ s_z\right) + \frac{\varepsilon}{1-\varepsilon}\varphi\mathbb{A}\left(\Xi\circ s_z\right), \\
\frac{A_- f_\ddagger}{1+\varepsilon} &= A(\widehat{f}\circ s_z) - \frac{\varepsilon}{1+\varepsilon}(1-\varphi)A_+\left(V\circ s_z\right) - \frac{\varepsilon}{1+\varepsilon}\varphi\mathbb{A}\left(\Xi\circ s_z\right),
\end{aligned} \tag{6.12}$$

(b) $f_\dagger, \widehat{f}_\dagger \in \mathcal{D}(B_+)$ *and* $f_\ddagger, \widehat{f}_\ddagger \in \mathcal{D}(B_-)$. *Suppose furthermore that* $\mathbb{B}$ *is convex semi-monotone, then, for any* x, y *such that* $z = x - y$, *we have*

$$\begin{aligned} \frac{B_+ f_\dagger}{1-\varepsilon}(x) &\leq \frac{1}{1-\varepsilon}\left(B_+ f_\dagger(x) - B_+\widehat{f}_\dagger(y)\right) + B\widehat{f}(y) \\ &\qquad + \frac{\varepsilon}{1-\varepsilon}(1-\varphi)B_+V(y) + \frac{\varepsilon}{1-\varepsilon}\varphi B_+\Xi(y), \\ \frac{B_- f_\ddagger}{1+\varepsilon}(x) &\geq \frac{1}{1+\varepsilon}\left(B_- f_\ddagger(x) - B_-\widehat{f}_\ddagger(y)\right) + B\widehat{f}(y) \\ &\qquad - \frac{\varepsilon}{1+\varepsilon}(1-\varphi)B_-V(y) - \frac{\varepsilon}{1+\varepsilon}\varphi B_-\Xi(y). \end{aligned} \tag{6.13}$$

Proof. The domain statements $f_\dagger \in \mathcal{D}(A_+)$, $f_\ddagger \in \mathcal{D}(A_-)$, $f_\dagger, \widehat{f}_\dagger \in \mathcal{D}(B_+)$ and $f_\ddagger, \widehat{f}_\ddagger \in \mathcal{D}(B_-)$ follow by Lemma 6.2. The four statements in (6.12) and (6.13) follow from linearity of A_+ and convex semi-monotonicity of B_+. □

Proof of Theorem 3.12. To prove inequality (3.6), and consequently the strong comparison principle for the Hamilton–Jacobi equation in terms of H, it suffices by Lemma 6.1 and Proposition 6.3 to establish (6.4), which we repeat for readability:

$$\liminf_{\alpha\to\infty} \frac{H_+ f_\dagger(x_\alpha)}{1-\varepsilon} - \frac{H_- f_\ddagger(x'_\alpha)}{1+\varepsilon} \leq \varepsilon\left(C^0_\varepsilon + C_{\varepsilon,\varphi}\right). \tag{6.14}$$

Let $\theta^*_{1,\alpha} \in \Theta_1$ be such that

$$\begin{aligned} H_+ f_\dagger(x_\alpha) &= \sup_{\theta_1\in\Theta_1}\inf_{\theta_2\in\Theta_2}\left\{\mathbb{A}_{\theta_1,\theta_2} f_\dagger(x_\alpha) + \mathbb{B}_{\theta_1,\theta_2} f_\dagger(x_\alpha) - \mathcal{I}(x_\alpha,\theta_1,\theta_2)\right\} \\ &= \inf_{\theta_2\in\Theta_2}\left\{\mathbb{A}_{\theta^*_{1,\alpha},\theta_2} f_\dagger(x_\alpha) + \mathbb{B}_{\theta^*_{1,\alpha},\theta_2} f_\dagger(x_\alpha) - \mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta_2)\right\}. \end{aligned}$$

Such optimizer exists by the compactness of Θ_1 and the lower semi-continuity of $\mathcal{I}$ in θ_1 assumed in (c). By Isaacs' condition (d), we can write

$$H_- f_\ddagger(x'_\alpha) = \inf_{\theta_2\in\Theta_2}\sup_{\theta_1\in\Theta_1}\left\{\mathbb{A}_{\theta_1,\theta_2} f_\ddagger(x'_\alpha) + \mathbb{B}_{\theta_1,\theta_2} f_\ddagger(x_\alpha) - \mathcal{I}(x'_\alpha,\theta_1,\theta_2)\right\}.$$

Then, by compactness of Θ_2 and the upper semi-continuity of $\mathcal{I}$ in θ_2 assumed in (c), we can find $\theta^*_{2,\alpha} \in \Theta_2$ such that

$$H_- f_\ddagger(x'_\alpha) = \sup_{\theta_1\in\Theta_1}\left\{\mathbb{A}_{\theta_1,\theta^*_{2,\alpha}} f_\ddagger(x'_\alpha) + \mathbb{B}_{\theta_1,\theta^*_{2,\alpha}} f_\ddagger(x'_\alpha) - \mathcal{I}(x'_\alpha,\theta_1,\theta^*_{2,\alpha})\right\}.$$

Consequently, we can estimate

$$\begin{aligned} \frac{1}{1-\varepsilon}H_+ f_\dagger(x_\alpha) - \frac{1}{1+\varepsilon}H_- f_\ddagger(x'_\alpha) &\leq \underbrace{\left[\frac{1}{1-\varepsilon}\mathbb{A}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} f_\dagger(x_\alpha) - \frac{1}{1+\varepsilon}\mathbb{A}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} f_\ddagger(x'_\alpha)\right]}_{(1)} \\ &\qquad + \underbrace{\left[\frac{1}{1-\varepsilon}\mathbb{B}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} f_\dagger(x_\alpha) - \frac{1}{1+\varepsilon}\mathbb{B}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} f_\ddagger(x'_\alpha)\right]}_{(2)} \end{aligned}$$

$$+\underbrace{\left[\frac{1}{1+\varepsilon}\mathcal{I}(x'_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha})-\frac{1}{1-\varepsilon}\mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha})\right]}_{(3)}.$$

We treat (1), (2), and (3) separately. Note, that due the compactness of Θ_1 and Θ_2, along subsequences, the optimizers $\theta^*_{1,\alpha}$ and $\theta^*_{2,\alpha}$ converge to some θ^*_1 and θ^*_2, respectively.

Estimate (1): Using the expansions of $A_+f_\dagger$ and $A_-f_\ddagger$ obtained in Lemma 6.4 we find

$$\begin{aligned}
\frac{\mathbb{A}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}f_\dagger(x_\alpha)}{1-\varepsilon}-\frac{\mathbb{A}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}f_\ddagger(x'_\alpha)}{1+\varepsilon} &= \frac{A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+}f_\dagger(x_\alpha)}{1-\varepsilon}-\frac{A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},-}f_\ddagger(x'_\alpha)}{1+\varepsilon}\\
&\le A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}f_1(x_\alpha)-A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}f_2(x'_\alpha)\\
&\quad+\frac{\varepsilon}{1-\varepsilon}(1-\varphi)A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+}\left(V\circ s_{x_\alpha-y_\alpha}\right)(x_\alpha)\\
&\quad+\frac{\varepsilon}{1+\varepsilon}(1-\varphi)A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+}\left(V\circ s_{x'_\alpha-y'_\alpha}\right)(x'_\alpha)\\
&\quad+\frac{\varepsilon}{1-\varepsilon}\varphi\mathbb{A}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}\left(\Xi_1\circ s_{x_\alpha-y_\alpha}\right)(x_\alpha)\\
&\quad+\frac{\varepsilon}{1+\varepsilon}\varphi\mathbb{A}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}\left(\Xi_2\circ s_{x'_\alpha-y'_\alpha}\right)(x'_\alpha).
\end{aligned}\tag{6.15}$$

We first consider the terms involving V and Ξ. By Proposition 5.2 (j), we have that, along subsequences, the optimizers $(x_\alpha,y_\alpha,y_{\alpha,0},y'_{\alpha,0},y'_\alpha,x'_\alpha)$ converge to (z,z,z,z,z,z) with $z\in\widehat{K}$ and $p_\alpha,p'_\alpha\in B_{1/\alpha}(0)$. Then, using the compatibility of $\mathbb{A}_{\theta_1,\theta_2}$, cf. Assumption 3.16, we find

$$\begin{aligned}
&\liminf_{\alpha\to\infty}\frac{\varepsilon}{1-\varepsilon}(1-\varphi)A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+}\left(V\circ s_{x_\alpha-y_\alpha}\right)(x_\alpha)\\
&\qquad+\frac{\varepsilon}{1+\varepsilon}(1-\varphi)A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+}\left(V\circ s_{x'_\alpha-y'_\alpha}\right)(x'_\alpha)\\
&\qquad+\frac{\varepsilon}{1-\varepsilon}\varphi\mathbb{A}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}\left(\Xi_1\circ s_{x_\alpha-y_\alpha}\right)(x_\alpha)+\frac{\varepsilon}{1+\varepsilon}\varphi\mathbb{A}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}\left(\Xi_2\circ s_{x'_\alpha-y'_\alpha}\right)(x'_\alpha)\\
&\qquad\le\frac{2\varepsilon}{1-\varepsilon^2}\left((1-\varphi)A_{\theta^*_1,\theta^*_2,+}(V)(z)+\varphi\mathbb{A}_{\theta^*_1,\theta^*_2}(\Xi_{z,0,z})(z)\right).
\end{aligned}\tag{6.16}$$

Next, we consider the second line in (6.15). Using that, for all θ_1, θ_2, $\mathbb{A}_{\theta_1,\theta_2}$ has a controlled growth coupling $\widehat{\mathbb{A}}_{\theta_1,\theta_2}$ with a modulus uniform in θ_1 and θ_2 satisfying the maximum principle and Theorem 5.10 (a), we find

$$\begin{aligned}
A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}f_1(x_\alpha)-A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}f_2(x'_\alpha) &= \widehat{\mathbb{A}}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}\left(f_1\ominus f_2\right)(x_\alpha,x'_\alpha)\\
&\le\widehat{\mathbb{A}}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}\left(\frac{\alpha}{2}d^2_{x_\alpha-y_\alpha,x'_\alpha-y'_\alpha}\right)(x_\alpha,x'_\alpha)\\
&\le\omega_{\widehat{\mathbb{A}},\widehat{K}}\Big(\alpha\left(d(x_\alpha,y_\alpha)+d(y_\alpha,y'_\alpha)+d(y'_\alpha,x'_\alpha)\right)^2\\
&\qquad+\left(d(x_\alpha,y_\alpha)+d(y_\alpha,y'_\alpha)+d(y'_\alpha,x'_\alpha)\right)\Big),
\end{aligned}\tag{6.17}$$

which, by Proposition 5.2 (g), converges to 0 as $\alpha\to\infty$.

Estimate (2): By using the expansions of $B_+f_\dagger$ and $B_-f_\ddagger$ obtained in Lemma 6.4, we find

$$\begin{aligned}
\frac{\mathbb{B}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}f_\dagger(x_\alpha)}{1-\varepsilon}-\frac{B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}f_\ddagger(x'_\alpha)}{1+\varepsilon} &= \frac{B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+}f_\dagger(x_\alpha)}{1-\varepsilon}-\frac{B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},-}f_\ddagger(x'_\alpha)}{1+\varepsilon}\\
&\le B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}\widehat{f}_1(y_\alpha)-B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}\widehat{f}_2(y'_\alpha)
\end{aligned}$$

$$+ \frac{1}{1-\varepsilon}\left(B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+}f_\dagger(x_\alpha) - B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+}\widehat{f}_\dagger(y_\alpha)\right)$$
$$+ \frac{1}{1+\varepsilon}\left(B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},-}\widehat{f}_\ddagger(y'_\alpha) - B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},-}f_\ddagger(x'_\alpha)\right)$$
$$+ \frac{\varepsilon}{1-\varepsilon}(1-\varphi)B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+}V(y_\alpha) + \frac{\varepsilon}{1+\varepsilon}(1-\varphi)B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+}V(y'_\alpha)$$
$$+ \frac{\varepsilon}{1-\varepsilon}\varphi\mathbb{B}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}\Xi_1(y_\alpha) + \frac{\varepsilon}{1+\varepsilon}\varphi\mathbb{B}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}\Xi_2(y'_\alpha).$$

Again, by sending $\alpha \to \infty$, using Proposition 5.2 (j), and the compatibility of $\mathbb{B}_{\theta_1,\theta_2}$, cf. Assumption 3.16, we obtain that

$$\begin{aligned}\liminf_{\alpha\to\infty} & \frac{\varepsilon}{1-\varepsilon}(1-\varphi)B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+}V(y_\alpha) + \frac{\varepsilon}{1+\varepsilon}(1-\varphi)B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+}V(y'_\alpha) \\ & + \frac{\varepsilon}{1-\varepsilon}\varphi\mathbb{B}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}\Xi_1(y_\alpha) + \frac{\varepsilon}{1+\varepsilon}\varphi\mathbb{B}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}\Xi_2(y'_\alpha) \\ & \leq \frac{2\varepsilon}{1-\varepsilon^2}\left((1-\varphi)B_{\theta^*_1,\theta^*_2,+}(V)(z) + \varphi\mathbb{B}_{\theta^*_1,\theta^*_2}(\Xi_{z,0,z})(z)\right).\end{aligned} \tag{6.18}$$

Using that, for all θ_1, θ_2, $\mathbb{B}_{\theta_1,\theta_2}$ is semi-monotone with $\mathcal{B}_{\theta_1,\theta_2}$ and the expressions for the gradients obtained in Proposition 5.10, we find that

$$\begin{aligned}&\frac{1}{1-\varepsilon}\left(B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+}f_\dagger(x_\alpha) - B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+}\widehat{f}_\dagger(y_\alpha)\right) \\ &\qquad + B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}\widehat{f}_1(y_\alpha) - B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}\widehat{f}_2(y'_\alpha) \\ &\qquad + \frac{1}{1+\varepsilon}\left(B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},-}\widehat{f}_\ddagger(y'_\alpha) - B_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},-}f_\ddagger(x'_\alpha)\right) \\ &= \frac{1}{1-\varepsilon}\left(\mathcal{B}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}(x_\alpha,\alpha(x_\alpha-y_\alpha)) - \mathcal{B}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}(y_\alpha,\alpha(x_\alpha-y_\alpha))\right) \\ &\qquad + \mathcal{B}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}(y_\alpha,\alpha(y_\alpha-y'_\alpha)) - \mathcal{B}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}(y'_\alpha,\alpha(y_\alpha-y'_\alpha)) \\ &\qquad + \frac{1}{1+\varepsilon}\left(\mathcal{B}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}(y_\alpha,\alpha(y'_\alpha-x'_\alpha)) - \mathcal{B}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}}(x'_\alpha,\alpha(y'_\alpha-y'_\alpha))\right).\end{aligned} \tag{6.19}$$

By the semi-monotonicity property of $\mathbb{B}_{\theta_1,\theta_2}$, (6.19) is bounded by

$$\begin{aligned}&\frac{1}{1-\varepsilon}\omega_{\mathcal{B},\widehat{K}}(d(x_\alpha,y_\alpha)+\alpha d^2(x_\alpha,y_\alpha)) + \omega_{\mathcal{B},\widehat{K}}(d(y_\alpha,y'_\alpha)+\alpha d^2(y_\alpha,y'_\alpha)) \\ &\qquad + \frac{1}{1+\varepsilon}\omega_{\mathcal{B},\widehat{K}}(d(y'_\alpha,x'_\alpha)+\alpha d^2(y'_\alpha,x'_\alpha)).\end{aligned} \tag{6.20}$$

Thus, by Proposition 5.2 (g), taking the $\liminf_{\alpha\to\infty}$ gives 0.

Estimate (3): We have

$$\begin{aligned}&\frac{1}{1+\varepsilon}\mathcal{I}(x'_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) - \frac{1}{1-\varepsilon}\mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) \\ &= \left[\mathcal{I}(x'_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) - \mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha})\right] - \frac{\varepsilon}{1-\varepsilon}\mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) - \frac{\varepsilon}{1+\varepsilon}\mathcal{I}(x'_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}).\end{aligned}$$

By assumption, $\mathcal{I}$ admits a modulus of continuity $\omega_{\mathcal{I},K}$, uniform in θ_1, θ_2, implying

$$\frac{1}{1+\varepsilon}\mathcal{I}(x'_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) - \frac{1}{1-\varepsilon}\mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) \\ \leq \omega_{\mathcal{I},\widehat{K}}(d(x_\alpha,x'_\alpha)) - \frac{\varepsilon}{1-\varepsilon}(1-\varphi)\mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) - \frac{\varepsilon}{1+\varepsilon}(1-\varphi)\mathcal{I}(x'_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}).$$

Sending $\alpha\to\infty$, using the lower semi-continuity of $\mathcal{I}$, and using Proposition 5.2 (j), we find

$$\begin{aligned}
&\liminf_{\alpha\to\infty} \frac{1}{1+\varepsilon}\mathcal{I}(x'_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) - \frac{1}{1-\varepsilon}\mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) \\
&\quad \leq \liminf_{\alpha\to\infty} \omega_{\mathcal{I},\widehat{K}}(d(x_\alpha,x'_\alpha)) \\
&\quad\quad + \limsup_{\alpha\to\infty}\left[-\frac{\varepsilon}{1-\varepsilon}(1-\varphi)\mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) - \frac{\varepsilon}{1+\varepsilon}(1-\varphi)\mathcal{I}(x'_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha})\right] \\
&\quad \leq -\frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\mathcal{I}(z,\theta^*_1,\theta^*_2).
\end{aligned} \tag{6.21}$$

Conclusion: Putting together (6.16), (6.17), (6.18), (6.20), and (6.21), we can conclude that

$$\begin{aligned}
\liminf_{\alpha\to\infty} \frac{H_+f_\dagger(x_\alpha)}{1-\varepsilon} - \frac{H_-f_\ddagger(x'_\alpha)}{1+\varepsilon} &\leq \frac{2\varepsilon}{1-\varepsilon^2}\left((1-\varphi)A_{\theta^*_1,\theta^*_2,+}V(z) + \varphi\mathbb{A}_{\theta^*_1,\theta^*_2}(\Xi_{z,0,z})(z)\right) \\
&\quad + \frac{2\varepsilon}{1-\varepsilon^2}\left((1-\varphi)B_{\theta^*_1,\theta^*_2,+}V(z) + \varphi\mathbb{B}_{\theta^*_1,\theta^*_2}(\Xi_{z,0,z})(z)\right) \\
&\quad - \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\mathcal{I}(z,\theta^*_1,\theta^*_2) \\
&\leq \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\left\lceil (A_{\theta^*_1,\theta^*_2,+} + B_{\theta^*_1,\theta^*_2,+})(V) - \mathcal{I}(\cdot,\theta^*_1,\theta^*_2)\right\rceil \\
&\quad + \frac{2\varepsilon}{1-\varepsilon^2}\varphi\left\lceil (\mathbb{A}_{\theta^*_1,\theta^*_2} + \mathbb{B}_{\theta^*_1,\theta^*_2})(\Xi_{\cdot,0,\cdot})\right\rceil_{\widehat{K}_\varepsilon} \\
&\leq \varepsilon\left(\frac{2}{1-\varepsilon^2}c_V^+ + \frac{2}{1-\varepsilon^2}\varphi\left\lceil (\mathbb{A}_{\theta^*_1,\theta^*_2} + \mathbb{B}_{\theta^*_1,\theta^*_2})(\Xi_{\cdot,0,\cdot})\right\rceil_{\widehat{K}_\varepsilon}\right) \\
&\leq \varepsilon\left(C^0_\varepsilon + C_{\varepsilon,\varphi}\right),
\end{aligned}$$

where $c_V^+ := (c_V \vee 0)$ with c_V given by (3.4), and C^0_ε and $C_{\varepsilon,\varphi}$ defined via the last two lines. The estimate on the difference of Hamiltonians (6.14) and thus (6.4) of Proposition 6.3 are satisfied. As a consequence our final estimate (3.6) and, consequently, the strict comparison principle follow. □

Appendix A. Convergence of integrals

Lemma Appendix A.1. *Let $\mathcal{X}$ be a Polish space, $W\colon\mathcal{X}\to(0,\infty)$ be a continuous function, and ν_n,ν_∞ be non-negative Borel measures with $\int_\mathcal{X} W\,\mathrm{d}\nu_n<\infty$ for all $n\in\mathbb{N}$ and*

$$\lim_{n\to\infty}\int\phi\,\mathrm{d}\nu_n = \int\phi\,\mathrm{d}\nu_\infty\in\mathbb{R} \tag{A.1}$$

for every function $\phi\in C(\mathcal{X})$ with $|\phi(x)|\leq W(x)$ for all $x\in\mathcal{X}$. Moreover, let $\phi_n,\phi_\infty\in C(\mathcal{X})$ with $\phi_n\to\phi_\infty$ uniformly on compacts and $\sup_{n\in\mathbb{N}}\sup_{x\in\mathcal{X}}\frac{|\phi_n(x)|}{W(x)}<\infty$. Then,

$$\lim_{n\to\infty}\int\phi_n\,\mathrm{d}\nu_n = \int\phi_\infty\,\mathrm{d}\nu_\infty.$$

Proof. By assumption, the family $\mu_n := W \mathrm{d}\nu_n$ satisfies $C_\mu := \sup_{n\in\mathbb{N}} \mu_n(\mathcal{X}) < \infty$ and

$$C_\phi := \sup_{n\in\mathbb{N}} \sup_{x\in\mathcal{X}} \frac{|\phi_n(x) - \phi_\infty(x)|}{W(x)} < \infty.$$

Using the fact that a function $\phi \in C(\mathcal{X})$ satisfies $|\phi(x)| \le W(x)$ for all $x \in \mathcal{X}$ if and only if $\phi = W\psi$ for some $\psi \in C_b(\mathcal{X})$, it follows that $\mu_n \to \mu_\infty := W \mathrm{d}\nu_\infty$ weakly. In particular, the family $(\mu_n)_{n\in\mathbb{N}}$ is tight. Hence, for all $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subseteq \mathcal{X}$ such that

$$C_\phi \mu_n(\mathcal{X} \setminus K_\varepsilon) < \frac{\varepsilon}{3} \quad \text{for all } n \in \mathbb{N}\,.$$

Now, let $\varepsilon > 0$. By (A.1) and since $\phi_n \to \phi_\infty$ uniformly on compacts and W is continuous, there exists some $n_0 \in \mathbb{N}$ such that

$$C_\mu \sup_{x\in K_\varepsilon} \frac{|\phi_n(x) - \phi_\infty(x)|}{W(x)} < \frac{\varepsilon}{3} \quad \text{and} \quad \left| \int \phi_\infty \, \mathrm{d}\nu_n - \int \phi_\infty \, \mathrm{d}\nu_\infty \right| < \frac{\varepsilon}{3}.$$

We thus obtain that

$$\begin{aligned} \left| \int \phi_n \, \mathrm{d}\nu_n - \int \phi_\infty \, \mathrm{d}\nu_\infty \right| &\le \int |\phi_n - \phi_\infty| \, \mathrm{d}\nu_n + \left| \int \phi_\infty \, \mathrm{d}\nu_n - \int \phi_\infty \, \mathrm{d}\nu_\infty \right| \\ &\le \int_{K_\varepsilon} |\phi_n - \phi_\infty| \, \mathrm{d}\nu_n + \int_{\mathcal{X}\setminus K_\varepsilon} |\phi_n - \phi_\infty| \, \mathrm{d}\nu_n + \frac{\varepsilon}{3} \\ &\le C_\mu \frac{|\phi_n(x) - \phi_\infty(x)|}{W(x)} + C_\phi \mu_n(\mathcal{X} \setminus K_\epsilon) + \frac{\varepsilon}{3} < \varepsilon \end{aligned}$$

for all $n \in \mathbb{N}$ with $n \ge n_0$. The proof is complete. □

Appendix B. Equivalent characterization of the definition of viscosity solutions

Lemma Appendix B.1. *Let $H_1 \subseteq C_l(\mathbb{R}^q) \times C(\mathbb{R}^q)$ and $H_2 \subseteq C_u(\mathbb{R}^q) \times C(\mathbb{R}^q)$ be two operators with domains $\mathcal{D}(H_1)$, $\mathcal{D}(H_2)$. Moreover, let $\lambda > 0$ and $h_1 \in C_l(\mathbb{R}^q)$ and $h_2 \in C_u(\mathbb{R}^q)$.*

(a) *Let $u\colon \mathbb{R}^q \to \mathbb{R}$ be u a viscosity subsolution to (2.1). Suppose $\delta > 0$ and $(f, g) \in H_1$ are such that $\{x \in \mathbb{R}^q \,|\, u(x) - f(x) \ge \lceil u - f \rceil - \delta\}$ is compact. Then there exists some $x_0 \in \mathbb{R}^q$ with*

$$\begin{aligned} u(x_0) - f(x_0) &= \sup_{x\in\mathbb{R}^q} u(x) - f(x), \\ u(x_0) - \lambda g(x_0) &\le h_1(x_0). \end{aligned} \tag{B.1}$$

(b) *Let $v\colon \mathbb{R}^q \to \mathbb{R}$ be a viscosity supersolution to (2.2). Suppose $\delta > 0$ and $(f, g) \in H_2$ are such that $\{x \in \mathbb{R}^q \,|\, v(x) - f(x) \le \lfloor v - f \rfloor + \delta\}$ is compact. Then there exists some $x_0 \in \mathbb{R}^q$ with*

$$\begin{aligned} v(x_0) - f(x_0) &= \inf_{x\in\mathbb{R}^q} v(x) - f(x), \\ v(x_0) - \lambda g(x_0) &\ge h_2(x_0). \end{aligned}$$

In particular, the outcomes of (a) and (b) hold if $H_1 \subseteq C_+(\mathbb{R}^q) \times C(\mathbb{R}^q)$ and $H_2 \subseteq C_-(\mathbb{R}^q) \times C(\mathbb{R}^q)$.

Proof. We only show Part (a). Part (b) follows analogously. Assume that u is a viscosity subsolution to (2.1) and let $(f,g) \in H_1$. We aim to establish the existence of x_0 such that (B.1) is satisfied.

Due to the subsolution property of u, there exists a sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^q$ such that

$$\lim_{n\to\infty} u(x_n) - f(x_n) = \lceil u - f \rceil ,$$

$$\limsup_{n\to\infty} u(x_n) - \lambda g(x_n) - h_1(x_n) \leq 0.$$

For n large, we have

$$x_n \in \{x \in \mathbb{R}^q \,|\, u(x) - f(x) \geq \lceil u - f \rceil - \delta\},$$

which is a compact set by assumption. Thus, there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}} \to x_0 \in \mathbb{R}^q$. Since $u(x_n) - f(x_n) \to \lceil u - f \rceil$, it follows that

$$\lceil u - f \rceil = \lim_{k\to\infty} u(x_{n_k}) - f(x_{n_k}) \leq u(x_0) - f(x_0) \leq \lceil u - f \rceil ,$$

where the inequality follows by the upper semi-continuity of $u - f$. The inequality is thus an equality, establishing the first statement of (B.1). Due to the continuity of f, we additionally find that

$$u(x_0) = \lim_{k\to\infty} u(x_{n_k}).$$

Since g and h_1 are continuous, we conclude

$$\begin{aligned} u(x_0) - \lambda g(x_0) - h_1(x_0) &= \lim_{k\to\infty} u(x_{n_k}) - \lambda g(x_{n_k}) - h_1(x_{n_k}) \\ &\leq \limsup_{n\to\infty} u(x_n) - \lambda g(x_n) - h_1(x_n) \leq 0, \end{aligned}$$

establishing the second statement of (B.1). □

References

[1] C.D. Aliprantis, K.C. Border, Infinite Dimensional Analysis, Third. Springer, Berlin, 2006, pp. xxii+703. A hitchhiker's guide.

[2] O. Alvarez, A. Tourin, Viscosity solutions of nonlinear integro-differential equations, Ann. Inst. Henri Poincaré C, Anal. Non Linéaire 13 (3) (1996) 293–317.

[3] M. Arisawa, A new definition of viscosity solutions for a class of second-order degenerate elliptic integro-differential equations, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 23 (5) (2006) 695–711.

[4] M. Arisawa, Homogenization of a class of integro-differential equations with Lévy operators, Commun. Partial Differ. Equ. 34 (7–9) (2009) 617–624.

[5] S. Awatif, Équations d'Hamilton-Jacobi du premier ordre avec termes intégro-différentiels. I. Unicité des solutions de viscosité, Commun. Partial Differ. Equ. 16 (6-7) (1991) 1057–1074.

[6] M. Bardi, I. Capuzzo-Dolcetta, Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations, Birkhäuser, Boston, MA, 1997 (English).

[7] G. Barles, C. Imbert, Second-order elliptic integro-differential equations: viscosity solutions' theory revisited, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 25 (2008) 567–585.

[8] R.F. Bass, Uniqueness in law for pure jump Markov processes, Probab. Theory Relat. Fields 79 (2) (1988) 271–287.

[9] C. Bertucci, Stochastic optimal transport and Hamilton–Jacobi–Bellman equations on the set of probability measures, Ann. Inst. Henri Poincaré C (2024).

[10] J. Blessing, R. Denk, M. Kupper, M. Nendel, Convex monotone semigroups and their generators with respect to Γ-convergence, J. Funct. Anal. 288 (8) (2025) 110841.

[11] R.C. Buck, Bounded continuous functions on a locally compact space, Mich. Math. J. 5 (2) (1958) 95–104.

[12] K. Burdzy, W.S. Kendall, Efficient Markovian couplings: examples and counterexamples, Ann. Appl. Probab. 10 (2) (2000) 362–409.

[13] P. Cannarsa, C. Sinestrari, Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control, vol. 58, Springer Science & Business Media, 2004.

[14] C. Costantini, T. Kurtz, Viscosity methods giving uniqueness for martingale problems, Electron. J. Probab. 20 (67) (2015) 1–27.
[15] M.G. Crandall, L.C. Evans, P.-L. Lions, Some properties of viscosity solutions of Hamilton-Jacobi equations, Trans. Am. Math. Soc. 282 (2) (1984) 487–502.
[16] M.G. Crandall, H. Ishii, The maximum principle for semicontinuous functions, Differ. Integral Equ. 3 (6) (1990) 1001–1014.
[17] M.G. Crandall, H. Ishii, P.-L. Lions, Uniqueness of viscosity solutions of Hamilton-Jacobi equations revisited, J. Math. Soc. Jpn. 39 (4) (1987) 581–596.
[18] 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) (1992) 1–67.
[19] M.G. Crandall, P.-L. Lions, Viscosity solutions of Hamilton-Jacobi equations, Trans. Am. Math. Soc. 277 (1) (1983) 1–42.
[20] R. Denk, M. Kupper, M. Nendel, Convex semigroups on L^p-like spaces, J. Evol. Equ. 21 (2) (2021) 2491–2521.
[21] S.N. Ethier, T.G. Kurtz, Markov Processes: Characterization and Convergence, Wiley, 1986, pp. x+534.
[22] G. Fabbri, F. Gozzi, A. Swiech, Stochastic optimal control in infinite dimension, in: Dynamic Programming and HJB Equations, in: Probability Theory and Stochastic Modelling, vol. 82, Springer, Cham, 2017, pp. xxiii+916. With a contribution by Marco Fuhrman and Gianmario Tessitore.
[23] A. Figalli, N. Gigli, A new transportation distance between non-negative measures, with applications to gradients flows with Dirichlet boundary conditions, J. Math. Pures Appl. 94 (2) (2010) 107–130.
[24] W.H. Fleming, H.M. Soner, Controlled Markov Processes and Viscosity Solutions, 2nd ed., Springer, New York, NY, 2006, pp. xvii + 428 (English).
[25] N. Guillen, C. Mou, A. Świech, Coupling Lévy measures and comparison principles for viscosity solutions, Trans. Am. Math. Soc. 372 (10) (2019) 7327–7370.
[26] J. Hollender, Lévy-type processes under uncertainty and related nonlocal equations, Ph.D. dissertation, Technical University of Dresden, 2016.
[27] L. Hörmander, The Analysis of Linear Partial Differential Operators. I, Classics in Mathematics, Springer-Verlag, Berlin, 2003, pp. x+440. Distribution theory and Fourier analysis. Reprint of the second (1990) edition, Springer, Berlin. MR1065993 (91m:35001a).
[28] M. Hu, S. Peng, G-Lévy processes under sublinear expectations, Probab. Uncertain. Quant. Risk 6 (1) (2021) 1–22.
[29] H. Ishii, Existence and uniqueness of solutions of Hamilton-Jacobi equations, Funkc. Ekvacioj 29 (2) (1986) 167–188.
[30] H. Ishii, Uniqueness of unbounded viscosity solution of Hamilton-Jacobi equations, Indiana Univ. Math. J. 33 (5) (1984) 721–748.
[31] H. Ishii, A. Roch, Existence and uniqueness of viscosity solutions of an integro-differential equation arising in option pricing, SIAM J. Financ. Math. 12 (2) (2021) 604–640.
[32] E.R. Jakobsen, K.H. Karlsen, A "maximum principle for semicontinuous functions" applicable to integro-partial differential equations, NoDEA Nonlinear Differ. Equ. Appl. 13 (2) (2006) 137–165.
[33] E.R. Jakobsen, K.H. Karlsen, Continuous dependence estimates for viscosity solutions of integro-PDEs, J. Differ. Equ. 212 (2) (2005) 278–318.
[34] R. Jensen, P.-L. Lions, P.E. Souganidis, A uniqueness result for viscosity solutions of second order fully nonlinear partial differential equations, Proc. Am. Math. Soc. 102 (4) (1988) 975–978.
[35] R. Jensen, The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations, Arch. Ration. Mech. Anal. 101 (1) (1988) 1–27.
[36] T. Lindvall, Lectures on the Coupling Method, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1992, pp. xiv+257. A Wiley-Interscience Publication.
[37] P.-L. Lions, Solutions de viscosité des équations de Hamilton-Jacobi du premier ordre et applications, in: Goulaouic-Meyer-Schwartz Seminar, 1983–1984, École Polytech, Palaiseau, 1984, Exp. No. 6, 13.
[38] M. Nendel, Lower semicontinuity of monotone functionals in the mixed topology on c b, Finance Stoch. 29 (1) (2025) 261–287.
[39] H. Pham, Continuous-Time Stochastic Control and Optimization with Financial Applications (Stochastic Modelling and Applied Probability), vol. 61, Springer-Verlag, Berlin, 2009, pp. xviii+232.
[40] F.D. Sentilles, Bounded continuous functions on a completely regular space, Trans. Am. Math. Soc. 168 (1972) 311–336.
[41] H.M. Soner, Optimal control with state-space constraint. II, SIAM J. Control Optim. 24 (6) (1986) 1110–1122.
[42] M. Spivak, A Comprehensive Introduction to Differential Geometry, vol. 1, M. Spivak, Brandeis Univ., Waltham, MA, 1970, iii+656 pp. (not consecutively paged) paperbound, Published by M. Spivak.
[43] D.W. Stroock, S.R.S. Varadhan, Multidimensional Diffusion Processes, Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences), vol. 233, Springer-Verlag, Berlin-New York, 1979, pp. xii+338.
[44] H. Thorisson, Coupling, Stationarity, and Regeneration, Springer-Verlag, 2000.
[45] H. Whitney, Analytic extensions of differentiable functions defined in closed sets, Trans. Am. Math. Soc. 36 (1) (1934) 63–89.
[46] J. Yong, X.Y. Zhou, Stochastic Controls, Applications of Mathematics (New York), vol. 43, Springer-Verlag, New York, 1999, pp. xxii+438. Hamiltonian systems and HJB equations.